Università di Firenze, Università di Perugia, INdAM consorziate nel CIAFM

DOTTORATO DI RICERCA
IN MATEMATICA, INFORMATICA, STATISTICA

CURRICULUM IN MATEMATICA
CICLO XXXIV

Sede amministrativa Università degli Studi di Firenze
Coordinatore Prof. Matteo Focardi

Proving by mathematical induction: an analysis from history and epistemology to cognition

Settore Scientifico Disciplinare MAT/04

Dottorando:
Bernardo Nannini

Tutori
Prof. Samuele Antonini
Prof.ssa Veronica Gavagna

Coordinatore
Prof. Matteo Focardi

Anni 2018/2021

Abstract

The study presented in this thesis investigates the Proof by Mathematical Induction as a research problem in mathematics education. Mathematical induction (MI) is characterized by a considerable complexity, both from a historical-epistemological and a cognitive-didactic point of view. On the one hand, in fact, MI has a fundamental and foundational role in modern mathematics, a role that has been reached after a long process of historical development. On the other hand, research in mathematics education is unanimous in highlighting how MI is extremely problematic from a didactic point of view, with difficulties of different and various nature that are recorded transversely from novice to expert students.

The purpose of this research is to shed some light on this complexity, by observing and analysing it from different standpoints. In line with this research objective, a conceptual framework is structured, obtained as a *combination* (Prediger et al., 2008) of different perspectives, in order to get a multi-faced insight into the research problem.

First of all, this study takes into consideration a historical-epistemological perspective. In particular, an analysis of the historical genesis of the proof by induction is conducted, focusing on a series of traces of proofs by induction that historiographic research has identified. Subsequently, starting from this historical-epistemological analysis, other different theoretical perspectives are considered, each with a different focus in relation to the study of MI from a cognitive and educational point of view. Preliminarily, the theoretical standpoint on 'Argumentation and Proof' adopted in this thesis is presented. This standpoint, which emerges within the studies on *Cognitive Unity* (starting from Boero et al., 1996), allows to extend the focus of this study not only to the formal proofs by MI, but also to those informal argumentations related to MI, the *recursive argumentations*, as defined in the thesis. Another perspective considered is the APOS Theory (Arnon er al., 2014), adopted in this study to investigate students' interiorization of crucial processes for the construction of the *Schema* of MI (i.e., in APOS terms, the *Genetic Decomposition* of MI). In addition to this, the theoretical framework of *Intuitions* according to Fischbein (1987), is considered, with the aim of investigating on the students' *intuitive acceptance* of the proof by MI. Finally, a multimodal semiotic perspective, with the construct of the *Semiotic Bundle* (Arzarello, 2006), is also considered in order to identify and analyse an extended variety of signs (speech, gestures, inscriptions) produced and used by students as a lens to observe their processes related to MI during problem solving activities.

Within this conceptual framework, the research questions of the study are then formulated, each in line with one of the adopted perspectives. Besides the research question related to the historical-epistemological analysis (R.Q.1), the others address, respectively, students' interiorization of the *Explain Induction Process*, one of the crucial processes involved in the MI Schema, in APOS terms (R.Q.2), students' intuitive acceptance of the proving scheme by MI (R.Q.3), and students' use and production of multimodal signs in the generation and construction of recursive argumentations and proofs by MI during problem solving (R.Q.4).

To investigate these research questions, two different empirical studies are conducted. The first one is based on an online survey with some close-ended and open-ended questions. It involves a total of 307 participants including undergraduate and master's students from various academic courses from different Italian universities. The second empirical study involves some task-based interviews with

experienced (Doctoral) and less experienced (second or third year undergraduate) students in mathematics and physics. The data collected are video and audio recordings, interview transcripts, and solvers' written inscriptions.

Analyses of the data collected in the two empirical studies are then presented and discussed, highlighting some interesting results in relation to the research objective. On the one hand, these results include the identification of some problematic aspects for students in relation to the Explain Induction process, in particular the construction of chains of logical inferences (Modus ponens and Modus tollens), or in relation to the intuitive acceptance of MI (e.g. in accepting the validity of the implication involved in the inductive step as independent from knowing the truth value of antecedent and consequent). On the other hand, these results also include the identification of two categories of signs (named Linking and Iteration signs) and the construction of an interpretative tool (the distinction between ground and meta-level signs) that allowed to identify, observe, and analyse some effective processes of students involved in the construction of recursive argumentations and proofs by MI during problem solving.

Lastly, these results are discussed in order to provide an answer to the research questions. The findings of this thesis are then contextualised within the existing literature, highlighting how some of them confirm and extend previous studies on MI, and how others offer an original contribution in relation to the research problem. Finally, this thesis concludes by presenting possible research directions that this study offers and some didactical implications that arise from it.

TABLE OF CONTENTS

1 INTRODUCTION

1.1 A RESEARCH PROBLEM

The first point to clarify to present this work is the starting problem that prompted this study. With this aim, I wish to start by quoting two voices as distant in time as in their content. The first voice is of Henri Poincaré who, in *La Science et l'hypothèse,* referring to mathematical induction as a method to define mathematical objects or to prove statements, claims that "it is therefore mathematical reasoning *par excellence"* (Poincaré, 1905, p.9, italics in original). The second voice is of Francesca[1], a grade 12 student at an Italian secondary school. Francesca's class had just concluded a series of activity conducted by me in January 2018 aiming to the introduction of the proof by Mathematical Induction[2]. In a conclusive survey, in which it was asked to express comments on the activities and on eventually encountered difficulties, Francesca wrote: "The involved topics related more to the logic than to the use of the canonical mathematics that we study at the secondary school. The difficulty is that you have to rely on the logic and on the method of reasoning rather than on the use of mathematical laws. The biggest difficulty, for me, was that I had to adapt myself to an 'out of the box' mathematics, at least for what I studied at school". Francesca's words, and together with her those of several other classmates, describe mathematical induction as something completely distant and foreign to the mathematics she daily encounters at schools. Poincaré and Francesca's voices are, in my opinion, the representants of two point of views which, as far as distant, coexist.

The first one is the scientific community's point of view, in which the fundamental role of mathematical induction is recognised in almost every branch of mathematics, either as a proving method, or as a way to define mathematical entities. An attempt to give credit to this transversality and richness of application of mathematical induction, can be found in Gunderson (2011), a monograph containing almost one thousand pages of exercises, problems, and more broadly applications of mathematical induction. The importance of mathematical induction, however, does not uniquely lie on the variety of its applications, but also in its foundational role for the modern mathematics itself. The study on foundations of mathematics, which from the end of the 19th century involved several of the most influent mathematicians and philosophers, elected the Principle of Mathematical Induction as one of the keystones on which much of the modern mathematics is built. Coherently with this aspect, mathematical induction is part of the syllabus of almost every course in mathematics during the first years of university. Moreover, at least in Italy, a reference to mathematical induction is also present within the general objectives for the teaching of mathematics at secondary school: In the national indications for university oriented secondary schools (MIUR, 2010), among the eight "groups of concepts and methods that the student will actively dominate", it is indicated the following: "a knowledge of the principle of mathematical induction and the capacity of applying it, possessing, moreover, a clear idea of the philosophical meaning of this principle" [3] (p. 22, translated by me).

The second point of view, highlighted by Francesca's words, is instead the point of view of many students and teachers and from which several studies in mathematics education has started. Later in this chapter,

[1] The name is a pseudonym.

[2] The teaching experiment, designed within the framework described by Harel (2001), was part of my Master's final thesis.

[3] "Gruppi di concetti e metodi di cui lo studente saprà dominare attivamente. […] una conoscenza del principio di induzione matematica e la capacità di saperlo applicare, avendo inoltre un'idea chiara del significato filosofico di questo principio" (MIUR, 2010, p. 22.).

the main results of these studies will be presented, here it is sufficient to observe that the research in mathematics education is concordant in recognising mathematical induction as a didactically problematic topic. First of all, we can register that, in most of the cases, mathematical induction is a 'niche' topic, encountered by students only during the first years of the university, and almost exclusively among scientific academic courses. In Italy, despite what written in the national indications, it is very rarely encountered during the secondary school. I could register this aspect in a survey distributed among around three hundred students from several Italian universities, in different scientific academic courses, in which less than one third of the students answered to have encountered mathematical induction during the secondary school.[4] As a further aspect related to this point, we can observe that if we look at the final national written exams for scientific secondary schools in Italy ("Esame di Stato per liceo scientifico"), from 2010, the year of publication of the "Indicazioni nazionali per i licei", and today, mathematical induction was never included. The only exception is in a question contained in the written exam of 2010, involving a problem for which an appropriate solution could be obtained by mathematical induction.[5] A second problematic aspect, highlighted by the research in mathematics education, is the frequent presence of several and deep difficulties of students dealing with mathematical induction, and these difficulties are registered transversely among students of different mathematical experience. To summarise, thus, the didactical problem related to mathematical induction is dual: from one side most of the students conclude their mathematical experience having never encountered mathematical induction, from the other side, among those ones who encounter it, several students reach only a partial and sometimes not mathematically correct knowledge of it.

In conclusion, thus, the problem that moves the study presented in this thesis is described by the distance between the two just described point of views. From one side, mathematical induction is described as an epistemologically central aspect of the modern mathematics, being at the same time an important proving scheme and an axiom for several mathematical theories, and thus providing at the same time structure and foundation to the mathematical building. From the other side, however, looking at mathematical induction from the educational and didactical point of view, we can see how it occupies a very marginal role, sometimes being ignored by the school practice, and how it is often problematic to be understood by students.

In the next pages of this chapter, starting from the just presented problem, I will define the research general objectives of this study. Before focusing on this, however, in the following two sections I will present the mathematical standing point of the thesis and an overview of the principal studies in mathematics education focusing on mathematical induction.

1.2 Mathematical induction, Principle, Proof, and Definition

In this brief section I will clarify what, from now on in this thesis, will be indicated with *Principle of Mathematical Induction*, with *Proof by Mathematical Induction,* or with *Definition by Mathematical Induction*.[6]

The principle of mathematical induction generally indicates one of the axioms for the construction of the arithmetic of the natural numbers. This axiom could have different formulations with different levels of formality. For the purposes of this study, the chosen formulation is the following:

---

[4] The survey will be presented and analysed in detail in this thesis.

[5] The text of the problem was the following: "Let p(x) be a polynomial of grade n. Prove that its nth-derivative is $p'(x)=n!a_n$, where a_n is the coefficient of x_n".

[6] A more formal approach to MI can be found in Lolli (2008).

Principle of Mathematical Induction (PMI).

Let U be a subset of $\mathbb{N}$ such as:

(i) $0 \in U$
(ii) $\forall x.\ (x \in U \rightarrow (x+1) \in U)$

Then, U = $\mathbb{N}$.

The principle of MI, as well as being a foundation for the formalization of the set $\mathbb{N}$ of natural numbers, provides formal correctness to a proving scheme extremely used which takes the name of 'Proof by mathematical induction'.

Proof by Mathematical Induction (Proof by MI).

Let P(n) a predicate on the set $\mathbb{N}$ of natural numbers. If the following conditions hold:

(i) $P(0)$,
(ii) $\forall n \in \mathbb{N}. (P(n) \rightarrow P(n+1)$,

Then, $\forall n \in \mathbb{N}.\ P(n)$.

In particular, from the statement (i) and (ii), the validity of the statement *P(n*)* for any natural number *n** follows. Practically, to prove a statement with the form $\forall n.\ P(n)$, a possible proving strategy is to prove (i) and (ii) of above, which are traditionally called respectively *the induction base* (or *base case*), and *the inductive step.*

It is possible to generalise what just said to a statement with the form $\forall n \geq n_0.\ P(n)$. In this case the base of the induction becomes $P(n_0)$ and for the inductive step it is sufficient to prove '$\forall n \geq n_0.\ (P(n) \rightarrow P(n+1))$'.

The structure of a proof by MI can be expressed as an inference law. In its general form, for instance, it would be:

$$\frac{P(n_0) \wedge \forall n \geq n_0. \big(P(n) \rightarrow\ P(n+1)\big)}{\forall n \geq n_o. P(n)}$$

As said, from a formal point of view, the validity of a proof by MI is guaranteed by PMI. Generally, however, an intuitive and non-formal justification for the correctness of a proof by MI is provided: starting from the base case and the inductive step it is possible to construct a series of logical inferences which starts form the base n_0 and reaches any natural number $n \geq n_0$. Poincaré himself describes this non-formal justification as it follows:

> The essential characteristic of reasoning by recurrence [MI] is that it contains, condensed, so to speak, in a single formula, an infinite number of syllogisms. […]. They follow one another, if one may use the expression, in a cascade. The following are the hypothetical syllogisms: The theorem is true for the number 1. Now if it is true of 1, it is true of 2; therefore it is true of 2. Now, if it is true of 2, it is true of 3; hence it is true of 3, and so on. (Poincaré, 1905, pp. 9-10).

Proof by 'Strong' mathematical induction

In this thesis, in a few cases, I will refer to another proving scheme, related to the previous one[7], which is generally called 'Strong' Mathematical Induction, whose structure is the following:

Let P(n) a predicate on the set $\mathbb{N}$ *of natural numbers. If the following condition holds:*

(i) $\forall n \in \mathbb{N}.\ [\forall k < n. (P(n)] \rightarrow P(n+1)$ [8]

Then $\forall n \in \mathbb{N}.\ P(n)$.

Definition by Mathematical Induction (or by Recurrence)

Formally PMI can be used to construct the structure of the set $\mathbb{N}$. In particular, with PMI it is possible to prove a theorem, traditionally called 'the recursion theorem', which assures the existence and the unicity of functions on $\mathbb{N}$ defined by recurrence. Starting from this, it is possible to define in $\mathbb{N}$ the classic operations of sum and product, and the total order relation ≤, with the respective properties.

Very often, in the mathematical practice, PMI is used to define mathematical objects even outside the formal context of axiomatisation of $\mathbb{N}$. Let us suppose, for instance, that we want to define a countable family of sets $\{A_i\}_{i \in \mathbb{N}}$. It can be done as it follows:

(i) We define the set A_0.
(ii) We define the set A_{n+1}, in terms of the definition of the set A_n, for a generic index *n*.

In this way, any element of the family $\{A_i\}_{i \in \mathbb{N}}$ is recursively defined from the previous elements.[9]

Let us observe, finally, that in the practise of the mathematician, what just said is used in a more or less formal way, depending on the context and on the objectives of the work in which it is used. As we will see, one of the aims of this study is to investigate the role that Mathematical Induction could have, not when it is used in formal contexts, but in other situations, still extremely common in the praxis of this discipline, as the exploration of problems or the formulation of conjectures.

1.3 Mathematical Induction in Mathematics Education

In this section I will present the principal studies on mathematical induction in mathematics education. Most of the studies here presented will be further explored later in this thesis. The following is a critical overview on these studies, aiming to organise them within a unitary discourse. With this objective, they will be presented following as divided in three categories: the studies focusing on epistemological aspects of MI, the studies investigating on the students' difficulties related to MI and the studies describing some effective teaching experiences for the learning of MI. The three proposed categories are not mutually exclusive, as several studies presented within one category also investigate some points of another. In these cases, the different aspects contained in the same study will be presented each in the respective category.

[7] For a formal discussion of the logical connections between MI and 'Strong' MI, see Lolli (2008).
[8] This also include the condition that P(0).
[9] Again, the formal correctness of this strategy is given by PMI and by the recursion theorem.

1.3.1 Epistemological aspects of MI

In a work from more than one hundred years ago, Young (1908) writes: "The process of mathematical induction is exceptionally well fitted to introduce the beginner to the philosophic study of mathematical thinking" (p. 146). Young's opinion is that the study of MI has a didactic potentiality because it can be the starting point to introduce some deep and crucial topics of the modern mathematics, such as those related to its foundations or to the different meanings of the term 'infinite'. The epistemological centrality of MI in mathematics is also analysed, more recently, by Ernest (1984), who, before focusing on some educational issues related to MI, presents a "Conceptual analysis of Induction". In this analysis, Ernest focuses on a series of different mathematical concepts underlying or related to MI. The Figure 1.1 shows the network diagram resulting from his analysis.

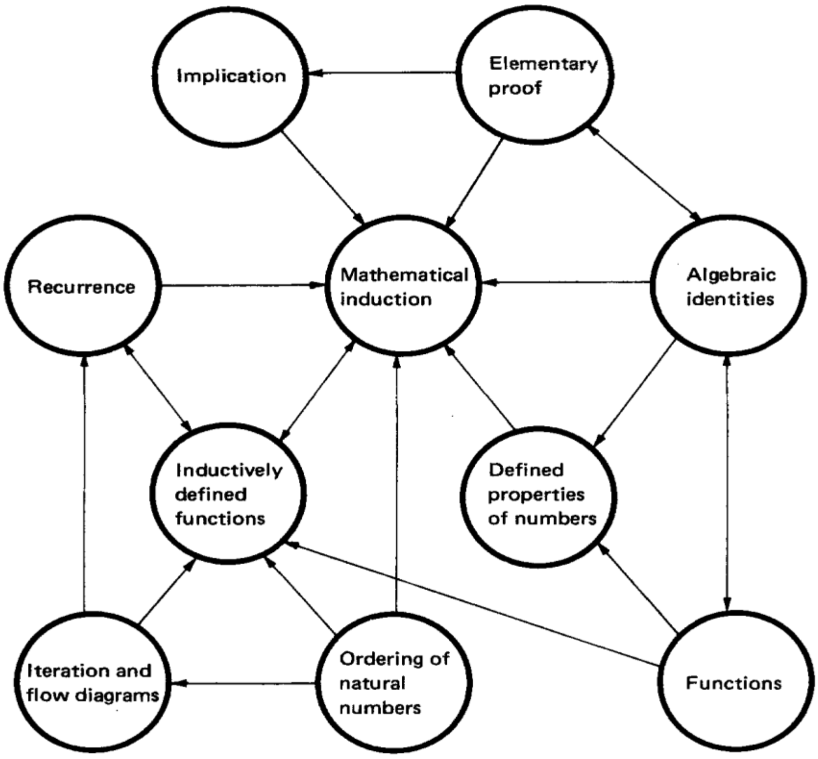


Figure 1.1. *Mathematical Concepts underlying or related to MI following Ernest's analysis. The figure is taken from Ernest, 1984, p 180.*

One of the aspects that Ernest's analysis highlights is the connection between mathematical induction and computational recurrence. Leron and Zazkis (1986) focus on this aspect, observing that there is a deep similarity between a proof by induction and a recursive process:

> A recursive process invokes a version of itself as a subprocess during execution; a recursive procedure contains a version of itself as a subprocedure in its definition; in a proof by mathematical induction, one version of the theorem is used to derive another version. (Leron & Zazkis, p. 25, 1986).

The authors observe that these similarities can be exploited from a didactical point of view. In particular they suggest that appropriate activities with the computational recursion can be used by teachers to provide to students a "steppingstone toward a better understanding of induction." (Ibidem, p. 28).

The fact that MI is central in aspects related to the foundations of the modern mathematics entails a logical-epistemological complexity which could be potentially problematic from a didactic point of view. Lowenthal and Eisenberg (1992) raise the issue of the nature of the PMI as a 'meta-theorem'. They observe that often, MI is presented to students simply as a routine procedure used to prove theorems, whose validity is justified with some metaphors (as the one of the line of falling dominos), whilst other mathematically and philosophically important aspects are ignored, such as the necessity for a proof to be finite or the role of PMI as a 'meta-theorem'. Lowenthal and Eisenberg state that this creates a sort of "illusion of rigor" (p. 237) which brings them to raise the issue of whether the introduction of MI at school, without a focus on its philosophical importance, is worth it or not.

A further aspect of complexity related to MI is its delicate relationship with empirical induction (i.e., generalising a statement from some particular cases). As we will see, this aspect has been discussed with a cognitive point of view in relation to students' difficulties with MI. Rips and Asmuth (2007), however, observe that even from a logical and epistemological point of view it is possible to register a tension between mathematical and empirical induction. The authors analyse some logical aspects of MI, observing that in two points some similarity with empirical induction seem to emerge. The first one is in the logical law known as Universal Generalization (UG), affirming that if a predicate is proved for an arbitrary member of a domain, it must be valid for all members of the domain[10]. This rule is used within a proof by MI, in the inductive step when one proves $P(n) \rightarrow P(n+1)$ for an arbitrary n in $\mathbb{N}$, in order to prove $\forall n \in \mathbb{N}\ (P(n) \rightarrow P(n+1))$. UG allows to prove a universal statement proving it for a generic case. Since this case is generic, arbitrary, we are not formally using an (empirical) inductive inference, however a similarity seems to be present since in both cases a general conclusion is inferred from a premise involving a particular case. The second point in which Rips and Asmuth see some similarities between MI and empirical induction is in the non-formal justification of the validity of MI as a series of syllogisms. They state:

> "[T]here is a tension between math induction and the usual constraint that proofs be finite. Math induction seems to rely on an infinite process of iterating through the natural numbers. […]. From a psychological point of view, of course, no such infinite process is remotely possible. It's this gap that creates the similarity with empirical induction, since in both cases the conclusion of the argument seems beyond what the premises afford." (Rips & Asmuth, p. 16, 2007).

1.3.2 Students' difficulties with MI

Several studies in mathematics education provide empirical evidence of students' difficulties or, more generally, problematic aspects related to MI. On this aspect, the literature is concordant in considering MI as a problematic topic from a cognitive and educational point of view. Students' difficulties related on MI have been investigated focusing, principally, on three aspects: the difficulties related to the logical structure of a proof by MI (or of PMI), the technical difficulties in the construction of a proof by MI, and by reinterpreting, with a specific focus MI, some classic studies on students' difficulties with mathematical proofs.

Difficulties related to the logical structure of MI

One of the most problematic aspects for student relates to the understanding of the inductive step, where it is necessary to prove the validity of the implication $P(k) \rightarrow P(k+1)$ for a generic natural number k. Avital and Lebeskind (1978) observe that often students seem to interpret the proof of the implication in the inductive $P(k) \rightarrow P(k+1)$ as the proof of $P(k+1)$, which poses the problem: "How can you establish the truth of $P(k+1)$ if you don't even know if $P(k)$ is true?" (p. 430). Similarly, Dubinsky and Lewin (1986) highlight that it could be problematic for students to interpret the implication $P(n) \rightarrow P(n+1)$ as a unique predicate $Q(n)$. Fischbein and Engel (1989) further investigate this aspect, observing that, from a cognitive

[10] For a more formal definition of UG, see van Dalen, 2008, pp. 91-93. The use of UG within a proof by MI will be analysed in detail in this thesis at 3.3.2.

point of view, it could be problematic to accept that the validity of the implication *P(k)→P(k+1)* as independent from the truth of *P(k)*, the inductive hypothesis:

> The difficulty is that the student has to build the entire segment of the inductive step (if P(k) is true then P(k+1) is also true), on a statement which, itself, has not been proven and cannot be proven in this segment of the reasoning process. As a matter of fact, we are absolutely not used to this way of reasoning. We may start from a given reality for formulating a certain theorem and then try to prove it by checking its consequences. […] What we usually do not do is to check the implication itself, in itself, without any concern whatsoever for the truth of the two statements involved. It is like building a bridge in the air without any support on both its ends." (p. 284, underlined in original).

Still focusing on this point, Ernest (1984) observes that a further complexity is given by the fact that the inductive step involves the same predicate *P(n)* of the theorem to be proved. In proving the implication of the inductive step, one assumes *P(n)* for a generic $n\geq 0$ (i.e., the inductive hypothesis), while the general statement corresponds to $\forall n\geq 0.\ P(n)$. This might bring to the misconception that in a proof by induction, with the inductive hypothesis, we are assuming what we want to prove, with obvious problems of circularity.

Beyond the inductive step, research focused on problematic aspects related to the base of the induction. Several studies (Ernest, 1984; Dubinsky & Lewin, 1986; Ron & Dreyfus, 2004; Garcia-Martinez & Parraguez, 2017, Larson & Petterson, 2018) register that often proving $P(n_0)$ is not recognised by students as necessary for a proof by induction, but it is simply considered as a preliminary verification of the truth of the statement for a particular case. Pang and Dindyal (2012) register that for some students the base case should be verified before the induction step for the proof to be valid. Moreover, other studies (Avital & Lebeskind, 1978; Stylianides et al., 2007), show that students could have difficulties in recognising that, in some cases, the base of the induction could involve a number n_0 different from 0 or 1.

Another critical aspect of mathematical induction for students is related to its non-formal justification, that is why once proved the validity of P(0) and of $\forall n\geq 0\,(P(n) \rightarrow P(n+1))$ one could conclude $\forall n\geq 0\ P(n)$.[11] As already observed this is could be obtained by constructing a series of syllogisms: P(0) and P(0) → P(1), implies by *modus ponens* P(1); P(1) and P(1) → P(2) implies by *modus ponens* P(2), and so on. Fischbein (1987) observes that in order to conclude that the predicate P(n) is true for the whole and infinite set of natural numbers it is necessary to pass from a potential to an actual view of infinity and this requires an "intuitive leap" (p.52) which could be problematic for a student. Harel (2001) highlights that often students are convinced of the validity of MI by ritual or authoritarian aspects instead of by deductive constructions. Moreover, Palla et al. (2012) register difficulties in students in justifying why a proof by induction 'works' on the set of natural numbers but not on other sets, such as on the set of real numbers.

Technical difficulties in proving by MI.

Other studies highlight that the difficulties with MI could also be related to the construction of a specific proof by induction. Avital and Lebeskind (1978) observe that often students cannot determine the predicate *P(n)* involved in the problem or the parameter *n* on which to apply induction. Moreover, they observe that often the predicate P(n) involves some algebraic expressions that need to be appropriately manipulated and this might cause some difficulties. Nardi and Iannone (2003) further investigate this last point showing that specific difficulties arise in students when proving statements involving finite sums

[11] For the sake of brevity, I have considered P(0) as the base case. Analogous reasoning can be done if the base case involves a different natural number n_0.

for which they often seem not to recognise the parameter on which applying induction within the several variables involved or when the predicate P involves non-symmetrical, one-way relationships (such as, for instance, algebraic inequalities).

Studies on students' difficulties with proof and proving reinterpreted with a focus MI.

Further research on students' difficulties with MI developed from other classical studies in mathematics education focusing on students' difficulties with mathematical proofs and proving processes. Hanna (1989) observes that, from a didactical point of view, it is possible to distinguish between proofs that 'prove' (i.e., that validate the statement) and proofs that 'explain' (i.e., that explicate "the mathematical properties that cause the asserted theorem or other mathematical statement to be true", p. 51). Hanna observes that proofs that prove without explaining could be rather problematic from a didactical point of view. As an example for this distinction, Hanna presents two proofs for the classic formula for the sum of the first *n* natural numbers ($1+2+3+\ldots\ n) = \frac{n(n+1)}{2}$), the first one by MI and the second one by showing that the same result can be obtained by summing $[1+2+3+\ldots\ n]$ with $[n+(n-1)+(n-2)+\ldots\ 1]$, and then dividing the result by 2. The author observes that the second one, differently from the first one, is an example of a proof that shows. Since "[i]n general, proofs by mathematical induction are non-explanatory" (ibidem, p. 52), some students' difficulties with them could be interpreted in terms of difficulties in accepting a non-explanatory proof.

Another critical aspect relates to the problematic relationship between empirical induction and mathematical induction. As observed by Harel and Sowder (1998), several students are completely convinced by empirically inductive argumentations and this contributes to the fact that they do not feel the need for a mathematical proof in general, and in particular for a proof by MI for those statement that they could have already tested for some specific numbers. A further aspect which contributes to this problem is that, as observed Ernest (1984) there is a "unfortunate ambiguity in the word 'induction'" (p. 180). The delicate relationship between empirical and mathematical induction has been also analysed for students aware of the necessity of a mathematical proof to prove a conjecture. Pedemonte (2007), within theoretical perspective of the *Cognitive Unity* (Boero et al., 1996) focuses on students' difficulties in the transition between the student's production of a conjecture and the successive construction of a proof by MI. The researcher observes that an empirical inductive argumentation supporting the conjecture, does not provide a deductive structure following which a subject could subsequently construct a proof by mathematical induction. Therefore, it is possible that a student does not recognise as convincing a proof by MI because of the distance (i.e., absence of *cognitive unity*) between the proof and the previous argumentation.

Transversality of the difficulties

A further aspect emerging from the just presented studies is that the difficulties related to MI or the presence of not mathematically correct conceptualisation of MI have been registered transversely in subjects with different experience in mathematics, from secondary school students to third year university students in mathematics and Secondary school mathematics teachers. The table 1.1 summarises this variety of results.

This aspect highlights how the difficulties related to MI could be rather problematic to overcome. Focusing on this point, Movshovitz-Hadar (1993) uses the term 'fragile knowledge' in relation to MI. In her empirical study, she registers that many of the involved students stated, at the beginning of the study, to consider MI a not very difficult topic and they evaluate their expected success to cope with a proof by

MI between 75% and 100%. However, when dealing with a paradoxical statement for which a (false) proof by induction was presented[12], not only could they not find the flaw in the proof, but some of them started doubting the validity of a proof by MI in general. Carotenuto et al. (2018) have replicated the study with some third-year university students in mathematics, this time showing a proof by MI for the statement 'All new-borns have eyes of the same colour' and obtaining similar results. The researchers, thus, conclude:

> MI remains a very hard topic, also in the case where there is a shared awareness of the necessity of proving by a deductive process the empirical conjectures and also when students have excellent technical capabilities. (Ibidem, p. 225).

Subjects involved	**Empirical Studies showing difficulties with MI**
Secondary School Students (11th grade)	Palla et al., 2012
Secondary School Students (11th grade)	Fischbein & Engel, 1989
Secondary School Students (12th grade)	Pang & Dindyal (2012)
Secondary School Students (12th and 13th grade)	Pedemonte, 2007
Undergraduate Students (1st year)	Dubinsky & Lewin, 1986
Undergraduate Students (1st year)	Nardi & Iannone, 2003
Undergraduate Students (1st year in Mathematics and Physics)	Larson & Petterson, 2018
Undergraduate Students	Harel, 2001
Undergraduate Students	Garcia-Martinez & Parraguez, 2017
Undergraduate Students (3rd year in Mathematics)	Carotenuto et al., 2018
Preservice Secondary School Mathematics Teachers	Movshovitz-Hadar, 1993
Preservice Elementary and Secondary School Mathematics Teachers	Stylianides et al., 2007
Secondary School Mathematics Teachers	Ron & Dreyfus, 2004

Table 1.1. The just presented studies investigating on difficulties related to MI. The table highlights the transversality of the empirically registered difficulties. In the table are reported only those works in which an indication of the level of the subjects involved in the study is provided.

1.3.3 Effective teaching experiences with MI

Some researchers in mathematics education reported on some teaching interventions related to MI with different groups of students that have resulted being effective from a didactic point of view.

Dubinsky (1986, 1989) presents a teaching experience in which MI was introduced to some undergraduate students in Computer Science with a series of different computer activities with a programming language called ISETL. During a series of lessons, students arrived to:

- Create an ISETLS function that, given a predicate $P(n)$ as input (corresponding to an empty array with infinite dimension), returns as output the predicate $Q(n) = P(n) \rightarrow P(n+1)$.
- Create an ISETL function that, given a predicate $P(n)$ as input determines, if it exists, the first natural number n_0 for which $P(n_0)$ is true.

[12]The proof is the one presented by Austin (1988), for the statement: 'Given any number N of coins, all identical except for one which is lighter, it is always possible, to determine the false coin with no more than 4 weighings of an equal-arm pan balance'.

- Finally, using the previous two functions, create a ISETLS program that, given a predicate *P(n)*, determines the first natural number n_0 for which $P(n_0)$ is true and then enters in a loop of applications of *modus ponens* through the predicate *Q(n)*. If the proposition $\forall n \geq n_0.P(n)$ is valid, the program would run forever.

Dubinsky observes that although the program cannot provide a formal proof for the statemen $\forall n \geq n_0.P(n)$, since it would need an infinite number of steps, nevertheless it supported students in constructing an intuitive justification for the validity of MI. Successively the students were involved in the construction of classic proofs by induction with a paper and pen. Finally, Dubinsky registers, considering the results of a conclusive test, that the teaching experience seems to have been effective for introducing MI to the students.

A teaching experience which does not involve computer activities is presented by Harel (2001), in which some undergraduate students from different academic courses are introduced to proving by MI with a series of problem solving activities. Harel starts from two theoretical premises, one involves the formulation of the problems and the other introduces a distinction in the possible argumentative solutions for the problem. For what concern the problems formulation, Harel, with a focus on *recursive problems* (i.e. when "the mathematical solution of the problem requires the formation of a recursive representation of a function", ibidem p. 190), distinguishes between *explicit recursion* and *implicit recursion* depending if a recursive representation is contained explicitly or not in the problem's text. For example, the problem 'Prove that $1 + 3 + \cdots + (2n - 1) = n^2$ for all positive integers n" is an explicitly recursive problem while problems such as the Tower Hanoi Problem[13] are implicitly recursive. Secondly Harel proposes a distinction between two particular kinds of argumentations that a subject might produce when solving a problem: the *result pattern generalization* and the *process pattern generalization*. I will introduce the with two examples reported by Harel himself (Ibidem, pp. 191-192). He compares the answers given by two students to the following problem: 'Prove that for all positive integers n, $\log(a_1 \cdot a_2 \cdot \ldots \cdot a_n) = \log a_1 + \log a_2 + \ldots + \log a_n$'. A student writes:

Response 1

log(4·3·7) = log 84 =1.924 log(4·3·6) = log 72 = 1.857
log 4 + log 3 + log 7 = 1.924 log 4 + log 3 + log 6 = 1.857
Since these work, then log(a1·a2·…·an) = log a1 + log a2 +… + log an.
(Ibidem, p.191)

While another student writes, instead:

Response 2

(1) $\log(a_1 \cdot a_2) = \log(a_1) + \log(a_2)$ by definition

(2) $\log(a_1 \cdot a_2 \cdot a_3) = \log(a_1) + \log(a_2 \cdot a_3)$. Similar to $\log(a \cdot x)$ as in step (1), where this time $x = a_2 a_3$.
Then
$\log(a_1 \cdot a_2 \cdot a_3) = \log(a_1) + \log(a_2) + \log(a_3)$

(3) We can see from step (2) any $\log(a_1 \cdot a_2 \cdot \ldots \cdot a_n)$ can be repeatedly broken down to
$\log a_1 + \log a_2 + \ldots + \log a_n$.
(Ibidem, p.192)

[13]"The Tower of Hanoi Problem: Three pegs are stuck in a board. On one of these pegs is a pile of disks graduated in size, the smallest being on top. The object of this puzzle is to transfer the pile to one of the other two pegs by moving the disks one at a time from one peg to another in such a way that a disk is never placed on top of smaller disk. How many moves are needed to transfer a pile of n disks?" (Harel, 2001, p. 195)

Harel observes that, even if the two answers are based on a generalization (i.e., "drawing a general conclusion from a number of individual cases", ibidem, p. 191), however in the first case the focus is on the regularity of the result of a particular operation, whilst in the second case the focus is on the regularity of the involved process. The first argumentation is called by Harel *Result Pattern Generalization*, and the second one *Process Pattern generalization*. He observes that the two argumentations are generalizations starting from a limited number of examples, however they have a very different structure which can be interpreted as (empirical) inductive, the first one, and deductive the second one. Harel writes:

> In Response 1 the reference rule that governs the evidencing process is empirical; namely, $(\exists r\in R)(P(r)) \rightarrow \forall r\in R)(P(r))$. In Response 2, on the other hand, it is deductive; namely, it is based on the inference rule $(\forall r\in R)(P(r)) \wedge (w\in R) \rightarrow P(w)$ (Here, r is any pair of real numbers a and x; R is the set of all pairs of real numbers; P(r) is the statement "log(ax) = log a + log x" and w in step n is the pair of real numbers, $a_1a_2 \dots a_{n-1}$ and a_n.) (Ibidem, p.193).

Moreover, Harel observes that, through an argumentation based on a process pattern generalization it is possible to construct an iterative argumentation which could provide a proof for a statement with the form $\forall$*n. P(n)*.

Let us suppose, for instance, that given the sequence '$\sqrt{2}$, $\sqrt{2+\sqrt{2}}$, $\sqrt{2+\sqrt{2+\sqrt{2}}}$, …' we need to prove that all its terms are lower than 2. A possible argumentation could be the following:

- The first term of the sequence $\sqrt{2}$ is lower than 2.
- From the previous point it follows that $2+\sqrt{2}$ is lower than 4, therefore also the second term of the sequence $\sqrt{2+\sqrt{2}}$ is lower than 2.
- From the previous point it follows that $2+\sqrt{2+\sqrt{2}}$ is lower than 4, therefore also the third term of the sequence $\sqrt{2+\sqrt{2+\sqrt{2}}}$ is lower than 2.
- And so on for all the terms of the sequence.

Harel uses the term *Quasi-induction* for this particular form of argumentation to highlight the connection between this one and mathematical induction.

At this point Harel presents his teaching intervention aiming to introducing MI as a proving scheme to students. The intervention was articulated in three different phases:

1) Students are firstly involved in some problem solving activities, specifically with implicit recursion problems. Harel highlights how this choice is different from a standard introduction to MI, in traditionally the implicit recursion problem are the last one to be presented. Harel justifies his choices stating that implicit recursion problems seem to "induce students to focus on process pattern generalization rather than on result pattern generalization". (Harel, 2002, p. 194). At the end of this phase, students seemed to recognise as not acceptable the use of result pattern generalization to prove a statement and they could construct autonomously quasi-inductions for this category of problems.
1) In the second phase, the explicit recursion problems were introduced. The didactical goal of this second phase was that the students recognised that quasi-induction could be used to solve these second category of problems as well.
2) In the third and last phase, the proofs by MI are introduced. The transition between quasi-induction and MI requires the transformation of the series of implications $P(0)\rightarrow P(1)$, $P(1)\rightarrow P(2)$, and so on, in a unique $P(n)\rightarrow P(n+1)$ which represents generically all the others. Harel observes that this transition is cognitively delicate:

> In quasi-induction the conviction that P(n) is true for any given natural number n stems from one's ability to imagine starting from P(1) and going through the inference steps, P(1)→P(2), P(2)→P(3), …, P(n−1)→P(n). This does not mean that one actually runs through many steps, but that he or she realizes that in principle this can be done for any given natural number n. In particular, in quasi-induction one views the inference, P(n−1)→ P(n), just as one of the inference steps—the last step—in a sequence of inferences that leads to P(n). In MI on the other hand, one views the inference, P(n)→P(n+1), as a variable inference form, a placeholder for the entire sequence of inferences. (Ibidem, p. 201).

Finally, Harel registers that the intervention seems to have been effective for introducing MI to students. In particular, he observes:

> The students seemed to have easily assimilated the principle of MI into their scheme of quasi-induction: About 75% of the problems assigned were solved correctly and in terms of the principle of MI; the rest were either solved by quasi-induction or by other means. (ibidem, p. 203).

Harel concludes observing that the effectiveness of the intervention was registered not only in relation to a correct conceptualisation of MI from many students, but also in relation to other more general aspects, as the transition from the use of empirical to deductive argumentations:

> The most significant result reported in this paper is that in this alternative treatment students altered their current ways of thinking, primarily from mere empirical reasoning—in the form of result pattern generalization—into transformational reasoning—in the form of process pattern generalization. (Ibidem, p. 206)

Cusi and Malara (2009) present another teaching intervention, still grounded on the use of quasi-induction, this time named *naïve induction:* "showing the passage from k to k+1 for particular values of k" (p.16). This study, differently from the previous ones, does not involve students with no experience with MI, but instead it involves subjects some pre and in service middle school teachers who had already encountered MI during their studies aiming at improving their awareness about the meaning of the principle of MI. The researchers start from the following theoretical premise:

> "The essential steps in a constructive path toward PMI should include:
> 1. A thorough analysis of the concept of logical implication.
> 2. An introduction of PMI through the naïve approach, drawing parallels between
> PMI and the ordering of natural numbers, and the use of reference metaphors.
> 3. A presentation of examples of fallacious induction to stress the importance of the inductive
> basis." (Cusi & Malara, 2009, p. 17).

The researchers register that the teaching experiment, framed on these three points, seemed to have obtained some positive results on the involved subject, in particular in supporting them in overcoming two widespread misconceptions: the fact that in the inductive step, through the inductive hypothesis we assume what we are trying to prove, and the fact that the base of the induction is not really necessary for the proof.

Finally, I conclude this literature review by presenting the study presented by Stylianides, Sandefur and Watson (2016). The study does not involve a teaching experiment, but it contains analysis of some effective processes activated by some undergraduate students when dealing with the resolution of problems that might involve MI. As written above, Hanna (1989) highlights that generally, a proof by MI is not explanatory for the reader and this could cause problems for students. Stylianides, Sandefur and Watson (2016), challenge this point by observing that if the focus is on the proving process instead of on the proof as a product, then proving by MI can be explanatory for a subject. In particular, the researchers observe that some conditions seem to be important for this aspect to realise:

(1) The mathematical problem should have an open formulation so that the students need to investigate it in order to produce a conjecture for it.
(2) Students need to have experience of the utility of generating examples for a proving activity. In particular it is important for them to be aware of "how constructing examples can be used not just to verify but also to expose structural relationships and generate conjectures." (Ibidem, p. 24).
(3) Students need to be able to manipulate the constructed examples with some familiarity in order to focus on the underlying structural relationships between them.
(4) The problem solution should not be reached by the subject with a simple symbolic manipulation, but it should require a subject's consideration of the mathematical relationships underlying the objects involved in the problem.

The authors conclude that when the point (1)-(4) of above are realised it is possible for students to be involved in the processes of exploration of the problem and of examples construction that could became effective for them to prove by MI their conjectures. In particular, as they write,

> [T]his exploration could, in turn, lead them [(the students)] to construct informally the inductive step and, eventually, see the utility and apply mathematical induction in their proving activity to formalize their work. In such situations, proving by mathematical induction becomes a codification of students' sense making and serves first as a method that explains and then as a proof that verifies. (Ibidem, p. 33).

1.4 Objectives and focus of this study

In section 1.1 I have highlighted the research problem that moves this study, which is that despite MI has a foundational and epistemologically central role for the modern mathematics, it is a widely problematic topic from an educational point of view. The just presented literature review has highlighted that the problem is extremely complex from a cognitive and didactical point of view. On the other side, also from the mathematical point of view, MI seems to have an epistemological complexity. A trace of this complexity can be registered, as we will see, within the history of mathematics observing the long process which preceded the genesis of the modern formulation of proofs by MI.

The main objective of this study, therefore, is to shed light on this this complexity, investigating it both from a historical-epistemological and a cognitive-didactical point of view. By saying 'shed light' I do not mean 'to solve' this complexity, but to 'put ourselves in the conditions to observe it'. This point of view is drawn from the philosopher Edgar Morin, who in his influent essay 'On complexity' (2008), writes:

> What is complexity? At first glance, complexity is a fabric (complexus: that which is woven together) of heterogeneous constituents that are inseparably associated: complexity poses the paradox of the one and the many. Next, complexity is in fact the fabric of events, actions, interactions, retroactions, determinations, and chance that constitute our phenomenal world. But complexity presents itself with the disturbing traits of a mess of the inextricable, of disorder, of ambiguity, of uncertainty. Hence the necessity for knowledge to put phenomena in order by repressing disorder, by pushing aside the uncertain. In other words, to select the elements of order and certainty, and to eliminate ambiguity, to clarify, distinguish, and hierarchize. But such operations, necessary for intelligibility, risk leading us to blindness if they eliminate other characteristics of the complexus. (Morin, 2008, p. 6).

On other terms, thus, by focusing on the complexity of a research problem, we need to accept the possible presence, in the problem itself, of contradictions, paradoxes, and uncertainties because these ones are proper characteristics of the complexity itself that we aim to observe. Any trial to reduce the "disorder" of this complexity could risk bringing to a simplification of the problem which underestimates it, or, to use Morin's words, to a "blind intelligence" (ibidem, p.4).

Starting from this idea, therefore, this study firstly aims to identify some theoretical perspectives in mathematics education to observe the research problem within its complexity. Different point of views could shed light on different aspects of MI as a didactical problem. Then, this study aims to adapt and refine theoretical lenses and constructs within the chosen perspectives in order to interpret and analyse some teaching-learning processes that constitute this complexity. Moreover, the development of these lenses and constructs aims to provide tools to formulate didactical objectives or to construct didactical activities.

With these objectives, in this study I will focus on investigating the processes in the learning and in the use of MI as an instrument to construct argumentations or proofs, define, generate conjectures, solve problems.

1.5 INTRODUCTION OF THE CONCEPTUAL FRAMEWORK

The conceptual framework of this thesis is composed by some theoretical perspectives. The motivations for the choice of each of them and the respective research foci will be briefly discussed in this section. Further details will be provided in Chapters 2-7, which are dedicated to the description of the theoretical perspectives and constructs used in this study and to the research questions posed within each theoretical perspective.

The first considered point of view is a historical-epistemological one. Although it is quite improper to refer to this perspective as a theoretical framework for Mathematics Education, it will be part of the conceptual framework of this study. This choice, as it will be further discussed, was made starting from two hypothesis. Firstly, I considered that an analysis of the historical development which brought to the genesis of MI could provide indications on its epistemology. Secondly, I considered that such a historical-epistemological analysis could provide elements to interpret processes in the use, in learning and in teaching of MI. These elements could be then used for the successive cognitive-didactical analysis of MI. Coherently with what just said, thus, the historical-epistemological perspective is considered as a part of the conceptual framework of this thesis. The research goal within this perspective is to analyse a series of voices belonging to the history of mathematics which, following several scholars, contain traces of the genesis of MI.

This study is inserted within the general line of research on proof and argumentation. Focusing on this theme, several and different theoretical perspectives have developed in mathematics education. Therefore, preliminarily I will explicate to what extent and on which theoretical point of views this study is framed in relation to this aspect. The theoretical starting points of this thesis will be the studies which traditionally go under the name of the construct of Cognitive Unity (starting from Boero et al., 1996). Within this perspective, a central aspect from a research and didactical point of view is to focus on the processes which bring students to the generation of a conjecture and to the construction of an argumentation supporting it, as preliminary processes for the successive construction of a proof. The studies within this perspective have highlighted that students' activity aiming to the generation of conjectures or to the construction of argumentations could be at the same time an effective didactical tool to introduce students to mathematical proofs, and a useful research lens through which to observe and interpret students' effective processes or difficulties when dealing with the construction of mathematical proofs. Within this study, this general perspective will be adopted focusing on MI. As a consequence of this perspective, therefore, in this study we will be interested in investigating not only on proofs by MI but also on other forms of less formal argumentations related to it.

The third point of view which this study takes into consideration is the APOS theory, a classical theoretical perspective in mathematics education (Arnon et al. 2014). As it will be presented, this perspective has been used other times to investigate MI from a cognitive and didactical point of view. In particular, the APOS theory is well fitted to provide a model for a possible cognitive counterpart of the logical structure of MI. The research goal of this study within this perspective will be dual: from one side to investigate if the conducted historical-epistemological analysis could provide new elements to deepen and expand the APOS analysis of MI in the literature, from the other side to focus on some points of this analysis with new experimental data.

A further aspect which will be addressed in this study is the intuitive acceptance of MI as a proving scheme. The theoretical references for this point are Fischbein's studies on intuition (1987) and on the intuitive acceptance of a proof for a subject (1982). As highlighted during the literature review, several studies register students' difficulties in being fully convinced of the validity of a proof by MI, also among students who seem to know how to construct a proof by MI. The research goal for this study within this perspective, thus, is to investigate which specific aspects of MI seem to present the major difficulties in being intuitively accepted by students and on which processes could instead support students in reaching an intuitive acceptance of this proving scheme.

As said, in this study we will investigate the students' processes of construction of proofs by MI or of argumentations related to MI. To analyse this, a semiotic perspective will be taken into consideration. In particular, focusing on the signs produced and used by a subject when constructing an argumentation or a proof by MI could provide a window through which to observe processes involved in these constructions. This study will use the theoretical construct of the Semiotic Bundle (Arzarello, 2006), a perspective in which a wide spectrum of signs is considered, including also less codified categories of sign, such as spoken utterances, written inscriptions, and gestures. This perspective will be used to observe and analyse the semiotic production of students involved in the resolution of problems potentially related to MI.

The use of different theoretical perspectives, as the just presented ones, is concordant with the complexity of the research problem. In particular, as said, this choice allows to take into consideration different aspects related to MI. However, a comment is necessary on the mutual relationship between these different theoretical perspectives within this study. The (meta-)theoretical reference for this point is the *Networking of theories* (Prediger et al., 2008; Bikner-Ahsbahs & Prediger, 2014). With reference to the networking strategies described in this meta-theoretical perspective, we can say that the considered theoretical perspectives will be *combined* in this study. The term 'combining' indicates the networking strategy considering different theoretical approaches as juxtapose and autonomously used in order to observe a problem from different point of views:

> Sometimes, the theoretical approaches are only juxtaposed […]. Then we speak of combining […]. Combining theoretical approaches does not necessitate the complementarity or even the complete coherence of the theoretical approaches in view. Even theories with conflicting basic assumptions can be combined in order to get a multi-faceted insight into the empirical phenomenon in view. (Prediger et al., 2008, p.173)

As presented above, the considered theoretical perspectives have different research foci in this study, however it is possible that they will intersect on some aspects. In particular, as we will see, a few elements of the conceptual framework of this study will be obtained by considering simultaneously two (or more) theoretical perspectives. In these cases, referring to the theories networking, these theoretical perspectives will be (locally) *coordinated*. The term 'coordinating' indicates the networking strategy which corresponds to the simultaneous use of more theoretical perspective to describe, analyse, or interpret the same phenomenon. More specifically,

> We use the word *coordinating* when a conceptual framework is built by well fitting elements from different theories. (Prediger et al., 2008, p.172)

This strategy will be used only locally in this study and when it is done, the motivations for the use of such a strategy will be explicitly presented together with the elements supporting it. As described by Prediger et al., (2008), in fact,

> Applying the strategy of coordinating usually should include a careful analysis of the mutual relationship between the different elements and can only be done by theories with compatible cores (p.172)

Examples of coordination of theoretical perspective in this thesis are the use of the historical-epistemological analysis together with the APOS theory to provide an analysis of the MI from a logical and cognitive point of view, or the use of the construct of Theorem as a triplet (Mariotti et al., 1997) together with the Fischbein's notion of intuitive acceptance of a theorem (1982).

Let us observe that the networking strategies of *combining* and *coordinating* is proper of those studies in which, instead of using a single theoretical framework, a conceptual framework composted by several perspectives is adopted. Moreover, in these cases, as for this thesis, the aim of this choice is not to create a unitary theory obtained by integrating the several perspectives but to obtain a multi-faced analysis for a research problem:

> The networking strategies of combining and coordinating are typical for conceptual frameworks, which do not necessarily aim at a coherent complete theory but at the use of different analytical tools for the sake of a practical problem or the analysis of a concrete empirical phenomenon. (Prediger et al., 2008 p. 172).

To summarise, thus, the research problem of this study will be mainly addressed by combining the just described different perspectives, each of them focusing on a particular aspect of MI. However, a few local points the framework will be obtained by coordinating elements of different perspectives.

1.6 STRUCTURE OF THE THESIS

This thesis can be mainly divided into three parts.

In the first one, the Conceptual framework of the study is described, starting from a historical epistemological analysis of the proof by MI (Chapter 2), then presenting the other theoretical perspectives considered in the thesis (Chapters 3-6), and finally summarising the resulting development of the whole framework, within which the research questions are formulated (Chapter 7).

In the second part, the empirical studies constituting this thesis are described, firstly focusing on the methodological choices (Chapter 8) and then presenting, analysing and discussing the results. In particular, Chapter 9 is dedicated to the first study of the thesis, involving an online survey with university students from different academic courses, while Chapters 10-12 are dedicated the second empirical study, involving some task-based interviews with expert and less expert university students in mathematics and physics.

Finally, in the last part, the conclusions of the thesis are discussed, providing an answer for the research questions, highlighting the main contributions that this study offers, and describing limits of the study and possible didactical implications and directions for further research (Chapter 13).

1.7 Ethical Rationale

The study has involved several moments of data collection, with different methods (task-based interviews and online surveys) which will be presented in detail in Chapter 8. In all the cases, however, the same deontological principles have been adopted. The involved subjects were unpaid volunteers, they were aware that their answers were recorded and how (depending on the times, with audio-video recordings and collecting what they wrote or saving their answers to an online survey). Subjects were informed that the elaborated data would have been used anonymously. Moreover, they were aware that they were participating to a study of mathematics education, but they were not informed on the specific content of the study (MI). All the names of the subjects reported in this study will be pseudonyms. For what concerns the interviews, in which audio and video were recorded, the involved subjects agreed by signing a consent form to the use of the collected materials for the purpose of research and of its presentation. In Appendix B of this thesis, the text of the consent form is shown.

2 A HISTORICAL-EPISTEMOLOGICAL ANALYSIS

2.1 PREMISES

The first point of view which this study takes into consideration is the historical-epistemological one. Before presenting the conducted analysis, I wish to clarify the motivations at the base of the use of such analysis in a study of mathematics education.

I would like to start presenting the position on this aspect of Hans Freudenthal an influent voice within the mathematics education community. As summarised by Van den Heuvel-Panhuizen and Drijvers "Freudenthal considered mathematics as a human activity" (2014, p. 522). This vision of the mathematics brings as a consequence that its epistemology is strongly influenced by the history in which mathematics, as a human activity, has existed and exists. The first two chapters of Freudenthal's book *Mathematics as an educational task* (1973) focus, respectively, on the thousand-year-old radices of the modern mathematics and, at the same time, on the deep changes that have occurred in the mathematics during the years of Freudenthal's own scientific activity. The author seems to highlight that mathematics has a sort of dual nature: mathematics as a process, which continuously develops together with the history in which it is embedded, and mathematics as a product, that is what the paradigms of a specific historical and cultural context have considered to be mathematics. These two faces of mathematics seem to be in a continuous and mutual dialog: from one side the product is determined by the previous historical process and from the other side the paradigms which characterise a certain mathematics as a product influence the directions of its development. Considering this, therefore, Freudenthal's point of view is that history and epistemology of mathematics are deeply intertwined.

This perspective, thus, supports the first hypothesis from which the analysis presented in this chapter has started: **a historical analysis of the genesis of MI could provide indications on its epistemology**. From now on, therefore, in this study I will refer to this analysis with the adjective 'historical-epistemological'.

Once accepted this point, we may ask ourselves which contribute could provide such an analysis for a study in mathematics education. To address this point, I wish quote an essay by Luis Radford, (1997). In a section of this work, the author reflects on the connections between the historical-epistemological development of a mathematical concept and its cognitive-didactical development in the teaching and learning processes. Radford observes:

> Although, at a first glance, it seems evident that historical mathematical developments must have something to inform us about the difficulties that modern students encounter when they learn mathematics, a closer look at the situation reveals that it is far from easy to link both domains - the historical and the psychological. (1997, p. 28)

Radford from one side distances himself from those who see a substantial analogy between the learning trajectory of a concept for a subject (ontogenesis) and the historical development of the concept (phylogenesis), however, from the other side, he observes that a historical-epistemological analysis could still provide an important contribute for the research in mathematics education. Indeed, he writes:

> I think that if, as the socio-cultural perspective suggests, knowledge is a process whose product is obtained through negotiation of meaning which results from the social activity of individuals and is encompassed by the cultural framework in which the individuals are embedded, the history of mathematics has a lot to offer to the epistemology of mathematics. Indeed, historical-epistemological analyses may provide us with interesting information about the development of mathematical knowledge within a culture and across different cultures and provide us with information about the way in which the meanings arose and changed; we need to understand the negotiations and the cultural conceptions that underlie these meanings. The way in which an

> ancient idea was forged may help us to find old meanings that, through an adaptive didactic work, may probably be redesigned and made compatible with modern curricula in the context of the elaboration of teaching sequences. (ibidem, p. 32).

Therefore, interpreting Radford's words, a historical-epistemological analysis could allow us to observe with modern eyes a mathematics which is proper of other historical, social, and cultural contexts. This observation could provide a new light through which (re-)observe and (re-)interpret teaching and learning processes within the context in which they occur.

This brings us to the second hypothesis underlying this chapter: **a historical-epistemological analysis of the genesis of MI could provide elements to interpret processes of use, teaching, and learning of MI.**

These premises were necessary to frame this analysis within the study presented in this thesis. In the following section I will introduce the conducted historical epistemological analysis of MI, focusing on its contents and on the methodological choices.

2.2 Introduction

At the beginning of the 20th century, within the mathematical community it was generally accepted that the first mathematician to have used MI to prove a proposition was Pascal in his *Traité du triangle arithmétique* (1654). This opinion was certified by the influent voice of Moritz Cantor in his *Vorlesungen über Geschichte der Mathematik:*

> The time has come in which the insufficient induction had to surrender its place to the so-called "Complete Induction", or to put it differently, the invention of the proof from n to n+1 was about to happen, and it was Blaise Pascal who delivered it. (Cantor, 1880, p. 684, translated by me).

However, as Bussey (1917) informs us, something was going to change. The mathematician and science historian Giovanni Vacca, at that time one of Giuseppe Peano's assistants in Turin, communicated privately to Cantor that already a few years before in Peano's *Formulario Matematico* (1895) there was a note in which it was claimed that the father of the modern MI was indeed Francesco Maurolico who had used a sort of MI in his *Arithmeticorum libri duo* of 1575.[14] Cantor informed the colleague Florian Cajori who was interested in writing a paper on the history of MI. Cajori, thus, wrote to Vacca whose answer was finally published in 1909 (Vacca, 1909). This publication started circulating internationally.[15]

Vacca's (1909) and Cajori's (1918) publications had the effect of opening a historical debate on the paternity of MI. Several and influent voices have taken part to this discussion during the last century. This debate, however, was as much fruitful as not conclusive. Several traces of the use of MI in the history of mathematics have been proposed by different historians but, at the same time, other scholars have expressed doubts on the fact that what proposed had sufficient characteristics to be properly considered as mathematical induction.

The aim of this chapter is to use the just presented debate to highlight a series of moments in the history of the western mathematics that, following the opinion of some scholars, have contributed to the genesis of the modern MI. In other terms, in this chapter I will analyse the different traces of MI that have been identified in the historiographic debate aiming not to support one opinion against the others, but to highlight those characteristics which convinced a researcher of the presence of MI in a source or those ones which, since missing, convinced another researcher to say that what proposed was not MI.

[14] Cantor himself will tell this thing in 1902 in his *Zeitschrift fur Mathemaschen und Naturwissenschaftlichen Unterricht,* Vol XXXIII p. 536.

[15] For a focus on this specific part of the story, see Moscheo (2011) in which part of the correspondence between the above quoted scholars is also presented.

To summarise, thus, in this chapter I will investigate the following question:

What mathematical aspects characterised the historical genesis of the proof by MI? In particular, what characteristics and turning points emerge from the traces of proofs by MI that the historiographic research has identified?

Firstly, the books VII-VIII-IX of Euclid's *Elements* will be analysed since most of the historical debate focused on these ones (Fowler, 1994; Heath, 1956; Mueller, 1981; Unguru, 1991, 1994; Vitrac, 1994). Then I will examine an extract from Plato's dialog of *Parmenides* addressed by Acerbi (2000), which will be followed by an analyses of the *Sorites paradox*, attributed to Eubilides of Miletus, which had an influent diffusion in the history of western logic. After this, I will consider some propositions from the texts from which the historical debate started: Maurolico's *Arithmeticorum libri duo*, with reference to some studies from the literature (Vacca, 1909; Bussey, 1917; Cassinet, 1988; Pasquotto, 1998; Moscheo 2011), and Pascal's *Traité du triangle arithmétique.* Lastly, I will analyse a letter by Fermat in which he describes the proving method known with the name of *Infinite descent* and that have been quoted in relation to the genesis of MI (Bussey, 1918, Ernest 1982).

Before continuing it is necessary to observe that a few traces of induction that have been identified by the historiographic research will not be presented in this analysis. In particular I am referring to Rashed's study (1972) and to Rabinovitch's study (1970) investigating on the possible presence of mathematical induction respectively in the Arabic and in the Hebraic traditions. This choice was made since both the authors enters the historical debate by saying that if the paternity of MI was attributed to Maurolico, then we should also take into consideration, respectively, al-Karaji and as-Samaw'al, and Levi Ben Gershon who, following Rashed and Rabinovitch, proved similar proposition of Maurolico's and in a similar way. Therefore the two studies, although having an important role in the historiographical debate, for what concern the aim of my analysis do not provide further elements in addition to what I will already highlight in relation to Maurolico. For this reason, these sources will not be presented in the analysis.

I conclude this introduction with two methodological considerations:

1) Each author has been analysed as it follows: firstly each text has been read looking for traces of mathematical induction; only then the quoted studies have been analysed and what discussed in those have been compared with my personal analysis of the tests. The presented analysis therefore emerged as a critical rereading of the studies to which a personal interpretation of the analysed texts is added.
2) Texts have been analysed in the original language, when possible, except for the text in Greek. For those ones, I relied on well accepted translations that will be indicated in the chapter. The texts in this chapter will be presented in English. This translation, if not differently indicated, was made by me.

2.3 Euclid – Elements VII-VIII-IX

The books VII-VIII-IX of the Elements are traditionally called the *arithmetical books.* They treat about numbers (i.e., natural numbers in which, as we will see, the unity is excluded), ratios between them, divisibility, powers, and geometrical progressions.

For the analysis of the text, I relied on two classic translations, respectively the English one by Heath (1956) and the French one by Vitrac (1994), both dependent on Heiberg's critical edition of the beginning of the last century.

The tradition of the Euclidean text, both the direct and the indirect one, is exceptionally complex due to its wide diffusion and this makes the construction of a *stemma codicum* extremely difficult. Moreover, the text is particularly stratified, an aspect which makes the attribution of the effective paternity of a sentence very problematic. These philological problematics, however, go beyond the purpose of this chapter. In this study, therefore, I will not deal with the philological aspects of the Element but only with the contents of the analysed text. In other terms, I will analyse the contents of a proposition or of a definition independently if it, or a part, is attributed to Euclid or not. From now on, thus, I will freely refer to Euclid as the author of the propositions even if I am aware that, from a historical point of view, this could be an inaccuracy.

In the appendix A of this thesis is presented a list of all the definitions of the book VII and all the statements of books VII-VIII-IX.

2.3.1 A first problem

The book VII is opened by 22 definitions that will be used during the following three books. Focusing on this definition it is possible to observe some interesting elements from our analysis. The first two definitions are the following:

<u>Definition 1.</u> *An unit is that by virtue of which each of the things that exist is called one.*

<u>Definition 2.</u> *A number is a multitude composed of units.*

A first and important consequence of these two definitions is that the unit is not a number (Mueller, 1981; Vitrac, 1994). This aspect has evident effects in several propositions involving the unit 1, for instance the propositions on geometrical progressions with 1 as first term (Propp. IX.8-13), in which it is treated differently from all the other terms of the sequence. Let us consider, for example, the proposition IX.8. The first part of the statement is the following:

If as many numbers as we please beginning from an unit be in continued proportion, the third from the unit will be square, as will also those which successively leave out one. [...]

The proof starts in this way:

Let there be as many numbers as we please, A, B, C, D, E, F, beginning from an unit and in continued proportion.

The number A, however, is not the unit, but it is the second element of the sequence. A letter is not assigned to the unit during the whole proposition, Euclid refers to it as '*μονάς' (unit)*, moreover in the figure corresponding to the proposition, the segment representing the unity is not drawn.

The same thing happens in several other propositions. There are however a few exceptions. Sometimes, in fact, the unity is indicated with a letter, and it is present in the figure attached to the proposition. For instance, in VII.15, which in modern formulation states that 'If $\frac{1}{n} = \frac{k}{kn}$, then $\frac{1}{k} = \frac{n}{kn}$', the unit is called A and it is used to measure a segment. However, it is interesting to notice that in a previous proposition (VII.13), Euclid has already proved that 'if $\frac{m}{n} = \frac{k}{kn}$ then $\frac{m}{k} = \frac{n}{kn}$' which would directly imply the VII.15 considering m=1. Euclid however proves the VII.15 without any reference to the previous proposition as if it needed to be considered as a different case.

A direct reference to the unit (with a letter indicating it and a drawing) is also present in VII.37 and in IX.9-10, which involve a proportion between the unit and three numbers, and in IX.20 in which the unit is added to the product of a series of numbers.

It seems that Euclid confers to the unit the role of number, designating it with a letter and drawing it together with the other numbers, only when it is operatively used in the proof of the proposition in some calculations. Anyway, even when this happens, Euclid keeps specifying that it is the unit (he writes '*the unit A', 'the unit E', 'the unit DF'*, …), as to highlight its different nature from the other numbers.

Let us focus now on the second definition of book VII:

Definition 2. A number is a multitude composed of units.

We can preliminarily observe that the term used by Euclid to indicate this composition of unity is '*πληϑος*' which both Heath and Vitrac translate as 'multitude', respectively in English and French. Vitrac, moreover, observes that the term '*πληϑος*' should be compared with the other Greek term '*ποσόν*', considering that the first one is generally used to indicate an indeterminate quantity while the second one a determinate quantity. Following Vitrac, thus, Euclid seems to use this term to express a genericity in its definition, referring to an indeterminate multitude of units.

A second aspect to highlight is what noticed by Mueller (1981) who observes that Euclidean definition of number reflects an important characteristic of the Greek arithmetic which distances it from the modern one: in the modern arithmetic the unit is the natural number 1 and all the other natural numbers are obtained with progressive applications of the successor operation starting from 1, in the Greek arithmetic, instead, a number is obtained as a multitude of units moreover "there are indefinitely many units and indefinitely many ways of combining them into multitudes. Clearly then, there is no unique 2 or 3" (Mueller, 1981, p.60). In other terms, the generic number *N* (or, following what said, all the possible numbers *N*) is obtained by grouping together *N* units (and not by adding 1 to *N*-1). Surely this does not mean that Euclid was not aware of the successor-predecessor relationships between natural numbers, but that the successor operation has not the same constructive and structural role for the natural numbers as it happened for the modern arithmetic of natural numbers.

Mueller attributes this characteristic of natural numbers to the whole Greek arithmetic, however we can find examples in which this seems to be different. For instance, in a note for the definition of above, Heath writes:

> Theon[16], in words almost identical with those attributed by Stobaeus […] to Moderatus, a Pythagorean, says […]: "A number is a collection of units, or a progression (προποδισός) of multitude beginning from an unit and a retrogression (αναποδισμός) ceasing at an unit. (1956, Vol.2, p. 280)

This second definition of number, differently from the Euclidean one, seems to refer to a progressive structure for the natural numbers more similar to the modern one which was described above.

These observations in relation to the definition of numbers in Euclid's *Elements* highlight a first problem. The role of the successor operation, fundamental for the recursive definition of natural numbers and consequently for MI, is not central in Euclid. Quoting Mueller (1981):

> [N]umbers are not characterized as generated from units in a serial order. They are simply finite aggregates of units. Of course, there are important relations between these aggregates, but the relation of successor is not one which plays an important role in the Elements. Thus it can be said that the integers themselves are not conceived in the structural way conducive to the use of induction, but that there is inductive reasoning about collections or sequences of positive integers. (p. 70).

[16] To avoid misunderstandings, the reference is to Theon of Smyrna, lived between the 1st and the 2nd century AD, and to his *Expositio rerum mathematicarum ad legendum Platonem utilium,* not to Theon of Alexandria, author of a fundamental revision of Euclid's *Elements*.

Therefore, this could represent an obstacle for the presence of traces of MI in Euclidean arithmetic since it seems to miss, in its foundations, that recursive structure which is fundamental for MI. However, as highlighted by Mueller, in some propositions it is still possible to observe some elements which seem to have a connection to MI, or, for saying with Mueller, a trace of "inductive reasoning". These propositions will be analysed in the following section.

2.3.2 A classification

It is possible to identify three categories of propositions among those that I will present. This classification is an attempt to group the propositions considering, in the statement or in the proofs, those elements which can be read as traces of MI. In particular:

1) The first category includes those propositions or couples of propositions in which a given property or construction is proved for two numbers and then the same thing is proved for three numbers by connecting this case to the previous one.
2) The second category includes those propositions in which the statement involves an arbitrary quantity of numbers. For those ones the proof is conducted for a specific quantity of numbers (from 2 to 6 numbers) in a way that it can be generalized to any given quantity.
3) The third category included those propositions which involve the well-ordering principle, in slightly different formulation than the modern one.

<u>First category</u>

This category, includes those propositions or couples of propositions with the form P(2) and P(3) where P(n) is a given property or construction involving *n* numbers and in which P(3) is proved by showing P(2)→P(3).

Probably the clearest example of a couple of proposition of this category is VII.2-3 where it is shown, in the first one, how to find the *Greatest Common Divisor* (*GCD)* of two numbers *a* and *b*, using the famous Euclid's algorithm, and in the second one how to find it for three numbers *a*, *b*, *c*, by showing that, in modern notations, *GCD*(*a,b,c*)=*GCD*(*GCD*(*a,b*), c) and then using the previous proposition. The same thing happens in the propositions VII.34-36 in which it is shown how to determine the *Least Common Multiple (lcm)* between two numbers and then between three numbers using the previous case. Euclid shows only the case P(2) and P(3), with no explicit reference to a possible generalisation for greater given quantity of numbers, however it is not unreasonable to believe that Euclid is implicitly showing hot to generalise the process to greater quantities of numbers. For saying it with Mueller: "It would seem perverse to deny Euclid's intention is to convey a general result" (1981, p.69).

To support this last point, Mueller observes that Euclid in the proof of VII.33 uses the GCD of an arbitrary quantity of numbers even if, previously, he only showed how to obtain the GCD for three numbers. The proposition VII.33 states: '*Given as many numbers as we please, to find the least of those which have the same ratio with them*' and it belongs to the second category of the analysed propositions. While the statement is expressed in general terms ("given as many numbers as we please") the proof is constructed for three numbers. As a consequence, when during the proof the GCD of the considered numbers is introduced, it involves only three numbers as in VII.3. However, these ones act as a generic example representing an arbitrary quantity of numbers (as indicated by the statement of the proposition), therefore in this application of VII.3 Euclid is implicitly referring to its generalisation to "as many numbers as we please".

In this category there are also two propositions (VII.27 and VIII.13) in which a generalization is explicitly present in the statement. For the two propositions, both Heath and Vitrac, following Heiberg's comment, express doubts about the real paternity of the sentence which generalises the statement. Let us analyse the first one; the second one is completely analogous.

Proposition VII.27. *If two numbers be prime to one another, and each by multiplying itself make a certain number, the products will be prime to one another; and, if the original numbers by multiplying the products make certain numbers, the latter will also be prime to one another [and this is always the case with the extremes].*

The statement can be expressed in modern notation as it follows:

If two numbers a and b are relatively prime then also a^2, b^2 are relatively prime, and a^3, b^3 are relatively prime [and, in general, a^n, b^n are relatively prime].

In the proof, firstly it is shown that if a and b are relatively prime then a^2, b^2 are relatively prime (which is a direct consequence of a previous proposition, VII.25) and then, starting from this conclusion and still applying the VII.25, it is proved that a^3, b^3 are relatively prime as well. This proposition belongs therefore to the first category of our analysis since a successive case (a^3, b^3 are relatively prime) is proved using the previous one (a^2, b^2 are relatively prime). We can notice that by iterating this strategy one could prove that in general a^n, b^n are relatively prime. This part, however, is not present in Euclid's proof even if it was indicated in the statement. This aspect, together with a linguistic analysis of the sentence between square brackets brought Heiberg to conclude that the sentence generalising the proposition is probably an interpolation of date earlier than Theon. In any case, independently from who wrote it, however, it is interesting to notice is that the proof made by Euclid showing a connection between two consecutive cases has been seen as generalisable to a general case.

Second category

This second group of propositions include those ones in which the statement refers to an arbitrarily great quantity of numbers. Most of the times the statement starts with the sentence: '*If there be as many numbers as we please […]*'.

As we will see, despite the general statement, the proof is conducted for a specific quantity of numbers, generally between two and six numbers. The proof of these propositions opens with the sentence '*Let there be as many numbers as we please, A, B, C, D […]*', if, for instance, four numbers are considered, and it continues with a construction involving them o with an application of a previous proposition to them. The proof therefore involves a specific quantity of numbers. However, it could be easily generalised in modern terms to a generic quantity N. This generalization is never present in Euclid's proof. Let us see some examples.

Proposition VII.12 states:

If there be as many numbers as we please in proportion, then, as one of the antecedents is to one of the consequents, so are all the antecedents to all the consequents.

In modern notations it could be written as[17]:

[17] It is necessary to add some clarifications on the used notations. In this and other following propositions I will translate the text in modern notation by using three suspension dots (…) to indicate the generalisation. This choice is distant from Euclidean style. As described by Unguru (1994), the difference is not simply syntactical, but semantical: the three dots (…), in a modern interpretation, are a representation of an infinity of numbers following the written ones, an idea which is not coherent with the Euclidean context.

If $\frac{a}{a'} = \frac{b}{b'} = \frac{c}{c'} = \ldots$ *then each of these ratios is equal to* $\frac{a+b+c+\ldots}{a'+b'+c'+\ldots}$

Euclid proves the proposition for four numbers for which he has at disposal a series of propositions to use. This aspect will be recurrent for all the propositions of this category: the quantity of numbers involved in the proof directly depends on the propositions that need to be used, and in particular on the quantity of numbers that they involved in their respective proofs. As an example of this aspect, we can consider the proposition VII.39, which involves the following problem:

To find the number which is the least that will have given parts.

The problem is equivalent to find the *lcm* of *N* numbers (i.e., the 'given parts'), which has been already solved by Euclid for the specific case of *N=3* in VII.36. Coherently with this, Euclid proves the proposition VII.39 for the specific case of three numbers, in order to directly apply VII.36: *"Let A, B, C be the given parts."*

In this category there are some propositions in which it is present an iteration. Let us start by analysing the IX.8:

If as many numbers as we please beginning from an unit be in continued proportion, the third from the unit will be a square, as will also those which successively leave out one; the fourth will be a cube, as will also all those which leave out two; and the seventh will be at once cube and square, as will also those which leave out five.

The proposition can be expressed with modern notations as:

Let 1, a_1, a_2, a_3, *… a geometrical progression. Then for every natural number n,* a_{2n} *is a square number,* a_{3n} *is a cube number, and* a_{6n} *is both a square and a cube number.*

The proof develops as it follows:

Let A, B, C, D, E, F as many numbers as we please beginning from an unit and in continued proportion. Since the unit is to A as B is to C, the unit measures the number A the same number of times that A measures B, but the unit measures the number A exactly A times, therefore A also measures B exactly A times. Therefore, A by multiplying itself has made B, thus B is a square.

Since B, C, D are in continued proportion and B is a square, then D is also a square [Prop. VII.22]. For the same reason F is also a square.

Similarly we can prove that all those which leave out one are square. […]

At this point Euclid, with an analogous argument proves the statement for the elements with the form a_{3n} and then he concludes:

[…] Therefore F is also cube. But it was also proved square: therefore the seventh from the unit is both cube and square. Similarly we can prove that all the numbers which leave out five are also both cube and square.

The structure of the proof is interesting. Firstly it is proved that if 1, A, B are in continued proportion, then $B=A^2$. Then the proposition VII.22 is applied ('If three numbers be in continued proportion, and the

---

Quoting Unguru (1994): "Neither Euclid nor other Greek mathematicians had them [the three dots] at their disposal and this is not unrelated, I think, to the lack of mathematical induction in their universe of discourse, which […] reflects a different ontology from the modern one." (Unguru, 1994, p. 271).

first be square, the third will also be square') [18], and the property of being a square number transits from B to D and then, applying another time the VII.22, from D to F. Euclid concludes with 'similarly we can prove that all those which leave out one are square'. We can say that Euclid's proof has some similarities with a proof by induction in which some instances of the inductive step (represented by the proposition VII.22) are iterated for the first cases. This proposition was one of the most discussed within the historiographic debate on the presence, or not, of induction in Euclid's arithmetic (see Unguru, 1991; Frowler, 1994; Unguru 1994). One of the most sceptical positions is kept by Unguru (1991, 1994), who writes:

> Euclid does not appear to have felt that he had given a demonstration covering all numbers leaving out one all those leaving out five, but rather that he had shown how the proposition could be proved for any assemblage of numbers in continued proportion, the proof in the Elements being a token, an example for any future proofs. And this, to my way of thinking, is not an instance of a proof by mathematical induction, a method that does not stand in need of reinforcement and supplementation by an indefinite number of proofs with identical conceptual structure." (Unguru, 1991, p. 278)

A second interesting example is the proposition IX.12 which, in modern notations, states:

Let 1, a_1, a_2, a_3, … be a geometrical progression. If the number a_n is divided by the prime numbers p_1, …, p_m, then a_1 is also divided by p_1, …, p_m.

The proof (conducted for n=4) develops by showing that if a prime p divides a_4 then p also divides a_3 and, from this, then p also divides a_2 and, from this, then p also divides a_1. Each of the previous implications is obtained as a consequence of the proposition IX.11, which in modern notations states that in a geometrical progression with ratio k, $a_n=k\cdot a_{n-1}$. Then Euclid concludes that the same thing can be proved for any prime which divides a_4.

This proof has an interesting structure with an iteration which involves decreasing terms of the geometrical progression. In particular a chain of deductions is constructed, in which every step, obtained with the proposition IX.11, ends on the term of the progression from which the following step will start. Moreover, it is interesting that Euclid repeats the same words and uses the same structure of the sentence for describing each step. This aspect, as we will see, will be registered also in the following author of this chapter, when they deal with a proof which involves an iteration.

The last proof that I present for this category is the famous IX.20:

Prime numbers are more than any assigned multitude of prime numbers.

In the proof, Euclid shows how, starting from a finite set of primes, one could construct a prime number which does not belong to the given set. The proof starts as it follows:

Let A, B, C be the assigned prime numbers. […] Let DE be their lcm and let the unit DF be added to DE, obtaining EF. Then EF is either prime or not.

If it is prime, then A, B, C, EF are prime numbers and they are more than A, B, C.

If it is not prime, then it is measured by some prime number G [VII.31].

[18] Let us note that, as observed at the beginning of this section, the unit necessitates a different treatment than the other numbers, so the proposition VII.22 is not applied directly to 1, A, B to conclude that B is a square number.

At this point Euclid shows, by absurd, that G must be different from A, B, C, and he concludes:

Therefore, the prime numbers A, B, C, G have been found which are more than the assigned multitude of A, B, C.

Euclid's proof can be reread in a modern key in three different ways:

1) Euclid is providing an algorithm which generates prime numbers, all different. The proof contains the (generic) example of an instance of this algorithm, in particular showing how, starting from three prime numbers, to construct the fourth one. The set of prime numbers, therefore, is infinite[19] because the algorithm could potentially never end.

2) The proof can be seen as a proof, with a generic example, of an inductive step. In fact, we could express the proposition as $\forall n \in \mathbb{N}$, *P(n)* where *P(n)* is the predicate '*The cardinality of the set of prime numbers is at least n'*. In this way the proposition could be proved by induction and Euclid's argument could represent a generic proof for the inductive step ($P(3) \rightarrow P(4)$).[20]

3) Considering *P* as the set of prime numbers, Euclid's proof shows that, for a given and arbitrary $n>0$, $|P| = n \rightarrow |P| = n+1$ which is impossible if $n < \infty$. This is, in substance, the structure of the classic proof by absurd of the infinitude of prime numbers.

Let us observe that all the three interpretations (and Euclid's proof itself) formally work with the implicit assumption that at least a prime number exists.

Before concluding the presentation of the second category of proposition, let us notice that, in four propositions (IX.21-13), Euclid does not use an iteration to prove them although they could be proved in this way. I will present the first one as a paradigmatic example:

If as many even numbers as we please be added together, the whole is even.

The proof does not develop by iteratively using the fact that the sum of two even numbers is an even number but considering a generic quantity of even numbers to be added together (four generic even numbers, in Euclid's proof) and then proving the statement directly for them.

In conclusion, to prove the propositions in this second category Euclid seems to use different strategies. Sometimes the general statement is proved for a specific quantity of numbers without any explicit conclusion to generalise the proof to a generic quantity of numbers (as in the first two shown examples). Some other times the general statement is proved constructing an iteration which is finite (as when it involves the decreasing position of the terms of a progression) but which in some cases (as in the third and fourth examples), even if Euclid does not refer to it in these terms, it could be interpreted as potentially infinite, when it can be used to reach a generic number greater than the base of the iteration.

[19] Let us observe that the sentence 'the set of prime number is infinite' is a modern interpretation of the proposition which is distant from the Euclidean context because it involves an idea of infinity which is actual and not only potential as in the Greek mathematics. (Heath, 1956; Mueller, 1981; Unguru, 1991; Vitrac, 1994).

[20] This second interpretation is also possible only in a modern key. In particular the use of functional notation for the predicate *P(n)* and the presence of a parameter (*n*) on which to apply MI are distant from the Euclidean context. This point is central for those scholars that critique the possibility of the presence of induction in Euclidean arithmetic. Among those, I quote Unguru (1991): "Without a number that can serve as an independent variable, it is impossible to formulate a true proof by mathematical induction, in which the claim requiring proof is a function of the natural numbers" (p.284).

Finally, other times, Euclid does not construct an iteration to prove a statement, even if it was possible, as in the last presented example.

Third category

This last group of propositions involves those cases which involve the well-ordering principle[21], as we will see, in slightly different formulation than the modern one. Considering the logical equivalence between the well-ordering principle and the principle of 'strong' mathematical induction[22], these propositions could provide useful elements for the purposes of our analysis.

The first two propositions of the 7th book involve the famous Euclid's algorithm for determining the GCD of two numbers. The algorithm develops with a series of successive divisions in which, at every step, the divisor becomes the successive dividend ant the remainder the successive divisor, until we obtain a last division with zero as remainder. The remainder of the penultimate division is the GCD of the two numbers. It is interesting to notice that the finiteness of the algorithm is guaranteed by the fact that the sequence of the remainders is strictly decreasing and thus must be finite. This point is a consequence of a principle, often called the *Infinite descent principle*, which states that any strictly decreasing sequence of natural numbers must be finite. Moreover, the infinite descent principle is formally equivalent to the well-ordering principle[23]. Thus, Euclid seems to be implicitly assuming this principle which, from a logical point of view, is extremely related to the principle of MI.

In VII.31 the connection with the *Infinite descent principle* is more evident. The proposition states:

Any composite number is measured by some prime number.

The proof develops as it follows. We take a generic number A which, since it is composite, is divided by a lower number B. If B is prime, the proof is concluded. Otherwise, B is composite and therefore it is divided by a number C, which, since it divides B, also divides A. Therefore, if C is prime the proof is concluded, otherwise it is divided by a lower number D. At this point Euclid writes:

Thus, if the investigation be continued in this way, some prime number will be found [...] For if it not found, an infinite series of numbers will measure the number A, each of which is less than the other: which is impossible in numbers.

In other terms Euclid explicitly affirms that it cannot exist an infinite sequence of natural numbers which is strictly decreasing, that is the principle of infinite descent.

Mueller (1981) suggests that the assumption of the well-ordering principle can be also registered in the proposition VII.20, this time not in the formulation of the infinite descent principle. The proposition, in modern terms, states:

If two numbers A and B are the lowest having the ratio $\frac{A}{B}$*, any other couple of numbers C and D with* $\frac{C}{D} = \frac{A}{B}$ *can be written as C = M·A and D=M·B for some M>1.*

The presence of the well-ordering principle can be seen not in the proof but in the statement itself, in particular in the fact that "the least numbers of those which have the same ration with them" (A and B, in our notations) exist. In other terms we are assuming that, given the set formed by the couples of

[21] I consider the following formulation: $\forall \mathrm{A} \subseteq \mathbb{N},\ \mathrm{A} \neq \emptyset \rightarrow\ \exists \mathrm{a} \in \mathrm{A},\ \mathrm{a} = \min(\mathrm{A})$.
[22] See Lolli (2008).
[23] For a proof, see Lolli (2008).

natural numbers having a fixed ratio, there exists a couple in which the first and the second number are the least between respectively the first and the second numbers of all the other couples of the set.

In conclusion the propositions of this third category seem to suggest that Euclid assumes in his *Elements* the well-ordering principle as an implicit axiom. As observed by Mueller (1981), however, despite the strong connection from a logical point of view, between this principle and the principle of mathematical induction, there is one important difference between the two of them:

> Unlike the principle of induction, the least number principle or the denial of infinitely descending chains does not seem to depend upon a genuinely structural conception of the positive integers. For these principles can be understood solely in terms of the Euclidean conception of numbers as finite concatenations of unit. (p.78).

Following Mueller, this aspect could contribute to understand why the principle of MI, even in a *naïve* formulation, does not seem to be present in Euclid, while other extremely related principles are instead present.

2.3.3 Conclusions

In the previous section I have presented a series of propositions which for different reasons are related to mathematical induction:

1) The proposition of the first category involves the proof of a statement or a construction for two and then for three numbers, using the first case for proving the second one. This strategy reminds, in a certain sense, an instance of the inductive step: P(2)→ P(3) instead of $\forall n \in \mathbb{N}$. P(n)→P(n+1). Euclid does not explicitly generalise the argument to bigger numbers, however in some propositions he seems to be aware that this thing is possible. For instance, in VII.33 he considers the GCD between an arbitrary quantity of numbers even if he showed only how to determine it for two or three numbers.

2) The propositions of the second category involve some iterations. In particular when the statement involves an arbitrary quantity of numbers, after having shown the steps of an iteration for the first cases, Euclid concludes the proof generalising what done to every number involved in the statement. The connection with MI seems to be in this point: in the presence of an explicit iteration which starts from an initial case and reaches 'as many numbers as we please'.

3) Lastly, with the third category of propositions, we observed that Euclid includes between the implicit axioms that he uses in his arithmetic the well-ordering principle (expressed infinite descent principle), which from a logical point of view is extremely related to the principle of MI.

These are the characteristics that brought some historians of mathematics to talk about traces of MI in Euclid's *Elements* (Freudenthal, 1953; Mueller, 1981; Fowler, 1994).

However, as observed, there are a few central points in the Euclidean arithmetic which seems to be very distant from MI, if not even an obstacle for its presence. First of all, the definition of number seems to reveal an Euclidean idea of number constitute as an agglomerate of units and not with a recursive application of the successor operation (Heath, 1956; Mueller, 1981; Unguru, 1991). Secondly, Euclid's idea of infinity is potential and still very distant from the actual infinity one (Mueller, 1981; Unguru, 1991, 1994; Vitrac, 1994). If proving by MI means, among other things, proving something for an infinite number of cases, this is something that cannot be present in Euclid. Finally, the absence of the algebraic symbolisms brings as a consequence the absence of a parameter (and of a parametric expression for the predicate P(n)) on which to apply induction (Unguru, 1991, 1994). Following Unguru, these three characteristics are sufficient for concluding "the historical impossibility of the existence of genuine inductive proof within the confines of Greek mathematics" (Unguru, 1991 p.273).

As said, the debate was open for all the previous century, reaching perhaps its apex in the sharp papers exchange between Unguru and Fowler on the journal *Physis* at the beginning of the '90s (Fowler, 1994; Unguru, 1991, 1994). This debate was later reopened by Acerbi (2000) who "to increase the confusion on this subject" (p.58), proposed a passage from Plato's *Parmenides* as an example of the possible presence of an ancient trace of MI and that will be analysed in the following section.

The following table contains the list of the propositions of the arithmetic book of Euclid's *Elements* which belong to the three proposed categories.

First category	**Second category**	**Third category**
- VII,2-3 - VII,5-6 - VII,27 - VII,34-36 - VIII,13 - VIII,18-19 - VIII,20-21	- VII,12; N=4 - VII,14; N=3 - VII,33; N=3 - VII,39; N=3 - VIII,1; N=4 - VIII,2; N=4 - VIII,3; N=4 - VIII,4; N=3 - VIII,6; N=5 - VIII,8; N=2 - VIII,9; N=2 - VIII,10; N=2 - IX,8; N=6 - IX,9; N=6 - IX,10; N=6 - IX,11; N=5 - IX,12; N=4 - IX,13; N=5 - IX,20; N=3 - IX,21; N=4 - IX,22; N=4 - IX,23; N=3 - IX,32; N=3	- VII,1-2 - VII,21 - VII,31

Table 2.1. Propositions of books VII-IX belonging to the three categories above presented. For the second category of propositions, which involve in the statement an arbitrary quantity of numbers, I also indicate, besides the number of each proposition, the specific quantity of numbers for which each poof is made.

2.4 PLATO – *THE PARMENIDES*

The *Parmenides* is a Platonic dialog belonging to the so-called 'later-middle dialogues. It describes a dialog between Parmenides, Zenon, and Socrates. The main theme of the discussion from which the excerpt is taken is *The One,* in the sense of the whole*,* of the entirety. In this part of the dialog, Socrates is arguing that that the One, as such, cannot be 'in contact' (i.e. physically linked) to something else (149 a7-c3). I analysed the text in its Italian translation by G. Cambiano (1981). The above reported English translation is an extension of the translation presented in Acerbi (2000). The parts that I added are the sentences (1) and (9), whose English translation is mine, (2)-(8) are taken without modification from Acerbi (2000, p. 64). The enumeration is introduced to support the following analysis.

"(1) We say that what is in contact, but remaining distinct, must be adjacent to the term with which it must be in contact, and no third term can be between them."
"True."
"(2) Then they must be two, at least, if there is to be contact."
"They must."
"(3) And if to the two terms a third be added in immediate succession, they will be three, while the contacts [will be] two."
"Yes."
"(4) And thus, one [term] being continually added, one contact also is added, (5) and it follows that the contacts are one less than the number of terms. (6) For the whole successive number [of terms] exceeds the number of all the contacts, (7) as much as the first two exceeded the contacts, for being greater in number than the contacts: (8) for afterwards when an additional term is added, also one contact to the contacts [is added]."
"Right."
"(9) Then whatever the number of terms, the contacts are always one less."
"True"
"(10) If instead there is only one term and there is not duality, it cannot be a contact."

After having defined what he means with 'to be in contact' (1), Socrates states that there must be at least two terms in order to have a contact (2). This sentence is repeated in its contrapositive form at (10) and this will be enough to prove what Socrates needs (that the One cannot be in contact to something else). However, in the central part of the excerpt (2-9) seems to prove a different statement: the number of contacts are always one less than the number of terms, as he summarise at the end (9). Socrates states since when there are two terms there is one contact (2) and since if we continuously add an additional term the number of contacts continuously increases by one (4), therefore the number of contacts is always one less than the number of terms (5). Successively (6-8) Socrates seems to generalise the statement saying that the difference between the number of contacts and the number of terms, independently from what it was at the beginning, does not change when adding a term. Finally, Socrates concludes to have proven that "whatever the number of terms, the contacts are always one less" (9).

Acerbi (2000) observes that in this excerpt, if we translate the rhetorical structure of Socrates' argument with an algebraic symbolism, it is possible to register some similarities with a proof by MI. The following is a paraphrase of Acerbi's analysis.

Socrates wants to prove that "the number of contacts is always one less than the number of terms" (5).

Let n be the number of terms in contact by the definition of (1) and let $C(n)$ indicate the number of contacts between these n terms. The statement to prove, thus, becomes '$C(n)=n-1$ for every natural number n'.

Let us follow Socrates' argument using the just introduced notation:

(2) Then they must be two, at least, if there is to be contact.	$C(2) = 1$
(3) And if to the two terms a third be added in immediate succession, they will be three, while the contacts [will be] two.,	$C(3) = 2$
(4) And thus, one [term] being continually added, one contact also is added,	$C(n+1) = C(n) + 1$
(5) and it follows that the contacts are one less than the number of terms.	$C(n) = n - 1$

In the successive lines Socrates proves, with the same argument, a more general statement: if the difference in the number of terms and in the number of contacts would be a different number, the same different would have remain constant when adding an additional term. With the notation of above Socrates' argument is the following:

(7) as much as the first two exceeded the contacts, for being greater in number than the contacts	$2 - C(2) = k$, With *k* arbitrary
(6) the whole successive number [of terms] exceeds the number of all the contacts.	$n - C(n) = k$
(8) for afterwards when an additional term is added, also one contact to the contacts [is added].	$C(n+1) = C(n) + 1$

Acerbi's position is that Socrates' argument contains some of the structural characteristics of a proof by MI. In particular:

(i) The base of the induction is proved (2).

(ii) The inductive step is stated, proved, and its generality is recognised (3), (4), (8).

(iii) The proof's generality is stated (5), (6), (9).

Finally, let us observe that, in addition to what noticed by Acerbi, it is possible to observe the presence of other characteristics that we already registered for Euclid:

a) The generalization of the (inductive) step occurs after some explicit steps. In particular, Socrates shows the truth of the statement for the case n=2 (2) and then for n=3 (3) using the previous case.

b) The generalization occurs by observing the generality of the process through which it was possible to deduce the truth of the statement for three terms using the truth of the statement for two terms. This first step, therefore, acts as a generic example.

In the following section I will consider another excerpt still belonging from the ancient Greek world: *The sorites paradox.* This passage is quoted by Acerbi himself (200, p.72) as a possible further trace of MI in the ancient Greek mathematics.

2.5 The *Sorites paradox*

The *Sorites paradox,* from the Greek 'σόρος', heap, and for this also known as the *Paradox of the heap*, is attributed by Diogenes Laertius (3rd century AD) to the Megarian philosopher Eubilides (4th century BC). The paradox is described by Laertius (1989) as it follows:

> It cannot be that if two is few, three is not so likewise, nor that if two or three are few, four is not so; and so on up to ten. But two is few, therefore so also is ten (p. 191).

In other terms, since adding 1 to a small number does not change it in a big number and because 2 is a small number, the number 10 (or every other number) will never be a big number.

Another formulation, analogous to the previous one, from which probably the paradox takes the name is the following:

- A grain of wheat does not form a heap.
- If a grain of wheat does not form a hep, two grains of wheat do not form a heap either.
- if two grains of wheat do not form a heap, three grains of wheat do not form a heap either.

…

- If 999 999 grains of wheat do not form a heap, a million grains of wheat do not form a heap either.

Therefore:

- A million grains of wheat do not form a heap.

This paradoxical argument had a wide diffusion in the Hellenistic philosophy, in particular within the Sceptical school, and it arrived then to the medieval logic. It encountered a new fortune at the end of the 19th century thanks to some Marxist philosopher and from then raised a considerable interest in philosophy (Hyde & Raffman, 2018; Oms & Zardini, 2019). From a logical-mathematical point of view the paradox has been analysed both by Frege and Russel and 'solved' by saying that the predicate '*to be small'* or '*to form a heap'* have a vague attribution of truth and therefore are not acceptable as legitimate predicates in the formal classical logic. The problem, however, was reopened in more modern times, in relation to *fuzzy* logic in which intermediate truth-false values are considered.

Independently from its formal logical analysis, we can observe that this paradox is structured in an argument in which some similarities with MI are present. In particular let us analyse the structure of the argument referring it to a generic predicate *P(n)*. The structure of the argument can be represented:

- With an iterative chain of *modus ponens* reaching an arbitrary natural number *n>1:*

$$
\begin{array}{l}
P(1) \quad \wedge \quad P(1) \rightarrow P(2) \\
\overline{\qquad\qquad\qquad\qquad} \\
\qquad\qquad P(2) \quad \wedge \quad P(2) \rightarrow P(3) \\
\qquad\qquad \overline{\qquad\qquad\qquad\qquad} \\
\qquad\qquad\qquad\qquad P(3) \quad \wedge \quad P(3) \rightarrow P(4) \\
\qquad\qquad\qquad\qquad \overline{\qquad\qquad\qquad\qquad} \\
\qquad\qquad\qquad\qquad\qquad\qquad \ldots \\
\qquad\qquad\qquad\qquad\qquad\qquad \overline{\qquad\qquad\qquad\qquad} \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad P(n-1) \ \wedge \ P(n-1) \rightarrow P(n) \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \overline{\qquad\qquad\qquad\qquad} \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad P(n)
\end{array}
$$

- With a Mathematical Induction form:

$$
\frac{\begin{array}{c} P(1) \\ \forall \mathrm{n} \in \mathbb{N}, P(n) \rightarrow P(n+1) \end{array}}{\forall \mathrm{n} \in \mathbb{N}, P(n)}
$$

The representation which is closer to Laertius' formulation is the first one. In the second formulation the series of steps are condensed in a unique one (i.e. the inductive step). Moreover, another important difference is that in the first one there is an idea of infinity which is potential: given a number *n*, with a finite number of steps it is possible to show P(n). In the second one, however the expression $\forall$ *n* ∈ ℕ reveals the reference to an idea of infinity which is actual: it is the whole entire and infinite set of ℕ which to be considered.

However, despite the differences between Laertius's formulation and the one explicitly involving MI, a new element not registered yet in the previous authors seems to be present: the chain of logical inferences which transmits the truth of the statement P(1) to the statement P(n) where n is an *arbitrary natural number greater than 1*. It is exactly this arbitrariness of number of steps what creates the paradoxical conclusion of the argument. It is not paradoxical that 'ten is few' or that 'a million grains of wheat do not form a heap', but the fact that *continuing the argument* it is not possible to find a number which is 'not few' or a number of grains of wheat that form a heap. In other terms to perceived as paradoxical the conclusion of the argument it is necessary to recognise that it is not possible to find a number which cannot be reached by the chain of modus ponens.

In the following section, with a jump of several centuries, I will analyse some excerpts from the works of the Sicilian mathematician of the 16th century Maurolico that, at the beginning of the last century were quoted by Vacca (1909) and Bussey (1917) as examples of a first use of mathematical induction to prove a statement.

2.6 MAUROLICO – THE *ARITHMETICORUM LIBRI DUO*

Francesco Maurolico (1494 – 1575) writes in 1557[24], already sixty years old, a compendium of arithmetic, the *Arithmeticorum Libri Duo,* that will be later published in 1575 in the collection *Opuscula Mathematica*[25].

In the book a proper theory of (natural) numbers is developed. The starting point, following the Euclidean tradition, is given by a series of definitions. Already in the first definition, the one of unit and number, however, Maurolico seems to disagree with Euclid:

The unity is the principle and it constitutes every number, moreover it firstly constitutes itself. [26]

Differently from what happened in the *Elements*, for Maurolico the unit also constitutes the number formed by the unit itself (the one). Following this definition, thus, it is now licit to talk about *the number one*. However, it continues having a special nature compared to the other numbers:

Each of these numbers, thus, is either the unit, which corresponds to the point […], or it is linear […], or it is plane […], or it is solid.[27]

After this definition, Maurolico explains how the unit can construct different kind of numbers and how these numbers, considering how they are constructed, are denoted. In figure 2.1 we can observe, as examples, how the unit can be used to construct the radices (linear numbers), the triangular and the square numbers (kinds of plane numbers).

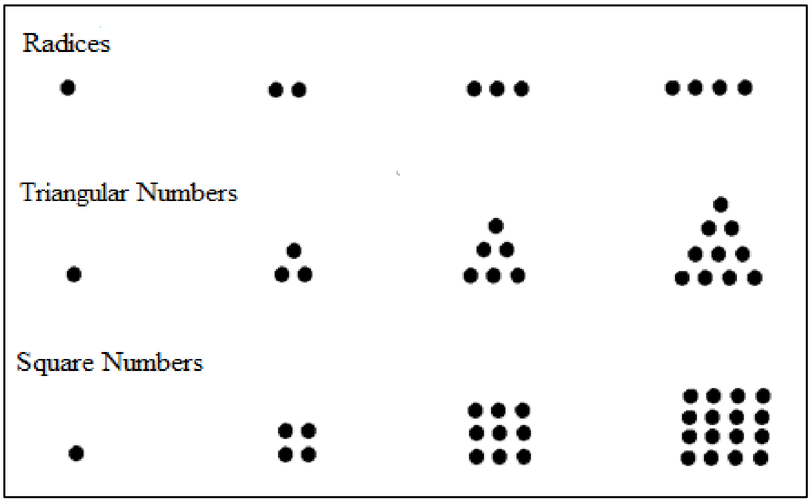


Figure 2.1. Drawings corresponding to the first four radices, triangular numbers, and square numbers. The drawings of the triangular and square numbers are a reconstruction of the original drawings which are reported beside the definitions at the margin of the page. The drawings of the radices are not present in original, I have constructed them in analogy with the other drawings.

Analysing these constructions, we can observe another important difference with the numbers of the Euclidean arithmetic. The structural role of the successor operation, as highlighted by the construction

---

[24] Maurolico started to write a compendium of arithmetic at least twenty years before and probably the version of 1557 is the result of a continuous modification of this first manuscript.

[25] The critical edition of the *Arithmeticorum Libri Duo* is available online, on the website of the "*Edizione Nazionale dell'Opera Matematica di Francesco Maurolico*: http://maurolico.elabor.biz/Maurolico/sezione.html?path=6.A.1,

[26] *Unitas est principium et constitutrix omnium numorum, constituens autem imprimis seipsam.* (Maurolico, 1575, p.2)

[27] *Omnis igitur numerus aut est unitas, quae respondet puncto[…] Aut est linearis, […]Aut superficialis, […] Aut solidus.* (Maurolico, 1575, p.2)

of the radices[28], and more in general the presence of a recursive construction of the numbers is evident here. This is an important aspect to notice in relation to the presence of MI in Maurolico.

Maurolico concludes the first page with a table which summarises the just given definitions. A part of this table is reconstructed in Table 2.2, to which I will refer in the following analyses of the propositions.

Radices	**Even numbers**	**Odd numbers**	**Triangular numbers**	**Square numbers**
1	0	1	1	1
2	2	3	3	4
3	4	5	6	9
4	6	7	10	16
5	8	9	15	25
6	10	11	21	36
R_n	E_n	O_n	T_n	S_n

Table 2.2. List of the first six radices, even numbers, odd numbers, triangular numbers, and square numbers, as reported by Maurolico. The last line, which is not present in the original table, introduces the notation that will be used in the following pages to indicate the n-th number of each category.

A few observations are necessary.

- The original table does not contain the symbolic labels of the last line. Maurolico's arithmetic is still essentially rhetorical, they have been added by me to better present the following analyses.
- The original table is more extended in the number of the lines (which are ten) and in the number of the columns (other categories of numbers are included); Columns and lines reported in the Table 2.2 are sufficient for the following propositions.
- The numbers 1 and 0 are present in the original table as well. In particular, 1 is the first radix, the first odd number, triangular number, etc., 0 is the first even number and the first of the *parte altera longiores*[29]*,* etc. This aspect, compared to the Euclidean arithmetic, represents a further step in the direction of the modern definition of natural numbers: the one and even the zero are numbers of some kinds.

[28] "Radices are formed by the unit, and by the continous addition of the unit" (*Radices formantur ab unitate, et per unitatis continuam additionem),* (Maurolico, 1575, p. a).
[29] The numbers with the form $(n-1) \cdot n$, for some natural number *n.*

- Finally, the table is used by Maurolico to introduce the following terminology, that will be used in our analysis as well: numbers of the same line are called *collateral numbers;* numbers belonging to two consecutive lines (even if in two different columns) are called *precedent* or *consecutive numbers,* depending on if they belong to the upper or the downer line. Thus, for instance, taken the fourth even number [6] the consecutive triangular number is 15, its collateral triangular number is 10, and its precedent triangular number is 6.

We are now ready to analyse a couple of propositions that, following some scholars (Vacca, 1909; Bussey, 1917), attest an evident trace of the presence of MI in Maurolico's arithmetic.

<u>*Proposition XIII*</u>

Every square number together with the consecutive odd number, constitute the consecutive square number.

<u>*Proof:*</u>

For example, the fourth square number 16 together with the fifth odd number 9 make the fifth square number [25]. In fact, for the proposition VI[30]*, the fourth radix together with the fifth radix make the fifth odd number. Therefore, for the previous proposition [XII*[31]*], the fourth square number together with the fourth and the fifth radix will make the fifth square number. Therefore, the fourth square number together with the fifth odd number, that is 16 together with 9, constitute the fifth square number, which is 25. This concludes the proposition.* [32]

Let us analyse the structure of Maurolico's proof. The proposition is proved through a generic example which, however, can be generalised to any square number. With our symbolic notations, the statement of the proposition is:

$$\forall n, \qquad S_n + O_{n+1} = S_{n+1}$$

In the Proposition VI, Maurolico proved that every radix together with its consecutive radix give the consecutive odd number [$\forall n,\ R_n + \mathrm{R}_{\mathrm{n+1}} = \mathrm{O}_{\mathrm{n+1}}$].

Moreover, in the Proposition XII, Maurolico proved that every square number together with its collateral radix and its consecutive radix give the consecutive square number [$\forall n,\ S_n + R_n + R_{n+1} = S_{n+1}$].

[30] Proposition VI: Every radix together with its precedent radix gives the collateral odd number; together with its consecutive radix, instead, it gives the consecutive odd number. (*Omnis radix cum radice praecedenti, facit sibi collateralem imparem; cum sequenti vero sequentem*). With our notation: $\forall n,\ R_n + \mathrm{R}_{\mathrm{n+1}} = \mathrm{O}_{\mathrm{n+1}}$.

[31] Proposition XII: Every square number together with its [collateral] radix and with its consecutive radix, gives the consecutive square number. (*Omnis quadratus cum radice sua et cum radice sequenti coniunctus, consummat quadratum sequentem*). With our notation: $\forall n, S_n + R_n + R_{n+1} = S_{n+1}$.

[32] *Propositio 13ª. Omnis quadratus cum impari sequenti coniunctus, constituit quadratum sequentem. Exempli gratia: quartus quadratus scilicet 16 cum impare sequentis loci scilicet cum 9 coniunctus, efficit quintum quadratum.*
Nam per sextam praemissarum, radix quarta cum quinta componunt imparem quintum; cumque per praecedentem, quadratus quartus, cum quarta et quinta radicibus, pariter sumptus, efficiat quadratum quintum, sequitur ut idem quadratus quartus cum impare quinto, hoc est 16 cum 9, constituat quadratum quintum scilicet 25, sicut concludit propositio. (Maurolico, 1575, p.7)

Thus, at this point, using these two previous propositions, it possible to obtain the thesis of the Proposition XIII:

$S_4 + O_5 = S_5$ [The statement to prove, for the particular case n=4]

$\updownarrow$

$S_4 + R_4 + R_5 = S_5$ [Proposition XII, for n=4]

For

$R_4 + R_5 = O_5$ [Proposition VI, for n=4]

This concludes the proof.

The just proved proposition XIII is then used by Maurolico to prove the proposition XV.

Proposition XV

By adding the odd numbers starting from the unit in a consecutive order, the square numbers starting from the unit will be continuously obtained, each of them being collateral of the last odd number.

Proof:

For the proposition before the previous one [XIII], the unit [the first square] together with the consecutive odd number [3] make the consecutive square number 4; and this 4, the second square number, together with the third odd number 5 make the third square number 9; and this 9, the third square number, together with the fourth odd number 7 makes the fourth square number 16. And so on, ad infinitum, by repeating the proposition XIII, the statement is proved. [33]

The structure of the proof is the following: to prove the statement $\forall n, O_1 + O_2 + \cdots + O_n = S_n$ it is constructed a series of implications which is potentially infinite:

$S_1 = O_1 = 1$ *[Definition]*

$S_1 + O_2 = S_2$ *[Proposition XIII, for n=1]*

$S_2 + O_3 = S_3$ *[Proposition XIII, for n=2]*

$S_3 + O_4 = S_4$ *[Proposition XIII, for n=3]*

...

and so on, ad infinitum.

[33] *Propositio 15ª. Ex aggregatione imparium numerorum ab unitate per ordinem successive sumptorum, construuntur quadrati numeri continuati ab unitate, ipsisque imparibus collaterales.*
Nam per antepraemissam, unitas imprimis cum impari sequente facit quadratum sequentem scilicet 4. Et ipse 4 quadratus secundus , cum impari tertio scilicet 5 facit quadratum tertium, scilicet 9. Itemque 9 quadratus tertius cum impari quarto scilicet 7 facit quadratum quartum, scilicet 16 et sic deinceps in infinitum, semper decima tertia repetita, propositum demonstratur. (Maurolico, 1575, p.7)

Vacca's (1909) and Bussey's (1917) opinion is that in the proof of the just presented couple of propositions (XIII and XV) it is possible to observe the structure of a proof by MI. In particular, the proposition XIII constitutes what would be called the inductive step, while the proposition XV contains the induction base and the explicit construction of the "cascade" of syllogisms (for saying it with Poincaré) which, starting from the base case and successively using the inductive step, reaches every number involved in the statement.

Cassinet (1988) instead expresses a more cautious opinion. He recognises, from one side, how Maurolico, often in the book, makes a "frequent use of reasoning by recurrence to demonstrate propositions" (p. 171), however, from the other side, he observes that Maurolico never presents a justification for the validity of this way of proving a proposition. In relation to this last aspect, as we will see, Pascal will provide a justification for its validity, instead. A similar cautious position is expressed by Pasquotto (1998) who, instead of referring directly to MI in Maurolico's works, calls it "method which uses the Principle of Mathematical Induction" (p.23, my translation) and he concludes that "however, it is clear that Maurolico's induction […] has formally few in common with the one to which we are used in mathematics" (p.42, my translation).

Despite similarities and differences of Maurolico's proofs with the modern MI, we can register the presence of new elements in relation to the previously analysed authors. Firstly, in Maurolico the word 'infinitum' finally appears. The iteration, thus, is not simply reaching an arbitrarily big number, but it continues 'ad infinitum'. We can notice that in the proposition XV of above there is not a reference to the *totality* of square numbers: it is said that the square numbers will be continuously obtained (i.e., without exclusions), and not that *all* the square numbers will be obtained. This aspect seems to highlight a view of infinite sets which is still potential, similar to the Euclidean 'as many as we please'. However, Maurolico himself states that the proposition is proved by iterating the reasoning "in infinitum", which instead seems to go in the direction of a view of infinity which is actual. Surely the line between the two aspects here is blurred.

A further aspect of novelty is given by the fact that the statement which constitutes the step of the iteration is previously proved in an *ad hoc* proposition (in our example, the proposition XIII). This provides a peculiar structure to Maurolico's proof, which I will call 'bipartite structure': a principal statement involving a (infinite) series of natural numbers is proved firstly with a proposition involving two generic numbers, one consecutive to the other, and then with a second proposition in which the previous one is (infinitely) applied starting from the first natural number for which the general statement is true. As we will see this structure will be registered in Pascal as well. Comparing this structure with the one of a classic proof by MI, we can see that here the inductive step is proved previously than the base of the induction and in separate proposition. However, the structure 'base of the induction + inductive step' as two independent parts of the proof seems to be slowly emerging.

Let us observe, finally, the presence of an element in Maurolico's proof that was already noticed in the previous authors. The structure of the sentences which describe the steps of the iteration, and which precede the 'and so on, ad infinitum', is repetitive: the involved number changes, but the grammatical structure and the other elements of the sentence remain the same. The presence of kind of a 'modular' structure of the sentences was observed in Euclid and Laertius. It seems that the rhythmic repetition of the worlds suggests the presence of *a pattern* in the steps of the proof which could be repeated and generalised toward an arbitrarily big number (in Euclid or Laertius) or *ad infinitum* (in Maurolico). In other terms, it seems that the repetitive structure of the sentences is used to represent the fact that the particular and specific steps of the iteration can be generalised.

2.7 PASCAL – THE *TRAITE DU TRIANGLE ARITHMETIQUE*

In 1665 the '*Traité du triangle arithmétique: avec quelques autres petits traitez sur la mesme matière'* is published posthumously in Paris, under the name of Blaise Pascal. The book treats, as said in the title, about the arithmetic triangle, nowadays also known precisely as the Pascal's triangle. Pascal's starts preliminarily with the construction of the structure of the triangle: firstly, two arbitrarily long sides of a triangle are constructed and with them an arbitrary number of cells which constitute the triangular array ("I will take as many cells I want" [34]), finally these cells are labelled with some letters. The results of this construction is represented in the Figure 2.2.[35]

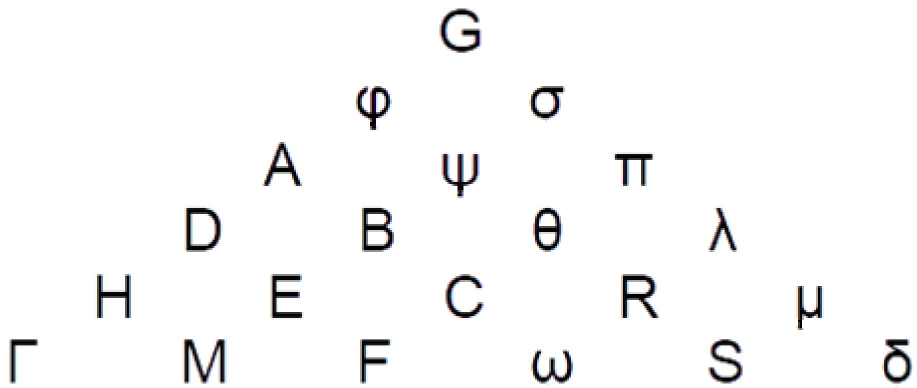


Figure 2.2. Pascal's arithmetic triangle. The letters are the ones originally used by Pascal and to those it will be referred in the following propositions.

At this point Pascal assigns a numerical value to every cell of the triangle:

The number of the first cell, the one which contains the right angle [G], is arbitrary, but once this is fixed all the others will be forced and for this reason it will be called the Generator of the triangle. Each of the other [numbers] is specified by the following rule.

The number of each cell is equal to the one of the cell which precedes it in the perpendicular line plus the one of the cell which precedes it in the parallel line. [36] *So, the cell F, namely the number of the cell F, is equal to the cell C plus the cell E, and so on for all the others.* [37]

[34] *Je prends tant que je veux* (Pascal, 1665 p.1)

[35] The construction made by Pascal results in a slightly different drawing compared to the one represented in Figure 2.2. In particular the two oblique sides are, in Pascal's drawing, vertical (the side GΓ) and horizontal (the side Gδ), forming a square angle in the vertex G. However nowadays the Pascal's triangle is represented as in figure 2.2. I chose to remain concordant with this modern representation to facilitate the comprehension of the following propositions.

[36] With reference to the Figure 2.2 the definition would be: The number of each cell is equal to the number of the precedent cell belonging to the previous line plus the number of the consecutive cell still belonging to the previous line.

[37] *Le nombre de la premiere cellule qui est à l'angle droit est arbitraire, mais celuy-là estant placé tous les autres sont forcez, et pour cette raison il s'appelle le Generateur du triangle. Et chacun des autres eſt specifié par cette seule regle.*
Le nombre de chaque cellule, est égal à celuy de la cellule qui la precede dans son rang perpendiculaire, plus à celuy de la cellule qui la precede dans son rang parallele. Ainsi la cellule F, c'est à dire le mombre de la cellule F, égale la cellule C, plus la cellule E; et ainsi des autres. (Pascal, 1665, section 1, p.2).

Let us observe that this construction has some aspects which seems related to MI. The value of each cell, in fact, is recursively defined starting from the generator G. In particular two aspects of the recursion are described: the starting point (i.e., the generator G which is arbitrary, and which determines the values of all the other terms) and the step of the recursion (i.e. how the cell of a line are obtained by the cells of the previous line). It is interesting that Pascal describes just one single step of the recursion (and for only one cell) which becomes a generic representant of all the other steps ("and so on for all the others").

Moreover, as a consequence of Pascal's definition, the cells constituting the two oblique sides of the triangle are all equal to the generator G, since they do not have, respectively, a precedent cell or a consecutive cell in the previous line. This property is proved in the *Cosequence Premiere,* the first proposition proved by Pascal after the definition of the arithmetic triangle.

Pascal concludes the construction of the arithmetic triangle by saying that in all the book he will explicitly refer only to the specific case of the triangles with G=1 (see Figure 2.3), but that all the propositions, if appropriately modified, will be valid for any arbitrary G.[38]

1

1 1

1 2 1

1 3 3 1

1 4 6 4 1

1 5 10 10 5 1

Figure 2.3. Arithmetic triangle obtained with G=1.

This historical analysis started by saying that, at the beginning of the 20th century, Pascal was commonly considered the father of proofs by MI in modern sense. In particular, one of the examples of the presence of MI in Pascal's work was registered in the *XII Consequence*, the 12th proposition after the definition of the arithmetic triangle. Before presenting it in detail, however, I wish to analyse a couple of other propositions (VII Consequence and VIII Consequence) which represent, in a certain sense, a bridge between what it was observed in Maurolico, and the aspects of novelty observed in Pascal.

VII Consequence

In all the arithmetic triangle the sum of the cells of any given base[39] is the double [of the sum] of the cells of the previous base.

Proof:

Let D B ϑ λ be any given base. I affirm that the sum of the cells is the double of the sum of the cells of the previous A Ψ Π.

[38] I consider the triangles in which the generator is the unit; but what will be said can be adapted for all the other [triangles]. (*Je considere les triangles, dont le generateur est l'unité , mais ce qui s'en dira conviendra à tous les autres*). (Pascal, 1665, section 1, p.2).
[39] A "base" corresponds for Pascal to a horizontal line of cells in our drawing.

Because the extremes D, λ are equal to the extremes A, Π.
And the others B = A + Ψ,
ϑ = Ψ + Π.
Therefore, D + λ + B + ϑ = [A + Π + A + Ψ + Ψ + Π =]
= 2A + 2Ψ + 2Π.
We can prove the same thing for all the other bases.[40]

<u>VIII Consequence</u>

In all the arithmetic triangle the sum of the cells of any given base is a number of the double progression which starts from the unit, whose exponent is the same [of the number] of the base.[41]

<u>*Proof:*</u>

Because the first base is the unit.
The second one is the double of the first one, thus it is 2. [VII Consequence]
The third one is the double of the second one, thus it is 4. [VII Consequence]
And so on, to infinity.[42]

This couple of propositions is extremely similar to the couple of propositions analysed for Maurolico.

First of all, also in this case, we have an initial proposition (the VII consequence) in which a property involving two consecutive natural numbers is proved (in this case, the two consecutive numbers are the ordinal numbers of two consecutive bases) and then a second proposition (the VIII consequence) in which the previous property is iteratively applied 'to infinity' starting from the unit. As it happened in Maurolico, Pascal explicitly refers to the infiniteness of the iteration ("and so on, to infinity") and to the generality of the statement ("In all the arithmetic triangle […] any given base").

Secondly the proof of the first VII consequence is again conducted for a generic example (the third and the fourth base of the triangle) with a process that can be generalise to the other bases. This generalisation is explicitly stated by Pascal ("We can prove the same thing for all the other bases") while, instead, the same thing remained implicit in Maurolico (and in Euclid).

---

[40] *Consequence septiesme. En tout Triangle Arithmetique, la somme des cellules de chaque base, est double de celles de la base precedente.*
Soit une base quelconque D B ϑ λ. Je dis que la somme de ses cellules, est double de la somme des cellules de la precedente A Ψ Π.
Car les extremes, D, λ, égalent les extremes, A, Π,
Et chacune des autres B, ϑ,en égalent deux de l'autre base, A + Ψ, Ψ+ Π.
Donc, D + λ + B + ϑ, égalent 2A + 2Ψ + 2Π.
La mesme chose se demonstre de mesme de toutes les autres. (Pascal, 1665, section 1, p.5).
[41] The statement can be expressed in modern notations as 'the sum of numbers of the n-th line is equal to 2^n'. Note that this proposition is valid only if G=1.
[42] *Consequence huitiesme. En tout Triangle Arithmetique, la somme des cellules de chaque base, est un nombre de la progression double, qui commence par l'unité, dont l'exposant est le mesme que celuy de la base.*
Car la premiere base est l'unité.
La seconde est double de la premiere, donc elle est 2.
La troisiesme est double de la seconde, donc elle est 4.
Et ainsi à l'infiny. (Pascal, 1665, section 1, p.5).

Lastly, the iterative structure of the argument which constitutes the proof of the VIII consequence is again expressed with a repetition of sentences all with the same structure: "The second one is the double of the first one, thus it is 2. The third one is the double of the second one, thus it is 4".

In conclusion, Pascal's proof for these two propositions seems to be completely analogous to what happened in Maurolico. Evident differences can be instead observed in the following proposition, the XII consequence, in which the proof given by Pascal was interpreted by many mathematicians as a proof by MI.

<u>*XII Consequence*</u>

In all the arithmetic triangle, [taken] two adjacent in the same base cells, the precedent one is to the consecutive one, as the number of the cells that are between the consecutive cell and the end of the base is to the number of the cells that are between the precedent one and the beginning of the base.

<u>*Proof:*</u>

Let E, C be two any given adjacent cells in a given base, I say that E is to C as 2 is to 3, because E is the precedent cell and C the consecutive one and because there are two cells before E, which are E and H, and because there are three cells after C, which are C R μ[43]*. Despite this proposition has an infinity of cases, I will provide a shorter proof, thanks to the use of two lemmas.*
The 1st, which is self-evident, [affirms] that this proportion holds for the second base; indeed is well visible that φ is to σ as 1 is to 1.
The 2nd [affirms] that if this proportion holds for a base, then it will necessarily hold for the successive base as well.
As a consequence, we will have that it holds for all the bases: in fact it holds for the second base, for the first lemma, therefore for the second [lemma] it holds for the third base as well, therefore for the fourth [base], and to infinity.
Thus, we only need to prove the second lemma, in this form.
If this proportion holds for a given base, as in the fourth one D λ[44]*, which is to say, if D is to B as 1 is to 3, and B is to ϑ as 2 is to 2, and ϑ is to λ as 3 a 1, etc.,*
I say that the same proportion will hold also for the successive base, H μ [45]*, which is to say that, for example, E is to C as 2 is to 3.*

Indeed D is to B as 1 is to 3, for the hypothesis.
Thus D+B is to B as 1+3 is to 3,
that is E is to B as 4 is to 3.

Similarly, B is to ϑ as 2 is to 2, for the hypothesis.
Thus B+ ϑ is to B as 2+2 is to 2,
that is C is to B as 4 is to 2.

[43] The cells E and C, taken from the base H E C R μ, constitute another time a generic example. In the specific case of these cells the statement of the precise translation of the proposition in modern notation would be C:E = 3:2, thus what written by Pascal (E:C=2:3) is not an exact transposition of the statement. Moreover, we can observe that in the case of a triangle generated by G=1, we have E=4 and C=6 whose ratio is indeed 2:3.
[44] The complete base is D B θ λ, from which the proportions written above by Pascal follow.
[45] *H E C R μ.*

But B is to E as 3 is to 4, for what we have proved above,
therefore C is to E as 3 is to 2, which is what we need to prove.[46]

The same will be proved in all the rest [of the triangle], since this proof is not based on other than the proportions of the precedent base and the fact that any cell is equal to the precedent one plus the upper one, this will be true for all of them.[47]

Pascal's proof has the structure of a modern proof by MI, with two exceptions. Firstly, Pascal proves the inductive step for a generic example (two specific cells of two specific lines), even though several times he clarifies that the proof should not lose of generality because of the process can be generalised to the other cells and lines. Secondly the classic algebraic symbolism of proof by MI is still absent; in particular the parameter n on which induction is applied (i.e., the line's number) remains implicit. Except for these two aspects, however, Pascal's proof could be considered as a classic proof by induction in modern terms, perhaps not completely formalized, but still correct.

Let us focus on the aspects which differentiate this proof with the previously analysed ones, in particular the couple of propositions still proved by Pascal and the ones by Maurolico of above. In this last proof it is possible to register some interesting new characteristics:

1) The proof is composed by two parts, the two "lemmas", one corresponding to the base case and the other to the inductive step, which instead of being proved in two separate propositions, as it happened previously, are proved within the same one. Moreover, the order in which these two propositions are proved is opposite from the previous propositions. For Maurolico (and in the first two propositions by Pascal) firstly an independent proposition involving an (archetype of) inductive step was proved and then, successively, a proposition was proved true for a first number and, starting from this, the previous proposition was iteratively applied. In this last proof by Pascal, instead, the inductive step is proved after the base of the induction.

[46] In fact, from $\frac{B}{E} = \frac{3}{4}$ and from $\frac{C}{B} = \frac{4}{2}$ it follows that $\frac{C}{E} = \frac{B}{E} \cdot \frac{C}{B} = \frac{3}{4} \cdot \frac{4}{2} = \frac{3}{2}$

[47] *Consequence douziesme. En tout Triangle Arithmetique, deux cellules contigues estant dans une mesme base, la superieure est à l'inferieure, comme la multitude des cellules depuis la superieure iusques au haut de la base, à la multitude de celles, depuis l'Inferieure iusques en bas inclusiuement.*
Soient deux cellules contigues quelconques d'une mesme base, E, C, je dis que E, inferieure, est à C, superieure, comme 2 à 3, parce qu'Il y a deux cellules depuis E iusques en bas sçauoir, E, H, parce qu'Il y a trois cellules depuis C iusques en éaur, sçauoir C, R, μ.
Quoy que cette proposition ait une infinité de cas, j'en donneray une demonstration bien courte, en supposant 2 lemme. Le 1. qui est evident de soy-mesme, que cette proportion se rencontre dans la seconde baſe; car il est bien visible que φ est à σ comme 1, à 1. Le 2. que si cette proportion se trouve dans une base quelconque, elle se trouvera necessairement dans la base suivante. D'où il se voit, qu'elle est necessairement dans toutes les bases: car, elle est dans la seconde base, par le premier lemme, donc par le ſecond elle est dans la troisiesme base, donc dans la quatriesme, et à l'infiny.
Il faut donc seulement demonstrer le second lemme, en cette sorte. Si cette proportion se rencontre en une base quelconque, comme en la quatriesme D λ, c'est à dire, si D est à B comme 1 à 3. Et B, à ϑ comme 2 à 2. Et ϑ à λ comme 3 à 1. etc. Je dis, que la mesme proportion se trouvera dans la base suivante H μ, et que par exemple E est à C comme 2 à 3.
Car D est à B comme 1 à 3. par l'hypothese. Donc D + B est à B comme 1 + 3 à 3. [Donc] E à B comme 4 à 3.
De mesme, B est à ϑ comme 2 à 2 par l'hypothese. Donc, B + ϑ à B comme, 2 + 2 à 2. [Donc] C à B comme 4 à 2.
Mais B à E comme 3 à 4 comme il est monstré. Donc par la proportion troublée, C est à E comme 3 à 2. Ce qu'il falloit demonstrer.
On le monstrera de mesme dans tout le reste, puisque cette preuve n'est fondée que sur ce que cette proportion se trouve dans la base precedente, et que chaque cellule est égale à sa precedente, plus à sa superieure, ce qui est vray par tout. (Pascal, 1665, section 1, pp. 7-8)

2) The statement of the proposition does not involve, explicitly in its formulation, a recursion. The recursion resides in the definition of the arithmetic triangle itself but not in the statement. Indeed, the statement only refers to any two adjacent cells of any given line. This is a new aspect in our analysis. In all the previous examples that we analysed, when an iteration or a recursion was constructed, the recursion was explicitly in the statement itself (as in the last two couples of propositions or in Euclid's propositions on geometrical progressions). This aspect extends, in a certain sense, the domain of the use of such recursive argumentations.

3) If in the precedent propositions by Pascal and by Maurolico the proof was *infinite*, here the proof becomes *finite*, as explicitly stated by Pascal's himself: "Despite this proposition has an infinity of cases, I will provide a shorter proof, thanks to the use of two lemmas". This is a crucial point, and it is perhaps one of the biggest steps toward modern induction made by Pascal. For the first time between the analysed authors, the proof is not composed by an infinite number of steps (as in Maurolico: "And so on, ad infinitum, by repeating the proposition XIII, the statement is proved"), but it is only composed by two steps. The second "lemma", as called by Pascal, becomes a generic representant of all this infinite number of steps, exactly how it happens in a modern inductive step.

4) The proof contains the justification of the validity of proof itself. This happens right after Pascal states the two lemmas. The justification is given by constructing a chain of inferences (formally, by *modus ponens*) which transfer the truth of the proposition (i.e., that the given proportional property holds in a line) from the second line, for which it is true for the first lemma, to the third line, then from the third line to the fourth one, and so on to infinity. The justification is thus constructed referring to an infinite iteration, a process that Pascal already used in the VIII consequence. The fact that Pascal provided a justification for the validity of his proof is an interesting point. It seems that he was aware of the novelty represented by this way of proving a proposition and thus he felt the necessity of explaining it to the reader. This thing happens all the time in which Pascal uses his MI to proof a proposition in which a justification of its validity is always provided.

5) Within the *Traité du triangle arithmétique,* the same proving scheme used in this proposition is present other times. In particular it is used in the proposition 1 of the 2nd section[48], *Usage du triangle arithmétique pour les combinaisons,* in the lemma 5of the 9th section[49], *Combinationes,* and in the resolution of a famous probability problem, *Le problème des partis,* object of the 3rd section[50]. This aspect suggests that Pascal uses this proving scheme as a useful heuristic to prove propositions or to solve problems.

---

[48] In all the arithmetic triangle, the sum of the cells of a parallel line is equal to the total number of the combinations of the number corresponding to the exponent of the line in the number corresponding to the exponent of the triangle. (*En tout Triangle Arithmetique, la somme des cellules d'un rang parallele quelconque égale la multitude des combinaisons, de l'exposant du rang dans l'exposant du Triangle)* (Pascal, section 2, p.7).

[49] The proposition is the same of the previous one, this time translated in Latin: *In omni triangulo Arith. summa cellularum seriei cujuslibet , aequatur multitudini combinationum exponentis seriei, in exponente trianguli.* (Pascal, section 9, p. 24).

[50] The problem is the following: to divide the prize of a game between two players who decide to interrupt the game before it is actually finished, in a way that the division of the prize respects the winning probability of each player at the moment of the interruption. The problem and the solution given by Pascal is analysed, with modern notations, in Bussey (1917).

To conclude this historical analysis of the genesis of proofs by MI, one last author is missing: Pierre de Fermat and his *method of infinite descent*, which will be analysed in the following section. As we will see, it will highlight other not yet considered aspects related to MI.

2.8 FERMAT – THE *METHOD OF INFINITE DESCENT*

Pierre de Fermat and his 'Method of infinite descent' are often quoted in relation to MI (Bussey, 1918; Ernest 1982). From a logical point of view the connection between MI and the principle of infinite descent, on which Fermat's method is constructed, and which we have already analysed for Euclid, is in the well-ordering principle[51]. In this section I will analyse in detail the Fermat's use of the principle of infinite descent. I particular I will focus on a letter, written to Pierre de Carcavi in 1659, in which Fermat affirms to have found a new and singular method for proving statements: the method of infinite descent. Let us analyse the first part of the letter, containing Fermat's description of the method. The French text that I took as reference is the one published in the second volume of the critical edition of Fermat's writings of 1894, cured by Tannery and Henry (Fermat, 1894). The numbers between parentheses have been added by me for the following analysis.

Summary of the new discovers in the science of numbers.

(1) And because the ordinary methods, which are on the books, were not sufficient to prove so difficult propositions, finally I found a very singular path toward this objective. I called this method of proving the infinite descent or indefinite, etc.; at the beginning I used it only for negative propositions, such as, for example: that there is no [natural] number which is lower than a multiple of 3 of a unit[52] and which is composed by a square number and by the triple of another square number[53]; that there is no right triangle whose sides are [natural] numbers and whose area is [equal to] the square of a [natural] number. (2) The proof is made by άπαγωγήν την έις άδύνατον[54] in this manner: If there is a right triangle whose sides are integers and whose area is equal to a square [natural number], then there is a second, triangle, smaller than the first one, which has the same property. If there is a second triangle, smaller than the first one, which has the same property, then, by the same reasoning, there is a third triangle, smaller than this second one, which has the same property, and then a fourth, a fifth, etc., descending ad infinitum. Now, it happens that, given a number, there cannot be an infinite descent of numbers smaller than the given one (I am still talking about integers [natural] numbers). From which, one concludes that it is therefore impossible that there is a right triangle whose area is a square [natural number]. (3) From this we can deduce that there are no [right triangles whose sides are] fractions whose area is a square; for, if there are [right triangles whose sides are] fractions, then there are [right triangle whose sides are] integers, which cannot be, as proved by the descent. (4) I do not add here how I deduced that, if there is a right triangle with this nature, then there is another one with the same nature smaller than the first one, because the reasoning on that would be too long and this is all the mystery of my method. I would be very glad if Pascal and Roberval and many other wise people will look for it on my indication.

(5) For a long time, I was not able to apply my method to affirmative propositions, because the manned and the idea to get at them is much more difficult than that which I use for the negative ones. So that, when I had to prove that every prime number which exceeds of a unite a multiple of 4 is composed by two square numbers[55], I found myself in much torment. (6) But at the end, a reflection many times repeated,

[51] For a formal analysis of the logical connection between these principles, see Lolli 2008.
[52] $n=3k-1$
[53] $n=a^2+3b^2$
[54] *Reductio ad absurdum*, i.e., by contradiction.
[55] For every prime number p with the form $p=4k-1$ for some k natural number, then $p=a^2+b^2$, for some a, b, natural numbers.

gave me the light that I needed, and the affirmative propositions worked with my method, with the help of some new principles which should be added for necessity. This new reasoning for these affirmative propositions is the following: if an arbitrary prime number, which exceeds of a unit a multiple of 4, is not composed by two squares, then there will exist a prime number of the same sort but less then the given one, and then a third still smaller, etc., descending ad infinitum util one comes to the number 5, which is the smallest of the kind in question, and which would not be composed by two square numbers, although in fact it is. From which one must deduce, by reductio ad absurdum, that all numbers of the kind in question are consequently composed by two square numbers.[56]

Let us analyse in order what written by Fermat. At the beginning of the excerpt (1), right after the anticipating title "Summary of the new discovers in the science of numbers", Fermat affirms to have encountered some propositions to prove which, the traditional methods ("which are on the books") were not sufficient. This initial difficulty, however, prompted him to find a new, "singular", method of proof. Fermat says to call this method the "infinite" or "indefinite" descent. The name is evocative for what happens when it is applied. Let us focus on the two terms that Fermat uses: "infinite" and "indefinite". As I have observed previously in this chapter, the two terms are not synonyms, from a modern mathematical point of view. It is not clear if Fermat considers them as such or not. As I will discuss later

[56] *RELATION DES NOUVELLES DÉCOUVERTES EN LA SCIENCE DES NOMBRES*
(1) Et pour ce que les méthodes ordinaires, qui sont dans les Livres, étoient insuffisantes à démontrer des propositions si difficiles, je trouvai enfin une route tout à fait singulière pour y parvenir. J'appelai cette manière de démontrer la descente infinie ou indéfinie, etc.; je ne m'en servis au commencement que pour démontrer les propositions négatives, comme, par exemple: Qu'iln'y a aucun nombre, moindre de l'unité qu'un multiple de 3, qui soit composé d'un quarré et du triple d'un autre quarré; Qu'il n'y a aucun triangle rectangle en nombres dont l'aire soit un nombre quarré.
(2) La preuve se fait par ἀπαγωγήν την ἐις ἀδύνατον *[reductio ad absurdum] en cette manière: S'il y avoit aucun triangle rectangle en nombres entiers qui eût son aire égale à un quarré, il y auroit un autre triangle moindre que celui-là qui auroit la même propriété. S'il y en avoit un second, moindre que le premier, qui eût la même propriété, il y en auroit, par un pareil raisonnement, un troisième, moindre que ce second, qui auroit la même propriété, et enfin un quatrième, un cinquième, etc. à l'infini en descendant. Or est-il qu'étant donné un nombre, il n'y en a point infinis en descendant moindres que celui-là (j'entends parler toujours des nombres entiers). D'où on conclut qu'il est donc impossible qu'il y ait aucun triangle rectangle dont l'aire soit quarrée. (3) On infère de là qu'il n'y en a non plus en fractions dont l'aire soit quarrée; car, s'il y en avoit en fractions, il y en auroit en nombres entiers, ce qui ne peut pas être, comme il se peut prouver par la descente. (4) Je n'ajoute pas la raison d'où j'infère que, s'il y avoit un triangle rectangle de cette nature, il y en auroit un autre de même nature moindre que le premier, parce que le discours en seroit trop long et que c'est là tout le mystère de ma méthode. Je serai bien aise que les Pascal et les Roberval et tant d'autres savans la cherchent sur mon indication.*
(5) Je fus longtemps sans pouvoir appliquer ma méthode aux questions affirmatives, parce que le tour et le biais pour y venir est beaucoup plus malaisé que celui dont je me sers aux négatives. De sorte que, lorsqu'il me fallut démontrer que tout nombre premier, qui surpasse de l'unité un multiple de 4, est composé de deux quarrés, je me trouvai en belle peine. (6) Mais enfin une méditation diverses fois réitérée me donna les lumières qui me manquoient, et les questions affirmatives passèrent par ma méthode, à l'aide de quelques nouveaux principes qu'il y fallut joindre par nécessité. Ce progrès de mon raisonnement en ces questions affirmatives est tel: si un nombre premier pris à discrétion, qui surpasse de l'unité un multiple de 4, n'est point composé de deux quarrés, il y aura un nombre premierde même nature, moindre que le donné, et ensuite un troisième encore moindre, etc. en descendant à l'infini jusques à ce que vous arriviez au nombre 5, qui est le moindre de tous ceux de cette nature, lequel il s'ensuivroi t n'être pas composé de deux quarrés, ce qu'il est pourtant. D'où on doit inférer, par la déduction à l'impossible, que tous ceux de cette nature sont par conséquent composés de deux quarrés. (Fermat, 1894, pp. 432-433).

in this section, the difference between 'infinite' and 'indefinite' is a crucial element to consider if we want to compare Fermat's method with the modern mathematical induction. Successively Fermat continues the letter by stating that he firstly used his method to prove some proposition involving a negative statement (i.e., with the form "there does not exist…"). He then reports two examples of this kinds which, in modern notations, can be written as 'there does not exist a natural number *n* such as n=3k+1 and n= a^2+3b^2 for some *m, a, b* natural numbers', and 'there does not exist a right triangle whose sides are natural numbers and whose area is a square number'.

In the subsequent paragraph (2), Fermat describes the method of infinite descent showing how it is possible to apply it to prove the second proposition of those reported as examples. The proof is made by contradiction. Firstly, we suppose that there exists a right triangle whose sides are natural numbers and whose area is a square number, then we show that from this it is possible to construct a second triangle, whose sides are all natural numbers, smaller than the first ones, and whose area is still a square number. At this point the procedure can be repeated to obtain a third triangle whose sides are smaller than the second one and which still has the given properties, and then a fourth triangle, and a fifth, and so on, "descending ad infinitum". This, as Fermat claims, is impossible because we would have an infinite descent of natural numbers (formally, an infinite and strictly decreasing sequence of natural numbers) corresponding to the decreasing lengths of the sides of the sequence of triangles. Having found a contradiction, the proof is concluded.

At this point (3), Fermat generalises the proposition to the case of right triangles whose sides are fractions. He states that also in this case there does not exist a triangle of that kind whose area is a square number, because this case can be reconducted to the previous one, which was proved to be impossible.

Then (4), Fermat says that he will not write in the letter the proof of the fact that taken a triangle with the properties of the proposition it is possible to construct another triangle, with smaller sides, with the same properties. Fermat says that the reasoning would be too long and that he does not want to reveal the "mystery" of his proof. We do not know if, when the letter was written, he was in possess of a correct proof of this part. The whole theorem, with a complete proof with the method of the infinite descent, would be published only in 1670 by Fermat's son in his edition of Diophantus' *Arithmétiques* which is substantially a copy of Bachet's edition of 1621 to which he added the notes written by the father. [57]

In the second part of the excerpt, Fermat affirms to have had many difficulties when trying to apply this method to prove "affirmative" propositions as, for example, the proposition which, in modern notations, states that 'For every prime number p with the form p=4k-1 for some k natural number, then $p=a^2+b^2$, for some a, b, natural numbers' (5). However, Fermat says that after "a reflection many times repeated", he finally managed to adapt his method to this second category of propositions. Fermat's argument for the previous proposition is the following: let us suppose by contradiction that there exists a prime p with the given form (p=4k-1) but without the given property (i.e., that can be written as the sum of two square numbers); then we can prove that it is possible to find a second prime number, lower than the first one, still with the same form and still without the given property. By iterating this reasoning "descending at infinitum" we come to the number 5 obtaining that it should be a prime number with the given form but without the given property, which is absurd since 5 can be written as the sum of two square numbers. Therefore, the statement is proved by contradiction. Let us observe that Fermat's argument is different from the negative propositions of above. In this case the descent of numbers ends in correspondence to a number for which the property to prove is true. Formally, despite Fermat writes "infinite descent", in

[57] The story of this theorem and of its tradition, and the interpretations that have been given in the years is discussed in Goldstein, 1995. In this study, the effective paternity of the proof is also investigated, showing how Bernard Frenicle de Bessy, a mathematician contemporary to Fermat and with whom he was in close correspondence, was in possess of an analogous proof.

this case an "indefinite" (but finite) descent is created. The contradiction is not given by the presence of a strictly decreasing infinite sequence of natural numbers, but by the fact that one can find a number (5, in Fermat's example) for which a given property is at the same time true and false.

The just presented letter offers some interesting elements to observe for our analysis:

- Firstly, we can register that the method described by Fermat, at least for the negative propositions, is analogous to the structure of Euclid's proof of the proposition VII.31 of the *Elements* ("Any composite number is measured by some prime number"). In both cases the contradiction is obtained by constructing an infinite (or, better, indefinitely long) strictly decreasing sequence of natural numbers.[58]

- An interesting point is that Fermat gives a name to this method. This highlights that Fermat recognised some paradigmatic characteristics in the proofs he had constructed and that he extracted from those a "method" which he describes in this letter. In other terms, he is describing a heuristic to prove statements. This is an aspect of novelty compared to the previously analysed authors: the "new discovers in the science of numbers", as written in the title of the letter, are not some specific propositions (for which, in fact, Fermat does not give a complete proof) but a new method for proving them. To summarise, the object of Fermat's discourse is the method itself, the components which structure it, and the way it can be applied.[59]

- A third element to observe is in relation to the second part of the text, when the affirmative propositions are considered. In this case, as seen, the absurd is not obtained, as in the negative propositions, by contradicting the "infinite descend principle", but in a different way: if we assume that for a generic given number n the proposition is false, then it is possible to prove that the proposition is false for a particular number n_0 as well (5 in Fermat's proof) for which, however, we know that the property is true instead. Thus, given that for an arbitrary number n the proposition cannot be false, it is necessary true for every number n. If we further analyse this argument, we can register that, from a logical point of view, it seems to have several similarities with mathematical induction. First of all, the argument is composed by two parts: in the first one, we prove that if $P(n)$ is false for a given n, then $P(m)$ is also false for some $m<n$, and we iterate this implication to obtain that $P(n_0)$ must be false; in the second one, we show that $P(n_0)$ is true. The parallelism with MI is evident: $P(n_0)$ true becomes the base case, while the first part can be seen as the proof of the statement '$P(n)$ false $\rightarrow \exists\, m<n$, $P(m)$ is false' which is the contrapositive statement of the inductive step of a proof by strong induction (i.e., with the form, '[$\forall\, m<n$, $P(m)$ true] $\rightarrow$ $P(n)$ is true').

- As seen in the previous point, Fermat's argument seems to involve, in a certain sense, the contrapositive of an inductive step. More generally, Fermat is using the infinite descent principle which, formally, is equivalent with the principle of strong MI. If from a logical point of view the

[58] As already observed, the principle by which this sequence is not possible is equivalent to the well-ordering principle and strongly related to the MI principle. See Lolli (2008) for a proof.

[59] I wish to make a clarification. I am not saying that Fermat was the first one to reflect and describe the structure of a proof. Moreover, we have also to take into consideration that, compared to the previous analysed texts, this is extremely peculiar because it is taken from a private letter in which Fermat describes and summarises some findings. The precedent texts, instead, had not this particular nature. In any case, however, this excerpt highlights a further element not yet considered in our analysis, which is that, from an epistemological point of view, it is a relevant step when a proving process becomes a "method" to prove theorems, with a proper name and a structure that can be described in generic terms.

two arguments are equivalent, from an epistemological point of view a certain distance between the two seems to be present. Fermat's method involves a number of steps which is indefinite, but finite. Fermat writes that "descending ad infinitum" one comes to the number 5, but formally the descent starts from a generic but fixed number and thus ends after a finite number of steps. The step proved by Fermat, which is the contrapositive of the inductive step, represents, generically, one of these finite steps which links the starting number to 5. In a proof by induction, one intuitively refers to an infinite number of steps which links the base of the induction to every greater natural number. The inductive step, thus, represents generically one of these steps. To further clarify this point, I wish to quote some words by Unguru (1994). He is referring to Euclid, but what said seems to be appropriate for this point as well:

> *Descente infinie* and Induction are *not* conceptually the same procedure. The former is […] finite, while the latter is infinite, i.e., to put this more scrupulously, the method of infinite descent deals with a finite number of cases (this is why it works) while the method of mathematical induction covers an infinity of cases. (Unguru, 1994, p.271, italics in original)

This concludes the analysis of Fermat's method of the infinite descent or indefinite descent. It is interesting to note that the name which came down through history is the first one, even though, as we observed, the descent is never really infinite: in the case of negative propositions because it is assumed that such an infinite descent cannot exists, and in the case of positive propositions because the descent, after an finite number of steps, stops to specific number. Lolli (2008) decides to break this tradition and call the principle underlying this method the *Finite descent principle*, affirming that: "in all the literature this principle is called "infinite descent", but we refuse to adhere to this incongruence of giving a principle the name of what the principle itself affirms cannot exists." (p.44, my translation).

This section concludes this historical journey on the traces of MI which has involved several and temporally distant authors, and which made us pass through more or less explicit recursions, finite, indefinite or infinite iterations, definitions by recursion to arrive to Pascal's and Fermat's proofs. In this continuous metamorphosis it was possible to identify some elements which seem to have contributed to characterize what, in modern terms, MI is. In the following and last section of this chapter, I will analyse these elements, summarising what emerged from the historical-epistemological analysis.

2.9 CONCLUSIONS

I concluded the introduction of this chapter with the following question, which I will discuss in this section:

What mathematical aspects characterised the historical genesis of the proof by MI? In particular, what characteristics and turning points emerge from the traces of proofs by MI that the historiographic research has identified?

What follows is a summary of the elements observed in the previous pages and which contribute to answer the previous question. The presentation is organised with this criterion: firstly, I will present a series of elements related to some 'foundational' aspects of MI which I will call *basic characteristics*; then I will present some elements emerging by the analysis of the presented proofs, focusing on differences and similarities in their structure, which I will call *structural characteristics;* finally I will analyse how the contents of the propositions changed in the analysed authors, highlighting elements which I will call *characteristics of the propositions.*

A - Basic characteristics

(A1) Natural numbers as aggregations or as progressions.

From one side we registered the Euclidean definition of number: arbitrary aggregation of units, from the other side there is the definition given by Maurolico for the linear numbers: successive addition of a unit to the precedent element. The connection between this aspect and MI, as observed, is evident: one of the elements on which MI is founded is the structural role of the successor operation for the natural numbers. This aspect was one of the main points addressed by those scholars who considered as impossible the presence of MI in Euclid (Unguru, 1991, 1994).

(A2) Potential infinity or actual infinity.

We registered an epistemological distance between the "Prime numbers are more than any assigned multitude" by Euclid and the "And so on, to infinity" by Pascal. In particular in the first case, arbitrarily big but finite quantities of numbers are involved, whilst in the second case there is a reference to an infinite quantity of numbers for which the proposition is proved. This aspect is related to MI and in particular to its necessity. If the statement involves a finite number of cases (even if arbitrarily big) an iteration with a finite number of step could be sufficient to construct a prove for it, but when the statement involves an infinite number of cases, an iterative argument would involve an infinite number of steps (as in Maurolico) or a new strategy, as said by Pascal: "Despite this proposition has an infinity of cases, I will provide a shorter proof, thanks to the use of two lemmas".

(A3) Iteration

One of the recurring themes of this historical analysis was the presence of iterations both within proofs and as a way to construct definitions. These iterations and their generalization were described by the different authors in a similar way from, a linguistic point of view. In particular they seem to overcome the absence of an algebraic symbolism through which to represent in a compact form the steps of the iteration, by using the rhetorical technique of the repetition: the same sentence, with the same structure and the same words (except for the numbers involved) is repeated three, four times, before concluding with “and so on”. In this way they seem to suggest to the reader a *pattern* in the processes described to highlight the generalization of the iteration.

Despite this similarity between the authors, we have also registered some differences in the iterations constructed by the authors. For what observe in the previous point, Euclidean iterations, for instance, are finite, while Maurolico and Pascal refer to an infinity of steps instead; in Fermat also refers to iterations which develops for an indefinite number of steps.

B – Structural Characteristics

(B1) <u>Absence or presence of the parameter.</u>

An important difference between the traces of MI in the presented authors and a modern version of MI is that in the first case the parameter to which to apply induction, is still absent. This difference, as observed in relation to Euclid, is not simply syntactical but could highlight an epistemological distance between them. In other terms, a to express the statement in a parametric for is a characterising element of a proof by MI. Quoting Unguru (1991):

> The fact that number is always determinate in Greek mathematics prevents it from playing the role of independent variable in any numerical expression. Greek number is always a function of the units of measurement, it is always a «number of». For a genuine, not a potential, proof by induction, a «pure», «unadulterated», indeterminate, general, abstract number is needed and not a «number of». In other words, without a number that can serve as an independent variable, it is impossible to formulate a true proof by mathematical induction, in which the claim requiring proof is a function of the natural numbers. (p.284).

This aspect is present also in Fermat and Pascal. In their proof the parameter on which the induction (or the iteration) is constructed is not explicit. They refer to a base of the arithmetic triangle and the successive one or to a first, a second, a third triangle. In a modern formalization of their argument, probably a preliminary passage would be made by a mathematician, which is to introduce a series of indices for, respectively, the bases of Pascal’s arithmetic triangle or the element of the Fermat’s sequence of triangles, in order to make explicit the parameter on which induction is applied. This passage is not to be underestimated if we look at MI from an educational point of view: as observed in the literature review, to express the statement to prove in a parametric for or to determine the parameter on which to apply induction is often problematic for students.

(B2) <u>Generic example as a general case</u>

In all the analysed examples even if the statement was general, the proof was conducted for a generic example, a construction or argument involving a specific case which however can be generalise to all the other cases. In these terms a proof by generic example is not so distant to when, in a modern proof, we start the argument with the sentence “Let *n* by a generic natural number” and then we conclude with “since the genericity of *n*, this is true for all the natural numbers”. This reasoning finds its formal validity in the logical law of the *Universal Generalization,* which roughly affirms that a proof for P(n) with n generic in the universe of the discourse constitutes a proof for $\forall$n. P(n) [60]. Coming back to the examples analysed in this chapter we can see that in these cases, the role of the parameter *n* was taken by a specific number which however represents all the other natural numbers.

[60] The law of Universal Generalization and its relationship with MI will be further analysed in 3.3.2.

(B3) Infinite proof or finite proof.

Comparing the analysed authors, it is possible to register differences in proofs they constructed in relation to the finite or infinite number of steps required to conclude the proof. In the Euclid the proofs are finite: the iteration ends after a finite number of steps. in Maurolico and in the first couple of propositions by Pascal the proofs are, in a certain sense, infinite since the author explicitly refer to an infinite number of steps necessary to conclude the proof. In the other Pascal's proposition, however, the mathematician states that he could make the proof finite by using two lemmas, but the justification of this strategy is again given through an infinite number of steps. The problem of MI as a proof which involves a finite, or an infinite number of steps is delicate even from a logical epistemological point of view and it will be formally solved only with the axiomatization of the natural numbers by Peano and the introduction of the Principle of MI: a proof by MI is not valid for the infinite chain of modus ponens that one could construct starting from base and inductive step but because the principle of MI assures its validity. In other terms is exactly the principle of MI which avoids the infinite number of steps. However, despite this formal aspect, Pascal seems to have highlighted a central point of MI: the fact that the proof of two "lemmas" (the base case and the inductive step) substitutes the construction of an infinite number of steps.

(B4) Inductive step as an independent proposition

When presenting Maurolico's propositions I observed the presence of a particular structure in the proofs which I called 'bipartite structure': a principal statement involving a (infinite) series of natural numbers is proved firstly with a proposition involving two generic numbers, one consecutive to the other, and then with a second proposition in which the previous one is (infinitely) applied starting from the first natural number for which the general statement is true. These two propositions have been interpreted, in modern terms, as the inductive step (the first one) and the base of the induction with the successive application of the inductive step (the second one). The same structure has also been registered in the first couple of presented propositions by Pascal. In these examples, thus, the statement that can be interpreted as the inductive step is proved in an independent proposition, which is then used in the successive proposition. This structure is then changed by Pascal who merges the two propositions (called "lemmas") within the same and inverts their order, proving the inductive step after the base of the induction. It is interesting that in this second case, Pascal decides to provide to the reader a justification of its proof, describing how the two lemmas can be combined to prove the general statement. To summaries, thus, Pascal's proof is composed by three independent parts: the two lemmas with their independent proof and the justification of why the validity of the general statement to prove follows from the validity of the two lemmas.

(B5) Increasing or decreasing iteration

In the analysed excerpts we could observe the presence of increasing or decreasing iterations, or, in modern terms, iterations involving an increasing or decreasing numerical sequence. For what concerns the increasing ones, two different kinds of iterations were observed, namely the ones involving a finite (sometimes generic) number of steps and the infinite ones. The two cases, as observed, are not epistemologically equivalent. In particular they seem to involve a different view of infinity (potential or actual). Similarity, for what concerns the decreasing iterations it was possible to register iterations which were potentially infinite, as in the cases in which a contradiction of the well-ordering principle is reached, or finite (and indefinite), as in Fermat's method for the affirmative propositions. This last case has a structure which, from a logical point of view, is extremely similar to a proof by strong MI: a chain of implications (each the contrapositive of an inductive step) is constructed to connect a generic number with a specific number n_0 for which the statement is known to be true. However, despite the logical analogies between the two proofs, important differences are still present: the iteration is decreasing and finite in one case and increasing and infinite in the other.

C – Characteristics of the propositions

(C1) Statement with an implicit or explicit recursion.

Focusing on the kind analysed propositions we can register that in the most of them a recursion is explicitly contained in the statements themselves. All the propositions on the geometrical sequences in Euclid, as well as Maurolico's proposition or the first ones of Pascal are example of this kind, since they explicitly involve the proof of a property for the terms of a sequence or a progression of numbers. However, it was also possible to register the presence of propositions in which an iteration was constructed within the proof even if the statement did not involve it explicitly. This is the case, for instance, of the proof of the existence of a factorisation in prime numbers in Euclid, VII.31, or in Fermat's propositions which involved a generic number (or triangle) in the statement, or in the last Pascal's proposition in which the recursion was implicit in the definition of the arithmetic triangle but not in the statement of the proposition to be proved. This point seems to be important from an epistemological point of view, in particular it highlights that some mathematicians recognised that iterative procedures could be used to prove statements in which an iteration was not present.

(C2) Not necessity for trivial cases.

A last but not least important aspect to observe is the presence of propositions which are proved directly even though they could have been rapidly proved by the authors using the same recursive strategies that they had already used in other propositions. This are, for instance, Euclid's propositions IX.21-23 or the VII-VIII Consequences in Pascal. In the first ones, Euclid shows that the sum of any quantity of even number gives as a result an even number, proving it for a generic example (four generic even numbers) and not by iterating the fact that the sum of two even numbers is an even number. Pascal, instead, despite having a method to prove propositions with an infinite number of steps "thanks to the use of two lemmas", proves a proposition (Consequence VIII) by constructing an infinite iteration, as previously done by Maurolico. Both the examples of Euclid and Pascal, despite their differences, highlight that the construction of a recursion (in Euclid) or of a proof by Induction (as in Pascal) was not always the preferred way to prove a statement by them.

The ten just presented points (A1-3, B1-5, C1-2) represents an answer to the question which prompted this analysis. As we will see, these points will be reread and interpreted from a cognitive and didactical point of view and they will provide element that will be integrated in the conceptual framework of this thesis, presented in the next chapter.

3 ARGUMENTATION AND PROOF

The objective of this chapter is to theoretically frame this study within the wide scope of research in ME that enters under the name *Argumentation and Proof.* In doing so, beyond presenting the theoretic standpoint that characterizes this study, I will clarify the meaning of some key terms within this work, such as Argumentation, Mathematical Proof, and Mathematical Theorem.

3.1 INTRODUCTION AND PRELIMINARY DEFINITIONS

Nicolas Balacheff, in a paper of 2008, asks: "Is there a shared meaning of "mathematical proof" among researchers in mathematics education?" (Balacheff, 2008, p. 501). The response that Balacheff gives is substantially negative; in fact, he observes that:

> Currently the situation of our field of research is quite confusing, with profound differences in the ways to understand what a mathematical proof within a teaching-learning *problématique* is but differences which remain unstated. (Ibidem, p. 501, italics in original).

One highly contested aspect in the literature regards the relationship between argumentation and mathematical proof. As well described in Reid & Knipping (2010, cap.8) and in Stylianides, Bieda & Morselli (2016), in this research field many and diverse points of view have been adopted, which rest in stark contrast, even amongst themselves. Here I limit myself to two classic examples, between themselves very far apart, which highlight the variety of the perspectives adopted.

Some scholars highlighted the differences between argumentation and proof, from a social point of view (Balacheff, 1988) and from a logical point of view (Duval, 1991). In particular, according to the latter, argumentation (argumentative reasoning) and proof (deductive reasoning) are highly distant from each other from a logical-epistemological point of view (he speaks of "rupture" between them). According to him, whilst the argumentation is founded substantially on the semantics of the statements involved, the focus on a proof is instead on the relationships between statements, that is on the validity of the rules of inference used, therefore on the syntax. According to Duval this profound difference often remains hidden from both argumentation and proof are expressed with similar verbal structures, typical of everyday language. [61] According to this perspective, the argument can therefore represent an epistemological obstacle to grasp the sense of a mathematical proof for students.

Harel and Sowder (1998), on the other hand, in describing their model of Proof Schemes, adopt a very different perspective from that of Duval. They call *proving:* "the process employed by an individual to remove or create doubts about the truth of an observation." (Harel & Sowder, 1998, p. 241). This process is in turn divided into two subprocesses called *ascertaining* (removing one's doubts), persuading (removing the doubts of others). Differently from Duval, in the work of Harel and Sowder argumentation and mathematical proof are not considered epistemologically different, but rather they are understood as analogous processes of a person characterised by the same purpose, that of removing one's own doubts or those of others about the truth of an observation.

Duval's and Harel and Sowder's examples of above show how, in relation to the meaning of Argumentation and Proof, extremely distant points of view coexist. Due to such complexity and variety of perspectives, I believe it is necessary to clarify the theoretical point of view of this thesis work in relation to this point.

---

[61] 'Everyday language' is used here with reference to Ferrari (2004). Specifically, Ferrari uses the term 'everyday' ('quotidiano, in Italian) "to refer to the most common registers of verbal [oral and written] language" (p. 110, my translation).

The point of departure is to clarify the meaning with which I will understand the terms 'Argumentation' e 'Proof'.

In this work we will adopt the definition of ***Argumentation*** given by Douek (1999a, 1999b; 2002; 2007). In particular for Douek, the term "Argumentation" refers to two related things:

> First, it denotes the individual or collective process that produces a logically connected, but not necessarily deductive, discourse about a given subject. […]
> Second, it points at the text produced through that process. (2002, p. 304, bold added by me)

Douek, moreover, introduces the term ***Argument,*** adopting the Webster's dictionary definition**: "A reason or reasons offered for or against a proposition, opinion or measure"** (cited in 2002, p.304, bold added by me). Following Douek, therefore, "**an 'argumentation' consists of one or more logically connected arguments**" (ibidem, p. 304, bold added by me).

Some comments are necessary on the terms just introduced:

- First of all, we observe that in the definition of Argumentation we distinguish the argumentation process and the product of argumentation. The first is the process of producing a 'discourse' with the characteristics of the definition, while the second is the text (written or spoken) resulting from this process. This distinction between process and product that involves argumentation, as we will see below, will also be present when we talk about mathematical proof.
- Secondly, the term 'discourse' is used by Douek with an inclusive meaning. In another work, in fact, she writes: "the discoursive nature of argumentation does not exclude the reference to non-discoursive (for instance, visual or gestural) arguments." (2007, p.169). Therefore, if we refer to an argument as a process, it could be composed of non-verbal arguments, such as gestures or inscriptions. This point, as we will see later, will be particularly important in this study, in particular when we use a semiotic perspective as a lens for observing argumentation processes.
- The various arguments that make up an argumentation are 'logically connected' to one another. By this it is intended that these could be connected by deductions as well as by other logically invalid inferences (such as inductions, analogies, abduction, etc.).

Douek also takes a position in relation to the term "Proof". The researcher proposes the distinction between "Formal Proof" and "Mathematical Proof". With the first term she refers to a "proof reduced to a logical calculation" (Douek, 1999a, p. 128), a series of steps connected by the application of a particular inference rule of the logic theory. For Douek, a "formal proof" is what Duval (1991) has called "démonstration". With the term **Mathematical Proof**, on the other hand, Douek means "**what in the past and today is recognized as such by people working in the mathematical field**" (ibidem, p. 128, bold added by me). As Douek herself writes (1999a), this definition of 'Mathematical Proof' is very general:

> "This approach covers Euclid's proof as well as the proof published in high school mathematics textbooks, and current modern-day mathematicians' proofs, as communicated in specialized workshops or published in mathematical journals". (pp. 128-129).

In this thesis when I use the term Proof, I intend it to mean "mathematical proof" as described above. In a few cases in which instead it is referring to a "formal proof" it will be explicitly indicated with the addition of the adjective "formal".

The definitions of argumentation and proof given so far highlight a further aspect, extremely important in relation to Douek's position in the research debate on the relationship between argumentation and proof. The researcher, in fact, distances herself from Duval's position about the "rupture" between argumentation and proof. Following Douek, "**mathematical proof can be considered as a particular case of argumentation**" (1999a, p.129, bold added by me).

In consequence, as for argumentation, for proof we will also distinguish between product and process. In particular, following Douek's distinction (1999a), I refer to:

- Proving, as the process of constructing a mathematical proof (as defined above).

- **Proof,** as the oral or written mathematical text resulting from proving.

This distinction is central from an educational point of view. In particular, also considering the proving process with which a proof was obtained by a student, it is possible to take into consideration some aspects that otherwise would risk being hidden. What has just been said is applicable, in general, to any activity of construction of proofs, however I believe that it is further important if reread in relation to the construction of proofs for MI. A Mathematical proof by MI, in fact, has so to speak a canonical and well codified form ("induction base" & "inductive step"). This sometimes results in a multi-step procedure for building a proof by induction. As well recorded in the literature (Harel, 2001; Movshovitz-Hadar, 1993; Carotenuto et al., 2018), sometimes the difficulties of students are not so much in constructing a proof for MI correctly, as in constructing the correct mathematical meaning of such a proof. Di Martino (2018) observes how often the understanding of a proof by induction is at an instrumental and non-relational level, where the reference is the classic one to Skemp (1976). In conclusion, therefore, if we limited ourselves to analysing only the mathematical proof by MI produced by one or more students, without taking into account the proving process, the analysis would be too limited with respect to the complexity of the research problem from which this work is born

Having said this, in the following paragraphs, I will present a classical theoretical construct, that of *Cognitive Unit* (Boero et al., 1996, Garuti et a., 1998) and, within it, the notion of mathematical Theorem (Mariotti et al. 1997), which they will form part of the conceptual framework of this work. The construct of the Cognitive Unit*,* as we will see, arises from a series of studies that has the aim of investigating the relationship between argumentation and proof, on the one hand trying to overcome the dichotomy highlighted by Duval and described above, on the other hand not ignoring it the profound differences, but trying to exploit possible analogies and meeting points.

3.2 COGNITIVE UNITY

As said, some scholars highlighted the differences between argumentation and proof, from a social point of view (Balacheff, 1988) and from a logical point of view (Duval, 1991). These positions seem to suggest an epistemological distance, a rupture, between conjecturing and proving. On the contrary, some other scholars hypothesised that a continuity may occurs between the process of producing a conjecture and constructing its proof. Boero (1999) highlights how for an expert mathematician, the construction of a mathematical proof occurs through a series of phases, often intertwined:

> I) Production of a conjecture (including: exploration of the problem situation, identification of "regularities", identification of conditions under which such regularities take place, identification of arguments for the plausibility of the produced conjecture, etc.). […]
> II) Formulation of the statement according to shared textual conventions […]
> III) Exploration of the content (and limits of validity) of the conjecture […]
> IV) selection and enchaining of coherent, theoretical arguments into a deductive chain, frequently under the guidance of analogy or in appropriate, specific cases, etc. […]
> V) Organization of the enchained arguments into a proof that is acceptable according to current mathematical standards. […]
> VI) Approaching a formal proof. (Boero, 1999, p.1).

Boero, et al. (1996), record how similar characteristics seem to be present also in students involved in the production of conjectures and construction of proofs. The authors, in their analysis, distinguish two phases: the phase of exploration of the problem that leads the students to a conjecture; the phase of construction of a mathematical proof for the statement corresponding to the previously produced conjecture. In relation to this, they observe how a continuity can be recorded between these two phases. Particularly:

> - During the production of the conjecture, the student progressively works out his/her statement through an intense argumentative activity functionally intermingling with the justification of the plausibility of his/her choices;
> - During the subsequent statement proving stage, the student links up with this process in a coherent way, organizing some of the justifications (“arguments”) produced during the construction of the statement according to a logical chain. (Boero, 1996, p. 113).

The authors use the term Cognitive Unity to express this hypothesis of continuity between the production phase of a conjecture and that of the construction of a proof. In a subsequent work, however, Garuti et al. (1998), observe how this continuity may also not be present and how some difficulties for students in the construction of proof can be interpreted precisely in the terms of a “gap between the exploration of the statement and the proving process” (p.347).

In light of these new results, therefore, the term Cognitive Unity has been reformulated and this second definition is the one referred to today in the literature. As described by Mariotti (2006), therefore, the term **Cognitive Unity indicates**

> the possible congruence between the argumentation phase and the subsequent proof produced, clearly assuming that congruence may or may not occur. (p.183, bold added by me).

As observe by Mariotti (2006), the construct of Cognitive Unity allows us to overcome the dichotomy between argumentation and proof highlighted by Duval:

> The main strength of this construct is that of providing a way to escape the rigid dichotomy setting argumentation against proof: The possible distance between argumentation and proof is not denied but also not definitely assumed to be an obstacle; in this perspective, the essential irretrievable distinction between argumentation and proof is substituted by focussing on analogies, without forgetting the differences.” (Mariotti, 2006, p. 184).

Subsequent studies have enriched the construct with further elements. In particular, Boero et al. (2010), following Pedemonte’s studies on cognitive unity (2005, 2007), highlighted that this possible congruence between argumentation and proof could be analysed from two different points of view:

1) Considering the *referential system,* i.e. the knowledge system available to the subjects and used during the construction of a conjecture and the consequent proof (Pedemonte, 2005),
2) Considering the *structure* of the logical connections between statements.

Following this, the authors provide the following two definitions:

> there is **continuity in the referential system between argumentation and proof** if some expressions, drawings, or theorems used in the proof have been used in the argumentation supporting the conjecture. There is **structural continuity between argumentation and proof** when inferences in argumentation and proof are connected through the same structure (abduction, induction, or deduction). (Boero et al., 2010, p. 5, bold added by me).

3.2.1 Cognitive Unity: a focus on Mathematical Induction.

In the literature review I observed how Harel and Sowder (1998) differentiate two particular types of generalization sometimes used by students to convince themselves of the truth of a statement of the form $\forall n.P(n)$: the *result pattern generalization* (RPG) and the *process pattern generalization* (PPG). Harel (2001) observes how, although both types of generalization are based on an empirical induction (a statement in universal form is supported by the verification of a finite series of particular cases), in the second the focus is on the process that allows passing from one case to the next, an aspect that recalls the inductive step $P(n) \rightarrow P(n+1)$ of a proof by induction. This leads Harel to believe that this second type of generalization can be the starting point for constructing the meaning of a proof by induction.

Pedemonte (2007) re-read Harel's analysis with the construct of cognitive unity. In particular, she observed how a conjecture justified by an argument based on a RPG has a purely inductive (empirical) form, while a proof by MI has a deductive structure. Pedemonte argues that therefore there is a structural distance between the two of them. The researcher concludes that:

> An inductive argumentation based on a result pattern generalisation seems to be unsuccessful for the construction of a mathematical induction. […] The distance that students should cover in the case of an inductive argumentation based on a result pattern generalisation may represent an obstacle for them. (Pedemonte, 2007, p. 38)

However, the case in which the argument is based on a PPG is different. Pedemonte observes that in this case the conjecture with the form $\forall n.P(n)$ can be obtained by generalizing the following series of implications $P(1)$, $P(1) \rightarrow P(2)$, $P(2) \rightarrow P(3)$, …, therefore the argument can provide a structure on which to subsequently construct the inductive step $P(n) \rightarrow P(n+1)$. Pedemonte therefore, argues that in this case there can be a structural continuity between an inductive argument and a proof by induction. In the experimental study she conduces, she also observes how for all the groups of students who have managed to prove their own conjecture by induction, this conjecture was previously supported by an argument based on a PPG.

In a more recent study, Pedemonte and Buchbinder (2011) examine the role of producing examples in the proving process, with a focus on argumentation via RPG or PPG. The work is related to mathematical induction since the solution of the problem involved can be proved by MI. The researchers observe that although the production of examples leads students both to build argumentation by RPG and by PPG, only in the second case the students were then able to build a proof, even if no one built a proof by induction. The researchers therefore observed that, even in the case of arguments using PPG, a proof by induction is not built autonomously by students if they are not familiar with this proving scheme.

3.2.2 Cognitive Unity within this work

The construct of cognitive unity proved particularly useful in ME. First, it can be used as a theoretical lens to analyse cognitive processes involved in students' proof-building activities. In particular, it is possible to interpret some difficulties in the construction of a mathematical proof by students in terms of absence of cognitive units, at the referential or structural level, between the argumentation phase and the construction phase of a mathematical proof.

A second application is to provide some principles for the task design. In particular this construct suggests that suitable activities to introduce students to the construction of proofs could involve open problems that require formulating a conjecture and then proving it (that is an exploration phase and a proving phase). According to this perspective, therefore, a direct didactic implication is that of not limiting the activities of constructing mathematical proofs only for statements not personally conjectured by students, or to a mnemonic reproduction of theorems proved by the teacher or the textbook, highlighting instead the importance of the exploratory phase and the formulation of a conjecture.

The study presented in this thesis seeks to take into account both of these possibilities of using this construct. In particular, I will look at the arguments produced during different types of activities in students with different levels of experience. This will have, in the first place, a research objective, that of determining, in these argument-construction activities, some processes that are effective for the construction of argumentations or proofs by induction. Consequently to this, a didactic objective is added, that of determining some specific characteristics of the activities and tasks involved in them that are favourable for the construction of these effective processes.

3.3 THE NOTION OF MATHEMATICAL THEOREM AS A TRIPLET

Within the studies on cognitive unity on the complex relationship between argumentation and proof, a further element has been taken into consideration: the notion of Mathematical Theorem (Mariotti et al., 1997). The starting point is the consideration that a mathematical proof cannot be considered isolated from two other elements: the statement object of proof and a reference theory, within which the proof makes sense. Following this reasoning, therefore, "what characterises **a Mathematical Theorem is the system of statement, proof and theory**." (Mariotti et al., 1997, p. 182, bold added by me). This characterization highlights some central aspects relating to theorems, highlighted by the studies on cognitive unity: the statement of the theorem must also be considered in relation to how it was constructed (conjecture), and the proof and the reference theory within which it is constructed are intertwined with the argument which supported the conjecture.

The explicit reference to a reference theory as a characterizing element for a mathematical theorem allows to take semantic aspects into consideration. As described by Mariotti (2006):

> In their practice, mathematicians prove what they call "true" statements, but "truth" is always meant in relation to a specific theory. From a theoretical perspective, the truth of a valid statement is drawn from accepting both the hypothetical truth of the stated axioms and the fact that the stated rules of inference "transform truth into truth". […] For a mathematician, the existence and the reliability of a theoretical framework within which the proof of a statement is situated is unquestionable and tacitly assumed, even when it is not made explicit. On the contrary, for novices, the idea of a truth as theoretically situated may be difficult to grasp; however, this way of thinking cannot be taken for granted and its complexity cannot be ignored. (p. 184).

What Mariotti wrote highlights two points. In the first place, there is an overcoming of the argumentation vs proof distinction proposed by Duval (1991), cited at the beginning of the chapter, according to which the first is centred on the semantics of the statements involved while the second on the validity of the logical inferences involved (syntax). Following Mariotti's words, however, even a proof has a semantic component and on this parallelism there can be continuity between argumentation and proof. Secondly, however, Mariotti observes how this semantic value is theoretically situated. In other words, the epistemic value of the statements involved depends on the axioms and other theorems of the theory. This aspect, as highlighted, is delicate from a cognitive point of view and cannot be ignored when one has a didactic objective.

In summary, therefore, within this perspective a *Mathematical Theorem* consists in the system of relations between a statement, its proof, and a theory within which the statement and the proof make sense. In accordance with the notation introduced by Antonini and Mariotti (2008) **I refer to the notion of Mathematical Theorem above described as a triplet (S, P, T) whereas the letters stand for *Statement, Proof, Theory.***

In the next paragraph I present an example of application of the Mathematical Theorem construct, taken from Antonini & Mariotti (2008). With this construct, the authors propose an interpretative model with which to analyse the complexity of the proving scheme of indirect proofs (by contradiction or by contraposition). This example will be paradigmatic for the construction of a similar model for proof by MI that I will present later.

3.3.1 The notion of Mathematical Theorem: A model for indirect proof

Antonini and Mariotti (2008) used the construct of Mathematical Theorem to construct an interpretative model of theorem with a proof by contraposition or by contradiction (which together they call indirect proofs). The authors, analysing two examples of indirect proofs, observe that in these cases in order to prove the *principal statement* $S=p\rightarrow q$ what it is actually constructed is a direct proof of a statement S* which they call *secondary.* In the case of a proof by contradiction, $S^*=p\wedge\neg q\rightarrow r\wedge\neg r$, where r is any proposition, in the case of a proof by contraposition $S^*=\neg q\rightarrow\neg p$. Therefore, from a logical point of view, it is necessary to justify why the proof of S* can be accepted as a proof for S. This can be obtained by a new theorem which affirm that $S^* \rightarrow S$. As observed by the authors:

> This theorem [$S^* \rightarrow S$] is not part of the theory in which the *principal* and *secondary statements* are formulated, but it is part of the logical theory. Referring to their meta-theoretical status, we call the statement $S^*\rightarrow S$ *meta-statement*, the proof of $S^*\rightarrow S$ *meta-proof*, and the logical theory, in which the meta-proof makes sense, *meta-theory.* (Antonini & Mariotti, 2008, p. 405, italic in original).

Finally, generalizing the previous analysis, the authors provide the following model of indirect proof:

> [I]n any theorem with indirect proof we can recognize two theoretical levels, three statements, and three theorems:
>
> (1) the sub-theorem (S*, C, T) consisting of the statement S* and a direct proof C based on a specific mathematical theory T (Algebra, Euclidean Geometry, and the like);
>
> (2) a meta-theorem (MS, MP, MT), consisting of a meta-statement MS = $S^*\rightarrow S$ and a meta-proof MP based on a specific meta-theory MT (that usually coincides with classic logic);
>
> (3) the principal theorem, consisting of the statement S and the indirect proof of S, based on a theoretical system consisting of both the theory T and the meta-theory MT.
>
> We call indirect proof of S the pair consisting of the sub-theorem (S*, C, T) and the meta-theorem (MS, MP, MT); in symbols P = [(S*, C, T), (MS, MP, MT)]. In summary, an indirect proof consists of a couple of theorems belonging to two different logical levels: the level of the mathematical theory and the level of the logical theory. (Ibidem, p. 405).

Therefore, a theorem with an indirect proof becomes represented as a triplet in the following manner:

Theorem = (S, [(S*, C, T), (MS, MP, MT)], (T, MT)).

The model of indirect proof described above has permitted the authors to analyse and interpret cognitive and didactical issues related to indirect proofs and indirect argumentation. In the following paragraph I present a model of proof by induction constructed in an analogue manner.

3.3.2 The notion of Mathematical Theorem within this study: A model for proof by MI

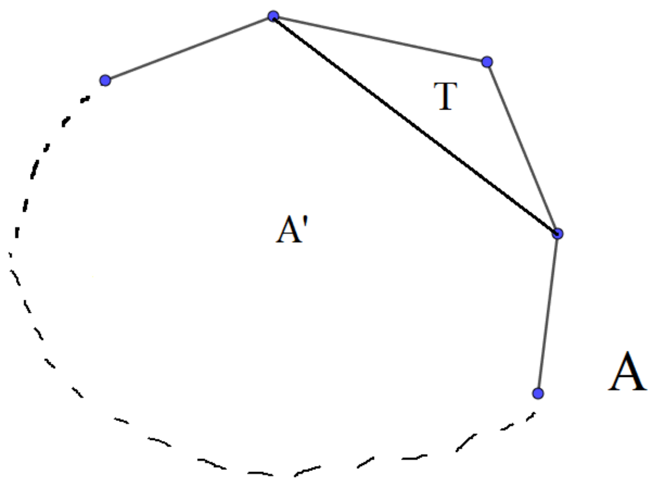
To facilitate the following analysis, let us take a paradigmatic example of a theorem with proof by induction.

<u>Statement.</u> The sum of the internal angles of a convex polygon of *n* vertices is equal to π*(n-2).

<u>Proof.</u> The proof is done for induction on n ≥ 3.

Base case: n=3. As noted in the previous theorem, the sum of the internal angles of a triangle is equivalent to π that, in turn, equivalent exactly to π*(n-2) with n=3.

Inductive step: we fix a generic n≥4, and we suppose that the sum of interior angles of a convex polygon of n vertices is equivalent to π*(n-2). We show that then the sum of interior angles of a convex polygon of n+1 vertices is equal to π*(n+1-2) = π*(n-1). Let A, then, be a convex polygon of n+1 vertices. Select one of these vertices and join with a segment the two consecutive vertices to the selected vertex. The polygon A is divided into a triangle T and a convex polygon A' with n vertices. The figure below shows this construction. Note that this construction works since n+1≥4.



By construction, the sum of the interior angles of A is given by the sum of the interior angles of T plus the sum of the interior angles of A'. Now the sum of the interior angles of T is π since T is a triangle, while since A' is a convex polygon with n vertices, the sum of its interior angles is π*(n-2), by inductive hypothesis. So, the sum of the interior angles of A is given by π + π*(n-2) = π*(1+n-2) = π*(n-1) which is what we wanted to prove.

From this we conclude that the sum of the interior angles of a convex polygon of n vertices is equal to π*(n-2). □

We can now analyse the theorem just proved with notations similar to those in the previous section.

The theorem involves a statement of the form ∀n≥3 A(n), where A(n) is the predicate "The sum of the interior angles of a polygon of *n* vertices is equal to π*(n-2)". The proof is developed by proving the conjunction of two statements S1 and S2, respectively the base case, corresponding to A(3), and the inductive step, corresponding to ∀n≥3 (A(n)→ A(n+1)). Statements S1 and S2 are proved with respect to a theory T corresponding, in this case, to Euclidean geometry.

In summary, schematizing what has been said, I observe how they are present:

- A statement S = ∀n≥3 P(n), which I will call *principal*.

- A statement S*= S1 ∧ S2 which I will call *auxiliary*[62], where S1 and S2 are two further statements independent of each other, S1=A(3) e S2=∀n≥3 (A(n) → A(n+1)) respectively.

[62] Antonini & Mariotti (2008), use the adjective 'secondary' for the statement S* which is proved in direct form within an indirect proof. In the model I propose in relation to MI, however, I choose to use the term 'auxiliary' because it is more evocative of the fact that the statement S*, corresponding to the conjunction of base and inductive step, serves to prove the main statement S.

- A proof P1 of statement S1, which in the text corresponds to the paragraph introduced by the expression "Base case".

- A proof P2 of statement S2, which in the text corresponds to the paragraph introduced by the expression "Inductive step".

- A theory T within which the statements S1 and S2 are proved.

From a logical point of view, deducing S from $S^*= S1 \wedge S2$ requires the validity of the statement $S^* \rightarrow S$. This new statement, however, is not proper to the theory T, in fact it corresponds to an application of the principle of induction, which more formally is a theorem in a particular logical reference theory (e.g., of ZFC set theory). In analogy with Antonini and Mariotti (2008), I will call the statement $S^*\rightarrow S$ *meta-statement*, the proof of $S^*\rightarrow S$ *meta-proof*, and the reference theory in which the proof makes sense *meta-theory*. I observe that the meta-theorem in question, which as mentioned is nothing but Principle of Mathematical Induction, does not refer to a specific predicate but to any P(n) defined in N. More formally, then, the *meta-statement* $S^*\rightarrow S$ should be written as 'For any predicate P(n) defined in N, $S^*\rightarrow S$' where S* and S are defined from P(n) as above. In order not to burden the notation, I will not make this second level of quantification explicit except when necessary.

A model for proof by mathematical induction

The analysis just made for the example above can be generalized to any theorem proved by induction, with reference to the (S,P,T) model of a mathematical theorem. In particular, in a theorem with statement $S=\forall n\ (A(n))$ and proof by induction it is possible to identify:

- Two *sub-theorems* (S1, P1, T) and (S2, P2, T), where $S1=A(n_0)$, $S2=\forall n\ (A(n)\rightarrow A(n+1))$[63], and P1, P2 are the respective demonstrations based on a specific theory T (Algebra, Euclidean Geometry, etc.).

- A *meta-theorem* (MS, MP, MT) where, respectively, *meta-statement* $(S1 \wedge S2) \rightarrow S$, and MP is the meta-proof of MS based on a meta-theory MT (e.g., ZFC or a theory for the arithmetic of N).

- The *principal theorem* (S, P, (MT, T)) consisting of the principal statement S and a *proof by induction* P of S, based on the theoretical system consisting of both the theory T and the meta-theory MT.

We call a *proof by induction* of S the triplet consisting of the two *sub-theorems* (S1, P1, T) and (S2, P2, T), and the *meta-* theorem (MS, MP, MT).

The model just described for a theorem by mathematical induction can be expressed in a more condensed way as it follows: **Theorem = (S, [(S1, P1, T), (S2, P2, T), (MS, MP, MT)], (T, MT)].**

What has just been described can be further expanded if we focus on the theorems involved in the model.

A focus on the theorem (S1, P1, T), the base case.

Frequently the basis of the induction corresponds to a simple verification of the predicate $A(n_0)$. Often this corresponds to verify a simple numerical equality or, as in the theorem taken as an example above, to recall a result already known. It should be noted, however, that even in these cases, it represents a theorem in a particular theory T of reference. In the proof by Mi presented above, for instance, this base

[63] If the basis of induction is n_0 then the statement S2 can become $\forall n\geq n_0\ (A(n)\rightarrow A(n+1)$. This is what happens, for example, in the polygon proof above. In order not to burden the notation I will denote the inductive step in universal form $\forall n\ (A(n)\rightarrow A(n+1)$. What I will say next will continue to apply in the other case.

case was the theorem stating that the sum of the internal angles of any triangle is equal to π. This aspect is explicitly taken into account by the presented model.

A focus on the theorem (S2, P2, T), the induction step.

The inductive step, i.e., the sub-theorem (S2, P2, T), has a complex statement: $\forall$n (A(n)→A(n+1)). It can be further rewritten as $\forall$n (Q(n)), where Q(n) is the predicate A(n)→A(n+1). The proof P2 of S2 suffers from this complexity. Generally, P2 has this structure: (a) first we fix a generic n$\in N$ greater than or equal to the base, (b) then we prove the validity of Q(n)=A(n)→A(n+1), showing how, assuming A(n), it is possible to deduce A(n+1). This concludes the proof of P2. From a logical point of view, steps (a) and (b) just described provide a proof of $\forall$n (A(n)→A(n+1)), based on some particular inference rules proper to classical logic. This point is discussed in more detail below.

A focus on the meta-theorem (MS, MP, MT), the Principle of MI.

As mentioned, a proof by induction is developed by proving two mutually independent statements, respectively the base case S1 and the inductive step S2. From this, it is then possible to apply the meta-theorem with statement (S1 ∧ S2) → S to prove the main statement S. It can be observed that, in practice, a proof by induction, generally, the reference to this meta-theorem is not present, except when it is stated that the proof will be conducted by MI. So, in a sense, the meta-theoretical issues related to the principle of induction are 'downloaded', more or less consciously, in the periphrasis "The proof is done by induction". This is fully within the practice of mathematicians and is accepted as legitimate by the scientific community. An important aspect that the model highlights, and which instead in practice might remain hidden, in brief, is that once the statements S1 and S2 have been proved, a further logical step is required to obtain a proof of S, and this step requires a particular (meta-)theorem.

A focus on the meta-theory

In a proof by induction, it is possible to identify additional meta-theoretical elements if we go to analyse some inference rules that are used to build the proof. These inference laws are particular theorems of the meta-theory of classical logic that formalizes deductive reasoning.[64] First, observe that the meta-theorem S*→S, which allows to prove the main statement, needs, in hypothesis, S1 ∧ S2; in the course of the proof by induction the statements S1 and S2 are proved independently. Thus, a logical step is necessary, one that states that a proof of S1 and one of S2 constitute a proof of S1 ∧ S2. Such a law of inference is traditionally called *introduction of the conjunction* and is generally represented as below:

$$\frac{\alpha\,,\,\beta}{\alpha \wedge \beta}$$

Secondly, we further analyse what happens during the proof of the inductive step S2. As noted above, the proof P2 generally occurs in two steps: (a) we fix a generic n greater than or equal to the base, (b) we prove the implication A(n)→A(n+1), taking A(n) as the hypothesis and deducing A(n+1). From a logical point of view, the fact that steps (a) and (b) provide a proof of n $\forall$n (A(n)→A(n+1)) rests on two laws of

[64] For the analysis that follows I will refer to *natural deduction* with classical inference laws. The theoretical reference I use is van Dalen, 2008.

inference: the *Universal Generalization* (or introduction of the universal) and the *introduction of the implication*.

The first states that if from a set of starting hypotheses Γ, which contain no assumptions about x,[65] it is possible to infer α(x), with x being a free variable in α, then from the same set of hypotheses it is possible to infer ∀x(α(x)). Such a law is formally represented as below:

$$\frac{\begin{array}{c}\Gamma \qquad \textit{[with } x \notin FV(\Gamma)] \\ \cdots \\ \alpha(x)\end{array}}{\forall x\,(\alpha(x))}$$

This inference rule is used in the proof of the inductive step applied to α(n)=A(n)→A(n+1). The inductive step is in fact proved for a generic *n*, where *n* corresponds to the free variable above and from that, by Universal Generalization, follows ∀n (A(n)→A(n+1)).

The second law of inference that characterizes the proof of the inductive step is the one known by the name of introduction of implication. It states that if from a set of starting hypotheses Γ, to which we add the hypothesis α, it is possible to infer β, then from Γ it is possible to infer A → B. What this law of inference tells us less formally is that in order to prove A→B we can take A among the hypotheses and deduce from that B. Formally this law of inference is represented as below; the square brackets around α in the premises of the inference indicate that α can be eliminated and is incorporated into α→β.

$$\frac{\begin{array}{c}\Gamma, [\alpha] \\ \cdots \\ \beta\end{array}}{\alpha \rightarrow \beta}$$

This inference rule is used in the proof P2 of the inductive step to prove the implication A(n)→A(n+1). In this step, in fact, we assume the hypothesis A(n), which is added to the hypotheses Γ already present in the main statement (for example, in the theorem presented above, one of these was that the polygon in question was convex). This step is in a sense evoked by the use of the expression *Inductive Hypothesis*, which shows that A(n) is an additional hypothesis to the hypotheses Γ of the statement S. At this point we deduce the validity of A(n+1) from the hypotheses Γ and A(n). Once this is done, through the inference rule of the introduction of the implication, we conclude that, starting from Γ, it is possible to deduce A(n)→A(n+1). As noted above, this does not conclude the proof of the inductive step, it is in fact necessary to intervene another inference rule, the Universal Generalization, so as to conclude that since the implication A(n)→A(n+1) has been proved for a generic n, then the statement S2=∀n(A(n)→A(n+1)) has been proved.

Finally, we note that this analysis does not exhaust all possible inference rules contained in a particular proof by induction. For example, the implication of the inductive step A(n)→A(n+1) could be proved indirectly, and this formally requires other inference laws. Moreover, the proofs of the theorems

[65] More formally, it means that x is not a free variable in none of the hypotheses Γ.

corresponding to the base case and the inductive step are constructed by relying on all the classical inference laws of natural deduction. The analysis just presented, while not exhaustive, nevertheless allowed to highlight a structural complexity from a logical point of view that a proof by induction has, regardless of the complexity of the predicate A(n) involved in it.

In the next paragraph I present a cognitive analysis that emerges from the model presented above and the logical analysis just conducted. This paragraph aims to draw attention to some aspects that I believe highlight the complexity, at the cognitive level, of a proof by MI.

3.3.3 A first cognitive analysis emerging from the model

A first delicate aspect that the model highlights is the relationship between the theorems (S1, P1, T) and (S2, P2, T). The proof of both is necessary for the proof of S since it allows to apply the meta-theorem S*→S, corresponding to the Principle of MI. I have already observed how, from an intuitive point of view, a possible justification of the correctness of the principle of induction can be obtained from a series of *Modus Ponens* inferences: from A(0) and from A(0)→A(1) follows A(1); from A(1) and from A(1)→A(2) follows A(2); and so on. Thus, this argumentation also needs both statements S1 and S2. So, when statements S1 and S2 are read within the meta-theorem S*→S they must be cognitively perceived in conjunction with each other and both necessary for the proof of S. As pointed out in the literature (Avital & Libeskind, 1978; Ernest, 1984) often, instead, the base case is not considered necessary by students for the proof but is considered a simple preliminary verification of the principal statement in a specific case. The model just presented allows this phenomenon to be described in terms of a non-conceptualization of the theorem (S1, P1, T) as independent of the main theorem and a non-conceptualization of the meta-theorem S*→S which requires the validity of S1 in order to be applied.

A second aspect that is intertwined with the previous one is that the theorems (S1, P1, T) and (S2, P2, T) are mutually independent and are proved independently of each other. The historical-epistemological analysis that I presented in Chapter 2 showed that this aspect was present in the "traces" of induction from history. I used the term *Bipartite Structure* to highlight how, in some propositions of Maurolico and Pascal[66] a particular proving structure is outlined for which:

- An archetype of the inductive step is proved in a self-contained proposition.
- In the next proposition, what has just been proved is applied from what today we would call the base of the induction, to prove a proposition of the form $\forall$n (A(n)).

These two aspects just observed highlight an important aspect that emerges from the model: from a cognitive point of view, the triplets representing the theorems described in it need to be conceptualized both as independent and in mutual relation.

A further point to note is the following: the statements S=$\forall$n(A(n)), S1=A(n_0) and S2=$\forall$n(A(n)→A(n+1)) all involve the predicate A. Although the three statements state different things, nevertheless there is an intertwining between some of their components that from a cognitive point of view could create a kind of *interference*[67] between the various statements and give rise to difficulties or misconceptions. Let's see in detail some aspects.

[66] See section 2.9.

[67] The term *interference* is used here with metaphorical allusion to the physical interference between plane waves: different waves that, when considered simultaneously, modify each other.

Interference between Statement S and Statement S1.

In the base case, we prove the statement $A(n_0)$ which however also corresponds to a particularization of the main statement $\forall n\ (A(n))$. From a cognitive point of view, this step can create some difficulties. For example, a student who finds a perfectly convincing inductive argumentation by RPG might consider the basis of the induction as an argument "reinforcing" the validity of the statement or as a part of a proof "by cases", or even as sufficient itself to prove the general statement. Conversely, however, a student who does not find the above RPG argument convincing might consider the demonstration of $A(n_0)$ unnecessary for the purpose of proving the general statement.

Interference between Statement S and Statement S2.

The statement $S2=\forall n(A(n)\rightarrow A(n+1))$ involves within it the predicate A(n) of the principal statement S. Moreover, as noted above, in the proof of the inductive step is generally done by showing that, for a generic n, it is possible to infer A(n+1) by assuming A(n) as hypothesis. We have already observed how, from a logical point of view, this is permissible for proving S2, however from a cognitive point of view if one does not have control over when the variable *n* is quantified (statement S) and when it is free (within the proof P2 of S2), the assumption of A(n) among the hypotheses could be interpreted the assumption of $\forall n(A(n))$ which is precisely the general statement. This aspect, as noted in the literature review, often leads to one of the most common misconceptions among students following which "a proof by induction is a proof in which one assumes what one wants to prove" (Avital & Libeskind, 1978, Ernest, 1984; Fischbein & Engel, 1989).

To summarise, thus, from a cognitive point of view it is necessary to conceptualize that the variable n is found to have different roles within a proof by MI:

- In the main statement S, it is quantified and represents **any natural number n**, thus the totality of the set $\mathbb{N}$.
- In the proof of $A(n)\rightarrow A(n+1)$ in P2, instead, it represents **a generic, particular, natural number**. It is **particular** because it is a **single number** so that *n+1* is **its** successive natural number, but it is also ***generic,*** in the sense that *n* could be any natural number.
- In the statement $S2=\forall n(A(n)\rightarrow A(n+1))$, the variable *n* is again quantified and represents again the totality of the set $\mathbb{N}$.

In this regard, I observe how, in a classic work, Mason and Pimm (1984) point out that determining when an algebraic expression indicates a *general, generic,* or *specific* case is extremely delicate and potentially problematic from a cognitive point of view. The authors' starting point is to analyse the following example. Generally, the algebraic expression 2*n* is used to denote a generic even number. However, they ask, "What does 2n stand for? What is n?" They observe that on the one hand, "2n is standing for any even number. It is a shorthand for, or a name for {2n: n a whole number}", on the other, "2n is a particular, but not specific even number", in the sense that *2n* is an even number but "not a specific even number like 2, 4 or 6." (p.280). In essence, then,

> There seems to be a dual perception invoked by the use of the symbols 2N, and it is these shifts of perception which may lie at the heart of some algebraic difficulties. (Ibidem, p. 280)

In particular, what the authors observe, is that in mathematics we often make use of expressions such as *2n* either to denote a fixed even number, a generic and any even number *2n*, or every even number. Depending on the context and the use that is being made of it, the mathematician knows how to dissolve this apparent ambiguity of expression. The authors, however, observe that this aspect is not obvious from a cognitive point of view, but rather that often the difficulties of students can be read precisely in relation to this. The authors bring as a paradigmatic example the case of a generic student who, struggling with the statement "the sum of any two even numbers is even", constructs the proof writing "Let 2N be any even number. Then 2N+2N=4N which is even" (p.283). Mason and Pimm analyse this situation as follows:

> The 'any' is still causing some difficulty. It is very reasonable to argue that since 2N stands for ANY even number, the sum of ANY two even numbers can be represented by 2N + 2N. What seems to be missing is an awareness of 2N as a PARTICULAR but non-specific even number. […] Thus using 2N and 2N is using the same particular even number twice, and not ANY two particular erven numbers. (Ibidem, p.283, capital in original)

Rereading what was said by Mason and Pimm, in relation to what was said above about the proof of the inductive step, it can be seen that in the transition from A(n) to A(n+1) a student is faced with a situation analogous to that just described. It is necessary then to be aware that with the inductive hypothesis we assume the validity of A(n) for a particular natural number, so we are not supposing A(n) for every natural number n, which corresponds exactly to the thesis of the main statement S. At the same time, however, in A(n)→A(n+1), the variable *n* is generic, so that we can use UG and conclude $\forall$n (A(n)→A(n+1)).

Interference between Statement S1 and Statement S2.

As noted, to prove the implication contained in the inductive step one also assumes among the hypotheses A(n) for a generic n, and from this one deduces A(n+1). As mentioned, assuming among the hypotheses A(n) is justified by the law of inference of the introduction of the implication. However, it is possible for a student to believe that assuming the truth of the inductive hypothesis is justified precisely from the basis of induction. In particular, the latter guarantees that the set A for which A(n) is true is other than empty, i.e., that there is at least one number for which A(n) is true. At this point, then, it becomes legitimate to assume that A(n) is true because at least one number for which this happens exists. According to this interpretation, then, the base of induction becomes a warrant for the assumption of the inductive hypothesis in the proof of S2. In truth, as we have already observed, from a mathematical point of view, the statements S1 and S2 are mutually independent and, in particular, the proof of S2 does not necessitate the validity of the base case. An example of this kind of interference can be found in Fischbein and Engel (1989), who, in a study that I will analyse in detail later (see section 5.2.4), record the following statement of a student in relation to the proof of the inductive step: "If the truth of the basis has been confirmed one may assume that the number k, which appears in the inductive hypothesis, is the same initial number and then the statement is true for k" (p.283). As noted in the previous point, in the proof of A(k)→A(k+1), the variable k represents a particular but generic number, in the answer of the student instead it represents a particular and specific number, just the basis of induction.

3.3.4 The use of the model in this study

Before concluding this section, I would like to clarify the use I intend to make of this model within this study.

First, the model, has allowed me to establish a terminology (*Principal Theorem, Auxiliary Theorem, Meta-theorem*, etc.) that will be used from now on, and to which, I will also refer in relation to other theoretical constructs.

Second, it will be used as a tool for research design. As we will see, some methodological choices, in particular on the tasks and problems used, are made taking into account some elements that the model just presented has highlighted, such as the conceptualization of the various triplets, both as independent and in relation to each other. In this regard, it will also provide a research tool with which to identify, describe, analyse, and interpret student difficulties. The analysis just presented has allowed us to highlight some points that seem to be most delicate from a cognitive point of view. In the second part of this thesis, I will further explore this analysis in light of the data collected in the experiments conducted.

Finally, the model will be used as an instructional design tool in designing activities with students to promote proof by MI. In particular, the model has highlighted some specific aspects of proof by induction, on which targeted activities can be built, e.g. with focus on Base case and Inductive Step as particular mutually independent theorems, on the meta-theorem S→S*, or on the different role of the predicate A(n) in the different statements S, S1, S2, S→S*.

3.4 Recursive argumentation

So far in this thesis, when discussing those argumentations that are interpretable as related to mathematical induction, we have generically used the term 'recursion argumentation'. The goal of this short section is to define, through the terminology introduced in this chapter, what I will more formally mean by 'recursive argumentation' from now on in this study.

I recall the definition of 'Argumentation' according to Douek (2002), described at the beginning of the chapter and adopted in this study:

> First, it denotes the individual or collective process that produces a logically connected, but not necessarily deductive, discourse about a given subject. […] Second, it points at the text produced through that process. (2002, p. 304)

Moreover, “an ‘argumentation’ consists of one or more logically connected arguments” (ibidem, p. 304), where “argument” is defined as “a reason or reasons offered for or against a proposition, opinion or measure” (from Webster Dictionary, cited in Douek, 2002, p.304).

Following this definition, therefore, an argumentation consists, among other things, **of one or more arguments**, **of one or more statements** to which each argument refers by offering "a reason", **and of possible links between each argument**.

On the basis of these elements, it is possible to analyse the structure of those arguments that have so far been referred to generically as 'arguments related to mathematical induction'. Two central features of these seem to be respectively:

- The repetition of a series of similar arguments in support of a series of statements.
- The fact that the various arguments are linked to each other, so that a statement supported by a given argument is then used in the argument to support the next statement.

We can consider, for example, the informal justification of the validity of MI as a chain of syllogisms: "P(0) and P(0)→P(1) therefore P(1), P(1) and P(1)→P(2) therefore P(2), and so on". In this case there is a repetition of arguments, each corresponding to an inference via Modus Ponens, and moreover, in this repetition, every statement (P(1), P(2), P(3), ...) is used in the argument that supports the next statement.

A similar analysis can be made for those arguments related to Fermat's method of infinite descent (as, for example, the justification given by Fermat himself, see section 2.8). In this case, within a proof by contradiction, it is possible to observe the presence of an argumentation with a structure similar to the previous one: from P(n*) and from P(n*)→ P(n*-1) it follows P(n*-1), then from P(n*-1) and from

P(n*-1)→P(n*-2) it follows P(n*-2), and so on indefinitely until obtaining a contradiction. Similarly to the previous case, it is possible to observe the presence of a repetition of arguments connected to each other so that a statement is used in the argument that supports the next statement.

Starting from these two paradigmatic examples and with reference to the definition of argumentation above, it is possible to give a definition of recursive argumentation.

First, I introduce the definition of **Repeated Argumentation**:

An argumentation which involves or refers to a repetition of an argument that support successive statements.

At this point, a recursive argument can be defined as a particular case of repeated argumentation in which each statement is used in the argument to support the next statement.

In particular, I define **Recursive Argumentation:**

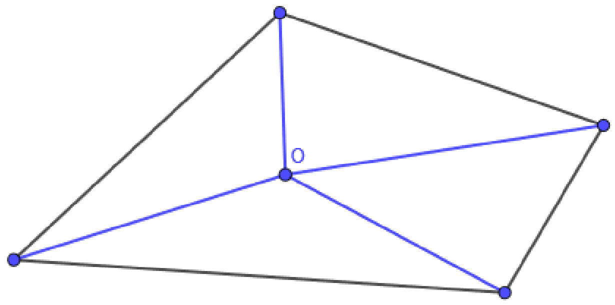
An argumentation which involves or refers to a repetition of an argument that supports successive statements, so that each of these statements is used in the argument supporting the following statement.

This one and the definition of repeated argumentation are inspired to the definition of Recursive Reasoning by Stylianides, Sandefur and Watson (2016): "Reasoning relating to or involving the repeated application of a rule or procedure to successive results" (p.21). In particular, in our case, the expression "repeated application of a rule or procedure to successive results" has been adapted to the case of argumentation and therefore taking into account those cardinal elements that emerge from Douek's (2002) definition adopted in this work (arguments, statements, and links between arguments).

To further clarify the distinction between repeated and recursive argumentations just defined, let us look at an example of a repeated argumentation that is not recursive. Consider the following argumentation for the statement 'The sum of the interior angles of a convex polygon with *n* sides is equal to π*(*n*-2)':

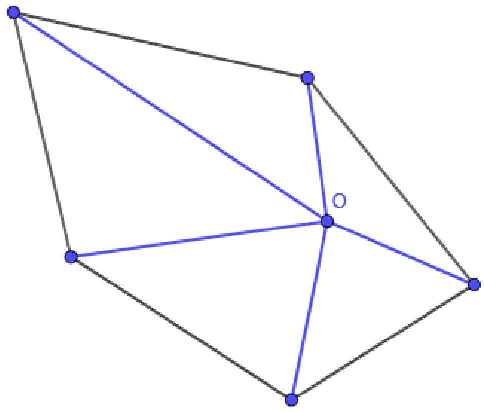
- if n=3, the polygon is a triangle, so the sum of the interior angles is exactly $\pi=\pi*(3-2)$, which is the formula to be proved for n=3.

- If n=4, as shown in the figure below, taking a point O inside the polygon we can draw the four segments connecting each vertex of the polygon to the point O, obtaining four triangles.



Now, the sum of the interior angles of the polygon with four sides can therefore be obtained by adding the interior angles of each of the triangles that compose it and subtracting 2π, which corresponds to the turn angle in O. Thus, we obtain $\pi*(4)-2\pi = \pi*(4-2)$ which is the formula to be proved for n=4.

- If n=5, as shown in the figure below, taking a point O inside the polygon we can draw the five segments connecting each vertex of the polygon to the point O, obtaining four triangles.



Now, the sum of the interior angles of the polygon with five sides can therefore be obtained by adding the interior angles of each of the triangles that compose it and subtracting 2π, which corresponds to the turn angle in O. Thus, we obtain $\pi*(5)-2\pi = \pi*(5-2)$ which is the formula to be proved for n=5.

- Continuing in the same way for any convex polygon of n sides, we obtain that the sum of its interior angles is $\pi*(n-2)$.

It is possible to observe that in this argumentation, each statement corresponding to the case of a polygon with a fixed number of sides, with the exception of the one involving the triangle, is supported by an argument, whose structure is analogous in the various cases. Therefore, it is a repeated argumentation. However, none of these statements is used in the argument to support the next one. Thus, the argumentation is not a recursive following the given definition.

The example just presented is also an example of an argumentation involving a process pattern generalisation, since the general statement is supported in the last point by generalising the process of decomposing the polygon into triangles to any convex polygon. This, therefore, highlights that the definition of recursive argumentation given above does not correspond to that of argumentation by process pattern generalization. This does not mean that the two forms of argumentation are mutually exclusive. In fact, later on, I shall present an example of an argumentation by process pattern generalization that is also a recursive argumentation.

Figure 3.1 below schematises the definitions of repeated argumentation and recursive argumentation given above.

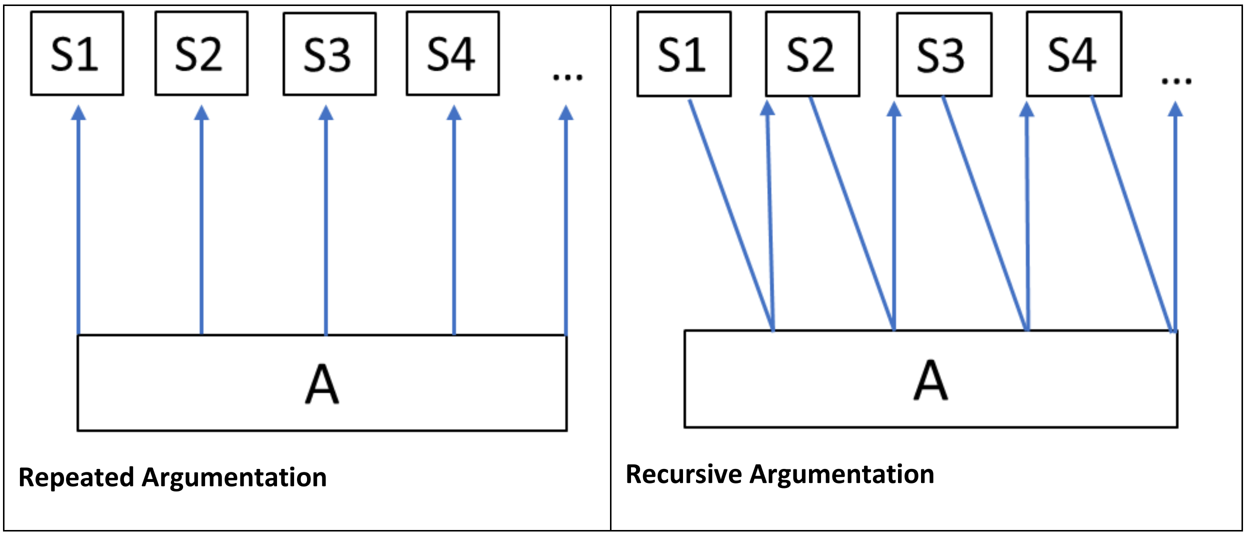


Figure 3.1. *The structure of a repeated argumentation on the left and of a recursive argumentation on the right. In both cases the argument A successively supports a series of statements (S1, S2, …), but in a recursive argumentation each statement is then used to support the following one. Note that the number of each statement refers to the order in which it is considered within the argumentation, and not to the number involved in the predicate P(k) to which the statement is referring.*

A few comments are needed on the above definition of recursive argumentation. Firstly, notice that it was written that it *'involves or refer to'* a repetition of an argument.

By this I mean that in the discourse containing the argument it may be that the repetition of an argument in support of subsequent statements is explicitly present, but it may also be that this repetition is only referred to. In the informal justification of the validity of mathematical induction as a chain of inferences, for example, the argument by MP is repeated twice and then a repetition of it is referred to with the words 'and so on', so in this example both a repetition of an argument is involved and a reference to it is made. In other cases, a recursive argumentation might only refer to a repeated argument, without this repetition actually being present. For instance, let us consider the following excerpt taken from Harel (2001), who presents it as an example of argumentation by Process Pattern Generalization, of a student dealing with the problem

'Prove that for all positive integers n, $\log(a_1 \cdot a_2 \cdot \ldots \cdot a_n) = \log a_1 + \log a_2 + \ldots + \log a_n$':

(1) $\log(a_1 \cdot a_2) = \log(a_1) + \log(a_2)$ by definition

(2) $\log(a_1 \cdot a_2 \cdot a_3) = \log(a_1) + \log(a_2 \cdot a_3)$. Similar to $\log(a \cdot x)$ as in step (1), where this time $x = a_2a_3$.
Then
$\log(a_1 \cdot a_2 \cdot a_3) = \log(a_1) + \log(a_2) + \log(a_3)$

(3) We can see from step (2) any $\log(a_1 \cdot a_2 \cdot \ldots \cdot a_n)$ can be repeatedly broken down to
$\log a_1 + \log a_2 + \ldots + \log a_n$.
(p.192)

Note that in this argumentation there is only a reference to a repetition of an argument. In particular, the argument in (2) is not subsequently repeated for later cases, but in (3) there is a reference to the fact that it can also be repeated for the higher cases, so that "any $\log(a_1 \cdot a_2 \cdot \ldots \cdot a_n)$ can be **repeatedly** broken down to $\log a_1 + \log a_2 + \ldots + \log a_n$.".

This example introduces another important aspect of the definition of recursive argumentation, namely that the reference to the repetition of an argument is something subjective. We, as researchers with experience of recursive argumentations and of proofs by MI, can observe that a student's argumentation is referring to the repetition of an argument, but it could be that the student is not aware of this. For instance, we could interpret the expressions "and so on", "etcetera" or "…" as referring to a repetition of the previous arguments, but a student could see in those something else, such as a generalization by empirical induction. This comment highlights, thus, that the awareness of the repetition of an argument is a central element to be considered when analysing a student's argumentation.

A further aspect to be clarified is that the fact of being 'recursive' for a repeated argument also depends on the statements considered. Indeed, if instead of the statements P(n) as n varies, one considers the statements P(n)→P(n+1) then a possible recursive argumentation can be interpreted simply as a repeated argumentation. In fact, each implication P(n)→P(n+1) is not used to prove the following one P(n+1)→P(n+2).. In particular, therefore, in the case of a statement of the form '∀n. P(n)', shifting the attention from the statements P(n) to those P(n)→P(n+1) we still have a repetition of a series of arguments (which eventually can be generalized to support the statement '∀n. P(n)→P(n+1)) without there being a recursion, that is, without each statement being used in the argument to support the next statement. In a certain sense in this way a complexity is added to the structure of the series of statements considered (now they are implications) but complexity is removed from the argumentation, which goes from being a recursive argumentation to being simply a repeated argumentation. This point, similarly to the previous one, highlights the central role of the subject interpreting the argumentation. The same argumentation can be interpreted as recursive or not by two different subjects.

It is also to observe that that, following the given definition, in a recursive argument it is possible to recognise a structure similar to that of the model theorem with a proof by induction developed from the construct of theorem as triplet. In a recursive argumentation, in fact, it is possible to recognise a main argumentation, an auxiliary argumentation (the argument which is repeated), and possibly also a meta-argumentation, which is the one that guarantees that the auxiliary argumentation supports the statement involved in the main argumentation [68]. Let us consider, for example, the following recursive argumentation, a *Quasi-induction* to say it with Harel (2001), that supports the statement

S = 'The terms of the sequence $\sqrt{2}$, $\sqrt{2+\sqrt{2}}$, $\sqrt{2+\sqrt{2+\sqrt{2}}}$, … are all lower than 2'

The statement S is true, for:

- The first term of the sequence $\sqrt{2}$ is lower than 2.
- From the previous point it follows that $2+\sqrt{2}$ is lower than 4, therefore also the second term of the sequence $\sqrt{2+\sqrt{2}}$ is lower than 2.
- From the previous point it follows that $2+\sqrt{2+\sqrt{2}}$ is lower than 4, therefore also the third term of the sequence $\sqrt{2+\sqrt{2+\sqrt{2}}}$ is lower than 2.
- And so on for all the terms of the sequence.

In this case, in the argumentation supporting the statement S, it is possible to recognise the repetition of an argument (auxiliary argumentation) which is the one supporting the fact that if one term of the sequence is less than 2 the next term is also less than 2. We can observe that this auxiliary argumentation is explicitly involved in the first cases and then referred to as 'and so on' in the last line. In this case, the

[68] I wish to clarify that I am not saying that these three elements must always be present in a recursive argument, but that they might be.

meta-argumentation (not present) would be the one supporting the fact that in this way all the terms of the sequence are actually reached, so that the truth of S can be concluded from the auxiliary argumentation and the fact that the first term of the sequence is lower than 2.

A further important point to be clarified is the relation between a recursive argument that emerges from the definition and a proof by MI. In particular, it is possible to consider a proof by induction as a particular case of a recursive argumentation, but some clarifications are needed. According to the given definition, a recursive argument involves or refers to a repeated argument. In a proof by MI this reference is present in the expert mathematician who proves by induction. In the latter, in fact, there is an awareness that the inductive step can be 'separated' into an infinite series of steps and that each of these can be applied from the base case until any natural number greater than the base is reached. In this case, then, a proof by MI can be recognised as a particular recursive argument. What is important to note, however, is that these aspects are not explicit in the text of a proof by induction, so the subject's interpretation and awareness of the chain of infinite syllogisms behind a proof by MI becomes central. In other words, the definition of recursive or repeated argumentation is subject-centred, in sense that different subjects might consider the same argumentation in different ways. A student who produces a correct proof by induction but does so 'blindly', without recognising the possible recursion behind it, is not arguing recursively. Examples of this type are well recorded in the literature. In section 1.3, for instance, I showed that Harel (2001) records how often students construct proofs by MI in a ritual way without being aware of its recursive nature. This last comment highlights a didactically relevant aspect, namely the importance for a student to recognise the fact that there is recursion behind a proof by induction. This point is crucial from the perspective of the cognitive unity, since if a subject 'sees' a continuity between a recursive argumentation and a proof by induction, then he or she may become aware that a proof by induction can be used to formalise a recursive argumentation that has been produced informally during the exploration of a problem. This is in line with the literature, and it allows us to reinterpret some of the proposed didactic interventions presented in the literature review (Section 1.3.3). These, in fact, seem to be based on the premise that, using our terminology, it is crucial from a cognitive and didactic point of view that a subject recognises the recursive argumentation structure in a proof by induction. For example, the didactic interventions of Dubinsky (1986, 1989) or Harel (2001) can be described as composed by activities in which students were involved in the construction of recursive argumentations as an introduction to the proofs by MI.

Before concluding this section, it is interesting to observe how, in an opposite way, a proof by induction might not be seen in connection with recursion even by an experienced mathematician. In fact, when the MI is considered as a formal law of inference, i.e. a particular theorem of the meta-theory of reference, then the connection with a recursion is not necessary from a logical-mathematical point of view. It is, in fact, justified as a theorem of this meta-theory. In this case, therefore, if we consider a proof by MI as a 'formal proof', it no longer corresponds to a recursive argumentation. The matter changes, instead, if we consider more generally a proof by MI as a 'mathematical proof' (we have dealt with the distinction between 'formal proof' and 'mathematical proof' in 3.1). In this case, in fact, the historical analysis of the development of MI has shown that from an epistemological point of view there is a deep connection between mathematical induction and recursive argumentation, with the former historically developing almost as a formalization of the latter. Therefore, even if from a formal logical point of view the connection between recursion and MI is not necessary, nevertheless from a historical-epistemological point of view it is present and thus it should not be ignored when considering MI from a didactical point of view.

4 APOS THEORY

4.1 PREMISES AND INTRODUCTION

The APOS Theory is a classical theory of mathematics education that has made it possible to bring deep contributions to the study of learning and understanding processes related to MI (Dubinsky 1986, 1989, 1991; Dubinsky & Lewin, 1986; Garcìa-Martinez & Parraguez, 2017). As we will see, within this theoretical perspective it is possible to construct “a hypothetical model that describes the mental structures and mechanisms that a student might need to construct in order to learn a specific mathematical concept” (Arnon et al., 2014, p. 27), or what in APOS Theory is called a *Genetic Decomposition (GD)* of a mathematical concept. In other terms a GD has the aim to offer a model to describe a cognitive counterpart to the logical-mathematical structure of a particular concept. In the specific case of MI, thus, the APOS perspective allows to take into consideration, from a cognitive point of view, its logical and epistemological complexity.

In this section I will firstly describe the fundamentals of the APOS theory and introduce its terminology; then I will present the GD of MI which is described in the related literature; finally, starting from the historical-epistemological analysis described in chapter 2, I will propose a GD of MI which expands and deepens the previous one, including some aspects that have been highlighted by the historical-epistemological analysis.

4.2 DESCRIPTION OF APOS THEORY

The APOS theory had a long development in mathematics education. In the mid-80s, in fact, Dubinsky and his collaborators, by examining Piaget’s works on the *genetic epistemology*, introduced what, later, would be the key points of APOS Theory. In the successive years these points have been further investigated and modified. In this thesis I will refer to the APOS theory as described in the monography “APOS Theory” (Arnon et al., 2014) which provides a comprehensive presentation of a unified version of the theory. A complete report of the historical development of the theory is presented in the second chapter the quoted monography.

As described by the authors,

> APOS Theory is principally a model for describing how mathematical concepts can be learned; it is a framework used to explain how individuals mentally construct their understandings of mathematical concepts” (Arnon et al. 2014, p.17).

APOS models the construction of mathematical knowledge from, as called by the authors, four **mental structures**: *Actions, Processes, Objects,* and *Schemas* (from which the acronym APOS), and several **mental mechanisms** that lead to their construction*: interiorization, coordination, reversal, encapsulation,* and *thematization.* Figure 4.1 shows an overview of the structures with the respective relationships between them.

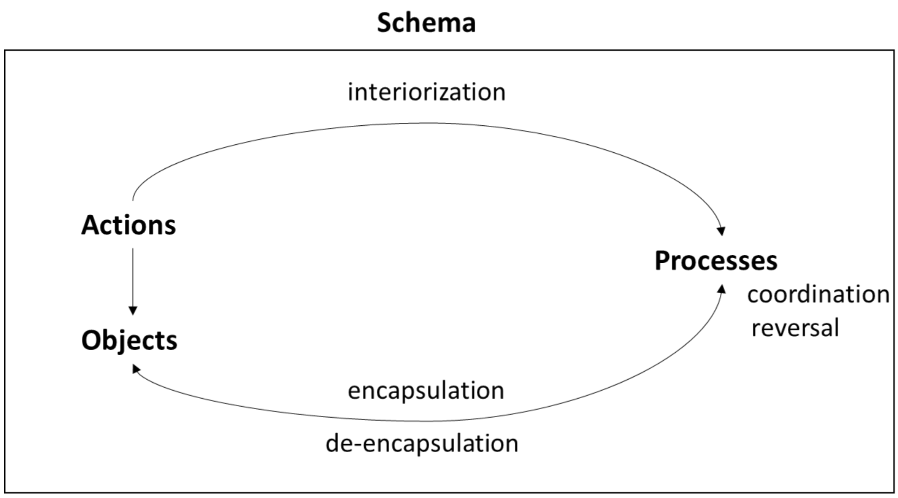


Figure 4.1. *Overview of APOS structures and their interactions. The figure is constructed as in Arnon et al., 2014, p.18.*

4.2.1 Actions

A subject firstly constructs a concept as an **Action, that is "an externally directed transformation of a previously conceived Object, or Objects."** (Arnon et al., 2014, p.19). The transformation is 'externally directed' in the sense that it relies on external cues, such as external instructions. An Action is a transformation which needs to be performed explicitly and the steps composing it "cannot yet be imagined and none can be skipped." (Ibidem, p.19). For instance, a subject could start constructing the concept of numerical function by the substitution of specific values for the variable in an expression. The expression in this case acts as an external guidance for the subject to perform the transformation. As said by the authors, "although the most primitive of structures […], Actions are fundamental to APOS Theory. An Action conception is necessary for the development of other structures" (Ibidem, p.20). As we will see in the next paragraph, performing an Action could lead the subject to construct a Process.

4.2.2 Interiorization of Actions into Processes

An Action could be **interiorized into a Process**, **i.e., "a mental structure that performs the same operation as the action being interiorized, but wholly in the mind of the individual"** (Dubinsky et al., 2005a, p.339). The mechanism of interiorization is a crucial transition from a cognitive point of view. It is described as follow:

> As Actions are repeated and reflected on, the individual moves from relying on external cues to having internal control over them. This is characterized by an ability to imagine carrying out the steps without necessarily having to perform each one explicitly and by being able to skip steps, as well as reverse them. (Arnon et al., 2014, p. 20).

With reference to the above example of a numerical function, a Process stage is reached when the subject perceived a numerical function as a process which transforms numbers in numbers and can imagine this process without actually performing the transformation involved.

A successive step in the cognitive construction of a concept after the Process stage is the **encapsulation** of a Process into an Object.

4.2.3 Encapsulation of Processes into Objects

A subject could then apply an Action to a Processes. In order do so an individual needs to see "a dynamic structure (Process) as a static structure to which Actions can be applied." (Arnon et al., p. 21). This mechanism, called **encapsulation**, is described as it follows:

> If one becomes aware of **the process as a totality**, realizes that transformations can act on that totality and can actually construct such transformations (explicitly or in one's imagination), then we say the individual has encapsulated the process into a cognitive object. (Dubinsky et al., 2005a, p. 339, bold added by me).

Referring again to the example of the concept of numerical function, the encapsulation of the Process $y=f(x)$ into the object f is required when dealing with some mathematical properties of f (such as it being injective, surjective, monotone, etc.). Perceiving functions as objects is also required to form sets whose elements are functions or defining arithmetic operations between them.

When a Process has been encapsulated into an Object, it is possible for a subject, if necessary, to go back to the Process underlying the encapsulated Object. This mechanism is called **de-encapsulation**.

4.2.4 Coordination and Reversal of Processes

Interiorizing Actions is not the only way to construct new Processes. Already conceived Processes could be **coordinated** to form new ones: for example, an Object resulting by a Process P_1 could be then transformed by a different process P_2. When interiorized this leads to the construction of a new process $P_2(P_1)$.

Moreover, a Process could be **reversed**, giving rise to a new Process. For instance, as noted by Dubinsky (1991), the reversal of a function as a Process is cognitively linked to the construction of the concept of inverse function:

> It is by reflecting on the totality of a function's process that one makes sense of the notion of a function being onto. Reflection on the function's process and the reversal of that process seem to be involved in the idea of a function being one-to-one. (Dubinsky, 1991, p. 115)

4.2.5 Schemas and Thematization of a Schema into an Object

Finally, Actions, Processes, Objects and the just described interactions, when seen symoultaneusly form a **Schema**. More precisely:

> [A]n individual's Schema for a certain mathematical concept is the individual's collection of Actions, Processes, Objects, and other Schemas which are linked by some general principles or relations to form a framework in the individual's mind that may be brought to bear upon a problem situation involving that concept. (Arnon et al., 2014, p.110)

Moreover, a Schema is a dynamic structure characterized by its "continuous reconstruction as determined by the mathematical activity of the subject in specific mathematical situations." (Arnon et al., 2014, pp. 24-25). In Arnon et al., (2014, chapter 7) several examples of schemas are provided. I quote here one of them, the Function Schema:

> It can be composed of different types of functions such as real-valued functions, multivariable functions, vector-valued functions, and/or proposition-valued functions. These different types of functions may have been constructed as Processes or Objects, together with the operations that can be applied to them. For some students, different types of functions may be related by the common idea of a function Process: an operation applied to a set of inputs to obtain a set of outputs. Functions differ in the types of inputs involved, the nature of the operations applied to those inputs, and the results of the operations. Although individuals' Schemas may include the same types of functions, their components or the types relations constructed among them may differ. (p. 111).

Finally, a Schema could be conceived by a subject as a static structure (Object) to which new transformations can be applied. This mechanism is called **thematization**. Through thematization a Schema can be used in the construction of a new one or, more in general, through it Schemas can be linked together. Conversely, a Schema can be **de-thematized** by an individual to obtain the Actions, Processes, Objects, and other Schemas from which the original schema was constructed.

This concludes the description of the constructs of the APOS theory, divided in different categories of mental structures (Actions, Processes, Objects, and Schemas) and several different mechanisms between them (interiorization, (de-)encapsulation, coordination, reversal, (de-)thematization). In the following paragraph I will present how these constructs can be used to create a genetic decomposition of a mathematical concept. Before doing so, however, let us observe that although Actions, Processes, Objects, and Schemas have been presented linearly, highlighting a certain hierarchical progression A→P→O→S, this does not imply that APOS theory considers the cognitive development of an individual to progress linearly: "the development does not always proceed linearly, one stage after another. Rather, an individual may move back and forth between stages as the situation requires." (Arnon et al., 2014, p.9). In this sense, thus, the mechanisms of de-encapsulation (return from Object to Process) or of de-thematization (return to one or more structures that compose a Schema) should not be interpreted as regressions in the cognitive development, but rather they are considered as necessary to the construction of mathematical knowledge.

4.2.6 Genetic Decomposition (GD) of a mathematical concept

One of the main applications of APOS theory is that it allows to provide

> a model for describing how mathematical concepts can be learned, […] used to explain how individuals mentally construct their understandings of mathematical concepts" (Arnon et al. 2014, p.17).

Within APOS Theory this model is called a **Genetic Decomposition (GD)** of a mathematical concept. A GD is defined as "**a hypothetical model that describes the mental structures and mechanisms that a student might need to construct in order to learn a specific mathematical concept**." (Ibidem, p. 27, bold added by me).

As described by Arnon et al. (2014, pp. 27-28) a GD of a concept is generally constructed by researchers as a theoretical hypothesis based on several aspects:

I. The researchers' experience in the learning and teaching of the concept.
II. The researchers' knowledge of APOS Theory.
III. The researchers' mathematical knowledge.
IV. The literature review of the research in ME on the concept.
V. The historical development of the concept.

This hypothesis is then tested with empirical studies, modified and refined, and then tested again. This continuous cycle of modifications and experimental tests of the hypothesized GD is carried on until the researchers reckon to have reached a stable version for it. This refining process is generally lead by two questions: "(1) Did the students make the mental constructions called for by the genetic decomposition? (2) How well did the subject learn the mathematical content?" (Ibidem, p. 28). Different answers to these two questions can provide elements to the researchers to consider as valid or not a preliminary GD.

As Arnon et al. (2014) acknowledge, the GD of a concept can be used by researchers with two main roles:

a) A GD can be an interpretative tool for analysing and describing students' processes and difficulties:

> A genetic decomposition acts as a lens, analogous to a diffraction grating that researchers use to explain how students develop, or fail to develop, their understanding of mathematical concepts. […] [A] genetic decomposition is a tool by which researchers try to make sense of how students

> go about learning a concept and to explain the reasons behind student difficulties. (Arnon et al., p. 38).

b) A GD can be used as a model for the design of teaching interventions:

> In addition to being a theoretical model for research, the genetic decomposition for a concept guides instruction. Since a genetic decomposition describes the constructions a student may need to make in order to learn a mathematical concept, it can be used to design activities to help students to make the proposed constructions. (Ibidem, p. 47)

Later on in this chapter, I will present a GD of MI obtained as a modification and refinement of the GD of MI from previous studies in the literature. In this study, the proposed GD of MI will be used with the role (a) of above and thus to describe and analyse processes and to interpret (part of) student's difficulties related to MI. As it will be discussed in the conclusions of this thesis (section 13.3), a possible research direction that might stem from this study could focus on the design of teaching interventions framed on the presented GD of MI.

Two observations need to be made at this point. First, a GD for a concept should not considered as unique:

> As with any general and descriptive theoretical model, several genetic decompositions can be designed by different researchers or even by the same group of researchers to describe the learning of a particular concept. If those genetic decompositions are supported by empirical studies of students' constructions, they could all be considered reasonable descriptions of students' constructions. (Ibidem, pp. 40-41).

This aligns with the fact that "APOS Theory acknowledges that different students can follow paths different from those described in a particular genetic decomposition." (Ibidem, p.40). Thus, in conclusion, a GD does not describe the cognitive path that every student needs to take to construct a specific concept, instead:

> [A GD] can serve as a powerful descriptive and predictive tool to describe an individual's mathematical thinking. By detailing the structures involved in learning a particular concept, a genetic decomposition can help an instructor to uncover sources of difficulty that arise in the learning process. (Ibidem, p.18)

The second observation, related to the previous one, is that a GD can be continuously modified and refined as new theoretical elements or empirical data are found which were not previously considered. The modifications which might occur in a GD can be both the addiction of new structures or mechanisms not previously considered, and a focus on some aspects already mentioned in the original GD which leads to introduce explicitly new elements in the GD. An example of such a modification of a GD will be presented later in relation to MI. As we will see, in a recent paper, Garcìa-Martinez and Parraguez (2017), proposed a modification to the GD of MI made in Dubinsky (1991).

4.3 Mathematical Induction within APOS Theory

As already mentioned, there are several studies on MI within the framework of APOS Theory. In Dubinsky & Lewin (1986), although most of the theoretical structure of the APOS Theory was not developed yet, a GD of MI is presented. This model is then further discussed and deepened in Dubinsky (1991) with more specific references to the elements of APOS. This second version is the one that has been retaken and described in Arnon er al. (2014). Together with these works, Dubinsky has published two studies (1986, 1989) which report on a teaching intervention, framed within the APOS theory, in which MI was introduced to students with a series of computer programming activities. Let us start the presentation of these studies by analysing in detail the proposed GD of MI.

4.3.1 A GD of MI

Figure 4.2 contains an overview of the GD presented in Dubinsky & Lewin (1986), Dubinsky (1991), and Arnon et al. (2014), whose details are described in this paragraph.

The proposed GD starts from the presence of two schemas, considered preliminary for the construction of the PMI schema: the Function Schema and the Logic Schema. Respectively, it is necessary that the Function Schema includes the process of evaluating a function for a given value of its domain, while the Logic Schema must include the Processes of construction of statements in the first-order propositional calculus. The Function Schema must also allow the construction of the Process that transforms natural numbers into propositions obtaining a proposition-valued function of the positive integers. In other words, this process allows us to construct a predicate P(n), whose truth value depends on the natural number n. The Function and the Logic Schemas has then to combine (i.e., the above mentioned Processes need to be coordinated), so as to be able to construct the process that given a predication A(n) allows to construct a function A that associates, to each natural number n*, the truth value of A(n*). Having reached this point, therefore, proving the validity of a proposition of the form $\forall n \in \mathbb{N}$ P(n) corresponds to the Process of determining that the function P on the natural numbers has a constant value 'True'.

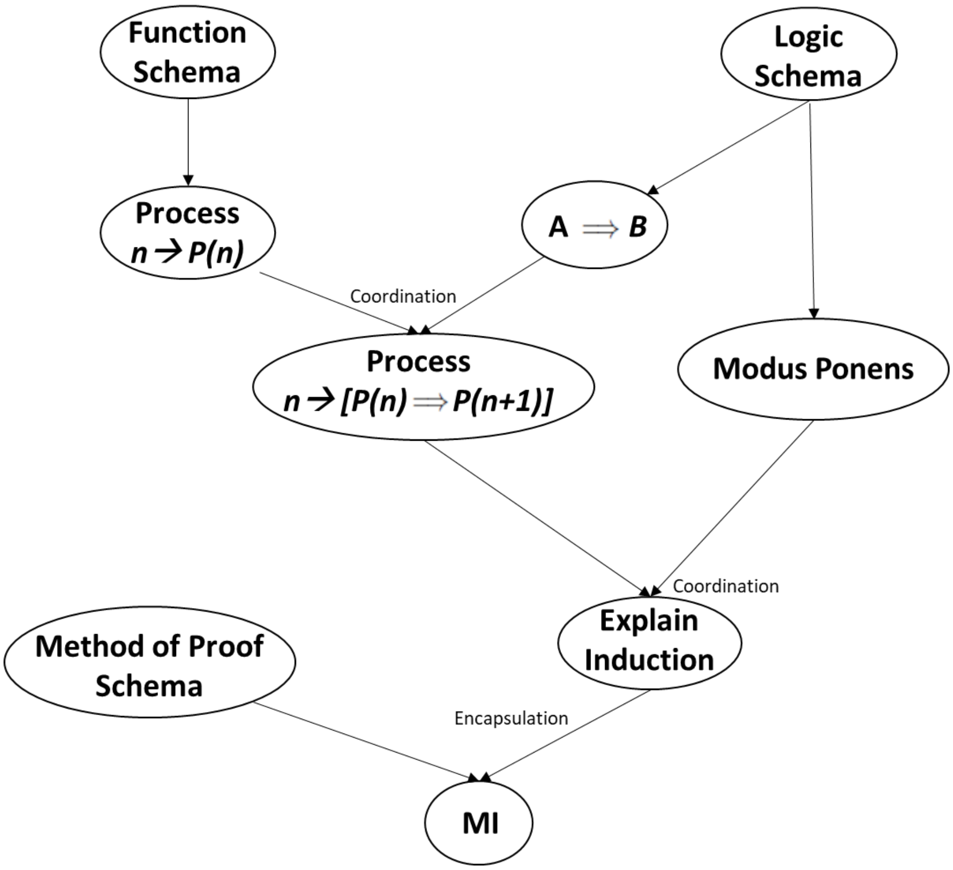


Figure 4.2. *GD of MI as presented in Dubinsky & Lewin (1986), Dubinsky (1991), and Arnon et al. (2014).*

As observed by Dubinsky (1991), "the essential point in an induction proof is that one does not prove the original statement directly, but rather the implication between two statements derived from it" (p.113), namely the inductive step. This point is explicitly described in the GD of MI as it follows. Once the predicate P(n) has been constructed, the next step consists in the process of constructing the predicate Q(n)=P(n)→P(n+1) that corresponds to the predicate involved in the inductive step. Dubinsky observes how this step is delicate and requires "a cognitive step" (Dubinsky 1991, p. 113), in fact the encapsulation of the implication is necessary in order to build Q(n). If this does not happen, it is possible that the subject considers $P(n) \rightarrow P(n+1)$ as two independent predicates.

At this point, again starting from the Logic Schema, a new Process intervenes, the Modus Ponens:

> This process is an interiorization of an action applied to implications [...]. The action consists of beginning at the hypothesis, determining that it is true, and then "crossing the bridge" to the conclusion and asserting its truth. (Ibidem, p.113).

In other words, the Process of Modus Ponens, given the implication $A \rightarrow B$, allows to transfer the truth of A to B.

At this point, a subject has to interiorize a new process called *Explain Induction* by which to construct a justification of the validity of MI or, in terms of what previously described, why once it has been establish that P(1) is true and that the predicate $Q(n)=P(n) \rightarrow P(n+1)$ is true for all the natural numbers, the predicate P(n) is necessarily true for all the natural numbers. Dubinsky (1991) describes the processes as in the following:

> Suppose that it has been established that Q has the constant value true. The first step in this new process which must be constructed is to evaluate P at 1 and to determine that P(1) is true (or, more generally, to find a value n_0 such that $P(n_0)$ is true). Next, the function Q is evaluated at 1 to obtain $P(1) \rightarrow P(2)$. Applying modus ponens and the fact just established) that P(1) is true yields the assertion P(2). The evaluation process is again applied to Q but this time with n=2 to obtain $P(2) \rightarrow P(3)$. Modus ponens again gives the assertion P(3). This is repeated ad infinitum, alternating the processes of modus ponens and evaluation. (p. 113).

From an APOS perspective, therefore, the construction of this iterative process which allows to transfer the truth value from $P(n_0)$ to P(n) for every $n>n_0$ is one of the fundamental points for the construction of the MI Schema.

Finally, the just described Explain Induction process can be encapsulated by a subject so as to construct the "Proof by Induction" object, which enriches the Schema Method of Proof, which can be used to demonstrate and solve problems.

Before continuing, two observations on the Explain Induction process are necessary.

First of all, according to the just presented GD, the interiorization of the Explain Induction process is cognitively crucial for the construction of the Schema of PMI. According to this model, in fact, **the 'Proof by MI' object can be built by a subject as encapsulation of the Explain Induction process**. This leads to a direct didactic implication that it may be effective to introduce proofs for MI with activities focused on building this process (for an example, see Dubinsky 1986, 1989, which will be briefly described above).

The second observation is that the Explain Induction process, as described in the GD in its generality, is the process by which a subject constructs a justification of the functioning of the MI both for a specific predicate and for a generic predicate P(n). In fact, **for a complete construction of the MI Schema, a subject must be able to construct this process even for a generic predicate P(n)**. This should not be underestimated from a cognitive point of view as it requires a student to conceptualise the predicate P(n) as an object even without knowing its specific content.

The just presented GD of MI was used by Dubinsky as a model for designing a teaching intervention focusing on MI. In this intervention, described in two works (Dubinsky 1986, 1989), MI was introduced to students through a series of computer activities with ISETL, a programming language based on classical set theory. During a series of lessons, students arrived to:

- Create an ISETLS function that, given a predicate *P(n)* as input (corresponding to an empty array with infinite dimension), returns as output the predicate *Q(n) = P(n)→P(n+1)*. This corresponds to the encapsulation of the implication described in the GD which allows to construct the predicate Q(n) from P(n).
- Create an ISETL function that, given a predicate *P(n)* as input determines, if it exists, the first natural number n_0 for which $P(n_0)$ is true.
- Finally, using the previous two functions, create a ISETLS program that, given a predicate *P(n)*, determines the first natural number n_0 for which $P(n_0)$ is true and then enters in a loop of applications of *modus ponens* through the predicate *Q(n)*. If the proposition $\forall n \geq n_0.P(n)$ is valid, the program would run forever. The functioning of this program corresponds to the Explain Induction process described above. Moreover, the fact that the program could be salved and used when necessary with other predicates corresponds to the encapsulation of this process into an object

Dubinsky observes that although the program cannot provide a formal proof for the statemen $\forall n \geq n_0.P(n)$, since it would need an infinite number of steps, nevertheless it supported students in constructing an intuitive justification for the validity of MI. Successively the students were involved in the construction of classic proofs by induction with a paper and pen. Finally, Dubinsky registers, considering the results of a conclusive test, that the teaching experience seems to have been effective for introducing MI to the students.

4.3.2 A modification of GD of MI from the literature

In a more recent study, Garcìa-Martinez and Parraguez (2017) proposed a modification to the GD presented above. The starting point of the two researchers is the observation that, coherently with the literature, often the base of the induction is considered as not necessary by students for the correctness of a proof by MI. In the GD of above, the fact that $P(n_0)$ is valid (i.e., the base case) is a premise of the Explain Induction process, however the two researchers consider that it is necessary to give a greater centrality to this point in the GD of MI. For this reason, they explicitly introduce a new process in the GD, the *Base of Induction Process,* corresponding to the determination of the initial number n_0 from which the explain induction process can start. Following this, it becomes cognitively relevant for the construction of the MI schema that the base of the induction does not simply correspond to verify the truth of $P(n_0)$ but may also involve the process of determining the value n_0. Garcìa-Martinez and Parraguez finally reports on an empirical study which highlights how the introduction of this aspect in the GD of MI from one side offers new elements for analysing students' difficulties related to MI and, on the other side, suggests the importance of teaching activities in which students themselves have to determine the value n_0, that later will correspond to the base of the induction, before constructing a proof by MI.

With the modification proposed by Garcia-Martinez and Parraguez (2017), the diagram representing the GD of MI becomes the one reported in figure 4.3.

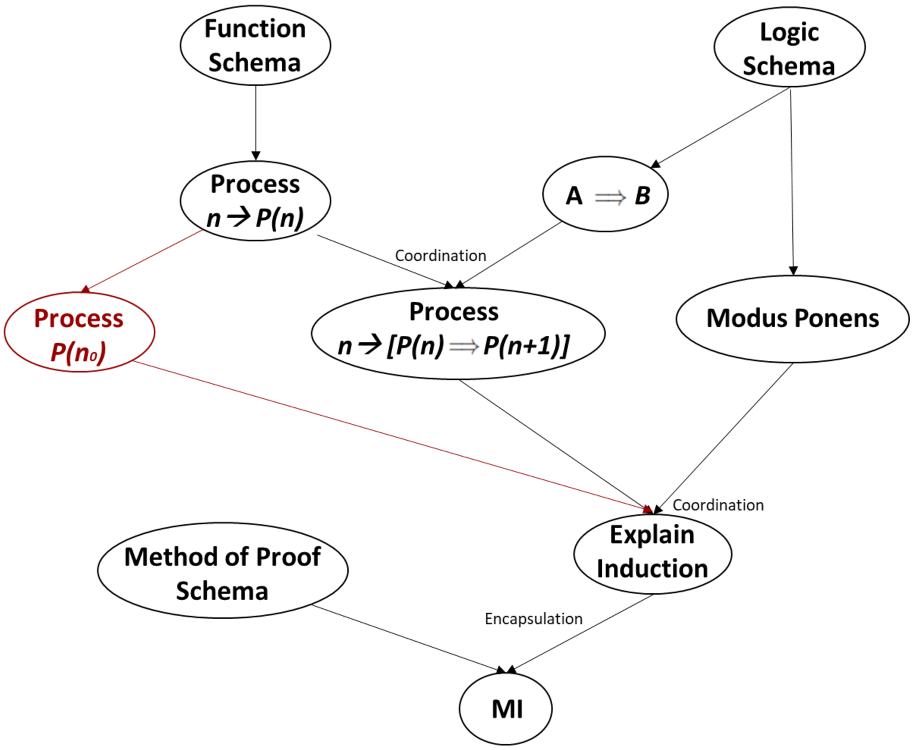


Figure 4.3. *GD of MI including the Base Case Process (in red), as proposed in Garcia-Martinez & Parraguez (2017).*

4.4 APOS Theory within this study

As we will see the just described model will provide elements useful to design research activities and to analyse some of the collected data. Preliminarily to this, however, the GD of above can be further deepened by considering the historical-epistemological analysis presented in chapter 2. As notices above, in fact, the APOS theory explicitly recognises that a GD of a concept can be constructed (or modified) by including elements emerging from an analysis of the historical development of the concept.

4.4.1 A contribution from the historical-epistemological analysis

In the conclusions of the Chapter 2, I have presented a series of characteristics highlighted by the analysis of the historical traces of MI. In particular these were summarised in the following points:

(A1) Natural numbers as aggregations or as progressions

(A2) Potential infinity or actual infinity

(A3) Iteration

(B1) Absence or presence of the parameter

(B2) Generic example as a general case

(B3) Infinite proof or finite proof

(B4) Inductive step as an independent proposition

(B5) Increasing or decreasing iteration

(C1) Statement with an implicit or explicit recursion

(C2) Not necessity for trivial cases

Excluding the last two points, which relate to the kind of propositions involved, let us observe that the other points can be re-read in relation to the APOS analysis of MI. Two points have already been taken into consideration in the GD of MI presented above.

1) The iteration (A3) as extremely related to MI is explicitly involved in the Explain induction process. In particular, starting from base case and inductive step, trough successive applications of MP it is possible to *iteratively* conclude that P(n) is true for each of the natural numbers.
2) The inductive step as an independent proposition (B4) seems to find its counterparts within the GD in the encapsulation of the proposition P(n)→P(n+1) which has to be conceptualised by a subject as a unique proposition Q(n), whose truth value corresponds to the validity of the implication. As described in the GD of MI the Explain induction process is constructed starting from the fact that Q(n) has a constant value 'true' on the set of natural numbers.

The remaining other points can provide new elements with which the presented GD of MI can be deepened and expanded. Let us start by summarising what these points referred to:

1) One of the first points highlighted by the analysis was the distinction between natural numbers as aggregations or as progressions (A1). It was observed how in the Euclidean perspective a natural number is an arbitrary aggregation of units, while in Maurolico the linear numbers (which corresponds to the natural numbers) are recursively constructed by adding a unit to the previous linear number. This aspect highlights a crucial characteristic of MI, which is the fact that it is strongly related to the structural role that the operation successor has in the set of natural numbers.
2) A second emerged aspect was the role of the generic example (B2) with which, in the analysed traces, the generality of a statement was proved. As observed this aspect is related to the absence of a parameter *n* (B1), proper of the modern algebraic symbolism, which allows to construct a proof for *a generic n*, and, more formally, to base a proof on the logical inference law of the Universal Generalisation.
3) The historical analysis allowed to highlight how different scholars registered a trace of MI in proofs involving increasing or decreasing iterations (B5). It was possible to observe, for instance, how Fermat's method of infinite descent or some Euclidean proofs based on equivalent formulation of the well-ordering principle had a structure which, although different, recalls the modern MI. It is supposed by contradiction that there exists a number n^* for which the proposition is false and, starting from this, a decreasing sequence of natural numbers for which the proposition is false is constructed. This sequence is infinite, and thus a contradiction with the infinite descent principle is reached, or the sequence reaches a natural number n_0 for which the proposition was known to be true, and thus a contradiction is again reached.
4) Finally, it was possible to register a tension between potential and actual infinity (A2) and, related to this point, the presence of proofs in which the authors, or the commentators, referred to a finite or an infinite number of steps (B3) composing the proof.

In the next paragraphs, each of the previous four points will be reinterpreted in APOS terms and in relation to the GD of MI presented above.

4.4.2 The Successor Schema

From a mathematical point of view, the successor operation has a structural role in the modern formalization of natural numbers. Let us consider, for example, Peano's famous axiomatization. The operation successor has a structural role also in relation to MI. In particular, the induction step corresponds to prove that the set of numbers n for which P(n) is true is closed with respect to the successor operation. In the GD of above, it was observed that with the Explain Induction process a subject can construct a justification of the functioning of MI as chain of syllogisms which, starting from the base of the induction n_0, reaches every natural number greater than n_0. Let us observe, however, that in this passage it is necessary that a subject has interiorized the fact that every natural number *n* can be reached by progressively adding 1 to 0 (or, more generally that any natural number $n>n_0$ by progressively adding 1 to n_0). Despite this aspect might look very intuitive, a deeper analysis in APOS terms of this point could provide further elements for the GD of MI.

Cognitively, following the APOS perspective, the construction of the *Successor Schema* seams to require some steps.

First of all, it is necessary for a subject to recognise that the action of 'adding 1' performed on a natural number gives as result its consecutive natural number. This aspect, which from a mathematical point of view trivially relies on the definition of consecutive number, from a cognitive point of view requires the awareness that during the enumeration the passage from one number to the next one corresponds to summing 1 to the previous one.

The action of 'adding 1', thus, must be interiorized in the process which transforms natural numbers, and which corresponds exactly to the process by which a natural number is transformed in its successor. This, in a compact form, is the process which, given a natural number *n,* allows to obtain the successor of *n* as $S(n)=n+1$.

Note that this expression for $S(n)$ is valid in N, where the successor of an element *n* is univocally identified by $n+1$. The successor Schema can be thus subsequently accommodated to include the successor operation also in other numerical sets, for example in the set 2N of even numbers S(n) is obtained as S(n)=n+2. A subject's successor schema could also allow them to realise that there are numerical sets where it is not possible to define a notion of successor of an element in a compatible way with the order of the sets, in the sense that it does not exist a third element between an element and its successor (such as in Q or R). This aspect is important in relation to MI, because it allows the subject to understand why MI is not a particularly well fitted proving schema in all numerical sets.

In relation to MI, moreover, the successor schema needs to be enriched by a subject's interiorization of some fundamental mathematical properties:

(a) The Process of continuously adding +1 from 0 is potentially infinite.
(b) The Process of continuously adding +1 from 0 does not contain loops.
(c) Any natural number can be reached by the process of continuously adding +1 departing from 0, in a finite number of step (i.e. $n=S^n(0)$).

Let us observe that, from a mathematical point of view, these properties are related to Peano's axioms, which can be formulated as:

1) There exists a number $0\in \mathbb{N}$.
2) There exists a function $S: \mathbb{N} \longrightarrow \mathbb{N}$.
3) The function S is injective.
4) For every $n\in \mathbb{N}$, $S(n) \neq 0$
5) If $U\subseteq \mathbb{N}$ such as $0\in U$ and $S(u)\in U$ for every $u\in U$, then $U= \mathbb{N}$

At this point, thus:

- Property (a) corresponds to the fact that every natural number has a successor, or that there exists a successor function from N to N. From this it follows that the successor function can be iteratively applied without never reaching a number for which it is not defined

- Property (b) is a direct consequence of axioms 3) and 4).

- Property (c) is related to the axiom 5) which, in other terms, affirms that the natural numbers are *exactly* those numbers obtained by iteratively applying the successor function starting from 0.

As anticipated above, the successor Schema seems to be central in relation to MI Schema. The GD of MI introduced above establishes that a subject might construct the Explain Induction Process by the coordination of the Induction Base Process, of the Modus Ponens Process and of the Evaluation Process applied to the function Q(n)=P(n)→P(n+1). Mathematically, this means to construct the following logical inferences: from P(1) and P(1)→P(2) it follows P(2); from P(2) and P(2) →P(3) it follows P(3), and so on. It is important to notice, however, that **to perceived that this process can reach any natural number *n*, it is necessary for a subject to have conceptualized that the Successor Process has the above-mentioned (a), (b), (c) properties.**

The construction of such a rich Successor schema should not be regarded as immediate from a cognitive point of view. Relaford-Doyle and Núñez (2018) highlighted that cognitively some steps are required to interiorize that:

> I. The natural numbers are unbounded, since one can always be added to produce the next number.
> II. Adding one is the only way to generate the next natural number.
> III. The same relation holds between any natural number and its successor (e.g., the pairs 3&4, 8&9; 9,381,763&9,381,764; and 10^{23}&10^{23} + 1 are all governed by the same +1 relation).
> IV. No natural number is more 'natural' than any other. (p.236).

In their empirical study the authors register that "Concepts of natural number—even among college-educated adults—are not best characterized by the Peano's axioms. Educated adults demonstrated conceptualizations that were at odds with the formal characterization [of N]" (Ibidem, p.249). The authors observe that if a subject's experience with natural numbers is mainly related to the process of counting it is possible that the points I-IV of above are not correctly conceptualized by them:

> In our everyday experience with counting numbers, large numbers *are* different from smaller numbers: they are encountered less frequently, are lexically and notationally more complex, and are harder to manipulate in calculations. These issues are all irrelevant in the domain of *natural number*, where the successor function generates a system in which all natural numbers, regardless of magnitude, are governed by the same logic (including deep implications such as I–IV). […] In our view, the assumption that people come to formally consistent understandings of natural number by virtue of their experience with familiar counting numbers is analogous to the claim that people come to scientifically valid understandings of the animal kingdom by virtue of their experience with their pet cats and dogs. (Ibidem, p. 249-250, italic in original).

In conclusion, both for the importance of the operation Successor in relation to MI emerging from the historical analysis, and the empirical evidence of its potentially problematic construction from a cognitive point of view, my point of view is that the Successor Schema should be explicitly included in the GD of MI. **The Explain Induction Process, in conclusion, can be modelized as a coordination of the already considered Processes together with the *Process n=S^n(0)*, related to the Successor Schema.**

Figure 4.4 represents the proposed modification to the GD of MI related to this point.

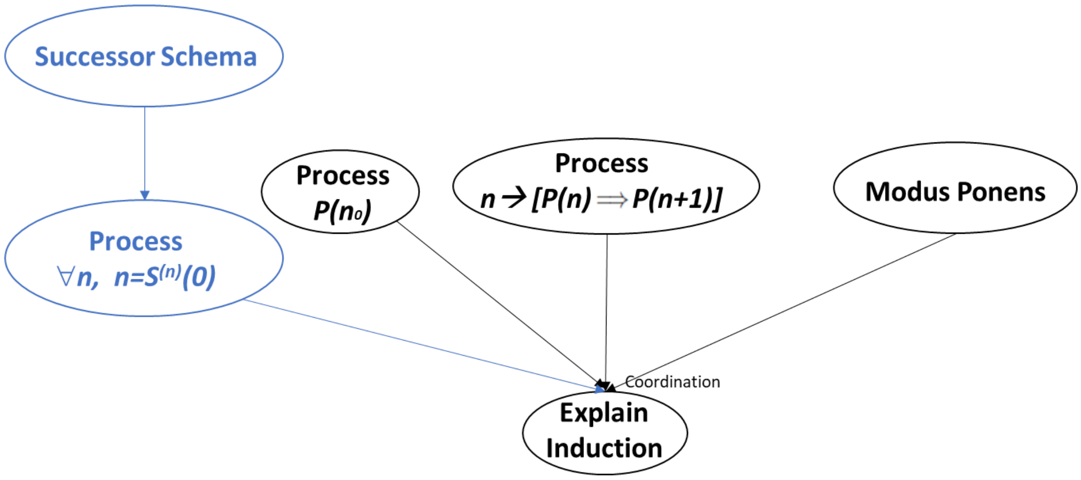


Figure 4.4. *The Successor Process (in blue) within the previous GD of MI.*

4.4.3 The Universal Generalization Process

From the historical-epistemological analysis it was possible to observe how, when the propositions involved a general statement, the proof was instead conducted for a generic example. For instance, in Pascal's proposition, considered by many as a clear historical example of a proof by MI (*Consequence XII,* analysed in 2.7), the inductive step although expressed in general terms is proved only for two specific bases of the arithmetic triangle, the fourth and the fifth one. In other terms, a proof for statement $\forall n[P(n) \rightarrow P(n+1)]$ is constructed by Pascal's, firstly by showing that $P(4) \rightarrow P(5)$ and then concluding that the same process could be used for all the other cases: "The same will be proved in all the rest [of the triangle]" (Pascal, 1665, section 1, p.8). This aspect, emerging from the historical analysis, highlights the connection between the role of the generic example to prove the generality of a statement and the inference law of the Universal Generalization (UG), generally used in a modern proof of a statement containing a universal quantification.[69] As written by Mason and Pimm (1984), in fact, "A generic example is an actual example, but one presented in such a way as to bring out its intended role as the carrier of the general"(p.287). Similarly in a proof where UG is used, a statement with the form $\forall x \in A.P(x)$ is proved by showing the validity of $P(x)$, where the variable x represents a generic fixed element of the set A (similar to what happens with a generic example).

This connection between generic example and UG can be analysed from a cognitive point of view in APOS terms. In particular, the Action of verifying the truth of a proposition for multiple cases can be interiorized by a subject into a Process when s/he recognises the common characteristics of these Actions. Once this has happened, it is not cognitively necessary to verify the proposition for more cases. Instead, it is sufficient to consider only one, the generic example, which becomes representative for all the others, and thus generalizable. Following this, therefore, a generic example can be seen by a subject as representative for the Process of 'verifying for any case'. In the above example of Pascal, for instance, he proves all the infinite cases corresponding to the inductive step by showing just a single step ($P(4) \rightarrow P(5)$) which represents all the cases.

The following step consists in using a parameter instead of a specific numerical example. This is what happens in a proof by MI when the inductive step is proved by showing that $P(n) \rightarrow P(n+1)$ for a generic natural number n. At this point, formally, UG can be applied, and one can conclude that $\forall n.[P(n) \rightarrow P(n+1)]$.

[69] I have already discussed the inference law of UG and its specific role in a proof by MI in 3.3.2.

From a cognitive point of view, the introduction of the parameter requires for a subject to recognise that the variable *n* can be considered as representant of a natural number which is, at the same time, generic and specific. It is generic since it is not any particular number, but it is specific because it is *a fixed* natural number. As described by Mason and Pimm (1984) and already analysed in section 3.3.3, this point is delicate from a cognitive point of view and in relation to that students' difficulties may emerge.

To summarise, thus, the conceptualization of the logical law of UG for a subject can be described in APOS terms as a **process obtained by the coordination of two further processes: a) the process of proving a universal proposition by a generic example and, b) the Process by which a letter becomes a parameter to represent a generic number**. By coordinating these two processes, thus, a subject can interiorize that when a proposition is proved for a generic $x \in A$, then the proposition is proved $\forall x \in A$. **I will call this process the Universal Generalization (UG) Process.**

As observed, in a proof by MI, the UG is not directly applied to the statement $\forall n.P(n)$ but to the induction step $\forall n.\ [P(n) \rightarrow P(n+1)$. This is a delicate stage**: firstly, for the subject it is necessary to recognize that UG is used, and thus what it is actually proved is $\forall n.\ [P(n) \rightarrow P(n+1)]$; secondly s/he has to consider that UG is applied to the implication and not to the proposition P(n)**, otherwise there is the risk for the subject to interpret it as a proof of $[\forall n.P(n)] \rightarrow P(n+1)$, with obvious problems of circularity.

In conclusion, the GD of MI presented above can be enriched by including the UG process in relation to the process of constructing the inductive step. As observed, the Explain Induction process was constructed as a coordination between base and inductive step. The just presented analysis, however, highlighted that to construct the latter it is important for a subject to have interiorized the UG process. This point is represented in figure 4.5, where the UG process is represented within the other elements of the GD of MI.

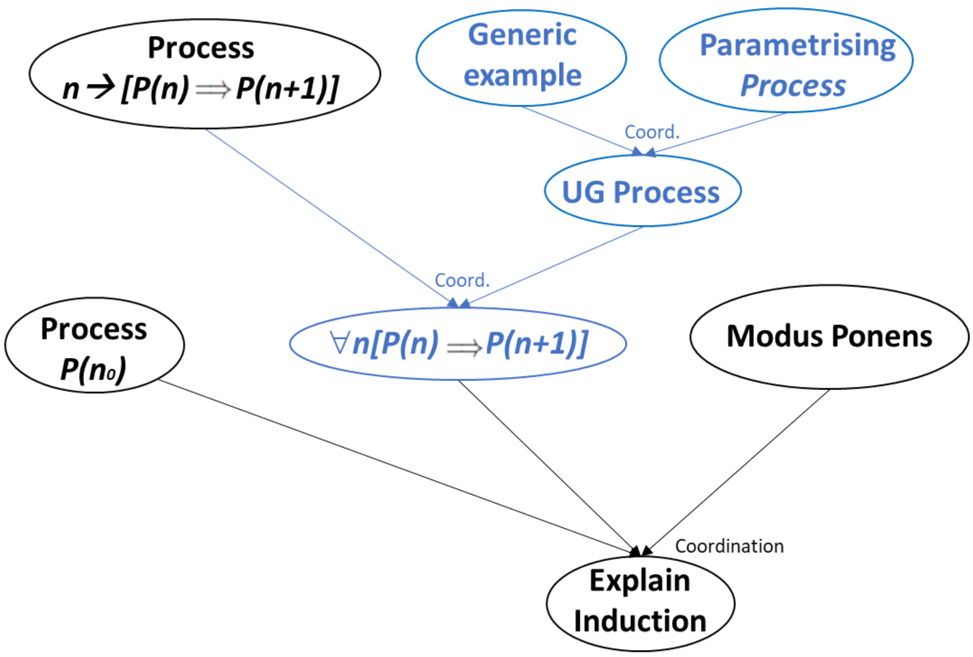


Figure 4.5. *The Universal Generalization Process (in blue) within the previous GD of MI.*

4.4.4 The Explain Induction Process – Direct and Indirect form

As described the Explain Induction process allows the subject to construct an (informal) justification of the validity of MI. In particular, starting from the validity of $P(n_0)$ and of $\forall n.\ [P(n) \rightarrow P(n+1)]$, a series of modus ponens (MP) are constructed to conclude the validity of P(n) for every $n \geq n_0$. The historical-epistemological analysis highlighted that something similar happened also in those indirect proofs involving recursive argumentations. For instance, in analysis of Fermat's letter to P. de Carcavi, I observed (see. 2.8) how the mathematician, to prove that every prime p with the form *4n+1* can be written as the sum of two square numbers, iteratively applies a sort of contrapositive of the inductive step to prove that if a generic prime *p** does not have the given property then neither 5 has it, which is absurd.

This point suggests a depending on the Explain induction process in the GD of above. In particular, the explain induction process could be formed in a different way. For instance, a similar process could be constructed to justify why, if the inductive base and the inductive step are valid, it cannot exist a natural number $n^*>n_0$ for which $P(n^*)$ does not hold. In fact, if we suppose that $P(n^*)$ does not hold, then necessarily $P(n^*-1)$ does not hold either, because otherwise, by Modus Ponens, $P(n^*)$ would hold. Thus, $P(n^*-2)$ does not hold either, because otherwise $P(n^*-1)$ would hold. By iterating this, we obtain that $P(n_0)$ does not hold, which however contradicts the induction base. Therefore, $P(n^*)$ must hold. This argument, from a logical point of view, is constructed by an iteration of the conditional inference *Modus Tollens* (MT): from $A \rightarrow B$ and $\neg B$, it follows $\neg A$. In particular, in this case MT is applied to the implication $P(n) \rightarrow P(n+1)$ and $\neg P(n+1)$, obtaining as a result $\neg P(n)$. This is, with similar terms, what happened in Fermat's proof of above.

We can interpret in APOS term the construction of this process by a subject. First of all, it is necessary that the subject has interiorized the process corresponding to the Modus Tollens. This process, that I will call the MT process, can be constructed by a subject as the coordination, within the Logic Schema, of the MP process, with the negation Process, i.e. the process by which a subject can transform a proposition Q in its negation $\neg Q$. In particular, given an implication $A \rightarrow B$ and a proposition C, a subject, by the negation process, can recognise that $C=\neg B$. At this point, the subject needs to recognise that from $A \rightarrow B$ and $\neg B$, it necessarily follows that $\neg A$, since otherwise, by modus ponens, B would be valid. Once the MT process has interiorised it becomes possible for a subject to construct a series of syllogisms which allows to conclude that if there exists a number $n^*>n_0$ for which $P(n^*)$ does not hold, then $P(n_0)$ does not hold either.

In conclusion, thus, together with the Explain induction described by Dubinsky, it is possible to include in the GD of MI also the analogous process but involving the MT instead of the MP. To distinguish these two analogous (but different) processes, **I will call them, respectively, the Explain induction process in direct form (the one involving the MP) and the Explain induction in indirect form (the one involving the MT).** The figure 4.6 of above represents the proposed modification to the GD of MI relatively to this point.

Let us observe that there is another difference between the two forms of the explain induction process. As described, the direct one necessitates the interiorization of the process '$\forall n.\ n=S^n(0)$' through which a subject recognises that any natural number can be reached by iteratively applying the successor operation starting from 0. In the explain induction in indirect form this process must be transformed in the *reversed* one. Indeed, in this second case, one need to recognise that by iteratively applying the *predecessor operation* ($S^{-1}(n)= n-1$) starting from *n** it is possible to reach any $n<n^*$, and in particular 0.

In the GD presented above what said in relation to the explain induction process in indirect form was not taken into consideration. My point of view, however, is that a sufficiently rich Schema of MI should include, together with the Explain induction process in direct form, also the one in indirect form. In particular, when this latter process is interiorized, the MI schema can be used by a subject to justify why, by the validity of the inductive step and by the fact that $P(n^*)$ is false, it is possible to conclude that $\forall n \leq n^*$ $P(n)$ is false.

In conclusion thus, **when the Schema of MI contains both the Explain Induction Process in direct form and the Explain Induction Process in indirect form, a subject can understand that given the validity of the inductive step, $\forall n.\ P(n) \rightarrow P(n+1)$**

- **if it exists a number n_0 such as $P(n_0)$ is true, then $P(n)$ is true also for every $n \geq n_0$,**
- **if exists a number n_1 such as $P(n_1)$ is false, then $P(n)$ is false also for every $n \leq n_1$.**

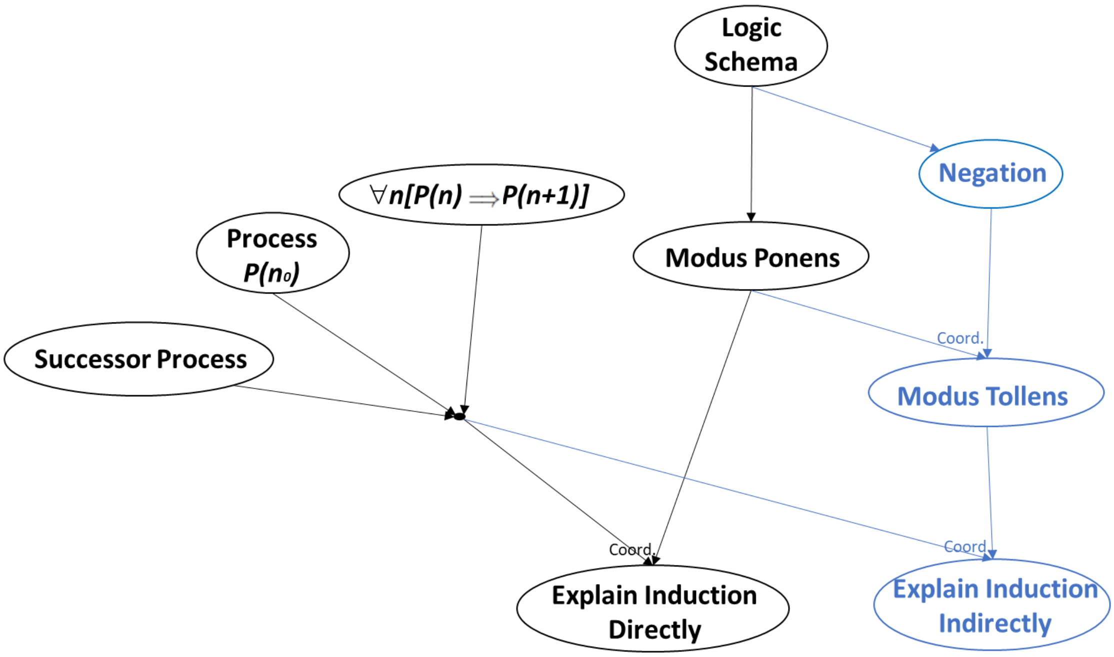


Figure 4.6. *The Modus Tollens and the Explain Induction Processes in indirect form (in blue) within the previous GD of MI.*

4.4.5 Potential and Actual Infinity in the encapsulation of the Explain Induction Process

A further aspect highlighted by the historical-epistemological analysis is the dichotomy between potential and actual infinity. This element, although just hinted in the conducted analysis, allowed to register an epistemological distance between the Euclidean propositions, in which the statement is proved for "as many numbers as we please" and the most recent ones, such as Pascal's consequence XII, were the proof involve "an infinity of cases". Moreover, it was observed that this distinction brings some consequences on the structure of the proofs themselves. When, as in Euclid, the statement involves a finite number of cases, even if arbitrary, the recursive argumentation which structures the proof has a finite number of steps; when the statement involves an infinite number of cases, instead, an infinite number of steps are required, as in Maurolico, or a different strategy which allows to prove an infinite number of cases with a finite proof, as in Pascal ("Despite this proposition has an infinity of cases, I will provide a shorter proof thanks to the use of two lemmas", Pascal, 1665, section 1, p.8).

The cognitive (and mathematical) distance between *potential* and *actual infinity* has been analysed within the APOS theory Dubinsky et al. (2005a, 2005b). As the authors write:

> APOS Theory can help us to understand the distinction between the potential and the actual [infinity] [...]. **Potential infinity is the conception of the infinite as a process**. This process is constructed by beginning with the first few steps (e.g., 1, 2, 3 in constructing the set N of natural numbers), which is an action conception. Repeating these steps (e.g., by adding 1 repeatedly) ad infinitum requires the interiorization of that action to a process. **Actual infinity is the mental object obtained through encapsulation of that process**. [...] Thus we would say that, through

encapsulation, the infinite becomes cognitively attainable. (Dubinsky et al, 2005a, p. 346, bold added by me)

As observed by the authors, thus, the construction of an actual conception of infinity can be cognitively attained through the encapsulation of an infinite process which, despite been considered as an endless process, is conceptualized in its totality. Starting from this analysis, the authors present a classic example: the equality $0.\overline{9} = 1$, for which several studies highlighted the student's difficulties in accepting its validity (Tall 1977; Tall & Schwarzemnberger, 1978; Cornu, 1991). Following the conducted analysis in APOS terms, the authors propose two different interpretations for the phenomenon. Firstly, it is possible that "[a] student may actually conceive of .999. . . as consisting of a string of 9s that is finite but of indeterminate length" (Dubinsky et al. 2005b, p. 262). In this case, since the process with which $0.\overline{9}$ is mentally constructed is not perceive as infinite, the resulting number cannot be equal to 1: there will always be a distance between the two numbers. Secondly, the authors observe that even when $0.\overline{9}$ is perceived as an infinite process by a subject, an encapsulation of this process into an object is necessary, because otherwise it cannot be compared with the object 1:

> An individual who is limited to a process conception of .999. . . may see correctly that 1 is not directly produced by the process, but without having encapsulated the process, a conception of the "value" of the infinite decimal is meaningless. (Ibidem, p. 261).

In other terms, if $0.\overline{9}$ is perceived by a subject as a process, it is possible that s/he considers that it only approximates 1, approaching it without never actually reaching it.

The just conducted analysis seems to suggest that, in the transition from infinite process to object, two aspects are to be considered: first of all, the process must be perceived by the subject as an infinite process and then an encapsulation must take place, by which the totality of this process becomes an object.

The authors conclude that the relationship between potential and actual infinity can be interpreted within APOS theory in the following terms:

> **Once the individual can see all of the steps of an infinite process as a single operation that can be carried out and finished, and present at a moment in time, he or she can conceive of the infinite as a completed totality. Once this totality is encapsulated, the notion of potentiality is transformed into an instance of actual infinity, a mathematical entity to which actions can be applied.** Hence, the existence of the one does not negate that of the other, nor is either a misconception with respect to the other. Instead, the potential and actual represent two different cognitive conceptions that are related by the mental mechanism of encapsulation. These conceptions and their relationship become part of the individual's infinity schema. (Dubinsky et al., 2005a, p. 346, bold added by me)

The just described transition between potential and actual infinity plays an important role in the MI schema. As described, through the Explain Induction Process, a subject might construct a chain of deductions which potentially reaches any natural number greater than the base of the induction. The next cognitive step is to perceive this as an infinite process which reaches all the natural numbers. **Finally, this infinite process needs to be encapsulated so that the infinite chain of deductions could be considered in its totality, as a unique deduction which starts from n_0 and covers the whole set of natural numbers $\mathbb{N}$.** If this encapsulation is not accomplished, and the subject is still considering this chain of deductions as a process, it is possible that s/he perceives a proof by MI as infinite and never ending. In this case, thus, s/he could think that to conclude, by MI, that P(n) is true on the whole and infinite set $\mathbb{N}$, it would need an infinite amount of time.

Let us observe that the transition from the infinite chain of modus ponens and MI as a finite proof is delicate not only from a cognitive point of view, but also from a logical-mathematical one. In my analysis of a mathematical theorem by MI as a triplet (Statement, Proof, Theory), see 3.3.2, I have observed that, mathematically, to conclude $S=\forall n(P(n))$ from the validity of the base S1 and of the inductive step S2, it is necessary to use the meta-theorem $S1 \wedge S2 \rightarrow S$, corresponding to an instance of the Principle of MI. The meta-theorem $S1 \wedge S2 \rightarrow S$, thus, allows to conclude a proof by MI without constructing the infinite chain of deduction which, otherwise, would be necessary to prove the statement for every natural number. In other terms, the Principle of MI allows the transition from the infinite number of statements:

(a) "Given any number $n \in \mathbb{N}$, $n \geq n_0$, 'the proposition P(n) is true'" – in fact, from S_1 and S_2, it is possible to build a finite chain of deductions which brings to the truth of P(n), starting from the truth of $P(n_0)$,

to the unique statement:

(b) "The proposition '$\forall n \in \mathbb{N}$, $n \geq n_0$, P(n)' is true".

In the case of (a), by varying n, we have an infinite number of different proofs, every time build with a different chain of deductions departing from n_0. In the case of (b), instead, by using the Principle of MI, with a unique (and finite) proof we can conclude that '$\forall n \in \mathbb{N}$, $n \geq n_0$, P(n)' is true.

To summarise, thus, the above described transition, which from a mathematical point of view is justified by the principle of MI, should be interiorized from a cognitive point of view as well. In particular, it is necessary that in a subject's Schema of MI, a proof by Mi is conceptualized as a proof that allows to prove a statement per the whole (infinite) set of natural numbers $\mathbb{N}$ but being, at the same time, finite. With reference to the GD of MI, the Explain Induction from an infinite process needs to be encapsulated into an object.

The theme of the dichotomy between potential and actual infinity has been also analysed with reference to MI by Fischbein (1987 pp. 51-52). This analysis is framed in a different theoretical perspective, the one of the *Intuitions,* presented in the following section where this point will be discussed in detail (see 5.2).

Following what just said, the GD of MI of above can be extended in relation to the encapsulation of the Explain induction process, by including the transition from potential and actual infinity part of the individual Infinity Schema (as described in Dubinsky et al., 2005a; 2005b). This transition, as observed, is both important from a logical-mathematical point of view, and delicate from a cognitive point of view within APOS theory. For this reason, it is important and useful that the GD of MI includes an explicit reference to this point. The figure 4.7 represents the proposed modification to the GD of MI in relation to the transition between potential and actual infinity.

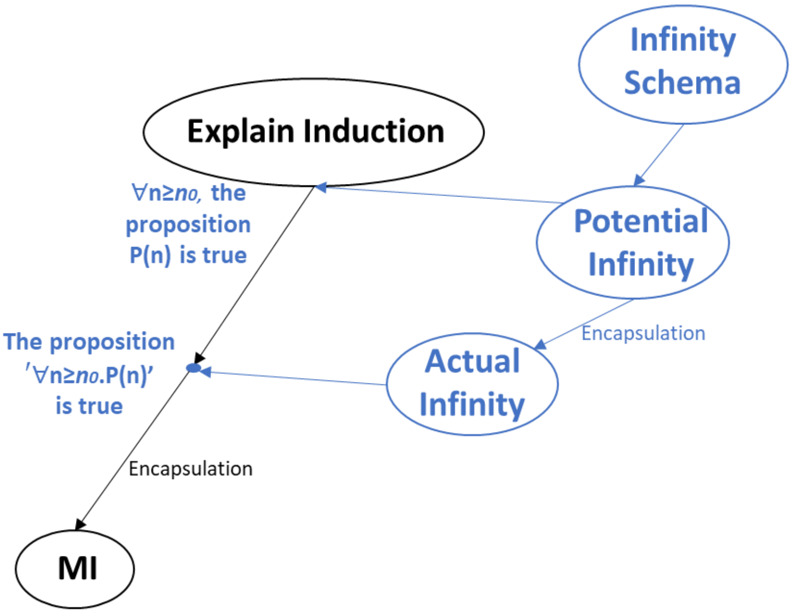


Figure 4.7. *The encapsulation to actual infinity within the previous GD of MI.*

In conclusion, the historical-epistemological analysis allowed me to deepen and expand some points of the GD of MI presented in the literature. In particular, four new points have been included in the GD: the successor process in the set of natural numbers ℕ as preliminary to the Explain induction process, the process corresponding to the logical law UG, the Explain induction process in indirect form, and the transition from potential and actual infinity in the encapsulation of the Explain induction process. The figure 4.8 summarises the whole GD of MI with these just described modifications.

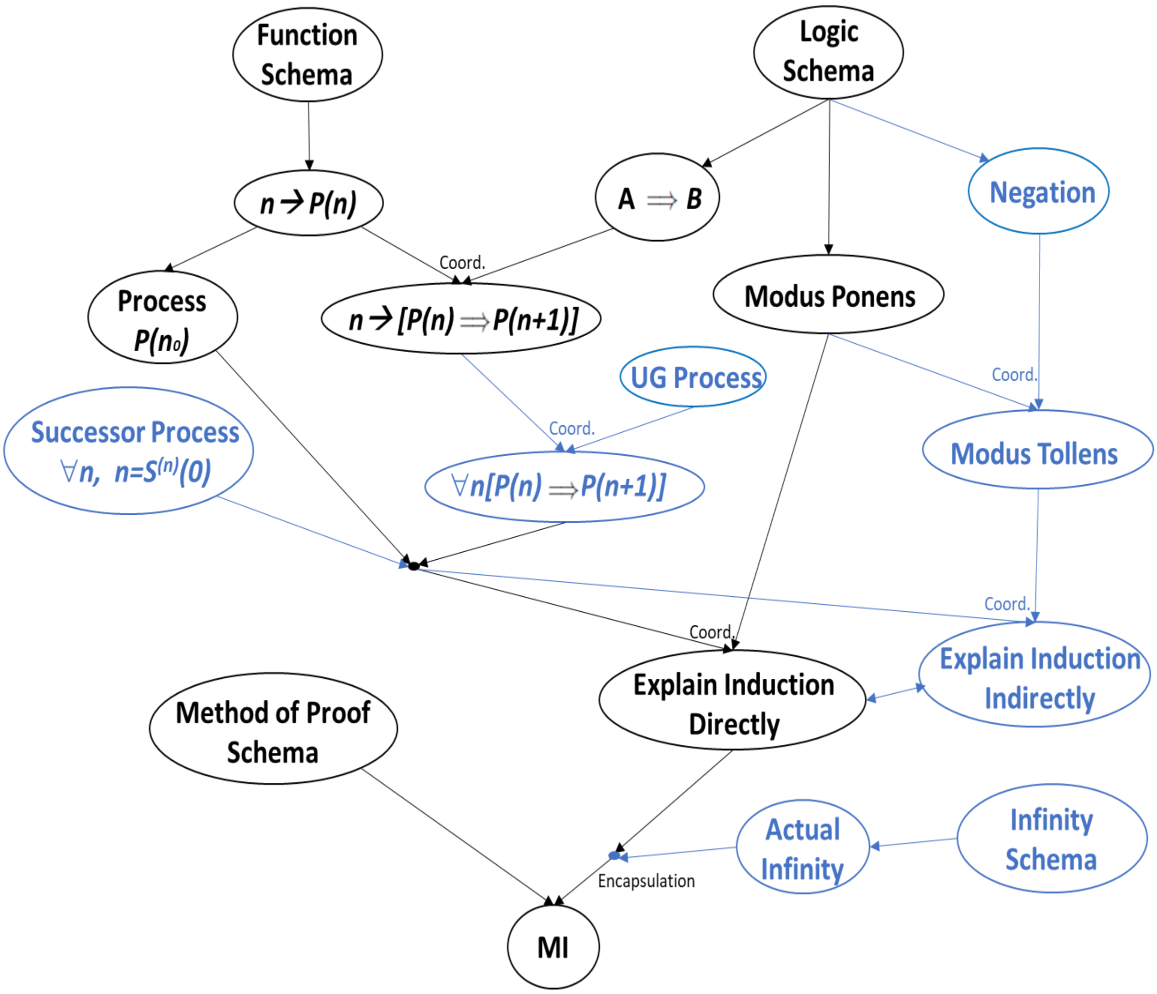


Figure 4.8. *GD of MI in which the just presented aspects (in blue) are integrated in the previous GD of MI.*

5 INTUITIONS AND INTUITIVE ACCEPTANCE

5.1 PREMISES AND INTRODUCTION

A further aspect that is taken into account in this thesis is the *intuitive acceptance of MI as a proving scheme*, that is up to which point such proving scheme is considered intuitively convincing. The theoretical reference is contained in the construct of the *intuitions* by Fischbein (1987) that I will present in this chapter. According to Fischbein (1982):

> There are frequent situations in mathematics in which a formal conviction, derived from a formally certain proof, is NOT associated with the subtle feeling of "It must be so", "I feel it must be so". (p.11, capital in original).

In other words, the proof of a theorem alone could not be enough for someone to get a feeling of evidence and certainty of the proved statement. In relation to this, Fischbein (1987) shows a famous example from the history of mathematics: the proof by Cantor of the fact that, using modern language, there is a bijection from $I=[0,1]\subset\mathbb{R}$ to $I^n = [0,1] \times \dots \times [0,1] \subset \mathbb{R}^n$. As mentioned in the correspondence between Cantor and Dedekind, the former, after giving a proof of the above fact, asks the latter an opinion on the found result, using the famous expression "I see but I do not believe" (reported by Cavaillès, 1969, p.169). Below Fischbein's comment to this episode:

> The fact of obtaining a proof supporting the equivalence should have strengthened his conviction. But four days after having written to Dedekind that his conviction is without restriction and after having obtained the formal proof, he still seemed shocked by his discovery. It seems that, while most of the mathematicians of his time were trapped in one intuition, Cantor was trapped in two contradictory intuitions: the old, "natural" intuition, according to which two continuous sets of points having a different number of dimensions cannot be equivalent and the new, the Cantorian intuition claiming the equivalence of the two sets. The fact of obtaining a proof supporting the second view (the equivalence of the two sets), does not solve the subjective conflict. Cantor is no less anxious and worried after obtaining the proof than he was before. He presses Dedekind to help him to overcome the difficulty:
>
> ". . . the result I have informed you about, appears to me so unexpected, so new that I would not be able to find my spiritual quietness before receiving, dear friend, your opinion about the exactitude of my finding." (Cavaillès, 1969, p. 169).
>
> (Fischbein, 1987, p. 26)

As highlighted by Fischbein, although extremely specific, this example shows that the feelings of obviousness, certitude, and consistency in mathematics, even for experts, are not necessarily based on formal logical deductions only. He writes:

> The drama of certitude is played at another level, with other mechanisms than those related to the dynamics of algorithms and logical connectives. (Ibidem, p.27).

This is confirmed by the voice of the mathematician Hadamard who in his essay 'The psychology of invention in the mathematical field' (1954) notes how in his activity as mathematician the informal and intuitive aspects have a fundamental role. For instance, he writes:

> The true process of thought in building up a mathematical argument is certainly rather to be compared with the […] the act of recognizing a person. […] [A]ny mathematical argument, however complicated, must appear to me as a unique thing. I do not feel that I have understood it as long as I do not succeed in grasping it in one global idea. (Hadamard, 1954, p.65)

As an example of this, he describes what appears in his mind as he reflects on the steps of the proof of the fact that prime numbers are infinite:

> I shall repeat the successive steps of the classic proof of that theorem, writing, opposite each of them, the corresponding mental picture in my mind. We have, for instance, to prove that there is a prime greater than 11.

STEPS IN THE PROOF	*MY MENTAL PICTURES*
I consider all primes from 2 to 11, say 2, 3, 5, 7, 11.	*I see a confused mass.*
I form their product *2 x 3 x 5 x 7 x 11 = N.*	*N being a rather large number, I imagine a pint rather remote from the confused mass.*
I increase the product by 1, say N plus 1.	*I see a second point a little beyond the first.*
That number, if not a prime, must admit of a prime divisor, which is the required number.	*I see a place somewhere between the confused mass and the first point.*

> (Ibidem, pp. 76-77)

Hadamard then observes that the mental images previously described, not useful for what concerns the formal validation of the proof, which depends solely on the correctness of the involved deductions, have in fact a key role in his understanding of the proof. Namely, he writes:

> [O]ne can easily realize how such a mechanism or an analogous one may be necessary to me for the understanding of the above proof. I need it in order to have a simultaneous view of all elements of the argument, to hold them together, to make a whole of them […]. It does not inform me on any link of the argument (i.e., on any property of divisibility or primes) ; but it reminds me how these links are to be brought together. (Ibidem, p. 77)

Hadamard's words thus highlight how to him, from a cognitive perspective, a proof is not made solely by the chain of logic deductions it consists of, but also by a global vision, an intuition, we could say, that summarizes the proof itself in his mind, and without which “one would be like the blind man who can walk but would never know in what direction to go.” (Ibidem, p. 105).

To recap what said before, then, Fischbein's position, supported by the above mentioned examples by Cantor and Hadamard, is that the logic and formal aspects of a proof, although ensuring its validity from a mathematical standpoint, could very well be separated from the feeling of evidence and certainty that bring someone to feel, intuitively, that “it is necessary true”, “it must be so”.

What we just described could have an important role for what concerns the research objectives of this study. Indeed, as noted in the literature review, several studies (Movshovitz-Hadar, 1993; Carotenuto et al., 2018) showed how some students, even among those expert and capable of constructing a proof by MI, react by doubting their acceptance of MI as a method of proof, if put in a potentially contradictory situation, like a (non-correct) proof by MI in which something clearly false is proved. In relation to this, Movshovitz-Hadar (1993) uses the term 'fragile knowledge'. Therefore, Fischbein's theoretical standpoint on intuitions allows to consider a sort of Cantorian “I see it, but I do not believe it”, which may be rewritten as "I can prove [by MI], but I do not believe it" and which can be perceived by some students.

Another, not secondary, aspect upon which this theoretical standpoint is built is that it has been already employed in relation to other types of mathematical proofs, as those by contradiction and by proof of the contrapositive statement (Antonini, 2004, 2019). Fischbein's standpoint turned out to be extremely effective in these studies to analyse and interpret student's difficulties in relation to these types of proofs. In section 3.3.2 I presented a model of proof by MI based on the construct of a theorem as triplet. This model allowed me to detect a parallel between the proving scheme of MI and that of the indirect proofs (particularly the presence of a meta-theorem and some auxiliary theorems through which the proof of the main statement is achieved). This analogy suggests that Fischbein's standpoint on Intuition could be useful for what concerns the analysis and interpretation of MI related difficulties as well.

Having said that, in the next paragraph I shall present the theoretical construct of *Intuitions* introduced before. Next, I shall highlight how this construct has been declined in relation to the theme of *intuitive acceptance* of a mathematical proof. Lastly, I shall present some aspects related to MI analysed by Fischbein within this theoretical standpoint.

5.2 THE FRAMEWORK OF INTUITIONS

As described by Fischbein (1987, pp. 3-5), the term *intuition* is used in the literature with several meanings often far from each other. To clarify his standpoint, the author affirms that he uses the term as a synonym of *intuitive knowledge*. Fischbein's starting point is that thinking is a form of behaviour and thus requires to stand upon solid reference points. Similar to what happens for practical behaviour, when someone relies on real tools and practical operations on them, when thinking involves ideal objects and operations, as in mathematics, some consistency and evidence features close to those of real tools must be assigned to these mental objects. According to Fischbein, the latter is precisely the role of intuitions. He writes:

> Our theory is that mental behavior (reasoning, solving, understanding, predicting, interpreting) including mathematical activity, is subjected to the same fundamental constraints. The mental "objects" (concepts, operations, statements) must get a kind of intrinsic consistency and direct evidence similar to those of real, external, material objects and events, if the reasoning process is to be a genuinely productive activity. **An intuition is, then, an idea which possesses the two fundamental properties of a concrete, objectively - given reality; immediacy - that is to say intrinsic evidence - and certitude (not formal conventional certitude, but practically meaningful, immanent certitude)**. (Fischbein, 1987, p.21, bold added by me)

Therefore, in other terms:

> Intuition fulfils, at the intellectual level, the function fulfilled by perception at the sensorial level: intuition is the direct, cognitive prelude to action (mental or practical). It organizes information in a behaviorally meaningful and intrinsically credible structure. (ibidem, p. 56).

Despite this parallel between intuitive knowledge and perception, both a direct form of knowledge that does not require justification, Fischbein stresses that they are not the same thing. Indeed, as opposed to perception, an intuition "always exceeds the given fact. […] [I]t implies an extrapolation beyond the directly accessible information" (Fischbein, 1987, p. 13). For instance, if a subject looks at two intersecting lines, s/he can perceive that pairs of opposite angles are equal, but this is different from the intuitive knowledge that every pair of intersecting lines necessarily forms pairs of equal opposite angles.

Thus, finally, **intuitive knowledge is:**

> a kind of knowledge which is not based on sufficient empirical evidence or on rigorous logical arguments and, despite all this, one tends to accept it as certain and evident. (Ibidem, p. 26).

5.2.1 General characteristics of intuitions

Fischbein (1987) determines a series of features that characterise intuitions as they are intended in his model. Their description, presented below, contributes to further characterise the meaning of *intuition* to Fischbein.

The "fundamental characteristic of intuitions" (p. 43) is **self-evidence.** This feature lets an intuition be considered valid by a subject "without the need for any justification" (p.43). As I will explain later, this can have subtle consequences for what concerns the proof of mathematical facts that are considered intuitively evident by a subject, and for which those subjects could not perceive the cognitive need for a proof.

Intuitions are then characterized by their **intrinsic certainty**, i.e. the property by which they are accepted as certain by the subject. Fischbein observes that self-evidence and certainty, although highly correlated are not identical. For instance, we are convinced that the theorem of Pythagoras is true, but it is not self-evident.

The above two characteristics of intuitions have an immediate and remarkable consequence: an intuition, whether or not it corresponds to a mathematically correct statement, is extremely robust and hard to change or remove. As highlighted by Fischbein, "Experience has shown that robust intuitions - no matter if they are correct or not - tend to survive even when contradicted by systematic formal instruction." (p. 47).

This comment introduces the third characteristic of intuitions, their **perseverance**: "intuitions, once established, are very robust" (p. 47). This aspect is delicate because it implies that **"**erroneous intuitions may survive together with correct, conceptual interpretations all our life" (p.47). This is related to a central aspect of Fischbein's theoretical standpoint, that is *overconfidence,* "a selection activity which is aimed to preserve, automatically, those data which seem to support a certain conception and at the same time, to ignore those contradicting it." (p.33). Thus, thanks to this, an intuition, for instance erroneous, could keep being fed and finding confirmations of its validity for someone even in those situations that could instead show it is not mathematically correct.

A further aspect that characterises intuitions is their **coerciveness**: "intuitions impose themselves subjectively on the individual as absolute, unique representation or interpretations" (p.47). This makes an intuition exerts a coercive effect on a subject's way of reasoning.

Perseverance and coerciveness, as noted by Fischbein, are potentially problematic characteristics of intuitions from a didactic point of view:

> The survival of such contradictions between intuitive, robust representations and scientifically acquired concepts is a permanent source of difficulties for the teacher. Very often the main recommendable procedure is to make the student aware of the conflict and to help him to develop control through conceptual schemas over his intuitions. (p.47).

The next characteristic of intuitions determined by Fischbein is the property of not being a mere skill or perception but having a **theory status** as well. Indeed, as he describes:

> An intuition is never confined only to stating the universality of a property nor to the perception of a certain fact. In an intuition one generally grasps the universality of a principle, of a relation, of a law - of an invariant - through a particular reality. (p.50).

At the same time, an intuition is not a pure theory, being it "expressed in a particular representation using a model: a paradigm, an analogy, a diagram, a behavioral construct etc." (p. 50).

As already described above and in relation to the previous point, an intuition is further characterized by its **extrapolativeness**: “an intuition always exceeds the data on hand” (p.51). It must be noted that an intuition is more than an extrapolative guess. Indeed, as described by Fischbein:

> a feeling of certainty is also a necessary characteristic of an intuition. Otherwise it is a mere guess. It is this particular combination of incompleteness of information and intrinsic certitude which best characterizes an intuition. (p. 51).

One more aspect that characterises an intuition is its **globality**: **“**Intuition is a structured cognition which offers a unitary, global view (or insight) of a certain situation” (p. 53). Following this, an intuition is a synthetic view, “as opposed to analytical thinking which is discursive in its very nature” (p.53). Hadamard’s words, already quoted above, are particularly fitting in relation to this, and are indeed mentioned by Fischbein himself:

> [A]ny mathematical argument, however complicated, must appear to me as a unique thing. I do not feel I have understood it as long as I do not succeed in grasping it in one global idea. (Hadamard, 1954, p. 65).

The last characteristic determined by Fischbein for intuitions is **implicitness**: “intuitive reactions are in fact the surface structure expression of tacit, subjacent processes and mechanisms” (Fischbein, 1987, p. 54). To illustrate this characteristic, Fischbein shows the following examples:

> The student who accepts the Euclidean postulate of the parallels is generally not aware that he extrapolates a limited experience beyond any possibility of verification. A person asked to define the notion of solids is generally not aware of the fact that he has in mind the paradigm of a certain solid object - for instance a piece of metal or a stone. The definition is inspired by the particular instance. (Ibidem, p.55).

This ends the description of those properties that characterize intuitions. In the next paragraph we will introduce a further aspect of this theoretical standpoint, discussed in Fischbein (1999), that is the relation between intuitive and formal knowledge of a mathematical statement.

5.2.2 Intuitive and formal knowledge

As said, an intuition related to a certain mathematical statement can give a subject that feeling of certitude and evidence of the fact that “it must be so”, without requiring any further justification. At the same time though from a mathematical point of view every statement needs to be proven inside some theory before being considered valid. Fischbein (1999), based on this distinction between acceptance of a statement at the intuitive or the logic-mathematical level, proposes the following differentiation to describe a subject’s knowledge of a mathematical statement:

- **Intuitive knowledge of a statement**. This is based on “[a] category of cognitions which appear directly acceptable as selfevident. These are intuitive cognitions.” (p. 18).
- **Formally based knowledge of a statement**. This in based on “[a] category of cognitions which are accepted indirectly on the basis of a certain explicit, logical proof. These are logical, or logically-based cognitions.” (p.18).

In light of this distinction, Fishbein (1999) remarks how, according to the specific mathematical statement, several situations can arise for what concerns the relation between a subject’s intuitive and formal knowledge of it.

Statements which are intuitive accepted and, even mathematically, they do not require a proof.

This is the case of axioms of specific theories which are intuitively evident. For example, the axiom of Euclidean geometry affirming that 'given two distinct points there is one and only one straight line determined by them'. Fischbein notes that this case should not be underestimated from a didactic standpoint. Indeed, it is worth stressing that these statements do not require a formal proof due to their epistemic role in their respective theory, and not because they are intuitively true. In modern mathematics, axioms upon which a theory is built "are no longer established by their apparent self-evidence, but in accordance with the axiomatic system established by the designer of the system." (Fischbein, 1999, p. 18). This aspect can also generate the opposite situation, that is when one or more axioms of a theory are not intuitively evident. This can be didactically and cognitively problematic because they must be accepted by a subject without any intuitive justification or mathematical proof.

Statements which appear intuitively to be true, nevertheless they require a mathematical proof.

This, for example, is the case of the statement 'In a triangle, the sum of two sides is always bigger than the third side'. In this case there is "coincidence between intuitive acceptance and a logically-based conclusion." (Ibidem, p. 21). Fischbein notes that, despite this coincidence between the intuitive and logic conclusion of the statement, these situations can still be didactically problematic because they raise the problem of the necessity of a proof. Often, in these situations

> The student's immediate reaction is: 'What for'? 'Why shall we prove what is evident'? The student does not accept that an intuitively evident statement should be formally proven. The proof seems to be superfluous to the student and the requirement to prove a statement which appears obvious may strengthen the student's feeling that mathematics is an arbitrary, useless, whimsical game. (Ibidem, p. 22)

Statements which are not self-evident and which have to be proved mathematically.

In this case, no intuition is linked to the involved statement. Fischbein notes that several mathematical statements fall in this category. For instance, he mentions the formula to solve quadratic equations, which "does not rely on any intuitive support" (Ibidem, p. 27). As noted by Fischbein, even these situations can potentially be problematic from a didactic point of view since in accepting these types of statements "the student has to rely only on the formal mathematical truth, on formal proofs and definitions" (ibidem, p. 27). Thus, in general, this aspect highlights a subtle point with respect to the relation between intuitive and formal knowledge:

> The student has to get used to the idea that mathematics is, by its very nature, an abstract, formal, deductive system of knowledge. Intuitive models are very often useful, but they are not always possible. (ibidem, p. 27).

Similar situations can often arise when a mathematical definition is extended in order to guarantee the logic consistency of the theory of reference. For instance, the definition of product of negative numbers and expressions like $a^0=1$ or $0!=1$.

Situations in which there is a conflict between intuitive and formal knowledge related to a statement.

This is, for example, the situation involving Cantor's "I see but I don't believe" described at the beginning of this chapter. As observed by Fischbein, these situations are extremely delicate in teaching:

> Such situations appear very often in the teaching of mathematics. The student may not be aware of the conflict. The effect is that the student does not accept, does not understand the formal statement, or even when he seems to understand initially, he tends to forget it and the intuitive interpretation is that which decides the student's solution. In our opinion it is helpful that the students should become aware of the conflict. Simply ignoring the conflict (that is, the intuitively erroneous reaction) leaves untouched the original intuition. Thus the conflict remains latent and finally the student will, probably, forget the formal, mathematically correct answer. (Ibidem, pp. 26-27).

<u>Situations in which two conflicting intuitions may appear.</u>

These are situations where a subject associates two intuitions that contradict each other. Even these situations, as the previous ones, are didactically subtle. Fischbein reports the example of the equivalence of the cardinality of the sets N and 2N, observing how a different representation of them could lead to two conflicting intuitions:

> Considering the set {1, 2, 3, 4, 5, 6 . . .} it seems intuitively that the set of natural numbers and the set of even numbers are not equivalent. But let us consider the following representation:
>
> {1, 2, 3, 4, 5, 6, . . .}
>
> {2, 4, 6, 8, 10, 12 . . .}
>
> (Fischbein, 1982, p. 21)

According to the author, the latter representation can intuitively hint at the bijective correspondence between the elements of the two sets. On the other hand, the former one highlights how 2N is a proper subset of N, which may lead to the intuition that the two sets cannot have the same cardinality.

Sometimes the reaction of a subject exposed to conflicting intuitions is to find a *compromise* between the two intuitions. This compromise could not correspond to the mathematically correct interpretation of the situation; thus, these cases should not be underestimated didactically. The following is an example taken by Fischbein (1987):

> if one compares the sets of points in two line segments of different lengths, one obtains two opposite conclusions, both intuitively valid: the two sets are equivalent because both are infinite; The two sets are not equivalent because the longer segment contains more points. (p. 208)

Reacting to these two conflicting intuitions, a student could find a compromise in the idea that points in the longer segment are bigger than those in the shorter one (Ibidem, p. 125). This intuition would allow the subject to overcome the contradiction of the two intuitions, but it would lead to a conclusion that contradicts the theory of Euclidean geometry.

The construct of intuitions just described is extremely general and it has been declined in relation to several research topics in ME. For what concerns this thesis objectives, I shall focus on those aspects related to mathematical proofs. In the next paragraph I will describe a classic work by Fischbein (1982) where the relation between intuition and proof is discussed.

5.2.3 Intuition and proof – Three levels of acceptance

The proof of a theorem within a specific theory is what ensures its acceptability from a logic-mathematical point of view. Nevertheless, this proof could not give to a subject the intuitive acceptance of the validity of the proven statement. Fischbein (1982) discussed this aspect emphasizing how the intuitive knowledge of a mathematical theorem can consist in three different levels:

1) An initial level that involves the intuitive acceptance of the statement itself. More specifically, the above classification showed how a subject may or may not agree, intuitively, with the statement involved in the theorem.

2) A second level that involves the intuitive acceptability of the mathematical structure of the proof. The intuitive acceptability of a proof is not limited to the understanding of each logical step involved in the proof, but it requires grasping the proof in a unique synthetical form.

3) A third level that involves the intuitive acceptability of the generality of the proof. In other words, it "refers to the fact of understanding the universal validity of the statement as guaranteed and imposed by the validity of the proof" (p. 18).

Fischbein observes how, although logically there is no difference between accepting a statement, accepting one of its proofs, or accepting that the universality of the statement is ensured by the proof, things are rather different from a cognitive point of view. He thus proposes an empirical study where a written (and correct) proof of the statement '$E=n^3-n$ is divisible by 6, for every n' is showed to some students. Students are then asked if they think that the statement is true, if the proof is correct, and if the proof guarantees the generality of the statement. What Fischbein notes is that there are several instances where the answers given by the students are in contradiction:

> We find here a revealing cleavage in the pupils' answers, leading to an interesting hierarchy: 81.5% affirm that the proof is fully correct; 68 5% agree with the theorem; for only 60% of the subjects the generality of the proof is justified by the theorem. It was found that only 41% of the subjects answered all of these three questions correctly. (Ibidem, p. 16).

Next a (false) counterexample (n=2357) is showed to the students, and their opinion on that is recorded. Fischbein notes that only the 32% of them questions the fact that there is a counterexample. Students are then asked if "additional checks (with other numbers) are necessary in order to increase your confidence in the theorem". Fischbein records the answers of those that had previously accepted the correctness of the proof and notes that only the 24.5% of them answers that additional checks are not necessary. In these cases, Fischbein writes, a student seems to be

> still oscillating between a formal - but not powerful enough - understanding of a mathematical statement and the intuitive, empirical tendencies which push him to seek new, additional facts which, he feels, may increase his conviction, his belief in the general validity of the already proved statement. (Ibidem, p.24)

Starting from this analysis, Fischbein observes how these two components, the intuitive and the logic-mathematical one, associated to the proof of a theorem must be cognitively blended for a subject to reach an intuitive acceptability on all the three levels described above.

> The main point, in my opinion, is that the logical form of necessity which characterizes the strictly deductive concatenation of a mathematical proof can be joined by an internal structural form of necessity which is characteristic of an intuitive acceptance. Finally, both can blend in an unique synthetical form of mathematical understanding. (Ibidem, p.15)

When this happens, it becomes possible for a subject to associate to a proof a series of features that makes the proof intuitively accepted. To summarise, thus,

> a mathematical truth can become really effective for productive mathematical activity if, together with a formal understanding of the respective truth, we can produce that kind of synthetic, sympathetic, direct acceptability of its validity. (Ibidem, p. 18).

The three levels of intuitive acceptance of a proof presented by Fischbein offer a rather general framework to decline the theme of intuitions in relation to mathematical proofs. Before moving on, I shall discuss to which extent this framework will be employed in this thesis. In particular, in this study, I will focus on the intuitive acceptability of proofs by mathematical induction, thus on the second level described by Fischbein. More precisely, I will investigate to which extent the structure of a proof by MI is intuitively accepted. This will also bring me to consider the third level of intuitive acceptability, that is the one related to the generality of the statement as univocally ensured by its proof. Indeed, an intuitive knowledge of why MI 'works' (a "cascade" of syllogisms, to quote Poincaré) seems to contain a justification of the reason why the proof ensures the generality of the statement. In fact, in a proof by induction every natural number *n* following the inductive basis is intuitively reached by the proof itself. Thus, this remark seems to suggest that the intuitive acceptability of how the MI works could be closely related to the intuitive acceptability of the fact that a proof by MI may guarantee the generality of a statement.

5.2.4 Intuitions and Mathematical Induction

In this last paragraph we introduce two specific aspects of MI that have been discussed by Fischbein within the perspective on intuitions just presented: the transition from potential infinite to actual infinite in the justification of the functioning of MI, and the proof of the inductive step in a proof by MI.

<u>The intuitive leap between potential and actual infinity.</u> – from Fischbein (1987, pp. 51-52)

As said before, an intuitive justification of MI's functioning can be built through a series of syllogisms: from the validity of the basis and the inductive step it can be derived that the predicate P holds for 0, from which, by modus ponens, P holds for 1 as well, from which, again by modus ponens, P holds for 2, and so on, up to every integer $n \geq 0$. Fischbein (1987) notes how, according to this justification, "an intuitive leap" (p. 51) is necessary in order to conclude that P holds on N. Indeed, Fischbein affirms that through intuition it is possible to conceive the potential infiniteness of the process only, but when dealing with the actual infinity of the whole set N, then a subject cannot rely on intuitive knowledge anymore. He writes:

> The principle on which mathematical induction relies is not gained by experience. Through experience one may learn that the inference under discussion is true for the first ten or for even the first hundred numbers but not for the whole infinite set of natural numbers. […] This kind of intuitive leap has to intervene always when dealing in mathematics with infinite processes or infinite sets. As long as one has to do with the dynamic form of infinity there is no apparent difficulty. It seems that one is naturally able to conceive of the indefinite continuation of a process like that of constructing always greater numbers or of extending indefinitely a line. […] **The notion of dynamic infinity expresses directly, in the purest way, the extrapolative capacity of intuition itself. Things change radically when one tries to pass to actual infinity.** […] **[T]he extrapolative capacity of intuition does not apply to actually given infinite sets.** The extrapolative capacity of intuition is dynamic, constructive in its very nature. Its predictive and explanatory capacity ceases if one refers to infinite sets considered as actually given. **Intuitively, such sets would mean the final state of an endless process, which is contradictory in natural, intuitive terms.** One may then assume that, psychologically, the universality grasped by intuition through a particular, given instance does not refer to an actual universality but rather to a potential one. […] An actually given infinity is a pure logical, conceptual construct, not intuitively acceptable - it has no behavioral meaning, it cannot be self-evident and intrinsically coercive. (pp. 51-52, bold added by me)

Thus, at the end, referring to the theoretical framework discussed above, Fischbein's position can be synthetized as follows: the intuitive acceptability of MI's scheme of proof can be reached at the potential infinite level, but in the transition to actual infinite that intuition must be paired up with formally based knowledge. We note how this position is consistent with the GD of MI presented in the APOS theory (see section 4.4) in which I highlighted how an encapsulation of an infinite process, necessary in order to go from a conceptualization of infinite as (potential) process to (actual) object, can be cognitively problematic.

<u>The inductive step in a proof by MI</u> – From Fischbein & Engel (1989)

Fischbein and Engel (1989) analyse some of the difficulties met by students in intuitively accepting the validity of a proof by MI. In their works, the authors focus on the proof of the inductive step in a proof by MI. In the proof of the implication P(k)→P(k+1), for a generic k, the inductive hypothesis P(k) is assumed, and P(k+1) is proved using this assumption. According to the authors, this step is cognitively problematic:

> What is difficult to understand is that P(k) (the inductive hypothesis) is postulated in the reasoning process (in the inductive step) not as a proved fact but as in hypothesis – that is with its initial status. The difficulty is that the student has to build the entire segment of the induction step (if P(k) is true then P(k+1) is also true), on a statement which, itself, has not been proven and cannot be proven in this segment of the reasoning process. (Ibidem, p. 284)

In the inductive step, the validity of the implication P(k)→P(k+1) itself is proven, and this is done without any concern of the truth value of the two involved statements. Fischbein and Engels affirm that this aspect is extremely delicate from an intuitive point of view:

> As a matter of fact, we are absolutely not used to this way of reasoning. […] **It is like building a bridge in the air without any support on both its ends.** […] The idea that one has to prove an implication p→q for which the problem of the objective truth of each of the two components, p and q, is totally irrelevant (in the realm of the induction step) seems to be intuitively unacceptable. This situation is complicated by the fact that the antecedent p includes the theorem to be proven. And then the student, being inclined to look for a truth value for the antecedent, is puzzled by the fact that the acceptance of the antecedent depends on the theorem which has to be proven. (Ibidem, pp. 284-285, bold added by me).

Thus, according to the authors, students can find themselves in a situation of intuitive conflict. When P(k)→P(k+1) is proven, P(k) is assumed as hypothesis, but why is this legit? From a logic perspective, this is guaranteed by the law of the introduction of the implication. This law states that, in order to prove an implication $p \to q$ from a set Γ of hypotheses, it is possible to add p to the set Γ, to then prove q from that.[70] However, Fischbein and Engel record how several students react to this situation of seeming intuitive conflict with compromises that do not align with the mathematically correct answer. Below I report the students answers as they were collected by the authors in the original work (Ibidem, pp. 282-283):

- <u>The inductive hypothesis has a limited validity.</u> "The inductive hypothesis should be considered true till the contrary has been proven".
- <u>The truth of the inductive step is granted.</u> "The inductive hypothesis is true and therefore we may rely on it".
- <u>The truth of the inductive hypothesis cannot be proven</u>. "It is totally impossible to prove the inductive hypothesis".

[70] This point has already been discussed in section 3.3.2.

- There is no relationship between the inductive hypothesis and the steps of the proof by MI. "It is possible to find that the inductive hypothesis is rejected and, despite this, the (inductive) proof is correct".
- The truth of the basis confirms the inductive hypothesis. "If the truth of the basis has been confirmed one may assume that the number k, which appears in the inductive hypothesis, is the same initial number and then the statement is true for k".
- The truth of the induction step confirms the inductive hypothesis. "If the inductive hypothesis had not been true, we could not have proven the induction step".

5.3 THE FRAMEWORK OF INTUITIONS WITHIN THIS WORK

As said, according to the framework on Intuitions by Fischbein, the intuitive acceptance of a theorem can involve three different stages:

1. The acceptance of the statement.
2. The acceptance of the structure of the proof.
3. The acceptance of the fact that the proof alone ensures the generality of the statement. Else, in other words, that the proof guarantees that there cannot be any counterexample.

Having said that, the research objective of this thesis within this perspective is to investigate the extent to which the proof by MI is intuitively accepted by students, that is item (b) in the above list.

This aspect can be further deepened by considering the model of a theorem proved by MI built upon the construct of mathematical theorem as a triplet, introduced in section 3.3.2. Indeed, this model allowed us to observe that three different theorems are simultaneously present in the proof by MI of a statement $S=\forall n \geq n_0\ A(n)$:

- The theorem (S1, P1, T) corresponding to the inductive base and its proof within the specific theory T.
- The theorem (S2, P2, T) corresponding to the inductive step and its proof within the specific theory T.
- The meta-theorem ($S^* \rightarrow S$, MP, MT), where $S^*=S1 \wedge S2$, whose proof lies in a specific meta-theory.

Starting from this, the intuitive acceptance of a proof by MI can be suitably reinterpreted as the intuitive acceptance of the three above theorems. We shall see how some aspects of each of them can be problematic from an intuitive point of view.

Intuitive acceptance of the theorem (S1, P1, T).

The statement $S1=A(n_0)$ is often trivial, that is in some sense P1 corresponds precisely to S1. Even in these, seemingly trivial, cases the intuitive acceptability of the statement and its proof should not be underestimated. Indeed, sometimes the inductive basis requires the use of specific properties that may not be intuitively evident (e.g. $0!=1$ or $a^0=1$). [71]

In most cases the statement S1 does not involve a universal quantifier. Here speaking of intuitive acceptance of the generality of the proof does not make much sense. However, this could not be case in some other situations. For instance, in the proof by MI of the formula for the sum of the interior angles

[71] See section 5.2.2

of an n-sides convex polygon (presented in 3.3.2), S1 is the statement according to which in any triangle the sum of the interior angles is equal to $\frac{\pi}{2}$. In this case, the intuitive acceptance at the third level of (S1,P1,T) is achieved when P1 is conceptualized as guaranteeing the validity of S1 for every triangle in Euclidean geometry.

Intuitive acceptance of the theorem (S2, P2, T).

The intuitive acceptance of the inductive step S2=$\forall$n(A(n)$\rightarrow$A(n+1)) in a proof by MI has been largely discussed in Fischbein and Engel's work (1987) presented above, where the authors record that some students had difficulties both in accepting the assumption of the inductive hypothesis as legit and in understanding that the inductive step has its own proof, and that within this proof only the inductive hypothesis is supposed to be true so to prove the validity of the implication A(n)$\rightarrow$A(n+1). Fischbein and Engel's analysis of these difficulties can be reinterpreted in terms of the theorem as a triple model. We can describe these difficulties in terms of a missing conceptualization of the theorem (S2,P2,T) as a theorem on its own, whose proof P2 does not depend on the validity of the main statement S or that of S1. As observed when we introduced the model of theorem as a triplet, the proof P2 of the inductive step relies on the logic law of the introduction of the implication, which in terms of the model belongs to the meta-theory level and it guarantees that to prove the statement $p\rightarrow q$ one can assume p and, from that assumption, then prove the validity of q. Thus, the intuitive acceptance of the proof P2 of S2 relies on the intuitive acceptance of this logic law as well. The third stage of intuitive acceptability for the inductive step corresponds to accepting that the proof P2 ensures the generality of S2, that is the predicate A(n)$\rightarrow$A(n+1) holds for each natural n. The logic law of the Universal Generalization (UG) guarantees that the proof of A(n)$\rightarrow$A(n+1), where n is free parameter, implies the validity of the statement S2=$\forall$n(A(n)$\rightarrow$A(n+1)). Thus, the third level of acceptance for the inductive step requires more in general the intuitive acceptance of the logic law UG as well.

Intuitive acceptance of the meta-theorem S*$\rightarrow$S.

Informally, what this meta-theorem states is that mathematical induction *works*, i.e. the fact that the validity of '$\forall n\geq n_0$ A(n)' can be deduced from the validity of the inductive base and of the inductive step. The proof of the meta-theorem is a specific formal proof of a meta-theory of reference[72]. In this study we will not focus on the intuitive acceptability of an eventual formal proof of the meta-theorem, but instead on the intuitive acceptance of the traditional recursive argumentation supporting the validity of the meta-theorem: from $A(n_0)$ and $A(n_0)\rightarrow A(n_0+1)$ it follows $A(n_0+1)$ by modus ponens; from $A(n_0+1)$ and $A(n_0+1)\rightarrow A(n_0+2)$ it follows $A(n_0+2)$, and so on. In this case, the third level of intuitive acceptability corresponds to the fact that this chain of syllogisms guarantees the validity of the statement '$\forall n\geq n_0$ A(n)', that is that there is no $n^*\geq n_0$ such that A(n*) does not hold. It is worth mentioning that at this meta-theory level the meta-theorem and the justification as chain of syllogisms just given do not necessarily refer to the specific predicate A(n), but they refer to a generic predicate instead. The intuitive acceptance of the meta-theorem S1$\wedge$S2$\rightarrow$S can thus involve a further degree of generality: it holds independently from the specific predicate A(n) involved.

The reinterpretation of the levels of intuitive acceptance of a proof by MI in relation to the model of theorem as a triplet highlights how, focusing on the intuitive acceptance of MI, we should keep into account the two different theoretical levels that arise from the model:

[72] For instance, the Principle of MI can be proved as a theorem of ZFC set theory.

- The level of the theory T in which the statements S1, S2 and the proofs P1, P2 are contained. At this level the intuitive acceptance consists in the intuitive acceptance of the specific theorems corresponding to the inductive basis and step, and thus it depends on a specific predicate P(n).

- A meta-theoretical level MT. At this level the intuitive acceptance involves the meta-theorem $S1 \wedge S2 \rightarrow S$ and the logic laws of introduction of the implication and universal generalization employed in the proof P2 of S2.

In conclusion, starting from this analysis, the research objective of this study within the theoretical perspective of intuitions can be rewritten as follows. I aim to investigate the intuitive acceptance of proof by induction seen as the intuitive acceptance of the three theorems (S1, P1, T), (S2, P2, T), (MS, MP, MT), both as independent theorems and as a system of theorems that are all necessary for the proof S of the main statement P. This objective shall be reformulated as a research question, together with others, at the end of the introduction of the conceptual framework.

6 A SEMIOTIC PERSPECTIVE

The last perspective considered in the conceptual framework of this thesis is a semiotic one, to whose presentation this chapter is dedicated.

6.1 PREMISES

In the last decades, the semiotic has been seen to have an increasing role in research in mathematics education. Research has showed that the analysis of signs can offer an access to mathematical thinking and allows to discover interesting processes with important didactical implications (for an overview of several classic semiotic perspectives, see Presmeg et al. 2016). Recently, the semiotic perspectives in ME have been enriched including the analysis of a wide spectrum of signs (mathematical symbols, diagrams, sketches, language, gestures, manipulation of artefacts, etc.). Arzarello (2006) uses the term "multimodal" to indicate a semiotic analysis which focuses on these several categories of signs.[73] In particular, he considers the different types of signs as an inseparable unit, and he introduces the notion of the *Semiotic Bundle* to take into consideration this aspect.

Before presenting in detail the construct of the Semiotic Bundle, I wish to present the motivations to include a semiotic perspective within the conceptual framework of this thesis. This perspective, first of all, will provide a lens trough which to observe processes involved in students' construction of recursive argumentations or proofs by MI. In particular the use of a multimodal semiotic perspective could highlight aspects that otherwise would remain unexplored. Recently, several studies showed, for instance, the importance of analysing gestures to investigate the processes involved in argumentation and proof (Edwards, 2010, Arzarello & Sabena, 2014, Krause, 2016, Sabena 2018). Such a semiotic perspective, thus, allows me to include in this study the analysis of the multimodal semiotic production of students involved in the construction of argumentations or proof. In particular, with this theoretical lens, I will observe the processes of students involved in problem-solving activities with the aim to investigate the roles of the use and production of multimodal signs for the construction of recursive argumentations or proofs by MI.

This chapter develops as it follows. Firstly, I will present the construct of the Semiotic Bundle, focusing on its definition and on its general characteristics; then I will briefly deepen the theme of the analysis of gestures, introducing a series of elements which will be used later in this thesis; lastly, I will introduce to what extent the presented perspective will be used in this study.

6.2 THE CONSTRUCT OF THE SEMIOTIC BUNDLE

The construct of the Semiotic Bundle (Arzarello, 2006) is theorised starting from some theoretical premises.

First of all, the construct is inserted within the tradition of the Peircean semiotic, for whom: **"[A] sign, or representamen, is something which stands to somebody for something in some respect or capacity"** (Peirce, CP, 2.228, bold added by me). From this definition two important characteristics emerge. The first one is that Peirce's definition is extremely general; in particular, "any perceivable (spatio-temporal) entit[y]" can be a sign (Arzarello, 2006, p. 271). The second characteristic is that an entity is a sign only in

[73] Initially the term "multimodal" is used by Arzarello (2006) with reference to the learning processes. In particular, he writes: "learning processes happen in a multimodal way, […] where more semiotic sets are active at the same moment". (p. 283). A multimodal semiotic perspective, therefore, is a semiotic perspective which takes into consideration this variety and simultaneity of signs produced and used during teaching-learning processes.

relation to a subject which interprets it as such. A classic example for this point is the one of a footprint in the sand, made by Peirce (CP, 4.531), and analysed by Arzarello (2006):

> For example, a footprint in the sand generally is not a semiotic representation in this sense: a person who is walking on the beach has no interest in producing or not producing it; however, the footprint that Robinson Crusoe saw on day was the sign of an unsuspected inhabitant of the deserted island, hence he gave it a semiotic function and for him the footprint became a semiotic representation. (p. 271)

The second theoretical premise necessary to frame the construct of the Semiotic Bundle is that it is also inspired by a Vygotskian perspective in which a subject's semiotic production is not seen as a passive external process which reflects the subject's thought, but it is considered as an active process which structure the thought itself. In particular, following Vygotsky there is an analogy between the role that signs have for the thought and the role that tools have for material labour

> [T]he invention and use of signs as auxiliary means of solving a given psychological problem (to remember, compare something, report, choose and so on) is analogous to the invention of tools in one psychological respect. The signs act as instrument of psychological activity in a manner analogous to the role of a tool in labour (Vygotsky, 1978, p.52)

Within this perspective, thus, signs are considered *resources* at disposition of the subjects involved in a particular activity.

A third preliminary assumption at the base of the construct of the Semiotic Bundle is that in order to be able to investigate the complexity of didactical phenomena it is necessary to extend the variety of signs which are traditionally considered, by including also less coded categories of signs such as gestures, written sketches, utterances, gazes, etc. In other terms, Arzarello's point of view is that the traditional notion of 'semiotic systems' (Duval, 2001, 2002, and 2006; Ernest, 2006) should be extended so to include a wide spectrum of different signs. In particular, Arzarello (2006) writes:

> [W]hen cognitive processes must be considered, […] it is the same notion of signs and of operations upon them that needs to be considered with a greater flexibility and within a wider perspective […]. For example, observing students who solve problems working in group, their gestures, gazes and their body language in general are also revealed as crucial semiotic resources. Namely, non-written signs and non-algorithmic procedures must also be taken into consideration within a semiotic approach. (pp. 274-275).

Starting from these premises, Arzarello introduces the construct of the *Semiotic Bundle*. To do so, he firstly defines the notion of **Semiotic Set** as system of three sets:

> a) A set of signs which may possibly be produced with different actions that have an intentional character, such as uttering, speaking, writing, drawing, gesticulating, handling an artefact.
>
> b) A set of modes for producing signs and possibly transforming them […].
>
> c) A set of relationships among these signs and their meaning embodied in an underlying meaning structure.
>
> (Ibidem, p. 281)

As he observes, this definition includes a wide variety of cases, "from the compositional systems, usually studied in traditional semiotics (e.g. formal languages) to the open sets of signs (e.g. sketches, drawings, gestures)." (p.281).

At this point, starting from the notion of Semiotic Sets, it is possible to define a **Semiotic Bundle** as:

(i) A collection of semiotic sets.

(ii) A set of relationships between the sets of the bundle.
(Ibidem, p.281, bold added by me).

Arzarello observes that a Semiotic Bundle is a dynamic structure which evolves during the time and depending on the semiotic activity of the subject(s). In particular, the different semiotic sets which compose the bundle can be enriched with new signs or with new modes for producing signs, as well as the set of relationships between the sets of the bundle can evolve. Another important characteristic is that

> a semiotic bundle must not be considered as a juxtaposition of semiotic sets; on the contrary it is a unitary system and it is only for the sake of analysis that we distinguish its components as semiotic sets." (Ibidem, p.281).

Within this perspective, thus, a sign may have different components depending on the semiotic sets involved in it. For instance, a subject's spoken utterance with a simultaneous gesture, referring to a certain written inscription, can be seen as a unique sign made by three components (the speech, the gesture, and the inscription).

6.2.1 Analysing the Bundle

The just defined construct of the Semiotic Bundle provides a powerful tool to observe and analyse teaching-learning phenomena. Following Arzarello (2006), the analysis of one or more subjects' semiotic production, through the construct of the Semiotic Bundle, can be conducted on two different levels:

Synchronic analysis of the bundle. The mutual relationship between the various semiotic sets is analysed, considering how they are simultaneously activated in the same temporal moment. This analysis can be interpreted as a section on the "time" dimension of the Bundle. In other terms, this, the focus is on 'a picture' of the Bundle at a fixed time. From a practical point of view the "time" dimension is not completely eliminated, but the analysis involves one or more multimodal signs produced or used in an interval of a few seconds. The focus, therefore, is on the multimodality of one or more specific signs, and thus on the different semiotic sets which are simultaneously activated.

Diachronic analysis of the bundle. The temporal development of the Bundle is analysed, considering how it evolves during the activity of the subject(s). In this case, the focus is on the evolution of the bundle itself, both in terms of the specific semiotic sets involved (as, for example, the repetition in time of similar gestures) and in terms of the mutual relationships between the activated semiotic sets. The diachronic analysis, therefore, is made considering a more extended interval of time than the synchronic one, and it allows to focus on the genesis and the development of specific signs during the subject(s) activity.

These two different analyses, synchronic and diachronic, are one transversal to the other and, together, they allow to take into consideration the two structural characteristics of the Bundle: its multimodality and its dynamicity.

In the study presented in this thesis, when the lens of the Semiotic Bundle will be used, **I will consider the Bundle formed by the three semiotic sets: Speech, Written inscriptions (e.g., words, mathematical symbols, drawings, sketches), and Gestures.** Moreover, it will be analysed bot synchronically and diachronically.

In the next paragraph I present a focus on the semiotic set of gestures, introducing a series of terms and theoretical elements that will be used later in this study.

6.3 A FOCUS ON GESTURES

The analysis of gestures in mathematics education is quite recent but it has gained an increasing attention (see, for example, Arzarello & Edwards, 2005). As expressed by McNeill and Duncan (2000), gestures can become "windows" for researchers on the reasonings of the subject performing them: "By virtue of idiosyncrasy, co-expressive, speech-synchronized gestures open a "window" onto thinking that is otherwise curtained." (p. 143). A further assumption supporting the relevance of such analysis for ME is that gestures are relevant not only in communication, but also in thinking processes. As McNeill (1995) writes: "Gestures do not just reflect thought but have an impact on thought. Gestures, together with language, *help constitute thought*" (p.245, italics in original). Recently, the study and analysis of gestures in ME has contributed to investigate processes involved in argumentation and proof. Before presenting this last point, it is necessary to introduce a series of terms and theoretical elements used to analyse gestures.

6.3.1 Theoretical elements for analysing gestures

In this thesis, in line with the so far quoted studies, as gestures I consider the spontaneous "movements […] of the arms and hands closely synchronized with the flow of the speech" (McNeil, 1992, p.11). Following McNeil's point of view, summarised by the given definition, gestures and speech constitute an inseparable unite, two faces of the same process:

> [Gestures] are tightly intertwined with spoken language in time, meaning, and function; so closely linked are they that we should regard the gesture and the spoken utterance as different sides of a single underlying mental process. (McNeill, 1992, p. 1).

However, at the same time, exactly for their different semiotic nature, speech and gestures reciprocally integrate themselves, each providing a different contribution to the subject's discourse:

> Each modality, because of its unique semiotic properties, can go beyond the meaning possibilities of the other, and this is the foundation of our use of gesture as an enhanced window into mental processes. (McNeill & Duncan, 2000, pp. 143-144).

This unity between speech and gesture has been analysed in detail by Goldin-Meadow (2003) who introduced the distinction between speech-gesture **match** or **mismatch** considering if a gesture refers to the same semantic content of the spoken utterance or not. Goldin-Meadow, in particular, highlights the cognitive potentials of mismatch-situations, affirming that trough a mismatch a subject is simultaneously considering and representing different aspects (the ones to which speech and gesture are referring) and this can have an important role in the representation of a new idea. For example, Sabena et al. (2005) register a speech-gesture mismatch produced by a student who, involved in an activity with some geometrical patterns, with her words describes the structure of the figure 5 while she is pointing to the figure 4. The authors observe how through this mismatch the student is considering symoultaneusly the two figures and thus it "provides a link between the two figures" (p. 131).

A traditional classification of gesture is the one provided by McNeil (1992) in relation to the semantic content of the speech accompanying the gesture. In particular the following (not mutually exclusive) categories of signs can be introduced:

- **Iconic gestures.** These are gestures which "bear a close formal relationship to the semantic content of speech" (p.12). These are, for examples, the gesture which mime an objects or actions on those.

- **Metaphorical gestures.** "These are like iconic gestures in that they are pictorial, but the pictorial content presents an abstract idea rather than a concrete object or event." (p. 14). A classic example of a metaphorical gesture is the one by which an expressed idea is gesturally represented as a physical object hold in one hand in front of the speaker.

- **Beats.** These are the gestures that mark the rhythm of the discourse. Typically, they are quick and short movements of the hands. As observed by McNeill:

> beats reveal the speaker's conception of the narrative discourse as a whole. The semiotic value of a beat lies in the fact it indexes the word or phrase it accompanies as being significant, not for its own semantic content, but for its discourse-pragmatic content. (p. 15).

- **Deictics.** These are the pointing gestures, generally performed with the fingers or with objects (such as with a pen). A gesture of this kind has "the obvious function of indicating objects and events in the concrete world, but it also plays a part even where there is nothing objectively present to point at" (p. 18). As written, thus, a deictic gesture can refer to an abstract idea not physically present, as the classic gesture of pointing of one's own back or to the left when talking about an event happened in the past. In examples like this one, the gesture is at the same time deictic and metaphorical.[74]

The just presented classification is not based on the physical characteristic of the gesture, but on its relationship with the semantic content of the related speech. For this reason, a gesture can belong to more than one category simultaneously. For instance, we have already observed that the gesture of pointing to the left one referring to 'the past' is simultaneously a metaphorical gesture and a deictic. Analogously an iconic gesture can also be a deictic if it also refers to the specific position of the represented object. Sabena (2008), showing the example of a gesture formed by thumb and pointing finger kept at a fixed distance as to hold the extremes of a little stick (the "Δ-gesture", as she calls it), observes that "this gesture can be considered an iconic gesture with respect to a segment-line in the Cartesian plane, and a metaphoric gesture, with reference to a numerical interval. These two referents come to be blended (or condensed in the same gesture)." (p. 29)

McNeill, in a successive study (2005), observes and analyses a phenomenon which sometimes occurs in a subject gestural semiotic production, which he calls **Catchment:**

> A catchment is recognized when two or more gesture features recur in at least two (not necessarily consecutive) gestures. (p.116).

Examples of catchments are the repetition of gestures with the same physical features or the repetition of a position in the space in a series of different deictics. Following McNeil, at the base of the presence of catchments in a subject's discourse there is "the logic […] that recurrent images suggest a common discourse theme, and a discourse theme will produce gestures with recurring features" (p.116). In particular, thus, a catchment indicates cohesion in a subjects' discourse, and it could be due to the presence of recurrent images in the speaker's thought. Following this, it becomes important to register the presence of catchments in a subject's semiotic production, because:

> By discovering the catchments created by a given speaker, we can see what this speaker is combining into larger discourse units – what meanings are being regarded as similar or related and grouped together, and what meanings are being put into different catchments or are being

[74] McNeil originally included a fifth category of gestures to the just presented four, the *Cohesive* gestures. These are gestures which "serve to tie together thematically related but temporally separated parts of the discourse" (1992, p. 16). This category will be eliminated by McNeill in his successive works and included within the notion of *Catchment*.

isolated, and thus are seen by the speaker as having distinct or less related meanings. (Ibidem, p. 117).

A further classification of gestures is the one introduced by Krause (2016) which refers to the relationships between a gesture and the system of inscriptions which is present in from of the subject producing the gesture. The researcher observes that "Gestures can refer to mathematical objects on three spatial levels" (p. 134):

1) **Gestures of 1st level.** They are gestures that refer "to something actually represented in a fixed diagram. It does not represent itself but works as an index to hint at something already represented." (p. 138). Examples of 1st level gestures are the pointing gestures which refer to a specific inscription or to part of it, or gestures which highlight an inscription (as circling or underlying with a finger an inscription). Krause refers this level as the level of the **Concrete**.

2) **Gestures of 2nd level.** In this case they are gestures that refer to the system of inscription but, at the same time, they also represent something new:

   On a second level, gesture is embedded in a fixed representation but does not merely refer to an already fixed concrete component. It represents itself but needs to be interpreted against the background of an inscription that provides relations. Such gestures may shape and highlight already fixed entities but also depict new entities in an established diagram as ephemerally embedded in it. (p. 138)

   2nd level gestures are, for examples, those with which a subject represents the movement of a figure (e.g., a rotation or a translation) ore those with which a subject draws with a finger something on the sheet without marking it (e.g. a height of a polygon already drawn). Krause refers to this level as the level of the **Potential*,*** because "a gesture that specifies on level 2 represents a 'hypothetical something' not there but potentially 'thought into' a present diagram" (p.138).

3) **Gestures of 3rd level**. They are gesture that are "free" in respect to the system of inscriptions:

   On this level, meaning is transported by the gesture itself such that it can be regarded as autonomous. Without being dependent on a present referential frame, the interpretation of the gesture is detached from the concrete. That is what makes it 'free'. (Ibidem, p.138)

   From a spatial point of view, they are gesture that are distant from the sheet or the screen which contains the system of inscriptions, and this aspect contributes to make them "autonomous". Krause observes that "these gestures may reveal a more conceptual than contextual idea of a mathematical situation or object." (p.138), and she refers to this level as the level of the **General.**

Krause's classification connects two different aspects of gestures: a spatial one, which is the physical distance between the gestures and the sheet (or the screen) containing the inscriptions, and a semantic one, which is the autonomy of the gesture in relation to the system of inscription. As observed by Krause, these to aspects seem to be intertwined: "The higher the level, the less the interpretation of the gesture depends on a fixed representation and the more detached it is from a concrete situation." (p.138).

The Table 6.1 shows a summary of the just presented theoretical elements for the analysis of gestures with reference to their possible integration in the construct of the Semiotic Bundle.

Theoretical Element	What it considers	How it can be determined within the Bundle	It can be interpreted as
Speech-Gesture Mismatch. (Goldin-Meadow, 2003)	Mismatch between the information expressed through gesture versus speech.	Synchronic analysis. Focusing on the relation between the Gesture and the Speech features of a multimodal sign.	It shows that the subject is symoultaneusly considering different entities.
Classification of Gestures in *Iconic, Metaphoric, Beats, Deictics*. (McNeill, 1992)	Relationship of a gesture with the semantic content of the speech.	Synchronic analysis. Focusing on the relation between the Gesture and the Speech feature of a multimodal sign.	Iconic and Metaphoric gestures give access to mental images accompanying the speech. Beats show part of the discourse that the subject feels relevant to stress.
Catchments (McNeill, 2005)	Recurrences of gestures features in two or more (not necessarily consecutive) gestures.	Diachronic Analysis. Focusing on the development of semiotic set of Gestures.	They show what meanings are being regarded by the subject as similar or related and grouped together, and what meanings are seen as having distinct or less related meanings.
Referential Levels of Gestures (Krause, 2016)	Detachment of a gesture from a concrete inscription.	Synchronic Analysis. Focusing on the relation between a Gesture and the semiotic set of Inscriptions.	It shows the level of generality of the information expressed through gesture by the subject.

Table 6.1. Overview of the just presented theoretical elements on gestures analysis with reference to the synchronic and diachronic analysis of the Semiotic Bundle. The last column on the right describes how each element can be interpreted following the reference literature.

6.3.2 Gestures' analysis in investigating on argumentation and proof

Recently in Mathematics Education, the analysis of gestures has been used to investigate processes involved in argumentation and proof. Edwards (2010), for instance, analysing the gestures in expert students in mathematics (PhD students) produced during the construction of proofs, observes that they represented the proof with a gesture shaping a path from two points in the air. This metaphorical gesture, following Edwards, highlights how the metaphor 'proof of A→B as a path from A to B' is particularly significative for those students.

Further studies highlight how the role of gestures in students' production of argumentation is not simply communicative but can actively support students in the structuration of the argumentation itself. Arzarello and Sabena (2014), for example present an empirical study supporting the fact that "gestures may also play specific roles in providing a logical structure to argumentation" (p. 99). In particular, the authors analyse the gestures produced and used by some 5th grade students dealing with the resolution of a problem. Arzarello and Sabena register that the students involved in the study use a series of repeated deictic gestures (catchments), pointing to different direction of the space around them while structuring an argumentation for solving the problem. More specifically the two researchers observe that these deictic gestures are used by students to refer, metaphorically, to the different cases of the problem considered by the argumentation. In other terms, through these gestures the students use the space around them to shape the argumentation itself. Arzarello and Sabena, thus, conclude that this gestural production contributed to support the students' efforts in producing a mathematical argumentation.

This aspect has been also observed by Krause (2015) who, analysing the gestures of some grade 10 students during their construction of argumentations, observes a repetition of a gesture which links two points in the air, co-timed with the sentences with the form "if…then…". Krause observes that this gesture, in which premise and conclusion of an "if…then…" sentence are linked by a metaphorical path in the air, allows students to represent the argumentation on a more general level in respect to what expressed with their words:

> The gesture itself implicitly realizes a connection of premise and conclusion and with this, supports the verbal connecting action of logical inference by illustrating it in a more general way. The visual access it provides does not refer to a mathematical object but to the structure of reasoning on a meta-level. […] In this case, the metaphorical character of the gesture is completely detached from the concrete content of the task. The structuring-gesture provides an implicit support by not enriching the utterance semantically, but on a meta-level. **Using a gesture this way may support the collective act of reasoning as it indicates the logical structure within an argument and may help to keep track. It makes traceable how the argument was organized as logical inference.** (p.1432, bold added by me).

Sabena (2018) further investigates this theme. The researcher, analysing with the lens of the Semiotic Bundle the argumentations of some primary school students, highlights that "gestures may contribute to carrying out argumentations that depart from empirical stances and shift to a hypothetical plane in which generality is addressed" (p. 554). In particular, Sabena registers how the contraction of two or more gestures into a unique one (for this reason called by her "blending or condensing gestures", p. 555) and the use of the gestures space in a metaphorical sense, is "supporting the students in structuring the entire argument at a global level" (p.556).

6.4 THE USE OF THE SEMIOTIC BUNDLE WITHIN THIS STUDY

As described the analysis of gestures and, more generally, of multimodal signs produced and used by students opened a new research direction for what concerns the study of processes involved in students' construction of argumentations and proofs. In particular, as registered by the studies presented above, recent research on this point highlighted that the use and production of those signs seem to actively support students in structuring their argumentation or in constructing proofs. In the study presented on this thesis I will investigate this research theme with a specific focus on recursive argumentations and on proofs by MI.

To conclude, thus, within the just presented theoretical perspective, this study will have to following research objectives:

- To register traces of the generation of recursive argumentations in students' (multimodal) semiotic production when involved in the resolution of problems.

- To investigate the role of students' production and use of multimodal signs in their construction of a proof by MI.

These two points will be further expanded in the chapter 7 of this thesis, where the research questions will be formulated. Moreover, the construct of the Semiotic Bundle will be also used as a lens trough which to observe elements proper of other theoretical perspective presented in this conceptual framework.

7 Summarising the Conceptual Framework and Research Questions

In the previous chapters I presented the different theoretical perspectives that compose the conceptual framework of this study. In this chapter I present an overview on the whole framework clarifying the specific mutual position of each perspective within the framework. After this, the research questions of this thesis will be presented, and their formulation will be discussed.

7.1 Overview of the Conceptual Framework

The conceptual framework was composed by some different theoretical perspectives.

The first adopted standpoint was a historical-epistemological one, to which chapter 2, preliminary to the others, was dedicated. This perspective, although it is not a “proper” theoretical perspective in mathematics education, is still a part of the conceptual framework since, as presented, the results obtained from the historical-epistemological analysis provided some elements with which the rest of the conceptual framework have been structured.

The other adopted theoretical perspectives are instead common perspectives in mathematics education and their use in this study has the purpose of conducting a cognitive-didactical analysis of MI. The starting point of this analysis was to clarify the position of this study within the research theme which traditionally goes under the name of ‘Argumentation and Proof’. In relation to this point, the theoretical perspective that I adopted is the one emerging from the studies on the construct of Cognitive Unity (starting from Boero et al. 1996, Garuti et al. 1998). Within this perspective, a great relevance is given to the argumentation. It is central both as a research theme, with the assumption that an analysis of the argumentations produced by students could provide element for interpreting their effective processes or difficulties in relation to mathematical proofs, and as didactical tool, suggesting that activities oriented to the students’ production of argumentations could support the teaching and learning of mathematical proofs. Aligning with this perspective, also in this study a central role will be given to those argumentations produced by students that are related to MI and not only to the classic proofs by MI. Starting from this framework, thus, I proposed a model to describe the structure of a theorem with a proof by MI using the construct of mathematical theorem as a triplet. Moreover, an operative definition of recursive argumentation has been given. To summarise, thus, this initial section represented the base on which the rest of the framework is articulated. Within it, I made explicit, in a certain sense, the theoretical position of this study regarding the relationship between argumentations and mathematical proofs and I clarified to what extent some key terms will be used in this thesis (‘proof’, ‘argumentation’, ‘theorem’, ‘meta-theory’, ‘Recursive argumentation’, etc.).

Starting from these premises, three different theoretical perspectives have been presented: The APOS Theory, the theoretical framework based on Fischbein’s studies on intuitions, and a multimodal semiotic perspective with the construct of the Semiotic Bundle. These three perspectives compose the framework each by providing a different and independent viewpoint to address the research problem. To use the terms proper of the *networking of theories* (Prediger et al., 2008) they are *combined* in the conceptual framework, where ‘combination’ stands for a simple juxtaposition of different theoretical perspective used one independently to the others. As I will present, one or more research questions will involve each of these three perspectives.

Within the conceptual framework there are some points that have been obtained with a tighter form of networking, called *coordination*, which corresponds to the simultaneous use of elements proper of different theoretical perspectives. In particular, the points of coordination in the presented framework are the following:

1) The theoretical considerations emerging from the adopted perspective on Argumentation and Proof have been taken into consideration also within the other theoretical frameworks. More specifically, these include the assumption on the relevance of focusing on argumentative processes, the operative definitions of argumentation and mathematical proof, the one of recursive argumentation, and the model of theorem with a proof by MI as a triplet. These are theoretical premises for all the study. Let us observe that these premises can be coherently coordinated with the other adopted theoretical perspectives. First of all, for what concern the multimodal semiotic, I have described that some recent studies showed how this perspective have been effectively used to observe and analyse the students' processes of argumentation and proof construction Arzarello & Sabena, 2014; Sabena, 2018). Analogously, Fischbein's framework on intuitions has been used in other studies in coordination with the construct of the Cognitive Unity: focusing on students' construction of argumentations and proofs to investigate their intuitive acceptance of proofs by contradiction or by contraposition (Antonini, 2004; Antonini 2019). Finally, within the APOS theory, although there is not an explicit stance in relation to the dialectic between argumentation and proof, it is possible to register that a focus on argumentative processes was given, at least in the studies on MI. For instance, in the interviews that Dubinsky and Lewin (1986) used to construct the original GD of MI, they asked students to explain why MI 'works'. The Explain Induction process that the authors register to be interiorized in some students, when described, corresponds specifically to a recursive argumentation supporting the validity of MI.

2) The historical-epistemological analysis, beyond providing an independent contribution to this study, highlighted elements which have been successively used in the development of the conceptual framework. First of all, the historical-epistemological analysis allowed me to expand and deepen the GD of MI within the APOS theory. The possibility of using such an analysis in this sense is openly expressed within the APOS theory:

   > The design of a genetic decomposition can also be based on the historical development of the concept. A study of the historical development of a concept may point to mental constructions that individuals might make. (Arnon et al., 2014, p.34).

   This possibility has been already used in some other APOS studies. For instance, Dubinsky et al. (2005a; 2005b) start from the analysis of some elements taken from the history of mathematics to propose a model, in APOS terms, of the cognitive conceptualization of potential or actual infinity in a subject.

   Secondly, the historical analysis of the genesis of MI allowed to highlight a deep epistemological connection between proving by MI and what I have called the recursive argumentation. In particular, the definition of recursive argumentation took into account those argumentations, coming from the history of mathematics, which have been considered by several scholars as traces of MI. Such use of a historical-epistemological analysis aligns with the theoretical viewpoint that I adopted in relation to the theme of Argumentation and Proof. Mariotti et al. (1997), for example, developed one of the first studies on the Cognitive Unity and mathematical theorem as a triplet starting also from a historical-epistemological analysis of the relationship between proof and argumentation. Similarly, Mariotti (2006) took into consideration historical elements in her epistemological analysis of the concept of 'mathematical proof' and, from this, she developed her study on 'Proof and Proving in Mathematics Education'.

3) The model of theorem with a proof by MI emerging from the construct of mathematical theorem as a triplet have been in coordination with the distinction of three levels of intuitive acceptance

of a theorem by Fischbein (1982). This allowed me to take into consideration different aspects of the intuitive acceptance of a proof by MI, in particular not only the aspects related to the particular general statement that it involves, but also some meta-theoretical aspects, as the intuitive acceptance of the meta-theorem which allows to conclude the validity of a general statement from the validity of base and inductive step.

In conclusion, thus, the conceptual framework of this study can be briefly summarised as follows: it is constructed starting from a historical-epistemological analysis of the genesis of the proof by MI and from a series of theoretical and general considerations on the relationship between Argumentation and Proof. Starting from these premises, three different theoretical perspectives have been considered, the APOS theory, Fischbein's framework on intuitions, and a multimodal semiotic perspective, which have been re-read so to be in line with the research objectives of this study. As I shall present in the next paragraph, from each of these perspective one or more research questions have been developed. The figure 7.1 of below summarises the just described structure of the conceptual framework, together with the respective positions in the framework of the different research questions presented in the next section.

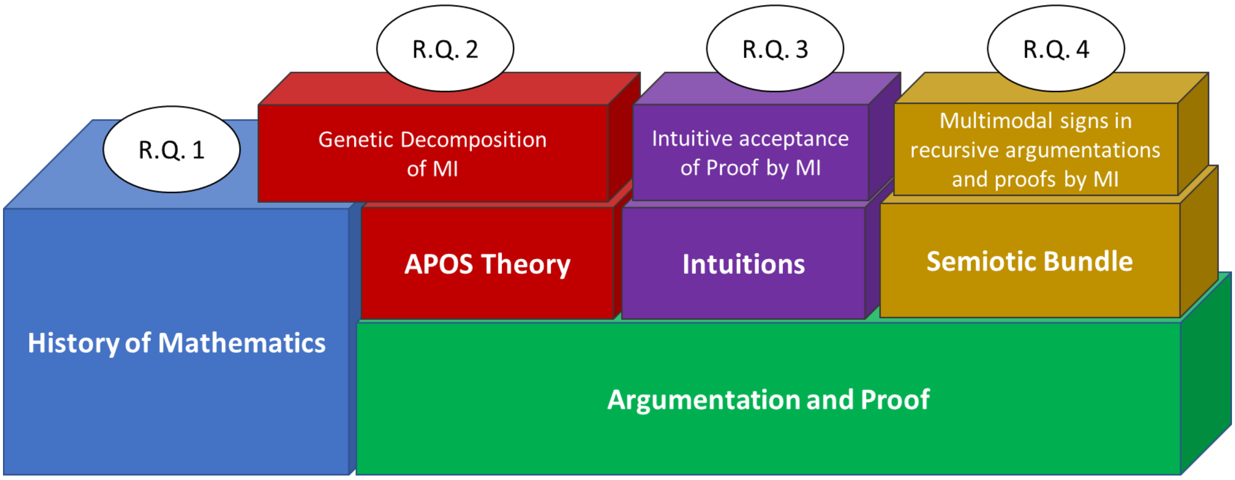


Figure 7.1. *Structure of the conceptual framework of this study. Different colours stand for different theoretical perspectives. The position of the several blocks alludes to the mutual position of the different perspectives within the framework. The research questions are positioned above the respective perspective in which they are*

7.2 Research Questions

In this section I present the research question of the study, each formulated within the specific theoretical perspectives composing the conceptual framework.

First R.Q. – Investigating traces of Induction in History of Mathematics

The first research question of this study has been already formulated and investigated in Chapter 2 in which I presented a historical-epistemological analysis of the genesis of the proof by MI. Specifically, I investigated the following research question:

R.Q. 1 - What mathematical aspects characterised the historical genesis of the proof by MI? In particular, what characteristics and turning points emerge from the traces of proofs by MI that the historiographic research has identified?

An answer to this question was presented in the conclusions of the chapter 2 (section 2.9). For the sake of clarity of this thesis, it was necessary to anticipate this question and the successive analysis because it contributed to the development of the framework, providing elements with which it has been enriched.

Second R.Q. – Investigating the Explain Induction Process

In the Genetic Decomposition of MI presented within the APOS theory a key aspect for the construction of the MI schema is the interiorization of the Explain Induction process. Through this process a subject can construct a justification of why, from the fact that $P(n_0)$ is true and the validity of the inductive step, it necessarily follows that $P(n)$ is true for every $n \geq n_0$. The historical-epistemological analysis suggested that a sufficiently reach schema of MI should also include the Explain Induction process in indirect form, through which a subject can justify why, starting from the validity of the inductive step and the fact that $P(n^*)$ is false, it necessarily follows that $P(n)$ is false for every $n \leq n^*$.

The research objective emerging from this analysis, thus, is to investigate the Explain induction process both in direct and in indirect form, scrutinizing to what extent this process is constructed by students who have some experience with proofs by MI and by students who have not. Note that in this point the term 'constructed' is used in APOS terms. In particular, within this perspective, we talk about 'process construction' both in relation to its construction starting from an action (i.e., interiorization) or from other processes (i.e., coordination, reversal), and in relation to a successive reconstruction of the process by the subject in a different situation. As Dubinsky (1991) states, referring in general to a Schema: "in the context of this theory, it is never clear (nor can it be) whether one is talking about a schema that is present or one that is being (re-)constructed" (p.112).

The specific research focus on the Explain induction process is motivated by two aspects. First of all, it is central in the GD of MI, both in the one presented in literature and in the one that I proposed in section 4.4. The interiorization of this process, in fact, is seen as a crucial step in the construction of the Schema of MI. In particular the object 'Proof by MI', following the GD of MI, is cognitively constructed by a subject exactly as the encapsulation of the Explain induction process. Therefore, to investigate students' interiorization of this process and on possible difficulties related to that could provide elements with which to interpret students' encapsulation of the object 'Proof by MI' and, more generally, their construction of the whole schema 'proof by MI'. Secondly, the choice of investigating this specific process of the GD of MI aligns with the perspective on Argumentation and Proof from which the conceptual framework was developed. Following this perspective, I observed that a key object of research of this study is represented by student's construction of recursive argumentations related to MI. Among those argumentations, there is the classic non formal justification for the validity of a proof by MI through ''an infinite number of syllogisms […], in a cascade" to quote Poincaré (1905, pp. 9-10). The APOS counterpart of this argumentation is precisely the description of the explain induction process in direct form. Similarly, the classic argumentation constructed on the principle of infinite descent, as the one described in Fermat's letter (see section 2.8), corresponds to what, in APOS terms, was described as the Explain induction process in indirect form.

In conclusion, I will investigate the following research question:

R.Q.2 – To what extent do students construct the Explain Induction Process, both direct and indirect form?

Third R.Q. – Investigating the intuitive acceptance of MI

The research objective within the framework of institutions is to investigate students' intuitive acceptance of the proof by MI. I observed that, using the model of theorem with a proof by MI as a triplet, the intuitive acceptance of a proof by MI can be reread in terms of the intuitive acceptance of the three theorems (S1,P1,T), (S2,P2,T), (MS,MP,MT) which constitute a proof by MI. In particular, these theorems need to be intuitively accepted by a student both as theorem whose validity is independent from the others and as a system of theorems, each necessary for the proof P of the general statement S.

To summarise, thus, I am interested in scrutinizing the students' intuitive acceptance of this system of theorems and on the eventual presence of elements which could influence this intuitive acceptance. In conclusion the involved research question will be the following:

R.Q.3 – To what extent a proof by MI, intended as a system of theorems (S1,P1,T), (S2,P2,T) and ((S1$\wedge$S2)$\rightarrow$S,MP,MT), is intuitively accepted by students? Moreover, which aspects can influence its intuitive acceptance?

Fourth R.Q. – Investigating multimodal signs in argumentation and proving by MI

In the description of the framework, I observed how recent research in mathematics education highlighted the effectiveness of using a multimodal semiotic perspective to investigate students' processes related to the construction of proofs and argumentations. In this study I am interested in deepening the research on this aspect, specifically by focusing on recursive argumentations and proofs by MI. In particular, the first objective is to investigate the presence of specific signs which characterise the process of generation of a recursive argumentation in students involved in a problems resolution activity. With 'process of generation' of an argumentation I mean both its emergence during the exploration of the problem, and its successive development to support a solution for the problem. The second research objective is to analyse the role that these signs can have for students in structuring a proof by MI. The theoretical standpoint here, supported by the study on the Cognitive Unity, is that there could be a parallelism (also in terms of the signs produced and used) in the process of construction of a proof and the previously constructed argumentation. Starting from this, therefore, the objective is to analyse the contribution that those signs can provide to student for the construction of a proof by MI.

To summarize, thus, I will investigate the following research question:

R.Q. 4 – Is it possible to identify crucial signs involved in the generation a recursive argumentation in students' exploration of a problem? What contribution can the production and use of these signs provide to students' construction of a proof by MI?

8 METHODS

8.1 PREMISES – HISTORY OF THIS STUDY

During the summer of 2019 a first research design for gathering the data for this study was planned. This a-priori project involved some successive moments to collect a series of qualitative data:

(1) A preliminary case study involving expert students in mathematics (PhD or master's students) dealing with the resolution of some problems and with the possible construction of proofs by MI to formalise their solutions.
(2) After an analysis of the case study (1), a new data collection, this time involving less expert students (1st year undergraduate students in scientific disciplines) dealing with the resolution of problems obtained as refinement of the ones assigned to expert students.
(3) Lastly, considering what emerging from (1) and (2) and in line with the presented framework, a teaching path in one (or more) secondary school classes with the aim of introducing students to recursive argumentations and to proofs by MI.

Unfortunately, after the case study (1), whilst I was analysing the collected data, the Covid-19 pandemic broke out. When it has been clear that the situation would not have come back to normality in short times, I decided to modify the research design previously planned.

In the summer 2020 the sanitary conditions in Italy still did not permit to carry out in presence interviews in an appropriate research environment: the access to offices and university rooms, as well as schools, was interdicted or strictly limited to few exceptions. Therefore, I sought for an alternative solution to collect some data that could provide a suitable contribution in relation to the research problem. With this aim, I chose to design an online survey to distribute among university students of different grades and fields of study. I conducted two first case studies with some second- and third-year students in mathematics and in engineering. Then, starting from the analyses of these first two case studies a further survey was structured and distributed in autumn 2020 to students of different grades and degrees from several Italian universities. The fact that lessons at the university were online in that period and that, after several months of online teaching, courses were developed in on-line platforms, allowed a rapid and widespread distribution of the survey to numerous students: in the ten days that the survey page was open, we collected 306 answers from 12 different universities.

Parallelly to this data collection, in autumn 2020 it was possible to carry out a few in person interviews with first-years university students, in line with the point (2) of the a-priori plan. Unfortunately, after a few weeks the sanitary situation worsened again, and the interviews were interrupted. It was still possible to collect some data that have been analysed in line with the research objectives.

The presentation of these different moments of data collection in the next part of this thesis shall not respect the chronological order in which they took place, but it shall develop, instead, by taking into account the methodological uniformity of each data collection. The order in which I will present them now is the same in which they will be presented in the next chapters:

- Firstly, I will present the online survey involving several university students of different grades and attending different degrees.
- Secondly, I will present the in-presence interviews with PhD and masters' students dealing with the resolution of some problems related to MI, together with the successive interviews with first-years university students.

In this chapter I will present in detail the methodological choices of the two studies (the survey and the interviews). In particular, I will analyse the proposed tasks, describe the sample of subjects involved, the methods of data collection and data analysis. In the next chapters of this thesis the collected data will be presented and discussed

8.2 OVERVIEW ON THE EMPIRICAL STUDIES

Before describing the details of the two empirical studies, we shall see how they are situated in relation to the research questions and to the theoretical perspectives considered in the framework. Although the different data collections were structured to mainly investigate different research questions, their answers at the end of the thesis, as we will see, will take into account all the collected data. For this reason, the distinction presented below is to be considered as an 'a-priori' organization of the empirical studies.

8.2.1 Online survey

The survey was designed with the aim of investigating on the research questions RQ2 and RQ3.

The RQ2 addresses the students' construction of the Explain induction process, both in direct and in indirect form. As described in 4.4, with this process a subject can justify why, if $P(n_0)$ is true and the inductive step is valid, on can conclude that $\forall n \geq n_0$ $P(n)$ is true (direct form) or why, if $P(n^*)$ is false and the inductive step is valid, one can conclude that $\forall n \leq n^*$ $P(n)$ is false (indirect form). Some tasks of the survey, thus, aimed to investigate this process and eventual problematic aspects related to it. Specifically, I chose to create tasks in which firstly it was stated, respectively, the truth or falsity of a property for a given number and the validity of a statement corresponding to the inductive step for the property and then subjects were asked to deduce from the given assumptions the truth of or the falsity of the property for a series of numbers and to explain their reasoning. The aim was to investigate the construction of the Explain induction process (direct and indirect) both in students who had already encountered MI during their studies and in students who had not. Moreover, I decided to present the statements of the tasks with a verbal formulation to not introduce difficulties related to the mathematical symbolism.

The RQ3 addresses the students' intuitive acceptance of the proof by MI, instead. From the model of theorem with a proof by MI as a triplet, I observed how its intuitive acceptance involves different aspects: the acceptance of the meta-theorem (from base, S1, and inductive step, S2, one can conclude the validity of the general statement S) and the acceptance of base case and inductive step as independent theorems. For what concerns the intuitive acceptance of the meta-theorem we can observed that the tasks above anticipated involve this aspect. In particular, in these tasks the validity of the statements S1 and S2 was given, where S1 was modified involving numbers different from 0 or 1. The assumption is that an intuitive acceptance of the meta-theorem allows students to re-construct the theorem in these particular situations and correctly answer the questions posed. The deviations from the correct answers and the written justifications provided could therefore give me indications related to students' intuitive acceptance of the meta-theorem. For what concerns the intuitive acceptance of base case and inductive step independent theorems, I introduced some tasks in which it is affirmed that base case and inductive step have different values of validity (inductive step valid with a false base case or a not present base case) and questions involving possible consequences of these situations are posed. Finally, to further investigate the intuitive acceptance of the proof by MI some open questions were posed to students at the end of the survey, asking to describe, informally, why MI 'works' for them and to describe, if there were, aspects of MI perceived by them as unclear or not convincing.

8.2.2 Interviews with students involved in a problem solving activity related to MI.

A further research objective of this thesis is to investigate the processes of construction of recursive argumentations and proofs by induction in students involved in activities of problem resolution. The perspective adopted to address this research objective is the multimodal semiotic perspective of the semiotic bundle. The research question RQ4 was formulated within this perspective. This research question firstly addresses the identification of crucial signs involved in the generation of a recursive argumentation in students' exploration of a problem. For this reason, since the observation of the student's exploration of the problem is essential to investigate the question, the problems were chosen and formulated in a way that an initial explorative phase would have been necessary. Secondly, the research question addresses the contribution of these multimodal signs in the process of construction of a proof by MI. For taking into consideration this research focus, thus, students were asked to provide a solution to the problem where with 'solution' I intended both an answer to the question involved in the problem and the proof supporting it.

In the interviews, after the part of problem resolution I decided to include some 'meta-level' questions, in which the object was the proof by MI itself and, in particular, the justification of its validity. This part of the interview aimed to consider the research question on the Explain induction process (RQ2) and on the intuitive acceptance of MI (RQ3).

The Table 8.1 of below summarises the just presented overview on the two empirical studies.

Empirical study	**Theoretical Perspectives mainly considered**	**RQs mainly involved**	**Types of tasks**
Online Survey	APOS Theory, Intuitions	RQ2 – RQ3	Multiple choice and open questions 'Meta' questions on MI
Interviews	Multimodal Semiotic	RQ4	Mathematical problems
	APOS Theory, Intuitions	RQ2 – RQ3	'Meta' questions on MI

Table 8.1. *Overview on the two conducted empirical studies, with indication of the theoretical perspectives framing them, of the RQs mainly involved, and of the types of tasks used.*

8.3 First empirical study – Online survey

The first empirical study that I present involved an online survey distributed among university students of different grade (both undergraduate and master's student) attending different academic courses from some Italian universities. The survey was designed after two precedent case studies with a limited sample of students. In this section I will directly present the definitive version of the survey from which the data later discussed have been collected. The survey contains a series of closed and open questions. The presence of some multiple choice questions partially moves away from the qualitative approach adopted in the rest of this thesis. However, despite this aspect, the survey allowed me to investigate some elements of the research objectives. In particular, as I will present in the next chapter, the results of the survey will highlight some problematic aspects for students in relation to MI. The survey offers a first picture of these problematics.

8.3.1 Content of the survey

The survey is divided in some separated sections which I will present here separately in the same order and written exactly as they were in the original survey, with the only exception of the English translation. For the sake of clarity, the comments on the questions involved will be presented at the end of each section.

Introduction

The survey starts with the following message which informs the participants on the contents of the survey and on the use of the collected data.

> *Dear student,*
> *The following survey is part of a study in Mathematics Education conducted by researchers of the University of Firenze and the University of Pavia.*
>
> *The survey is completely anonymous and in none of the phasis of the data gathering process it is possible to attribute the answers to a specific person. The collected data are exclusively the ones explicitly requested in the questions with the only exception of the date and time of the survey completion.*
>
> *You will be asked to invent a nickname which will be only used to collect the data and will not be used to get your identity.*
>
> *The survey is composed by some tasks containing several questions; please do not pass to the following task until you have answered all the questions of a task. The tasks should be done in order, and we ask you to not modify the answers once given.*
>
> *Answer the questions with a cross on the answer you consider correct, without leaving back any questions. We are also interested to know how sure you are of your answer and how much effort it took you. Please answer sincerely!*
>
> *The data collected in this study will be treated in respect to the 'Regole deontologiche per trattamenti a fini statistici o di ricerca scientifica pubblicate ai sensi dell'art. 20, comma 4, del d.lgs. 10 agosto 2018, n. 101 - 19 dicembre 2018 e ss.'*
> *In any case the information elaborated in the study will be used in publications in a totally anonymous form.*
>
> *The completion of the survey takes approximately 20 minutes.*
> *Thank you!*

After this message, participants are asked to insert a nickname (a pseudonym) which does not contain indication of their real names, to write the university they are attending, their academic course with the year and selecting between undergraduate and master's.

Comments on introduction

In the introduction message students are warned that they will be asked to answer the question respecting the order in which they are presented, to not go back to a previous section, to not change an answer once given, and to not leave any answer blank. These requests were prompted by the following motivations.

- Firstly, the request to answer the question respecting the order of questions and sections was made to uniform, for how it was possible, the possible influences of previous questions. In fact, I cannot exclude that the order in which the questions were posed had influenced the given questions, thus, to have homogeneous data from this point of view, it seems reasonable to select a fixed order in the questions and sections for each participant. Moreover, as I will present, the order of the questions within a section and the order of the sections was made to specifically designed to take into account some specific aspects of the research goal.
- The request of not changing an answer once given, and the one of not going back to a previous section once completed, was prompted by the fact that I am interested to register the first answer given to a question, the answer that was considered intuitively correct once the question was posed. It is possible that, moving on in the survey, a subject realises that s/he wants to change a given answer however, if I would have let this happen, since the survey was not on paper, I would have lost the trace of the first answer. At the end of the survey an open question was posed asking the participants to say if there were answers that they would have changed if they could.
- Lastly, the request of not leaving blank any question was made since I expected that there could be questions for which a participant would have been strongly undecided. By not having the possibility of leaving blank an answer, thus, the participant would have obliged to take a decision that, even if not considered as surely correct, would provide me some indications on the answer considered by the student the most likely to be correct.

As written in the message, I was interested in registering the 'level' of certainty of each answer and the perceived effort to answer. After every question, with the exception of the ones of the sections 'Introduction' and 'Conclusive questions', participants were asked to select 'very little [0]' / 'little [1]' / 'enough [2]' / 'a lot [3]' for the questions "How much are you sure about your answer?" and "How much effort did it take you to answer?". For this point, I took inspiration form a study by Fischbein (1981) in which he suggests to 'measure' in a qualitatively way the intuitiveness of an answer for a subject by considering the perceived confidence with which the subject answered (i.e. how certain the subject feels about the answer) and the perceived obviousness of the question (i.e. how much effort to answer was perceived by the subject). These questions were inserted to eventually register if some questions of the survey were considered extremely problematic by the students from an intuitive point of view (in terms of perceived certainty and effort to answer). I can anticipate that, at least for the registered data, none of the question of the survey obtained, on average in the considered groups of students, a level of perceived certainty lower than 1.5 or a level of perceived effort higher than 1.5 (both the 50% level in the 0-3 scale of above). Therefore, at least for the collected data, it was not possible to register the presence of questions which were stated to be problematic (in terms of perceived certainty and effort to answer) by the average of the students. This, of course, does not mean that none of the question was problematic, but that this aspect does not emerge from what stated by the students themselves. For this reason, the results for these questions will not be presented later.

The request of adding a nickname (or pseudonym) was not really necessary for data collection (a random number was in fact assigned to each participant at the moment of the submission of the answers) however was made to make further stress the fact that each survey is anonymous. The information on the academic course and year of course attended by each participant were needed, as I will present, to divide the participants in some groups considering the number of courses in mathematics at university level they have followed before taking the survey. The indication of the university they were from was asked to keep trace of the diffusion of the survey in the various university to which it was sent.

Task A – Aarney's property

The first task of the survey, hereinafter 'Task A', starts with the following text:

> A mathematician named Aarney noticed that the number 103 has a particular numerical property that he called Aarney's property.
>
> A colleague of his, intrigued by this property, proved that:
> *Taken any natural number, if for this number Aarney's property is true, then Aarney's property is also true for its consecutive natural number.*
>
> What can we deduce from this? For each of the following statements indicate with a cross the answer that you consider correct. The option "We cannot know", where present, means that we do not have enough information to answer.

The following questions are posed:

1) Indicate the correct answer:
   - ☐ Aarney's property is true for all the natural numbers.
   - ☐ Aarney's property is false for all the natural numbers.
   - ☐ There exist some numbers for which Aarney's property is true and there exist some numbers for which Aarney's property is false.
   - ☐ We cannot know (We do not have enough information to know if the previous statements are true or false.)

   Why?
   [Space for open answer]

2) For the number 102, Aarney's property is:
   - ☐ True
   - ☐ False
   - ☐ We cannot know

3) For the number 104, Aarney's property is:
   - ☐ True
   - ☐ False
   - ☐ We cannot know

4) For the number 101, Aarney's property is:
   - ☐ True
   - ☐ False
   - ☐ We cannot know

5) For the number 105, Aarney's property is:
   - ☐ True
   - ☐ False
   - ☐ We cannot know

6) For the number 94, Aarney's property is:
   - ☐ True
   - ☐ False
   - ☐ We cannot know

7) For the number 112, Aarney’s property is:
☐ True
☐ False
☐ We cannot know

8) For the number 4, Aarney’s property is:
☐ True
☐ False
☐ We cannot know

9) For the number 542384, Aarney’s property is:
☐ True
☐ False
☐ We cannot know

10) For the number 102, Aarney’s property is:
☐ True
☐ False
☐ We cannot know

11) For the number beginning with 1875249 followed by a million zeros and ending with 325593, Aarney’s property is:
☐ True
☐ False
☐ We cannot know

12) Explain your reasoning in answering the previous questions:
[Space for open answer]

Task B – Bertrand’s Property

As in the Task A, the second task of the survey, hereinafter ‘Task B’, starts with the following text:

A mathematician named Bertrand conjectured that the number 37 has a particular numerical property that he called Bertrand’s property; however, after having checked it, he noticed that instead Bertrand’s property is FALSE for the number 37.

A colleague of his, intrigued by this property, proved that:
Taken any natural number, if for this number Bertrand’s property is true, then Bertrand’s property is also true for its consecutive natural number.

What can we deduce from this? For each of the following statements indicate with a cross the answer that you consider correct. The option “We cannot know”, where present, means that we do not have enough information to answer.

The following questions are posed:

1) Indicate the correct answer:
   ☐ Bertrand’s property is true for all the natural numbers.
   ☐ Bertrand’s property is false for all the natural numbers.
   ☐ There exist some numbers for which Bertrand’s property is true and there exist some numbers for which Bertrand’s property is false.
   ☐ We cannot know.
   Why?
   [Space for open answer]

2) For the number 36, Bertrand’s property is:
   ☐ True
   ☐ False
   ☐ We cannot know

3) For the number 38, Bertrand’s property is:
   ☐ True
   ☐ False
   ☐ We cannot know

4) For the number 35, Bertrand’s property is:
   ☐ True
   ☐ False
   ☐ We cannot know

5) For the number 39, Bertrand’s property is:
   ☐ True
   ☐ False
   ☐ We cannot know

6) For the number 28, Bertrand’s property is:
   ☐ True
   ☐ False
   ☐ We cannot know

7) For the number 46, Bertrand’s property is:
   ☐ True
   ☐ False
   ☐ We cannot know

8) For the number 5, Bertrand’s property is:
   ☐ True
   ☐ False
   ☐ We cannot know

9) For the number 683747, Bertrand’s property is:
   ☐ True
   ☐ False
   ☐ We cannot know

10) For the number beginning with 34749 followed by a million zeros and ending with 561193, Bertrand's property is:
☐ True
☐ False
☐ We cannot know

11) Explain your reasoning in answering the previous questions:
[Space for open answer]

12) Since it is known that Bertrand's property is false for the number 37, Bertrand's colleague has **SURELY** mistaken his proof, because it is **NOT** true that taken any natural number (for instance, **the number 37**), if for this number Bertrand's property is true, then Bertrand's property is also true for its consecutive natural number.
☐ I agree ☐ I do not agree

13) Since it is known that Bertrand's property is false for the number 37, Bertrand's colleague has **SURELY** mistaken his proof, because it is **NOT** true that taken any natural number (for instance, **the number 36**), if for this number Bertrand's property is true, then Bertrand's property is also true for its consecutive natural number.
☐ I agree ☐ I do not agree

Task C – Coleman's Property

The third task, hereinafter 'Task C', starts with the following text:

A mathematician named Coleman noticed that the number 75 has a particular numerical property that he called Coleman's property; he then noticed that Coleman's property is FALSE for the number 45.

A colleague of his, intrigued by this property, proved that:
Taken any natural number, if for this number Coleman's property is true, then Coleman's property is also true for its consecutive natural number.

What can we deduce from this? For each of the following statements indicate with a cross the answer that you consider correct. The option "We cannot know", where present, means that we do not have enough information to answer.

The following questions are posed:

1) Indicate the correct answer:
☐ Coleman's property, without considering the numbers 45 and 75, is true for all the natural numbers.
☐ Coleman's property, without considering the numbers 45 and 75, is false for all the natural numbers.
☐ Besides the numbers 45 and 75, there exist some numbers for which Colemans's property is true and there exist some numbers for which Colemans's property is false.
☐ We cannot know.

Why?
[Space for open answer]

2) For the number 44, Coleman’s property is:
☐ True
☐ False
☐ We cannot know

3) For the number 46, Coleman’s property is:
☐ True
☐ False
☐ We cannot know

4) For the number 74, Coleman’s property is:
☐ True
☐ False
☐ We cannot know

5) For the number 76, Coleman’s property is:
☐ True
☐ False
☐ We cannot know

6) For the number 12, Coleman’s property is:
☐ True
☐ False
☐ We cannot know

7) For the number 156380, Coleman’s property is:
☐ True
☐ False
☐ We cannot know

8) For the number 60, Coleman’s property is:
☐ True
☐ False
☐ We cannot know

9) For the number beginning with 34371 followed by a million zeros and ending with 58933, Coleman’s property is:
☐ True
☐ False
☐ We cannot know

10) Explain your reasoning in answering the previous questions:
[Space for open answer]

Comments on Task A, Task B, and Task C.

The situation described in the texts which open the Task A, B, and C can be schematised as follows:

a) In Task A, it is stated that a given (unknown) numerical property A holds for the number 103 and that what we can call 'the inductive step' is valid for A, i.e. that taken any natural number, if A holds for this number then it holds for the consecutive natura number as well. In other terms, if we indicate the property A in terms of the predicate A(n): 'The property A holds for the number n', the text of the task states that A(103) is true and that $\forall n(A(n) \rightarrow A(n+1))$. From these statements, from a logical point of view, we can conclude that A(n) is true$\forall n \geq 103$, whilst we cannot know if A(n) is true or false for the natural numbers lower than 103.

b) In Task B, analogously to the previous one, it is stated that B(37) is false and that $\forall n.(B(n) \rightarrow B(n+1))$. Also in this case, I have expressed the test of the task in terms of the predicate B(n): 'The property B holds for the number n'. From these statements, from a logical point of view, we can conclude that B(n) is false $\forall n \leq 37$, whilst we cannot know if B(n) is true or false for the natural numbers greater than 37.

c) By indicating with C(n) the predicate 'The property C holds for the number n', in Task C it is stated that C(75) is true, C(45) is false, and that $\forall n.(C(n) \rightarrow C(n+1))$. From these statements, from a logical point of view we can deduce that C(n) is true $\forall n \geq 75$ and that C(n) is false $\forall n \leq 45$ C(n), whilst we cannot know if C(n) is true or false for the natural numbers strictly between 45 and 75.

These three tasks were designed to investigate the Explain induction process both in direct form and in indirect form. In particular, in Task A an Explain induction process in direct form could be constructed to conclude that A(n) is true for every n≥103, In Task B an Explain induction process in indirect form could be constructed to conclude that B(n) is false when n≤37, and in Task C both the processes could be constructed to conclude that C(n) is true for every n≥75 and C(n) is false for every n≤45. Let us focus with more detail on the questions posed in these three tasks.

The question (1) of each of the three tasks relates a proposition involving the truth value of the given property on the whole set of natural numbers. For what noticed in a)-c) of above the correct answer is "We cannot know" for Task A and B, and "Besides the numbers 45 and 75, there exist some numbers for which Colemans's property is true and there exist some numbers for which Colemans's property is false" for Task C.

After this question, each task contains a series of questions involving the truth value of the properties A, B, and C for a series of specific numbers. For the sake of brevity, from now on, I will indicate these questions by referring directly to the predicate they involve. For example the question 'A(102)' corresponds to the question 'For the number 102, Aarney's property is'. With this compact notation, the numerical questions of these first three tasks were:

- The numerical questions of Task A were, in this order: A(102), A(104), A(101), A(105), A(94), A(112), A(4), A(5423844) and lastly A(N*), where N* is "The number beginning with 1875249 followed by a million zeros and ending with 325593".
- The numerical questions of Task B were, in this order: B(36), B(38), B(35), B(39), B(28), B(46), B(5), B(683747), and lastly B(M*), where M* is "The number beginning with 34749 followed by a million zeros and ending with 561193".
- The numerical questions of Task C were, in this order: C(44), C(46), C(74), C(76), C(12), C(60), C(156380), and lastly C(L*) where L* is "The number beginning with 34371 followed by a million zeros and ending with 58933".

In Task A and B, the numbers involved in the questions follows a specific pattern. If n_0 is the number for which the text provides a known truth or false value, the questions involve the numbers: $n_0 \pm 1$, $n_0 \pm 2$, $n_0 \pm 9$, a "small" natural number (<10), a "big" natural number ($>10^6$) and a "very big" natural number ($> 10^{10^6}$).

In Task C the same pattern has been used but eliminating the questions on the numbers $n_0 \pm 2$, $n_0 \pm 9$ (where now n_0 can be both 45 and 75) and adding a question on the number 60, where 60 is equidistant from 45 and 75 (the known values of the property C). The Figures 8.1, 8.2, and 8.3 summarise the structure of the numerical questions of these three tasks, with indication of the correct answers.

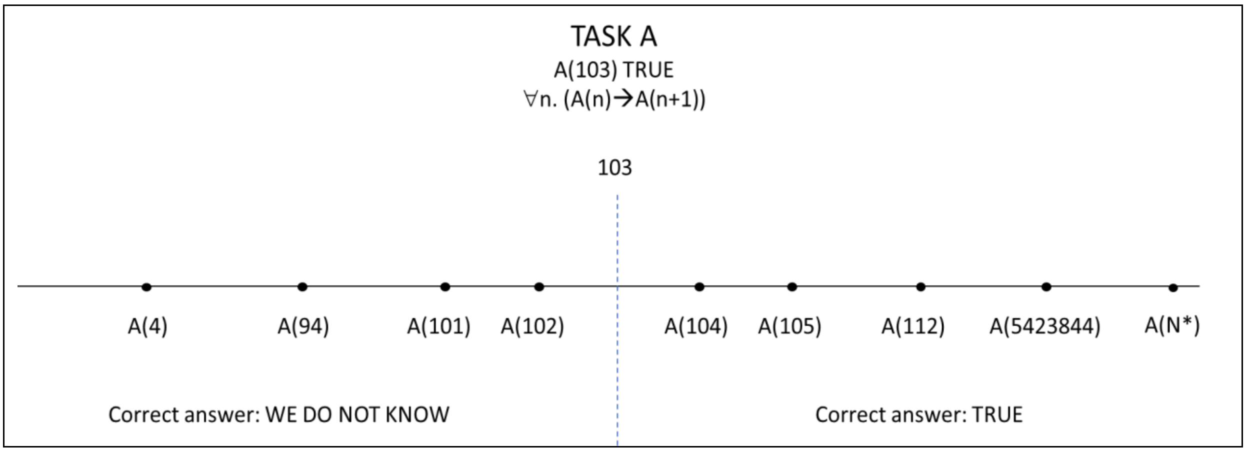


Figure 8.1. *Schema representing the questions on specific numbers for Task A. The order of the questions was A(102), A(104), A(101), A(105), A(94), A(112), A(4), A(5423844), and A(N*)*

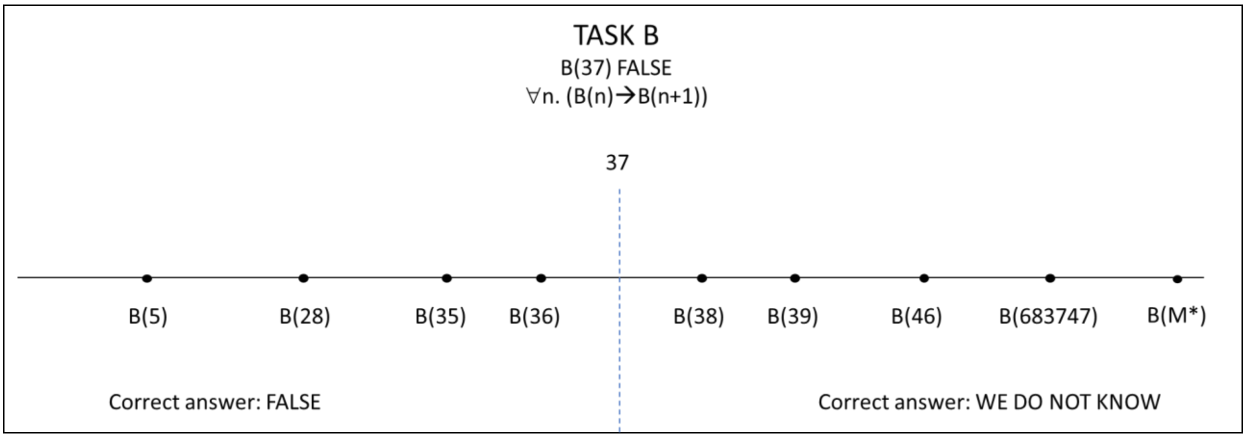


Figure 8.2. *Schema representing the questions on specific numbers for Task B. The order of the questions was B(36), B(38), B(35), B(39), B(28), B(46), B(5), B(683747), and B(M*).*

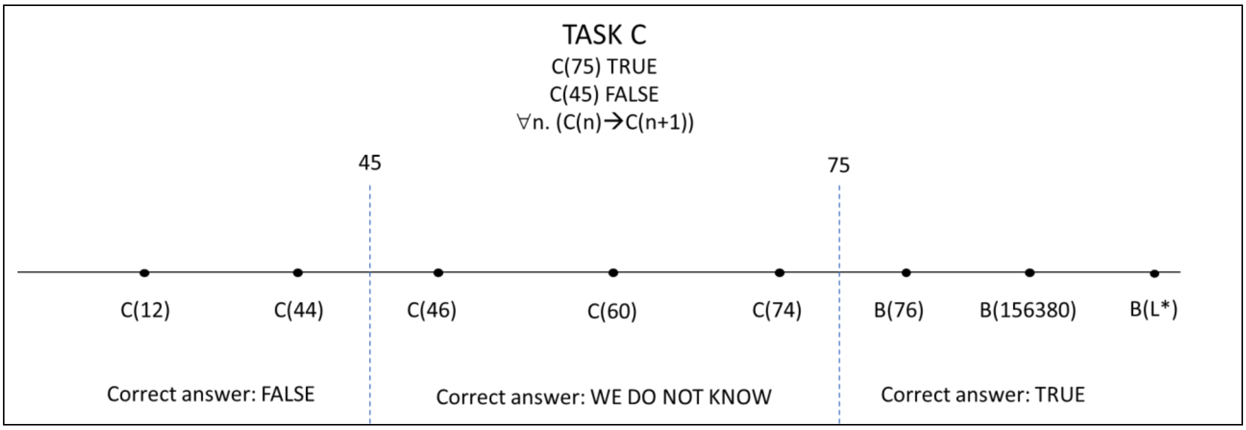


Figure 8.3. *Schema representing the questions on specific numbers for Task C. C(44), C(46), C(74), C(76), C(12), C(60), C(156380), and C(L*).*

Some further comments on the questions are necessary.

The choice of posing questions involving numbers progressively more distant from the number(s) n_0 of each task is coherent with the aim of investigating the Explain induction process (both forms) when the distance from n_0 is progressively increased. More specifically with these questions I was interested in investigating the construction of this process when, starting from the base n_0, it involves numbers progressively more distant from n_0. The choice of alternating the questions on the numbers lower than n_0 and greater than n_0 was made to progressively increase the distance from n_0 to both directions.

Task C was added for some reasons. Firstly, it was added to confirm, eventually, the results of Tasks A and B, respectively for the questions involving numbers $n \geq n_0$ with n_0 a truth value for the property, and the questions involving the numbers $n \leq n_0$ with n_0 a false value for the property. Moreover, supported by the previous case studies made for the survey, I hypothesised that some participants would have answered "True" to all the questions in Task A and "False" to all the questions in Task B. Thus, Task C enabled the posing of questions involving numbers ($45<n<75$) that are, at the same time, greater than a value for which the property is false and lower than a value for which the property is true.

I decided to change the numbers involved in every task to avoid, possibly, that a subject, in a task, would give an answer in analogy or contraposition to the other tasks since the question involved the same number. Moreover, this was decided to present the tasks to participants as independent one to the others.

Finally, the open question posed after the first question and at the end of the numerical questions of each task was inserted to have some qualitative elements to interpret the answers to the multiple choice questions of the tasks.

Two further questions need to be analysed, the questions 12) and 13) of Task B, which are present only in this task. In these ones it is asked to express a binary judgment ('I agree' / 'I do not agree') on the following statements:

[Question 12] "Since it is known that Bertrand's property is false for the number 37, Bertrand's colleague has **SURELY** mistaken his proof, because it is **NOT** true that taken any natural number (for instance, **the number 37**), if for this number Bertrand's property is true, then Bertrand's property is also true for its consecutive natural number."

[Question 13] “Since it is known that Bertrand’s property is false for the number 37, Bertrand’s colleague has **SURELY** mistaken his proof, because it is **NOT** true that taken any natural number (for instance, **the number 36**), if for this number Bertrand’s property is true, then Bertrand’s property is also true for its consecutive natural number.”

The two statements are both invalid: the fact that it exists a number for which Bertrand’s property is false does not imply that the inductive step, proved by Bertrand’s colleague, cannot be valid for every natural number. The two questions, thus, aimed to investigate to what extent the validity of the implication in the inductive step is intuitively accepted, independently from the truth value of antecedent or consequent. In particular the two questions ask to evaluate the validity, respectively, of the implication $B(37) \rightarrow B(38)$ and of the implication $B(36) \rightarrow B(37)$, knowing that B(37) is false. I decided to include both the questions because the validity of the implication could be put under discussion both if the antecedent is false or if the consequent is false. Thus, the aim is to investigate whether both the situations can be problematic to accept the validity of an implication.

Task D – Evelin’s Property

As previously, the fourth task, hereinafter ‘Task D’, starts with the following text:

> A mathematician named Dustin proved that, for a particular numerical property, called Evelin’s Property, for which it is not known for which numbers it is true or false, the following theorem holds:
>
> *Taken any natural number, if for this number Evelin’s property is true, then Evelin’s property is also true for its consecutive natural number.*
>
> By the fact that we do not know any number for which Evelin property is true or false and by the theorem proved by Dustin, indicate whether each of the following statements is CERTAIN, POSSIBLE, IMPOSSIBLE

The following statements to be evaluated are given:

1) Evelin’s property is TRUE for all the natural numbers
   - ☐ Certain
   - ☐ Possible
   - ☐ Impossible

2) Evelin’s property is FALSE for all the natural numbers
   - ☐ Certain
   - ☐ Possible
   - ☐ Impossible

3) Evelin’s property is FALSE for all the even natural numbers and at the same time is TRUE for all the odd natural numbers
   - ☐ Certain
   - ☐ Possible
   - ☐ Impossible

4) Evelin’s property is TRUE only every other couple of natural numbers. That is, for instance, that it is TRUE for the numbers 1 and 2, FALSE for the numbers 3 and 4, TRUE for the numbers 5 and 6, and so on.

☐ Certain
☐ Possible
☐ Impossible

5) Evelin's property is TRUE for the number 40 and at the same time TRUE for the number 41.
☐ Certain
☐ Possible
☐ Impossible

6) Evelin's property is TRUE for the number 40 and at the same time FALSE for the number 41.
☐ Certain
☐ Possible
☐ Impossible

7) Evelin's property is FALSE for the number 40 and at the same time TRUE for the number 41.
☐ Certain
☐ Possible
☐ Impossible

8) Evelin's property is FALSE for the number 40 and at the same time FALSE for the number 41.
☐ Certain
☐ Possible
☐ Impossible

9) Evelin's property is FALSE for the number 23 and at the same time TRUE for the number 49.
☐ Certain
☐ Possible
☐ Impossible

10) Evelin's property is TRUE for the number 12 and at the same time FALSE for the number 54.
☐ Certain
☐ Possible
☐ Impossible

11) Evelin's property is TRUE for small numbers and at the same time FALSE for the number with millions of digits.
☐ Certain
☐ Possible
☐ Impossible

12) Evelin's property is FALSE for small numbers and at the same time TRUE for the number with millions of digits.
☐ Certain
☐ Possible
☐ Impossible

13) Explain your reasoning in answering the previous questions:
[Space for open answer]

Comments on Task D

Analogously to what happened in the previous tasks, in Task D it was stated that for a given numerical property E, a mathematician proved the proposition $\forall n.(E(n) \rightarrow E(n+1))$, corresponding to an inductive step where E(n) is the predicate 'The property E holds for the number n'. However, differently from the other tasks, the text does not provide a natural number n_0 for which it is known if the property E is true or false.

As a consequence, from what stated in the text of this task, we can only conclude that if there exists a natural number n_0 for which the property is true then E(n) would be true for every $n \geq n_0$ and if there exists a natural number n* for which the property is false then E(n) would be false for every $n \leq n^*$.

A series of statements are then presented to which participants are asked to state if they are 'Certain' / 'Possible' / 'Impossible'. The following table (8.2) summarises the given statements, introducing a series of labels that will be used to indicate them, and contains indication of the mathematically correct answer for each of them.

Label	Statement	Mathematically correct answer
T	Evelin's property is TRUE for all the natural numbers.	Possible
F	Evelin's property is FALSE for all the natural numbers.	Possible
FTFT…	Evelin's property is FALSE for all the even natural numbers and at the same time is TRUE for all the odd natural numbers.	Impossible
TTFF…	Evelin's property is TRUE only every other couple of natural numbers. That is, for instance, that it is TRUE for the numbers 1 and 2, FALSE for the numbers 3 and 4, TRUE for the numbers 5 and 6, and so on.	Impossible
40T-41T	Evelin's property is TRUE for the number 40 and at the same time TRUE for the number 41.	Possible
40T-41F	Evelin's property is TRUE for the number 40 and at the same time FALSE for the number 41.	Impossible
40F-41T	Evelin's property is FALSE for the number 40 and at the same time TRUE for the number 41.	Possible
40F-41F	Evelin's property is FALSE for the number 40 and at the same time FALSE for the number 41.	Possible
23F-49T	Evelin's property is FALSE for the number 23 and at the same time TRUE for the number 49.	Possible
12T-54F	Evelin's property is TRUE for the number 12 and at the same time FALSE for the number 54.	Impossible

ST-BF	Evelin's property is TRUE for small numbers and at the same time FALSE for the number with millions of digits.	Impossible
SF-BT	Evelin's property is FALSE for small numbers and at the same time TRUE for the number with millions of digits.	Possible

Table 8.2. *Statements of Task D with the corresponding labels and indication of the mathematically correct*

The statements 'T' and 'F' were inserted to register possible problematics in accepting the universal (i.e. on the whole set $\mathbb{N}$) validity of the inductive step as independent from the universal validity of the general statement E(n). In other terms, the fact that $\forall n \in \mathbb{N}.(E(n) \rightarrow E(n+1))$ is valid does not imply that E(n) is true for every natural number n, and it neither implies that E(n) cannot be false for every natural number n.

The statements 'FTFT…' and 'TTFF…' are analogous and they can be answered by observing that they are in contradiction with the inductive step. However, there is a difference between the two statements since in the second one it is true that *for some* numbers n, $E(n) \rightarrow E(n+1)$ is valid whilst in the first one this is not. This aspect could make the answer to the statement 'TTFF…' more problematic.

The statements '40T-41T', '40T-41F', '40F-41T', and '40F-41F' involve the four different truth or false values for E(40) and E(41), while the last four statements ('23F-49T', '12T-54F', 'ST-BF', 'SF-BT') involve alternated truth or false values for $E(n_0)$ and $E(n_1)$ where n_0 and n_1 are natural numbers such as $n_0<n_1$. These last eight questions are analogous from a logical point of view since they can be answered by observing that, as a consequence of the inductive step, it cannot exist a number for which the property E is false which is greater than a number for which the property E is true. However, they involve numbers whose distance is different, from consecutive numbers (40 and 41) to indefinitely distant numbers (small numbers and numbers with millions of digits). This choice aligns with the previous tasks and aims to register possible differences in the answers when the distance between the involved numbers increases.

Finally, as previously, at the end of the task, I posed an open question asking participants to justify what they selected, aiming to register further elements with which to interpret the previous answers.

Overall comments on Task A-B-C-D

Before presenting the concluding questions of the survey, two general remarks on the previous four tasks are necessary. Firstly, we should notice that the questions in the tasks have a particular logical structure that could be problematic for students. Formally, in most of the questions It is asked to decide if a given statement S is a logical consequence of some given premises. In order to do so it is asked if the statement S is "true" or not, but this "not" is further divided in "false" or "we cannot know". This is done to cover all the different possibilities: as a consequence of the premises the statement S is true; as a consequence of the premises the statement S is false; the given the premises are not sufficient to conclude that S is true or that S is false. This could be a critical point, since it is necessary to recognise that if the truth of statement S is NOT a logical consequence of the premises, then we CANNOT conclude that S must be false. This complex logical aspect could be problematic, and we cannot exclude that it might create some difficulties in the participants. When the results of the survey will be presented in the next chapter, this aspect should not be neglected. It is possible that some not correct answers are (also) due to this logical complexity of the questions.

A further general aspect to consider is that the four previous tasks refer to four generic numerical properties, without an explicit definition for them. This choice was made to avoid students to answer to the questions by evaluating the property on the given numbers. However, in this way the questions are situated on a 'meta' level, whose object are generic properties and predicates. This could have added a further level of complexity to the survey. In particular, in APOS terms, to construct actions or processes on generic properties requires that they have been encapsulated into objects, something that should not be taken for granted in students. For this reason, the results emerging from the survey should be considered as limited by this particular structure of the tasks. When I presented the GD of MI within the APOS Theory (section 4.3), I observed that the Explain Induction process is the process by which a subject constructs a justification of the functioning of the MI both for a specific predicate and for a generic predicate P(n). For a complete construction of the MI Schema, a subject must be able to construct this process even for a generic predicate P(n). In this sense, thus, to investigate the construction of the Explain induction process also on this 'meta' level is relevant for the research objectives of this study.

Conclusive questions

Finally, the last section of the survey contains the following questions:

1) Has the proof by Mathematical Induction ever been object of study for you?
   ☐ Yes
   ☐ No
   ☐ I do not remember

[The following five questions (from 2 to 6) were not presented for the ones who answered 'No' at the previous question.]

2) When has the proof by Mathematical Induction been object of study for you? (You can select more options)
   ☐ Secondary school
   ☐ University
   ☐ I do not remember
   ☐ Other: _______.

3) What is, for you, a 'proof by mathematical induction'? (Do not worry about using a formal language; if you wish you may refer to mental images, feelings, memories).
   [Space for open answer]

4) Explain with your own words why a proof by mathematical induction assures that a proposition is true for all the natural numbers. (Do not worry about using a formal language; if you wish you may refer to mental images, feelings, memories).
   [Space for open answer]

5) Are there any aspects of the proof by mathematical induction that you do not understand or that you are not convinced by? (Do not worry about using a formal language; if you wish you may refer to mental images, feelings, memories).

[Space for open answer]

6) In the questions on Aarney's, Bertrand's, Coleman's, and Evelin's properties it was not possible to state anything certain because the proof of the base case of the induction was missing.
☐ True
☐ False
☐ I do not know

7) Would you like to change any questions that you gave in the survey? If yes, try to explain which ones and how; do not worry if you do not remember the details.
[Space for open answer]

8) Please leave a comment:
Express freely your thoughts and your doubts on the questions, on your answers, and in general on the topic of the survey
[Space for open answer]

After a participant clicks the 'submit the survey' button this message appears:

Your answer has been successfully registered. Thank you for your time!

Comments on the conclusive questions

The question 1) was added to divide the collected data by considering the group of participants who say to have encountered MI during their studies and the others. This allowed me to avoid posing the successive questions on MI to those participants who stated not to have encountered MI during their studies. Moreover, as we will see in the data analysis, it was possible to consider possible differences in the answers between the groups of participants.

The question 2) was added to register how many students, from what they can remember, encountered the proof by MI during the secondary school.[75]

In the successive questions 3), 4), and 5), I asked participants for some personal reflections on MI. These questions aim to register aspects related to the intuitive acceptance of the proof by MI. In relation to these three open questions a remark is necessary. They can provide elements with which to investigate the intuitive acceptance of mathematical induction by students (and as we shall see in the results, this did happen). However, we should be aware of the limits in investigating intuitions with questions of this kind. Indeed, they require explicit reflection on mathematical induction and on difficulties or unclear aspects in relation to it, but we cannot know to what extent the students who are answering are aware of their own difficulties and intuitions. In other terms, it is important to recognise, a-priori, that with

[75] The result for this question was anticipated in the introduction of this thesis (section 1.1). I registered that less than one third of the participants (31.0%) state to remember to have encountered the proof by MI at school. This result, as observed in the introduction, is in contrast with the centrality and importance that is given to the proofs by MI in the 'Indicazioni Nazionali per i Licei' (MIUR, 2012).

these questions it will be possible to collect the data related only to those students who are both aware of their own difficulties, uncertainties and, in general, aspects that could escape a full comprehension, and who, at the same time, are able, and willing, to put these aspects into a written answer. Therefore, it is certainly possible that there are elements related to the intuitive acceptability of induction that will not be registered by these questions.

The question 6) was added since, coherently with the literature (Avital & Libeskind, 1978; Harel, 2001), I thought that it was possible that some students consider that the base of the induction must be, necessarily, 0 or 1. Since in none of the previous task, information on the truth of the property for 0 or 1 is given, it is possible that some students answered "True" for this question.

The question 7) was added to give participants the possibility to describe possible given answers that they would want to change. With the answers to this question, I could register the potential presence of 'critical' moments in which subjects have realised they wanted to change a given answer (not necessarily into a mathematically correct one). For instance, in relation to the previous question, it is possible that some students, after having read it, would want to change the given questions of Tasks A-D.

Lastly, the question 8) was added to leave the space for free comments on the survey.

8.3.2 Participants

The survey was put online and the link to access it was distributed to 14 different Italian universities, in several academic courses. During the lessons or through the online page of the course, students were asked to participate, voluntarily, to the survey. They were generically informed that it was an anonymous survey for a study of mathematics education with an approximate duration of 20 minutes that they could complete during their free time. The link to the survey was kept active for 10 days.

8.4 Second empirical study – Task-based interviews

The second conducted empirical study involved a series of interviews with some university students dealing with the individual resolution of some problems. The research focus of this study was to investigate the processes involved in the exploration and the resolution of problems which could involve MI. Coherently with this focus the approach is purely qualitative. In detail, I decided to make use of the methodological tool of *clinical interviews,* more specifically of *task-based interviews.* In the next section I shall briefly present the general characteristics of these particular types of interviews.

8.4.1 Clinical interviews and Task-based interviews

With the term 'Clinical interview' is generally indicate the methodological approach developed by Piaget (1929) in his studies on the 'genetic epistemology' with children. In particular, as described by Ginsburg (1981), a clinical interview is

> a flexible method of questioning intended to explore the richness of children's thought, to capture its fundamental activities, and to establish the child's cognitive competence. (p. 4)

During the years this methodology have been widen to studies not necessarily involving children. More recently, diSessa (2007) describes a **clinical interview** as

> **A one-to-one encounter between an interviewer, who has a particular research agenda, and a subject**. […]. The interviewer proposes usually problematic situations or issues to think about and the interviewee is encouraged to engage these as best he/she can. The focal issue may be a problem to solve, something to explain, or merely something to think about. An interviewer may encourage the subject to talk aloud while thinking and to use whatever materials may be at hand to explore the issue or explain his/her thinking. (p.525, bold added by me)

The aim of a clinical interview, thus, is to give the researcher the possibility to observe the interviewee when involved in a particular situation which was designed a-priori by the researcher. The focus of the observation is on the process which characterise the development of the interview. It is important to notice that, in a clinical interview, the interviewer is at the same time an observer and a participant to the interview. The interviewer, thus, is active and can intervene to direct the interview toward those aspects that s/he is more interested in investigating. Among the interventions that the interview generally makes are the requests of clarification, of explanation, or the request of describing aloud what the interviewee is thinking about.

As described above, during a clinical interview, the interviewer and the interviewee are involved in a particular situation. This situation could correspond to a problematic situation, as a task to accomplish or to reflect on. In this case the term **Task-based interview** is used. This methodological tool is widely used in mathematics education since, as described by Goldin (2000),

> task-based interviews make it possible to focus research attention more directly on the subjects' processes of addressing mathematical tasks, rather than just on the patterns of correct and incorrect answers in the results they produce. Thus, there is the possibility of delving into a variety of important topics more deeply than is possible by other experimental means. (p. 520).

In the case of a task-based interview, the role of the interviewee characterised by a further element: s/he becomes a *problem solver*. Thus, the focus of the observation is on the processes involved in the problem solving situation created by the task.

A further methodological aspect to be discussed involved the choice, as anticipated above, of interviewing expert students (PhD). The analysis of experts' processes is considered a valuable and important research focus in studies on problem solving for the richness and effectiveness of the processes activated by such students. For example, Carlson and Bloom (2005) hypothesise that

> by observing individuals with a broad, deep knowledge base and extensive problem-solving experience, we would learn more about the problem-solving process and interactions of various problem-solving attributes (e.g., cognitive processes, metacognitive behaviors, and affective responses). (p. 52)

Furthermore, the analysis of experts' processes could provide elements with which to interpret phenomena (such as difficulties) involving non expert students. More specifically, as observed by Antonini (2011):

> The analysis of experts' processes may frame the study of students' difficulties, since these may be interpreted through non-activated processes. In other words, experts' processes and strategies can be useful in order to identify those processes that are not activated when non-expert subjects get stuck. (p. 207).

In conclusion, thus, coherently to what just said, in this study I initially focused on expert students' processes when dealing with the resolution of problems which could involve MI. As we will see, with the conducted analysis it will be possible to register a series of processes, highlighted by some 'crucial' signs, which seem to be crucial in the generation of a recursive argumentation or in the construction of a proof by MI. The absence of those signs in less expert students could be then interpreted in terms of non-activated processes.

Having made these brief methodological premises in the next sections I shall present in detail the research design of this study.

8.4.2 General characteristics of the interviews in this study

The students participating to the study were interviewed individually by me. The interviews involved different tasks (presented in the next section) which were presented to the participants one at a time. There was not a fixed limit of time for each task, but I decided to conclude each interview within 90 minutes in total. I tried to create an informal, encouraging, and supporting environment during the interviews. The first minutes were dedicated to some routine questions to break the ice ("How is it going at the university?", "What are currently studying?"). Successively the modalities of the interview were presented to each participant which were informed how the data for my study would be collected (vide-audio registrations and the sheet used during the interviews) [76]. I said to each participant that I was not interested to register if they could solve correctly or not the assigned problems but that I was interested to the resolution process. Thus, they were asked to "reason aloud" when possible. During the resolution of a task, I limited my interventions to the requests of clarification ("Could you explain to me again what you said please?") or, in case of prolonged silence, the requests of describing their thoughts ("Try to describe to me what you are thinking about"). Lastly, at the end of the interview, the various assigned problems were commented and, if not solved, a possible solution was provided.

8.4.3 The tasks

The tasks used in the interviews are presented below and, in the problems for which it is possible, I shall also propose a possible solution that makes use of mathematical induction (note that this solving strategy is not unique). At the end of each task some considerations about the choice to use the given problem and its formulation are made explicit.

<u>The Chessboard problem</u>

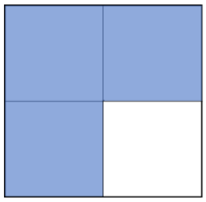
A 2^nx2^n squares chessboard is given. What is the maximum number of squares that can be covered with L-shape tiles composed by 3 squares? Justify your answer.

Let's note that, since 3 does not divide 2^nx2^n, it will never be possible to cover the whole chessboard with the given tiles. The best case, then, is that we can cover the entire chessboard, except for one square. If this were true, for every *n*, then the answer for the question of the problem would be that the maximum number of squares that can be covered is 2^nx2^n-1.

We prove by induction that what just written is true. Actually, we will prove something stronger, namely that given a chessboard of size 2^nx2^n it is possible to cover it all except for a small square occupying one of the vertices, from which follows trivially the statement above.

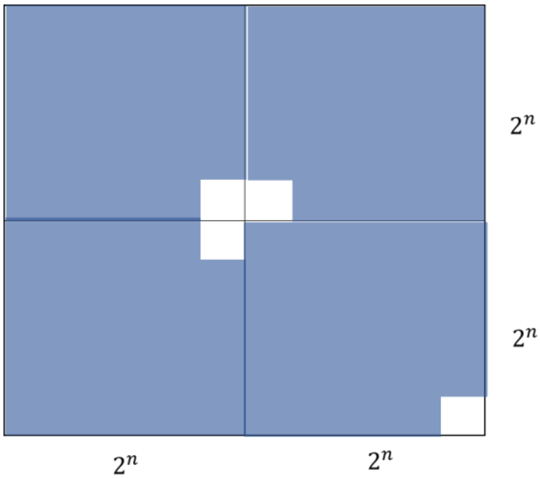
If n=1, we have a 2x2 checkerboard that can trivially be covered entirely except for a small square in any of the vertices (the figure of above shows one of the cases; the others can be trivially obtained by rotating the given figure).

[76] This information was already provided to the students when I asked for their availability to participate in the interviews. In particular, they were informed in advance that the interviews would have been audio and video recorded.



Now we suppose that it is possible to tessellate a chessboard of size 2^nx2^n except for a small square in one of the vertices, and we show that then it is possible to tessellate a chessboard of size 2^{n+1}x2^{n+1} except for one of the vertices.

A tessellation for a 2^{n+1}x2^{n+1} chessboard can be obtained as follows: we divide it into four 2^nx2^n sub-chessboards. By inductive hypothesis, the four sub-chessboards can be tessellated entirely except for a square at one of the vertices. Let the squares left out of 3 sub-chessboards concur in the centre of the 2^{n+1}x2^{n+1}chessboard, and let the square left out of the fourth one occupies one of the vertices of the 2^{n+1}x2^{n+1} chessboard. The figure below represents this situation, in the case where the bottom right vertex of the chessboard is left out; the other cases can be obtained by rotating this chessboard.



At this point we can use an L-shaped tile to cover the three squares in the centre and we have thus tessellated a chessboard of side 2^{n+1}x2^{n+1} except for a square at one of the vertices.

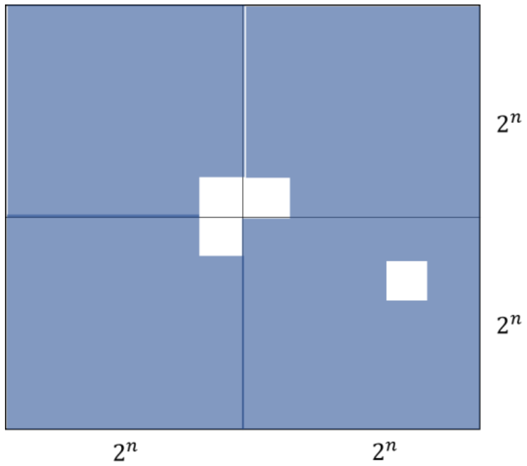
This concludes the proof.

■

Let us observe, first, that the requirement that the square left out of the tessellation occupies one of the vertices is not necessary. It is in fact possible to show that, given a chessboard like the one in the problem, it is always possible to find a tessellation that covers the entire chessboard except for one of the squares occupying any position. This can be proved again by induction. The proof is very similar to the previous one.

The base of the induction is the same as in the previous case.

Instead, for the inductive step, once the square to be left out is fixed, in any position, we divide the 2^{n+1}x2^{n+1} chessboard into four 2^nx2^n sub-chessboards, one of which will contain the square fixed at the beginning. At this point, the inductive hypothesis is applied for each of them. In particular, for the sub-chessboard that contains the fixed square we construct a tessellation that leaves it out, for the other three sub-chessboards we construct a tessellation like the one in the previous proof, so that the respective squares that remain empty concur in the centre of the $2^{n+1}x2^{n+1}$ chessboard.



At this point we cover the three squares in the centre with a tessellation and get the tessellation for the desired $2^{n+1}x2^{n+1}$, which concludes the proof.

■

The chessboard problem was taken from Harel (2001):

> Let n be a positive integer. Show that any $2^n \times 2^n$ chessboard with one square removed can be tiled using L-shaped pieces, where these pieces cover three squares at a time. (p. 195)

Compared to Harel's formulation, however, a more open formulation was chosen that did not already provide a statement to be proved. This aligns with the premises on cognitive unity adopted in the framework and with the research objectives of investigating also the processes involved during the exploration of the problem and the generation of a conjecture. I hypothesized that looking for a possible tessellation for different checkerboards of size 2^nx2^n, as n varies, could play an important role in solving the problem, possibly in the construction of a recursive argument. This hypothesis is in agreement with the work of Stylianides, Sandefur and Watson (2016), analysed in section 1.3.3, where they record the key role in exploring some examples for constructing a proof by induction.

<u>The false coin problem</u>

N identical coins are given. One of these, however, is fake and it weighs less than the others. There is a traditional two plates weighing scale at our disposition. What is, in function of N, the minimum number of weighings necessary to determine the fake coin?

Let us first consider a particular case of the problem, where $N=3^m$. In this way we can divide the group of coins into three equal groups and weigh two of these groups. With a weighing we can then determine in which group the fake coin is (either one of the two groups on the balance weighs less, in which case it contains the fake coin, or it is in the group left out). After one weighing, we then have a new group of 3^{m-1} coins in which there is a counterfeit coin. We can iterate this reasoning and, after a total of m weighings, determine the counterfeit coin.

Thus, we have shown that given a group of 3^m coins, one of which is false, it is possible to determine the false one with *m* weights. The argument given above can be easily formalised with a proof by induction on *m*.

The general case of the problem is a bit more complex and can be solved with a strategy similar to the previous one. Let N>3. If N is divisible by 3 then we divide the group into three equal groups and, with the strategy of before, we determine the group containing the fake coin. If N is not divisible by 3, then we divide the pile into three smaller piles of which two have the same number of coins and the third differs by only one coin from the other two (this is always possible because $N=3a+1$ or $N=3a-1$, for some natural a). We then weigh the two piles with the same number of coins and determine in which of the three piles is the fake coin. By using this strategy, we can determine the fake coin. In the worst case it would correspond to the last coin left out. We observe that it is possible to estimate how much the size of the pile of coins decreases at each step of the procedure. If at a certain point the pile is composed of L coins, with L between 3^k and 3^{k+1}, then at the next step there will be a pile of M coins with M between 3^{k-1} and 3^k. This observation ensures that after a number of weighings equal to the upper integer part of $\log_3(N)$ we arrive at determining the fake coin.[77] ■

A problem similar to this one is among those proposed by Harel (2001):

> You are given 3^n coins, all identical except for one which is heavier. Using a balance, prove that you can find the heavy coin in *n* weighings. (p. 195)

Again, changes were made from Harel's version. First, I decided to formulate the problem in an open form, for the same reasons as the previous problem. Also, I decided to give the problem in the general form, where the pile of coins was not necessarily a power of 3, as in Harel's formulation. This was done to make it necessary an exploration of the problem. In fact, I hypothesized that the students to whom the problem would be given (PhD students in mathematics) were already familiar with the version of the problem involving a power of 3 and would therefore try to solve the problem by tracing back to this case.

The problem was assigned not so much with the objective of seeing if the students could formalise a general proof, but if they could construct the solving strategy in the case 3^m and eventually generalise this strategy also in the general case N. A priori, therefore, I decided to suggest students considering the case $N=3^m$, in case of block in the resolution of the problem, and in any case not to ask for a proof of the problem in the general case but to stop at the description of the solving strategy.

<u>The banknotes problem</u>

In a far away country, there are only banknotes which value 3 or 5 money. Prove that any quantity of money, sufficiently big, can be obtained using only these two kinds of banknotes.

It can be shown that, any amount greater than or equal to 8 can be obtained with an appropriate amount of banknotes valuing 3 or 5.

This corresponds to prove that $\forall n \geq 8$ natural number, $n=3a+5b$ for some $a, b \geq 0$ natural numbers. We prove this statement by induction on n.

[77] For a more formal solution of the problem and a more general version of it (in this case we do not know if the counterfeit coin is lighter or heavier than the others) is presented in Manvel (1977).

- Base of the induction. If n=8, then n=5+3.

- Inductive step. Suppose that n>8 is written as n=3a+5b, with a, b ≥ 0. We want to show that n+1=3c+5d with c, d ≥ 0. We consider two cases: b=0 or b≠0.

- If b=0 it means that n=3a and since n>8 it means that a≥3. So n+1 can be obtained from n as follows:
  n+1 = 3a+ 1 = 3(a-3) + 9 +1 = 3(a-3) +10 =
  = 3(a-3) + 5·2 = 3c +5d, where c=a-3 which is a natural number since a≥3 and d=2.

- Se b≠0, thus b≥1, n+1 can be obtained from n as follows:
  n+1= 3a + 5b +1 = 3a + 5(b-1) +5 +1 = 3a + 6 + 5(b-1) =
  = 3(a+2) +5(b-1) = 3c+5d, where c=a+2 which is trivially a natural number and d=b-1 which is a natural number since b≥1

This concludes the proof. ■

The problem is a reformulation of a problem contained in Dubinsky (1986):

> Prove by mathematical induction that any sufficiently large number of dollars can be obtained by using only bills of denomination $5 and $3 (assuming that such bills exist).

Compared to Dubinsky's formulation, I chose to remove the reference to mathematical induction. This agrees with the goal of observing possible signs of a genesis of a recursive argument without explicitly asking the student. As in Dubinsky's formulation, in our case the problem is given in closed form "Prove that...". Despite this, however, the boundary value to which the proposition refers is not given. It was hypothesized, therefore, that despite the formulation "prove that", students would still initiate a phase of exploration of the problem, until eventually conjecturing a value corresponding to the numerical limit beyond which the proposition is true, and then proving the statement.

We can also observe that the problem can be modified by adding the possibility that a quantity of money can be obtained by using 3 and 5-value bills, and possibly receiving other 3 and 5-value bills as change. In mathematical terms this corresponds to asking what natural numbers can be obtained as a linear combination with integer coefficients of 3 and 5. The problem can be easily solved by observing that in this way it is possible to obtain the value 1 (giving two banknotes valuing 5 and receiving 3 as change, for example) and in this way, having obtained 1, we can obtain every other natural number. A priori, therefore I decided, to modify the problem formulation in this second way in case a student could not solve the banknote problem in its original formulation.

<u>The diagonals problem</u>

Determine the number of diagonals in a convex polygon with N vertices.

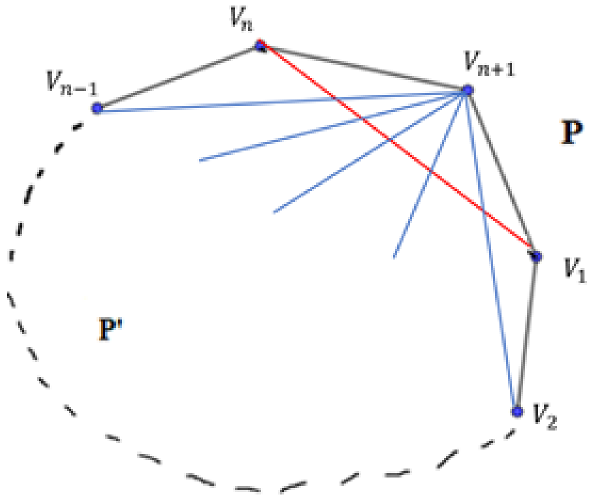
It is possible to prove by induction that, in a convex polygon with N vertices, the number of diagonals is given by the formula $\mathrm{D(N)} = \frac{(N-3)\cdot N}{2}$

- Base of the induction. If N=3 then $D(3) = 0$ which is indeed the number of diagonals of a triangle.

- Inductive step. We suppose that the number of diagonals of a polygon of N vertices is given by the formula D(N). Then we consider a polygon P of N+1 vertices that we call, in order V_1, V_2, ..., V_n, V_{n+1}. We want to show that the number of diagonals of P is

$$D(N+1) = \frac{(N+1-3)\cdot(N+1)}{2} = \frac{(N-2)\cdot(N+1)}{2} = \frac{N^2-N-2}{2}$$

Now, the number of diagonals of this polygon will be given by the number of diagonals of the polygon P' of vertices V_1, ..., V_n to which we add the segment V_1V_n (which is a side in the polygon P', but a diagonal in P) and all the segments that connect the vertex V_{n+1} to the vertices of P' forming a diagonal for P, which will be *n*, as the vertices of P', minus the two segments V_nV_{n+1} and $V_{n+1}V_1$ that are sides in P. The figure below represents what has been said.



We then have that the number of diagonals of P is given by the number of diagonals of P' (which by inductive hypothesis is $\frac{(N-3)\cdot N}{2}$), by the segment V_1V_n and by n-2 other segments.

In formula we have that the number of diagonals of P is equal to

$$D(N) + 1 + n - 2 = \frac{(N-3)\cdot N}{2} + n - 1 = \frac{N^2-3N+2N-2}{2} = \frac{N^2-N-2}{2}$$

that it is just the corresponding expression to D(N+1) drawn above. ■

We observe that the problem can be solved even without a recursive strategy: in particular we can observe that from each vertex of a polygon of N sides N-3 distinct diagonals start. If we count the diagonals of the polygon, without worrying about repetitions, simply as N-3 diagonals for each of the N vertices, we get $(N-3)\cdot N$. By doing so, each diagonal has been counted twice (in correspondence of the two vertices that are its extremes). From this, then, it follows that the number of diagonals of a polygon of N sides is $\frac{(N-3)\cdot N}{2}$.

Note that, also in this problem, as in the previous ones, a formulation was chosen that required, first, the production of a conjecture and then its demonstration.

A divisibility problem

Prove the following statement: For any n, prime, (n-1)! + 1 is divisible by n.

The problem corresponds to an implication of a number theory result known as Wilson's theorem: 'Any natural number *n>1* is prime if and only if $(n-1)! \equiv -1 \ mod \ n$', whose proof is not immediate.

This problem was given not to investigate its complete resolution by students, but instead to see if someone wanted to try to prove it by induction. In case this did not happen, I decided a priori to intervene and to explicitly ask "Why don't you try to prove it by induction?". Mathematical induction, in this case cannot work because the statement involves only prime numbers, for which we do not have a recursive form to express a prime as a function of the previous (which happens instead in N, 2N, etc ...).

A finite sum

Find the value of the following sum and prove the result:

$$\sum_{k=1}^{n} \frac{1}{k(k+1)}$$

One possible solution to the problem is to conjecture the value of the sum, for instance by generalising empirically from some cases, and then to prove the statement obtained by induction. This extremely standard problem was included to observe students involved in a 'standard' proof by induction. However, also in this problem, it was chosen not to assign the value of the sum, but to have the student conjecture it. This aligns with the theoretical premises and research objectives.

A (false) proof by induction[78]

I will prove by induction that there exists an algorithm that, given N coins, of which one is false and lighter than the others, and a traditional two plates weighing scale, determine in no more than 4 weightings the false coin.

Proof

- Base of induction

If N=2, with only one weighting, we can determine the false coin.

- Induction step

We assume, for the inductive hypothesis, that there exists an algorithm that, given N coins determine the false one in no more than 4 weightings. Let us consider a group of N+1 coins, of which one is false. We can determine the false coin with the following strategy:

(a) We take out 1 coin from the group
(b) We apply the algorithm that we have assumed to exist by inductive hypothesis to the group of N coins and in no more than 4 weightings we find the false coin, if it was in the group.

[78] The adjective "false" in the title is used to easily recall the problem in this thesis, but it was not written in the text given to the students. During the interviews the page containing this problem was titled as "A proof by induction".

(c) If, after 4 weightings, we have not found the false coin, it means that it was not in the group of N, therefore it is the one left out.

■

Comment the previous proof.

The task is an adaptation of the task used by Movshovitz-Hadar (1993) to investigate the fragility of knowledge related to MI. The proof given above is a less condensed version of the (false) proof given by Austin (1988) to the 'false coin problem', cited by Movshovitz-Hadar (1993, p. 254).

The aim of the task is to create a conflicting situation: on one side the proof seems to be correct, but on the other side the statement which has been proved is counterintuitive (it is false, indeed).

The flaw in the proof is not easy to find. In point (b), when the inductive hypothesis is applied, we have a group of N coins which may or may not contain the false coin. However, the inductive hypothesis involved a group of N coins, one of which must be false. Therefore, it was not possible to use the inductive hypothesis.

'Meta' questions on MI

Matteo claims that:
'For every *n*, $n \geq 1$, $n^3 - n$ is divisible by *6*'.

Matteo proposes the following proof:
By induction on n*:*

- *n* = 1, thus $(1)^3 - (1) = 1 - 1 = 0$ which is divisible by 6.

- We assume the claim true for $n = k$, that is $k^3 - k$ is divisible by 6,
  we prove the claim for $n = k + 1$.

$$(k+1)^3 - (k+1) =$$

$$= k^3 + 3k^2 + 3k + 1 - k - 1 =$$

$$= k^3 - k + 3k^2 + 3k =$$

$$= k^3 - k + 3k(k+1)$$

Now, $k^3 - k$ is divisible by 6 for the inductive hypothesis.
Also $3k(k+1)$ is divisible by 6 in fact:
$k(k+1)$ is even because in two consecutive numbers one is always even.
Therefore $3k(k+1)$ is a multiple of 6.

Therefore, since the sum of numbers which are divisible by 6 is still divisible by 6, $k^3 - k + 3k(k+1) = (k+1)^3 - (k+1)$ is divisible by 6.

Therefore, by induction, we have proved the proposition.

■

Answer the following questions:

1) Is Matteo's claim true?

2) Is Matteo's proof correct?

3) Is the generality of Matteo's claim guaranteed by his proof?

4) Marco affirms that he verified the proposition for $n = 2357$ he found that $2357^3 - 2357 = 105\,514\,228$ is not divisible by 6.
What is your opinion on this?

5) In relation to the previous point, comment the following claims:

a) From a theoretical point of view, Marco could be right because Matteo's proof was based on an inductive hypothesis, that we supposed true, but if it was false instead, the proof would not have been correct.

b) The statement is true, except for some particular cases, such as the one found by Marco. The proof given by Matteo is correct because it was said "If it is true for *n=k*, then it is true also for *n=k+1*", but if it was not true for *n=k*, then there could be some counterexamples like Marco's one.

c) Matteo proved that from 1 to 2357 the theorem is true, but thanks to Marco's counterexample we know that from that number on it becomes false because the inductive hypothesis does not hold anymore.

6) Comment the following claims in relation to Mathematical Induction:

a) A proof by Mathematical Induction is based on the Principle of Mathematical Induction, an axiom that can be accepted or not. In particular, in a proof, when we use the inductive hypothesis, we assume the thesis that we want to prove. This strategy, even if it is counterintuitive, works if we accept the Principle of Mathematical Induction.

b) A proof by Mathematical Induction is an infinite proof: to prove the claim for all the natural numbers it is necessary to make an infinite number of steps and, formally, it would end after an infinite amount of time.

Matteo's statement is true and the given proof is correct. The proof and the questions 1-4 are an adaptation of the first questions of Fischbein's study (1982), presented at 5.2.3. The first three are intended to investigate the three levels of intuitive acceptability of the proof provided, while the question 4) is included to investigate the students' resistance to the presence of any counterexamples to the proof.

In 5) three statements are made, which from a mathematical point of view are not correct, but which could be problematic for a student to refute since they involve delicate aspects related to MI. Specifically, statements (a), (b), (c) confuse the truth value of the inductive hypothesis (which for the purposes of proving the inductive step could be true or false) with the truth value of the principal statement (an example of what I called 'interference between S and S2' see 3.3.3). These questions were included to investigate the intuitive acceptability of the inductive step as a theorem independent from the principal theorem.

The point 6) involves two statements, which are still not correct from a mathematical point of view and may be problematic to confute for students. Specifically, point (a) refers to the Principle of MI as an axiom of a particular theory, stating that it is counterintuitive since it assumes what it is intended to prove. The request to comment on this sentence is meant to stimulate a reflection on the justification of mathematical induction. In other terms, I am interested in investigating the reaction to being told that the principle of induction is counterintuitive. To go against this claim, a student might give his/her own intuitive justification for how induction works. Assertion (b), on the other hand, involves another delicate aspect of MI, which is the dialectic between finite and infinite proof, analysed in the GD of MI (4.4.5). In this analysis, it was hypothesized the transition from a potential view of infinity to an actual one is involved in the encapsulation of the Explain induction, from an infinite process that potentially reaches every natural number greater than the base, to the object 'Proof by Induction' that guarantees the truth of the statement for all natural numbers.

Comments on MI

Let's talk about Mathematical Induction. What do you think about this method?

This question was posed, asked verbally, at the conclusion of the interview. The question was asked of those students who were not involved the previous task (see table 2, below).

The objective of the question is to investigate the interviewee's experience with mathematical induction, whether he/she is convinced by it and if there are aspects that are not clear and if s/he remembers of having difficulties with it.

The last two tasks presented were aimed at getting students to think about mathematical induction itself, how it works, its correctness and intuitive acceptability. For this reason, from now on, I will refer to the questions in these two tasks as "meta-questions about mathematical induction".

8.4.4 Participants

Nine students have been individually interviewed by me for approximately 90 minutes each. The students knew that I was a PhD student in mathematics education, but they did not know the specific focus of my study. The table 8.3 of below presents, for the interviewed students, the pseudonyms with which I will refer to them, their university grade, and the tasks used in each interview.

Student's pseudonym and University grade	**Tasks involved in the interview**								
	The chessboard problem	The false coin problem	The banknotes problem	The diagonal problem	A divisibility problem	A finite sum	A false proof by induction	'Meta' questions on MI	Comments on MI
Lorenzo - PhD Student in Mathematics	X	X	X		X		X	X	
Silvio - PhD student in Mathematics	X						X	X	
Lucrezia - PhD student in Mathematics	X				X		X	X	
Guido - PhD student in Mathematics	X		X		X	X		X	
Giuditta - Master student in Mathematics	X		X		X		X	X	
Tommaso - Undergraduate student in Physics			X	X					X
Dario - Undergraduate student in Physics	X		X						X
Claudio - Undergraduate student in Mathematics	X			X					X
Valentina - Undergraduate student in Mathematics	X		X						X

Table 8.3. *With reference to the interviewed students, the pseudonyms with which I will refer to them, their university grade, and the tasks used in each interview.*

8.4.5 Data analysis

Data consist of audio and video recordings, and of written inscriptions produced by the students. The camera was positioned in a way to register both the table containing the sheet on which the interviewee was writing and part of her/his body in order to register the "gesture space" (McNeill, 1992, pp. 86-91).

The transcription of the recordings was made considering the semiotic bundle composed by the tree semiotic sets: the speech, the gestures, and the inscriptions. Specifically, the transcripts were organised in tables with the following structure:

Line number	**Time**	**Who**	**Speech**	**Gesture**	**Inscription**

In the description of the data analysis that will follow in this thesis, transcripts will be presented with reference to this structure, but containing only those columns necessary for the analysis and, eventually, with an extra column labelled as "Comment".

Moreover, in the transcripts the following notations will be used:

- The interviewer (me) is indicated with 'I', the student with the initial of her/his pseudonym.

- The pauses are indicated respectively with a full stop [.] for short pauses which seem to correspond to the end of a sentence; three dots […] for longer pauses up to 5 seconds, also in the middle of a sentence; a square parenthesis with indication of time [*# of sec.*] for pauses longer than 5 seconds.

- For what concerns the description of a gesture, in particular to decide where it starts and it ends I considered, when present, its coherency with the simultaneous speech.

- When more than one gesture or inscription is produced symoultaneusly to the speech in the same line, in the column "Gesture" of the transcript, each gesture or inscription is described with reference to the spoken utterance simultaneous to the gesture. For instance, if a subject says "This is true for what written here" while pointing to a first inscription A and then to another inscription B, in the "Gesture" column I will write: "[This] *S/he points to the inscription 'A'.* [here] *S/he points to the inscription 'B' […]*".

Data have been analysed several times, focusing on different aspects, depending on the research question under investigation. The conducted analysis will be presented in this thesis following the different research objectives and not the chronological order of the interviews or of the tasks within an interview.

9 RESULTS, ANALYSIS AND DISCUSSION OF THE SURVEY

This chapter is divided in three main sections. In the first one, the results for all the questions of the survey are presented. Then, in the second section, these results will be analysed and discussed in relation to the research objectives of the study. Finally in the third section I shall draw some general conclusions on the analysed and discussed results.

9.1 RESULTS

9.1.1 Participants

We registered a total of 307 answers to the survey. For what concerns the Tasks A, B, C, and D, in this study I will present and analyse the answers of 252 participants, coming from the following groups of students:

- **Group G1.** First year undergraduate students in Natural Sciences, Biology, Agricultural Sciences, and Education Sciences. We collected 80 answers to the survey from this group.
- **Group G2.** First year undergraduate students in Mathematics, Physics, or Engineering. We collected 106 answers to the survey from this group.
- **Group G3.** Master's students in Mathematics. We collected 46 answers to the survey from this group.

This group division was made due to the different number of courses in Mathematics, Physics, or Computer Science that the students have encountered during their studies. In the specific case of the undergraduate students the distinction between G1 and G2 was introduced since in Italy, during the first year of university, the degrees listed in G2 include a high percentage of lectures in Mathematics, Physics, or Computer Science (about 60% for Engineering. close to 100% for the others), while instead this does not exceed 20% for the degrees listed in G1. The group G3, instead, is composed by master's students in Mathematics, therefore by students with a degree in Mathematics. [79]

In presenting the results, each group will be further divided considering the students who claimed to have encountered the proof by MI during their studies (G1+, G2+, G3+) and the ones who stated to have never encountered it or not to remember (G1-, G2-, G3-). From the data collected, the group G3- results empty, thus for G3 the division in + and - will not be made. The following table shows the number of participants to the survey for every group and subgroup.

	+	**-**	**Total**
G1	29	51	80
G2	101	25	126
G3	46	0	46

Table 9.1. *Groups' numerical composition.*

The survey was distributed also to other groups of students, in particular to second or third year undergraduate students in one of the degrees listed in G1 or G2, as well as master's students in Architecture or in Engineering. However, for these groups we collected a small number of answers (<10 participants), which therefore I decided not to include in the quantitative analysis of the results. This caused the difference between the 307 collected answers and the 252 answers actually presented and

[79] More precisely, this group could also include students with a first degree in Engineering or in Physics, attending then a master's in Mathematics. However, even in this case, they would be students with several more courses in mathematics in their curricula than the students in G1 or G2.

analysed. For what concerns the open answer questions of the last section of the survey, however, I will consider all the 307 collected answers.

The answers to these questions, containing comments or reflections on the proof by MI, will be analysed with a qualitative approach, thus every answer could be interesting independently from its statistical relevance.

In the next sections I will present the results to the multiple-choice questions of the Tasks A, B, C, and D, divided for the groups G1 (+ and -), G2 (+ and -), and G3. The answers for the open questions are presented and analysed in the discussion section.

9.1.2 TASK A – Results

In this task it is described a situation in which, for a predicate A(n), corresponding to "The property A holds for the natural number n", it is known that:

- A(103) is true;
- It is true that $\forall n(A(n) \rightarrow A(n+1))$.

From this it follows that A(n) is necessarily true for every natural number greater or equal to 103, whereas there is not enough information to conclude about the truth or the falsity of A(n) for the numbers that are lower than 103.

First question – The property A is: True for all the natural numbers / False for all the natural numbers / True for some and false for some natural numbers.

The answers to the first question, divided by groups and subgroups, are presented in the following tables.

Group G1	True for all the natural numbers	False for all the natural numbers	True for some and false for some natural numbers	We cannot know
G1	33.8%	0.0%	38.8%	27.5%
G1+	20.7%	0.0%	44.8%	34.5%
G1-	41.2%	0.0%	35.3%	23.5%

Table 9.2. *Percentage of answers for the first question of Task A for group G1, further divided in G1+ and G1-. The green colour indicates the mathematically correct answers.*

Group G2	True for all the natural numbers	False for all the natural numbers	True for some and false for some natural numbers	We cannot know
G2	20.6 %	0.8 %	42.9 %	35.7 %
G2+	18.8 %	1.0 %	42.6 %	37.6 %
G2-	28.0 %	0.0 %	44.0 %	28.0 %

Table 9.3. *Percentage of answers for the first question of Task A for group G2, further divided in G2+ and G2-. The green colour indicates the mathematically correct answers.*

Group G3	True for all the natural numbers	False for all the natural numbers	True for some and false for some natural numbers	We cannot know
G3	6.5%	2.2%	21.7%	69.6%

Table 9.4. *Percentage of answers for the first question of Task A for group G3. The green colour indicates the mathematically correct answers.*

<u>Numerical questions</u> – The property A(n) is: True / False / We cannot know.

The answers for **group G1,** and for **its subgroups G1+** and **G1-,** to the numerical questions (with the form A(n), for several natural numbers) are reported in the following three figures.

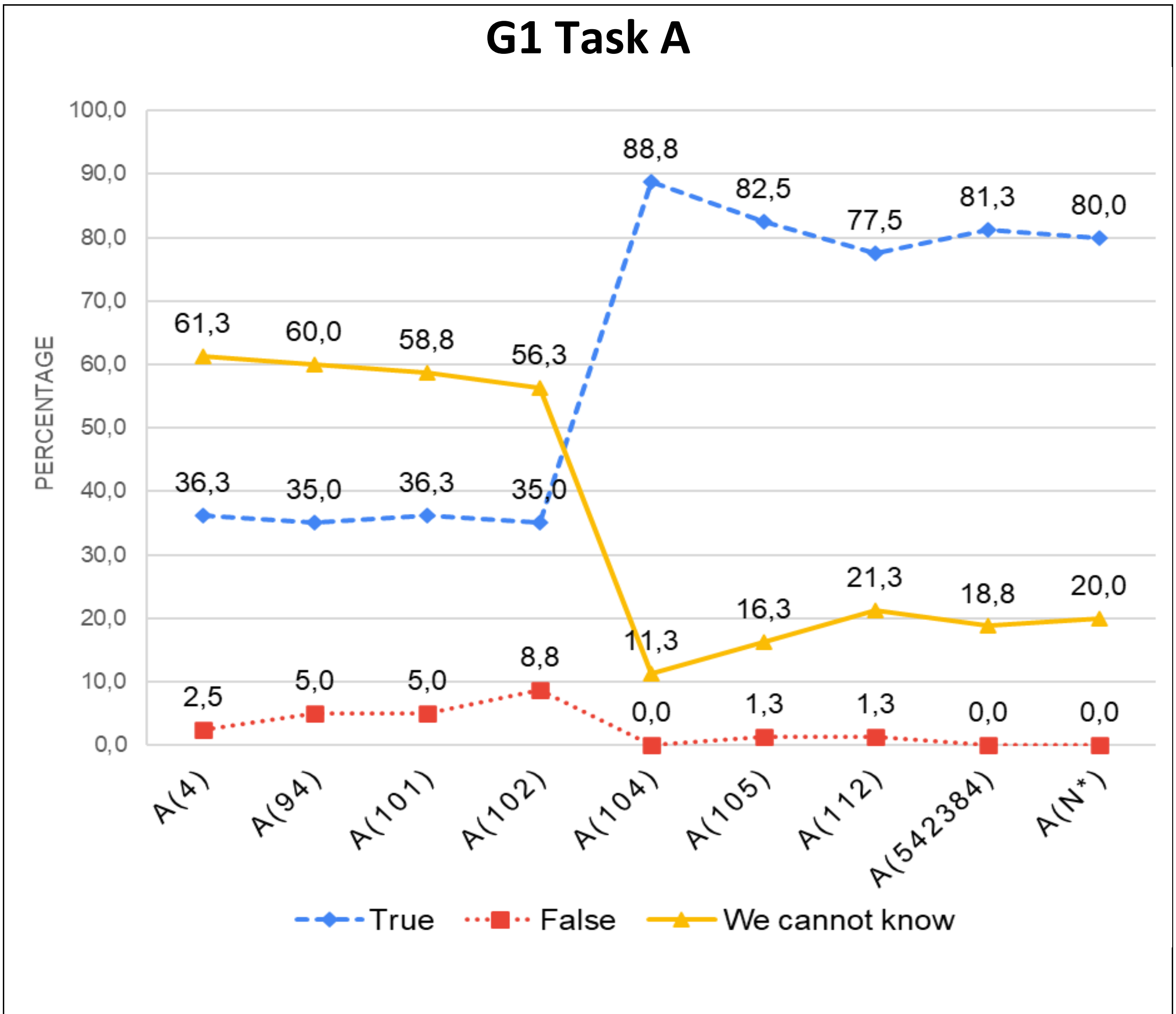


Figure 9.1. The chart shows the *percentage of answers for the questions A(n) for group G1 and the trends of these percentages when n varies.*

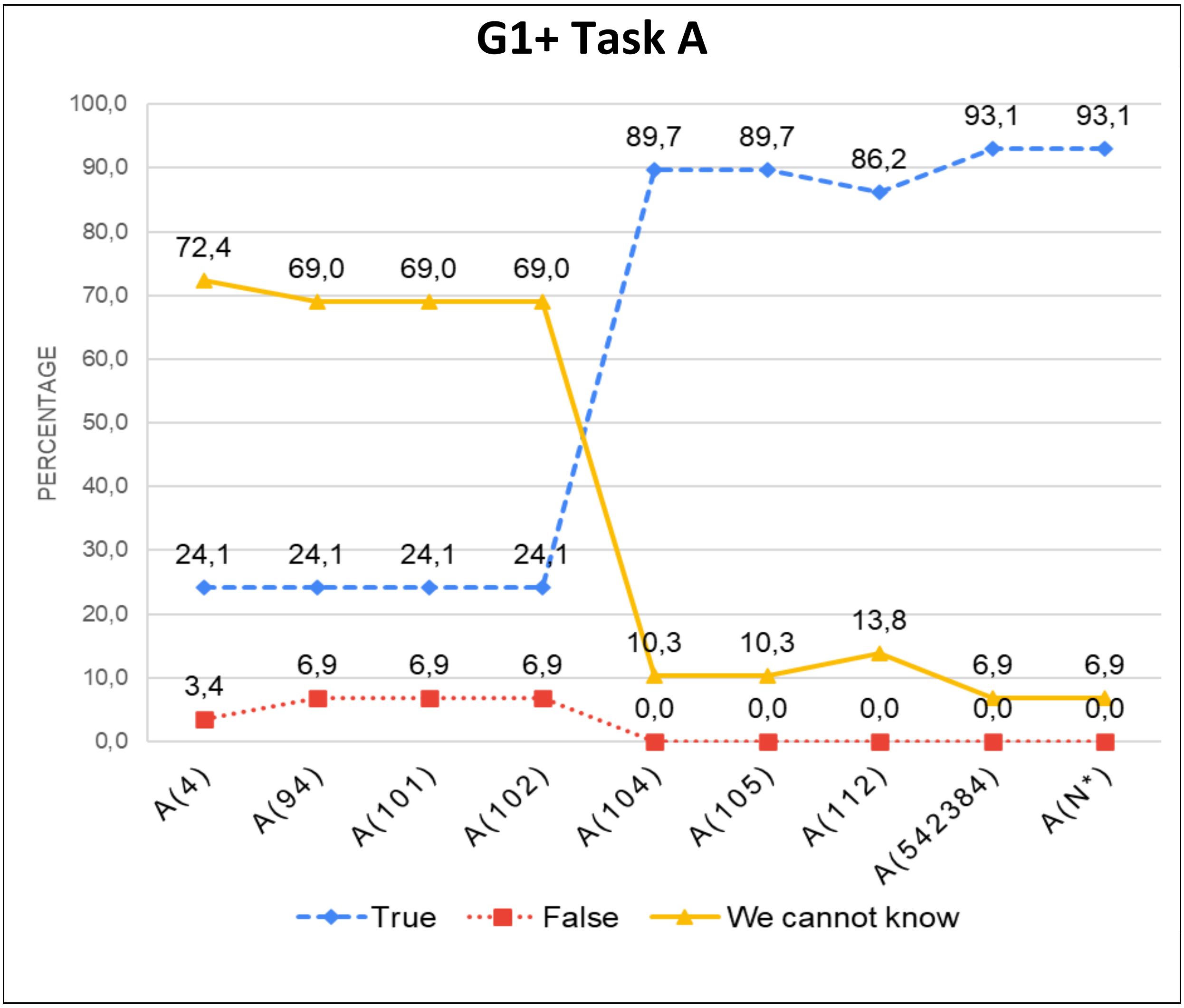


Figure 9.2. The chart shows the *percentage of answers for the questions A(n) for group G1+ and the trends of these percentages when n varies.*

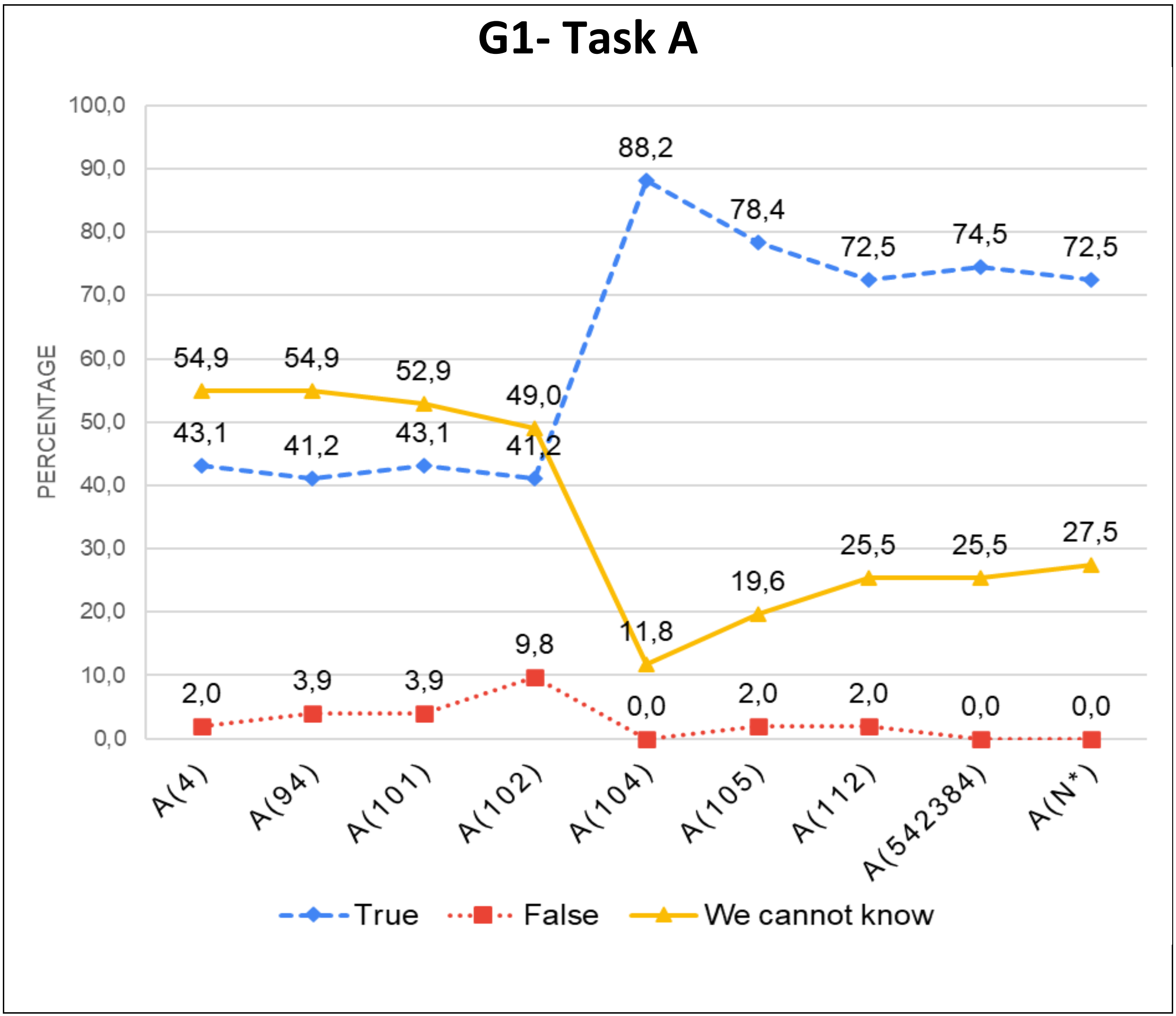


Figure 9.3. The chart shows the *percentage of answers for the questions A(n) for group G1- and the trends of these percentages when n varies.*

In the same way as for group G1, in the next three figures we will show the answers for this group of questions for the **group G2** and its **subgroups G2+** and **G2-.**

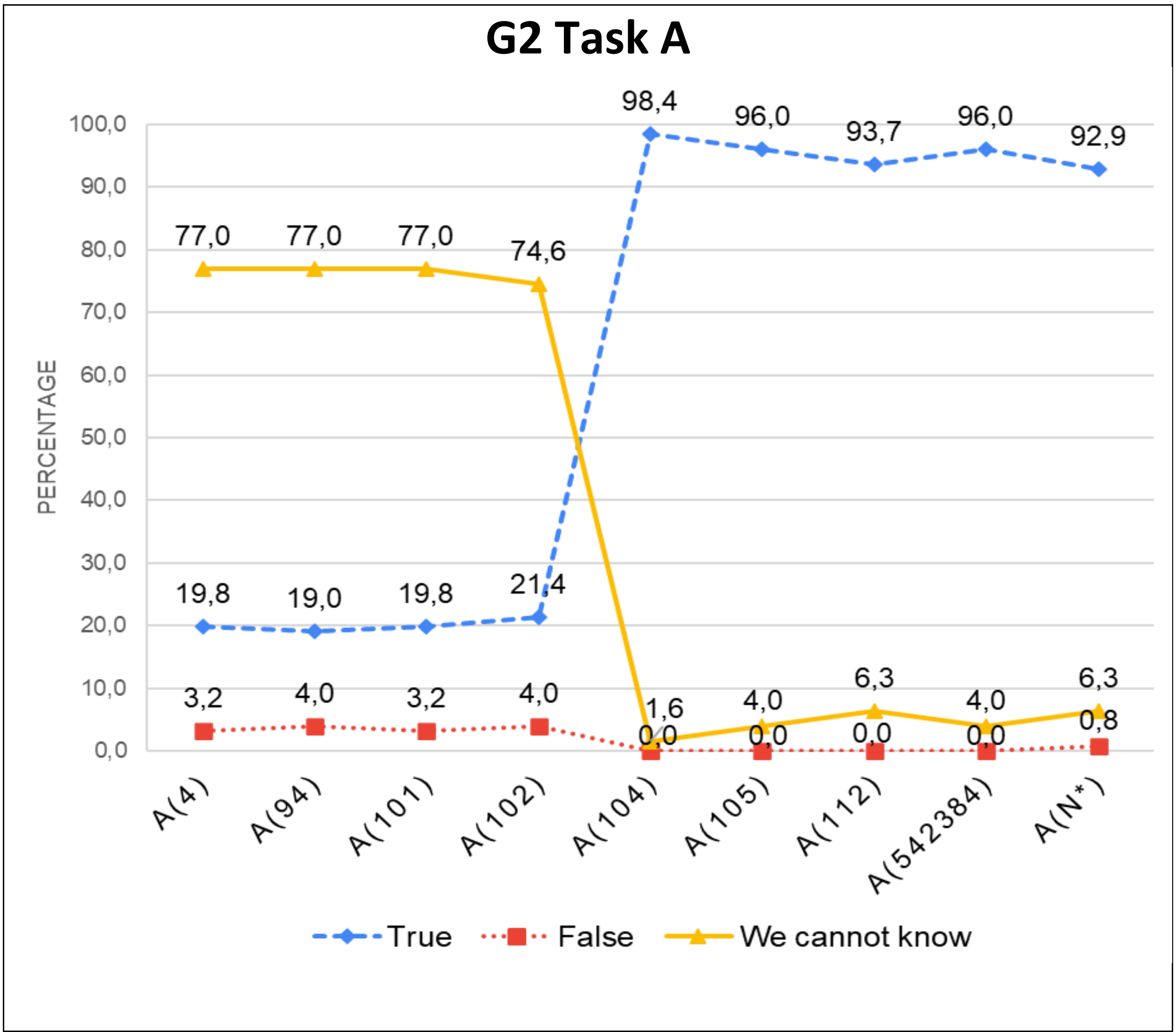


Figure 9.4. The chart shows the *percentage of answers for the questions A(n) for group G2 and the trends of these percentages when n varies.*

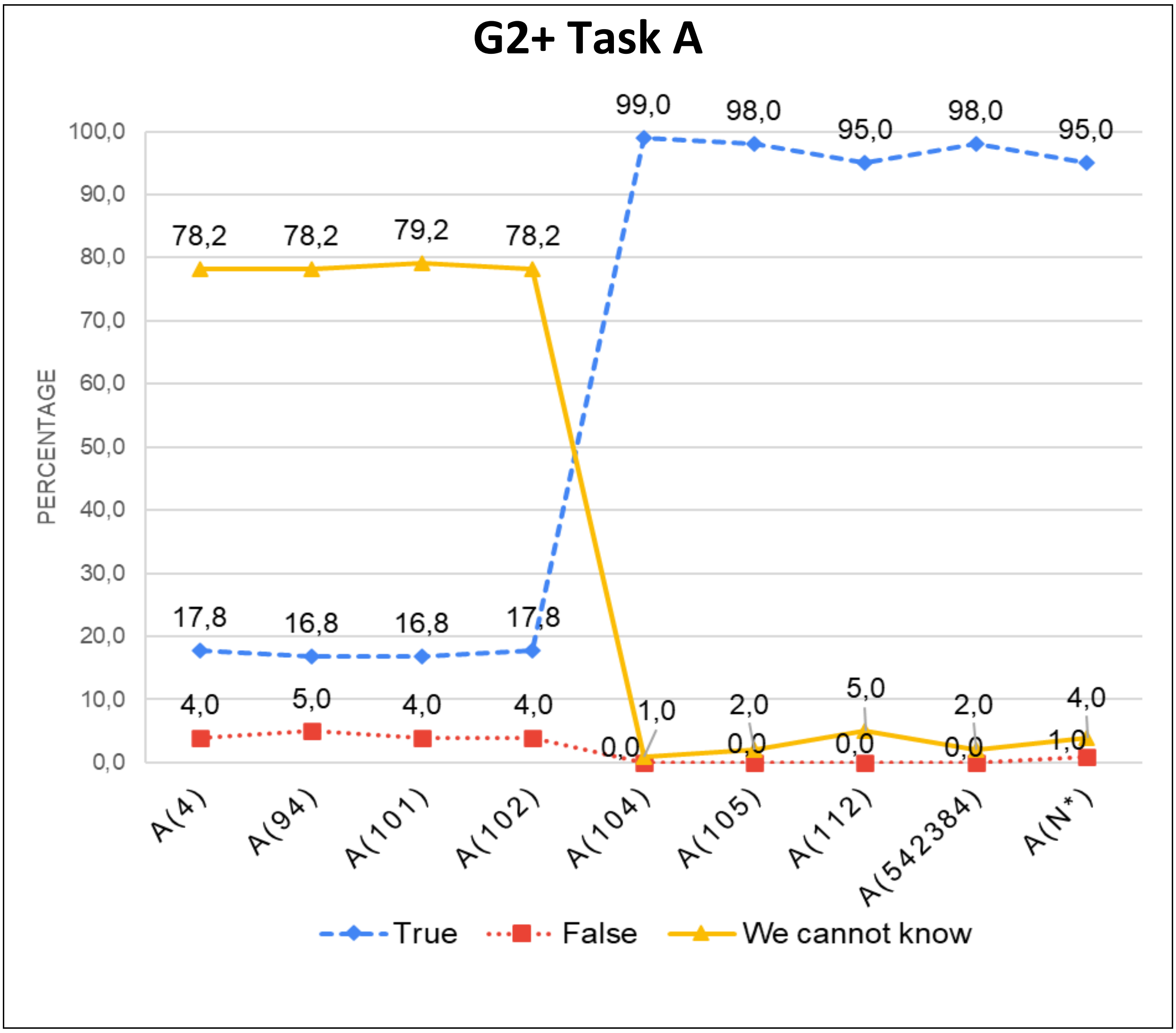


Figure 9.5. The chart shows the *percentage of answers for the questions A(n) for group G2+ and the trends of these percentages when n varies.*

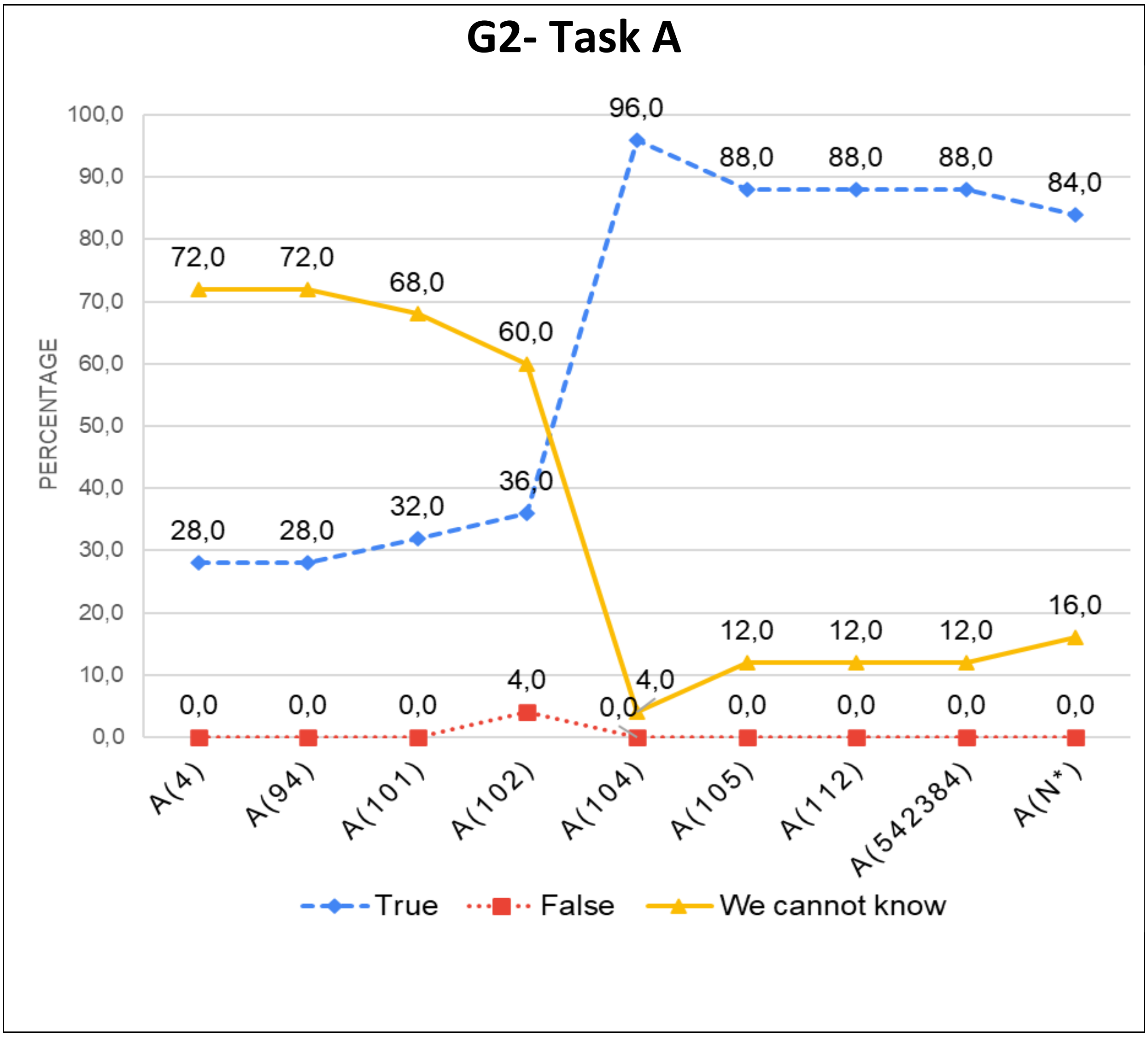


Figure 9.6. The chart shows the *percentage of answers for the questions A(n) for group G2- and the trends of these percentages when n varies.*

Finally, the following figure contains the answers for this series of questions for the **Group G3.**

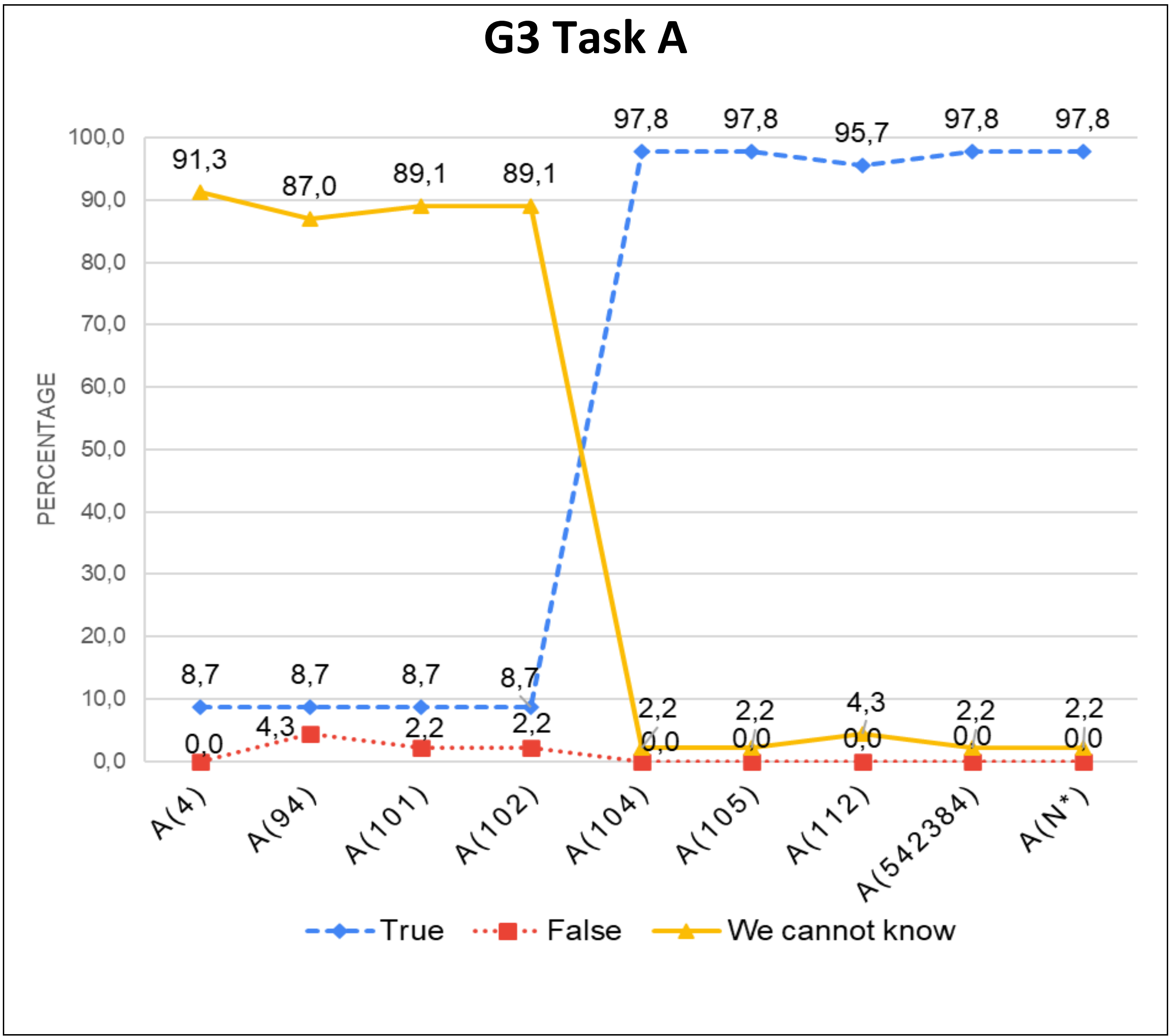


Figure 9.7. The chart shows the *percentage of answers for the questions A(n) for group G3 and the trends of these percentages when n varies.*

From the just presented results for Task A, we can extract the data corresponding to the **progression** of correct answers (TRUE) for the questions involving numbers increasingly greater than 103. In other terms, I will show the percentages of students who answered TRUE for every question from A(104) to some A(n) with n≥104. These data, presented for every group and subgroup, are shown in the following chart.

Group	**A(104)**	**A(104) & A(105)**	**A(104) & A(105) & A(112)**	**A(104) & A(105) & A(112) & A(542384)**	**A(104) & A(105) & A(112) & A(542384) & A(N*)**
G1	88.8%	82.5%	77.5%	77.5%	75.0%
G1+	89.7%	89.7%	86.2%	86.2%	86.2%
G1-	88.2%	78.4%	72.5%	72.5%	68.6%
G2	98.4%	96.0%	93.7%	93.7%	90.5%
G2+	99.0%	98.0%	95.0%	95.0%	92.1%
G2-	96.0%	88.0%	88.0%	88.0%	84.0%
G3	97.8%	97.8%	95.7%	95.7%	95.7%

Table 9.5. *Progression of correct answers (TRUE) for the questions A(n), where n increases from 104 for every group and subgroup.*

9.1.3 TASK B – Results

In this task it is described a situation in which, for a predicate B(n), corresponding to "The property B holds for the natural number n", it is known that:

- B(37) is false;
- It is true that $\forall n(B(n) \rightarrow B(n+1))$.

From this it follows that B(n) is necessarily false for every natural number lower or equal to 37, whereas there is not enough information to conclude about the truth or the falsity of B(n) for the numbers that are greater than 37.

<u>First question</u> – The property B is: True for all the natural numbers / False for all the natural numbers / True for some and false for some natural numbers.

The answers to the first question, divided by groups, are presented in the following tables.

Group G1	True for all the natural numbers	False for all the natural numbers	True for some and false for some natural numbers	We cannot know
G1	2.5%	22.5%	42.5%	32.5%
G1+	3.4%	17.2%	41.4%	37.9%
G1-	2.0%	25.5%	43.1%	29.4%

Table 9.6. *Percentage of answers for the first question of Task B for group G1, further divided in G1+ and G1-. The green colour indicates the mathematically correct answers.*

Group G2	True for all the natural numbers	False for all the natural numbers	True for some and false for some natural numbers	We cannot know
G2	1.6%	15.9%	35.7%	46.8%
G2+	2.0%	14.9%	29.7%	53.5%
G2-	0.0%	20.0%	60.0%	20.0%

Table 9.7. *Percentage of answers for the first question of Task B for group G2, further divided in G2+ and G2-. The green colour indicates the mathematically correct answers.*

Group G3	True for all the natural numbers	False for all the natural numbers	True for some and false for some natural numbers	We cannot know
G3	0.0%	2.2%	19.6%	78.3%

Table 9.8. *Percentage of answers for the first question of Task B for group G3. The green colour indicates the mathematically correct answers.*

<u>Numerical questions</u> – The property B(n) is: True / False / We cannot know.

The answers for **group G1,** and for **its subgroups G1+** and **G1-,** to the numerical questions (with the form B(n), for several natural numbers) are reported in the following three figures.

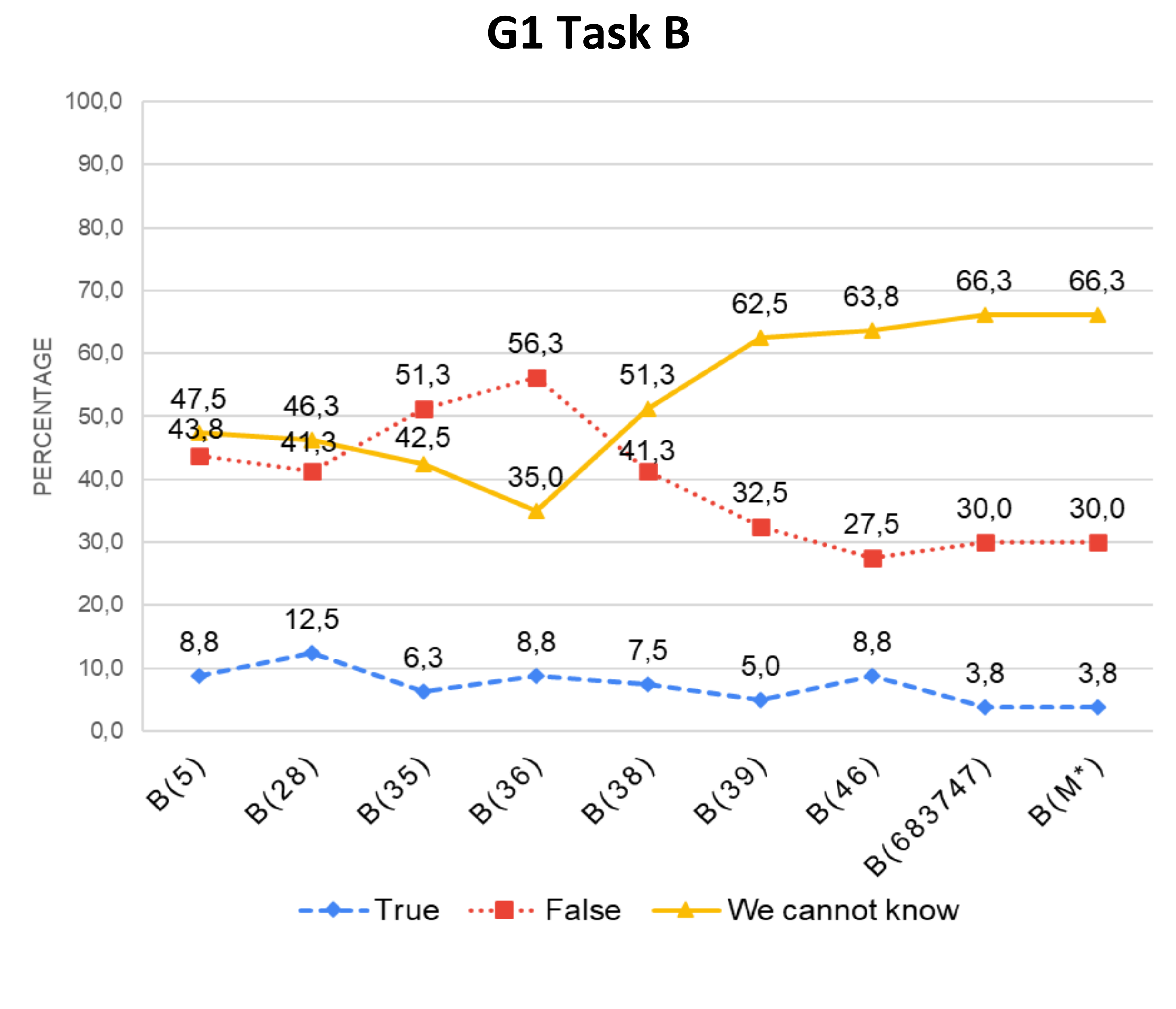


Figure 9.8. The chart shows the *percentage of answers for the questions B(n) for group G1 and the trends of these percentages when n varies.*

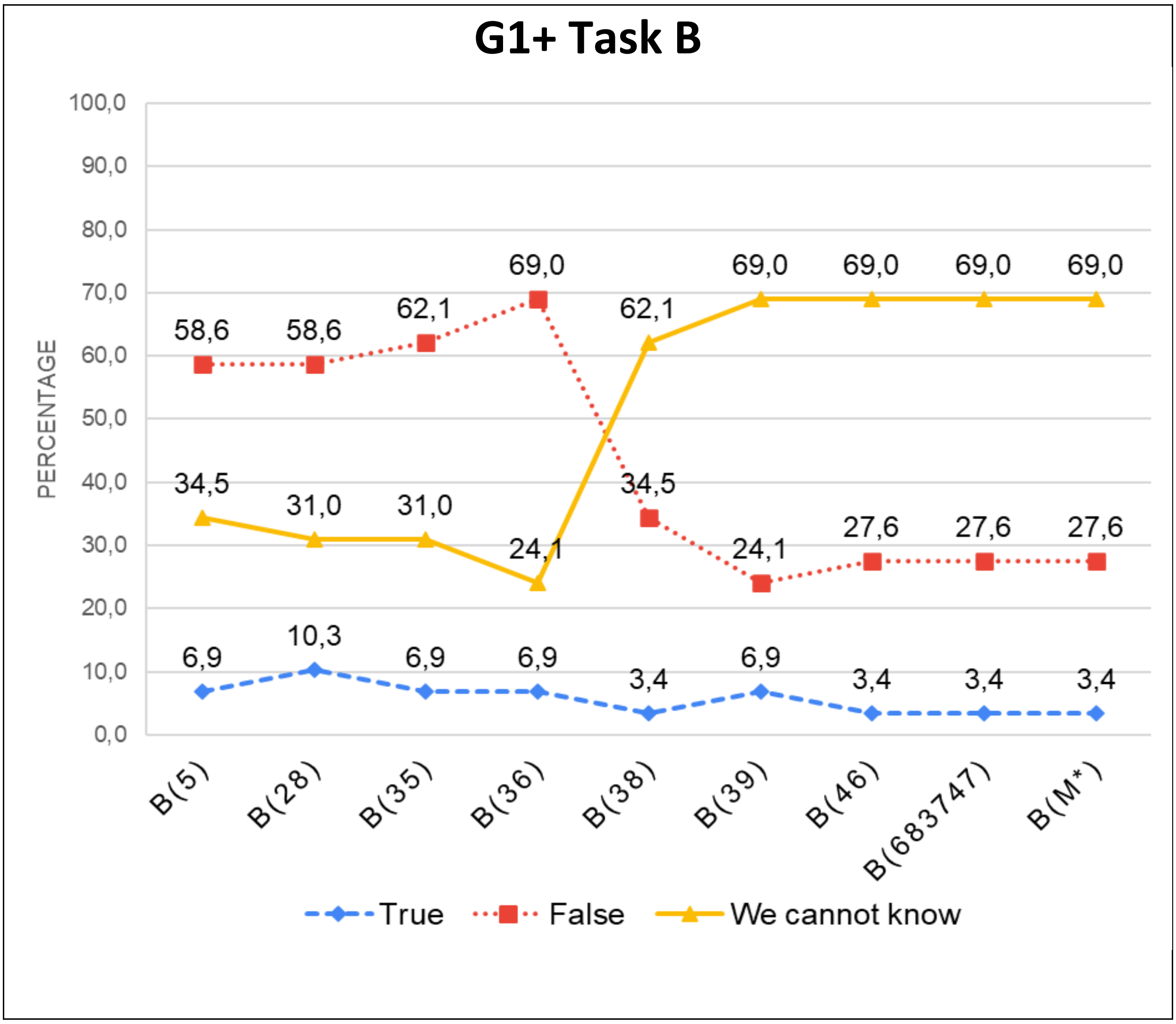


Figure 9.9. The chart shows the *percentage of answers for the questions B(n) for group G1+ and the trends of these percentages when n varies.*

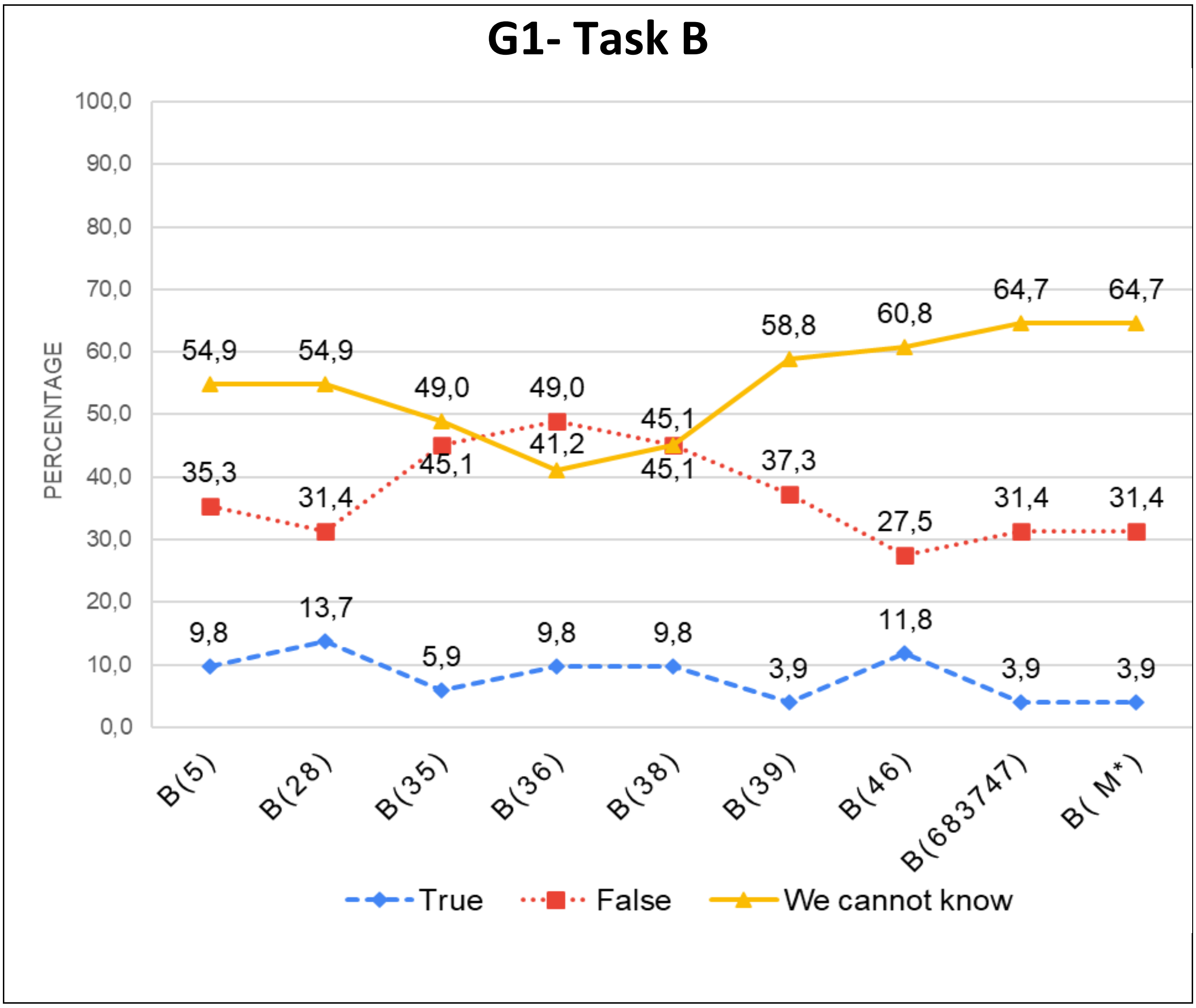


Figure 9.10. The chart shows the *percentage of answers for the questions B(n) for group G1- and the trends of these percentages when n varies.*

In the same way as for group G1, the next three figures will show the answers for this group of questions for the **group G2** and its **subgroups G2+** and **G2-.**

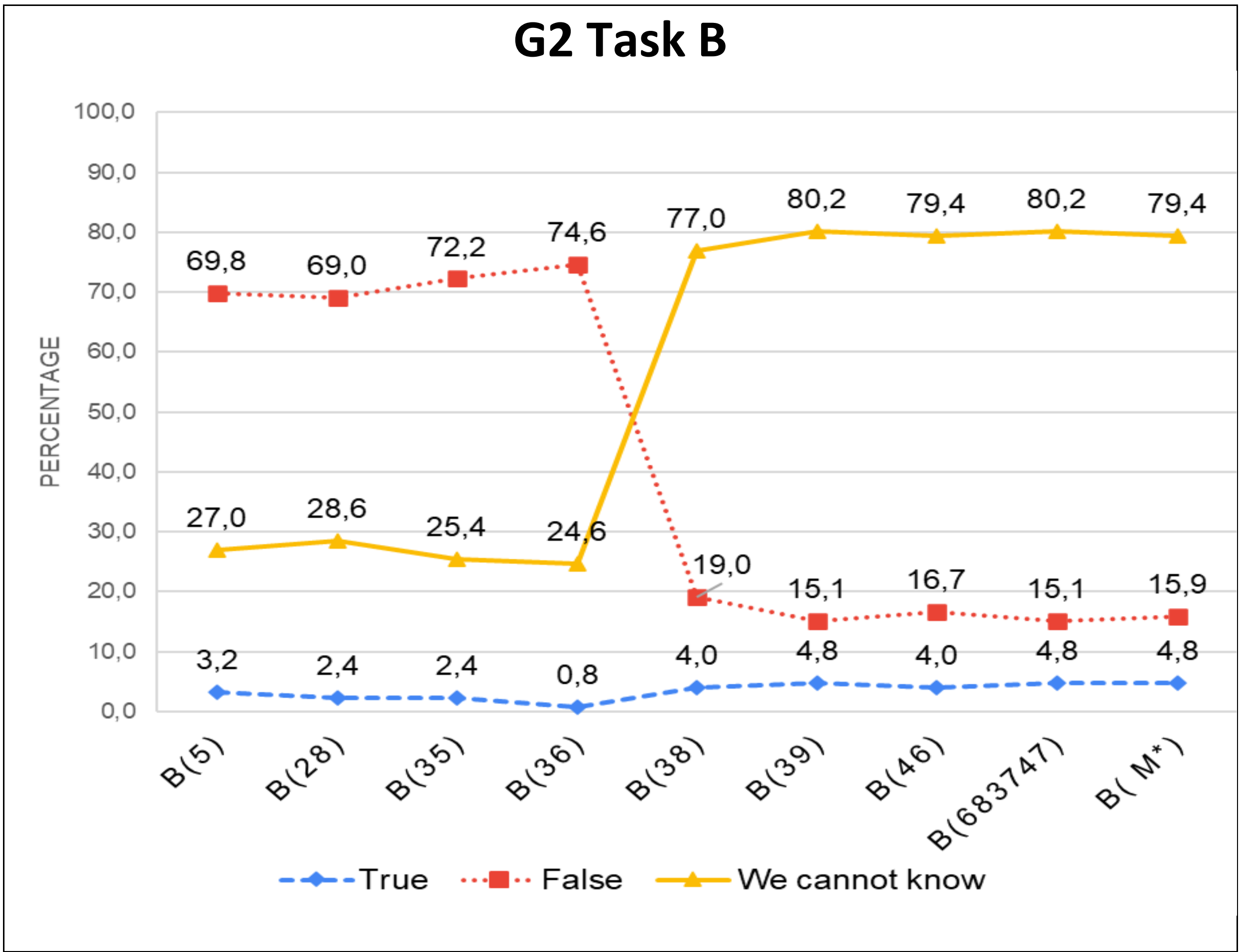


Figure 9.11. The chart shows the *percentage of answers for the questions B(n) for group G2 and the trends of these percentages when n varies.*

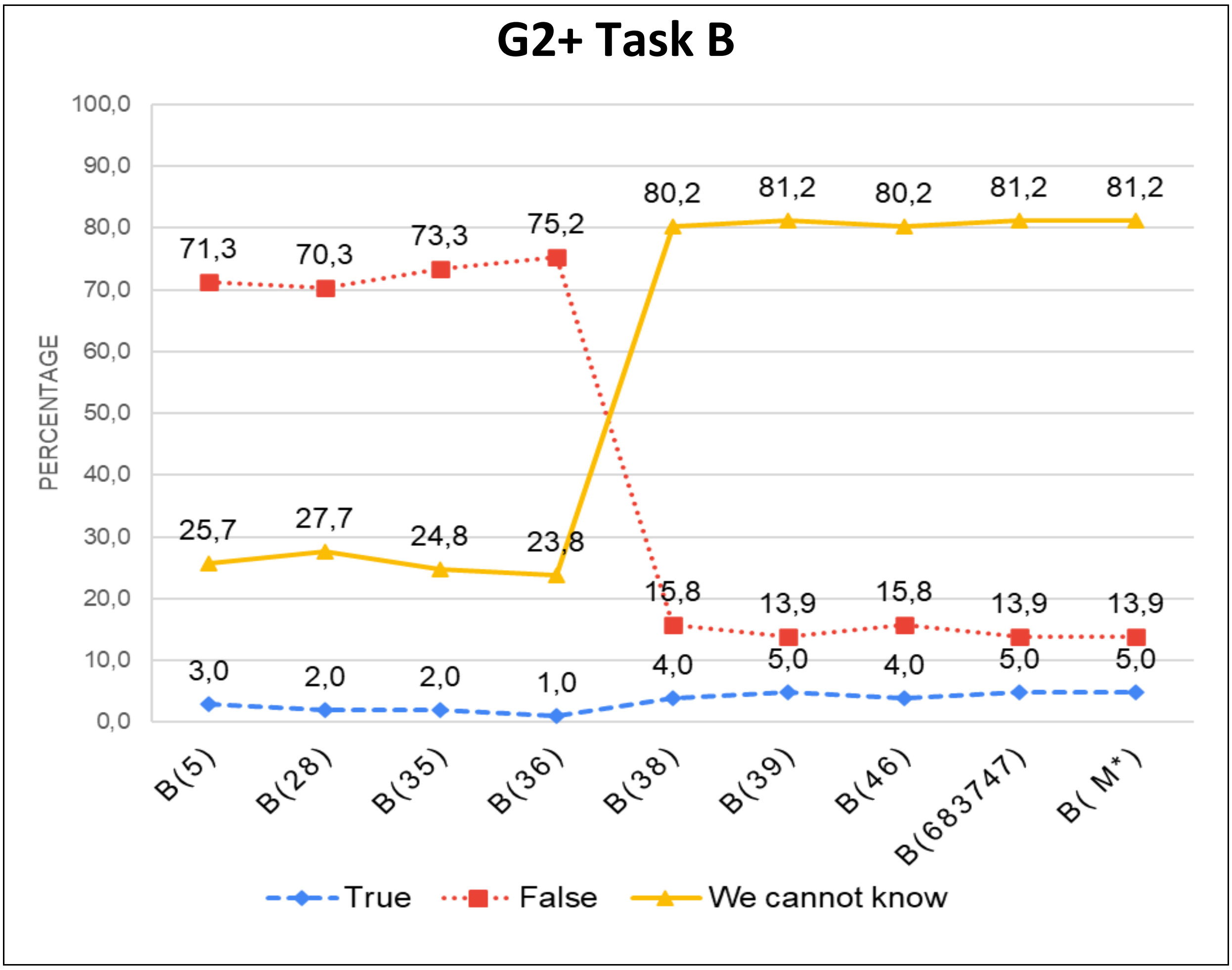


Figure 9.12. The chart shows the *percentage of answers for the questions B(n) for group G2+ and the trends of these percentages when n varies.*

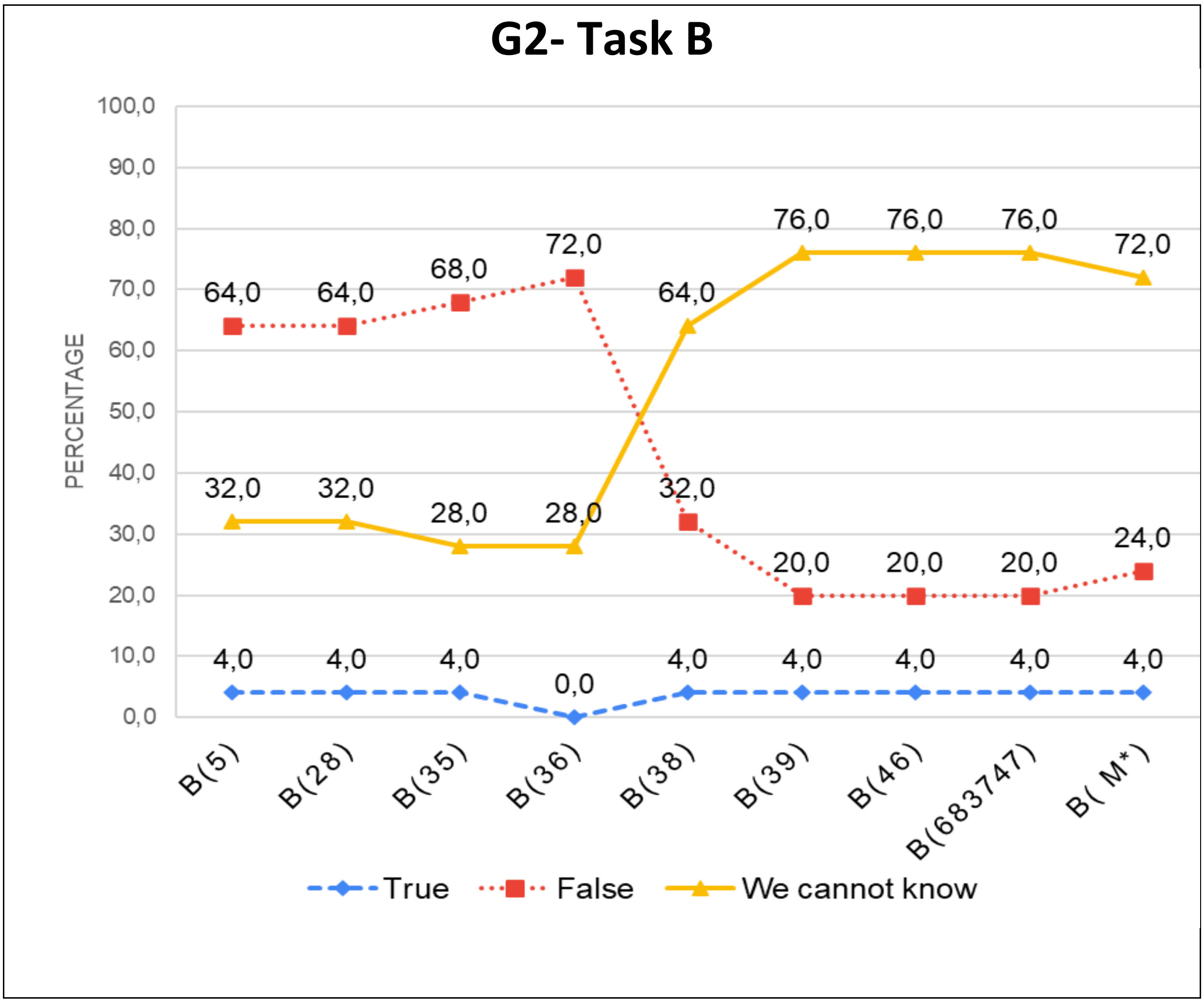


Figure 9.13. The chart shows the *percentage of answers for the questions B(n) for group G2- and the trends of these percentages when n varies.*

Finally, the following figure contains the answers of this group of questions for the **Group G3.**

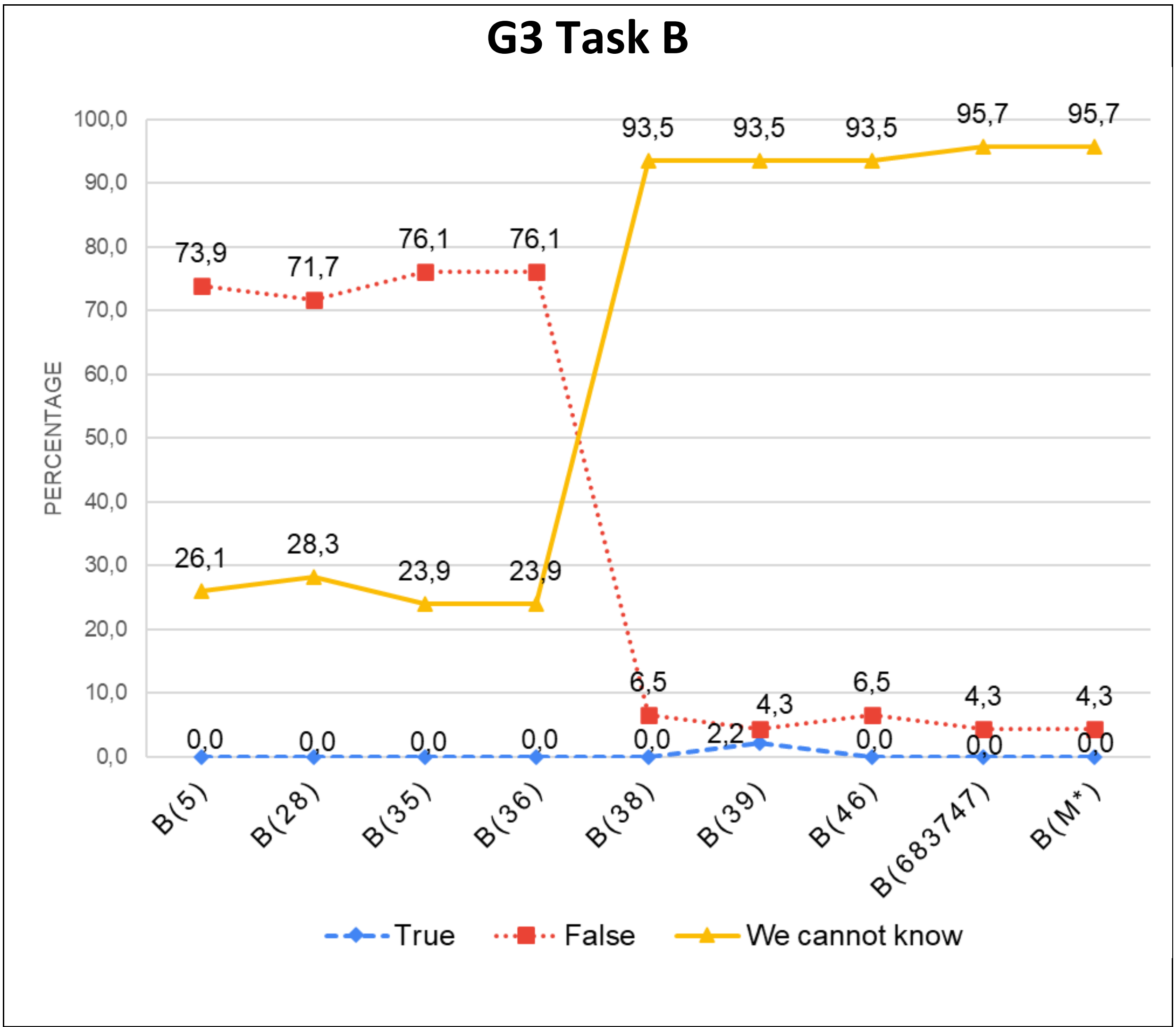


Figure 9.14. The chart shows the *percentage of answers for the questions B(n) for group G3 and the trends of these percentages when n varies.*

For Task B, similarly to what I have done for Task A, I also present the data corresponding to the **progression** of correct answers (FALSE) for the questions involving numbers increasingly lower than 37. In other terms, I show the percentages of students who answered FALSE at every question from B(36) to some B(n) with n≤36. These data, for **every group and subgroup**, are shown in the following chart.

Group	**B(36)**	**B(36) & B(35)**	**B(36) & B(35) & B(28)**	**B(36) & B(35) & B(28) & B(5)**
G1	56.3%	47.5%	41.3%	41.3%
G1+	69.0%	62.1%	58.6%	58.6%
G1-	49.0%	39.2%	31.4%	31.4%
G2	74.6%	71.4%	68.3%	67.5%
G2+	75.2%	72.3%	69.3%	68.3%
G2-	72.0%	68.0%	64.0%	64.0%
G3	76.1%	76.1%	71.7%	71.7%

Table 9.9. *Progression of correct answers (FALSE) for the questions B(n), where n decreases from 37 for every group and subgroup.*

<u>Last questions</u> – The theorem proved by Bertrand's colleague (the inductive step) is not correct for n=37 or for n=36.

The last two questions of the Task B asked to agree or disagree with the statement claiming that the inductive step (namely, the theorem of Bertrand's colleague) is not valid for n=37 and with the statement claiming the same thing but for n=36. More specifically the two statements were:

- "Since it is known that Bertrand's property is false for the number 37, Bertrand's colleague has **SURELY** mistaken his proof, because it is **NOT** true that taken any natural number (for instance, **the number 37**), if for this number Bertrand's property is true, then Bertrand's property is also true for its consecutive natural number."
- "Since it is known that Bertrand's property is false for the number 37, Bertrand's colleague has **SURELY** mistaken his proof, because it is **NOT** true that taken any natural number (for instance, **the number 36**), if for this number Bertrand's property is true, then Bertrand's property is also true for its consecutive natural number."

In other terms, the two questions asked to evaluate the validity, respectively, of the implication B(37)→B(38) and of the implication B(36)→B(37), knowing that B(37) is false. As observed in the a-priori analysis of the survey in the 'Methods' section, the two statements are both invalid. The fact that it exists a number for which Bertrand's property is false does not imply that the inductive step, proved by Bertrand's colleague, cannot be valid for every natural number. The two questions, thus, aimed to investigate to what extent the validity of the implication in the inductive step is intuitively accepted, independently from knowing the truth value of antecedent or consequent.

The following three tables contain the result for these two questions for **group G1** and its **subgroups G1+ and G1-.**

G1	B(n)→B(n+1) is false for n=37	B(n)→B(n+1) is false for n=36
I agree	36.3%	47.5%
I do not agree	63.8%	52.5%

Table 9.10. *Percentage of answers for the last questions of the Task B for the group G1. The green colour indicates the mathematically correct answers.*

G1+	B(n)→B(n+1) is false for n=37	B(n)→B(n+1) is false for n=36
I agree	27.6%	37.9%
I do not agree	72.4%	62.1%

Table 9.11. *Percentage of answers for the last questions of the Task B for the group G1+. The green colour indicates the mathematically correct answers.*

G1-	B(n)→B(n+1) is false for n=37	B(n)→B(n+1) is false for n=36
I agree	41.2%	52.9%
I do not agree	58.8%	47.1%

Table 9.12. *Percentage of answers for the last questions of the Task B for the group G1-. The green colour indicates the mathematically correct answers.*

The following three tables contain the result for these last two questions for **group G2** and its **subgroups G2+ and G2-.**

G2	B(n)→B(n+1) is false for n=37	B(n)→B(n+1) is false for n=36
I agree	23.0%	31.7%
I do not agree	77.0%	68.3%

Table 9.13. *Percentage of answers for the last questions of the Task B for the group G2. The green colour indicates the mathematically correct answers.*

G2+	B(n)→B(n+1) is false for n=37	B(n)→B(n+1) is false for n=36
I agree	17.8%	23.8%
I do not agree	82.2%	76.2%

Table 9.14. *Percentage of answers for the last questions of the Task B for the group G2+. The green colour indicates the mathematically correct answers.*

G2-	B(n)→B(n+1) is false for n=37	B(n)→B(n+1) is false for n=36
I agree	44.0%	64.0%
I do not agree	56.0%	36.0%

Table 9.15. *Percentage of answers for the last questions of the Task B for the group G2-. The green colour indicates the mathematically correct answers.*

Finally, the following table shows the results for these questions for **group G3**

G3	B(n)→B(n+1) is false for n=37	B(n)→B(n+1) is false for n=36
I agree	13.0%	10.9%
I do not agree	87.0%	89.1%

Table 9.16. *Percentage of answers for the last questions of the Task B for the group G3. The green colour indicates the mathematically correct answers.*

9.1.4 TASK C – Results

In this task it is described a situation in which, for a predicate C(n), corresponding to "The property C holds for the natural number n", it is known that:

- C(75) is true and C(45) is false;
- It is true that $\forall n(C(n) \rightarrow C(n+1))$.

From this it follows that C(n) is necessarily false for every natural number greater or equal to 75, that C(n) is false for every natural number lower or equal to 45, whereas there is not enough information to conclude about the truth or the falsity of C(n) for the numbers strictly between 45 and 75.

<u>First question</u> – Without considering the number 45 and 75, the property C is: True for all the natural numbers / False for all the natural numbers / True for some and false for some natural numbers.

The answers to the first question, divided by groups, are presented in the following tables.

Group G1	True for all the natural numbers	False for all the natural numbers	True for some and false for some natural numbers	We cannot know
G1	5.0%	2.5%	63.8%	28.8%
G1+	3.4%	0.0%	75.9%	20.7%
G1-	5.9%	3.9%	56.9%	33.3%

Table 9.17. *Percentage of answers for the first question of Task C for group G1, further divided in G1+ and G1-. The green colour indicates the mathematically correct answers*

Group G2	True for all the natural numbers	False for all the natural numbers	True for some and false for some natural numbers	We cannot know
G2	3.2%	1.6%	82.5%	12.7%
G2+	2.0%	2.0%	82.2%	13.9%
G2-	8.0%	0.0%	84.0%	8.0%

Table 9.18. *Percentage of answers for the first question of Task C for group G2, further divided in G2+ and G2-. The green colour indicates the mathematically correct answers.*

Group G3	True for all the natural numbers	False for all the natural numbers	True for some and false for some natural numbers	We cannot know
G3	2.2%	0.0%	82.6%	15.2%

Table 9.19. *Percentage of answers for the first question of Task C for group G3. The green colour indicates the mathematically correct answers.*

Numerical questions – The property C(n) is: True / False / We cannot know.

As done for the Task A and B, in the following figures I show the results for the numerical questions with the form C(n), for several natural numbers, **for each group and subgroup**.

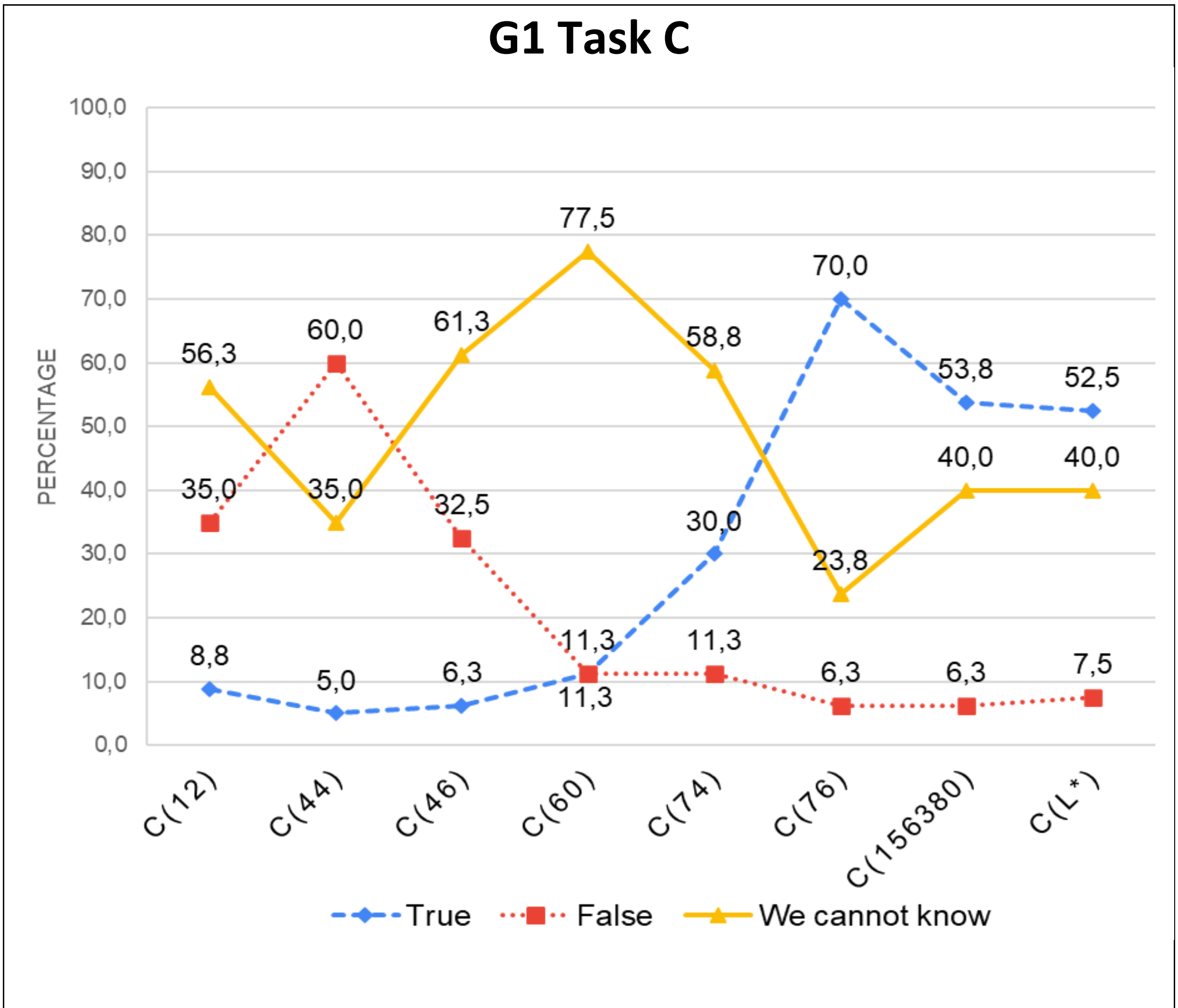


Figure 9.15. The chart shows the *percentage of answers for the questions C(n) for group G1 and the trends of these percentages when n varies.*

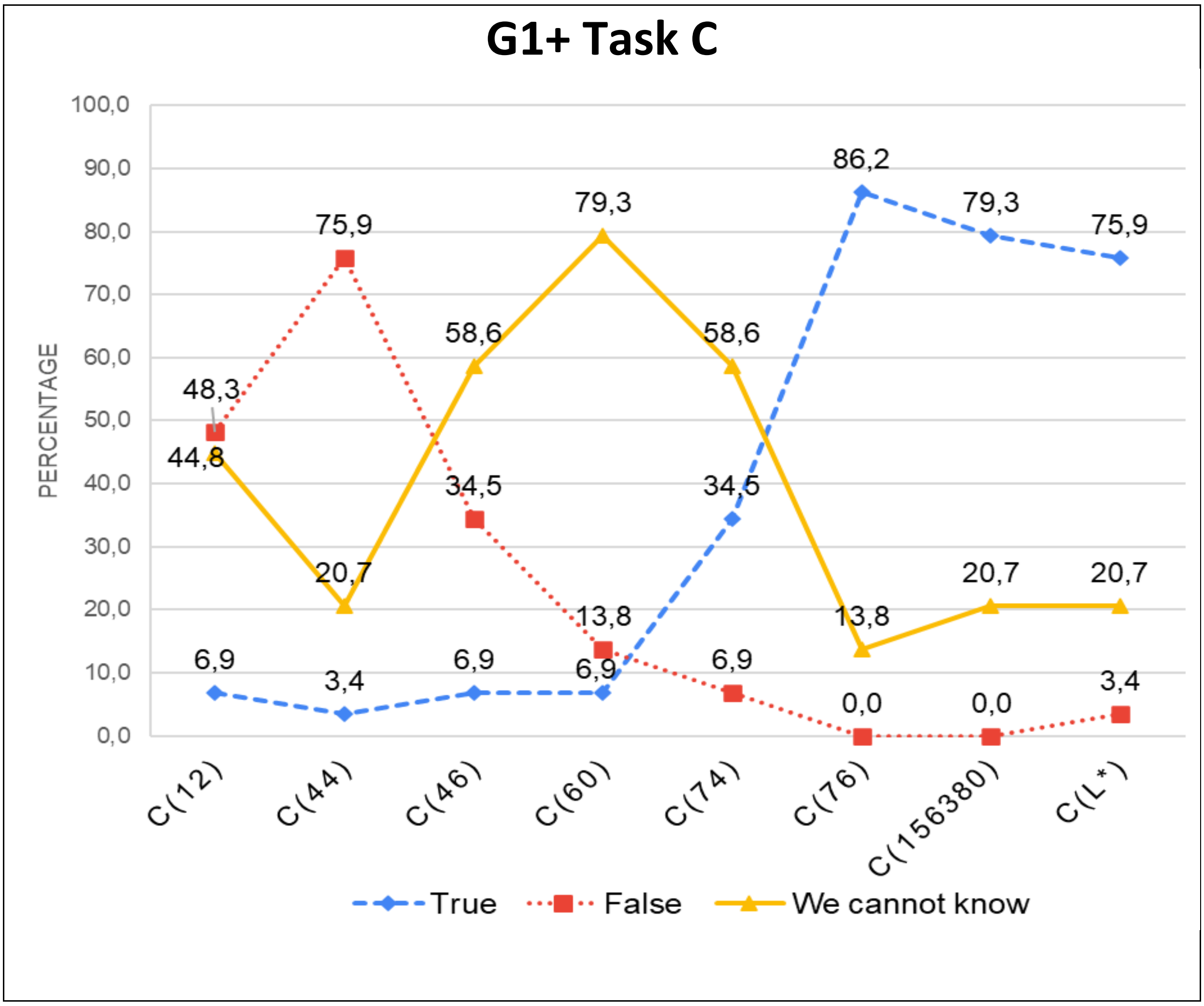


Figure 9.16. The chart shows the *percentage of answers for the questions C(n) for group G1+ and the trends of these percentages when n varies.*

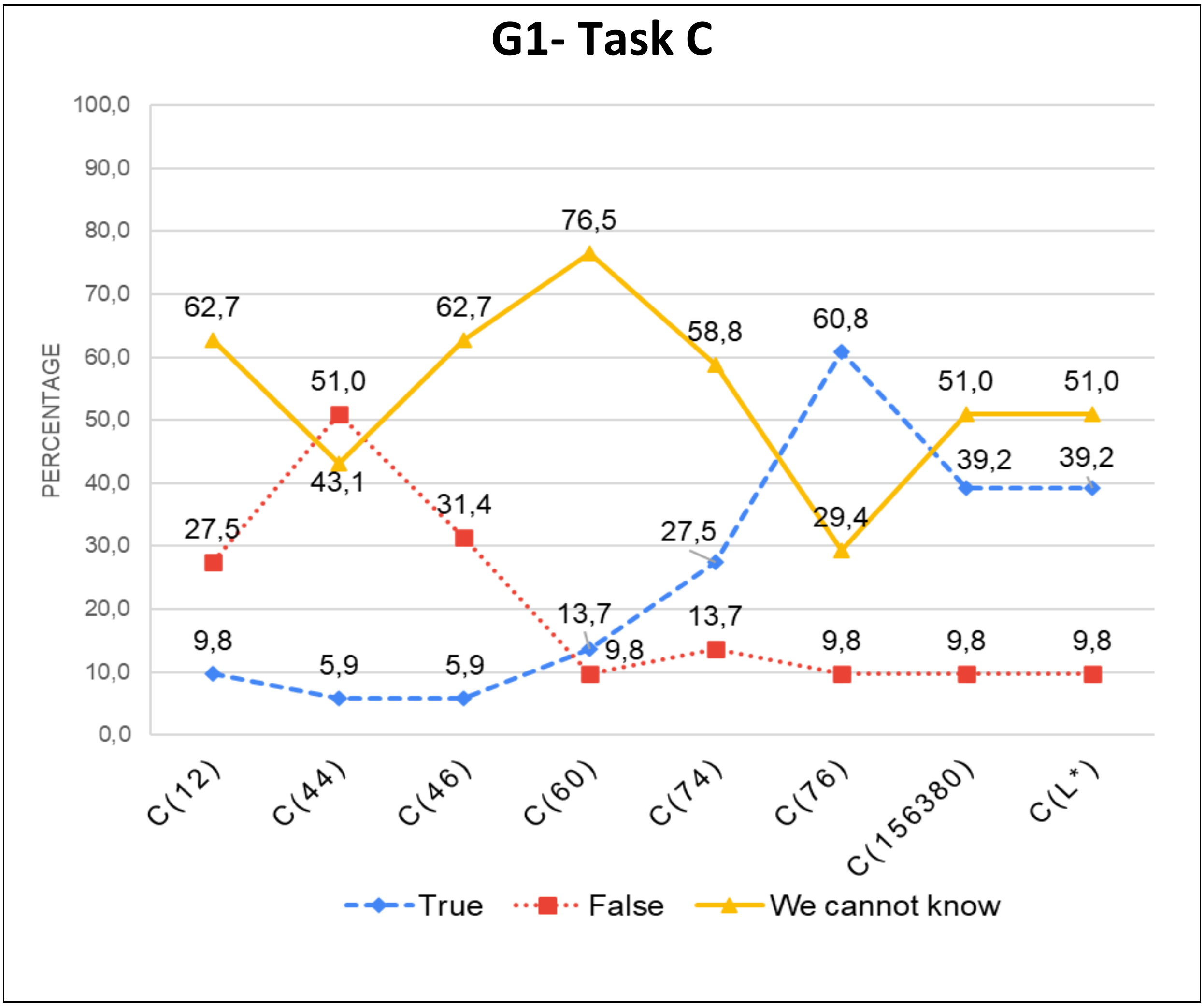


Figure 9.17. The chart shows the *percentage of answers for the questions C(n) for group G1- and the trends of these percentages when n varies.*

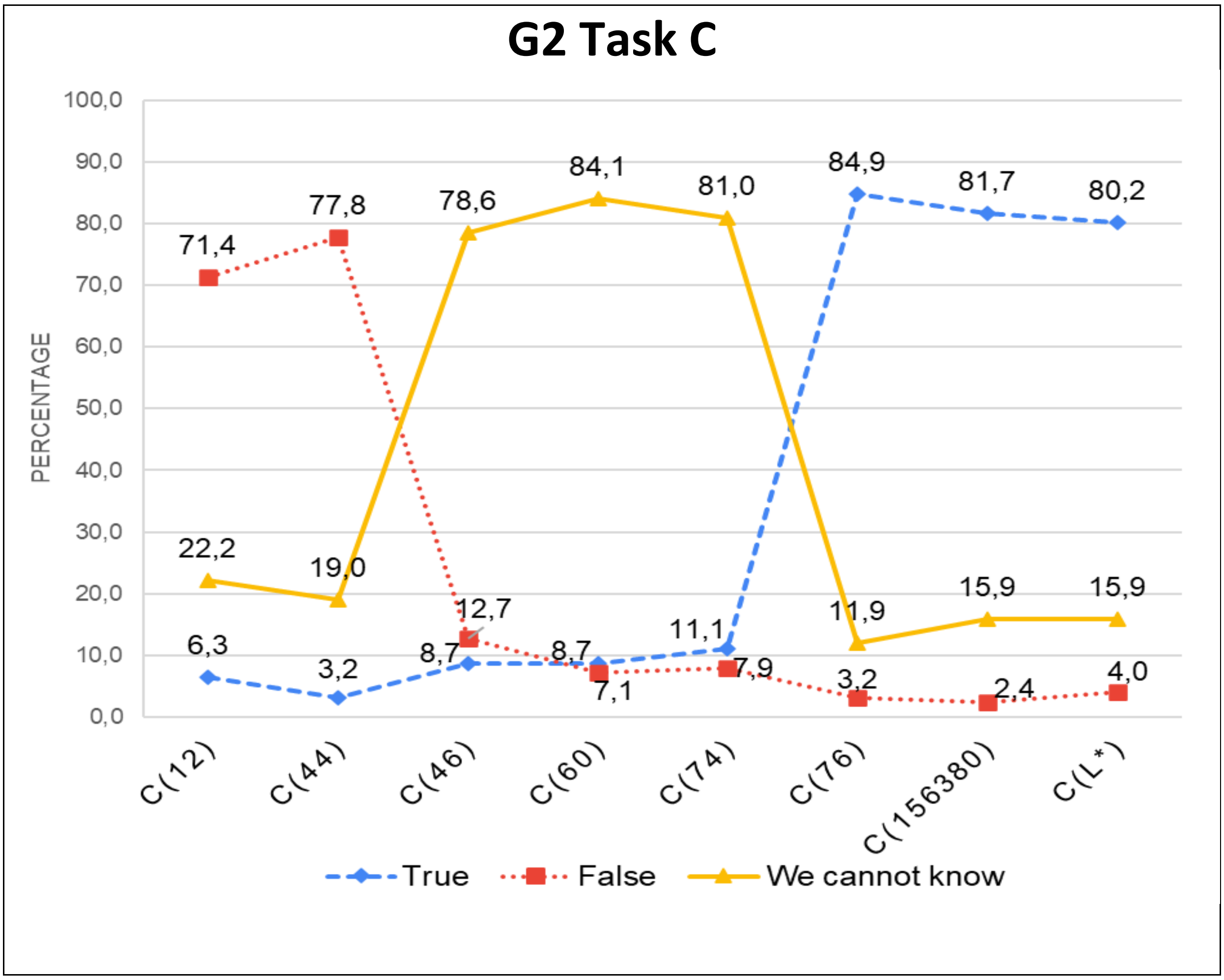


Figure 9.18. The chart shows the *percentage of answers for the questions C(n) for group G2 and the trends of these percentages when n varies.*

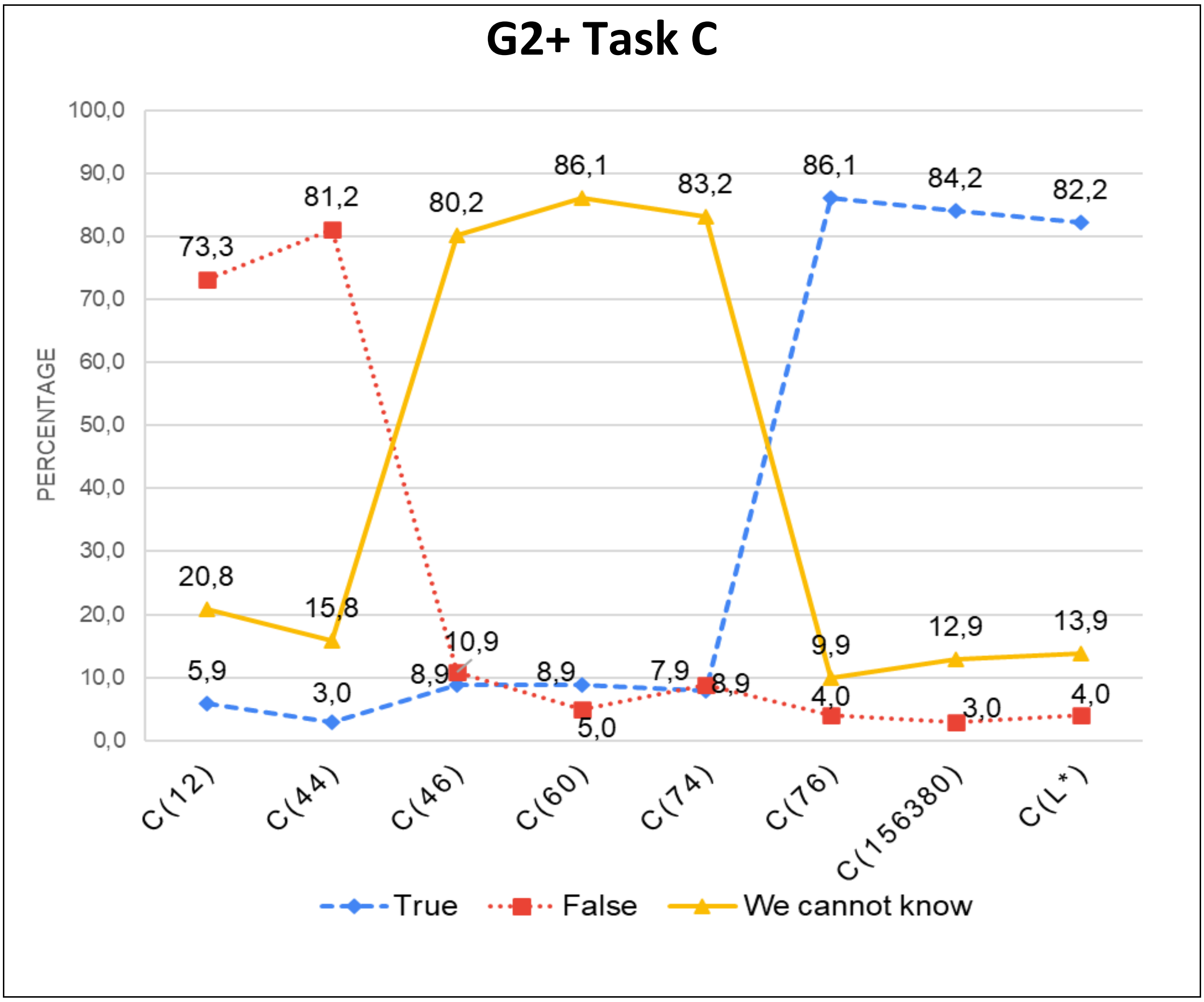


Figure 9.19. The chart shows the *percentage of answers for the questions C(n) for group G2+ and the trends of these percentages when n varies.*

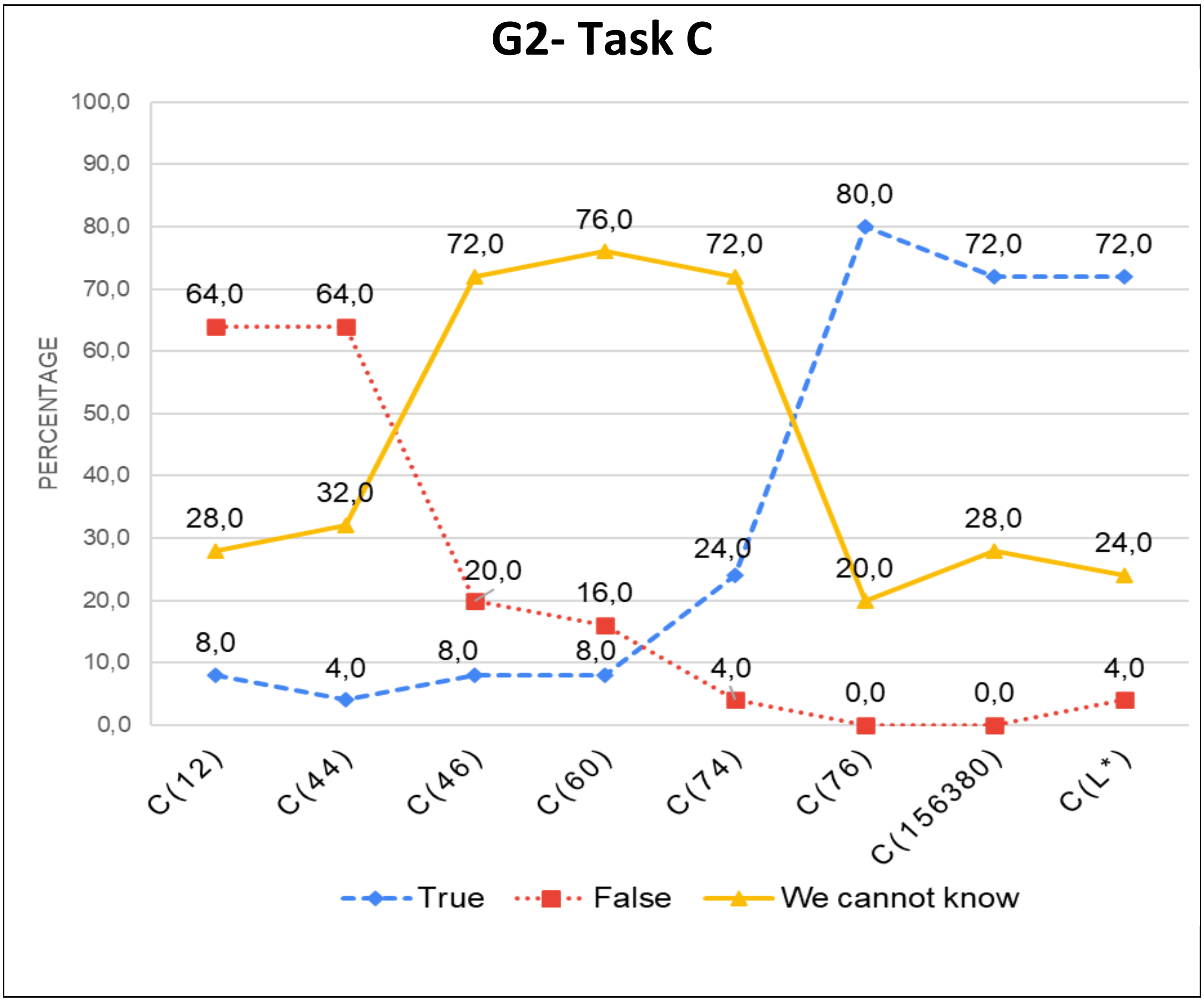


Figure 9.20. The chart shows the *percentage of answers for the questions C(n) for group G2- and the trends of these percentages when n varies.*

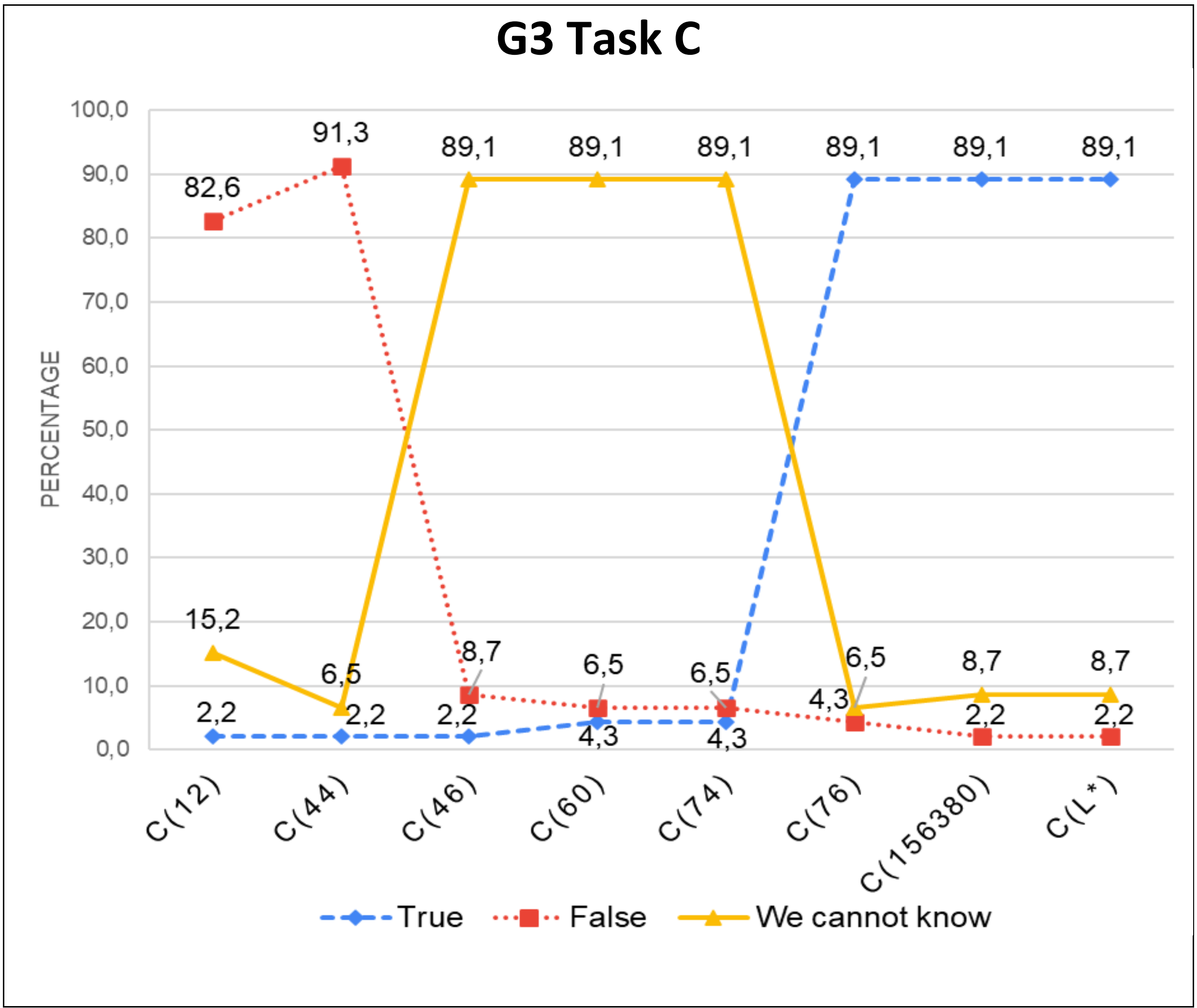


Figure 9.21. The chart shows the *percentage of answers for the questions C(n) for group G3 and the trends of these percentages when n varies.*

Finally, as done for Task A and B, from the overall results of Task C it is possible to extract the data corresponding both to the **progression** of correct answers (TRUE) for the questions involving numbers increasingly greater than 75 and the **progression** of correct answers (FALSE) for the question involving numbers increasingly lower than 45. In other terms, we show the percentages of students who answered TRUE for every question from C(76) to some C(n) with n≥75 and the percentages of students who answered FALSE for every question from C(44) to some C(n) with n≤45. These data, for **every group and subgroup**, are shown in the following table.

Group	**True for: C(76)**	**True for: C(76) & C(156380)**	**True for: C(76) & C(156380) & C(L*)**	**False for: C(44)**	**False for: C(44) & C(12)**
G1	70.0%	52.5%	47.5%	60.0%	35.0%
G1+	86.2%	79.3%	72.4%	75.9%	48.3%
G1-	60.8%	37.3%	33.3%	51.0%	27.5%
G2	84.9%	81.7%	80.2%	77.8%	70.6%
G2+	86.1%	84.2%	82.2%	81.2%	73.3%
G2-	80.0%	72.0%	72.0%	64.0%	60.0%
G3	89.1%	89.1%	89.1%	91.3%	82.6%

Table 9.20. *Progression of correct answers (TRUE) for the questions C(n), whereas n increases from 75, on the left, and progression of correct answers (FALSE) for the questions C(n), where n decreases from 45, on the right. All groups and subgroups are included.*

9.1.5 TASK D – Results

In this task it is described a situation in which, for a predicate E(n), corresponding to "The property E holds for the natural number n", it is know that:

- It is true that $\forall n(E(n) \rightarrow E(n+1))$.

Then it was asked to indicate if some statements about the property E were CERTAIN (C), POSSIBLE (P), or IMPOSSIBLE (I).

In presenting the results, as done in the Methods section where the task has been described, I will use the following labels to indicate each of the statements involved.

Label	Statement
T	Evelin's property is TRUE for all the natural numbers.
F	Evelin's property is FALSE for all the natural numbers.
FTFT…	Evelin's property is FALSE for all the even natural numbers and at the same time is TRUE for all the odd natural numbers.
TTFF…	Evelin's property is TRUE only every other couple of natural numbers. That is, for instance, that it is TRUE for the numbers 1 and 2, FALSE for the numbers 3 and 4, TRUE for the numbers 5 and 6, and so on.
40T-41T	Evelin's property is TRUE for the number 40 and at the same time TRUE for the number 41.
40T-41F	Evelin's property is TRUE for the number 40 and at the same time FALSE for the number 41.
40F-41T	Evelin's property is FALSE for the number 40 and at the same time TRUE for the number 41.
40F-41F	Evelin's property is FALSE for the number 40 and at the same time FALSE for the number 41.
23F-49T	Evelin's property is FALSE for the number 23 and at the same time TRUE for the number 49.
12T-54F	Evelin's property is TRUE for the number 12 and at the same time FALSE for the number 54.
ST-BF	Evelin's property is TRUE for small numbers and at the same time FALSE for the number with millions of digits.
SF-BT	Evelin's property is FALSE for small numbers and at the same time TRUE for the number with millions of digits.

The next three tables contain the different percentage of answers to these questions for **each group** and **each subgroup.**

Question	Group	Certain (%)	Possible (%)	Impossible (%)
T	G1	8.8	83.8	7.5
	G1+	3.4	93.1	3.4
	G1-	11.8	78.4	9.8
F	G1	2.5	70.0	27.5
	G1+	6.9	69.0	24.1
	G1-	0.0	70.6	29.4
FTFT…	G1	3.8	35.0	61.3
	G1+	3.4	34.5	62.1
	G1-	3.9	35.3	60.8
TTFF…	G1	2.5	42.5	55.0
	G1+	3.4	37.9	58.6
	G1-	2.0	45.1	52.9
40T-41T	G1	25.0	70.0	5.0
	G1+	34.5	65.5	0.0
	G1-	19.6	72.5	7.8
40T-41F	G1	6.3	27.5	66.3
	G1+	10.3	17.2	72.4
	G1-	3.9	33.3	62.7
40F-41T	G1	5.0	67.5	27.5
	G1+	6.9	75.9	17.2
	G1-	3.9	62.7	33.3
40F-41F	G1	10.0	80.0	10.0
	G1+	13.8	86.2	0.0
	G1-	7.8	76.5	15.7
23F-49T	G1	2.5	81.3	16.3
	G1+	0.0	100.0	0.0
	G1-	3.9	70.6	25.5
12T-54F	G1	3.8	50.0	46.3
	G1+	10.3	34.5	55.2
	G1-	0.0	58.8	41.2
ST-BF	G1	3.8	36.3	60.0
	G1+	3.4	27.6	69.0
	G1-	3.9	41.2	54.9
SF-BT	G1	5.0	67.5	27.5
	G1+	6.9	79.3	13.8
	G1-	3.9	60.8	35.3

Table 9.21. *Percentage of answers for the Task D for the group G1, further divided in G1+ and G1-. The green colour indicates the mathematically correct answers.*

Question	Group	Certain (%)	Possible (%)	Impossible (%)
T	G2	4.8	92.9	2.4
	G2+	2.0	96.0	2.0
	G2-	16.0	80.0	4.0
F	G2	1.6	81.7	16.7
	G2+	1.0	85.1	13.9
	G2-	4.0	68.0	28.0
FTFT…	G2	1.6	21.4	77.0
	G2+	1.0	18.8	80.2
	G2-	4.0	32.0	64.0
TTFF…	G2	3.2	15.9	81.0
	G2+	2.0	13.9	84.2
	G2-	8.0	24.0	68.0
40T-41T	G2	23.0	74.6	2.4
	G2+	20.8	79.2	0.0
	G2-	32.0	56.0	12.0
40T-41F	G2	1.6	7.9	90.5
	G2+	1.0	5.9	93.1
	G2-	4.0	16.0	80.0
40F-41T	G2	4.0	77.8	18.3
	G2+	5.0	78.2	16.8
	G2-	0.0	76.0	24.0
40F-41F	G2	8.7	89.7	1.6
	G2+	9.9	89.1	1.0
	G2-	4.0	92.0	4.0
23F-49T	G2	2.4	84.1	13.5
	G2+	3.0	88.1	8.9
	G2-	0.0	68.0	32.0
12T-54F	G2	0.0	19.8	80.2
	G2+	0.0	17.8	82.2
	G2-	0.0	28.0	72.0
ST-BF	G2	0.8	17.5	81.7
	G2+	0.0	14.9	85.1
	G2-	4.0	28.0	68.0
SF-BT	G2	0.8	84.1	15.1
	G2+	1.0	86.1	12.9
	G2-	0.0	76.0	24.0

Table 9.22. *Percentage of answers for the Task D for the group G2, further divided in G2+ and G2-. The green colour indicates the mathematically correct answers.*

Question	Group	Certain (%)	Possible (%)	Impossible (%)
T	G3	2.2	97.8	0.0
F	G3	0.0	93.5	6.5
FTFT…	G3	0.0	13.0	87.0
TTFF…	G3	0.0	10.9	89.1
40T-41T	G3	10.9	87.0	2.2
40T-41F	G3	0.0	13.0	87.0
40F-41T	G3	0.0	93.5	6.5
40F-41F	G3	2.2	97.8	0.0
23F-49T	G3	0.0	97.8	2.2
12T-54F	G3	2.2	13.0	84.8
ST-BF	G3	0.0	15.2	84.8
SF-BT	G3	0.0	93.5	6.5

Table 9.23. *Percentage of answers for the Task D for the group G3. The green colour indicates the mathematically correct answers.*

9.1.6 Other results

The last section of the survey, called 'Conclusive questions', was mainly composed by open answer questions, that will be analysed later. However, there was also a multiple-choice question investigating on the inductive base. In particular, it was asked to select 'True', 'False', or 'I do not know' in response to the following statement:

In the questions on Aarney's, Betrand's, Coleman's, and Evelin's properties it was not possible to state anything certain because the proof of the base case was missing.

The results for this question are shown in the following table. The question was posed only to the participants that in a previous question indicated that "the proof by mathematical induction has been object of study" for them, which are the students of the groups G1+, G2+, and G3.

Group	True (%)	False (%)	I do not know (%)
G1+	31.0	31.0	37.9
G2+	29.7	58.4	11.9
G3	34.8	45.7	19.6

Table 9.24. *Percentage of answers for the question on the base case in the section 'Conclusive Questions' of the survey. The green colour indicates the mathematically correct answers.*

Finally, I present the data corresponding to the percentage of students, for each group and subgroup, who answered correctly to **every question** in Task A, B, C, or D. For the first three Tasks I have considered only the numerical questions, i.e. the ones with the form A(n), B(n), or C(n) with n an assigned natural number. The following table contains these results.

Group	All correct answers for the numerical questions A(n)	All correct answers for the numerical questions B(n)	All correct answers for the numerical questions C(n)	All correct answers for the questions of Task D	All correct answers for the previous questions
G1	35.0%	17.5%	17.5%	11.3%	3.8%
G1+	55.2%	31.0%	27.6%	10.3%	6.9%
G1-	23.5%	9.8%	11.8%	11.8%	2.0%
G2	65.1%	46.8%	55.6%	42.9%	31.0%
G2+	70.3%	49.5%	59.4%	47.5%	33.7%
G2-	44.0%	36.0%	40.0%	24.0%	20.0%
G3	82.6%	65.2%	76.1%	67.4%	50.0%

Table 9.25. *Percentage of correct answers for all the numerical questions of TA, TB, TC, and for all the questions of TD. The last column shows the percentage of correct answers for all these questions together.*

9.2 ANALYSIS AND DISCUSSION OF THE RESULTS

9.2.1 Task A and Task B – Numerical questions

One of the aims of these first two tasks was to investigate the (re-)construction[80] of the Explain induction Process, both in direct (Task A) and indirect form (Task B). More precisely, with the numerical questions in Task A, we were interested to observe whether the students, starting from a base case true (A(103)) and the inductive step, concluded the truth of A(n) when n≥103. Similarly, with the numerical questions in Task B, we were investigating whether the students concluded the falsity of B(n) when n≤37, given that B(37) is false and the validity of the inductive step.

Results from Task A to these questions show an interesting phenomenon. As presented in the Table 4, **we can observe a decreasing trend in the progression of correct answers when the numbers involved increase from 103**. In particular, from these progressions, we can see that some students answered "True" only for A(104), others answered "True" for A(105) too, but not for greater numbers, and others still selected "True" for A(104), A(105), A(112), and A(542384), but not for A(N*). This phenomenon is registered in every group and subgroup considered. There are some numerical differences if we compare the groups, in particular the most experienced students and the ones who had encountered MI during their studies obtained, overall, higher percentages of correct answer. **However, the same decreasing trend can be registered in every group and subgroup.** To further highlight this point, Table 9.26 of below shows how the percentage of "True" answers decreases if we consider the first question A(104) only, or all the questions from A(104) to A(N*).

Group	A(104)	A(104) & … & A(N*)	Difference
G1	88.8%	75.0%	-13.8%
G1+	89.7%	86.2%	-3.5%
G1-	88.2%	68.6%	-19.6%
G2	98.4%	90.5%	-7.9%
G2+	99.0%	92.1%	-6.9%
G2-	96.0%	84.0%	-12.0%
G3	97.8%	95.7%	-2.1%

Table 9.26. *Percentage of correct answer (TRUE) for A(104) only, and for all the questions A(n) with n≥103. The last column on the right contains the difference between the previous two columns.*

Let us observe how the APOS Theory framework allows to provide a possible interpretation for this phenomenon.

In the Genetic Decomposition of MI, when describing the Explain Induction Process, we observed that through this process a subject constructs a series of Modus Ponens (MP), starting from the base $P(n_0)$ and potentially reaching any $P(n^*)$ with $n^*\geq n_0$. This requires a *coordination* of successive processes: once the first MP is applied, from $P(n_0)$ and $P(n_0)\rightarrow P(n_0+1)$, we obtain $P(n_0+1)$; at this point the second MP must be applied from the result of the previous one to conclude that $P(n_0+2)$ holds. In the same way, the next MP is then applied, and so on. In terms of processes, each subsequent MP necessitates coordination with the previous one. Finally, if this coordination is interiorized into a new process, the subject can conceive that it is possible to construct a chain of MP which reaches any natural number $n>n_0$. For this last point, the interiorization is necessary, because the subject needs to be able to imagine carrying out the steps and eventually to skip some of them. I will call **Chaining of Modus Ponens** this successive coordination of several MP within the Explain Induction Process. In APOS terms, the interiorization of the

[80] The term is here used in APOS sense, as described in section 7.2.

Chaining of MP process could be not immediate in a subject. First of all, in fact, a subject need to recognise that after the first MP, from $P(n_0)$ to $P(n_0+1)$, a second one could be applied from $P(n_0+1)$ to $P(n_0+2)$. In particular, the subject needs to apply a second MP starting from the result of his/her previous application of a MP. In this way, the second MP is *chained* to the first one. When this happens, we can say that the subject has constructed a Chaining of MP as an action, by explicitly performing two MP inferences, chained together. Similar actions of chaining a successive MP to the previous one can be constructed to reach successive numbers. However, it is possible that at a certain point the subject starts perceiving the Chaining of MP as a process, which does not necessitate to be explicitly performed but which can be imagined, eventually skipping some steps, to reach bigger numbers, even very distant from n_0. When the interiorization of the Chaining of MP as a process is complete, the subject recognises that with a finite number of steps (MP) from n_0 it is possible to reach any given number $n^*>n_0$.

At this point, with the just described model, the different categories of answers registered for Task A and presented above can be interpreted in terms of Chaining of MP:

a) The students who answered "True" only for A(104) but not for greater numbers, do not seem to have constructed a chaining of MP, not even as an action, because a second MP is not applied from the first one. They only applied one single MP.

b) The students who answered "True" for A(104), for A(105) but not for greater numbers may have constructed a chaining of MP as an action (they chained two successive MP), but not as a process since they correctly answered only to the questions which require a small number of steps from 103. It seems they did not imagine carrying out a higher number of MP steps.

c) Finally, the third category of answers, the ones of students who selected "True" for every A(n) with n≥103 except for A(N*) could be interpreted as a still not complete interiorization of the Chaining of MP process. This subjects, in fact, seem to be able to imagine the process to reach the previous numbers, but not to reach N*.[81]

Results from Task B show an analogous phenomenon. As presented in the Table 9.9, if we look at the percentage of correct answers ("False") for B(n) when n≤37, **we can observe a decreasing progression when the numbers involved decrease from 37.** In particular, there are students who answered "False" only for B(36), others still who answered that also B(35) is false, but not for lower numbers and some who selected "False" until B(5). As it happened for the Task A, there are some numerical differences, but **the trend is the same in every group and subgroups considered**. To further highlight this point, in Table 9.27 of below, I report how the percentage of "False" answers decreases if we consider the first question B(36) only, or all the questions from B(36) to B(5).

[81] This is of course only one possible interpretation for this last category of answers. The particular form of the number N*, a very 'uncommon' and big natural number, may have affected negatively the answers. For instance, in Relaford-Doyle & Nunez, 2018, that I already discussed in section 4.4.2, it is shown that even if from a mathematical point of view there are no differences in the "naturalness" of natural numbers, however from a cognitive point of view numbers such as N* could be considered as "unnatural" and thus problematic to deal with.

Group	B(36)	B(36) & … & B(5)	Difference
G1	56.3%	41.3%	-15.0%
G1+	69.0%	58.6%	-10.4%
G1-	49.0%	31.4%	-17.6%
G2	74.6%	67.5%	-7.1%
G2+	75.2%	68.3%	-6.9%
G2-	72.0%	64.0%	-8.0%
G3	76.1%	71.7%	-4.4%

Table 9.27. *Percentage of correct answer (FALSE) for B(36) only, and for all the questions B(n) with n≤37. The last column on the right contains the difference between the previous two columns.*

Similarly to what we have done for Task A, these results could be interpreted in terms of APOS Theory with reference to the Explain Induction Process.

The Explain Induction Process in indirect form has been described (see. 4.4) as the process by which a subject, through the construction of a series of Modus Tollens (MT), can justify why, if the inductive base and the inductive step are valid, it cannot exist a natural number $n^*>n_0$ for which $P(n^*)$ does not hold. In fact, if we suppose that $P(n^*)$ does not hold, then necessarily $P(n^*-1)$ does not hold either, because otherwise, for the inductive step, $P(n^*)$ would hold. Thus, $P(n^*-2)$ does not hold either, because otherwise $P(n^*-1)$ would hold. By iterating this, we obtain that $P(n_0)$ does not hold, which however contradicts the induction base. Therefore, $P(n^*)$ must hold. As it happened for the Explain Induction Process with direct form, also this process is characterised by a successive coordination of processes, chained one to another, each starting from the result of the previous one. In analogy with before I will call **Chaining of MT** this coordination of processes within the Explain Induction Process.

At this point, as done for Task A, the different category of answers for these questions of Task B, could be interpreted in terms of Chaining of MT:

a) The students who answered "False" only for B(36) but not for lower numbers do not seem to have constructed a chaining of MT, not even as an action, because they only applied a second MT is not applied from the first one. They only applied one single MT.
b) The students who answered "False" for B(36), for B(35) but not for lower numbers may have constructed a chaining of MT as an action (they chained two successive MT), but not as a process since they correctly answered only to the questions which require a small number of steps from 37. It seems they did not imagine carrying out a higher number of MT steps.

After all the numerical questions, both in Task A and in Task B, an open question was posed asking to explain the reasoning for the answers to the previous questions. The analysis of students' response to this question highlights elements which support the interpretation of the previous results in terms of construction of the Chaining of MP/MT processes. Let us analyse here three examples. Some others will be presented in the discussion of the results for Task C.

The first is the case of a student of G2-[82] who, for the Task A, answered "True" only to A(104) and "We cannot know" to all the other questions. As an explanation the student wrote:

[82] In this and all the other presented examples that will follow, each student's name is coded with a number, randomly assigned when the answers were collected, and the indication of the student's group.

Stud. 22/G2-: We only know that if a number, natural, has this property then also its consecutive number will have it. Then for all the other numbers, except for 103 and 104, we cannot know if they have the property or not.

As we can observe from the student's words, a chaining of MP is missing. The student seems to apply a first MP ones (from 103 to 104) and then to stop after one single step.

The second example involves a student of G2+ who answered correctly "True" to all the questions A(n) with n≥103. However, in the open question instead of providing a justification for the answers, the student posed a question:

Stud. 2/G2+: If it is true for a number and it is also true for its consecutive number, then is it also true for the consecutive of the second one?

With this question the student seems to express some doubts about the possibility of iterating the inductive step to successive numbers. Probably s/he decided to use this argument to answer the previous questions, however the student does not seem to be fully convinced by that. This extract is particularly interesting because, besides the correct answers the student gave, it shows that constructing the chaining of MP could be not so intuitive. Indeed, the student's words align with Fischbein description of absence of intuitive knowledge, described in 5.1. As said, when a person knows something intuitively, (s)/he feels that "it must be so", and that it does not require further justification. The student 2/G2+, instead, by posing a question on the correctness of her/his reasoning is showing that this feeling of "it must be so" is not present.

Lastly, I show an example of answer from the Task B. The student, from the group G2+, selected that B(36) is "False", that B(35) is "False", and "We cannot know" for all the other question. Then s/he wrote:

Stud. 43/G2+: Given that it does not hold for 37, surely it does not hold for 36 (otherwise 37 would have the property) and for 35. We cannot know for 28, for 16, and for 5 since we don't know if the numbers preceding them have the property.

The students' words appear enlightening for our analysis and highlight that, starting from 37, the student correctly applies two MT inferences, for the numbers 36 and 35. However, for 28, 16, and 5, s/he affirmed that "We cannot know" because it is not known the truth/false value for the numbers preceding them. For those numbers, the student refers to the impossibility to apply MP, instead of iterating the MT from 37, as already done for 36 and 35. This is an example of Chaining of MT which does not seem to be interiorized as a process because the student constructed only a few steps of the chain (two MT) without carrying out and eventually skip some of these steps to reach further numbers. It is interesting to observe that the number 16 was not present in the questions of Task B, but it was mentioned by the student and so it seems to represent a generic number n<35.

9.2.2 Task C – Numerical questions

In Task C, as well as the validity of the inductive step, it was stated that C(45) is false and C(75) is true. The results for the questions C(n) with n≥75 and C(n) with n≤45 confirm the just presented result of Task A and Task B. In particular, as shown by the Table 19, **the progression of correct answers "True" for C(n) when n≥75 and "False" for C(n) when n≤45 is decreasing when the distance with the known value 75 or 45 increases.** The only exception is for the progression of "True" answers for C(n) when n≥75 for group G3, which remain constant at 89.1%. Also in this task, therefore, we have registered a relevant number of students in every group, except for the just mentioned exception, who seems not to have constructed the Chaining of MP/MT or to have constructed it only partially for a few steps.

In analysing the results for the numerical questions to this task I will focus on the questions C(n) where 46≤n≤74, that are C(46), C(60) and C(74). Considering the answers to these questions a further interesting result emerges. **Both for groups G1 and G2 (and in each respective subgroups) there are less correct answers “We cannot know” for 46 and 74 than for 60.** This can be observed in the figures 9.15-9.20. In these graphs, considering the line indicating the percentage of the “We cannot know” answers, in fact, there is a peak in correspondence to the question C(60). As highlighted by the Table 9.28 of below, this result is due to the fact that there is a relevant number of “False” answers for C(46) and a relevant number of “True” answers for C(74), and then, for C(60), less students selected “True” or “False”, preferring “We cannot know” instead. This created the peak in the graph of the “We cannot now” answers in correspondence to the question C(60).

Group	C(46)			C(60)			C(74)		
	True	False	We cannot know	True	False	We cannot know	True	False	We cannot know
G1	6.3%	32.5%	61.3%	11.3%	11.3%	77.5%	30.0%	11.3%	58.8%
G1+	6.9%	34.5%	58.6%	6.9%	13.8%	79.3%	34.5%	6.9%	58.6%
G1-	5.9%	31.4%	62.7%	13.7%	9.8%	76.5%	27.5%	13.7%	58.8%
G2	8.7%	12.7%	78.6%	8.7%	7.1%	84.1%	11.1%	7.9%	81.0%
G2+	8.9%	10.9%	80.2%	8.9%	5.0%	86.1%	7.9%	8.9%	83.2%
G2-	8.0%	20.0%	72.0%	8.0%	16.0%	76.0%	24.0%	4.0%	72.0%

Table 9.28. *Percentage of answers for the questions C(46), C(60), and C(74) of Task C for the groups G1, G2, and their subgroups. The green colour indicates the mathematically correct answers*

This result highlights two different aspects:

a) The fact that many students decided to choose “False” for C(46) or “True” for C(74) instead of “We cannot know” confirms the phenomenon, registered in literature (Inglis & Simpson, 2008), that many students, given an implication A→B, seem to use two laws of inferences which are not valid from a logical point of view. The first one is known as Affirmation of the Consequent (AC), and states that from A→B and B, it follows A. In our case this corresponds to conclude that, given C(74)→C(75) and the truth of C(75), C(74) is also true. The second one, known as Denial of the Antecedent (DA), states that from A→B and not-A, it follows not-B. In Task C, this corresponds to conclude that, given C(45)→C(46) and the falsity of C(45), C(46) is also false.

b) The fact that for C(60) less students answered “True” or “False” shows, moreover, that some of these students did not construct the Chaining Process for these invalid laws of inferences (AC and DA). In fact, for instance, by iterating AC from C(75) one might erroneously conclude that C(60) should also be true. Similarly, instead, by iterating DA from C(45) one might erroneously conclude that C(60) should also be false. Note that, if both DA and AC are iteratively applied C(60) would be simultaneously true and false. The number 60, in this task, is quite distant from 45 and 75 (the known values) and this seems to bring several students to answer "We cannot know" for C(60) even if they previously answered "True" for C(74) or "False" for C(46). This result confirms the results of the previous tasks, showing that when the distance from the known (true or false) value increases, the percentage of "We cannot know" answers increases as well. In particular, these results show that the Chaining process can present similar problematics also for these (invalid) inferences. This aspect can also be observed in the graphs representing the answers for the Task A and B. In the graph "G1- Task A" (Figure 9.3), for instance, if we look at the answers for A(n) where n<103 (it was stated that A(103) is true) we register a descending trend in the percentage of "True" answers: 36.0% for A(102), 32.0% for A(101), 28.0% for A(94) and A(4). This highlights that a descending trend similar to the ones registered for the questions involving a chaining of MP an MT can be observed also for what concerns the chaining of AC. This if even more evident for Task B, in relation to the chaining of DA. As shown in the graphs representing the percentages of answers of groups G1(+/-) and G2(+/-) (Figures 9.8-9.13), the percentage of the answer "False" is always the highest for the question B(38) among all the questions B(n) where n>37 (it was stated that B(37) is false).

As in Task A and B, at the conclusion of Task C there was an open question asking to explain the reasoning behind the given answers. I present here some of them, focusing on the presence (or absence) of the chaining of conditional inferences, both the valid ones (MP and MT) and the not logically valid ones (AC and DA).

The first answer that I will present is of a student of G1+ who previously selected "True" for every question C(n) with n≥75, "False" only for C(44), and "We cannot know" for all the other questions. Then the student wrote:

Stud. 11/G1+: I know that for the consecutive number of 75 it is true and, as a consequence, for its consecutive number and so on, then I know that for the number 45 it is not true, therefore surely it is not true for 44 either, whereas it could be for 46 but I cannot know it.

The student seems to construct the Chaining of MP: firstly s/he applies a MP and then a second one chained to the previous one ("as a consequence"), then they refer to an iteration of this argument ("and so on"). When dealing with the number 45, s/he argues by MT that for 44 "surely it is not true", but then s/he does not iterate the argument and the Chaining of MT is not constructed (indeed s/he answered "We cannot know" for C(12)). This is therefore an example of answer in which the Chaining of MP seems to be present, while the Chaining of MT does not.

The second example is very similar to the previous one, but now it involves a student from G2+. The student selected the same options for the questions C(n) as the previous one, then s/he wrote:

Stud. 43/G2+: It does not hold for 44 because otherwise it would have been true for 45 (which is not). We cannot know if for 46 can hold (we do not have enough criterions to say it), the same for 74. Surely it holds for 76, because it holds for 75, same reasoning for the integer numbers greater than 76 (e.g. 156380).

The first sentence corresponds to an application of an argument by MT. However the student does not iterate the argument to reach the lower numbers (the selected "We cannot know" for C(12), indeed). The second sentence shows that correctly the student is not using AC and DA. Finally, in the third sentence a chaining of MP seems to be present ("same reasoning").

As a third example, I will show the case of a student from G3 who answered correctly "True" to every question C(n) with n≥75, and then correctly "False" for C(44) but not for C(12). Then s/he wrote:

Stud. 38/G3: For the numbers greater than 75 the aforementioned property holds, for the others it is not possible to know, except for 44.

It is not clear if the student has constructed a chaining of MP when s/he writes "for the numbers greater than 74 the aforementioned property holds", since s/he answered correctly but without a justification. Then without explaining why s/he says that for the number 44 the property does not hold. We cannot be sure that at this point the student used a MT inference, however the fact that s/he did not write anything more and that for C(12) s/he selected "We cannot know" seems to suggest that a Chaining of MT was not constructed. This example, even if less explicit than the previous ones, is very interesting because it shows that the Chaining of MT process could be problematic also for an expert student (master's in Mathematics).

The previous three examples showed the cases of answers in which we could register a difference in the chaining of MP or MT construction. The next example, instead, will show a different category of answer. The student, belonging to the group G1-, previously selected "False" to C(44) and to C(46), "True" to C(74) and C(76), and "We cannot know" for the other questions. This is what s/he wrote as a justification:

Stud. 25/G1-: The comparison happens only on 2 natural numbers, we do not have other elements. The first 4 numbers are the precedent and the consecutive of 75 and 45 (numbers of which we have correct information) for the other numbers we do not have elements to make a hypothesis.

The student, starting from the "numbers of which we have correct information", i.e. the fact that C(45) is false and C(75) true, seems to apply the four different conditional inferences. Indeed s/he concluded respectively that C(44) is false (MT), C(46) is false (DA), C(74) is true (AC), and C(76) is true (MP). The student, moreover, did not chained any of the inferences which s/he seemed to have used, in fact, s/he answered "We cannot know" for all the other questions. The first sentence may say something more about this point. The student wrote: "The comparison happens only on two natural numbers, we do not have other elements". It is not clear to what s/he is referring. "The comparison"[83] could mean the tests for the property C on the numbers 45 and 75 made by the mathematician Coleman, as described in the text of the task. However, it could also mean that the inductive step, namely the theorem proved by Coleman's colleague, involves only two numbers (a natural number and its consecutive natural number) and "we do not have other elements" on other numbers. In this second case with the sentence "the comparison happens only on two natural numbers, we do not have other elements" the student could be saying that, given the inductive step and a truth or false value for $C(n_0)$ for some n_0, it is possible to conclude the truth or false only of $C(n_0-1)$ and $C(n_0+1)$. This is exactly what happened in his/her answers in Task C.

Another interesting example of answer is the one given by a student of G3. The student in Task A, where A(103) is known to be true, answered correctly "True" for every question A(n) with n≥103 and, as a justification, s/he wrote: "From 103 on, all the numbers should have the colleague's aforementioned

[83] "Il confronto", in Italian.

property”. In Task B, where B(37) is known to be false, however, the student selected “False” for B(36) and B(38) and “We cannot know” for all the other questions. Then s/he wrote:

Stud. 20/G3: 37 has not the property, thus neither the consecutive 38 has it. If 36 had this property, then 37 would have it too, but it does not have it, therefore 36 has not the property either.

What the student writes as a justification for her/his previous answers is extremely interesting. First of all s/he justifies the answer “False” for C(38) by using the conditional inference of DA: “37 has not the property, thus neither the consecutive 38 has it”. Secondly using a MT argument, constructed as an indirect argument and a MP, s/he justifies why C(36) should be false. Moreover, the student does not construct a Chaining process for these two conditional inferences (DA and MT) and does not iterate the arguments to further numbers. This situation is enriched by a new aspect if we consider the student’s answers in Task C. Here the student selected “False” to C(44) and C(46), “False” again to C(74) and C(76) and “We cannot know” for all the others. Then the student wrote:

Stud. 20/G3: I reasoned on those numbers that were close to the considered ones, for the other numbers there is not much to evaluate, there are not enough data.

As written by the student, s/he “reasoned on those numbers that were close to the considered ones”, that are 44, 46, 74 and 75. The answers for C(44) and C(46) may have been given by applying a DA and a MT from the falsity of C(45), as it happened in the previous task. It is interesting to notice that also in this task the student did not construct a chain of DA or of MT. Then the student answered “False” also for C(74) and C(76). Unfortunately, s/he did not provide any justification for this choice, leaving us no elements to try to interpret her/his answers. In any case, we can register that the student did not seem to have constructed a Chain of MP to conclude that for every $n \geq 75$ C(n) must be true, since C(75) is true.

As the last example of an interesting not correct answer, let us analyse the justification written by a student of G1- who, in Task C, answered “We cannot know” for the questions C(n) with $n \leq 44$, “False” for the questions C(n) with $46 \leq n \leq 74$, and “True” for the questions C(n) with $n \geq 76$. Then s/he wrote:

Stud. 33/G1-: For all the numbers greater than 75, Coleman’s property is true, by applying the rule of the consecutives. For those numbers between 45 and 75 it should be false, while for those lower than 45 we do not have enough information.

When dealing with the number 75, for which the property was known to be true, the student seems to construct a chaining of MP: “by applying the rule of the consecutives” (note the plural “consecutives” which may suggest an idea of iteration for this rule). For the number 45, for which the property was known to be false, the student does not apply a MT to conclude that C(44) is false, but s/he seems to apply the inference of DA (from C(45) false it follows C(46) false) and then to apply the same “rule of the consecutives” to conclude that “for those numbers between 45 and 75 it should be false”. It is interesting to notice that this full argument creates a logical contradiction, of which the student does not seem to be aware, because if the “rules of consecutives” could be applied for transmit both a truth or false value from C(n) to C(n+1), then the predicates C(n) when $n \geq 75$ would be, at the same time, true and false since n is, at the same time, bigger than 45 and 75.

Before analysing the results of the other parts of the survey, I wish to present some examples of correct answers which seems to contain an indication of the presence of the Chaining of MP/MT processes in the student’s answer. As first example, I show the answer of a student from G1-. This excerpt is particularly interesting since in it we can register the presence of a chaining of MP process, in a student who has not encountered MI during her/his studies. At the end of Task A, as a justification for the answers, the student wrote:

Stud. 12/G1-: Any number which comes after the number 103 will have Aarney's property because every number is influenced by the previous one which, in a chain starting from 103, will have this property.

The student claims that the property must be true for every natural number greater than 103. S/he justified this claim firstly referring to the theorem proved by Aarney's colleague (i.e., the inductive step) by saying "every number is influenced by the previous one". Then the student says that every number "in a chain starting from 103, will have this property". The student seems to have interiorized the chaining of MP and to use it to support her/his argument. It is interesting that the student talks autonomously about a "chain starting from 103" when describing the series of number for which the property is true. This seems to indicate that s/he is referring to the fact that the truth values of these numbers for the property are linked together. Moreover, by saying that that the chain *starts* from 103, the student is also referring to a direction in which these links develop. We can thus interpret the student's justification as containing a trace of the chaining of MP process.

The second example is of a student from G3 who, after answering correctly to every question of the Task A, wrote the following justification:

Stud. 42/G3: Since the property is valid for the number 103, then the property is valid for the consecutive natural number (104). Since it is valid for 104, it is valid for the consecutive one (105) and so on for all the natural numbers greater or equal to 103.

The student's answer contains a first application of a MP from 103, then repeated starting from 104. The first part of the answer highlights a Chaining of MP, where the second one is applied from the result of the previous. In the second part of the answer the student describes the chaining of MP to reach every natural number greater than 103 ("and so on for all the natural numbers greater or equal to 103"), namely what it has been called the Chaining of MP process.

As last example I show what written by a student G2+ who answered correctly to every question of task B. In the justification s/he gave at the end of the task, we can see the trace of the Chaining of MT process:

Stud. 40/G2+: We can say that for all the numbers which precede 37 it is surely false because it is false for 37 and therefore for 36 as well, otherwise it would have been true for 37, and so on.

The student's answer contains an application of a MT ("it is false for 37 and therefore for 36 as well") which is then generalised to all the numbers lower than 37 ("and so on").

9.2.3 Task A, B, C – Other results

The first question of Task A, B, and C involved the validity of the properties A, B, and C, respectively, on the whole set of natural numbers ℕ. In particular, it was asked to say if, by the information given in the text, it is possible to conclude that the property involved is true for all the natural numbers, false for all the natural numbers, true for some and false for some natural numbers, or "We cannot know (we do not have enough information to know if the previous statements are true or false)". For tasks A and B, the correct answer was "We cannot know" since we could only conclude that A(n) is true for $n \geq 103$ and that B(n) is false for $n \leq 37$ respectively, with no information on the remaining numbers. For Task C, instead, the correct answer was "Besides the numbers 45 and 75, there exist some numbers for which Colemans's property is true and there exist some numbers for which Colemans's property is false.", since we could conclude that C(n) is false for n≤45 and true for n≥45.

As shown in Tables 9.2-9.4 (for Task A), and 9.6-9.8 (for Task B), **the percentages of correct answers for this first question of task A and B were significantly low for every group and subgroup.** For what concerns Task C, from the Tables 9.17-9.19, we can register a percentage of correct answers which, although higher than in the previous tasks, still indicates that **a relevant number of students did not chose the correct option**. The data corresponding to this result are reported again summarised in Table 9.29 (below).

Group	∀n. A(n) We cannot know	∀n. B(n) We cannot know	∀n. C(n) True for some, false for others
G1	27.5%	32.5%	63.8%
G1+	34.5%	37.9%	75.9%
G1-	23.5%	29.4%	56.9%
G2	35.7%	46.8%	82.5%
G2+	37.6%	53.5%	82.2%
G2-	28.0%	20.0%	84.0%
G3	69.6%	78.3%	82.6%

Table 9.29. *Percentage of correct answers to the first question of Tasks A, B, and C for every group and subgroup*

These percentages seem to highlight that this first question of the first three tasks was problematic for many students. There are a few necessary comments for these results.

Firstly, the relevant percentage of incorrect answers for these questions could also be due to the linguistic complexity of the questions, which from a logical point of view involve both existential and universal quantifiers and negations. In particular, for Task A and B the correct answer is "We cannot know", however for some numbers we actually know the truth or false value of the property (every n≥103 for Task A and every n≤37 for Task B). Therefore, it is possible that some students decided not to select "We cannot know" as the answer because they were interpreting the statement as "For every number, we cannot know", which is not true since there exist numbers for which we do know the value of the property. On the other side, for Task C, a flipped situation may have occurred. In this case there exist numbers for which we know that the property C is true (n≥75), some others for which we not that it is false (n≤45), but others still for which we do not know if it is true or false (46≤n≤74). Therefore, a subject may be tempted to answer "We cannot know" because, exactly as it happened in Tasks A and B, there is a series of numbers for which we cannot know if the property is true or false.

A second aspect to notice about the results for these questions is that if we look only at the percentage of correct answers, we have a very limited insight of what could be behind these answers. We already saw, for example, that several students, for the numerical questions of Task A, selected "True" only for A(104) and "We cannot know" for all the other questions. A coherent answer to the first question of Task A, in these cases, would be "We cannot know", which is the correct answer even if the other answers of the task were not correct. The same can be said for Task B, where the correct answer "We cannot know" could have been selected coherently also by students for which B(37) is false and it is not possible to know anything else. Similarly in Task C the correct answer "Besides the numbers 45 and 75, there exist some numbers for which Colemans's property is true and there exist some numbers for which Colemans's property is false" could have been selected also by students who correctly applied a MP, concluding that C(76) is true, and a MT, concluding that C(44) is false, but without constructing the Chaining process, or by those students applying the logically invalid inferences of DA or AC. This is to say that the results for the first question of Task A, B, and C cannot take into consideration the extremely variegated collection of answers resulting from the numerical questions.

The results of the numerical questions A(n), B(n), and C(n) may be also used to investigate a different aspect, not considered yet. As seen, some of these questions were related to the Chaining of MP, whilst some others to the Chaining of MT. A wide literature has highlighted how the MT is often cognitively more problematic than MP (Inglis & Simpson, 2008; Antonini, 2004; Fischbein, 1987, pp. 72-81). This brings us to ask whether the Chaining of MP/MT present similar problematics. To have a first insight on this, we can compare the results for the questions related to the chaining of MP and of MT in Task A and B, or in Task C.

For every group and subgroup, all the questions in Task A involving an application of MP, or a chaining of MP, i.e. A(104), A(105), A(112), A(542384), A(N*), obtained a higher percentage of correct answers than all the respective questions in Task B involving an application of MT, or a chaining of MT, i.e. B(36), B(35), B(28), B(5).

An analogous result is obtained if we compare the progressions of correct answer for A(n) when n increases from 103 with the progressions of correct answers for B(n) when n decreases from 37 (see Table 9.5 for A(n) questions, and Table 9.9 for B(n) questions). In this case t**he percentages in the progression of correct answers in Task A (i.e., chaining of MP) are always higher than the percentages in the progression of correct answers in Task B (i.e., chaining of MT).** This aspect has been registered in every group and subgroup.

The results of Task C confirm these aspects. As we can see in each graph representing the percentage of answers for the numerical questions C(n) (Figures 9.15-9.21), **the percentage of "True" answers for C(76), C(156380) and C(L*) are always higher than the percentage of "False" answers for C(44) and C(12).**

These results, first of all, confirm other studies registering that the inference MT may present greater difficulties than MP (Inglis & Simpson, 2008; Antonini, 2004; Fischbein, 1987 pp. 72-81), showing that the questions involving one application of MT obtained a lower percentage of correct answers than the one involving one application of MP. Moreover, these results show also that the Chaining of MP/MT presents similar problematics. The data, in fact, suggest that the Chaining of MT could be more problematic than the Chaining of MP, since in every group and subgroups not all the students who seemed to have constructed the Chaining of MP, also seemed to have constructed the Chaining of MT, even if they can correctly apply one MT. As a confirm of that, we can consider that in the analysis of both the open and closed questions of all the 252 participants, we have not found any answer in which the Chaining of MT seems to be present without the presence of the Chaining of MP as well, whilst the opposite situation was registered instead. In particular we registered the answers of students who correctly constructed a chaining of MP, correctly applied one single MT, but did not construct a chaining of MT (see, for instance the answer of the student 11/G1+ analysed in a previous paragraph). This leads to formulate the hypothesis that a greater complexity of an action (i.e. Modus Tollens vs Modus Ponens[84]), even once it has been built, also makes it more complex to build the process even if it is similar to one already built.

[84] In the GD of MI presented at 4.4, the process MT was modelled as a coordination of the process MP with the Negation process within the Logic Schema. From this point of view, thus, the MT process is more complex than the equivalent process with MP.

The last result to analyse for these first three tasks is the one related to the last two questions of Task B. In these questions students were asked to take a position (I agree/I do not agree) with a statement claiming that, since B(37) is false, the theorem of Bertrand's colleague, i.e. the inductive step ∀n(B(n)→B(n+1)), must contain a mistake. In the first question it was stated that the number 37 is a possible counterexample for the theorem, whilst in the second question the same thing was stated but for the number 36.

The results for these two questions, presented in Tables 9.10-9.16, show that a relevant number of students claimed to agree with the first statement or with the second statement, even though both the statements are false. These data are reported again in Table 9.30 below.

Group	I agree that B(n)→B(n+1) is not valid for n=37.	I agree that B(n)→B(n+1) is not valid for n=36.
G1	36.3%	47.5%
G1+	27.6%	37.9%
G1-	41.2%	52.9%
G2	23.0%	31.7%
G2+	17.8%	23.8%
G2-	44.0%	64.0%
G3	13.0%	10.9%

Table 9.30. *Percentage of "I agree" answers for the statements claiming that since B(37) is false then the theorem of Bertrand's colleague, i.e. the inductive step ∀n(B(n)→B(n+1)), must contain a mistake.*

These results, beyond the difference between the groups, highlight **that many students do not seems to accept the validity of the implication of the inductive step in itself independently from knowing the value of antecedent and consequent**. This offers a confirm to the study made by Fischbein and Engels (1989) on student's difficulties in accepting the validity of the inductive step, discussed in detail in the section 5.2.4. Moreover, a further aspect was registered for the groups G1 and G2. The statement claiming that B(36)→B(37) is a possible counterexample obtained higher percentages of "I agree" answers than the one claiming that B(37)→B(38) is a possible counterexample. In other terms, in these groups, there were more students convinced by the fact that, since 37 is false, it cannot be that B(36)→B(37) than students convinced by the fact that, since 37 is false, it cannot be that B(37)→B(38). Unfortunately, we did not register any qualitative data in the open question at the end of this task to interpret this aspect. A first explorative hypothesis could be that accepting that B(36)→B(37) is valid despite the fact that B(37) is known to be false requires for a subject to accept as valid a theorem (B(36)→B(37)) where the thesis is known to be false, whilst in the other case (B(37)→B(38) knowing that B(37) is false) it requires to accept a theorem in which the hypothesis is false. It could be, therefore, that a subject does not accept as valid a theorem in which a false thesis is concluded since in it something that is not true is proved, whilst instead s/he could perceive a theorem which starts from a false hypothesis simply as *useless*, since it cannot be applied, but not as invalid. This is of course not correct from a logical point of view; in a theorem, in fact, we prove the validity of the implication hypothesis→thesis and not the validity of the thesis itself. Further research with new and qualitative data would be necessary to deeper investigate this possible interpretation.

In addition to these quantitative results, I will also present some answers given by a student from G1+, that seem to be extremely related to this point. In Task A, the student correctly answered "True" for every question A(n) with n≥103 and correctly "We cannot know" for all the questions A(n) with n≤102. In the justification provided s/he seems have recognised that the implication proved by Aarney's colleague (the inductive step) can be iterated to conclude the truth of A(n) for every n≥103:

> Stud. 29/G1+: Aarney said that the property is true for 103 and since a colleague proved that the consecutive of a number with this property has itself the property, we can deduce that for all the numbers greater than 103 (and the number 103 itself) the property is true.

However, in Task B, the student's answer has completely changed. S/he selected "We cannot know" for every question in Task B; then for the two questions stating that the theorem of Bertrand's colleague was mistaken, s/he selected "I agree". In the open question at the end of the task the student wrote:

> Stud. 29/G1+: We do not know if the proof of Bertrand's colleague is true, so it is not possible to decide if those numbers make Bertrand's property true or not. The only certain thing is that for the number 37 the property is false.

The student is not convinced by the validity of proof of the inductive step, as we can see both from the first sentence s/he wrote ("We do not know if the proof of Bertrand's colleague is true") and for the fact that s/he selected "I agree" for the claim that the proof must be mistaken. At this point, as s/he says, "it is not possible to decide if those numbers make Bertrand's property true or not", indeed s/he answered "We cannot know" for every question of this task. From these answers we can see that the student is not accepting as valid the inductive step because C(37) is known to be false, and therefore s/he concludes that "the only certain thing is that for the number 37 the property is false". A similar thing happens in the subsequent task. In Task C, beside the validity of the inductive step, the fact that C(75) is true, it was also said that C(45) is false. At this point the student, coherently to what happened in the previous task, questions the validity of the proof of the inductive step made by the colleague. S/he wrote:

> Stud. 29G1+: If the proof made by Coleman's colleague was true, then the number 45 should make Coleman's property true, but this is not. Therefore it is possible that the colleague's proof is mistaken and not that we cannot use it.

The student clearly states that the inductive step is not valid, otherwise C(45) would not have been false. At this point, since the colleague's proof is possibly mistaken, s/he says that the theorem is useless ("we cannot use it"). In fact, coherently with that justification, the student selected "We cannot know" for all the questions of the task C. This choice is interesting because in Task A, as we observed, the student had correctly used the colleague's theorem to conclude that A(n) is true for every n≥103, but now s/he has not done the same for C(n) when n≥75. The fact that C(45) is false implies, for the student, that the inductive step is not valid, and therefore it cannot be used. This is an example of what I have called "Interference between the inductive step and the case base" (see section 3.3.2): the theorem corresponding to the inductive step is seen as valid only if paired with a truth base case. To conclude, this example showed clearly that recognising the validity of the implication in the inductive step $\forall n(P(n) \rightarrow P(n+1))$ as an independent theorem, unrelated to the truth/false value of a given P(n*), could be rather problematic for students, even for those who encountered MI during their studies.

9.2.4 Task D

In this task it was only given the validity of the inductive step for a certain property E and then it was asked to indicate if some statements about the property E were CERTAIN (C), POSSIBLE (P), or IMPOSSIBLE (I). The complete results for these questions were presented in Tables 9.21-23.

The first result that I will analyse is the one related to the statement labelled as 'F' in the results: "Evelin's property is FALSE for all the natural numbers". For this statement we registered, in every group, several students who answered "Impossible" (see Table 9.31, below).

	G1	G1+	G1-	G2	G2+	G2-	G3
It is IMPOSSIBLE that "Evelin's property is FALSE for all the natura numbers"	27.5%	24.1%	29.4%	16.7%	13.9%	28.0%	6.5%

Table 9.31. *Percentage of "IMPOSSIBLE" answers for the statement "Evelin's property is FALSE for all the natural numbers*

This aspect shows that for many students it was problematic to recognise that the inductive step is not in contradiction with the fact that the property E is false on the whole set of natural numbers. This confirms the result, discussed above, of the students' difficulties in accepting the validity of the implication in the inductive step as independent from knowing the truth value of antecedent and consequent and may suggest that for some students the inductive step $\forall n.(P(n) \rightarrow P(n+1))$ is considered valid only for the numbers for which the predicate P(n) is actually true.

Two other interesting results emerging from this task can be seen in the percentage of answers for the two statements labelled as 'FTFT…' and 'TTFF…'. The two statements claimed that the property E has an alternating truth/false value respectively on odd and even numbers (FTFT…) or every other couple of natural numbers (TTFF…). The validity of these two statements, from a formal point of view, is in contradiction with the validity of the inductive step, however the percentage of students who answered "Possible" is quite relevant (see Table 9.32, below).

	G1	G1+	G1-	G2	G2+	G2-	G3
It is POSSIBLE that 'FTFT…'	35.0%	34.5%	35.3%	21.4%	18.8%	32.0%	13.0%
It is POSSIBLE that 'TTFF…'	42.5%	37.9%	45.1%	15.9%	13.9%	24.0%	10.9%

Table 9.32. *Percentage of "POSSIBLE" answers for the statements 'FTFT…' and 'TTFF…'.*

These students, as these results show, did not seem to have realised that as a consequence of the inductive step it cannot exist a number n^* for which $E(n^*)$ is false, if n^* is greater than some n_0 for which $E(n_0)$ is true.

About this point, an interesting answer was registered, written by a student of G2+ who selected "Impossible" for 'FTFT…' but "Possible" for 'TTFF…'. In the open question at the end of the task, s/he wrote:

Stud. 5/G2+: Since we do not know for what numbers Evelin's property is true, but we know that it is true for couples of numbers, we cannot know for which number is true, but I can affirm that if it is true for a number (x) then it is also true for the number (x+1)

The answer seems to indicate that for the student the property must be true for couples of numbers, since if the property is true for a number x, then it is also true for the number x+1. Note that the student is not saying that if the property is true for a number x, then it will be true for every natural number greater then x, but s/he just refers to x+1. Following this answer, therefore, it could be that the student thought that the statement 'TTFF…' is "Possible" (as s/he selected) since in this statement it is affirmed that the property is true for a series of couples of consecutive numbers (x, x+1), and for these couples it is, indeed, valid that E(x)→E(x+1). In other terms, it could be that the student is interpreting the inductive step as a statement which claims that Evelin's property "is true for couples of numbers", as s/he wrote. Coherently with this interpretation s/he may have selected "Impossible" for the statement 'TFTF…' because in this case the property E is false for every couple of natural numbers (x, x+1).

Another interesting result registered in this task can be found in the answers for the statement '40T-41T', claiming that "Evelin's property is TRUE for the number 40 and at the same time TRUE for the number 41". For this statement, many students selected "Certain" as answer (see Table 9.32, below).

	G1	G1+	G1-	G2	G2+	G2-	G3
It is CERTAIN that '40T-41T'	25.0%	34.5%	19.6%	23.0%	20.8%	32.0%	10.9%

Table 9.32. *Percentage of "CERTAIN" answers for the statement '40T-41T'.*

A possible interpretation is that these students have interpreted the statement as hypothetical: "IF the property is true for 40, THEN is true for 41". This can be observed in the answer written by a student from G2-, who selected "Certain" for this statement and then s/he wrote:

Stud. 3/G2-: Since we do not know the numbers for which the rule is true or false, we cannot know if the previous questions are true or false. However if for a number x it is true, it is also true for the consecutive number, therefore, for example, 40 and 41 it is true. But it is not said that it is true. Because we do not know if for 40 and/or 41 the rule holds.

The student seems to state that, because of the inductive step ("if for a number x it is true, it is also true for the consecutive number"), the statement 40T-41T is true ("for example, 40 and 41 it is true"). S/he, in fact, selected "Certain" for this statement. However, as the student wrote "It is not said that it is true. Because we do not know if for 40 and/or 41 the rule holds". With this answer, thus, the student seems to refer to a possibility for the statement 40T-41T to be valid rather than to a certainty. In conclusion it seems that the student interpreted the statement '40T-41T' as hypothetical and therefore they selected "Certain".

Some results from this task seem to confirm the results of the previous tasks on the Chaining of MP/MT processes. The statement '40T-41F' is "Impossible". An argument for this conclusion is that if the property was true for 40, then, by MP, it would be true for 41 as well. Analogously we could also say that if the property was false for 41, then, by MT, it would be false for 40 as well. In both cases we conclude that the statement '40T-41F' cannot be "Possible". Similar arguments can be constructed to conclude that the statements '12T-54F' and 'ST-BF' are "Impossible", where '12T-54F' is the statement "Evelin's property is TRUE for the number 12 and at the same time FALSE for the number 54" and 'ST-BF' the statement "Evelin's property is TRUE for small numbers and at the same time FALSE for the number with millions of digits". However, for what concerns these last two statements, a Chaining of MP/MT could be involved, since the numbers involved are not consecutive. If we compare the results to these questions, we register that the statement '40T-41F' obtained a higher percentage of "Impossible" answers than the other two statements. This happened for every group and subgroup (see Table 9.33, below).

	G1	G1+	G1-	G2	G2+	G2-	G3
It is IMPOSSIBLE that '40T-41F'	66.3%	72.4%	62.7%	90.5%	93.1%	80.0%	87.0%
It is IMPOSSIBLE that '12T-54F'	46.3%	55.2%	41.2%	80.2%	82.2%	72.0%	84.8%
It is IMPOSSIBLE that 'ST-BF'	60.0%	69.0%	54.9%	81.7%	85.1%	68.0%	84.8%

Table 9.33. *Percentage of "IMPOSSIBLE" answers for the statements '40T-41F', '12T-54F', and 'ST-BF'.*

These data, showing that the percentage of correct answers decreases when comparing questions involving consecutive numbers with questions involving distant numbers, confirm that some students seem to have correctly applied one single MP/MT but did not construct a Chaining process for those.

One of the aims of Task D, in general, was to investigate whether students could recognise the inductive step for a given property as an independent theorem, from which it is possible to infer some consequences on the validity of the property itself on the set ℕ. In relation to this point the answers of two students of G3 are interesting. Both the students selected "Possible" for every statement in the task, then they wrote:

Stud. 5/G3: Since they did not prove that the property holds or not on any number, I preferred to leave open every option.

Stud. 22/G3: We do not know for which numerical values the property is true (or false), therefore all the options are plausible.

These answers seem to indicate that both the students are perceiving the inductive step itself as an 'empty' statement which, when considered as alone does not provide any information on the given property. In other terms, the students do not seem to recognise that the validity of the inductive step itself has some logical consequences on the property itself. With their words, in fact, the students seem to say that it is necessary to have other information to conclude something from it.

An almost opposite situation can be observed in the answer of a student from G1- who selected "Certain" for the statement which affirms that the property E is true for every natural number and for the statement which affirms that the property E is true for the numbers 40 and 41. Note that these are the only two statements in which it is not affirmed that there is a number for which the property is false. As an explanation for this choice, the student wrote:

Stud. 19/G1- : All the numbers have the property.

With this very concise answer the student seems to indicate that, as a consequence of the inductive step, we can conclude that the property is true for all the natural numbers. S/he, in fact, selected "Impossible" for all the statements in which it was affirmed that the property is false for at least one number. This situation is, in a certain sense, the opposite of the previous one. For the two students of G3 (5 and 22 of above) the inductive step was not recognised as a theorem itself, but as an empty statement, so that nothing was possible to conclude from it. For this student of G1-, instead, there is almost an overlapping between the inductive step $\forall n(E(n) \rightarrow E(n+1))$ and the general statement $\forall n\ E(n)$, so that the first one automatically implies the second one without the need of other information.

This last result can also be traced in the quantitative data collected from Task D. If we look at the answers for the statement 'T' (The property E is true for all the natural numbers) we can see that some students selected "Certain" as the answer (see Table 9.34, below).

	G1	G1+	G1-	G2	G2+	G2-	G3
It is CERTAIN that 'The property E is true for all the natural numbers'	8.8%	3.4%	11.8%	4.8%	2.0%	16.0%	2.2%

Table 9.34. *Percentage of "CERTAIN" answers for the statement 'The property E is true for all the natural numbers'.*

The fact that for many subjects the inductive step alone implies the truth of the general statement confirms other studies reported in literature (Ernest, 1984, Dubinsky & Lewin, 1986; Ron & Dreyfus, 2004; Garcia-Martinez & Parraguez, 2017) which show that the base of the induction $P(n_0)$ is often considered by students as not necessary in a proof by MI.

9.2.5 Overall results on Tasks A, B, C, D

A first necessary comment on the just discussed results relates to the group division that I adopted in the presentation of the result. As indicated at the beginning of the "Result" section, the division in groups G1(+/-), G2 (+/-), and G3 was made by considering, firstly, students' different "experience" with mathematical courses at university level and, secondly, their "encounter" with MI during their study. The goal of keeping track of this distinction whilst presenting the results was not to create a ranking between the groups involved, focusing on the differences between the percentages of correct answers. **However, it was made to show that some problematic phenomena related to MI could be traced transversely in vary and different groups of students.** What the presented results showed, in fact, was that similar categories of problematic answers have been registered among students with different experience with MI or with Mathematics in general. This aspect straightens the results that we discussed because it might suggest that some difficulties or problematic aspects with MI are quite deep and hard to overcome so that they could remain unsolved even in expert students (such as the one in G3).

As a further confirm of this point, in Table 9.25 (reported a second time below), I presented the results corresponding to the percentage of students who answered correctly to every question of Tasks A, B, C, or D. For Tasks A, B, and C, I considered only the numerical questions (A(n), B(n), C(n)) and not the most complex questions. The data collected showed low percentages of completely correct answer for each task. Note that if we considered all the questions of the survey, the percentages would be even lower.

Group	All correct answers for the numerical questions A(n)	All correct answers for the numerical questions B(n)	All correct answers for the numerical questions C(n)	All correct answers for the questions of Task D	All correct answers for the previous questions
G1	35.0%	17.5%	17.5%	11.3%	3.8%
G1+	55.2%	31.0%	27.6%	10.3%	6.9%
G1-	23.5%	9.8%	11.8%	11.8%	2.0%
G2	65.1%	46.8%	55.6%	42.9%	31.0%
G2+	70.3%	49.5%	59.4%	47.5%	33.7%
G2-	44.0%	36.0%	40.0%	24.0%	20.0%
G3	82.6%	65.2%	76.1%	67.4%	50.0%

Table 9.25. *Percentage of correct answers for all the numerical questions of TA, TB, TC, and for all the questions of TD. The last column shows the percentage of correct answers for all these questions together.*

Surely these are quite 'extreme' data, since they consider only the number of completely correct answers for a survey which was quite long, and perhaps repetitive for some participants. However, even with these limits, these results are quite interesting. Let us consider, for instance, the answers of the group G2 (+/-) for the numerical questions of Task C. In the graph contained in Figure 9.18, we can see that the percentages of correct answers to each question C(n) are all higher than 70% (min. 71.4% for C(12), max. 84.9% for C(76), mean 80.0%). However, the situation is a bit more complex if we look at the percentage of students of this group who answered correctly to every question of this task (55.6%). This means that there are students who answered correctly only to some of the questions and, more interestingly, that these questions were not the same for every student. This is to say that **the results on the percentages of correct answers to every question of a task show that the different categories of correct/not correct answers are rather various and complex.** This aspect could have remained hidden if we had only considered the percentages of correct answers separately for each question of a task.

A second general comment to make relates to the kind of questions posed in the survey. In the a-priori analysis of the survey, I observed that the fact that Tasks A, B, C, and D involved some generic and unknown properties could have added a further level of complexity to the survey since the questions were posed, in a certain sense, on a 'meta' level, whose objects were propositions and properties themselves. In the collected data it was possible to observe several answers which seem to be extremely related to this point.

As a first example let us consider the answers of a student from G1- given in Task A. The student in the first question ("Aarney's property is:") selected "We cannot know", then s/he wrote:

Stud. 34/G1-: Since we do not know the statement of the property we cannot know for what numbers it holds.

Then in all the subsequent numerical questions the student selected "We cannot know". A similar thing happened with the same student in Task C where, after selecting "We cannot know" for every question of the task, s/he wrote:

Stud. 34/G1-: We cannot know with certainty if we do not know what the theory says

In both cases the student's reasoning seems not be situated on that 'meta' level necessary to answer correctly to the questions, and the fact that the involved properties are unknown is indeed enough for the student to conclude "We cannot know" for every question.

A similar situation to the previous one can be seen in the following example. The student, from G1-, in Task C answered "We cannot know" for every question and, as a justification, s/he wrote:

Stud. 42/G1-: Since we do not know what this property is about it is not possible to give an answer with certainty

As in the previous example, the student seems to say that s/he needs to know "what this property is about" to answer the questions of the task. Again, s/he seems to have difficulties to talk about a general property and to reason on the 'meta' level of the task.

The last example of this kind is even more explicit than the previous ones. The student, now from group G3, answered "We cannot know" for every question of tasks A, B, and C, and "Possible" for every question of Task D. All the justifications that the student provided for these choices refer to the fact that the properties involved were unknown. In task A, for instance, s/he wrote:

Stud. 5/G3: Because I have no information on the property which verifies that number.

In Task B the justification clearly refers to the fact the property is unknown and therefore it was not possible to answer:

Stud. 5/G3: I have no information on the property. We can think about many properties which are false for the number 37, for instance of not being an even number, or of not being a multiple of 4, and so on. Therefore, since we do not know the property, I did not know how to justify the truth or the falsity of the property for the other numbers

The answer of the student is rich of interesting elements. The student claims again of having no information on the property. Then s/he suggests some different properties which could be false for the number 37, as the unknown Bertrand's property, such as the property of being an even number or of being a multiple of 4 (note that the student seems to be puzzle by double negations in this point and, more importantly, that both the properties are not inductive). The student thus concludes that, since there exist different numerical properties as the one described in the task, it is not possible to answer the previous questions with "True" or "False" ("I did not know how to justify the truth or the falsity of the property for the other numbers"). As in the previous cases, the student does not seem to situate the discourse on a meta level, in fact s/he needs to consider a specific and known property, to construct an argumentation on it. As a further confirmation of that, let us observe what the student wrote in a subsequent question at the end of the survey asking: "What is, for you, a 'Proof by Mathematical Induction'?". The student wrote:

Stud. 5/G3: It is a proof which we, initially, verify for an initial value, then we suppose true for $n=k$ and we prove for $n=k+1$

The sentence is grammatically ambiguous. It is not clear if “which” is the object of the verbs “verify”, “suppose”, and “prove” or if it stands for “in which” (in that case the objects of the three verbs are missing). However, we can notice that, in any case, there is not a reference to any generic property, predicate or proposition. Again, it seems that the student is not situating the discourse on a ‘meta’ level: s/he describes the actions to be performed in a proof by MI, but s/he has, literally, no words for the predicate involved. It is exactly the absence of the object “a predicate”, what contributes to the grammatical ambiguity of the sentence: What do “we verify”? What do “we suppose” and “prove”?

This last example, extremely interesting especially if we consider that the student is a Master student in Mathematics, showed that talking and reasoning on a ‘meta’ level, which is necessary, for instance, if we want to refer to the Principle of MI or to describe a generic proof by MI, could be an extremely problematic aspect for students.

Before presenting the analysis for the open answers of the last section of the survey, a comment on a further aspect is necessary. As said when the survey was presented, after each multiple-choice question the students were asked to indicate their perceived level of certainty and the effort they made to answer (very little [0] / little [1] / enough [2] / a lot [3]). These questions were posed to eventually register if some questions of the survey were considered extremely problematic by the students from an intuitive point of view (in terms of perceived certainty and effort to answer). In other terms I was interested in registering the possible presence of questions which obtained, on average in each group, a low level of perceived certainty (<1.5, i.e. 50% in the 0-3 scale) or a high level of perceived effort to answer (>1.5, i.e. 50% in the 0-3 scale). These results were not presented above since all the questions of Task A-D, in every group and subgroup obtained an average level of perceived certainty higher or equal to 1.5 and an average level of perceived effort to answer lower or equal to 1.5. Therefore, at least for the collected data, it was not possible to register the presence of questions which were stated to be problematic (in terms of perceived certainty and effort to answer) by the average of the students. This, of course, does not mean that none of the question was problematic, but that this aspect does not emerge from what stated by the students themself. These data, however, even if not conclusive, still highlight a worth noting element for our analysis, considering that some questions registered an extremely low percentage of correct answers. This indicates that many students who selected a not mathematically correct answer stated to be “enough” or “a lot” certain about the answer and that the effort was “little” or “very little”.

9.2.6 ‘Meta’ Questions on MI

In this section I will present and analyse some of the answers to the questions of the last part of the survey, the ‘Meta’ questions on MI, as called in the description of the survey. These questions were not posed to the students who stated to have never encountered “The Proof by Mathematical Induction”.

9.2.6.1 The question on the base case

One of these questions was still a multiple-choice question. The students were asked to select “True”, “False” or “I do not know” in response to the statement “In the questions on Aarney’s, Bertrand’s, Coleman’s, and Evelin’s properties it was not possible to state anything certain because the proof of the base case was missing”. The results for this question were presented in Table 9.24.

Firstly, we can observe that there is a relevant number of students who selected “I do not know” (37.9% for G1+, 11.9% for G2+, 19.6% for G3). This could be related to the fact that for several students it was not clear what “it was not possible to state anything certain” meant. Indeed, in the comments section at the end of the survey some students stated exactly that since this question was ambiguous and not clear, they answered “I do not know”.

Looking at the other answers, instead, it is possible to notice that a high percentage of students selected "True" for this question: 31.0% for G1+, 29.7% for G2+, and 34.8% for G3. These students agreed that since the base case was missing in the previous tasks all the answers could not be certain (note that a base case was missing only in task D, while in task B a "false" base case was given). This, for instance, is the comment of a student who, after selecting "I do not know" in the previous question, wrote:

> Stud. 12/G2+: I did not consider the base of the induction. Probably almost all my answers will be wrong. I am not sure at all of what I just did since proofs are something new to me.

The question on the base case seems to have prompted serious doubts to the students, who, in fact, states not to be sure anymore of what s/he just did and that "probably almost all my answers will be wrong". This answer is particularly interesting if we consider that the student answered correctly to every numerical question A(n), B(n), and C(n) and to all the questions of Task D, instead. Moreover, in the question asking to describe why a Proof by MI 'works', the student wrote:

> Stud. 12/G2+: Because through the base of the induction we can prove that the proposition is true for the first natural number, that is 1. After that, through the inductive step, we prove that if it is true for a given natural number n, it is also true for the consecutive one. In conclusion, if the proposition is valid for 1, it will be for the consecutive 2 as well, then 3 and so on.

The student's explanation is complete, both base case and inductive step are correctly described, and in the last sentence we could also see a trace of an Explain Induction Process ("In conclusion, if the proposition is valid for 1, it will be for the consecutive 2 as well, then 3 and so on. "). An element which is not present in the students' answer, however, is the fact that the inductive base could be a natural number n_0 different from 1. This could be caused by the fact that the question asked to explain why "a proof by mathematical induction assures that a proposition is true for all the natural numbers", which happens only if the case base is, indeed, 1 (or 0, depending on the definition of $\mathbb{N}$). However, it could also be that the student is not aware of the fact that in a proof by MI the base case could be a given n_0 different from 1. It could be that the student did not recognized A(103) or C(75) as possible base cases of an induction and therefore s/he concluded that "I did not considered the base of the induction" in the base case question of above.

This answer and the collected data of above highlight something in relation to the students' intuitive knowledge of MI. They show, in fact, that some students, despite having a formal knowledge of MI, being able to describe or apply it, could lose every certainty about MI when some components are modified (such as the base is moved to a number n_0 different from 0 or 1, or when $P(n_0)$ is said to be false), showing that they have not an intuitive knowledge of MI yet. In this sense this result is also a confirmation of what Movshovitz-Hadar (1993) called students' "knowledge fragility" in relation to MI.

In general, moreover, the data collected for this question confirm other results, documented in the literature, on student's difficulties when the base case involves a number n_0 greater than 1 (Avital & Lebeskind, 1978; Stylianides et al., 2007) and that some students may perceived MI as a fixed and "ritual" way of proving (Harel, 2001), so that they do not accept as correct and convincing a slightly modified version of a proof by MI.

9.2.6.2 Open questions on MI

In the final part of the survey three open questions on MI were posed to the participants, asking them respectively to describe what a proof by mathematical induction is, to explain why 'it works', and to say if there are any aspects that they did not understand or that they were not convinced by. A comment followed every question: "Do not worry about using a formal language; if you wish you may refer to mental images, feelings, memories". Although with the limitations already discussed in the a-priori analysis of these questions (section 8.3.1), in their answers it was possible to register interesting elements in relation to the research objectives.

In the model of Theorem by MI as a triplet (see. Section 3.3), I observed that a theorem proved by MI can be modelled as system of triplets corresponding to three different theorems: the theorem corresponding to the base case, the theorem corresponding to the inductive step, and the meta-theorem (formally represented by the principle of MI) which provides a justification of why from base case and inductive step it follows that the general statement must be true. As observed a crucial didactical aspect emerging from the model was that a student needs to conceptualize each of these three theorems both as independent and together within the same general theorem.

With reference to this model, I have analysed the students' answers to the open questions on MI investigating on how MI was described by them and to what extent the components of the model emerge in these descriptions. What follows is a series of paradigmatic examples representing different categories of answers.

<u>The proof by MI as a system of three theorems</u>

As a first example I show the answer of a student from G2 in which the reference to all the three theorems of the model and to their mutual relationships is quite evident. The student wrote the following answer to the question asking why MI 'works':

Stud. 39/G2+: If we prove that a property is valid for the smaller natural number (1 or 0) we have a starting base. Then the remaining part to prove is that taking any natural number "n", for which the property is valid, then the property holds for the number "n+1" as well. At this point we know that if we find a number for which the property holds, all the consecutive numbers will have the same property. Putting together these two little bricks we know with certainty that all the natural numbers bigger or equal to 1 (or 0) have the property (that is all the natural numbers).

The base case and the inductive step are well described as two independent parts of the proof, each corresponding to a particular theorem to prove (note, in this sense, that the student is both describing the statements and referring to their proofs). In the second half of the answer the student describes what we can interpret as the meta-theorem, however without giving a justification for it. It is interesting to notice that, in doing so the student metaphorically refers to the base case and to the inductive step as "two little bricks"[85] that can be put together. The use of this metaphor firstly highlights that the student is objectifying the base case and the inductive step, describing them as similar (the two bricks are equals from a physical point of view), independent (the two bricks were firstly separated objects), and equally necessary for concluding the proof (the two bricks must be put together). Secondly, it is interesting that the student is using a metaphor to talk about the base case and the inductive step when s/he is talking about the meta-theorem. We can thus see a parallelism between the student's metaphor, where the case base and the inductive step become two physical objects, and what happens in the meta-theorem from a logical point of view, where the case base and the inductive step become the involved "objects". This excerpt, moreover, could be interpreted, in APOS terms, as an example of encapsulation of the inductive step and the base case which becomes, in the student's description, two objects (the "two little bricks").

[85] "due mattoncini", in Italian.

Absence of the case base

In several answers, the proof by MI was described without any reference to the base case. This, for example, is what a master's student in Mathematics wrote to describe MI:

Stud. 1/G3: If a property is supposed to be true for n and it is true for n+1, then it is true for every n.

The answer is very concise. We can observe the presence of the inductive step, but there is not reference to the base of the induction at all. Moreover, the inductive step does not seem to be described as an implication ("a property is supposed to be true for n and it is true for n+1"). Note the student's use of "and" as s/he is describing the inductive step as $P(n) \wedge P(n+1)$.

Similarly, a student from G1 wrote:

Stud. 4/G1+: [MI] is a proving rule, useful to establish valid properties for every element in a given set. That is, if P is a property which holds for n, then P holds for the consecutive of n, and so on

The whole proof is described in terms of the inductive step ("If P is a property which holds for n, then P holds for the consecutive of n") and to an iteration ("and so on"). The base case and its role for the iteration are however absent.

The base case as unrelated to the general theorem

In the previous examples the whole proof by MI was described in terms of the inductive step without any reference to the base case. In the following example, the situation is, in a certain sense, opposite, since the reference to the base case is present whilst the inductive step seems to be missing:

Stud. 55/G2+: In the initial step, we prove a theorem for the most banal and simple case. Successively, we try to prove the theorem in general for any number/parameter, perhaps using some algorithms

The base case (called "the most banal and simple case") seems to be just a preliminary test for the student: only after having proved that the theorem is true for a simple case, we can start looking for a general proof for it. The case base, therefore, is seen as unrelated to the general theorem, whose proof is independently constructed after it. The expression "perhaps using some algorithms" could refer to the inductive step, but we do not have further elements to interpret the student's sentence.

The inductive step as proving $P(n) \wedge P(n+1)$

In many answers we found that the inductive step was described not as the implication $P(n) \rightarrow P(n+1)$, but as the conjunction $P(n) \wedge P(n+1)$. Moreover, in several of these answers the base case was absent (as in some examples of above). In these cases, therefore, proving by MI was described uniquely as proving that P holds simultaneously for n and n+1.

The following is an example of an answer of this kind. The student, from G2, in the question asking to describe MI, wrote:

Stud. 44/G2+: Proved that a property holds both for n and for n+1, then it holds for all the n

As the answer clearly indicates, the student describes MI as proving $P(n) \land P(n+1)$ to conclude that $\forall n\, P(n)$ ("it holds for all the n"). This description of MI is interesting since, from a logical point of view, it is true that once proved $P(n) \land P(n+1)$ for a generic n, we can conclude $\forall n\, P(n)$. This, in fact, can be obtained simply applying UG to $P(n)$, which can be directly deduced from $P(n) \land P(n+1)$. The justification given by the student is the following:

Stud. 44/G2+: If a property P holds for n and for n+1. I can substitute n with 0 and continue to infinity proving that the property holds for all the natural numbers.

With this answer the student seems to say that, since both $P(n)$ and $P(n+1)$ hold, by substituting n with every natural number starting from 0, we obtain $P(0)$ and $P(1)$, $P(1)$ and $P(2)$, $P(2)$ and $P(3)$, and so on. Therefore, we can conclude that $P(n)$ holds for all the natural numbers. The argumentation contains some superfluous repetitions, but from a logical point of view is valid. It is, of course, based on the fact that in a prove by MI we prove $P(n) \land P(n+1)$, which is not what happens instead. In a certain sense, the student is giving a justification of why UG works.

The following example is similar to the previous one. The student, an undergraduate student in engineering, described MI as it follows:

Stud. 38/Extra: It is a procedure which tries to establish a universal law starting from some single particular cases.

The description seems to refer to an empirical induction more that to a MI. The justification s/he gave, however, starts from $P(n) \land P(n+1)$, as in the previous case:

Stud. 38/Extra: Because if the proposition holds for n and n+1, if we substitute to n its consecutive, we obtain that the proposition holds for (n+1) and (n+1)+1. And as n goes to infinity, we have that the property holds for all the natural numbers.

As noted in the previous case, the student constructs an argumentation for supporting that, since $P(n) \land P(n+1)$ holds we can conclude that $\forall n\, P(n)$. In order to do so, the student refers to the process of substituting n and n+1 with n+1 and (n+1)+1 and to continue this process "as n goes to infinity".

<u>MI as proving $\forall n(P(n+1))$</u>

Some answers were registered in which MI is described as 'proving that a property is true for any consecutive natural number'. To better explain this kind of answers, let us consider the following example. The student, from group G2, gave the following explanation of why MI works:

Stud. 4/G2+: Simply because if we verify that a property is true for the consecutive number of any natural number, then it will be true for any natural number as well. If we take k=36 and we verify that it is valid for any consecutive natural number, 36 is also the consecutive of 35.

The logical structure that the student describes is the following: we prove $\forall n(P(n+1))$ ("we verify that a property is true for the consecutive number of any natural number"), therefore $\forall n(P(n))$ since any natural number is the consecutive of a natural number (note that the inference used by the student is valid, except for n=0). The student, then, presents a generic example to support her/his argument: given 36 and proved that P is "valid for any consecutive natural number" it follows that P(36) is true, since 36 is the consecutive of 35. We can interpret answers of this kind noticing that it seems that for the student the inductive step $\forall n(P(n) \rightarrow P(n+1))$ is collapsed in $\forall n(P(n+1))$. Since in proving the implication of the inductive step we prove P(n+1) after assuming P(n), it could be that the student is interpreting this as a proof of P(n+1) for a generic natural number n, therefore a proof for $\forall n(P(n+1))$. This is an example of a possible interference between inductive step and general statement, as it was described in section 3.3.3.

Interference between the base case and the inductive step

In the following answer we can see an example of interference between the base case and the inductive step, as described in section 3.3.3. In particular the inductive step is described as proving P(n+1) from P(0).

Stud. 11/G1-: Proving that p(n) is true for n=0, therefore I can prove that it is true for n+1

The student describes MI in two steps. Firstly they refer to the base case ("Proving that p(n) is true for n=0"), then to what could see as an inductive step "therefore I can prove that it is true for n+1" which seems not to be described simply as proving P(n+1), as in the previous example, but as proving P(n+1) using P(0), as the logical structure seems to suggest ("Proving that […], therefore I can prove […]"). It is like, for the student, the inductive step consisted in the implication $P(0) \rightarrow P(n+1)$.

The justification of the meta-theorem: the explain induction process

In some answers we found descriptions of why MI 'works', (namely, justifications for the validity of the meta-theorem) which seem to refer to the Explain Induction Processes. I will present here three examples which show how this process can be constructed with different features.

The first example involves an Explain induction Process with direct form (i.e., chaining of MP):

Stud. 16/G3: If a property has been proved by induction with base case 0, therefore taken any natural number n, starting from 0 we can deduce that the property holds for 1, so, for the inductive step, it holds for 2, then it holds for 3 and so on until we arrive to n.

In this answer two processes seem to be present: the UG process and the Explain Induction process. The student's argumentation, in fact, is constructed as it follows: firstly, we fix a generic natural number n, then by iteratively applying the inductive step from the base case 0 we conclude that P(n) must hold (i.e., the Explain induction process). Therefore, since n was generic, we can conclude that P(n) must hold for every natural number (the UG process). This last part is not explicitly written by the student who, however, seems to be aware of that since s/he is answering to the question asking to explain why MI assures that a predicate P(n) is true for every natural number.

The following is an example involving an Explain induction process still with direct form but now involving a generic example, instead of UG. This is what a student wrote to justify why MI 'works':

Stud. 14/G3: By contradiction, let us say that it does not hold for 123, for instance. However it hold for 1, and therefore for 2, and therefore for 3, and in a finite number of passages we arrive to 123.

The student constructed an Explain induction process to prove, as a generic example, that P(123) must hold. Note that despite the reference to a proof by contradiction, the student's argumentation has a direct form: s/he has proved that MI works in general by showing that "in a finite number of passages" we can see that the property is true for a given number, for instance for 123.

Lastly, the following is an example involving an Explain induction process, this time with indirect form (i.e., chaining of MT) constructed within an indirect argumentation[86]:

Stud. 15/G3: If it is true for 0 and it is valid that 'true for n' implies 'true for n+1', then if it were false for a given number it would be false for the precedent one and thus for precedent of that, and so on to zero, and therefore we would obtain that it is false for n=0… contradiction

The justification of the meta-theorem: the successor process

In some answers we registered a justification for the validity of MI which involves what I have called the Successor Process. This process was described in the refinement of the Genetic Decomposition of MI that I proposed in Section 4.4.2. With this term, I indicated the process through which a subject recognises that any given natural number n* greater than n_0 can be obtained by iteratively applying the successor function ($x \mapsto x+1$) from n_0 a finite number of times.

As said, we found traces of this process in a few students' justification for the validity of MI. This, for instance, is the explanation given by a student from G3 of why MI 'works':

Stud. 13/G3: Any natural number is obtained from 1 as a finite sum of 1. Therefore if adding 1 preserve a certain property, any number, reachable with a finite number of steps, satisfies the property, thus any natural number satisfies it.

As we can see from the student's answer, the Successor process is used to justify the validity of MI: "Any natural number is obtained from 1 as a finite sum of 1. Therefore…".

The following example contains a similar description of the Successor process, but with a more complex mathematical symbolism. The student, again from G3, wrote:

Stud. 12/G3: [A] proof by induction is developed in two steps: in the base case we prove that a certain property holds for a fixed integer N, in the inductive step we show that if it holds for an integer M, then it holds for M+1. We notice that taken any given integer H greater or equal to N, we can write it as N+k for some integer k, thus H = (…((N+1)+1)…+1)+1 where we summed 1, k times. The property is true for N, thus for N+1, thus for N+2 = (n+1)+1 and so on until H.

After describing base case and inductive step, the student justifies the functioning of MI showing that it is possible to deduce that the property is true for any given number H. As in some previous examples, the student constructs the Explain Induction process to connect the validity of the property for the case base (the number N, in student's notation) to the validity of the property for a generic number H greater than N ("The property is true for N, thus for N+1, thus for N+2 = (n+1)+1 and so on until H"). However differently than in the other analysed examples, here the student seems to give a justification of MI also by constructing a Successor process. Indeed s/he explicitly states that "taken any given integer H greater or equal to N, we can write it as N+k for some integer k, thus H=(…((N+1)+1)…+1)+1 where we summed 1, k times". With this sentence the student highlights an aspect which, although fundamental from a logical point of view, generally remains implicit: in a proof by MI a crucial point for its functioning is that any number for which we want to prove the proposition P must be reachable from the base n_0 with a finite number of applications of the successor function.

[86] The term "indirect argumentation" is intended as in Antonini (2019).

The justification of the meta-theorem: images, metaphors, and analogies

In many students' answers, the Proof by MI, or its justification, was described with reference to some images, metaphors, or analogies with other topics in Mathematics. The study of metaphors and of their roles within teaching-learning processes is a wide research direction in mathematics education. For a general review of this theme, with the references of the primary studies see Soto-Andrade (2014). In this brief section of the thesis, I focus on how images and metaphors are used by students to describe the proof by MI or to explain why MI 'works'. In a certain sense, thus, these metaphors can be seen as some subjective and non-formal justifications of the meta-theorem involved in the Proof by MI.

These are some expressions used in some students' answers when referring to MI:

- "A chain reaction"
- "A line of dominoes"
- "A cascade effect"
- "Going up on a ladder step by step"
- "An iterative algorithm"
- "It is like writing a never-ending program"
- "It is a kind of a 'bootstrap' proof" [87]
- "Pushing a property" from the case base to the subsequent numbers
- "A spark that let us start a fire"

I will focus here on a few examples of these answers which contain some interesting elements for our analysis. The first example is of a student from G2 who described MI both with reference to a "Chain logical process" and to some falling dominoes:

> Stud. 81/G2+: A proof by induction works because the inductive step says that a proposition holds for n then it holds for n+1. We have then to find a natural value for which this proposition is true, and a chain logical process begins so that the property holds, or the proposition is true, for every value bigger than the found number: very intuitive thinking to some domino tiles that if the first one falls then all the consecutive tiles will fall inevitably"

Let us focus on two aspects of this answer. Firstly, we can notice that it shows how the student connects the falling dominoes image to the "chain logical process", which seems to correspond exactly to the Chaining Process. The student clearly says that one can "intuitively" thinks to this process as a falling domino which makes all the consecutive dominoes fall. A second aspect to notice is the student's use of the adverb "inevitably" when describing the falling dominoes. With this expression the student is indicating that if the first tile falls, it is not possible to avoid the other tiles to fall. Out of the metaphor, with this term, the student seems to refer to $\forall n\, P(n)$ as a logical ("inevitably") consequence of inductive step and case base. Therefore, we can interpret this answer as a sign of an intuitive acceptance of the meta-theorem by the student who perceives it as a certain and inevitable.

As a second example I will show the answer of another student from G2, who still refers to some falling dominoes, but adds some other elements to the metaphor:

> Stud. 97/G2+: The mental image that was useful for convincing myself the first times that I dealt with a proof by induction is the following: let us image to have an infinite line of dominoes which are progressively numbered. The first hypothesis of the principle of induction is equivalent to say that a certain tile x falls, the second is equivalent to say that all the tile after the x are at the

[87] In Computer Science the term 'Bootstrapping' usually indicates "to use the output of one piece of software as input to another, so that the results of the first process help to improve the results of the next one, and so on". (https://www.macmillandictionary.com/dictionary/british/bootstrap).

right distance so that a tile's fall causes the consecutive tile's fall. Therefore, if the tile x falls, a sort of "chain reaction" will begin letting all the tiles after the x fall. However, if the tile x does not fall, the ones after, even if at the right distance, will not fall if there is not a falling tile before them.

The student's answer is extremely interesting. Firstly, we can see the student's attention to express the semantic connections with MI in the metaphor of the falling dominoes: the case base corresponds to a first falling tile, the inductive step to the fact that two tiles are "at the right distance". The student says that all the tiles are "at the right distance" instead of saying that a tile's fall will cause the consecutive tile's fall and in this way s/he enriched the metaphor with a further physical component. By doing so, some logical components (the "if…then" structure of the inductive step) are represented by a spatial feature of the image: the distance between two tiles which is "right", allowing a tile to push the following one when falling.

A second interesting point of the student's answer is in the last sentence: "However, if the tile x does not fall, the ones after, even if at the right distance, will not fall if there is not a falling tile before them". As a first interpretation we could say that the student is describing the necessity of the case base in a proof by induction saying that even if the tiles are at the right distance (the validity of the inductive step) it is not possible to conclude that they will fall. However, if we analyse the sentence from a logical point of view it may say something different. One could interpret the sentence "If the tile x does not fall, the ones after […] will not fall" as indicating the fact that if a tile falls (i.e. $P(n^*)$ is true) we can be sure that all the previous tiles have fallen (i.e. $P(n)$ is true for every $n \leq n^*$), which is indeed logically equivalent to what the student wrote. This is true for the rules of the domino game since we can only push the first tile of the line, however it is not mathematically coherent with MI. The validity of the inductive step and of $P(n^*)$ does NOT imply that $P(n)$ is true for every $n \leq n^*$. Let us consider, for example, the Task A of the survey in which it was given $A(103)$ true and the validity of the inductive step. Using the falling dominoes metaphor to solve this task could be misleading. Indeed if one considers $A(103)$ as "the tile 103 has fallen", and the fact that, since it is a domino, if a tile has fallen it means that the first one has been pushed, one may conclude that also the first 102 tiles have fallen, therefore that $A(n)$ is true for every $n \leq 103$. I wish to clarify that I am not saying that this is what the student considered in this example is meaning. In Task A and B, in fact, the student has answered correctly to every questions. What just said, however, shows that the falling dominoes metaphor could bring with itself some implicit meanings which are not mathematically coherent with MI. The just presented example is also interesting because it shows that the student in the previous tasks has answered in a way which is only partially coherent with the metaphor s/he used to describe MI: the part which does not align the mathematical meaning of MI ("If the tile x does not fall, the ones after […] will not fall") is not used by the student who, in fact, did not answered "True" for the questions $A(n)$ with $n<103$ and did not answered "False" for the questions $B(n)$ with $n>37$, as the sentence of above would suggest. This last observation highlights another point about the metaphors related to MI which is that it is not said that the metaphor used by a subject to describe MI matches with the way the subject actually uses MI when solving a problem, but instead it is possible to register a distance, even some inconsistencies, between the two of them.

A further of a metaphor used to describe the functioning of MI is of a student from G1+. In this example the "Chain reaction" metaphor is used in a way which is not mathematically coherent with MI. The student, from G1, as a justification of why MI works wrote:

Stud. 40/G1+: Because it is like a chain reaction. If the property is true for the number 53 it must be true for 52 as well so that it could be for 53, and so on

The student is referring to a chain reaction, however it seems that the direction of this reaction is not from 53 to bigger numbers but rather the opposite. The student says that since for 53 the property is true, "it must be true" for 52, "and so on" such as s/he was saying that the property must be true for every number lower than 53. As a confirm of this we can notice that the same student, in Task A, answered "True" for every question A(n), even for those involving the numbers lower than the known truth value n_0=103. As a justification for these answers, the student wrote:

> Stud. 40/G1+: […] the property is true also for all the numbers that precede it [103], because in order to be true for 103 it must be true for 102, 101, etc…

The "chain reaction" metaphor could help in interpreting this and the previous answer. If MI is seen as a "Chain reaction" which necessarily starts from 1 (or 0), and if we know that for 53 (or for 103) the property is true, it means that the chain reaction has actually started and reached 53, therefore it has also reached 52, otherwise it would not have reached 53, thus it has also reached 51, otherwise it would not have reached 52, and so on. This example provides us a further possible interpretation for the several answers of students who like this student, in the Task A, selected that A(n) is true for all the natural number lower than the number corresponding to the base case (i.e. 103).

These last two examples, respectively involving the 'falling dominoes' and the 'chain reaction' metaphors, highlight some issues about metaphors in relation to MI. First of all, as observed, these metaphors might bring with themselves some implicit features which do not align with the mathematically correct meaning of MI. Secondly the fact that a subject is able to correctly describe MI with the use of images such as the falling dominoes one, does not guarantee by itself that the subject's conceptualisation of MI is coherent with the mathematically correct one.

<u>MI as a not convincing way of proving</u>

The last category of answer that I present is not specifically referred to some components of MI as described in the model of MI as a triple, but it refers to the intuitive acceptance of MI by some students. In particular, in some answers we registered a comment that seems to highlight that MI was not intuitively accepted by the student, even if it was correctly described. This aspect is coherent with the framework of intuitions by Fischbein, described in chapter 5, following which the intuitive acceptance of a theorem is transversal to the formal knowledge of it. This means that a student could have a formal knowledge of a theorem (being able to correctly describe it or reconstruct) without having the feeling of certitude that 'it must be so' related to a theorem (the intuitive acceptance).

The first example that we will present is taken from a G2 student's answer. S/he described MI as it follows:

> Stud. 93/G2+: In case this proposition is true from n=0 to x=n and we prove that it holds for x=n+1 as well, then it holds for all the natural numbers

The student describes MI as proving P(0), …., P(n) and then P(n+1). This seems to be similar to some previous answers where MI was described as proving separately P(n) and P(n+1), but here there is also a reference to the proof of P for every number from 0 to n. It could be that the student is describing the inductive step in a "strong" MI: $\forall n\ [(\forall k \leq n\ P(k)) \rightarrow P(n+1)]$, but there are not further elements to support this point. After having described MI, the student wrote the following answer to the question asking to say if there were any unclear or not convincing aspects of MI:

Stud. 93/G2+: In this moment no. Even if for very big numbers, I think, there could be some cases where this property does not hold.

The student, who is now talking about MI, states that it is possible that for very big numbers MI could stop working. This is an example of absence of intuitive acceptance on the third level (the level of generality of a proof), since the student clearly indicates that "for very big numbers" some possible counterexamples may be found.

The following example shows a similar case, but now it involves a student whose description of MI seems to align with a mathematically correct one. The student described MI as it follows

Stud. 90/G2+: We verify that the related claim is true for a number and then, with the assumption that it is true for a not specified number n, we prove that it is true for n+1 as well.

Then, in the justification of the why MI works, s/he wrote:

Stud. 90/G2+: I think that the difficulty of proving that a claim is true for all the natural numbers is that these numbers are infinite. Thanks to this method we know that if it is true for a number, it will be true for all its consecutive numbers, solving therefore the problem of the infinity.

However, in the final comment, the student seems to reveal to be not completely convinced by MI:

Stud. 90/G2+: I think that, sometimes, this proof brings to some inexplicable paradoxes.

Unfortunately, we do not know what the student is referring to when saying "inexplicable paradoxes", however the student's answer is a sign of a not complete and intuitive acceptance of MI since it indicates that the student is experiencing some sort of cognitive conflicts when dealing with MI (note the term "inexplicable"). The student may have a formal knowledge of the proof by MI, being able of describing it and of constructing in for proving some statements, however s/he does not seem to be intuitively convinced by it ("sometimes, this proof brings to some inexplicable paradoxes").

This concludes the analysis of the answers to the open questions at the end of the survey. The next section contains a summary of the results that I just analysed. In Chapter 13 of this thesis, these results will be further discussed with reference to the research questions.

9.3 OVERALL CONCLUSIONS OF THE FIRST EMPIRICAL STUDY

One of the main aspects emerging from the results of the survey is the confirmation of how mathematical induction could be problematic from a cognitive and didactical point of view. With the data analysis it was possible to observe that some problematic aspects seem to involve transversely the vary groups. Despite having registered a higher number of correct answers in the groups of students with more experience with MI, however we have still registered a relevant number of not correct answers in every group, in master's students in mathematics (G3) as well. This seems to suggest that some difficulties or problematic aspect related to MI could be particularly robust and could persist even in expert students.

In relation to the problematic aspects of MI, the results of the survey allowed us to highlight some of them, partly confirming other results in the literature, partly identifying new ones.

An interesting aspect, registered in the answers from the first three tasks, is the fact that the percentages of correct answers decreased when the numbers involved in the questions were progressively more distant from the base of the induction n_0. This highlights how the Explain Induction process can be problematic to be interiorized. In particular, for some students it seems not to be immediate to recognise that, starting from the inductive base and the inductive step, it is possible to construct a (potentially infinite) chain of conditional inferences (MT/MP). This phenomenon, which I interpreted in terms of interiorization of a complex process obtained as a coordination of other processes (chaining of MP/MT), has been registered transversely in every group. A critical aspect emerging in the GD of MI presented in the framework was the transition from potential to actual infinity in the encapsulation of the explain induction process which allows the subject to recognise that the chain of MP, starting from the base n_0, reaches the whole set of natural numbers greater than n_0. The results that we just discussed, however, highlighted something more. Indeed, they showed that the construction of this chain of inferences could be problematic even when it involves only a finite number of steps, as it happened in the numerical questions of the first three tasks. Generally, when MI is introduced to students, its functioning is explained by explicitly constructing the first steps of a chain of MP inferences starting from the base of the induction to then conclude with a "and so on". The above presented results, however, highlight that the process which underlies this "and so on" could be (re-)constructed by students in some very different ways, with chains of MP that do not reach every natural number, that stop after a few steps, or that do not even start.

In the GD of MI presented in the framework, beside the Explain induction process in direct form, I introduced the Explain induction process in indirect form. With this process a subject recognises that, starting from the validity of the inductive step and the fact that $P(n_0)$ is false for some n_0, it is possible to conclude that $P(n)$ must necessarily be false for every $n \leq n_0$. Firstly, the survey highlighted that, also for this second process, we registered an analogous phenomenon to what it was registered for the Explain induction in direct form: the progression of correct answers decreases when the numbers involved in the questions are more distant from the known value n_0. However, the data analysis allowed us also to register some differences between the two processes. In particular, comparing the questions involving the chaining of MP with those involving the chaining of MT, we registered that, in every group, the second ones have always obtained a lower percentage of correct answers that the first ones. This point firstly confirms other studies which show that the logical inference of MT can be more problematic than MP for students. Secondly, this shows that similar problematics seems to be also present in the chaining process for these conditional inferences.

Further problematics aspects for the students involved in the survey relate to the inductive step. First of all, we registered the difficulty of several students in recognising the validity of the implication contained in the inductive step as independent from knowing the truth value of antecedent and consequent. This result confirms what already registered by Fischbein and Engel (1989). Secondly, we observed that some students did not recognise the inductive step as theorem itself. In particular, when an inductive step was presented alone, without a truth/false value for some n_0, as it happened in task D, we registered two almost opposite categories of not correct answers: on one side, for some students the inductive step was interpreted as an 'empty' statement from which it is not possible to infer any information regarding the involved property (indeed these students selected "we cannot know" for every statement of the task); on the other side, for other students the inductive step alone implies the truth of the property for every natural number (indeed these students selected "impossible" for every statement addressing the possible presence of some numbers for which the property is false). Thirdly, in the final open questions asking to describe MI, we observed that in several students' answers the description of the inductive step was not correct, showing some inconsistencies, from a mathematical point of view with the logical structure of MI. This aspect confirms other studies which highlight that one of the most problematic aspects for student in relation to MI is to correctly manage the complexity of the logical structure of the inductive step (Avital & Lebeskind, 1978; Ernest, 1984; Dubinsky & Lewin, 1986; Nardi & Iannone, 2003).

In the description of the model for a theorem by MI as a triplet, described as a system of three different triplets (the two sub-theorems corresponding to the base case and the inductive step, and the meta-theorem), it has been observed that a crucial aspect emerging from the model is that a subject conceptualises these theorems both as independent and in a mutual relationship. Some of the open answers at the end of the survey, highlighted how this aspect could be particularly difficult for students. Indeed, we registered several answers in which some components of the model were totally absent (as happened, for instance, in the extracts showing the absence of the case base or of the inductive step in students' description of MI), and other answers in which we registered some interferences between the triplets composing a proof by MI (as called in section 3.3.3). The presented model of a theorem by MI, allowed us to describe the logical complexity of MI, in terms of relationships between the involved statements, proofs, theory and meta-theory. The collected results from the survey showed how this complexity could be problematic for students.

One of the aims of the survey was to investigate students' intuitive acceptance of MI. The results seem to provide us a strong indication on this aspect, highlighting that even if for some students MI and its justification is described as obvious and self-evident, for some others this is not the case. This indicates that MI could be seen as not intuitive by students. This has been registered, firstly, in some open answers at the end of the survey in which some students stated themselves to have doubts about the validity of MI and, in some cases, they were students who had correctly described MI or correctly answered the tasks' questions. Secondly, traces of a not intuitive acceptance of MI can be also registered by analysing the answers for the tasks' questions. Here, some aspects of MI were modified in relation to how traditionally MI is encountered by students: the base was moved to number distant from 0 or 1, or it was given a false valued base case, or an inductive step was given without any information of the base case. Many students even with some experience of MI could not re-construct a proof by MI for these different cases. Other students instead, even after having answered correctly to some questions, stated not to be sure about their answers. In relation to the intuitive acceptance of MI, moreover, a further element to observe is what we have already said about the (re-)construction of the explain induction process. Results, in fact, highlighted how for several students it was not intuitive that, starting from the validity of base case and inductive step, one can conclude the truth of a proposition for every natural number greater than the base. As observed, this aspect involved both students with some experience with MI and students without.

Another problematic aspect emerging from the results is the difficulty (perceived or not) of some students in expressing themselves and reasoning on the meta-level involved in the survey, where the objects were generic (and unknown) propositions on the natural numbers. Indeed, we registered that some students answered "We cannot know" to some or even to all the questions of the survey claiming that, since the involved properties were unknown, it was not possible to answer to the questions. As highlighted in the model of theorem by MI as a triplet, any proof by MI involves a meta-theoretical level, the one of the meta-theorem through which one can prove the general statement starting from base case and inductive step. To describe and reason on MI as a proving schema, thus, is also necessary to situate the discourse on this meta-level. When MI is introduced to students, generally, it is described in meta-level terms, by referring to a generic predicate P(n). The results of the survey suggest that a particular attention should be paid to this aspect, because for some students it could be problematic to recognise that when we talk about MI the discourse may situate on a meta-level and that in this level the justification of the validity of MI is described.

Finally, the results of the survey showed that some issues for students about MI could be related to the traditional images and metaphors which generally are paired with the description of MI, as, for example, the falling dominoes metaphor. As observed the problematic aspect of these metaphors could be dual. From one side these metaphors might bring with themselves some implicit features which do not align with the mathematically correct meaning of MI. On the other side, the fact that a subject can correctly describe MI with the use of images such as the falling dominoes one, does not guarantee by itself that the subject's conceptualisation of MI is coherent with the mathematically correct one. Metaphors, however, seem to also have an important role in the construction of meaning of MI, specifically allowing the subject to represent and talk about MI as an object, that corresponds, in APOS terms, to the encapsulation from process to object of MI. This was noticed, for example, in one of the examples of above in which the student used the metaphor of the two little bricks to be put together to represent as objects base case and inductive step within a proof by MI. These comments suggest that, from a didactical point of view, it is important to be aware of both the potentialities and the limits of images and metaphors related to MI, in particular paying attention to those (often implicit) meanings that a metaphor brings with itself and that are inconsistent with the mathematical meaning of MI.

10 FIRST ANALYSIS OF THE INTERVIEWS – SIGNS OF INDUCTION

In this and the next two chapters I will present the results emerging from the analysis the second empirical study of this thesis, the task-based interviews. The interviews, as introduced in the methodology, have been analysed addressing mainly the research question RQ4 which investigates the use and production of signs by students involved in the construction of recursive argumentations and proofs by MI.

In this chapter I will focus on presenting two particular categories of signs which I identified, and which seem to have a crucial role in the generation and in the construction of a recursive argumentation or of a proof by MI: the Linking and Iteration signs. In the next chapter (Chapter 11) I will present a second aspect emerging from the data analysis, which is that, focusing on students' semiotic production, it is possible to distinguish signs which can be interpreted as belonging to two categories: the ground level signs and the meta level signs. Finally, in Chapter 12, other results emerging from the analysis of the interview related to other research questions will be presented and the overall conclusions of the whole empirical study with the task-based interviews will be discussed.

10.1 SIGNS OF INDUCTION: LINKING AND ITERATION SIGNS

The interviews were firstly analysed focusing on the students' production and use of signs involved in the generation and construction of recursive argumentations. Specifically, I looked for the presence of 'crucial' signs which seem to reveal the genesis of such argumentations or to support the students' construction of them. It can be helpful to remember the definition of recursive argumentation adopted in this thesis: an argumentation involving or referring to a repetition of an argument that supports successive statements, so that each of these statements is used in the argument supporting the following statement. As a result of this analysis, two particular kinds of such signs have been identified: **The Linking Signs and the Iteration Signs**. To introduce them I will present a paradigmatic example which will show how, adopting a multimodal semiotic perspective, it is possible to see a recursive argumentation emerging from a student's exploration of the problem. This example has been presented in Antonini & Nannini (2021) where the definition of Linking and Iterations signs as adopted in this thesis was introduced.

10.1.1 A first case study

Excerpt 10.1

Giuditta[88] is dealing with the chessboard problem. In the first minutes of the interview, she has explored the problem, producing some drawings and she has recognised that, for reasons of divisibility, it is not possible to completely tile any 2^nx2^n chessboards with L-shaped tiles covering three squares each. By minute 10:00 she has drawn an 8x8 chessboard and determined a tessellation which covers every square except for one (figure 10.1).

[88] In this and all the subsequent cases, when the subject is an expert student (PhD or last year master's in mathematics) it will be indicated only the student's pseudonym. The complete indication of the student's university grade and the field of studies is presented in table 8.3, section 8.4.4.

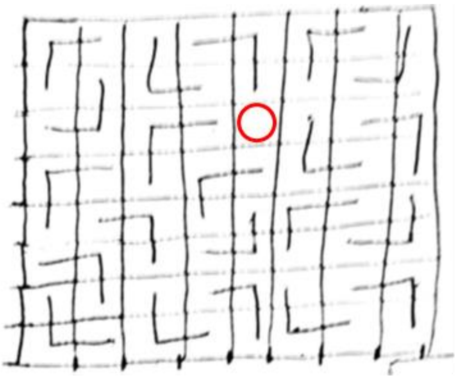
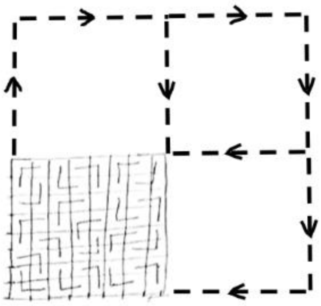


Figure 10.1. *Giuditta's drawing of the 8x8 chessboard completely tiled, except for one square. The red dot is added by me to indicate the square which has remained out of the tessellation.*

After this, the interviewer asks Giuditta if this property is also valid for other cases, for example for a 16x16 chessboard.

	Who	Speech	Gesture and inscription
1	**G**	Sixteen by sixteen…	*With her left middle finger and the tip of the pen in the right hand, she points to two vertices of the 8x8 chessboard drawing.*  *Fig. 10.2*
2	**G**	But then I have another three of these squares.	*She keeps her left middle finger on the vertex, and with the pen in the right hand she indicates respectively to the right, upper right, and above the drawing of the 8x8 chessboard.*
3	**G**	Here.	*Pointing with the left hand to the drawing of the 8x8 chessboard, she follows with the pen (without marking) the perimeter of 3 squares.*  *Fig. 10.3. The arrows indicate the trajectory of the pen over the sheet.*

4	**I**	Ok.	
5	**G**	And then there would be left out one, one, one and one.	*[one, one, one and one] She points to the drawing of the 8x8 chessboard on the sheet and to three points corresponding to the three other squares she just represented with the movement of the pen.*
6	**I**	Ok.	
7	**G**	And so, I would think to put three of them together, somehow. And then, there would always be one left out?	*She points to the sheet where the inscription of above is.*

This and all the other excerpts of this chapter will be analysed following their diachronic development line by line (diachronic analysis of the bundle), focusing when necessary on the composition of the bundle in a specific line (synchronic analysis of the bundle). Moreover, for what concern the analysis of gestures I will also refer to the specific theoretical elements introduced in section 6.3.1 (See Table 6.1 for a summary).

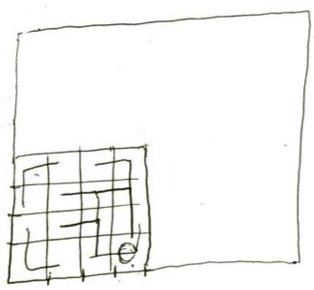
The bundle produced by Giuditta in line 1 reveals an interesting element. In this moment, on the sheet there is the drawing of the 8x8 chessboard (fig.10.1) and no other written inscriptions referring to a 16x16 chessboard. Giuditta says "sixteen by sixteen" and at the same time she points to two vertices of the drawing of the 8x8 chessboard (fig.10.2). This is a case of a speech-gesture mismatch: she refers to something through her speech (a 16x16 chessboard) and to something else through her gesture (an 8x8 chessboard). As observed in the presentation of the conceptual framework related to the semiotic perspective (Chapter 6), Goldin-Meadow (2003) highlights the cognitive potential of a mismatch in the representation of a new idea. In this case, the semiotic and the mismatch offer Giuditta the possibility to represent simultaneously two different chessboards (8x8 and 16x16). Then, in lines 2-3, Giuditta describes how these two different chessboards can be connected, noticing that a 16x16 chessboard can be obtained by adding three other 8x8 chessboards to the previous one. To represent this, she produces a series of signs: while saying "I have another three of these squares", with one finger she points to the inscription of the 8x8 chessboard and with the pen in the other hand she points the sheet respectively to the right, upper right, and above the inscription; then with the left finger still pointing to the inscription she follows with the pen the perimeter of three other square (fig.10.3). In summary, four 8x8 chessboards are represented: one by a written inscription, and three by Giuditta's speech and gestures. The speech-inscription-gesture bundle (lines 1-3) represents a 16x16 chessboard composed by four 8x8 chessboards. As a unit, it can be seen as a particular sign referring to two chessboards and to their relationships. This sign allows Giuditta to access the link between the tessellation problem in the case n=3 (8x8) and in the case n=4 (16x16).

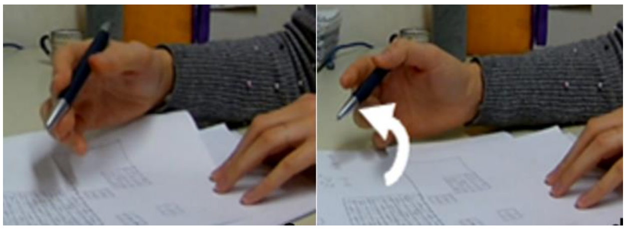
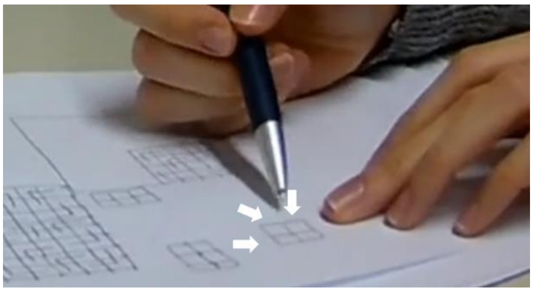
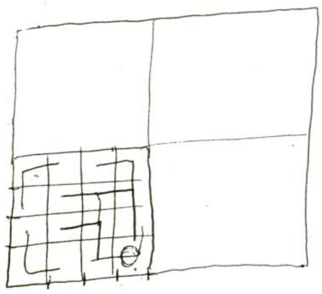
In the subsequent part of the excerpt (lines 5-7), Giuditta conjectures that the 16x16 chessboard can be tiled except for one little square (a square 1x1) and she imagines doing it by using the tessellation of four 8x8 chessboards. In each of them one small square would be left out, thus 4 squares in total, ("there would be left out one, one, one and one”), but three of them can be covered with an L-shape tile (“I would think to put three of them together”). Therefore, also the 16x16 chessboard would be tiled except for one little square. The sign representing the 16x16 chessboard as composed by four 8x8 chessboards that she has produced in the first part of the excerpt has a crucial role in the conjecture generation. In particular it enables Giuditta to anticipate the fact that the 16x16 chessboard can be tiled using the tessellation of the smaller one. Giuditta still does not know in which way three little squares remaining out from the tessellation of three 8x8 chessboards can be put together (she says “somehow”), however she seems to generalise her conjecture to every chessboard of the problem, expressing it as a question: “And then, there would always be one left out?”.

In the following part of the interview, Giuditta focuses on verifying her conjecture for other cases, respectively the 2x2 chessboard (n=1), the 4x4 chessboard (n=2) and, finally, the 1x1 chessboard (n=0). Figure 10.4 shows her drawings. Differently from her previous reasoning with the 16x16 chessboard, each of these is tiled independently, without connection between them.

Figure 10.4. *Giuditta’s drawing of a 2x2 chessboard (no tessellation was drawn) and of a 4x4 chessboard with a tessellation which leaves out one square. No drawing was made for the case 1x1.*

After having tested her conjecture for these cases, Giuditta claims to be convinced of its truth. The interviewer asks to explain her reasoning.

[…]			
8	**G**	So, what I was thinking was …	*She extends the drawing of the 4x4 chessboard of fig.10.4 into a new drawing corresponding to a bigger square.*  *Fig. 10.5*
9	**G**	That to come, to move forward from n equal to one, to n equal to two,	*She makes an arc-shaped gesture in the air from left to right.*

			 *Fig.10.6*
10	**G**	practically,	*With her left middle finger, she points to a drawing of a 2x2 chessboard, next to the just drawn inscription.*
11	**G**	I have to put another three identical little squares.	*With the pen, kept in the right hand, she points specifically to three squares of the drawing of the 2x2 chessboard, still pointed by her left middle finger (fig. 10.7). Then she rotates the pen around the inscription as to circle it (fig. 10.8).*  *Fig. 10.7*  *Fig. 10.8*
12	**G**		*She draws two lines on the drawing in fig. 10.5 obtaining the drawing of fig. 10.9*  *Fig. 10.9*

Giuditta produces a series of interesting signs. Firstly, she draws a big square (fig.10.5) widening the drawing of the 4x4 chessboard. With this she is representing the 4x4 chessboard previously drawn as embedded in a new 8x8 chessboard. This sign can be seen as an evolution of the bundle produced by

Giuditta in lines 2-3 to connect the 8x8 chessboard with the 16x16 chessboard (fig.10.3). Then, while saying “to move forward from n equal to one, to n equal to two” she makes an arc-shaped gesture in the air from left to right (fig.10.6). This is an interesting sign. The gesture is iconic and refers to a path. With reference to Krause’s classification (2015), it is a gesture which belongs to the level of generality (3rd level) since it is detached from the concrete inscriptions of the chessboards. In this, it is an independent and general gesture. It appears here for the first time and does not refer to any drawings, any chessboards, or tessellations. With this, Giuditta does not refer to the specific aspects of the relationship between a smaller chessboard and a bigger one, but she refers to the relationship itself. Thus, the gesture can also be seen as metaphoric: it represents metaphorically the relationship between two cases of the problem corresponding to n=1 and n=2, as Giuditta says. From this we can see that the structure of a recursive argumentation is emerging. Indeed, in lines 10-11 Giuditta shows concretely how the 2x2 can be formed from the 1x1 chessboard. She says: “I have to put another three identical square” while pointing to three squares of the drawing of the 2x2 chessboard. Finally, while staying silent she draws two new lines on the drawing corresponding to the 8x8 chessboard she previously made (fig.10.5) obtaining a new drawing (fig.10.9) which represent an 8x8 chessboard obtained from four 4x4 chessboards.

In the next minutes Giuditta explores the inscriptions she has just produced, in order to find a way to use four 4x4 chessboards, each with a tessellation which leaves one little square out, to tesselate the 8x8 chessboard. She recognises that, if the little square reaming out is in one of the vertexes of the 4x4 chessboard, then, when putting together the four 4x4 chessboards to obtain the 8x8 chessboard, they can be disposed so that three little squares are in the centre of the big chessboard and therefore they can be covered with a new L-shaped tile. She further modifies the drawing of fig. 10.9 to represent this aspect, obtaining the drawing showed in fig. 10.10.

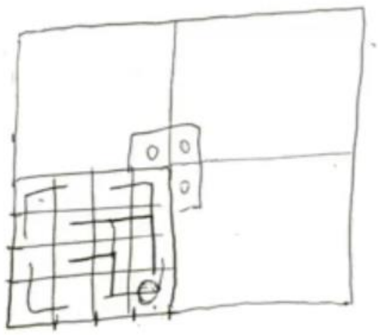
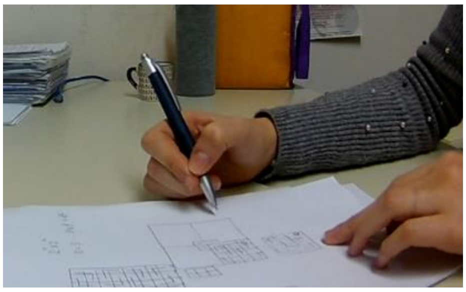
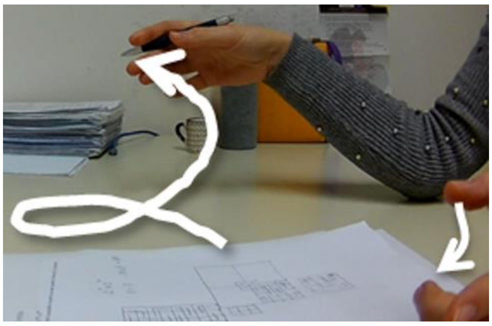


Fig.10.10

With this drawing, Giuditta is finally representing the complete connection between a 4x4 chessboard with a tessellation which leaves a little square out and an 8x8 chessboard with a tessellation which leaves a little square out. Remembering that Giuditta’s conjecture was that such a tessellation was always possible to create, we can see that now Giuditta has not simply proved that her conjecture is true for two different cases, but that the truth of the conjecture of two consecutive cases is connected: the fact that the conjecture is true for the 4x4 chessboard implies that it is true for the 8x8 chessboard as well. In other terms, a generic example for the inductive step has been constructed.

At this point, Giuditta concludes her argumentation.

[…]			
13	**G**	And this,	*With the pen in the right hand, she points to the inscription of fig. 10.10.*

			 *Fig. 10.11*
14	**G**	I can do it in general	*Starting from the sheet, with the pen in the right hand she makes a spiral movement in the air that widens as the right hand rises and concludes with spreading both the hands. Then she stays still in this in the final position for a few moments.*  *Fig. 10.12*

Giuditta does not write anything, and she uses very few words: “and this, I can do it in general”. However, analysing the whole semiotic bundle, her discourse is full of elements. In particular, her gesture seems to reveal the structure of her argumentation (a recursive argumentation). The gesture is articulated, and it develops in several components.

- Firstly (line 13), while saying “this”, she points to the drawing of fig.10.10. This is a contraction of the sign representing the connection between the tessellation of the 4x4 and of the 8x8 chessboards, which is now represented with a simple deictic gesture.
- Then (line 14), the gesture develops in a rather complex way: with her right hand Giuditta makes a circle around the inscription of the 8x8 chessboard and, while doing this she moves away her right hand from the sheet in a both upwards and outwards direction. In summary her gesture has a spiral shape in the air (fig.10.12). The upward direction takes the gesture from level 2 to level 3, following Krause’s classification. It is the first time that Giuditta produces this gesture in the air. The shift through levels and her words (“I can do it in general”) indicate the generality of the actions of tessellation.
- Moreover, the gesture grows wider away from her body to indicate the continuous construction of bigger and bigger chessboards (in mathematical terms, n is increasing). It is interesting to notice that, until now, her left hand has remained still with a finger of the drawing of the 4x4 chessboard (as shown in fig. 10.11), which could represent the starting point of the recurrence (in fact she has already directly verified the cases of the smaller chessboards).

If we look at her whole gesture together with the inscriptions on the sheet it refers to, we can interpret it as a unique sign representing an iteration which, starting from a 2x2 chessboard, allows to tesselate every bigger chessboard. The gesture in its totality is a gesture of level 2-3: it starts on the sheet, in which the base of the recursion is represented, and rapidly moves away becoming a gesture of the level of generality. The gesture, in summary, represents the entire argumentation that Giuditta has constructed to support her conjecture.

Finally, Giuditta concludes her argument keeping her hands still in the air for a few moments as if they contain the space in which her gesture took place. This space, to use an expression of McNeil (1992, p. 173) when describing an iconic gesture that indicates a point in the space, is not empty but "full of conceptual significance". In our case, this space is the location that contains the argumentation and its logical structure.

This conclude Giuditta's excerpt, in which we have seen her exploring the problem (lined 1-5), making a conjecture for it (line 7), and constructing a recursive argumentation to support her conjecture (lines 8-14). Analysing diachronically the evolution of the bundle in the excerpt we can notice that several times Giuditta has produced and used a particular category of signs. These are those signs representing the connection between two chessboards and their relationships. These signs have evolved during the excerpt, representing the connection between two specific chessboards (as in line 3 or 8), or the link between two cases of the problem (as the metaphorical arc-shaped gesture in line 9), or a link between generic chessboards of different sizes (line 14). Through the production and the exploration of these signs Giuditta seems, firstly, to generate a conjecture, and then, to construct and organise a recursive argumentation for it. These signs, therefore, seem to be crucial for her solution of the problem. The whole excerpt showed a repetition of signs of this kind, which is an example a what McNeill (2005) calls a *catchment* (see 6.3.1) which, following him, indicated and provides the discourse cohesion. Moreover, this repetition of sign over the time seems to support Giuditta structuring her argumentation. This is a confirm of a study by Arzarello and Sabena (2014), already analysed in 6.3.2, showing that catchments can contribute to support students in organise a mathematical argumentation. These signs produced by Giuditta are examples of what will be defined Linking Signs.

Another interesting sign that we have observed in Giuditta's excerpt, is her spiral gesture in the air at line 14, through which she seems to represent the entire iteration of actions of tesselate bigger and bigger chessboards using the tessellation of the previous one. As said this gesture allowed us to interpret her argument as a recursive argumentation. This is an example of those signs that will be defined Iteration Signs.

10.1.2 Linking ad Iteration Signs

The above presented example allows us to introduce two particular categories of signs which seem to have an important role in students' generation and construction of a recursive argumentation. These are:

- Signs produced or used to refer to two or more entities (objects, mathematical objects, problems, situations, etc.) and to their relationships, where these entities are seen in connection with two or more (consecutive) natural numbers. **These signs will be called Linking Signs (LS).**

- Signs that refer to iteration, or that are composed by a repetition (in time or in space) of linking signs, or that refer to a repetition of them. **These signs will be called Iteration Signs (IS).**

In Giuditta's excerpt, for instance, the LS were those signs produced or used to refer to two tessellated chessboards (the ones of dimension 2^nx2^n and the one of dimension 2^{n+1}x2^{n+1}, where n was specific or generic) and to their relationships (the fact that the tessellation of a chessboard be used to tesselate the subsequent one). Following the above definition, the two chessboards correspond to the two entities, whilst the 'n' and 'n+1' in their dimensions to the natural numbers connected to them. Note that in this case the two numbers were consecutive, but a LS could also involve entities seen in connection with numbers which are not consecutive, as some of the next examples will show. An IS, instead, was present in the last line of the excerpt when Giuditta made the spiral gesture while saying "and this I can do it in general", which refers to the iteration of actions of tesselate bigger and bigger chessboards.

In the next pages I will present other examples of LS and IS taken from the interviews, and I will discuss the contribution that their use and production might have for students during the problem solving activities. In other terms I will them to investigate the process of construction of a recursive argumentation or of a proof by MI. Before doing so, however, it is interesting to show that the presence of signs of this kind can be observed in other situations related to mathematical induction, such as in a proof by MI as a product, or in its communication.

Firstly, if we analyse a proof by MI as a product, we can see that a LS can be observed in the development of the inductive step. Let us consider, for instance, the usual proof by MI of the formula for the sum of the first n consecutive natural numbers.

<u>Statement:</u>

$$\sum_{k=1}^{n} k = \frac{\mathrm{n(n+1)}}{2}$$

<u>Proof by induction on n:</u>

The statement is true for n=1, moreover:

$$\sum_{k=1}^{n+1} k = 1 + 2 + \cdots + n + n + 1 = (1 + 2 + \cdots + n) + n + 1 =$$

$$= \frac{n(n+1)}{2} + (n+1) = \frac{(n+1)(n+2)}{2}$$

Therefore, […]

■

A crucial part in the proof of the inductive step is the expression $(1 + 2 + \cdots + n) + n + 1$. In this expression, with the use of some mathematical symbols, the parentheses, the sum of the first n numbers and the sum of the first n+1 numbers are represented and linked together, where the first one can be seen as embedded in the second one. This is a crucial part for the proof of the inductive step because it anticipates the use of the inductive hypothesis. In every proof by induction of this kind, some algebraic manipulations are required when proving the inductive step. These manipulations can be interpreted as oriented to the construction of a LS, where these signs are specific mathematical symbols. Note that in other contexts, such as if the proof was orally presented, instead of using the parentheses a similar LS could be obtained with an inscription or with a gesture, as shown in figure 10.13.a-b. It is interesting to notice that the sign represented in figure 10.13.b becomes a LS only if we consider the bundle composed by inscription (the sum written on the board) and gesture (the iconic gesture representing the two parenthesis). With this bundle the subject is representing the sum '1+…+n' as embedded in '1+…+ n+1'.

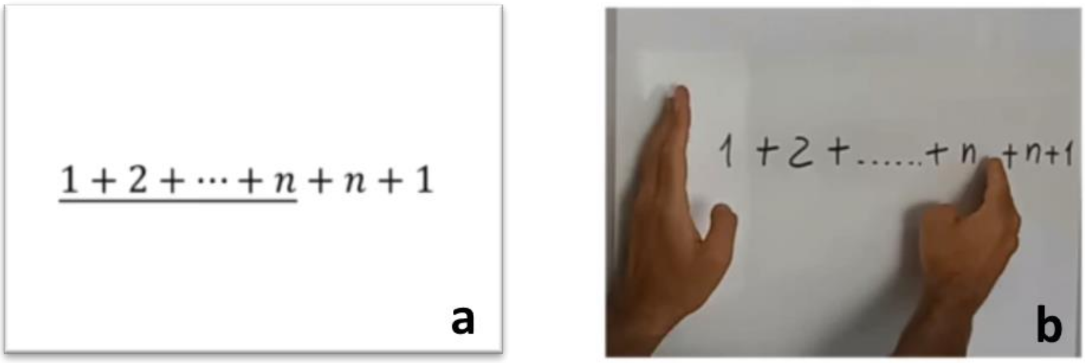


Figure 10.13. *Linking signs representing the sum of the first consecutive n+1 natural numbers as the sum of the first consecutive n natural numbers, plus n+1. The LS on the left (a) is only composed by inscriptions, whilst the other (b) is a bundle of inscriptions and gesture.*

Similarly, the presence of LS can be observed in other proofs by induction, not necessarily the ones involving an algebraic manipulation. For instance, in 8.4.3, I have presented a proof by induction for the formula for the number of diagonals of a convex polygon. In the proof of the inductive step, I used the following figure.

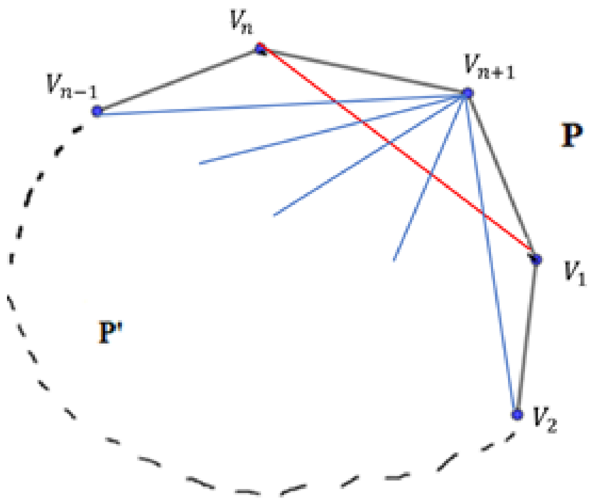


Figure 10.14. *Linking sign representing a n-vertices polygon as embedded in a n+1-vertices polygon.*

This is again a LS, representing a n-vertices polygon embedded into a n+1-vertices polygon together with some diagonals of the second one which are not diagonal of the first one. The figure was used to show that a n+1-vertices polygon has as many diagonals as a n-vertices polygon, plus (n+1)-3 (the blue segments of the figure), and plus 1 (the red segment). As it happened in the previous example, this LS makes it possible the proof of the inductive step.

If we consider the proof by MI as mathematical proof in broader sense, as in Douek's terminology (1998) (presented in section 3.1), and thus also extended to those proofs which belong to the history of mathematics, then it is possible to observe the presence of signs that can be interpreted as IS. In Pascal's proof of the Consequence XII, addressing a certain property of every line ("base") of the arithmetic triangle and considered by many as a clear example of proof by MI (see 2.7), the mathematicians, after having proved base case and inductive step (which he calls 'the two lemmas'), writes:

As a consequence, we will have that it holds for all the bases: in fact it holds for the second base, for the first lemma, therefore for the second [lemma] it holds for the third base as well, therefore for the fourth [base], and to infinity. (Pascal, 1665, section 1, pp. 7, my translation).

In Pascal's justification of the functioning of his proof we can observe the presence of an IS. Firstly, he constructs two steps of an iteration with a series of sentences, each of them can be interpreted as a LS which connects the truth of the proposition for a base of the triangle to the truth of the proposition for the following base (specifically from the second base to the third one and from this one to the fourth one. Then Pascal refers to the fact that the just described connection between consecutive bases of the triangle can be iterated 'to infinity'. The whole sentence, therefore, can be interpreted as an IS obtained firstly as a repetition of two LS and then with a sentence explicitly referring the repetition of those up 'to infinity'. It is interesting to notice that in a modern proof by MI such IS is generally not present since the proof is conducted only by proving base case and inductive step without any justification of the functioning of MI, which is not necessary for the validity of the proof and thus it remains implicit.

Moreover, the presence of LS and IS can be observed when considering other aspects of MI, not necessarily the proof by MI as a process or as a product, as in the previous examples. For instance, we can interpret, from a semiotic point of view and in terms of LS and IS, the classic images which generally accompany the non-formal justification of MI as a chain of infinite syllogisms, such as the one of the falling dominos. The whole image (either pictured or verbally described) is a rather complex sign composed by several other signs. Of those, one is the image of two dominos, one of each falling, potentially falling, or already fallen onto the other one. Out of the analogy this sign represents two propositions, P(n*) and P(n*+1), for which the truth of the first one implies the truth of the second one (i.e. an instance of the inductive step). This is therefore a LS. Moreover, the whole line of dominos (fallen, falling, or standing), representing the infinite syllogisms obtained by the base case and the inductive step, is a repetition of LS and, globally, can be interpreted as a unique IS.

The just presented examples of LS and IS refers principally to communicative aspects related to mathematical induction. It could be interesting to investigate the use of these signs in teaching, for instance analysing textbooks or their use by teachers. In this study however, coherently with the research questions, I will limit the investigation of LS and IS to the thinking processes related to students' resolution of problems.

In the next section I will present a gallery of episodes involving the use and production of LS and IS registered during the interviews. The aim of this section is firstly to provide a sort of phenomenology of these signs in order to further describe them and to analyse the characteristics they might present and secondly to highlight with series of examples how the construct of LS/IS allows the observation and analysis of processes.

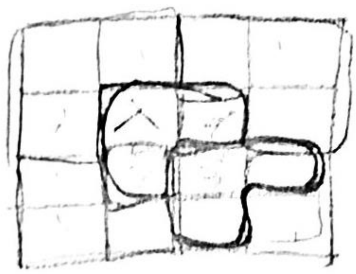
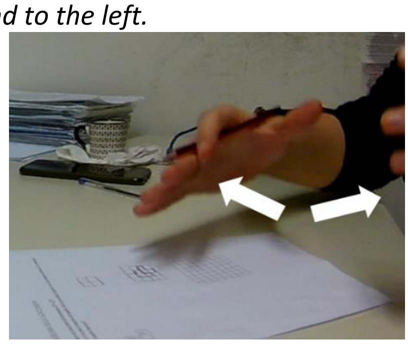
10.1.3 A gallery of episodes

10.1.3.1 A multimodal linking sign

In the perspective adopted in this study a sign is, potentially, a bundle of components belonging to different semiotic sets. In the following excerpt we can see a clear example of this aspect. In particular we will see an example of a LS obtained as a unit of different semiotic components.

Excerpt 10.2

Guido is dealing with the chessboard problem. In the first five minutes, after having read the text of the problem, he has observed that for a 2x2 chessboard and for a 4x4 chessboard it is possible to find a tessellation which leaves out only one square. Then he draws an 8x8 chessboard and he looks for a tessellation for this case.

	Who	Speech	Gesture and inscription
1	**G**	Ah! If I could do… if I put this in this way, here I can do like that…	*He changes the drawing previously made representing the tessellation of the 4x4 chessboard so that the little square remaining out occupies the bottom-left corner of the chessboard.*  *Fig. 10.15*
2	**G**	Then the one left out is here on this corner, which is better.	*With the pen he points to the left-down corner of the drawing, corresponding to the little square left out of the tessellation.*
3	**I**	Tell me again this, please. I did not understand.	
4	**G**	Because I noticed that…widening…	*He keeps the two hands open and close to each other over the sheet; then he moves the right hand to the right and the left hand to the left.*  *Fig. 10.16*

<table>
<tr><td>5</td><td>G</td><td>I would need to change a bit how I drew it here.</td><td>He points to the drawing of the 4x4 chessboard he just made in fig.10.15</td></tr>
<tr><td>6</td><td>G</td><td>So that I can do the successive step in a better way</td><td>He rotates forwards both the hands in the air.

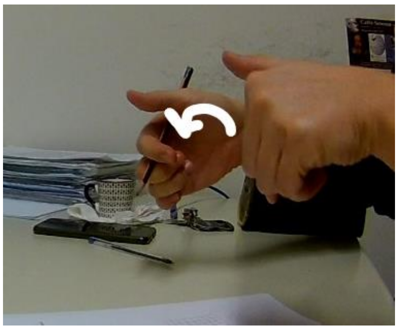



Fig. 10.17</td></tr>
</table>

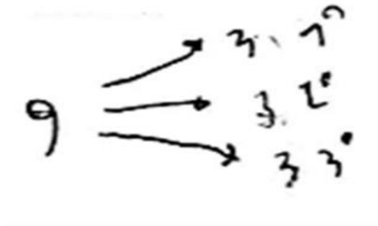
In line 1 and 2, Guido modifies his drawing representing the tessellation of a 4x4 chessboard. In the new drawing the little square remaining out of the tessellation occupies one of the corners of the chessboard. Then he says that this new configuration is better than the previous one. The interviewer asks for an explanation. In his answer, Guido produces and uses a series of sign that, all together, form a LS. He firstly makes an iconic gesture when saying "widening", as indicating the increasing dimension of the two considered chessboard (by saying "widening" he seems to refer to the passage from the 4x4 and the 8x8 chessboard). Then he points to the drawing of the 4x4 chessboard, when saying "I would need to change a bit how I drew it here", showing the modification in the drawing he just made. Finally, in line 6, he says: "I can do the successive step in a better way" and he makes a gesture which seems to represent metaphorically "the successive step" he is referring to. In summary in lines 4-6, Guido has described (without details) that the 4x4 chessboard, if tiled in a proper way (as in fig.10.15) can be used in 'the successive step', i.e. to tile a 8x8 chessboard. Analysing the bundle in its totality we can see it revealing a LS which would not be revealed by each semiotic set separately: the inscription, for instance, contains only the drawing of the two separate chessboards (the 4x4 and the 8x8). The gesture with the widening hands needs to be interpreted together with the inscription (it is a gesture of the 2^{nd} level in Krause's classification); finally, the utterance "then I can do the successive step" does not refer to any chessboards or tessellation, but only to a step of an iteration. Considering the whole bundle, however, we can observe that he is representing two chessboards and the link between them, in particular referring to the fact that the 4x4 chessboard can be used to tile an 8x8 chessboard.

10.1.3.2 An iteration sign obtained as a repetition of linking signs

As said in the definition, an IS can be obtained as a repetition of LS. The following excerpt shows an example of IS of this kind.

Excerpt 10.3

Lorenzo, in this part of the interview, is dealing with the false coin problem. After reading the text he claims to remember the solution of a similar problem in which nine weights are given, all identical except for one which is lighter. Lorenzo describes the solution of this second problem.

	Who	Speech	Gesture and inscription
1	**L**	In this case the game was: you split in three groups, then three coins, first group, three coins, second group, three coins, third group.	*He writes '9' on the sheet, then he draws three arrows starting from it and pointing to the right, then he writes '3' at the end of each arrow and, finally, '1st', '2nd', and '3rd'.*  *Fig. 10.18*
2	**L**	Then you say: we take two groups, we weight them and since there are three possible results, that are one is lighter, the other is lighter, or even, I can determine in which group the lighter one is. Then I iterate the procedure on the others.	

Later, Lorenzo tries to generalise the just described solution for a group of n coins:

[…]			
3	**L**	So, the reasoning would be… n coins, I split in three groups,	*He draws a point and three lines starting from it and pointing down on the sheet.*  *Fig. 10.19*
4	**L**	I select one of them with one weighting, I split in other three groups,	*He draws a second point at the end of the most left line of the just drawn ones and then three other lines starting from this new point.*  *Fig. 10.20*

5	**L**	I select one of them with another weighting.	*He draws a third point at the end of the most left line of new ones and then three other lines from this point.*  *Fig. 10.21*

In the first part of the excerpt (line 1), when describing the solution of the problem for a group of nine coins, Lorenzo produces a sign (the inscription in figure 10.19, together with his speech) which represents how the group of 9 coins is linked to three groups of 3 coins each. This is a LS which allows Lorenzo to connect the problem of determining the false coin in a group of 9 with the same problem for a group of 3 coins. In the second part of the excerpt (lines 3-5) Lorenzo describes a possible strategy to solve the general problem, and to do this he draws an inscription which is composed by the repetition of the previous LS, slightly modified: this time it develops vertically and without any indication to the number of coins (fig. 10.21). This is an IS and with this, Lorenzo, describes the iterative solution of the problem. In the remaining part of the interview related to this problem, Lorenzo explores the IS he just drew, which corresponds to a mathematical tree, in order to find, in function of n, its height. The height of the tree, in fact, corresponds exactly to the number of weightings necessary to determine the false coin. Lorenzo concludes the problem with a solution similar to the one described in the a-priori analysis of the interview.

10.1.3.3 Contractions of linking or iteration signs

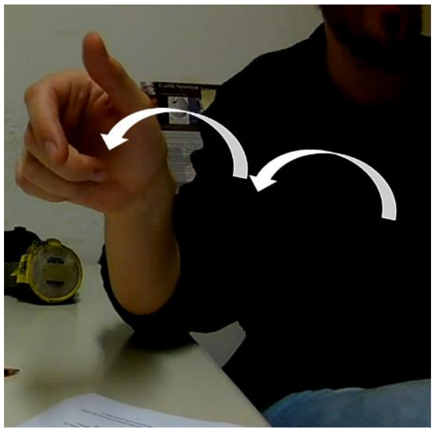
The contraction of signs is an interesting phenomenon to be considered within a semiotic perspective. Following Radford, "contraction is the mechanism for reducing attention to those aspects that appear to be relevant […] We need to forget to be able to focus" (Radford, 2008, p. 94). In other terms producing a contracted sign allows the subject to focus on some aspect of the sign considered as relevant, forgetting other aspects that appear not to be relevant in that moment.

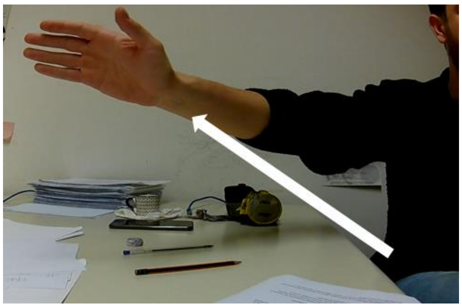
We have already seen an example of semiotic contraction involving a LS. In the excerpt 10.1, Giuditta firstly produces an elaborated LS to represent that four 8x8 chessboards can be used to tile a 16x16 chessboard (fig.10.3) which is then contracted in a rapid circular gesture when referring to the link between other two chessboards (fig.10.8). This contraction might show that in that moment Giuditta is focusing on the link between two chessboards, a link between two cases of the problem, rather than on the way this link can be obtained.

In the following excerpt we show an example of semiotic contraction involving an IS.

Excerpt 10.4

Guido, at the end of the interview, when commenting a proof by MI written by Matteo, an invented student of the task 'Meta questions on MI', provides the following justification for the validity of MI.

	Who	Speech	Gesture and inscription
1	**G**	You do the base case, which is true, and you verify it.	*He puts the right hand in front of him at the level of the table, touching his leg with the four fingers.*  *Fig. 10.22*
2	**G**	Then…	*He makes two consecutive arc-shaped gestures rotating the right hand in the air keeping thumb and pointing finger at a constant distance and moving the hand from left to right.*  *Fig. 10.23*
3	**G**	the inductive step	*He repeats the previous gesture of line 2 a second time.*

4	**G**	guarantees that it is always true	*He moves rapidly the right hand starting from his leg to an up-right direction in the air.*  *Fig. 10.24*

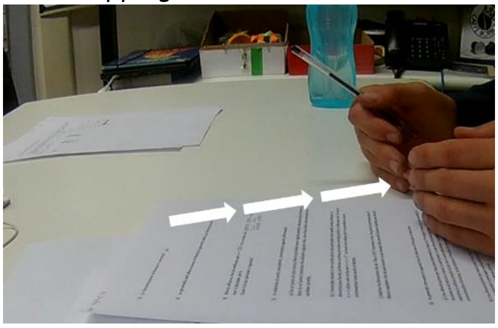
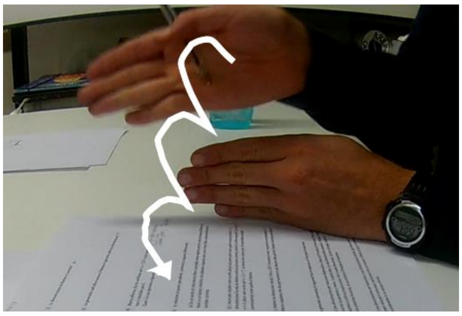
Guido's speech alone does not seem to provide a justification for the validity of MI. He only says that after the case base is proved true (line 1), the inductive step assures the truth of the proposition for every natural number (lines 3-4). However, if we look at his gestures, we notice that his discourse is enriched by other semantic elements. The base case is represented by a point on Guido's leg, which he touches with his right hand (fig. 10.22). Before saying "the inductive step", Guido makes a quite complex gesture: he moves his right hand in the air from left to right and simultaneously he rotates it twice, keeping thumb and pointing finger at a constant distance, forming two arcs (fig. 10.23). He then repeats the same gesture a second time when saying "the inductive step". These gestures are interesting because with them Guido seems to represent the inductive step itself as a series of arcs in the air, where each arcs indicate, metaphorically, the link between two cases of the proposition to be proved by induction (the implication $P(n)\rightarrow P(n+1)$). The gesture in line 2 and its repetition in line 3 are two examples of IS. Together with the gesture previously made in line 1, in fact, Guido is referring to an iteration that starts from the base of the induction. Finally, in line 4, while saying "it is always true", Guido performs a new gesture which is a repetition of the previous ones, but now more rapid and contracted: he starts touching his leg (as when he was referring to the base case), and then he moves fast to an up-right direction in the air. This time the hand moves in a straight line without shaping any arcs (fig. 10.24). This is again an IS but now obtained as a contraction of a previous ones. With this metaphorical gesture Guido seems to describe the whole iteration itself as a continuous and smooth path (there is not a reference the steps composing the iteration anymore) through which it is possible to conclude the truth of P(n) for all the natural numbers greater that the base.

Focusing on semiotic contractions it is possible to observe an interesting phenomenon which is the contraction of an IS into a new LS. As shown in the next excerpt, an IS, representing an iteration can be contracted and transformed into a LS which links the start and the end of the iteration.

Excerpt 10.5

In the final part of the interview, Lorenzo is explaining the functioning of a proof by MI, in particular the role of the inductive step.

	Who	Speech	Gesture and inscription
1	**L**	The structure is… you have the first step which is true.	*He puts the right hand open perpendicularly to the table and touches the top part of the sheet.*

			 *Fig. 10.25*
2	**L**	And every time that you use it…	*He moves the right hand down on the sheet stopping the movement twice.*  *Fig. 10.26*
3	**L**	The inductive hypothesis	*He puts the right hand perpendicularly on the sheet.*  *Fig. 10.27*
4	**L**	you know that it is true for the chain of previous steps.	*Keeping the left hand still on the paper, he moves the right hand to the left, starting from the position of the left hand, shaping some arcs in the air.*  *Fig. 10.28*

At this point the interviewer asks to Lorenzo to explain again what he was saying. Lorenzo describes again how, using the inductive step, the truth of the base case can be transferred to the subsequent cases. He speaks about a “chain of implications”. In doing it he firstly draws the inscription reported in fig.10.29.a which then he completes adding some numbers as in fig. 10.29.b.

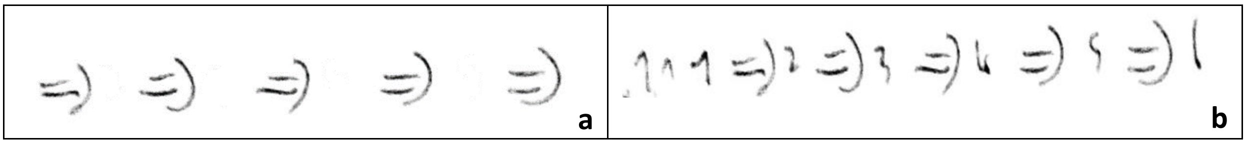
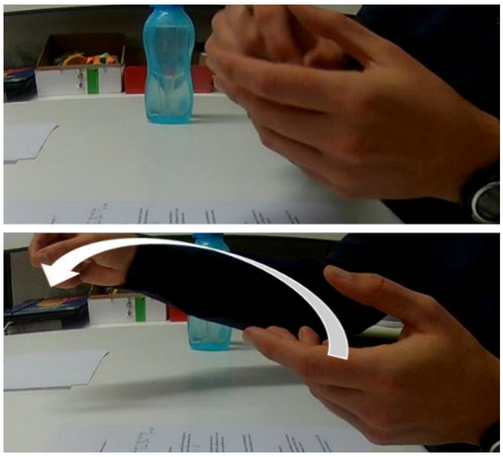


Figure 10.29. *Lorenzo’s inscriptions.*

Since now, Lorenzo has represented the structure of MI with three different but similar IS in which the iteration is represented as a series of steps. The first one, is the gesture shown in fig. 10.25 and 10.26. Here, the base case is a point on the sheet starting from which Lorenzo moves the right hand down, stopping the movement twice as representing three steps of the iteration. The second LS is gesture shown in the fig. 10.27 and 10.28. It develops in a similar way of the previous one: the right hand starts from a position on the sheet, indicated by the position of the left hand, which now represents a generic case P(n), and moves to the left, shaping some arcs in the air, which represent the steps of an iteration. In this case, differently than before, the iteration is not from the base case to greater numbers, but from P(n) to smaller cases. He is saying that the case P(n) (which he calls the inductive hypothesis) is true for the previous steps of the iteration. The third IS is the inscription in fig. 10.29.a, where the iteration underlying MI is represented by Lorenzo with series of arrows drawn on the sheet (he says “a chain of implication”). The inscription is then enriched with the reference to some consecutive numbers, from 1 to 6, fig. 10.29.b. In this second inscription, we can notice the presence of the base case: before the chain of implications ‘1→2→3→4→5→6’, Lorenzo wrote ‘1 ∧’.

The interview continues and Lorenzo produces a new sign to describe the structure of MI. Lorenzo is referring now to the proof by induction made by Marco, an invented student subject of the task ‘Meta questions on MI’:

[…]			
5	**L**	Matteo has shown that the base is true	*He puts the left hand perpendicularly on the sheet, similarly to the gesture of line 1.*
6	**L**	and with the other one	*He points to the sheet where Marco’s proof of the inductive step is contained.*
7	**L**	he uses it to show that for every…	*He touches the palm of his left hand with the fingers of his right hand, then, keeping the left hand still, he moves the right hand to his right stopping to a point in the air. In the movement, the gesture shapes an arc in the air.*

			 *Fig. 10.30*
8	**L**	for every number it is true.	*He repeats the same gesture of the previous line, but now reaching a different point in the space.*

Lorenzo is describing another time the functioning of MI. Now, however, instead of representing it as an iteration composed by several steps, he makes an arc gesture in the air linking two points, one on the palm of his left hand and the other in the air on his right side (fig. 10.30). The two points represent, metaphorically, the base case (the first one) and a generic number greater that the base. In line 8, he repeats the same gesture, but now reaching a different point in the air, as showing a second still generic number different from the previous one. These two gestures are a development and a contraction of the previous IS. They preserve some features of the previous gestural IS (the left hand perpendicular to the sheet to represent the base case, and the movement of the right to represent the direction of the iteration) and of the IS corresponding to the inscription in fig.10.29.a-b (the right hand moves in the same direction of the arrows in the inscription). However, the new gestures in line 8 and 9 do not represent the steps of the iteration anymore (they are smooth and continuous), but simply a link between the base case and a generic number. In other terms, with these two new gestures Lorenzo seems to represent the idea that, using the inductive step it is possible to connect the truth of $P(n_0)$ with the truth of $P(n^*)$ for a generic $n^* \geq n_0$. In summary, the IS of the first part of the excerpt, have developed and contracted into a new LS. Following the definition of IS, they can be obtained as a repetition of LS; this example has shown that an opposite phenomenon can also occur: a LS can be obtained as a contraction of an IS.

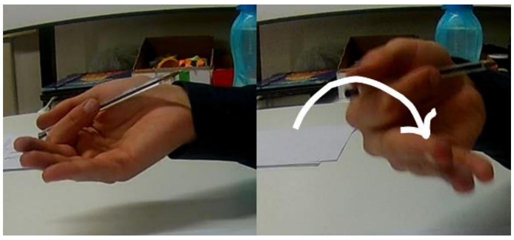
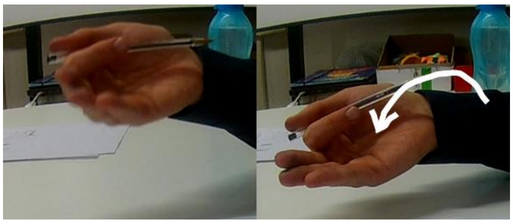
In addition to what just said, this last example also shows another characteristic that a LS might have. As observed when the definition of LS was given, the entities that such a sign links are seen in connection with natural numbers which do not have to be necessarily consecutive. The LS produced by Lorenzo does not link a natural number with its consecutive one, but a natural number (corresponding to the base case) with a generic one greater than the base. A similar situation happens in the next example.

10.1.3.4 A 'strong induction' linking sign

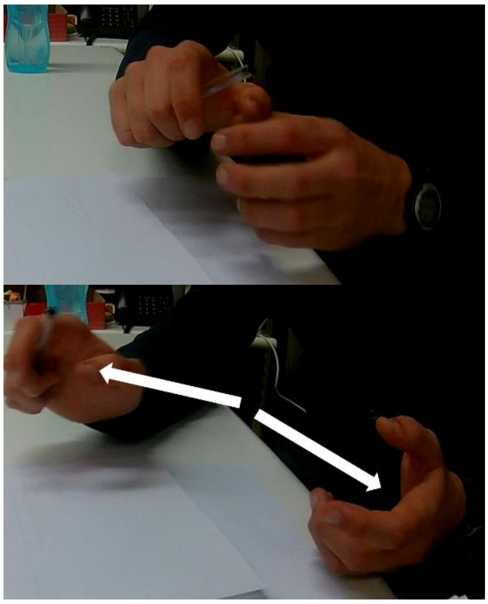
In the following excerpt we can see the production of a LS when referring to the inductive step of a proof by 'strong' induction[89]. In this case, the inductive step consists in proving P(n+1), assuming the truth of P(k) for every $n_0 \le k \le n$. To prove this, generally, one tries to connect the proposition P(n+1) with a P(k) for some $n_0 \le k \le n$ which is true for the assumption of the inductive hypothesis.

Excerpt 10.6

Lorenzo is dealing with the 'A divisibility problem' which involves a divisibility relationship between some expressions to be proved for every prime number *n*. As described in the a-priori analysis in section 8.4.3, the problem was given to investigate whether the subjects could recognise that for this problem, using MI was not a successful strategy. Lorenzo, indeed, is trying to prove the proposition writing the expressions in the field $\mathbb{Z}_n$. When the interviewer asks if MI could be a possible proving strategy, Lorenzo says that it would be very complicated, but in case it would be a proof by strong induction. Lorenzo describes how, in this hypothetical case, the inductive step would be:

	Who	Speech	Gesture and inscription
1	**L**	When I use the inductive hypothesis, I would have	
2	**L**	that for every prime number lower than n, it holds	*Starting with the right hand open with the palm facing upwards, he makes an arc-shaped gesture in the air from the right to the left.*  *Fig. 10.31*
3	**L**	and I would need to prove that it holds for n…	*He makes a similar gesture than the previous one but now moving the right hand from the left to the right.*  *Fig. 10.32*
4	**L**	But the bad thing of primes is that you cannot factorise them to exploit this fact.	*Starting with the two hands together, he moves the right one to the right and the*

[89] See section 1.2.

			left one to the left, as separating an object in two parts.  *Fig. 10.33*

In this short episode we can see the presence of three LS. The first one is the bundle produced in line 2 and used by Lorenzo to represent the connection between the generic prime number n with "every prime number lower than n". In line 3 a second LS is produced. This time it seems to represent the inductive step itself, as proving the proposition for the prime number n starting from the fact that it holds for every prime lower that n. In this case the direction of the link is the opposite of the previous one, as the direction of the gesture confirms. The third LS is the one produced in line 4. Here, while saying "you cannot factorise them", Lorenzo makes the gesture of separating an object hold in his hands in two parts (fig.10.33). This is a metaphorical gesture, representing the (impossible) factorisation of a prime number in two smaller prime numbers. Moreover, it is also a LS connecting a prime number with two other prime numbers lower than the latter. The whole excerpt, with these three different LS, even if representing a situation that Lorenzo says to be not possible for this specific problem, shows us how Lorenzo would structure and organise the proof of an inductive step in proof by strong induction: to connect the generic case P(n) with one (or more) P(k) for $n_0 \leq k \leq n$ (as represented with the LS of line 2 and 4) and then to deduce the truth of P(n) using the fact that P(k) is true (as represented with the LS of line 3).

10.1.3.5 A 'zoom' on a linking sign

The next excerpt shows how a subject might expand a LS enriching it with new aspects not considered yet. In a certain sense the new LS is produced as a 'zoom' on a previous one.

Excerpt 10.7

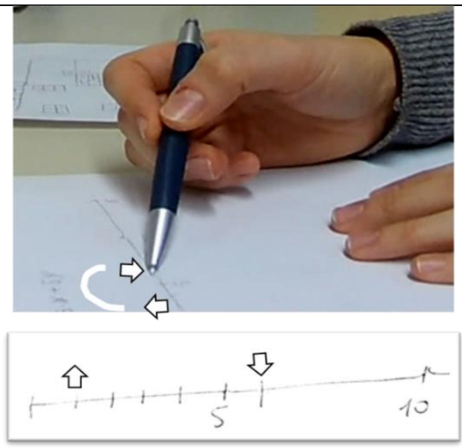
Giuditta is exploring the banknotes problem. In the first ten minutes she has produced several calculations without reaching any conjectures. Then, she notices that "every three fives there are five threes". When saying this she represents a numbered line divided in some intervals, drawing the inscription of figure 10.34.

Figure 10.34. *Giuditta's inscription when saying "every three fives there are five threes"*

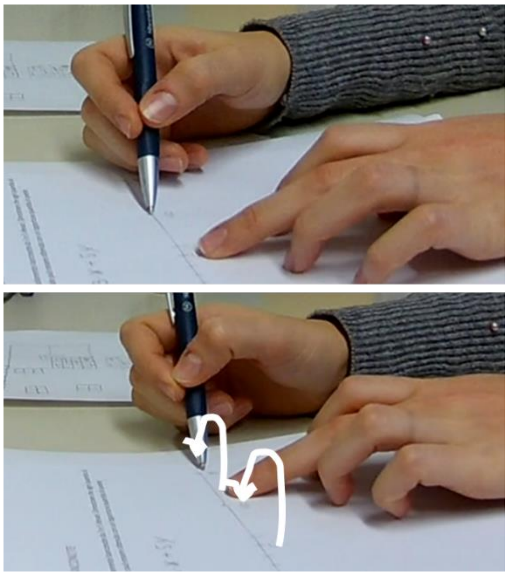
Giuditta seems to be blocked, thus the interviewer asks Giuditta to consider a slightly different problem in which only banknotes valuing 3 or 5 can be used to pay, but now one can have a change back (still with banknotes valuing 3 or 5). Giuditta observes that in this case one can obtain the value of 1, giving 10 and receiving 9 back. Note that from this one could easily conclude that every value can be obtained, since 1 was obtained and every natural number is a positive multiple of 1. However, Giuditta follows a different strategy. Firstly, she observes that 2 can be obtained as well, giving 5 and receiving 3 back. Then, Giuditta, referring to the inscription of fig. 10.34, starts constructing an argumentation of the fact that every value can be obtained:

	Who	Speech	Gesture and inscription
1	**G**	So… if I can do it in an interval of five,	*With the left middle finger and the pen in the right hand, she points to the two extremes of the 0-5 interval of the line of fig. 10.34*  *Fig. 10.35*
2	**G**	then I can do it…	*Keeping the left middle finger still on the left extreme of the line, she moves the pen in an arc-shaped trajectory stopping where '10' is written. Starting from there she makes the same gesture a second time, now stopping where '15' is written.*

			 Fig. 10.36
3	**I**	Why this?	
4	**G**	Because I can divide in small intervals of five…	*She puts the two hands parallel to each other and perpendicular to the sheet, as representing the extremes of an interval, then she moves both hands keeping them at a fixed distance to the right. Each hand shapes two arcs in the air.*  Fig. 10.37
5	**G**	Because if I can do it for one, two, three, four and five	*While counting, she points at the positions of the numbers 1, 2, 3, 4 and 5 respectively on the drawing of fig. 10.34*
6	**G**	Then I can do it also…	*Staring on the point corresponding to 1 on the interval 0-5, she moves the pen on the interval 5-10 of the drawing and she marks with a line a point inside this interval, corresponding to the position of 6.*

			 Fig. 10.38
7	**I**	Yes, in the other interval… why?	
8	**G**	Because they are in the same residue class.	

In the successive two minutes Giuditta finds a way to combine the 3 and 5 banknotes (with changes) to obtain the values 1, 2, 3, 4, and 5. Then she concludes that it is possible to obtain every value:

[…]			
9	**I**	We can do it for all of them [1, 2, 3, 4, 5], so, for what we said…	
10	**G**	We can shift this, in this way.	*With the left middle finger and the pen in the right hand, she points to the two extremes of the 5-10 interval. Then she moves the left middle finger and the pen in the right hand on the extremes of the 10-15 interval, each hand shaping one arc in the air.*  *Fig. 10.39*

The presented excerpt contains, firstly an IS, composed by two LS, produced by Giuditta in lines 1-2 when saying "If I can do it in an interval of five, then I can do it". With this sign Giuditta seems to link the first quintuplet of numbers (1-5) with the second one (6-10) and then with the third one (11-15). Note that now the link involves the whole intervals, as highlighted by Giuditta's pointing at their extremes. Then, in line 4 she produces a similar sign, still an IS composed by two LS (fig. 10.37). With these signs Giuditta represents the division of the line in intervals, but also the connection between them, showing that each interval starts where the previous one ends. After, Giuditta explains why, once having proved the conjecture for 1, 2, 3, 4 and 5, one can conclude that it is true also for the other numbers (lines 5-6). To do so, she produces an interesting sign: when saying "then I can do it also…" she makes a gesture linking a point of the first interval (corresponding to the number 1) with a point on the second interval (corresponding to the number 6) which she marks with a line. Looking at the bundle of gesture and inscription, we can see that Giuditta is again linking two intervals as she did before, but now this link is enriched by a new component: the number 1 is linked to the number 6. With this new LS, which is an evolution of the previous ones, Giuditta seems to represent the fact that adding five (i.e. a banknote valuing 5) one can link 1, 2, 3, 4 and 5 with respectively 6, 7, 8, 9 and 10. The previous LS represented the link between two consecutive intervals of five numbers, this new LS however represents the connection between two numbers belonging to these two intervals. This LS, therefore, allows Giuditta to represent with more details the first (but also a generic) step of the iteration that she has rapidly described with the IS of before: not only the intervals of five numbers are iteratively linked, but also every number of an interval are linked with the respective numbers of the successive interval. This point is further clarified in line 8 by Giuditta who, answering to the question of the interviewer, says that is possible "because they are in the same residue class". With this she is referring to the fact that 1 and 6 (which seems to be used as a generic example also for other couples of numbers of other intervals) are in the same congruence class modulo 5.

After this, Giuditta finds a way to combine the 3 and 5 banknotes (with changes) to obtain the values 1, 2, 3, 4, and 5. Finally (line 10) she concludes that "shifting this" (i.e. the interval of five numbers) one can obtain every number starting from the first five numbers. In this last line, Giuditta represents again the iteration using a sign (fig. 10.39) similar to the signs produced at the beginning of the interview. It is interesting to notice that while talking about the interval 0-5 she points at the extremes of the interval 5-10. This is another example of speech-gesture mismatch which allows Giuditta to refer simultaneously to two intervals. With this Giuditta seems to anticipate with her gesture the result of the "shifting" of the first interval on the second one. Then she represents a successive step of the iteration (from the second to the third interval). The whole sign, therefore, considered as a bundle of speech, gesture, and inscription is an IS representing the iterative solution of the problem constructed by Giuditta.

10.1.4 Linking and iteration signs as resources

As said at the beginning of this chapter, analysing the interviews it was possible to observe that the production and the use of LS seems to be crucial in the generation of a recursive argumentation, in its construction and also in the process of constructing a proof by induction. In the first analysed excerpt (10.1), we noticed how Giuditta's production of LS and IS, from one side enabled us to see a recursive argumentation emerging in her discourse, but from the other side allowed Giuditta to better explore the problem itself and to recognise a possible solving strategy. Giuditta, in fact, explores and develops the LS she has produced and from this exploration she recognises that the tessellation of a chessboard can be used to tessellate the subsequent one and that this strategy can be iterated. Similarly, in the excerpt 10.3, I showed how Lorenzo drew a mathematical tree (fig. 10.21) in order to describe the iterative solution of the problem. As observed, however, the drawing of the tree has not just a descriptive role, but it is used and explored by Lorenzo to provide a solution for the problem. He noticed, in fact, the tree's height corresponds to the number of weightings necessary to find the false coin. By the investigation of the structure of this tree, when N varies, Lorenzo reaches a solution for the problem.

These two examples, thus, showed that LS and IS are not just signs to be registered by an observer to have a "'window' onto [students'] thinking" (McNeil & Duncan, 2000, p.143), but they are also resources for the students themselves, supporting them in organising and structuring a recursive argumentation. As we have observed in the two examples of above, by the exploration of some of these signs a student could develop some aspects of the problem's solution (such as the inductive step or the step of an iterative solution). In other terms some of these signs could become *diagrams* in Peircean sense (see Dörfler, 2016, for a focus on this point) on which the subjects' reasonings can develop. Following Dörfler (2016), diagrams can be understood as signs on which actions can be performed, moreover "diagrams are not to be understood in a figurative but rather in a relational sense (such as a circle expressing the relation of its peripheral points to the midpoint)" (Dörfler, 2016, p. 25). A diagram can also suggest new modes of action, which cannot be deduced directly from its properties, but only from the relational aspects highlighted. A subject's activity on a diagram can also produce new knowledge. In tune with this, some of the LS/IS analysed above, thus, as the mathematical tree drawn by Lorenzo or the drawing of the chessboards one nested into the other by Giuditta, become diagram in the moment in which they represent the relationships between some cases of the problem. In this perspective the activity of a subject on these signs through which a recursive argumentation or a proof by MI is developed can be interpreted as a form of *diagrammatic thinking* (Dörfler, 2016).

In the following section I will present other paradigmatic examples of the just described aspect. The excerpts are chosen in order to provide examples of different situations involving linking and iteration signs as resources for the subjects.

10.1.4.1 Linking and iteration signs to organise a recursive argumentation

In the following example, we will see how a student's production and use of some LS and IS seem to support the organisation and the construction of a recursive argumentation for the problem.

Excerpt 10.8

Silvio is dealing with the chessboard problem. In the minutes that anticipate this excerpt, he has explored some cases, finding that for the chessboards corresponding to n=0, n=1, and n=2, it is possible to create a tessellation which leaves out only one little square. At this point, on the sheet there are several drawings. One of these represents a 4x4 chessboard and it has just been used by Silvio to describe a tessellation which leaves out only one little square (fig. 10.40). The tessellation has been described drawing only one tile and indicating with gestures where to put the other tiles.

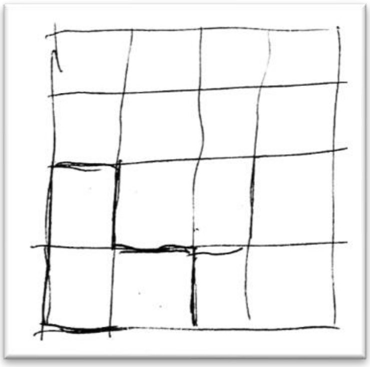
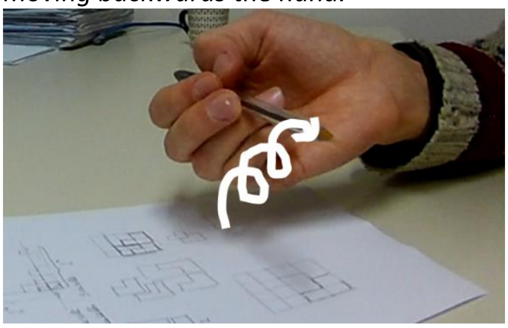
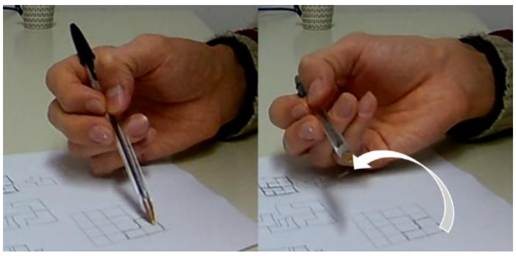
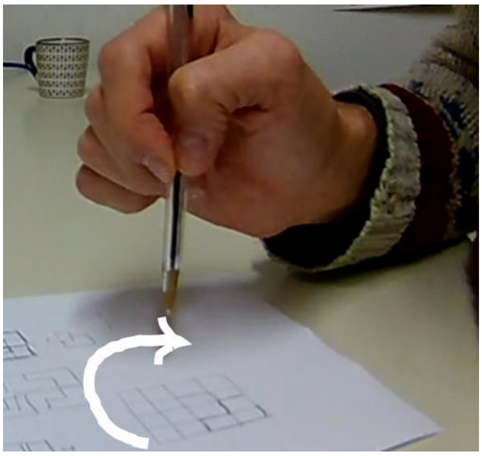
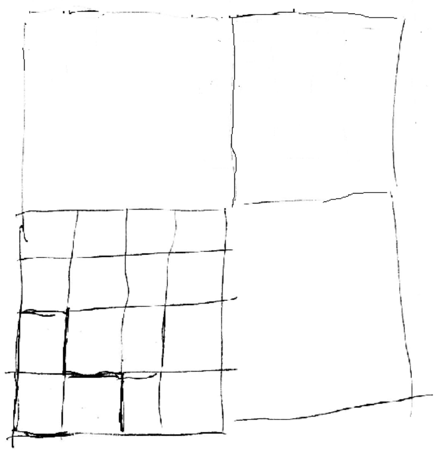
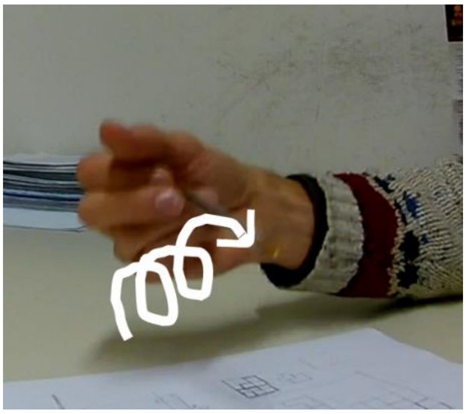


Figure 10.40. *Silvio's inscription representing a 4x4 chessboard. For this case he found a tessellation which covers all the squares except of one.*

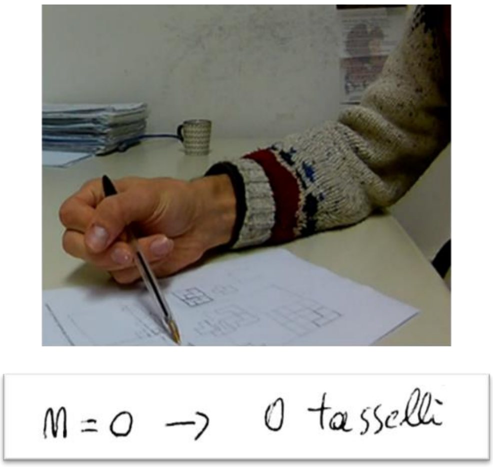
The interviewer asks Silvio if it is possible to generalise the result to other chessboards.

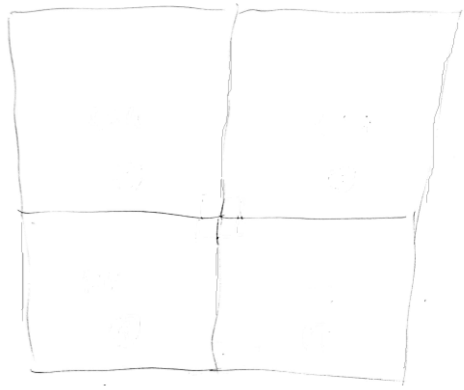
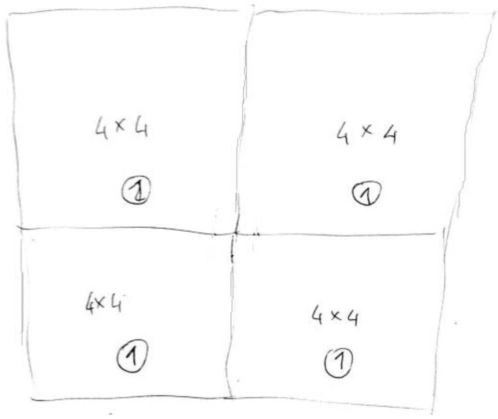
	Who	Speech	Gesture and inscription
1	**S**	How could I do? … I mean … Ah, yes.	
2	**S**	Once you have established… I mean, every case is reconducted to the previous one, right?	*Keeping the pen in his right hand, he rapidly draws some circles in the air, while moving backwards the hand.*  *Fig. 10.41*
3	**S**	In the sense that, if I go to the next case, so two to the third,	*Touching with the pen in his right hand the inscription of fig. 40, he draws an arc in the air moving the hand forwards.*  *Fig. 10.42*
4	**S**	Basically, this square here is created again three times.	*He points to the inscription of the 4x4 chessboard and then he draws an arc in the air over the sheet around the*

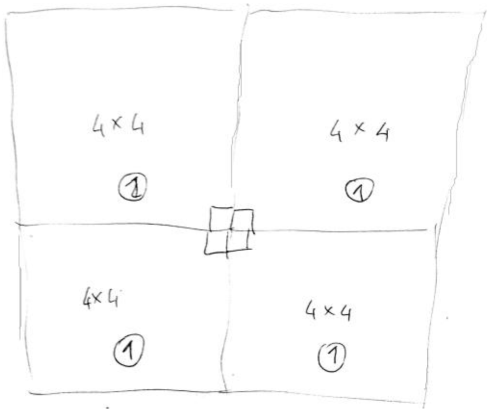
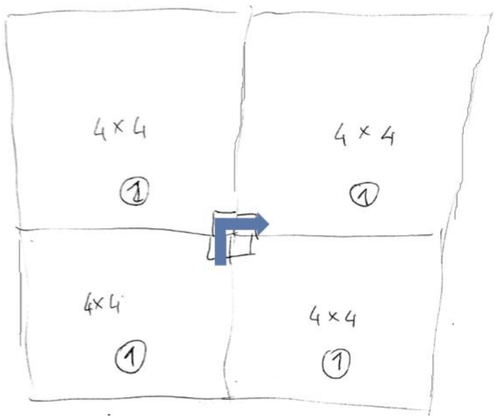
			inscription from the top left corner to the right side of the square.  *Fig.10. 43*
5	**I**	Other three times.	
6	**S**	Exactly, in this sense…I have another square here, another square here and another here.	*He extends the inscription of fig. 40 drawing the sides of three squares constructed respectively from the top side of square representing the 4x4 chessboard, from the top right vertex and from the right side.*  Fig.10.44
7	**I**	You can turn the sheet if you prefer.	
8	**S**	Yes, better.	*He turns the sheet on the other face which is blank.*
9	**S**	So, in practice, the case two to the third.	*He draws a square on the sheet.*

10	**S**	So, we agree that all of them are reconducted to a previous case.	*He makes a gesture similar to the one of line 2, but this time higher in the air.*  *Fig. 10.44*
11	**S**	To the immediately previous case, in practice.	
12	**I**	Ok.	
13	**S**	I could do something like… but it is everything a bit chitchatted[90], I don't know if I could formalise it in this moment.	*He turns again the sheet on the first face where the inscriptions of the first part of the interview were.*
14	**I**	It's all right.	
15	**S**	I mean… I think that one could do something like by induction.	*He rapidly draws again some circles in the air as in line 2.*  *Fig. 10.45*

[90] He says: "Chiacchierato", in Italian.

16	**S**	Because, since in the case zero I have only one little square, and it remains out,	*With the pen he touches the sheet where he wrote* 'n=0 → 0 tiles' *during the first part of the interview.*  Fig. 10.46
17	**S**	then, let's say … in a sequential way, this little square will always remain out.	*With the pointing finger of his right hand, he draws several circles in the air whilst moving up his hand. The gesture is wide in the air and it lasts for almost 3 seconds.*  *Fig. 10.47*
18	**S**	Yes… but it must be like this. Surely.	*He turns the sheet on the side were, in line 9, he drew a square.*
19	**S**	In the best option only one square remains uncovered.	*He extends the drawing of the square adding three other squares attached to it, obtaining the following inscription.*

			 *Fig. 10.48*
20	**S**	Because if I… let's say that this is the four by four, this is the four by four… and this is the four by four.	*He writes '4x4' in the centre of the four squares in the inscription of the previous line.*
21	**S**	If here it remains one out, and here one, and here one, and here one.	*He writes '1' and circles it, in each of the four squares, obtaining the following inscription.*  *Fig. 10.49*
22	**S**	In the best option, I have the empty one here, the empty one here, the empty one here… and the other here.	*He draws four little squares, one for each of the previous squares, in correspondence to the centre whole inscription. He obtains the following drawing.*

			 *Fig. 10.50*
23	**S**	And therefore, in any case, only one remains out.	*He rapidly draws, without marking the sheet, an 'L' over three little central squares of the inscription of line 22.*  *Fig. 10.51. The arrow indicates the trajectory of the pen over the inscription.*
24	**I**	Because you put another tile there?	*He points at the centre of the inscription.*
25	**S**	Exactly.	

In the remaining part of the interview, Silvio tries, successfully, to find a tessellation of the 4x4 chessboard which leaves out one square in one of the corners, so to recreate the situation represented in fig.10.50.

<u>Analysis</u>

To better analyse Silvio's excerpt, we can divide it in three parts which mark the development of Silvio's argumentation (from line 1 to 12, from 13 to 17, and finally from 18 to 25).

In the first part (lines 1-12) Silvio introduces the idea that there is a connection between the cases of the problem. In line 2, when saying "every case is reconducted to the previous one" he produces an interesting gesture (fig. 10.41) which he repeats in line 10 when saying "all of them are reconducted to a previous case". These two gestures, together to his speech are IS. With these, Silvio is referring to the fact that there is a link between two consecutive chessboards and that this link can be iterated to involve every chessboard of the problem. It is interesting to notice that these gestures seem to represent the idea of “backward iteration”, they are metaphorical in this sense. The gesture in line 10 is detached from the sheet, high in the air, and it does not refer to any inscription on the sheet. In Krause’s classification it is a gesture on the level of generality (3rd level). With this gesture Silvio is describing the iteration itself and its logical structure without any reference to the chessboards. In lines 3-6, Silvio describes in detail how two consecutive chessboards are linked together. To do so, he produces three different LS.

- The first one occurs in line 3, when, whilst saying “to go to the next case, so two to the third”, he makes an arc shaped gesture in the air over the inscription of the 4x4 chessboard. With this LS, Silvio is referring to the ‘forward’ transition from the 4x4 chessboard to the 8x8 chessboard.
- Right after this, in line 4, he produces a new LS, which this time contains some references of how it is possible to link the 4x4 chessboard with the 8x8 chessboard: while saying “this square here is created again three times” he rapidly makes a circular gesture (fig. 10.43) on the inscription. With this he seems to refer to the action of using three other 4x4 chessboard to form, with the first one, an 8x8 chessboard.
- Finally, in line 6, he draws the inscription of an 8x8 chessboard composed by four 4x4 chessboards (fig. 10.44). This is a new LS which represents the situation he just described.

Analysing diachronically the evolution of these LS we can see how they convey different semantic aspects of the link between two chessboards: the first one involves the fact that passing from the 4x4 to the 8x8 chessboard corresponds to a step of an iteration (Silvio indeed says “to go to the next case”). The second one represents the action of constructing an 8x8 chessboard adding three new 4x4 chessboard to the previous one. Finally, the last LS represents a static situation in which an 8x8 chessboard is composed by four 4x4 chessboards. In the end, with these three LS, therefore Silvio has justified why he was saying that “every case is reconducted to the previous one”.

In all this first part of the excerpt, Silvio has represented the structure of a possible iterative solution of the problem, prompted by the observation that a $2^n x 2^n$ chessboard is composed by four $2^{n-1} x 2^{n-1}$ chessboards (shown by Silvio for the case n=3). It is interesting to notice that, until now, Silvio has not talked about tessellation at all. He indeed, has not found yet a way to tessellate the 8x8 chessboard using the tessellation of the 4x4 chessboard. This is a very interesting aspect since, in a certain sense, Silvio is anticipating how a recursive argumentation could develop without knowing the details yet. The LS and IS he produced, thus, seem to prepare Silvio’s argumentation providing a logical structure for it. They are crucial signs, in this sense, for the generation of a recursive argumentation: in this first part, Silvio has organised and planned the recursive argumentation which he will develop later.

In the second part of the excerpt (13-16), Silvio enriches his discourse with a new element: the tessellation of the chessboards. Firstly (line 15), he produces a new IS similar to the previous ones. It is interesting that in this point the gestural IS is co-timed with the words “something like by induction”. This shows that Silvio is indeed recognising a connection with mathematical induction. In the successive lines, he clarifies what he meant with “something like induction”. He says that the chessboard corresponding to n=0 is composed only by one little square which thus cannot be tiled (line 16). Then he says that from this, “in a sequential way”, it could be possible to show that the little square will remain out of the tessellation for all the bigger chessboards as well. In describing this strategy, Silvio is apparently using his previous observation of the fact that it is possible to construct a $2^n x 2^n$ chessboard with four $2^{n-1} x 2^{n-1}$ chessboards. Silvio’s speech contains elements which refer to the tessellation of the chessboard (“this little square will always remain out”), something that was not present in the first part of the excerpt. However, observing the whole semiotic bundle, we can see that he is also referring to the structure of the argumentation itself. He firstly touches the sheet where the inscription for the 1x1 chessboard is, which he calls the “case zero” (fig. 10.46). Then, while saying “in a sequential way”, he draws several circles in the air whilst moving up the hand (fig. 10.47). Note that in this gesture there is not any reference to chessboards or tessellations. This gesture, in fact, seems to repeat the “something like by induction” gesture previously made, but now it is longer (both in time and space). However, if we consider the whole movement of his right hand (fig. 10.46 and 10.47), we can see it starting from the sheet, concretely touching with the pen the inscriptions which refer to the chessboards, and then moving up, drawing a sort of helix in the air, as representing the iterative structure of the reasoning by induction itself. This is therefore an IS which starts as a 1st level gesture (deictic) in Krause’s classification and then it develops in the air as a third level gesture. With this IS Silvio therefore seems to connect the chessboards involved in the problem (pointed at the beginning of the gesture) with the logical structure of the iteration that he has described in the first part of the excerpt, represented by the helix in the air.

In order to conclude the argumentation, Silvio still needs to find a way to use a tessellation of a chessboard which leaves out only one square, to obtain a tessellation of the subsequent chessboard which leave out only one square (namely, a sort of the inductive step). This is what Silvio tries to obtain in the next lines.

In the third and final part of the excerpt (lines 18-25), Silvio focuses on the connection between the 4x4 and the 8x8 chessboard, in particular on finding a way to tessellate the 8x8 chessboard using the tessellation of the 4x4 chessboard. Firstly (line 19) he extends the drawing of a square he previously made, obtaining a big square composed by four squares (fig. 10.48). This, as Silvio says in line 20, represents four 4x4 chessboards forming an 8x8 chessboard. This is a LS which reproduces the inscription of line 6 (fig. 10.44). In the successive two lines, he enriches the inscription with new elements to represent firstly the fact that every 4x4 chessboard is tessellate except for one little square (fig. 10.49) and then the fact each of these little squares “in the best case” are in the centre of the 8x8 chessboard (fig. 10.50). This is an evolution of the previous LS: not only the two chessboards are linked, but their tessellations are also linked now. Finally, in line 23, the LS evolves again: Silvio draws rapidly, without marking the sheet, an ‘L’ over three little central squares of the just obtained inscription (fig. 10.51). With this gesture, as confirmed by his answer to the interviewer’s question (lines 24-25), he mimes the action of putting a L-shaped tile in the centre of the 8x8 chessboard to cover three of the four blank squares. The LS has now fully developed. Silvio has finally represented with a generic example how to use tessellation of a chessboard which leaves out only one square, to obtain a tessellation of the subsequent chessboard which leave out only one square.

Finally, the iterative solution of the problem which he anticipated in the first two parts of the excerpt is supported by the details of how a single step can be obtained. Note that, following Silvio argumentation, the empty little square should be in a corner in every chessboard so that the subsequent chessboard can be tessellate adding an L tile in the centre. However, in the 8x8 chessboard obtained by Silvio this is not happening (the empty square is indeed in the centre), therefore, the step of the iteration is not complete. This will be noticed and corrected by Silvio at the end of the interview, after the intervention of the interviewer.

The whole excerpt shows how the production and use of LS and IS support Silvio’s development of a recursive argumentation. In the first two parts, Silvio has not found a way to tessellate a chessboard using the tessellation of a previous one yet, however he describes how the problem could be solved “doing something like by induction”. In these parts Silvio seems to plan and organise a possible iterative solution of the problem. In a certain sense this happens ‘a-priori’ since a step of this iteration has not been constructed yet. This happens only later, in the third part of the excerpt, when Silvio successfully obtains a tessellation of an 8x8 chessboard using the tessellations of four 4x4 chessboards. This is accomplished by Silvio by exploring, modifying, and enriching some previous LS, in particular the inscriptions he already drawn. The whole excerpt, in conclusion, highlighted how some LS and IS can be semiotic resources for the subject who has produced them: supporting both the planning and the organisation of a subsequent argumentation, and the construction of the argumentation itself.

10.1.4.2 The exploration of a linking sign to construct the inductive step

In the last excerpt we saw an example of how the production and use of LS and IS might support a student in generating and structuring a recursive argumentation. In the following excerpt, instead, we will see how the production and the successive investigation of a LS supported a student involved in proving the inductive step during the construction of a proof by MI. This shows that these kinds of signs can be important not only during the phase of exploration of a problem or of production of a conjecture, but also in the successive phase where the argumentation is formalised in a mathematical proof (in our case by induction).

Excerpt 10.9

Claudio, a second-year undergraduate student in physics, is dealing with the problem of determine the number of diagonals of a (convex) polygon with N vertices. During the first twenty minutes he analysed the problem for the cases corresponding to N=4, 5 and 6. In this last case, he noticed that it is possible to count all of its diagonal in the following way: we start from one vertex, for which we have three different diagonals; then we move clockwise to the successive vertex, for which we have again three different diagonals; then if we keep moving clockwise to successive vertices and we count only the diagonal that we have not counted yet we obtain, respectively 2, 1, 0 and 0 new diagonals. Claudio drew the inscription reported in fig. 10.52.

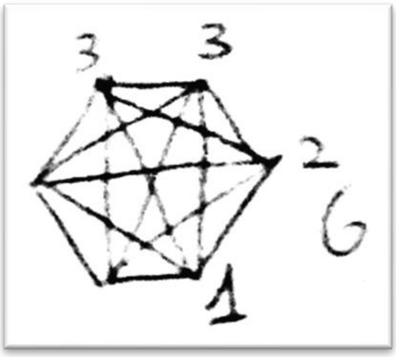


Figure 10.52. *Claudio's inscription representing a hexagon with its diagonals. The numbers on the vertices the number of not already counted diagonals starting from the considered vertex. The number '6' on the right indicates that the drawing refers to the case N=6 of the problem.*

Then Claudio generalised the solution for a polygon with N vertices. After a few attempts he wrote the following expression for the number of diagonals (which he indicates with 'nd') for a polygon with n vertices:

$$\boldsymbol{nd = 2(n-3) + \sum_{k=1}^{n-3}(n-3-k)}$$

The original inscription made by Claudio is reported in figure 10.53.

n numero di vertici

$$nd = 2(n-3) + \sum_{k=1}^{n\,3} n - 3 - k$$

Figure 10.53. *Claudio's inscription containing the formula, expressed as a sum, for the number of diagonals of a polygon of n vertices. In the first line there is written 'n number of vertices'.*

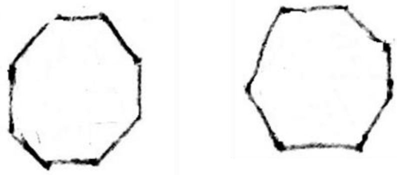
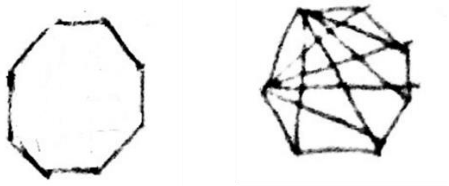
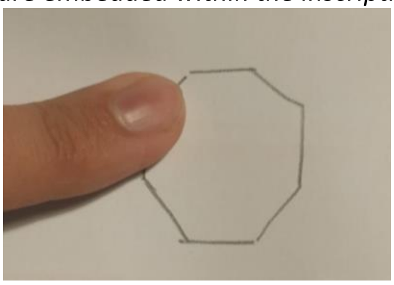
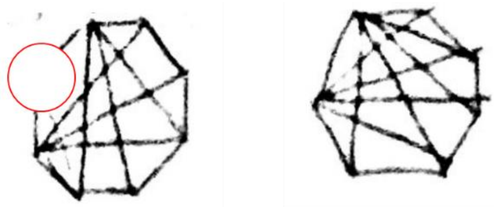
After this, at the request of the interviewer to prove what conjectured, Claudio starts the construction of a proof by MI for the fact that the just written expression corresponds to the number of diagonals of a polygon with n vertices. The choice of proving the formula by MI was not suggested by the interviewer, but unfortunately Claudio does not make explicit the reason of this choice, he only says that "it is the first thing that comes into mind". Claudio rapidly proves the base case, noticing that the expression becomes zero when n=3 which is coherent with the fact that a triangle has zero diagonals. Then he starts proving the inductive step. He writes the expression of above substituting n with n-1, obtaining the expression of fig. 10.54, which is the inductive hypothesis, and he claims that he needs to show that, with this assumption, for a polygon with n vertices the expression of above holds.

$$nd_{n-1} = 2(n-1-3) + \sum_{k=1}^{n-4} n-4-k$$

Figure 10.54. *Inscription containing the expression that, for the inductive hypothesis, corresponds to the number of diagonals of a polygon with n-1 vertices.*

At this point, however, Claudio starts having some difficulties in writing an expression representing the number of diagonals of a polygon with n vertices (for which he said "I add a vertex") using the expression 'nd_{n-1}' of fig. 10.53. The interviewer, thus, asks Claudio to make a drawing to represent this situation.

	Who	Speech	Gesture and inscription
1	**C**	Let's do this…	*He draws a polygon with 8 vertices.*
2	**C**	I take also the one with seven… one, two, three, four, five, six… it is very ugly.	*He draws a polygon with 7 vertices on the right of the previous drawing. On the sheet now there are the two inscriptions of fig.10.55, corresponding to an 8 sides polygon (on the left), and a 7 sides polygon (on the right).*

			 *Fig. 10.55*
3	**I**	No worries, they do not have to be regular.	
4	**G**	Ok, ok. So, I said… between these… this one will have a number of diagonals…	*He draws 8 diagonals starting from two consecutive vertices of the 7 sides polygon.*  *Fig. 10.56*
5	**G**	Which was given by this formula here.	*With the top of the pen, without marking, he circles the expression for ‘nd_{n-1}’ in the inscription of fig.10.54.*
6	**G**	Then, what did I do?	
7	**G**	In practise, without considering the last vertex…	*He puts the left pointing finger to cover a vertex of drawing of the 8 sides polygon. The following picture recreates the gesture embedded within the inscription.*  *Fig. 10.57*
8	**G**	I will go to count all the diagonals that there are between all the angles except for that one.	*Keeping the left pointing finger still on a vertex of the 8 sides polygon, he draws 6 diagonals, respectively 3 for each of the two consecutive vertices of the one which is covered.*  *Fig. 10.58. Claudio’s drawing of six diagonals of the 8 sides polygon on the left, together*

			with the already made inscription on the right. The red circle indicates the part covered by Claudio's finger whilst he is making the drawing.
9	**I**	Yes, which is the case n minus 1, right?	
10	**C**	Exactly! And I counted them all.	
11	**I**	Then you said: I add a vertex.	
12	**C**	Yes, I add a vertex.	*With the pen he points to the vertex of the 8 sides polygon that he was previously covering with his finger.*  *Fig. 10.59*
13	**C**	And this vertex… I can link it with n minus three… with other n minus three angles.	*He rapidly draws, without marking, some lines that start from the vertex of the 8 sides polygon he was pointing to go to the direction of other vertices.*
14	**I**	Ok	
15	**C**	Ok, good. So, how many are these?	*He draws a circle in the air over the inscription of the 8 sides polygon with the diagonals.*
16	**C**	They are given from this formula here.	*He draws a circle in the air over the inscription containing the expression for* nd_{n-1} *of fig.10.53*
17	**C**	Plus, these n minus three.	*He draws a little circle in the air over the vertex of the 8 sides polygon pointed at line 12.*

Claudio then writes on the sheet the expression for the number of diagonals of a polygon with n vertices as the expression nd_{n-1} plus n-3. More explicitly he writes:

$$\boldsymbol{nd = \left(2(n-1-3) + \sum_{k=1}^{n-4}(n-4-k) \right) + n-3}$$

Where the expression in the first parenthesis corresponds to the expression of nd_{n-1} that he previously wrote. This expression is coherent with Claudio's argumentation of before, however it is not completely correct in terms of the solution of the problem. In fact, this expression does not count another diagonal which Claudio has not considered yet: when adding a vertex to transform an 'n-1'-sides polygon into an n-sides polygon, the side connecting the two vertices which are consecutive to the new vertex becomes a diagonal in the new polygon[91]. The interviewer, thus, intervenes to point this aspect to Claudio who realises and rapidly corrects the error, finally writing the correct expression of nd_n as $(nd_{n-1}) + (n-2)$. The original inscription written by Claudio is reported in figure 10.60.

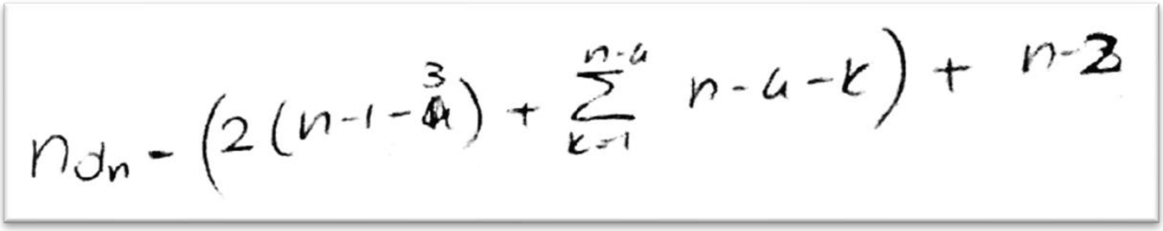


Figure 10.60. *Inscription containing the expression that corresponds to the number of diagonals of a polygon with n vertices, written in terms of the expression that, for the inductive hypothesis, corresponds to the number of diagonals of a polygon with n-1 vertices.*

Finally, Claudio, after some algebraic manipulations, correctly completes the proof of the inductive step showing the equivalence between the expression of fig. 10.60 and the one of fig. 10.53.

<u>Analysis</u>

Claudio's excerpt starts with his attempts in writing an expression for the number of diagonals of a polygon with n vertices using the inductive hypothesis. This is the first important part for proving the inductive step. For writing this expression Claudio, after the suggestion of the interviewer, produces some inscriptions. Firstly (lines 1-2), he draws two polygons one next to the other, respectively an 8-sides polygon and a 7-sides polygon (fig. 10.55). Previously, whilst trying to construct the inductive step, he was talking about the polygon 'n' and the polygon 'n-1' which now he represents, with a generic example, fixing n=8. Then in lines 4 and 5 he links the drawing of the 7 sides polygon with the inductive hypothesis, saying that this polygon has a number of diagonals given by the expression nd_{n-1}. He draws some diagonals for this polygon starting from two vertices. Until now the drawings of the two polygons are representing two separated cases of the problem. However, in the next lines, Claudio links the case of the 8 sides polygon with the previous one. To do so he produces a rather complex LS. First of all, he covers with his left pointing finger a vertex of the 8 sides polygon (fig. 10.57). This is not simply a deictic gesture. With this, Claudio is transforming the 8 sides polygon into a 7 sides polygon (as the one on the right). This gesture, co-timed with what he says ("Without considering the last vertex") is a LS: it represents a 7 sides polygon as embedded into an 8 sides polygon. It is worth it to remind that following Arzarello's definition (2006), a sign within the semiotic bundle is not simply a juxtaposition of components belonging to different semiotic set, but it is also formed by the mutual relationships between these components. In the case of the just analysed sign produced by Claudio this aspect is quite evident: the relation between the gesture and the inscription provides semantic elements to the whole sign.

Claudio then 'applies' the inductive hypothesis: he says that without considering a vertex the 8 sides polygon has the same number of diagonals of the 7 sides polygon. He draws some diagonals on the drawing, still keeping the finger to cover a vertex (fig. 10.58). In other terms he has recreated the drawing for the polygon with 7 sides together with the diagonals inside the drawing of the 8 sides polygon. This is another LS which is an evolution of the previous one since now the link involves the diagonals of the polygons as well. Note how the whole semiotic bundle is necessary for Claudio to represent this situation:

---

[91] This passage was shown in the a-priori analysis of the task in the 'Methods' section.

the inscriptions on the sheet, his speech to clarify that the number of diagonals is the same (indeed he has not drawn the same number of diagonals) and his action of covering one vertex to transform the 8 side polygons into a copy of the 7 sides polygon. In the successive lines Claudio's discourse continues as follows: firstly, (line 12) when saying "I add a point" he points to the vertex of the 8 sides polygon that he was covering before; then (line 13) he says that from this vertex other n-3 new diagonals can be formed, which correspond to the number of all the vertices of the polygon except for the vertex itself and the two vertices consecutive to it. Finally, in lines 15-17, he summarises what he has done counting again the total number of the diagonals of the polygon with n sides (represented in the drawing with 8 sides). He says that this number is equal to the number of the diagonals of a polygon with n-1 sides (represented in the drawing with 7 sides) which is given by the inductive hypothesis ("They are given from this formula here") plus n-3 which represent the new diagonal which start from the eighth vertex. After this, Claudio is convinced of his reasoning, and he writes the algebraic expression which corresponds to what he just described. This point is extremely interesting because at the beginning of the excerpt Claudio was having some difficulties in writing an expression for nd_n using the expression for nd_{n-1}, namely a LS involving the algebraic expression. However, with the help of the drawings he managed to construct a LS to connect the number of diagonals of a n-1 sides polygon with the number of diagonals of a n sides polygon and to express this relationship (and this LS) also in terms of an algebraic expression. The expression he writes is coherent with his explanation with the drawings, but it is not completely correct in terms of the solution of the problem; indeed, one diagonal has not been considered. However, after the intervention of the interviewer the expression is rapidly corrected in the one reported in fig. 10.60 and, from this, Claudio correctly concludes the proof of the inductive step.

In summary, the whole excerpt showed how the production and the exploration of some LS have helped and supported Claudio's construction of the proof of the inductive step. In particular, with the LS of lines 7 and 8 he found how to count the number of diagonals of a polygon with n sides *assuming* to know the number of diagonals of a polygon with n-1 sides. This relationship is then expressed by Claudio with an algebraic expression which, once corrected, finally allows him to conclude the proof of the inductive step.

10.1.4.3 Suggesting a linking sign as a possible 'unblocking' intervention

The last two examples that I will show refer to two particular situations in which the interviewer suggested what can be interpreted as a LS to help two students that seemed to be blocked in the problem solving process. As we noticed in some previous examples, in fact, linking and iterations signs seem to be crucial signs for the generation of recursive argumentations and their use and exploration seem to support the student's construction of such argumentations.

The first excerpt will show an example in which this intervention seems to be effective for the student's construction of a recursive argumentation. In particular, after the student has been supported by the interviewer in producing a LS, he starts proving by induction his conjecture, referring to the just produces LS.

The second excerpt, instead, will show an example of the almost opposite situation. To help a student who is having some difficulties in proving the inductive step within a proof by MI, the interviewer guides the student in producing a sign (which is a LS for the interviewer). However, the student does not seem to interpret the sign in a useful way for constructing the proof of the inductive step.

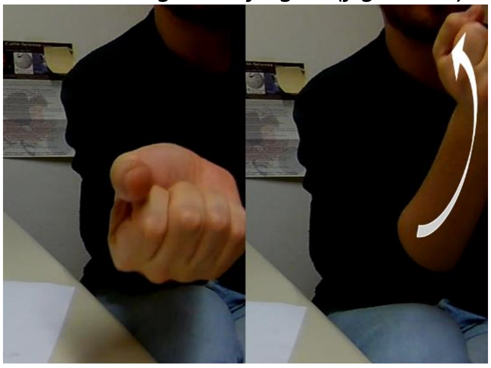
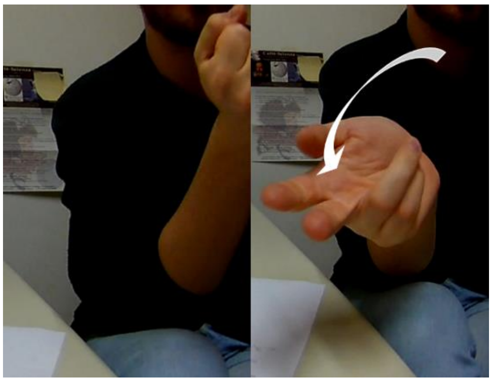
In summary, both the excerpts will highlight that suggesting a LS is only a possible strategy of intervention and that an important aspect for its effectiveness is how the student interprets the suggested sign.

Excerpt 10.10

Guido is dealing with the banknotes problem. During the first fifteen minutes he wrote the expression '$\alpha p+\beta q$' to represent the linear combination of $p=3$ and $q=5$ where α and β are integers. Then he determined the values of α and β to obtain the numbers from 1 to 7, concluding that sometimes α and β must be negative. After that he continued for the numbers from 8 to 13, finding a combination where $\alpha,\beta\geq 0$. He conjectured that 8 could be the limit after which every natural number can be reached with banknotes valuing 3 or 5. Then he tried to prove his conjecture saying that all the numbers which are multiple of 3, 5 or 8 can be reached but then he stopped following this strategy.

After over two minutes of full silence, the interviewer decides to intervene asking Guido what he is thinking about.

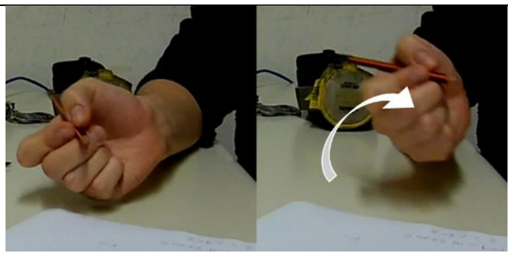
	Who	Speech	Gesture and inscription
1	**G**	I was still making some examples because nothing comes to my mind.	
2	**I**	Good, let's remain on the examples, it is fine. Let's start, I don't know… fifteen, no?	
3	**G**	Fifteen is three by five. Three times five or five times three.	
4	**I**	Yes, three times five or five times three, good.	
5	**I**	Think of this particular case. You are in a shop and the shopkeeper says: "This costs fifteen" so you give five [banknotes] of three, let's say.	
6	**G**	Yes, let's say three of five.	
7	**I**	Ok, three of five is better. Five, five, and five. The shopkeeper takes the money then gives them back to you and says: "Sorry, I made a mistake, it was sixteen"	
8	**G**	But I only have banknotes…	
9	**I**	Yes, you only have banknotes of five or of three.	
10	**G**	So you need to make two of five and two of three.	

11	**I**	Is there a way to change the banknotes you have already used without starting the count from the beginning?	
12	**G**	Ah, yes! You mean…	
13	**I**	I mean, you put three banknotes of five…	
14	**G**	Then I need to take away one banknote of five and give two of three and I think it is the only option…	*With the right hand, he makes the gesture of taking something (fig. 10.61) and then the gesture of putting something down whilst showing three fingers (fig. 10.62).*  *Fig. 10.61*  *Fig. 10.62*
15	**I**	Yes, probably the only option…	
16	**G**	Given these data.	
17	**I**	Ok.	
18	**G**	*[5 seconds]* Ah… I could use… *[5 seconds]*	

19	**I**	I don't know if it has helped…	
20	**G**	I think it has helped, instead… *[5 seconds]*. Yes, yes, this helps…	

After one minute of silence, Guido writes on the sheet the expression '$3^n \cdot 5^m$ with n, m ≥ 0'. The interviewer asks what the expression represents and, after some calculation Guido realises that it does not represent what he wanted. He cancels the previous expression, and he writes 'n3 + m5' and 'n, m ≥ 0'. Then, he continues:

[…]			
21	**G**	So, we do it for eight. It is n equal to one, m equal to one.	*He writes the following inscription.*  *Fig. 10.63*
22	**I**	Ok	
23	**G**	I assume… I assume to know… let t be strictly greater than eight… I can write t minus one	*He writes 't>8' and then 't-1' under the previous inscription.*
24	**I**	Ah, wait, are you proving by induction?	
25	**G**	Yes…	
26	**G**	I can write t minus one as n three plus m five… so this one is t minus one.	*He writes 't-1 = n3 + m5'.*
27	**I**	Ok.	
28	**G**	One is equal to two times three minus five.	*He writes '1 = 2 · 3 - 5'.*
29	**G**	Thus, I take off a banknote of five…	*He makes a similar gesture to the one of fig. 10.61.*

			 *Fig. 10.64*
30	**G**	Which I can do it… yes… ok, assuming that m is greater or equal to one, I have finished.	*He circles the m in the expression of line 26 and then he draws an arrow and writes 'm≥1' on the right.*

In figure 10.65 it is reported what Guido wrote on the sheet from line 20 to line 30.

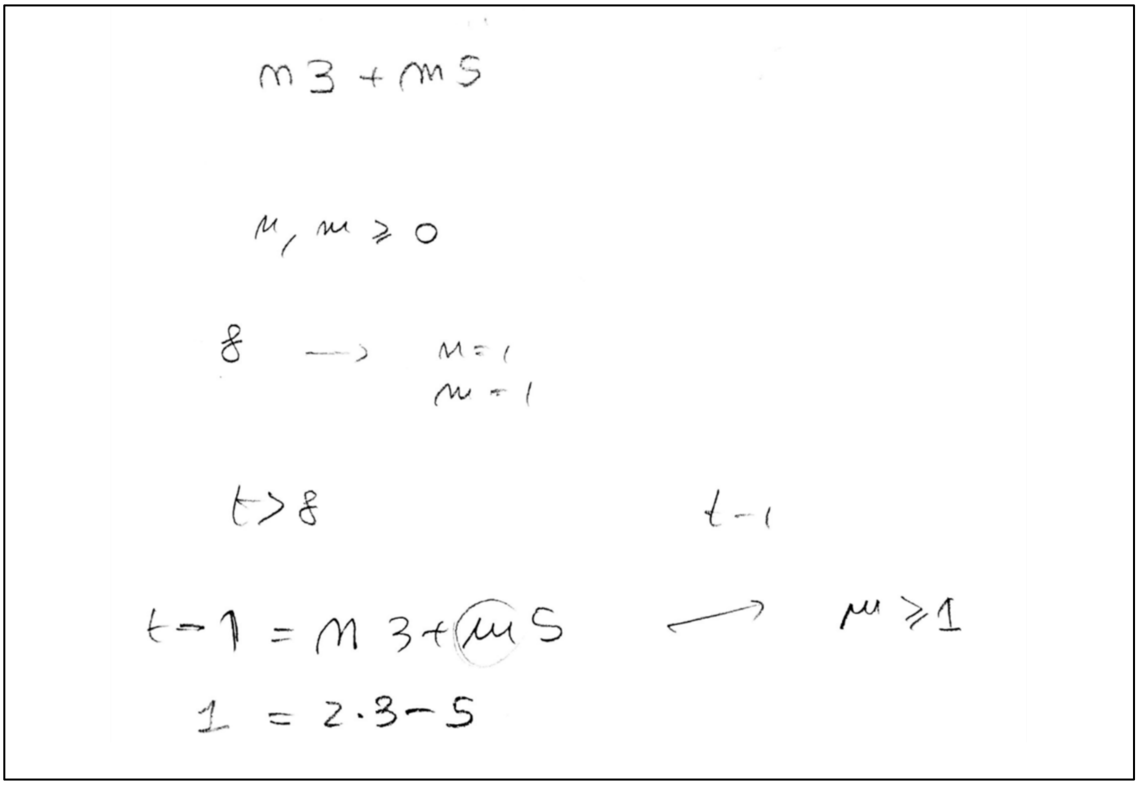


Figure 10.65. *Guido's inscriptions from line 20 to line 30.*

In the successive minute Guido explains to the interviewer that he just proved the inductive step when t-1 is written as n3 + m5 with m≥1. Note that Guido has not written explicitly the expression for t but only the expression for t-1 and for 1 (see last lines of the fig. 10.65). The passage from the expressions for t-1 and 1 to the expression for t remains implicit. After this, Guido says that to complete the proof he needs to prove the inductive step also when, in the expression for t-1, m is equal to 0.

31	**G**	Let's suppose that… so t minus one is equal to n times three. Ok?	*On a new page, he writes 't-1 = n3'*

32	**I**	Ok	
33	**G**	Now I have to do this case. I'll think a second…	
34	**I**	Sure	
35	**G**	This means that t is equal to n three plus one and I need to find…	*He writes 't = n3 +1'.*
36	**G**	It won't be difficult… If I take… yes, I take off three times three… ah but then I have the same problem…	
37	**G**	Yes, but here we are in the case t strictly greater than eight thus n must be greater or equal to three.	*He writes 't>8' and 'n≥3' underneath.*
38	**I**	Ok	
39	**G**	So, I can take off three times three, add two times five… and therefore I also solved this case.	*He draws a circle in the air with the pen over the expression he wrote at line 35.*
40	**I**	Because you take off nine and you add ten	
41	**G**	Yes. And so, we have proved it.	

In figure 10.66 it is reported what Guido wrote on the sheet from line 31 to line 41.

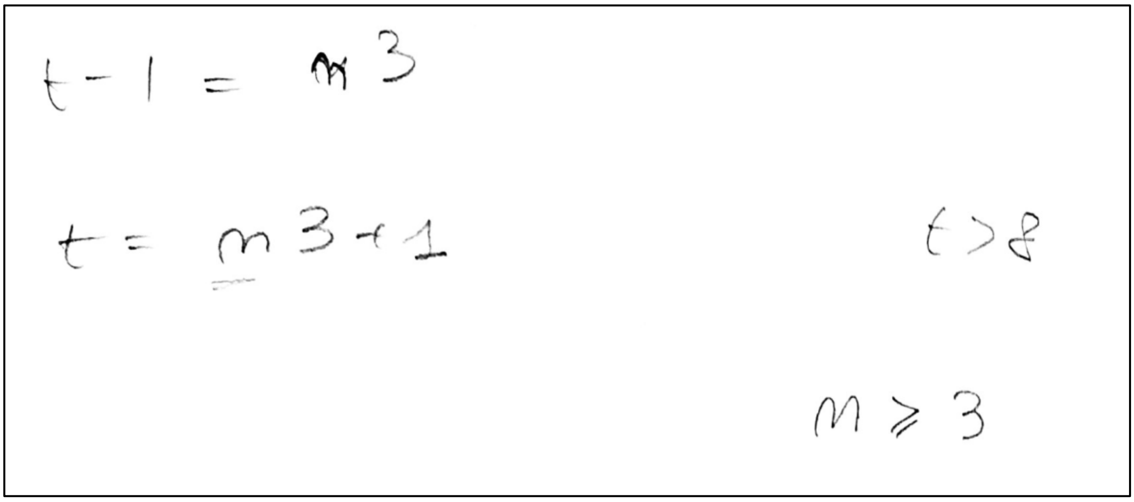


Figure 10.66. *Guido's inscriptions from line 31 to line 41.*

Analysis

The excerpt starts with Guido's claiming of having no ideas of how to prove his conjecture which, as said, was that every number greater or equal to 8 can be obtained using a positive number of banknotes valuing 3 or 5. Indeed, in line 1, He says: "I was still making some examples because nothing comes to my mind". The interviewer, thus, starting from this claim, tries to help Guido. Firstly (line 2) he says to consider the number 15 that, as Guido immediately says, can be obtained either with three banknotes valuing 5 or with five banknotes valuing 3. Then the interviewer asks Guido to imagine to be in a shop, and that he needs to pay 15 'money' for something. After Guido says that he would do it by using three banknotes valuing 5 (line 10), the interviewer asks what one can do to modify 15 into 16: "Is there a way to change the banknotes you have already used without starting the count from the beginning?". With this question the interviewer is asking Guido to find not a direct way to obtain 16 (which he has already found) but a way to obtain 16 starting from the banknotes used for obtaining 15. In other terms, the interviewer's aim is that Guido focuses on 15 and 16 as two linked cases. Guido immediately answers the question (line 14) producing what we could interpret as a LS: he says that for passing from 15 to 16 one can substitute a banknote valuing five of the ones used for making 15 with two banknotes valuing 3. In describing this Guido makes two iconic gestures (fig. 10.61-10.62) which mime the actions of taking back and putting down on the table something. Guido, guided by the interviewer's question, has therefore linked the two cases, finding a way to obtain 16 from 15, "given these data", i.e., that 15 was obtained with three banknotes valuing 5. The production of this LS seems to unblock Guido. Firstly (line 18) he says "Ah, I could use…", unfortunately without adding other information. Considering how the interview develops (Guido in the next lines will construct a proof by induction), It could be that in this line he is referring to MI, but we cannot know it with certainty. However, in any case, as Guido himself claims in line 20, he is convinced that the observation of before (namely, the LS between 15 and 16) could help him to prove his conjecture ("I think it has helped […] Yes, yes, this helps").

Guido, after writing the expression 'n3+m5' with 'n, m ≥0', starts constructing a proof (that later he will reveal to be by induction) of the fact that every number greater or equal to 8 can be written as in the expression of above. Note that this is not explicitly said by Guido, but it can be reconstructed from what he does and what he says. Firstly (line 21) he proves the base of the induction showing that 8 can be obtained with n=1 and m=1 in the expression of above (fig. 10.63). Then he starts proving the inductive step (as he says in response to the interviewer's question, lines 24-25). He firstly assumes that t-1=n3+m5 for some n, m ≥ 0 (lines 23 and 26). Then (line 28), he observes that 1 can be obtained as *'2 · 3 - 5'*. Guido, in this passage, seems to recall the LS produced before: he says: "I take a banknote of five…" and he performs a gesture (fig. 10.64) similar to the one done in line 14. The numerical expression *'2 · 3 - 5'* just written is indeed the numerical equivalence of substituting a banknote valuing 5 with two banknotes valuing 3. Guido seems therefore to reconstruct the LS which previously connected 15 with 16 in a generic LS connecting t-1 with t. In doing this, however, Guido realises that the strategy previously used for 15 and 16 was possible only because 15 was obtained with some banknotes of 5, one of which could be substituted. Within the proof of the inductive step, this corresponds to a situation in which t-1= n3+m5 with m≥1. This is observed by Guido in line 30 when he concludes that the inductive step is proved for this first case. Note that, in truth, if we only consider what Guido wrote (fig. 10.65) or what he said, he has not formally concluded the inductive step. In fact, he did not show how, using the expression for t-1 of above (the inductive hypothesis) and the fact that 1=2 · 3 – 5, one could conclude that t=a3+b5 for some a, b ≥0.

In the final part of the excerpt (lines 31-41), Guido proves the inductive step for the case in which t-1= n3, that is the expression of above when m=0. For this second case, Guido has not previously constructed a LS to represent a way for passing from t-1 to t, as instead it happened before. Indeed, he says to need some time to think about it (line 33). After a few seconds he realises that in in this second case one can transform t-1 in t by substituting three banknotes valuing 3 with two banknotes valuing 5. This is possible, as he observes in line 37, because t-1 = n3 and t-1 ≥ 8, from which it follows that n≥3 or, in other terms, that in t-1 there are at least three banknotes valuing 3. In line 39, Guido says that this concludes the proof of the inductive step also for this second case. Considering what he wrote (fig. 10.66), we can notice that, similarly to before, the inductive step has not been completely proved. Guido would still need to show that using the inductive hypothesis and the fact that 1 can be obtained as $2 \cdot 5 - 3 \cdot 3$, one can conclude that t=a3+b5 for some a, b≥0.

What this excerpt shows is that the intervention of the interviewer which prompted Guido's production of a LS, seems to have been effective in helping Guido. It is interesting to notice that, right after the production of the LS, Guido said that "this helps" and started proving his conjecture by induction. It seems that Guido recognised in the LS the structure of an inductive step. Indeed, when proving by induction he constructs the proof of the inductive step (in the first case) referring to the LS that he previously produced. The proof by MI that Guido has written at the end misses a few details, as observed, but this seems to be related to the fact that he was not writing a poof by MI as a product yet, but instead he was still trying to find a way for proving his conjecture. What it is important to notice, however, is that the excerpt started with Guido's claim of having no ideas of how to prove his conjectures and it ends with Guido who is convinced to have proved it. The LS of the first part of the excerpt (suggested by the interviewer and constructed by Guido) seemed to have had an important role in this transition, in moving the attention from the single cases to the connection between these cases.

Excerpt 10.11

Lucrezia is dealing with the chessboard problem. In the first twenty minutes of the interview, she explored the problem finding a tessellation for the 2x2 chessboard and for the 4x4 chessboard which leaves out only one little square. The interviewer showed to Lucrezia that, for reasons of divisibility it cannot be obtained a better tessellation, i.e., a tessellation which covers all the squares of the chessboard. Then, Lucrezia conjectured that, maybe, for a generic 2^nx2^n chessboard one could find a tessellation which covers all the squares except for one. This conjecture seemed to be obtained by empirical induction ("It is true for one, it is true for two…") and not supported by any recursive argumentation.

After this, Lucrezia says that she could try to prove by induction her conjecture. Unfortunately, she does not make explicit the reasons of this choice, however it is important to notice that she has not previously constructed any LS involving two consecutive chessboards. Lucrezia shows that the base case was already proved (i.e., a 2x2 chessboard, completely tiled except for one square). Then she starts proving the inductive step ("I assume it true for n and I prove it for n+1), but then she stops writing. After a few minutes of silence, the interviewer asks Lucrezia if there is something she is not sure about. Lucrezia says that she does not know how to use the inductive hypothesis. The interviewer decides to directly intervene and to help Lucrezia in constructing the inductive step. Firstly, he guides Lucrezia to draw the inscription of figure 10.67, representing a generic 2^{n+1}x2^{n+1} chessboard.

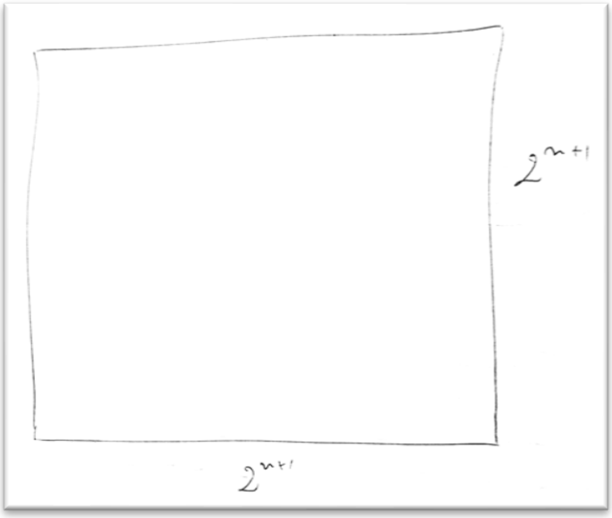
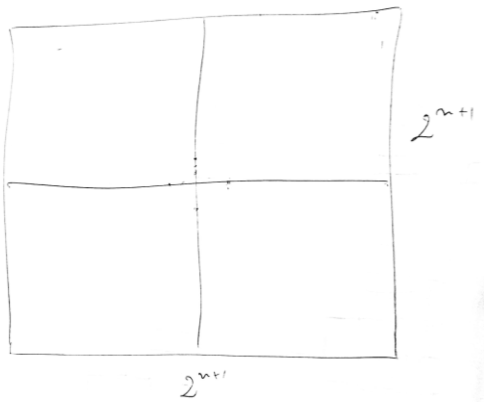


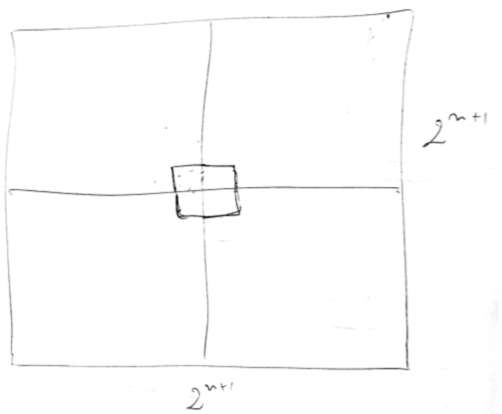
Figure 10.67. *Lucrezia's inscriptions representing a 2^{n+1}x2^{n+1} chessboard.*

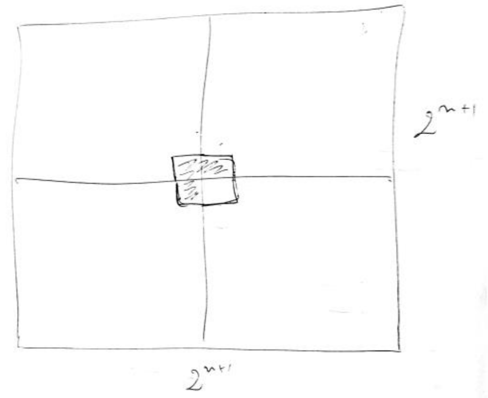
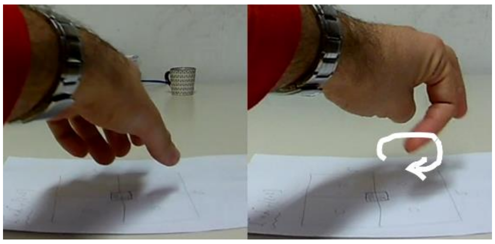
The interviewer, then, tries to support Lucrezia in recognising that the 2^{n+1}x2^{n+1} chessboard can be seen as composed by four 2^nx2^n chessboards.

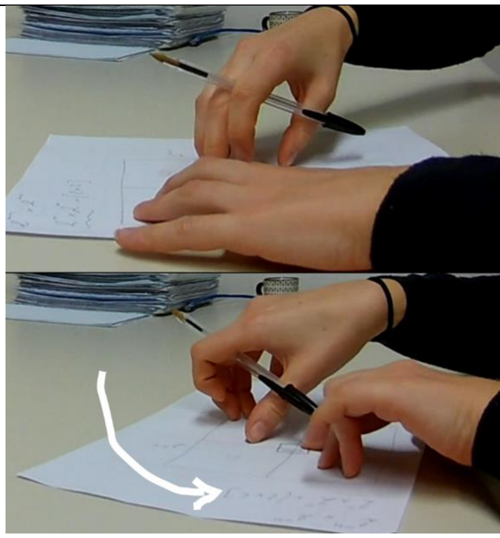
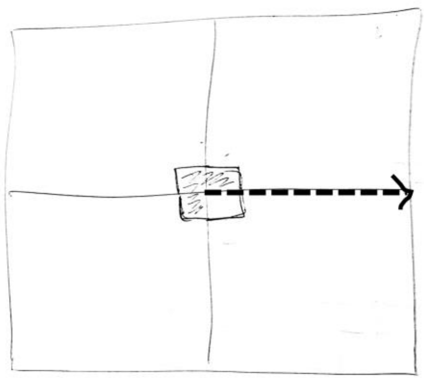
	Who	Speech	Gesture and inscription
1	**I**	For instance, if you divide in half this and this… I mean you divide it in four.	*He draws with the finger a cross over the inscription of fig.10.67 as to divide it in four squares.*
2	**L**		*She draws the lines indicated by the interviewer in the previous line, obtaining the following inscription.*  *Fig. 10.68*
3	**I**	Now each of these…	

4	**L**	It is ours. I mean two to the n by two to the n.	

Lucrezia, however, seems to have some difficulties in the construction of the inductive step, in particular in using the inductive hypothesis. After a couple of minutes, the interviewer decides to make another intervention. He suggests that one could try to prove the inductive step in this form: assuming that a 2^nx2^n chessboard can be completely tiled except for one square which occupies one of its corners, then a 2^{n+1}x2^{n+1} chessboard can be tiled completely except for one square which still occupies on its corners. Lucrezia does not seem to be convinced about this possibility but then, after having observed that for the 2x2 and for the 4x4 chessboard such a tessellation was possible, she seems to accept the suggestion. Then the interviewer asks Lucrezia to continues the proof of the inductive step.

[…]			
5	**L**	Good. So, I put one here, one here and one here. You want to do this, right?	*She draws four little squares in the centre of the drawing of the 2^{n+1}x2^{n+1} chessboard, obtaining the following inscription.*  *Fig. 10.69*
6	**I**	Yes.	
7	**G**	Then I put an L-shaped tile here, right?	*She colors with the pen three of the four squares just drawn, obtaining the following inscription.*

			 *Fig. 10.70*
8	**I**	Ok. But this one is not in one of the corners.	*He points the little white square at the center of the inscription of fig. 10.70.*
9	**L**	Mhm. Yes, and we want it in one of the corners…	
10	**I**	Yes, because we said: "I can do it, with one square which remains empty in one of the corners". But I can do this… I rotate it… and I put in in one of the corners.	*He makes the gesture of rotating the bottom right square of the inscription of fig.10.70.*  *Fig. 10.71*
11	**L**		*She makes a gesture as to say to wait.*
12	**I**	I mean…	
13	**L**	Yes, I know what you want to do. You want to do this.	*She puts the right hand on the bottom left square of the inscription of fig.10.70, as if she was holding it, and symoultaneusly she rotates the whole sheet anticlockwise.*

			 *Fig. 10.72*
14	**L**	But I don't know if here I will break some L-shaped pieces,	*With the pen she follows the line which separates the bottom right square with the top right square in the inscription of fig.10.70.*  *Fig. 10.73. The arrows indicate the trajectory of the pen over the inscription.*
15	**I**	while rotating.	*She makes the gesture of rotating the bottom right square, similar to the interviewer's gesture of fig. 10.71.*

Finally, the interviewer says that each of the four squares represent a 2^nx2^n chessboard whose tessellation, for the inductive hypothesis, is independent from the others; therefore, there cannot be any L-shaped tiles which belongs to two different 2^nx2^n chessboards. Lucrezia seems to accept this explanation and the resolution of the problem is concluded by the interviewer.

<u>Analysis</u>

In the first part of the excerpt (lines 1-4), the interviewer guides Lucrezia in drawing a generic 2^{n+1}x2^{n+1} chessboard (fig. 10.67) and then in recognising that this one can be divided in four squares, each representing a 2^nx2^n chessboard (fig. 10.68). This second inscription is, for the interviewer, a LS: it represents the fact that the tessellation of a 2^{n+1}x2^{n+1} chessboards can be obtained using four independent tessellations of four 2^nx2^n chessboards. However, as we can observe from how the excerpt develops, Lucrezia does not seem to have interpreted it in the same way. Indeed, in the second part of the excerpt, when Lucrezia constructs the inductive step together with the interviewer, she seems not to see that the four parts in which the big chessboard is divided represents four independent chessboards with their own tessellations. In fact, when in line 22 the interviewer says that to conclude the inductive step one needs to rotate one of the smaller chessboards, she says that some L-shape tiles may break in this rotation, pointing at the line which separates two of the four chessboards (lines 25-27). This comment might reveal that Lucrezia is not seeing the four $2^n x 2^n$ chessboards each independently tessellated which form a tessellation for the $2^{n+1} x 2^{n+1}$ chessboard, but that instead she is only seeing the $2^{n+1} x 2^{n+1}$ chessboard with a whole tessellation independent from the four $2^n x 2^n$ chessboards. In other terms, the inscription in figure 10.70, which for the interviewer is a LS connecting the tessellation of the $2^{n+1} x 2^{n+1}$ chessboard with the tessellation of four $2^{n+1} x 2^{n+1}$ chessboards, does not seem to be a LS for Lucrezia. For her this inscription does not link two chessboards of the problem and their tessellations but it seems that for her it only shows that a $2^{n+1} x 2^{n+1}$ already tessellated chessboard can be divided in four parts whose dimensions are those of four $2^n x 2^n$ chessboards, but which are not four independently tiled chessboards. In addiction of what already said, we can notice that both the interviewer and Lucrezia made a similar rotating gesture on the inscription of the chessboard (respectively the interviewer in line 10, and Lucrezia in lines 13 and 15). If seen from outside the gestures look the same, however they represent two different situations for the two subjects: for the interviewer it represents the action of rotating a $2^n x 2^n$ independently tiled chessboard which composes part of the tessellation of the $2^{n+1} x 2^{n+1}$ chessboard; for Lucrezia it represents the action of rotating a part of the tiled $2^{n+1} x 2^{n+1}$ chessboard and this part is covered by the tessellation of the $2^{n+1} x 2^{n+1}$ chessboard in a way that rotating it could break some tiles (the ones which might belongs both to the part which is rotating and to the rest of the chessboard).

In conclusion, the whole excerpt contained an example of an intervention by the interviewer who, to help Lucrezia in the construction of the inductive step suggested a sign (a LS for the interviewer) which however has not been interpreted as such by Lucrezia. She, indeed, does not seem to use this sign properly to construct the inductive step. In terms of the proof construction, thus, the suggestion of the LS by the interviewer does not seem to have been effective for supporting Lucrezia in the process of proving by induction. This excerpt, moreover, highlighted an important aspect of linking (and iteration) signs. Coherently with the Peircean definition of sign adopted in this study (see 6.2), every LS/IS is personal, it requires the presence of a subject interpreting it as a linking (or iteration) sign. In other terms a gesture, an inscription, a spoken utterance, or a bundle of them, could be a LS/IS for a subject without being a LS/IS for another subject, as it happened in this excerpt between the interviewer and Lucrezia.

10.2 First conclusions on Linking and Iteration signs

In this chapter, I have presented two categories of signs which emerged from the interview's analysis: the Linking Signs and the Iteration Signs. In particular, the analysis of these signs, of their production and use by students involved in a problem solving activity, highlighted some interesting aspects.

First of all, we could observe how, often, the students' production of LS and IS seems to characterise the generation of a recursive argumentation during the phase of exploration of the problem. As observed in some examples of above, in fact, **the production of LS and IS highlights those moments in the exploration of the problem related to the generation of a recursive argumentation to support a conjecture.** In a certain sense, LS and IS can be read as signs which could anticipate a student's construction of a recursive argumentation. Thus, a researcher or a teacher in registering the presence of LS and IS in a student's semiotic production could observe the trace of a recursive argumentation when the student is still exploring the problem or generating a conjecture for it.

Secondly, we observed how LS and IS may become resources for the students during the whole problem resolution activity and, specifically, also during the proving process. In particular, **the production, use, and exploration of these signs seem to support students in planning, organising and constructing a recursive argumentation (or, more specifically, a proof by induction), both at the level of its general logical structure and of the specific problem involved.** As we observed, LS and IS can be resources both when autonomously produced by a student and, sometimes, when their production is suggested by an external intervention. Moreover, analysing the interviews we registered how LS and IS can become resources for a subject in different ways. Sometimes a student recognises in a LS the structure of an inductive step and, starting from it, s/he succeeds in constructing a recursive argumentation or the inductive step itself within a proof by induction. Some other times, instead, the inverse process seems to happen: a student anticipates, through the production of LS and IS, the structure that a recursive argumentation should have, and then s/he constructs it for the problem under investigation with reference to those signs. In this case the production of LS and IS seems to have supported the student in planning and organising the successive construction of a recursive argumentation.

A third noticed aspect is that the observation of LS and IS can become a methodological instrument with which to register the possible students' cognitive unity between argumentation and proof, with a focus on mathematical induction. In particular, as observed, the production of those signs can be registered across the argumentative phase and the construction of a proof phase. This can happen when a student investigates on a LS, developing and enriching it to construct an argumentation, and then uses this sign in the construction of a proof by induction. Otherwise, it is possible that a student, when describing the structure of the argumentation s/he is constructing, produces (linking and iteration) signs which then are repeated when describing the structure of the proof by induction used to formalise her/his argumentation. In both cases, the observation of this semiotic continuity might reveal a student's cognitive unity between her/his argumentation and the subsequent proof by induction.

Finally, a further registered aspect involves the specific use of gestures in constructing recursive argumentations. As said in the presentation of the conceptual framework, recent research (Arzarello, Sabena, 2014; Sabena 2018) highlighted how gestures actively contribute to support students in structuring their argumentations. What this study has registered confirms these results. In particular, as observed in the second point of above, **the production of gestures**, which can be interpreted as linking or IS, **does not seem to have only a communicative role** (which is, of course, present), **but it also supports students in planning and organising, a recursive argumentation**, both in reference to its logical structure and in reference to the specific problem involved.

This last point, i.e. the fact that some signs could refer to the structure of a generic recursive argumentation while some others could refer to the specific problem considered, has been one of the starting points for a second analysis of the interviews. The result of this analysis will be the object of the next chapter.

11 SECOND ANALYSIS OF THE INTERVIEWS – LEVELS OF SEMIOTIC PRODUCTION

In this chapter, I will present a second aspect emerging from the data analysis, which is that, focusing on students' semiotic production, it is possible to distinguish signs which can be interpreted as belonging to two different levels: ground level signs and meta level signs.

11.1 LEVELS OF SEMIOTIC PRODUCTION: GROUND AND META LEVEL

During the analysing of the interviews, it was possible to register that, in some moments of the problem resolution, the students seem to pause calculations, symbolic manipulations, geometrical transformations, etc., in order to observe, reflect on, or describe what they have done, what they are planning to do, the structure of their own argumentation or the connections between some aspects of the problem with some mathematical theories. Through an analysis of students' semiotic production, it was possible to observe these 'reflective' moments and to register how they seem to play an important role in the problem resolution. In particular two different 'levels' of semiotic production have been distinguished in this analysis: **the ground-level and the meta-level semiotic production**. To better introduce them, I will borrow some voices of influent scholars in Mathematics Education.

Kilpatrick (1986), analysing the terms "reflection" as a particular process of a subject, writes:

> we somehow move into another dimension when we reflect on what we have done. In reflecting on our experience, we move out of the plane of our everyday existence. We give meaning to experience by getting outside the system. (p. 9).

In an analogous way, Sierpinska (2005), in describing what she means for "Theoretical Thinking", says:

> For theoretical thinking to even begin, the thought and its object must belong to different planes of action. Thinking-in-action must become thinking-about-action. The moment the philosopher reflects back on his thinking, verifying if it is well founded, his relation to the object of his thinking becomes theoretical. (p.120).

In line with these two quotations, analysing the interviews with a semiotic perspective it has been possible to see traces of these two "dimensions" or "planes of action", to say it respectively with Kilpatrick or with Sierpinska. In particular, observing students' semiotic production during the resolution of a problem, it was possible to recognise signs which did not refer to some specific objects of the problem or to some transformations of them, but to the problem itself or to its resolution. In these moments, students produce a discourse which refers to the problem's structure, to its resolution, or to heuristics that have been used or to be used. In other terms, in these moments students seemed to reason not *in the problem* but *about the problem,* paraphrasing Sierpinska's quotation of above.

These two different levels of discourse involved during a problem resolution (the one referring to the objects of the specific problem, and the one referring to the problem itself and to its resolution) can also be found in Polya's famous book (1945), *How to solve it.* The first part of the book is structured as a dialog between the author, playing the role of a teacher, and some imaginary students involved with some problems. Polya proposes some questions that might guide students during the resolution of the problems ("What is the unknown?", "Do you know a related problem?", "Can you check the argument?", etc..). Polya's questions are expressed in general terms and in fact are used in the same form for a series of different problems. In the second part of the book, Polya describes a series of different heuristics that might be helpful to solve a problem. These heuristics before being shown on some examples are described in general terms, referring to a generic problem, to its hypothesis, conclusions, figures, etc... The questions of the first part of the book and the heuristics' descriptions in the second part are examples of a discourse situated on a different level that the one of the specific objects of a specific problem. This level, as it will be defined below, will be called meta-level.

Similarly, a trace of these different levels can also be found in some famous works by Schoenfeld, which focus on the role of metacognition in problem solving processes. In Schoenfeld (1982), for instance, the researcher presents a teaching experiment with a group of undergraduate students with the aim of improving their competences in problem solving. Whilst involved in some problem resolution activities, students are asked by the teacher to answer the following questions, written on a poster in the classroom: "What (exactly) are you doing? (Can you describe it precisely?) Why are you doing it? (How does it fit into the solution?) How does it help you? (What will you do with the outcome when you obtain it?)" (Schoenfeld 1982, p. 34). Schoenfeld registers that these questions seem to support the students during the problems resolution, prompting the activation of some control and planning strategies. Note that the questions do not refer to any specific object of the problem, but to the resolution process itself. In a different study (1985), Schoenfeld describes the potential complexity of the resolution of a problem, metaphorically described as a path from an initial state and a goal, using the diagram reported in figure 11.1. The author claims that this path is very often rich of crossroads and, to be fully crossed by a subject, it often needs some 'steppingstones' (i.e., other problems whose resolution is helpful for the resolution of the first problem).

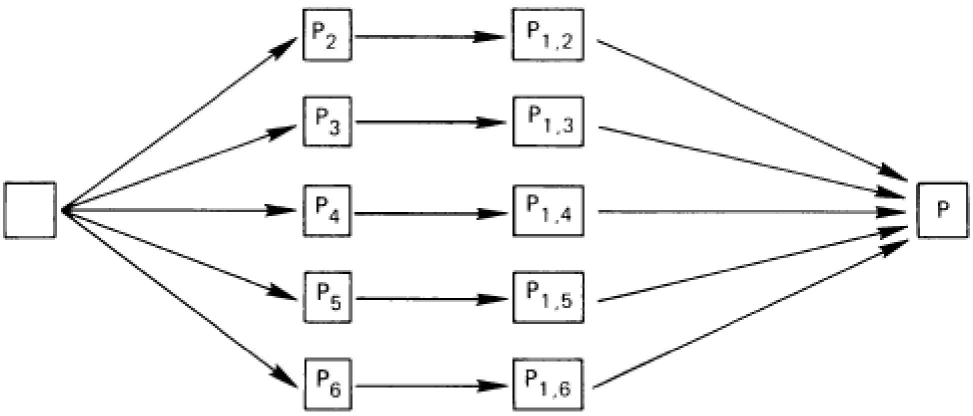


Figure 11.1. Schoenfeld's diagram (1985, p. 88) representing a problem as a series of different paths (with different steppingstones) to reach the goal P starting from the initial state (represented by the blank square).

Even without getting into detail of Schoenfeld's diagram, we can still observe how with this he is representing a problem and the structure of its possible resolution. This diagram is composed by several signs (the labels and the arrows) which do not refer to some specific objects of a problem but to the problem itself, to other problems and to their possible relationships within the problem P. Both the questions written on the poster described in Schoenfeld (1982) and the diagram in Schoenfeld (1985) are other examples of signs belonging to that meta level introduced above.

Having said that, in the following section I will give a definition for these two levels of semiotic production to which I have alluded in the previous paragraphs.

11.1.1 Ground and meta level

In relation to a subject's resolution of a problem, I introduce the following distinction for the subject's semiotic production:

Ground Level. I will say that the subject's semiotic production is *Ground Level* when it refers to the (mathematical or not) objects of the problem (formulas, polynomials, matrixes, graphs, chessboards, coins, etc.), both those explicitly described in its formulation and those introduced by the subject during the resolution, to their properties and to their mutual relationships.

Meta Level. I will say that the subject's semiotic production is *Meta Level* when it refers to the problem itself or to its components, to the resolution process, to heuristics, to argumentations or to semiotic representations, as objectified entities.

For instance, if we refer to a specific problem with a diagram such as the one of figure 11.1, or we pose questions similar to Polya's and Schoenfeld's questions of above, the level of our semiotic production is meta level, whilst if we write some calculations during a problem resolution the level of our semiotic production is ground level.

More specifically, the Meta level semiotic production can be described in terms of four different, but not mutually exclusive, categories. This description helps in characterising what I mean, more specifically, with meta-level:

Meta-Heuristic. The semiotic production refers to a heuristic as an object (as in the book of Polya, for instance), to its development or to its comparison with other heuristics. For instance, in these cases, the subject refers to the "cases of the problem", to the "goal", to "exploring the problem", to "look for analogous problems", to "split the problem in sub-problems". I will consider the term heuristic in a broader sense so to include the case in which the problem corresponds to construct a proof for a statement. In this case a proving scheme can become a heuristic for a subject; therefore, they belong to this category also all the considerations that a subject might do in relation to the structure of a particular proof. For instance, in relation to mathematical induction, an expression such as "I'll do the case base and then the inductive step" will be considered as a meta-level (meta-heuristic, in particular) expression.

Meta-Theoretical. The semiotic production refers to the connection between the problem and a particular mathematical theory in which the problem can be situated. The term "meta-theoretical" instead of simply "theoretical" is used to stress that I am not referring to some applications of theoretical results to the objects of the problem (for instance the sentence "This function is increasing because its derivative is positive" is not meta-theoretical), but to some considerations about the connection between the problem (or parts of it) and some theoretical results ("This problem seems to be linked to Lagrange's theorem"). The subject, therefore, reflects on the problem and on theorems as particular objects of a mathematical theory, and on the connections between them. To this category they also belong a subject's considerations about the epistemic value of a statement ("This is surely true", "This is probably false").

Logical. The semiotic production refers to the underlying logical theory (which most of the time corresponds to the classical logic). Thus, it refers to some results of this theory (for instance, in the case of the classical logic, the equivalence between a conditional statement and its contrapositive one, or the conversion rules for quantifiers). To this category they also belong a subject's considerations about the validity or not of an argumentation or of a proving scheme (for instance, in our case, on the validity of MI).

Meta-Semiotic. It refers to semiotic representations involved during the resolution of the problem, to their use or transformations. In other terms it refers to the representations themselves. Using a Peircean terminology, it refers to the "representamen" of signs. Examples of a semiotic production of this kind are when a subject explicates the choice of using a particular representation ("I'll sketch the graph of this function") or an evaluation for it ("This drawing is not correct").

In table 11.1 I present a series of paradigmatic examples of ground and meta level semiotic productions that have guided the analyses of the excerpts presented later on in this chapter.

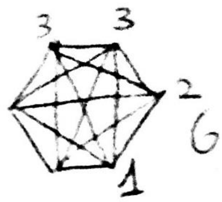
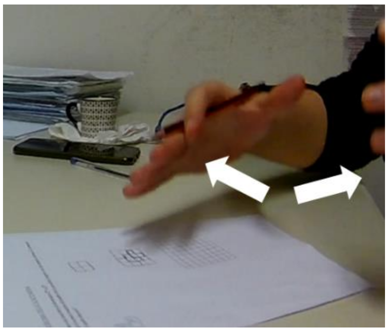
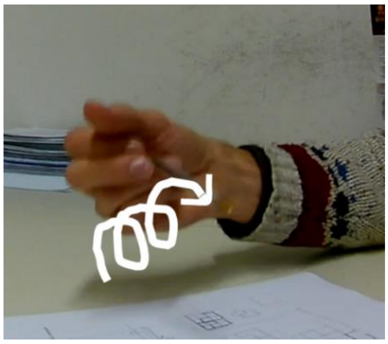
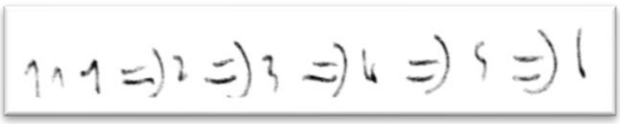
Ground level	"Drawing this diagonal here, two triangles will appear!"
	"We can divide everything by n and then this fraction will simplify"
	
	 "Widening this chessboard…"
Meta level	"Every case is linked to a previous one"
	"I know an analogous problem"
	"This problem is not easy"
	"The thesis can be written in a better way"
	 "I could do something like by induction"
	

Table 11.1. Paradigmatic examples of ground level and meta level semiotic productions. The examples are taken from excerpts that will be analysed in detail later on in this chapter.

A few observations on the definitions of ground and meta levels are necessary:

1) In these distinction between ground and meta level I would like to adopt the same point of view of Sierpinska (2005) about her distinction between practical and theoretical thinking:

   > These distinctions should be regarded as epistemological, not ontological distinctions. They are our simplified ways of knowing human cognitive activity in mathematics; they are not kinds of human cognitive activity. (p. 118).

   Moreover, as she writes:

   > Thinking takes place simultaneously at several planes of action, which can be considered separately only in theory, and even then, hypothesizing about the thinking at all of these planes in a subject at a given moment of an observation or interview may easily overwhelm even the most assiduous of researchers. (p. 134).

   Therefore, the proposed distinction in ground and meta level should be seen uniquely as a methodological tool for analysing the semiotic production of a subject and not as an affirmation of the existence of an ontological distinction between the two levels of signs.

2) Ground and meta level are to be intended as relative to the specific problem under resolution. In this sense, it is possible that the formulation of a problem outlines a series of ground level signs which, for another problem would be interpreted as meta level signs. For instance, in the problem "Prove that, in $\mathbb{N}$, the Well Ordering Principle implies the Mathematical Induction Principle", involves the two principles, which becomes objects of the problem. Therefore, expressions referring to them can be seen as ground level, whilst a reference to the principle of mathematical induction in the chessboard problem, for instance, could be interpreted as meta level.

3) The distinction between ground and meta level is not to be intended simply as a distinction between cognitive or metacognitive aspects. In particular some meta level signs are related to some metacognitive aspects (for instance a sentence like "I have to check a second time what I have just written" will be interpreted as meta level), but some others does not necessarily involve metacognitive aspects, such as in the description of the logical structure of a proving scheme or in a meta level spoken utterance such as "I can do the next case of the problem in the same way as before". This is the reason why in the characterisation of meta level semiotic production of above an explicit reference to the metacognition was not provided. In summary thus, we can say that a meta-level semiotic production could highlight some meta cognitive processes, but it could also highlight other non-metacognitive processes.

11.1.2 A first example

To better clarify the distinction in ground and meta level and how this was used to analyse the interviews I present a first brief example in which we can recognise the two levels of semiotic production.

Excerpt 11.1

Guido is dealing with the banknotes problem. In the excerpt 10.10, we have already presented and analysed Guido's resolution of the problem. The excerpt presented here, instead, is taken from the first minutes, when he is still exploring the problem. Right after reading the text, he writes the expression '$\alpha p+\beta q$' where α, β represent two integer numbers and $p=3$, $q=5$. Then he wants to find which natural numbers can be written with this expression if α, β are positive.

	Who	Speech	Gesture and inscription	Level[92]
1	**G**	So, one is necessarily negative. I mean, either alpha or beta is necessarily negative.	*He makes the counting gesture for 1 and then he points to the inscription* 'αp + βq' on the sheet.	*S: Ground* *G: Ground*
2	**G**	Two… either alpha or beta is negative… three, no… four, I would say… that either alpha or beta is negative.		*S: Ground*
3	**G**	Five, no… For six, no… Seven, either alpha or beta is negative		*S: Ground*

Guido continues the calculations in the same way for the numbers 8, 9, 10, 11 finding that they can be obtained with positive values for α, β in the expression of above.

[...]				
4	**G**	Twelve… Sure, ok. Thirteen… let's see… yes, thirteen is ok.		*S: Ground*
5	**G**	I would say that we can try to prove that after the eight… that after the eight it should be fine.		*S: Meta*
6	**G**	However, I don't know if… I mean it is still just a hypothesis.	*He rotates the left hand high close to his head.*	*S: Meta* *G: Meta*
7	**G**	Do I need to do a formal proof now?		*S: Meta*
8	**I**	It doesn't need to be formal, but try to prove it a bit… I		

92 In this and the following tables containing the transcript, an extra column, labelled as 'Level', was added. This column contains the indication of the level (ground or meta) of the semiotic production of the line. Note that the label will not be present if the line corresponds to an intervention from the interviewer. If in the line there are elements interpreted both of ground and meta level, I use the label ground/meta. The different semiotic sets are kept distinguished in this column, which means that different indications for spoken utterances (S), gestures (G), or inscriptions (I) are provided.

		mean at least you should be convinced.		
9	**G**	Yes, yes, I'll just think a second. Now I have done… I mean, generally I do the base case,	*He puts the two hands together as holding something over the sheet.*  *Fig.11.2*	*S: Meta* *G: Meta*
10	**G**	just to figure it out the situation and then… I think a bit better.	*He moves both the hands rotating them in the air with the fingers pointing down.*  *Fig.11.3*	*S: Meta* *G: Meta*

The first lines of the excerpt (1-4) present a ground level semiotic production: Guido has represented the problem's statement with the algebraic expression 'αp + βq' and he is making a series of calculation to explore to problem. The signs that he produces in these lines are all ground level. At this point Guido interrupts the considerations on numbers and moves his discourse on a different level, the meta level. Firstly, in line 5, Guido formulates what it will be his conjecture: "after the eight it should be fine", meaning that all the numbers grater or equal to 8 can be written with the formula of above with both α and β natural numbers. The utterance is interpreted as meta-level since, as highlighted by the use of the conditional ("It should be true"), it contains Guido's consideration on the epistemic value of the conjecture (called "it" by Guido). In line 6, he states that what he just said is only a hypothesis to be proved, then (line 7) he asks to the interviewer the level of formalism required for the proof, and finally (lines 9-10) he describes the heuristics that he is using: "Generally I do the case base, just to figure it out the situation, and then… I think a bit better”. With this last sentence, Guido explains what he “generally” does with a problem: firstly, he explores it in some examples “to figure it out the situation” (which he calls the case base, now), as he did in lines 1-4, and then starting from these he “think[s] a bit better” about the problem in general. Guido’s gestures of lines 9 and 10 are interesting and contribute to the meta level of Guido’s semiotic production. In line 9, when saying “the base case”, Guido makes the gesture of holding something with the two hands: the exploration of the problem for the first numbers of lines 1-4, is metaphorically represented as a material object hold with the two hands. In line 10, when saying “to figure it out the situation” he moves both hands rotating them around the position of the previous gesture. With this gesture, Guido seems to use his hands to explore the space in front of him, metaphorically representing the exploration of the problem for some examples. The gestures made by Guido do not refer the banknotes or numbers (the problem’s objects) but to the heuristics itself that Guido is using. Therefore, they are meta level signs.

This excerpt also introduces an important aspect of the meta level. Signs of this level might refer to some objects or material actions (as Guido’s gestures of lines 9-10) not to represent some specific objects of the problem, but to metaphorically represent something on the meta-level (in Guido’s excerpt, this is the strategy he is using). This is to say that the meta-level signs can have a certain materiality even if they do not refer to some specific objects.

11.1.3 Linking and iteration signs in the two levels

LS and IS have been defined in very general terms, the first ones in terms of connection between two entities the second ones in terms of composition of LS or of a reference to an iteration. This general definition, thus, leaves open the possibility for these signs to be either ground or meta level. For instance, the LS used in the proof of the inductive step for the formula of the sum of the first n natural numbers, reported at the beginning of Chapter 10 (fig.10.13), or the LS representing two consecutive chessboards in the chessboard problem (see Fig. 10.10, for example) are ground level signs. On the other side the LS produced as an arc-shaped gesture in the air while saying something like “I move to the next case” (as shown, for example, in fig.10.16) is a meta level sign. It is worth it to remind that the meta/ground level semiotic production is always relative to the problem under resolution. If for instance the problem asks to justify the validity of MI, then a LS composed by the utterance “Since the proposition is true for 0 then it is true for one” together with an arc-shaped gesture will be a ground level sign, while the same utterance and gesture produced in the chessboard problem, for instance, will be a meta level LS.

Having notices this, we can therefore combine the two semiotic analyses focusing both on LS and IS and on ground and meta-level. As observed in the previous chapter, LS and IS could become resources for a subject in the generation of a recursive argumentation and in its construction. The distinction between ground and meta level could be helpful in providing further insight on the role of LS and IS in the problem solving activity. In other terms, as we will see in some of the following analyses, LS and IS of different level (ground or meta) might become different kind of resources for the subjects, supporting them in different aspects of the problem resolution.

11.1.4 The dialectic process between ground and meta level

As said when the distinction of ground and meta level was introduced, analysing the interviews of expert students we registered a rich meta level semiotic production. This semiotic production is extremely intertwined with the ground level semiotic production in a way that it was possible to observe continuous transitions between the two levels. This dialectic process seems to support students' resolution processes, such as the process the process of construction of a recursive argumentation or of a proof by induction. More specifically, the semiotic production of ground level seems to be related with the construction of a specific recursive argumentation for the problem under resolution, whilst the semiotic production of meta level seems to be related with the subject's awareness of the fact that this argumentation can be effective to solve the problem, and with the awareness of the argumentation itself (of its structure, of its validity, of how it should be constructed).

To better describe the above mentioned dialectic between ground and meta level, I will introduced a series of terms corresponding to different phenomena that have been registered during the analysis of the interviews.

<u>Level's transitions.</u>

A level transition is registered when, analysing the diachronic development of the semiotic bundle, the semiotic production passes from one level to the other. When this passage happens without ruptures in the subjects' discourse, it will be called a **lifting**, for a transition from ground to meta level, or a **projection**, for the opposite transition.

An example of lifting has been already seen in the previous excerpt of Guido (Excerpt 11.1) from line 5 to 6. He expresses the conjecture in ground terms (line 5: "I would say that we can try to prove that after the eight… that after the eight it should be fine.") and then he refers to it with meta level sign (lines 6: "I mean it is still just a hypothesis."). A classic example of projection is when a subject describes a heuristic in general terms (for instance the structure of a generic proof by induction) producing meta level signs and then constructs it with reference to the objects of the problem producing ground level signs. I will present several examples of this kind later in this chapter.

Let us observe that not all the level's transitions are liftings or projections. For instance, let us consider the case of a subject who is performing some algebraic transformations (ground level) and who then, suddenly, interrupts the action to observe that s/he just recognised that the problem under resolution is a direct consequence of a known theorem (meta level). In this case the meta-theoretical observation of the subject has moved the discourse on the meta level, however there is not a continuity with the previous ground level semiotic production which, in fact, has been abruptly interrupted. In other terms, the semiotic production passed from ground to meta level, but with a rupture in the subject's discourse.

<u>Level's mismatch.</u>

In analogy with the term "gesture-speech mismatch" (analysed in 6.3.1), a level's mismatch is registered when, with a synchronic analysis of the semiotic bundle produced and used by a subject, it is possible to observe a bundle with a component of a semiotic set which is ground level and another component of another semiotic set which is meta level.

We will see examples of level's mismatches taken from the interviews later in this chapter. Since the excerpts from where they are taken will necessitate a too long analysis for this point, I would instead present the following invented example of level's mismatch. A subject dealing with a problem mimes the action of rotating a geometrical figure drawn on the sheet and whilst doing this s/he says: "I am trying to see if there are any symmetries that can help us". The gesture is ground level, whilst the speech contains a meta-heuristic observation, thus it is meta level. Thus, looking at the speech-gesture bundle we can register a level's mismatch.

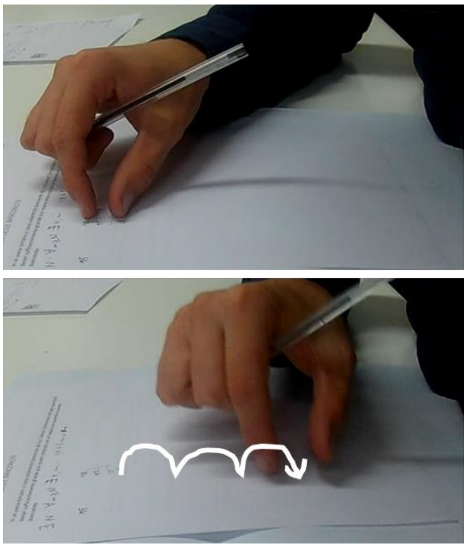
Merging point.

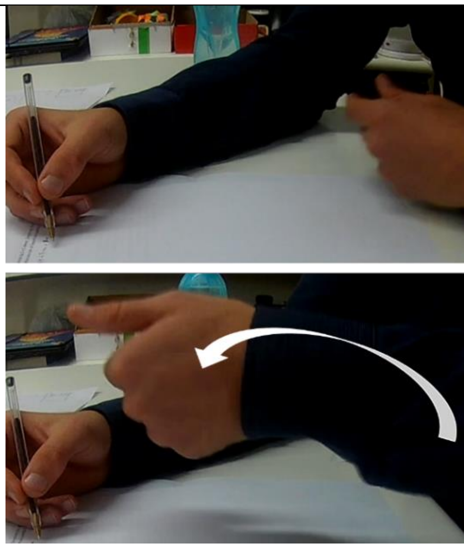
A merging point is registered when, analysing the bundle produced and used by a subject, it is possible to observe synchronically or in rapid succession, in the same semiotic set, a component which is ground level together with a component which is meta level. A merging point is thus similar to a level's mismatch with the difference that now the components of different levels belong to the same semiotic set, whilst in the previous case they belong to two different sets instead.

In the following short excerpt, we can see an example of merging point.

Excerpt 11.2

Lorenzo is dealing with the banknotes problem. The excerpt presents the moment in which Lorenzo conjectures that the problem can be solved by showing how to obtain the values corresponding to an interval of 10 consecutive numbers and then moving to the following intervals of 10 numbers adding, iteratively, two banknotes valuing 5. In the previous second he has just written in a column the numbers '10, 15, 20, 25'.

	Who	**Speech**	**Gesture and inscription**	**Level**
1	**L**	If I can do it for an interval of length ten, then I can do it for any successive interval of length ten,	*He puts the pointing finger and the thumb of the right hand on the sheet, respectively on the just written numbers 10 and 20. Then keeping pointing finger and thumb at a fixed distance he moves the right hand down on the sheet shaping some arcs in the air.*  *Fig.11.4*	*S: Meta* *G: Meta*
2	**L**	because I take other two banknotes of five.	*With the left hand, He makes the gesture of putting a banknote on the table.*	*S: Ground* *G: Ground*

			 *Fig.11.5*	

In these two lines we can observe a merging point in Lorenzo's gesture. Whilst with his words he argues that iteratively it will be possible to obtain any quantity of money, he performs in rapid succession two different gestures. The first one is the gesture of figure 11.4. It is a repetition of LS which represents the steps of the iterative solution that Lorenzo is describing. Within this gesture, each step of the iteration is represented by an arc in the air which starts and finishes on the sheet. With this gesture is not referring to any banknotes, but to what it looks like an imaginary vertical line of numbers. The whole gesture can be interpreted as meta level since it is used to describe the iteration structuring Lorenzo's argumentation. Immediately after having concluded this gesture, Lorenzo makes a second gesture: he mimes the action of putting something (a banknote) down on the table (fig.11.5). The gesture is co-timed with the words: "I take other two banknotes of five". With this gesture Lorenzo is focusing on a specific step of his iteration, which indeed is obtained adding 10 (i.e., two banknotes valuing 5) to the previous interval. This gesture is another LS since it represents a step of the iteration, but this time it specifically refers to some banknotes therefore it is ground level. Looking at the two lines of the excerpts together we can see how Lorenzo firstly describes the iterative strategy with an IS formed by a meta level gesture (fig.11.4) and then he describes the detail of a single step of this iteration with a ground level gesture, representing the action of by adding two more banknotes valuing 5. In conclusion, focusing on the gestures semiotic set, in these two lines we can register the presence of a merging point.

In conclusion, the just defined terms can be summarised as it follows. Using the theoretical lens of the semiotic bundle and focusing on the ground or meta level semiotic production, different aspects can be registered:

- With a synchronic analysis of the bundle, it is possible to observe the ground or meta level of the different components of the bundle at a fixed time. By doing this it is possible to register moments in which a subject's semiotic production contains simultaneously ground and meta level elements, within different semiotic sets (level's mismatch) or within the same semiotic set (merging point). The presence of a merging point can also be observed by analysing two or more signs produced in a rapid succession (thus, formally, with a diachronic analysis conducted in a short interval of time).

- With a diachronic analysis of the bundle, it is possible to observe the development over the time of the two levels semiotic production, and to register the level transitions (and between them, the eventual liftings and projections).

Finally let us observe that even if liftings/projections, and mismatch/merging points are highlighted by different analyses of the semiotic bundle, however it is not to be excluded that the phenomena that they describe can be intertwined. In particular, for instance, it could be that in an episode in which we register a lifting or a projection, we can observe punctually the presence of level's mismatch or merging points which contribute themselves to the transition between one level to the other in the subject's semiotic production.

To present in a more condensed form the following analyses involving quite long excerpts I will use a particular diagram. The diagram represents the temporal development of the excerpt, modelling and providing an overview on the subject's semiotic production with reference to the ground and meta level and to the eventual presence of liftings or projections as well as level's mismatches or merging points.

In this diagram, the ground and meta levels are represented as two parallel horizontal lines. Every time a spoken utterance (S), a gesture (G) or an inscription (I) is produced by the subject it is indicated in the diagram with a capital letter put on one of the two lines considering whether it was interpreted as ground or meta level. The excerpt's lines are presented from left to right, with reference to the transcript's numeration. The diachronic and synchronic analysis of the bundle can be read by observing respectively the horizontal or vertical development of the diagram.

The following table contains the legend of the used symbols.

S, **G**, **I**	Respectively, a spoken utterance, a gesture, or an inscription.
↗ ↘	Respectively, a lifting or a projection.
$\boxed{\begin{matrix}S\\ \\G\end{matrix}}$	A level's mismatch (in this example between the semiotic sets of Speech and Gesture).
$\vee$ **G** $\wedge$	A merging point (in the example, relatively to the semiotic set of Gesture).

Table 11.2. Legend of symbols used in the diagrams for the following excerpts.

For example, figures 11.6 and 11.7 contain the diagrams corresponding to the analyses of Guido's and of Lapo's episodes (excerpts 11.1 and 11.2) of above.

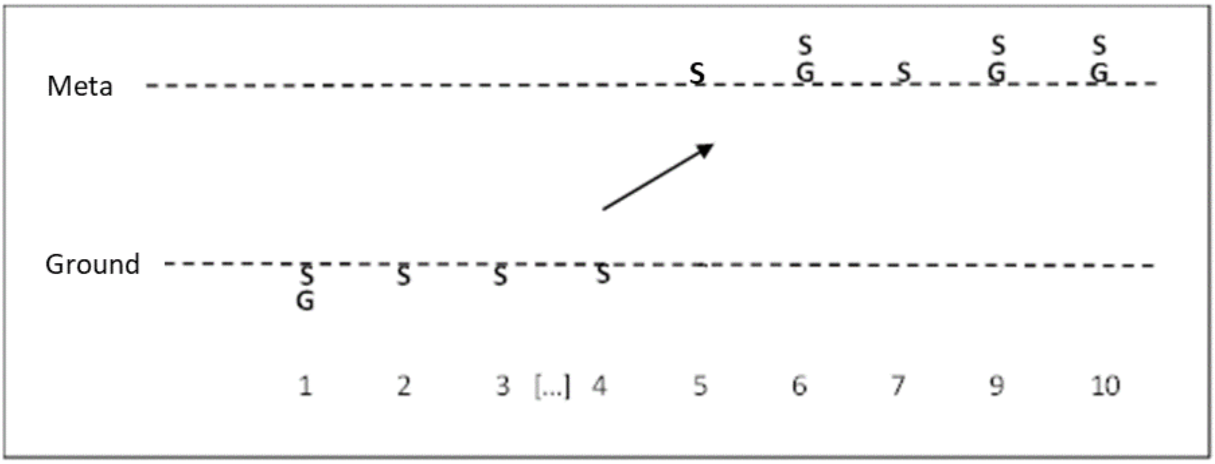


Figure 11.6. Diagram for the analysis for the excerpt 11.1.

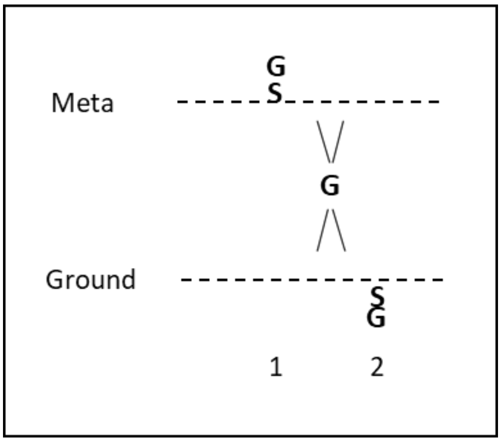


Figure 11.7. Diagram for the analysis for the excerpt 11.2.

11.1.5 Two paradigmatic examples: Lifting and projection of a recursive argumentation

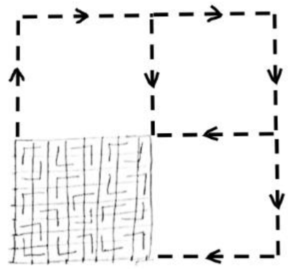
In this paragraph I will show two examples in which it is possible to observe some transitions between levels. The two episodes, if analysed in their complete development, show two processes which are, in a certain sense, opposite. In the first one it will be possible to see how a recursive argumentation, which starts concretely from the sheet involving the objects of the problem (ground level), is then lifted on the meta level. In the second one we will observe, instead, how a recursive argumentation, firstly expressed on the meta level, is then projected on the ground level involving the objects of the problem.

Both of the following two excerpts refer to the chessboard problem. The two students, respectively Giuditta and Silvio, conjectures that it is always possible to tessellate a 2^nx2^n chessboard except for one little square. The interviewer asks to the students to justify their conjecture. The two examples show how differently Giuditta and Silvio structure their argumentation.

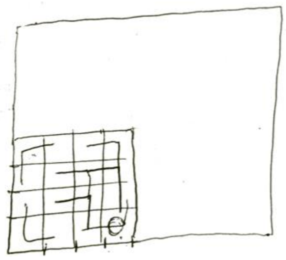
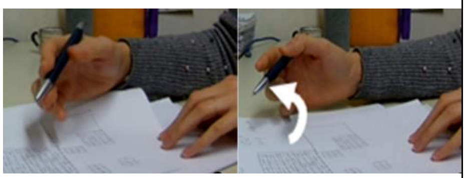
Part of the following excerpts have already been presented with a fine-grained analysis focusing on LS and IS in the previous chapter, respectively in the excerpt 10.1 for Giuditta and in 10.8 for Silvio. The two excerpts are reported here ex-novo, with a different line enumeration.

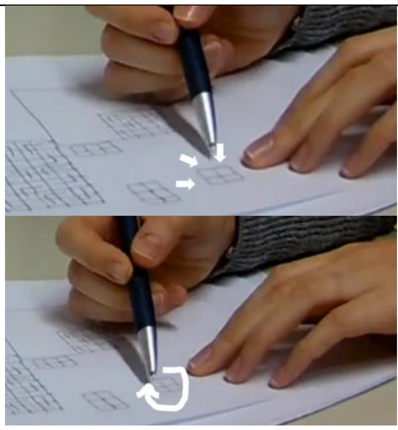
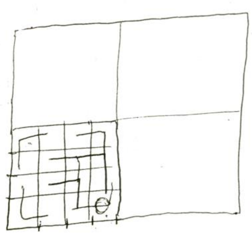
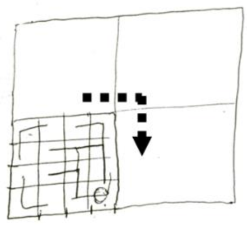
Excerpt 11.3

Giuditta has just found a tessellation for the 8x8 chessboard with the related drawing. The interviewer asks to consider the 16x16 chessboard.

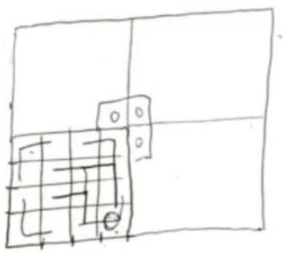
	Who	Speech	Gesture and inscription	Level
1	**G**	Sixteen by sixteen… But then I have another three of these squares.	*(a) [Sixteen by sixteen]. With her left middle finger and the tip of the pen in the right hand she points to two vertices of the 8x8 chessboard drawing.* *(b) [three of these]. She keeps her left middle finger on the vertex, and with the pen in the right hand she indicates respectively to the right, upper right, and above the drawing of the 8x8 chessboard.*	*S: Ground* *Ga: Ground* *Gb: Ground*
2	**G**	Here.	*Pointing with the left hand to the drawing of the 8x8 chessboard, she follows with the pen (without marking) the perimeter of 3 squares.*  *Fig.11.8*	*S: Ground* *G: Ground*
3	**I**	Ok.		
4	**G**	And then there would be left out one, one, one and one.	*[one, one, one and one]. She points to the drawing of the 8x8 chessboard on the sheet and to three points corresponding to the three other squares she just represented with the movement of the pen.*	*S: Ground* *G: Ground*
5	**I**	Ok.		
6	**G**	And so, I would think to put three of them together, somehow. And then, there would always be one left out?	*She points to the sheet where the inscription of above is.*	*S: Ground* *G: Ground*

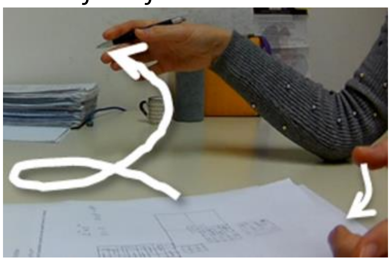
Giuditta abandons the case 16x16 and she tries to tile the 2x2, the 4x4 and the 1x1 chessboards, finding in every case a tessellation which leaves out one little square only. Giuditta then claims to be “convinced enough” of the fact that it is always possible to entirely tile a 2^nx2^n chessboard except for one little square. The interviewer asks to Giuditta to try to justify her conjecture.

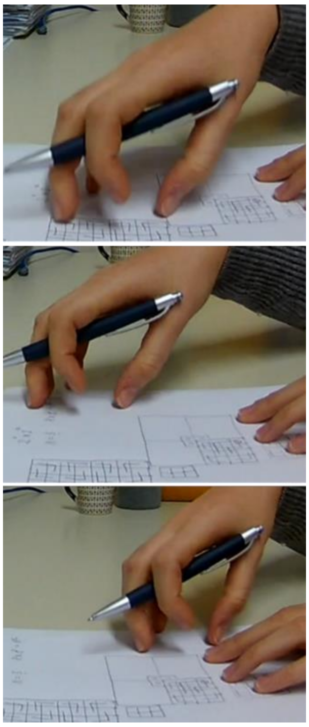
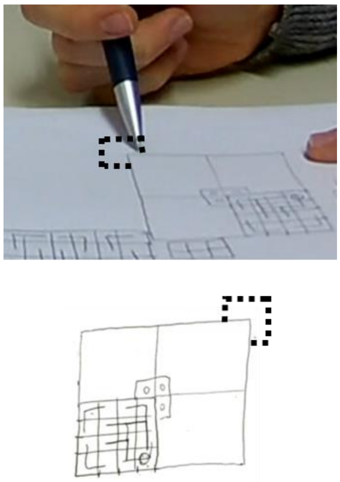
[…]				
7	**G**		*Starting from the just made drawing of the tiled 4x4 chessboard, she extends two sides and she completes the drawing obtaining a bigger square.*  *Fig.11.9*	*I: Ground*
8	**G**	So, what I was thinking was… That to come, to move forward from n equal to one, to n equal to two, practically, I have to put another three identical little squares.	*(a) [to move forward]. She makes an arc-shaped gesture in the air from left to right.*  *Fig.11.10* *(b) [I have to put]. With the pen, kept in the right hand, she points specifically to three squares of the drawing of the 2x2 chessboard, while it is pointed by her left middle finger. Then she rotates the pen around the inscription as to circle it.*	*S: Ground/meta* *Ga: Meta* *Gb: Ground* *Merging Point (Speech)* *Merging Point (Gesture)*

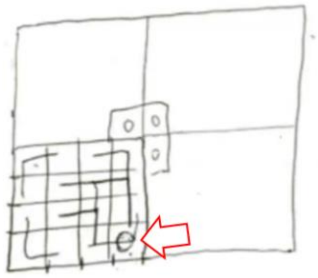
			 *Fig.11.11*	
9	**G**		*She draws two lines on the drawing of line 7 obtaining the following drawing.*  *Fig.11.12*	*I: Ground*
10	**I**	Ok.		
11	**G**	And so, if I analysed them separately, it will be left out one in each of them…	*She points to the centre of the three blank squares of the inscription of line 9.*	*S: Ground* *G: Ground*
12	**I**	Ok. But the ones remaining out will be separated. How can you do it?		
13	**G**	Eh, I could… isolate the separated one here in the centre and so that I can do this.	*Without marking with the pen, she draws an L-shaped inscription on the drawing of line 9, in the centre.*  *Fig.11.13*	*S: Ground* *G: Ground*
14	**I**	Ok.		

After this Giuditta produces a new drawing representing a 4x4 chessboard and she creates a tessellation which covers it completely except for the little square corresponding to the bottom left corner.

[…]				
15	**G**	Ok, but this is not a proof of the fact that I can isolate it in a corner.		*S: Meta*
16	**G**	But I can do the same in every corner. And so, If I isolate…	*She points to the inscription of line 9.*	*S: Ground* *G: Ground*
17	**I**	You can mark the isolated ones, only.		
18	**G**	This one, this one, and this one.	*She draws three little circles in the centre of the drawing, one for each blank square.*	*S: Ground* *I: Ground*
19	**I**	Yes		
20	**G**	Then we said that the rest… we can do it.	*She draws a circle in the air all around the drawing.*	*S: Ground* *G: Ground*
21	**I**	Yes.		
22	**G**	And this becomes a little block of three, an L of three.	*She draws the edge of an L-shaped tile in the centre of the drawing, covering the three circles drawn in line 18, obtaining the following inscription.*  *Fig.11.14*	*S: Ground* *I: Ground*
23	**I**	Ok.		
24	**G**	And still one is left out.		*S: Ground*

25	**I**	Ok		
26	**G**	And this, I can do it in general…	*(a) [this]. With the pen in the right hand, she points to the inscription of line 22.*  *Fig.11.15* *(b) [I can do it in general]. Starting from the sheet, with the pen in the right hand she makes a spiral movement in the air that widens as the right hand rises and concludes with spreading both the hands. Then she stays still in this in the final position for a few moments.*  *Fig.11.16*	*S: Meta* *Ga: Ground* *Gb: Ground/Meta* *Merging Point (Gesture)*
27	**G**	Because to pass from… moving forward… from n equal to one to n equal to two…	*[moving forward]. With the pen she rapidly draws an arc in the air. The gesture is similar to the gesture of line 8.*	*S: Meta* *G: Meta*
28	**G**	I always have powers of two. And so, also here… the next one will be… constructing other three big square like this one.	*(a) [will be…]. She draws without marking some lines as extending the drawing of line 22 to a bigger square. The gesture is similar to the one of line 2.* *(b) [constructing]. She makes the gesture of positioning three squares respectively on the top, on the top right, and on the right of the inscription.*	*S: Ground* *Ga: Ground* *Gb: Ground*

			 *Fig.11.17*	
29	**G**	And so, again… in these three I would have one as a rest, and this rest, I can isolate it again in the corners…	*(a) [in these three]. She points to the sheet where she just mimes to put three squares.* *(b) [I can isolate it]. With the pen, without marking, she draws the edge of an L shaped tile in correspondence of the top right corner of the inscription representing the 8x8 chessboard.*  *Fig.11.18. The dotted line indicates the trace of the pen on the sheet.*	*S: Ground* *Ga: Ground* *Gb: Ground*

30	**G**	And so, the initial one remains out.	*With the pen she points to the square marked as blank in the drawing of the 4x4 chessboard within the bigger inscription.*  *Fig.11.19. The red arrow indicated where the pen is pointing to*	*S: Ground* *G: Ground*

<u>Analysis</u>

The excerpt starts with Guiditta who individuates a possible connection between the tessellation of the 8x8 chessboard and the one of the 16x16 chessboard (lines 1-6). From the successive exploration of the tessellations for the chessboards with dimensions 1x1, 2x2, and 4x4, Giuditta formulates the conjecture that every chessboard of the problem can be completely tiled except for one little square. Then, after the request of the interviewer, Giuditta starts construing an argumentation supporting the conjecture. First of all, in lines 7-9, she further explicates the link between two consecutive chessboards (in this case the 4x4 and the 8x8 chessboards), showing how to construct the bigger one using four smaller chessboards. Differently than in the first part of the excerpt, however, now her explanation is slightly more general. In particular she refers to the link between the 4x4 and the 8x8 chessboard in line 7, however in line 8 the considered link is the one between the 2x2 and the 4x4 chessboards. Giuditta is therefore representing a series of different LS, involving chessboards with different sizes. In a certain sense she is generalising the link between two consecutive chessboards. Until now Giuditta's semiotic production was ground level, however in line 8 we can register the presence of some meta level aspects. Let us focus on her speech first. She says: "to move forward from n equal to one, to n equal to two, practically, I have to put another three identical little squares". With this sentence she seems to refer simultaneously to the iterative structure of her argumentation (the words "to move forward" do not refer to the chessboards or to the tessellations, but to the steps of the iteration itself) and to the chessboards of the problem ("to put another three identical little squares"). Thus, her speech contains both meta and ground level elements. Completely analogous is what happens in Giuditta's gestures. Firstly, in the gesture (a) of line 8, she produces an arc shaped gesture co-timed with the words "to move forward from n equal to one, to n equal to two". The gesture is meta level, representing the step from n=1 to n=2 of the iteration, metaphorically as a path in the air. Then, in the same line, she produces another gesture (b), which is ground level, since it deictically refers to the inscription representing the 2x2 chessboard, and it is used to show how three 1x1 chessboards can be used to extend the 1x1 into a 2x2 chessboard. In summary, in line 8, we register the presence of a merging point in both the speech and gesture semiotic sets.

In the successive lines (11-24), Giuditta explores the drawings she produced in order to find a way to connect the tessellation of the 4x4 chessboard with the one of the 8x8 chessboard. In this exploration she observes that if the square left out of the tessellation is in one of the corners of the chessboard, it is possible, in the successive chessboard, to cover three left blank squares with a new L shaped tile. Finally, in line 24 she produces an inscription to represent this situation, specifically a LS between the tessellation of the 4x4 chessboard and the one of the 8x8 chessboard. Giuditta's semiotic production is ground level in all this part of the excerpt, with the only exception of line 15. Here Giuditta, after having found a tessellation for a 4x4 chessboard which leaves out a square in a corner, states that "this is not a proof". With this comment, Giuditta clarifies that what she has found (a tessellation for a chessboard with the left out square in one of the corners) has not been proved in general, but only for 4x4 chessboard.

In line 26, an important passage happens: Giuditta generalises the link between the 4x4 chessboard and the 8x8 chessboard, with their tessellation. She says: "And this, I can do it in general" and she accompanies her words with a spiral gesture in the air which starts from the drawing of the 8x8 chessboard and that widens while the right hand rises and concludes with the two hand still over the sheet. A fine-grained analysis of this line, and in particular of this gesture, was presented in the excerpt 10.1 at the beginning of the chapter. Here it is sufficient to remember how, in this gesture, the LS corresponding to an arc-shaped movement around the drawing of a chessboard, representing the transformation of a chessboard in the successive one, has evolved: it becomes simultaneously a gesture of 3rd level in Krause's classification (i.e. a 'general' gesture) by detaching from the sheet and an IS in the moment in which the arc becomes a spiral representing the iterative transformation of a chessboard in all the subsequent ones. Finally, the hands still and high over the sheet with which the gesture concludes, contain the space in which the previous gestures were made and with this Giuditta seems to refer to the whole argumentation itself. We can re-read this line now, with reference to the ground and meta levels. The words "And this, I can do it in general" are meta level. In particular, the deictic "this" highlights how the entire process through which a tessellation of a chessboard is used to tessellate the subsequent chessboard has become a word ("this") accompanied by a deictic gesture pointing at the sheet. The spiral gesture, instead, starts as a ground level gesture explicitly referring to the chessboards but then becomes a meta level gesture, in the moment in which the hands stop over the sheet: the recursive argumentation has had a physical development (the enlarging of the chessboards) which Giuditta is containing with her hands. The gesture, in summary seem to refer both to the chessboards and tessellations, and to the recursive argumentation itself. Therefore, in this, we can register the presence of a merging point.

If we observe the development of Giuditta's semiotic production from line 1 to line 26, it is possible to observe a lifting in her discourse from the ground level to the meta level. Giuditta initially starts her argumentation on ground level (lines 1-6). Then, when she recognises that the link between two consecutive chessboards their tessellations could be generalisable to every couple of consecutive chessboards starts referring to the iterative structure of her argumentation (line 8, line 26). In these moments Giuditta lifts her argumentation on the meta level, as highlighted by the presence of the merging points in these two lines, representing the structure of the argumentation itself.

In the last lines of the excerpt (27-30), Giuditta projects the argumentation of the ground level again. Firstly, the line 27 contains another verbal expression which can be interpreted as meta level ("moving forward… from n equal to one to n equal to two") co-time with a meta level gesture (an arc in the air similar to the one of line 2). Even if this expression refers only to one specific step of the iteration (from n=1 to n=2), Giuditta seems to refer to a generic step, in fact in the subsequent lines the sentence continues by referring to two other cases (specifically from n=3 and n=4). Then Giuditta projects her discourse on the ground level. Indeed, she repeats a single step of the recursive argumentation to construct a tessellation for the 16x16 chessboard using four 8x8 chessboards with a tessellation which leaves out only one square in one of the corners. This step is presented by Giuditta in detail in lines 28-30. In particular, using the inscriptions present on the sheet and several gestures she recreates the LS previously produced in line 22 to connect the 4x4 chessboard with the 8x8 chessboard, this time to link the 8x8 chessboard with the 16x16 chessboard. This concludes Giuditta's argumentation.

In conclusion, Giuditta's excerpt showed how she produced an argumentation that starts with a ground level semiotic production, how she recognised its iterative structure and then how she referred to it with a meta level semiotic production. Focusing on this development, therefore, the whole episode showed a lifting in Giuditta's semiotic production which shifted her argumentation form the ground level to the meta level.

The figure 11.20 contains the diagram modelling the just presented analysis.

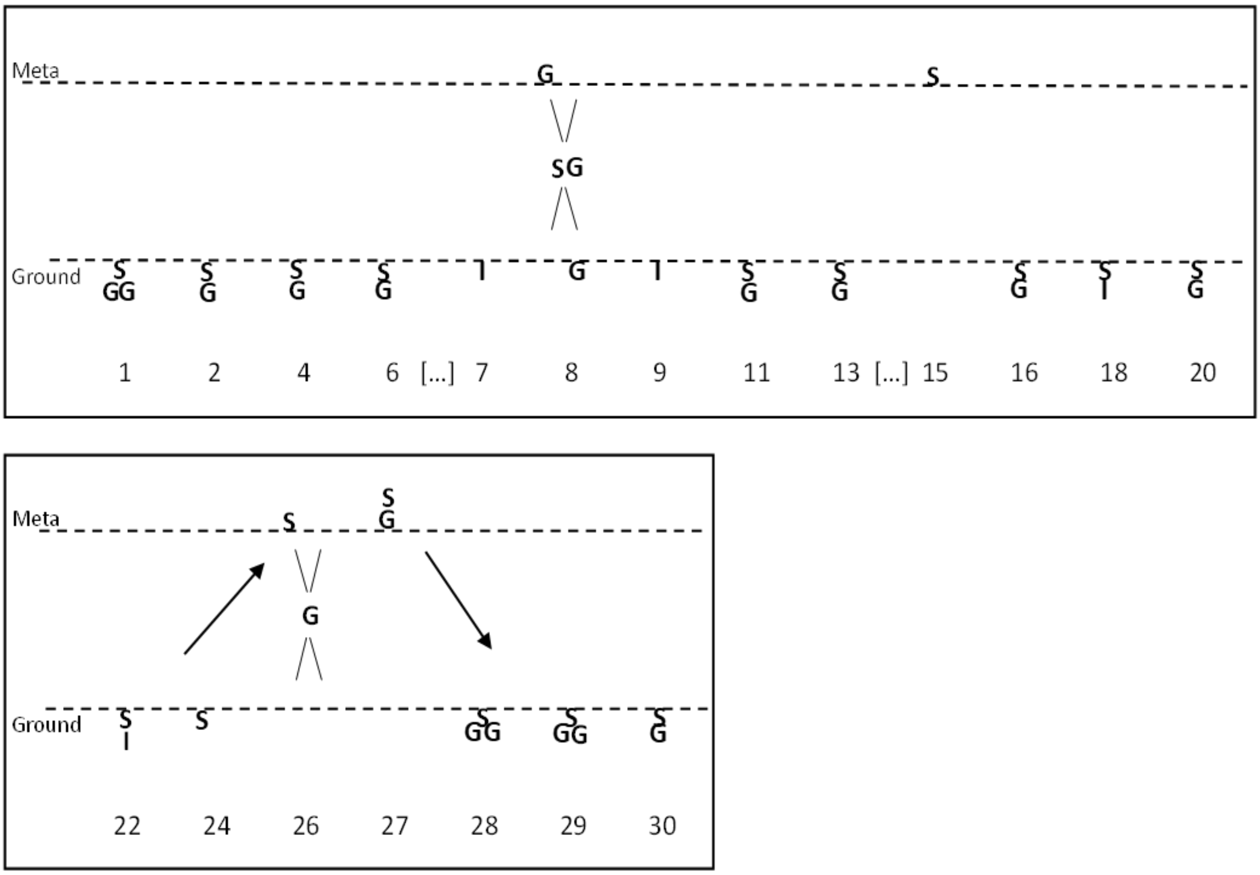
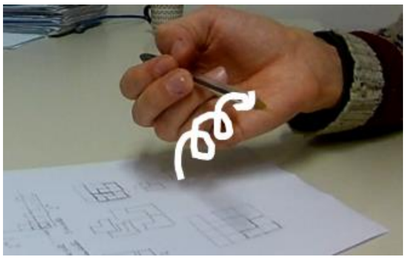
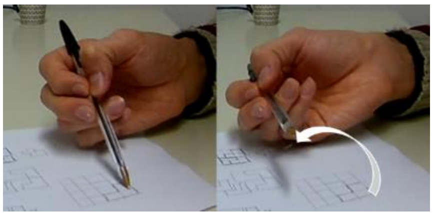
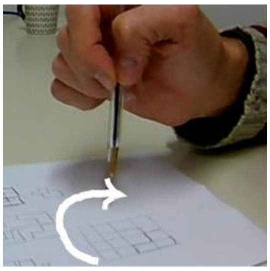


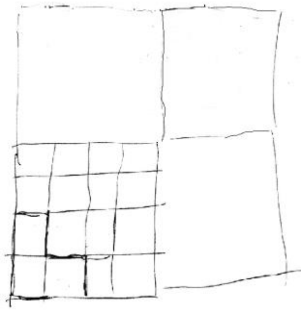
Figure 11.20. Diagram for the analysis for the excerpt 11.3.

Excerpt 11.4

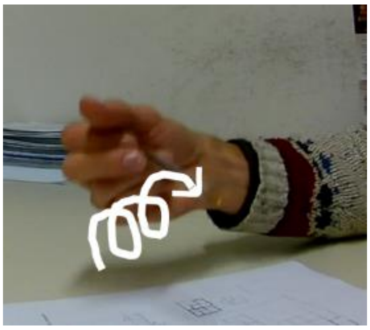
During the first fifteen minutes, Silvio has explored the problem, finding that for the chessboards corresponding to n=0, n=1, and n=2, it is possible to create a tessellation which leaves out only one little square. The cases have been solved separately. Then Silvio hypothesises that, perhaps, the same thing happens for every chessboard of the problem ("I would say that, maybe, I can always find the way for leaving out only one square"). The interviewer asks Silvio to try to justify this statement.

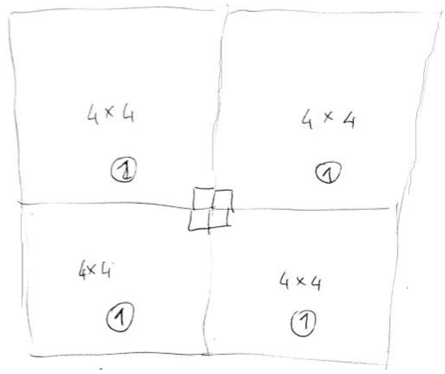
	Who	Speech	Gesture and inscription	Level
1	**S**	How could I do? … I mean … Ah, yes. Once you have	*Keeping the pen in his right hand, he rapidly draws some circles in the air, while moving backwards the hand.*	*S: Meta* *G: Meta*

		established… I mean, every case is reconducted to the previous one, right?	 *Fig.11.21*	
2	**I**	Ok		
3	**S**	In the sense that, if I go to the next case, so two to the third, basically this square here is created again three times.	*(a) [next case]. He draws an arc in the air moving the hand forwards starting from the sheet.*  *Fig.11.22* *(b) [this square here]. He points to the inscription of the 4x4 chessboard and then he draws an arc in the air over the sheet around the inscription from the top left corner to the right side of the square.*  *Fig.11.23*	*S: Ground/Meta* *Ga: Meta* *Gb: Ground* *Merging Point (Speech)* *Merging Point (Gesture)*
4	**I**	Other three times		
5	**S**	Exactly, in this sense… I have another square here, another square here and another here.	*He extends the inscription of the 4x4 chessboard drawing the sides of three squares constructed respectively from the top side of square representing the 4x4 chessboard, from the top right vertex and from the right side.*	*S: Ground* *I: Ground*

			 *Fig.11. 24*	
6	**I**	You can turn the sheet if you prefer.		
7	**S**	Yes, better. So, in practice, the case two to the third…	*He turns the sheet on the other face, which is blank, then he draws a square on the sheet.*	*S: Ground/Meta* *I: Ground* *Merging Point (Speech)*
8	**S**	So, we agree that all of them are reconducted to a previous case.	*He makes a gesture like the one of line 1, but this time higher in the air.*  *Fig.11.25*	*S: Meta* *G: Meta*
9	**I**	Ok		
10	**S**	To the immediately previous case, in practice. I could do something like… but it is everything a bit chitchatted[93], I don't know if I could formalise it in this moment.	*[previous]. He turns the hand backward high in the air.*	*S: Meta* *G: Meta*
11	**I**	It's all right.		

[93] "Chiacchierato" in Italian.

12	**S**	I mean… I think that one could do something like by induction.	*He rapidly draws some circles in the air as in line 1.*  *Fig.11.26*	*S: Meta* *G: Meta*
13	**S**	Because, since in the case zero I have only one little square, and it remains out…	*With the pen he touches the sheet where he wrote* 'n=0 → 0 tiles' *during the first part of the interview*.	*S: Ground/Meta* *G: Ground* *Merging Point (Speech)*
14	**I**	Ok.		
15	**S**	Then, let's say, in a sequential way… this little square will always remain out.	*(a) [in a sequential way]. With the pointing finger of his right hand, he draws several circles in the air whilst moving up his hand. The gesture is wide in the air and lasts for almost 3 seconds.*  *Fig.11.27* *(b) [this]. He points to the inscription of the 4x4 chessboard.*	*S: Ground/Meta* *Ga: Meta* *Gb: Ground* *Merging Point (Speech)* *Merging Point (Gesture)*
16	**S**	Yes… but it must be like this. Surely.		*S: Meta*
17	**S**	In the best option only one square remains uncovered.	*He extends the drawing of the square made at line 7 adding three other squares attached to it.*	*S: Ground* *I: Ground*

18	**S**	Because if I… let's say that this is the four by four, this is the four by four… and this is the four by four.	*He writes '4x4' in the centre of the four squares in the inscription of the previous line.*	*S: Ground* *I: Ground*
19	**I**	Ok		
20	**S**	If here it remains one out, and here one, and here one, and here one… In the best option, I have the empty one here, the empty one here, the empty one here… and the other here.	*(a) [one out… one… one… one]. He writes '1' and circle it, in each of the four squares of the inscription.* *(b) [one here… one here… one here… other here]. He draws four little squares, one for each of the previous squares, in correspondence to the centre whole inscription.* *He obtains the following drawing.*  *Fig.11.28*	*S: Ground* *Ia: Ground* *Ib: Ground*
21	**S**	And therefore, in any case, only one remains out.	*He draws rapidly, without marking the sheet, an 'L' over three little central squares of the inscription of line 20.*	*S: Ground* *G: Ground*
22	**I**	Because you put another tile there?	*He points at the centre of the inscription.*	
23	**S**	Exactly… let's say… the k plus one tile, and thus one remains necessarily empty.	*[let's say]. He rapidly draws some circles in the air rotating forward the right hand.*  *Fig.11.29*	*S: Ground* *G: Meta* *Level Mismatch*

<table>
<tr><td>24</td><td>I</td><td>Ok</td><td></td><td></td></tr>
<tr><td>25</td><td>S</td><td>Beyond the fact that I am not sure that I can have this thing, eh. Even though, this situation here, a bit like a spiral, bodes well, let’s say.</td><td>(a) [this thing]. He points to the centre of the inscription of line 5.

(b) He points to the inscription of line 5 representing the extended 4x4 chessboard. Then he draws in the air the shape of a spiral which enlarges whilst going up.

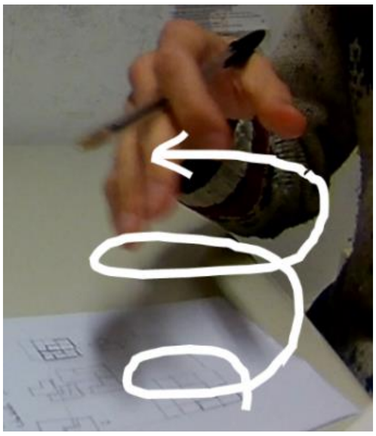



Fig.11.30</td><td>S: Ground/Meta

Ga: Ground
Gb: Ground

Merging Point (Speech)</td></tr>
</table>

In the remaining part of the interview, Silvio tries, successfully, to find a tessellation of the 4x4 chessboard which leaves out one square in one of the corners, so to recreate the situation represented in line 20.

<u>Analysis</u>

The whole episode has already been analysed in the excerpt 10.8 of the previous chapter focusing on linking and iterations sign and how their use and production supports Silvio in plan and structure a recursive argumentation. In the following analysis the focus will be on the two levels of semiotic production and to their dialectic.

In the first line, Silvio highlights the presence of a possible connection between two cases of the problem (the chessboards to tile) which he describes in general terms. His speech (“every case is reconducted to the previous one”) and his gesture are meta level. With these, in fact, Silvio refers to a possible heuristic to solve the problem (i.e., to link every case with the previous one). In line 3, Silvio starts justifying his previous statement, and in doing this his speech refers both to the iterative structure of the heuristic (“If I go to the next case”) and the chessboards (“this square here is created again three times”). Thus, we can register a merging point in Silvio’s speech in this line. Analogously, the gestural production in this line seems to have both meta and ground level components: the gesture (a) is a meta level LS (an arc in the air, representing a step of an iteration), whilst the gesture (b) is a ground level LS (an arc around the drawing of the 4x4 chessboard, representing the extension of the chessboard into an 8x8 chessboard).

In the successive lines (5-7), Silvio shows again the connection between two chessboards of consecutive sizes, but this time his discourse is fully ground level. Silvio draws some squares on the sheet and with his speech and gestures he refers to them. In line 7, however, we can see a trace of the meta level in Silvio’s speech in the moment in which he says, “The case two to the third”. The reference to a “case” of the problem can be seen as meta level. To summarise, these first lines (from 1 to 7) contain what we called a projection in Silvio's semiotic production. In particular Silvio has firstly described his idea of a link between two chessboards with a meta level semiotic production which he has then projected onto the ground level. The projection is further highlighted by Silvio words "in the sense that" in line 3 and "in this sense" in line 5 which anticipate his ground level discourse.

The lines 8-12 contain again a meta level semiotic production. Silvio explains that since all the cases "are reconducted to a previous case" (line 8), then "one could do something like by induction" (line 12). This is a meta level utterance. Moreover, he makes some meta level gestures as well. In particular in line 8 and 12 he produces two meta level IS, each formed by a rotating gesture in the air, and in line 10 a meta level LS representing the connection between two cases of the problem with an arc-shaped path in the air.

The sentence at line 12 ("I think that one could do something like by induction") opens to a new projection in Silvio's discourse which develops in the following lines. In lines 13 and 15 we register two merging points in Silvio's speech while he is explaining what he meant for "something like by induction". In particular in line 13, while saying "case zero" he points to the inscription 'n=0→0 tiles' previously made when exploring the 1x1 chessboard. He is explicitly referring to the "case zero" of a recursive argumentation (meta level) and, simultaneously, to the chessboard corresponding to this case (ground level). In the same way, in line 15, Silvio says "in a sequential way… this little square will always remain out", referring both to the structure of the argumentation he is constructing ("in a sequential way" – meta level) and to the tessellation of the chessboards ("this little square" – ground level). The two gestures in line 15 are respectively meta level, the IS representing the "sequential way" as a series of arcs in the air (a), and ground level, the pointing to the drawing of the 4x4 chessboard (b). Considering the whole semiotic bundle in this line, thus, we can register the presence of an IS which contains both reference to the chessboard (ground level) and to the structure of the argumentation itself (meta level). The analysis of this specific IS has been reported in Nannini (2022) where this sign has been called a *hybrid level* IS since it refers both to some objects of the problem and to the structure of a recursive argumentation itself. In the subsequent lines (17-21) Silvio shows in detail the link between the tessellations of two consecutive chessboards (the 4x4 with the 8x8) with a ground level semiotic production in all of the three semiotic sets. Specifically, he produces the inscription of line 20 together with the gesture of line 21 representing the action of putting a tile in the centre of this inscription. This is a ground level LS representing the tessellation of an 8x8 chessboard using the tessellation of four 4x4 chessboards. In these lines, to summarise, Silvio has projected on the ground level (referring to chessboards and tiles) his previous meta level discourse of lines 8-12 ("all of them are reconducted to a previous case", "to the immediately previous case", "something like by induction").

The line 23 concludes Silvio's argumentation. His discourse is still on the ground level, however it contains an interesting level's mismatch: while Silvio speaks about positioning a tile in the centre of the chessboard, he makes a rotating gesture in the air similar to those produced in lines 1, 8, 12 and 15. The gesture is different from the previous one since the rotation of the hand is forward instead of backward, which is coherent with the fact that Silvio is representing to a forward step in the iteration (from the case n=3 to n=4), however analogously to the previous ones, it refers to the iterative structure of his argumentation, thus it is meta level. The presence of this gesture that repeats the previous ones, forms a *catchment* in McNeill (2005) terms. The repetition of such gestures seems to suggest that during the whole episode Silvio is thinking to the structure of a particular argumentation ("something like by induction"), to which he refers with his gestures, and which he is trying to construct for the chessboard problem.

The episode concludes in line 25 with Silvio who says not to be fully convinced that what he has just described (a recursive argumentation) could be obtained for this problem, but that he thinks it could work because of “this situation here, a bit like a spiral”. The semiotic production of this last line is interesting. He makes two gestures which are both ground level, since they refer to the chessboards: (a) he points to the inscription produced in line 5 representing a 4x4 chessboard extended to an 8x8 chessboard and then (b) he makes a spiral gesture over this inscription as representing the increasing dimensions of the chessboards involved in the iteration. However, in his speech, we can register both some ground and meta level elements. He refers to the chessboards (ground level) when saying “a bit like a spiral”. Instead, the words “I can have this thing”, where ‘thing’ refers to the tessellation for a chessboard which leaves out one square in one of the corners, and “this situation here”, where ‘situation’ refers to the extension of a chessboard in the following one, can be interpreted as meta level. With these, in fact, Silvio seems to be referring to the problem itself and to a possible heuristic for solving it: there might be a “thing” which would let Silvio conclude the problem and “a situation” that might help to obtain the “thing”. In conclusion these last two sentences of line 25 contain another merging point.

In all this episode Silvio has described, both in ground and meta level terms, a possible recursive argumentation to prove his initial conjecture, showing a step of it (from the 4x4 to the 8x8 chessboard). It is interesting to observe, however, that Silvio, until now, has not found yet a way to tessellate a 4x4 chessboard so that the remaining square occupies one of the corners. In other terms, Silvio has planned and structured his argumentation which however still needs to be explicitly constructed.

The graph in figure 11.31 models the just presented analysis. The two projections, the several merging points and the level’s mismatch which have been registered seem to highlight an interesting aspect of Silvio’s reasoning. It seems that Silvio has in his mind a sort of ‘a priori’ structure for his (recursive) argumentation and that he tries to construct such an argumentation involving the objects of the problem. The whole excerpt, therefore, can be seen as an example of a projection of Silvio’s argumentation from meta to ground level.

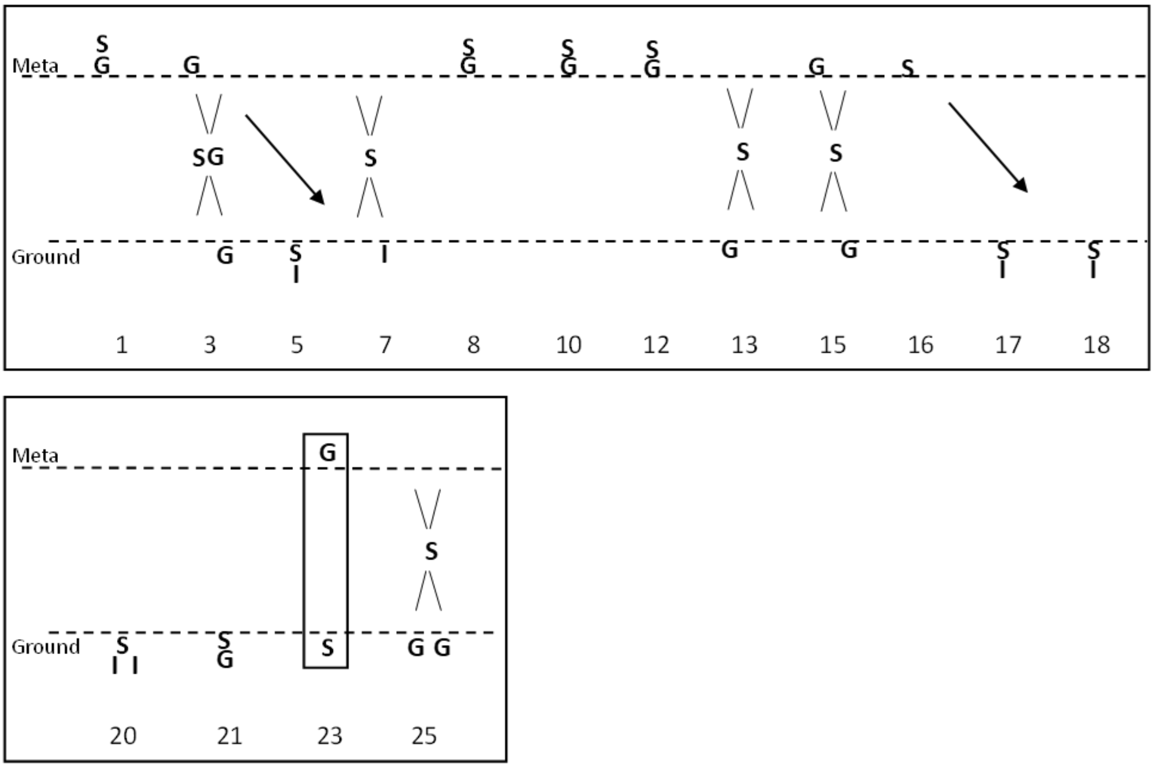


Figure 11.31. Diagram for the analysis for the excerpt 11.4.

Giuditta’s and Silvio’s episodes have shown two possible level transitions (from ground to meta and opposite) that might involve a subject’s semiotic production during the resolution of a problem. The two episodes have been presented together since they have some important common characteristics: the interviewed subjects have a similarly high level of experience in mathematics (and specifically in proving by induction), both the two excerpts involve to the same problem, and they refer to a moment in which both the subjects are constructing a recursive argumentation for the same conjecture. Considering this premises of symmetry between the two episodes, it is thus interesting to notice that Giuditta and Silvio shape their argumentation with two different processes. Giuditta explores some cases of the problem and, only after having explicitly found what is the link between them, she lifts her own argumentation from the ground to the meta level, recognising in it some general elements, such as its recursive structure. On the other side, Silvio, after having explored the problem, hypothesises that it can be solved by reducing every case to the previous one. This hypothesis is made before having explicitly found the link between to cases of the problem. Silvio starts describing the structure of a recursive argumentation referring to it with a meta level semiotic production and then tries to explicitly construct such argumentation referring to the problem’s objects, thus with a ground level semiotic production. The conclusion of the two episodes is interesting. Both Giuditta and Silvio conclude their argumentation with a spiral gesture starting from the sheet and widening while the hand moves up (line 26 for Giuditta, line 25 for Silvio). The same gesture, however, as the analyses have highlighted, has an extremely different story and genesis within the two episodes.

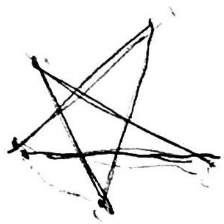
The two just presented experts have also shown how the two levels of semiotic production are extremely ‘permeable’ for the two students. In fact, it was possible to register a series of different liftings or projections, merging points and level’s mismatches. In other terms in these two examples, we registered a rich dialectic between the two levels with continuous ‘up and down movements’ in the students’ discourse. The two students, in particular, during their problem exploration seem to combine a ground level semiotic production with a meta level semiotic production and this combination has an effect on the development of their argumentation. In the following paragraph I will further discuss this point, focusing on possible roles that the meta level semiotic production might have for students during the resolution of a problem.

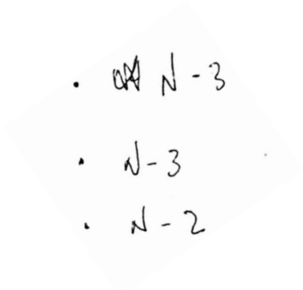
Before this, however, I wish to present another example taken from the interviews, this time from a less expert student (first year undergraduate student in physics), in which these transitions between levels are not present. Specifically, the student’s exploration of the problem is accompanied by a semiotic production which is almost uniquely ground level. As we will see, after a series of different calculations, the student will say to be “stuck”.

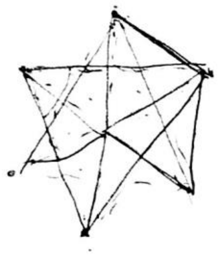
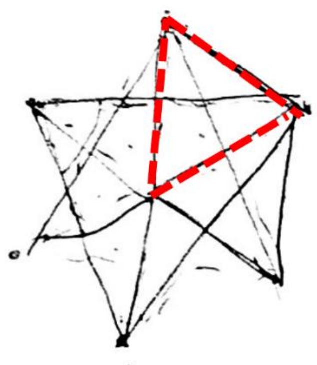
Excerpt 11.5

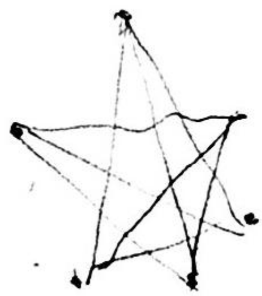
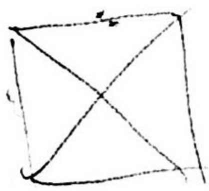
Tommaso is dealing with the diagonals problem. The excerpt corresponds to the first seven minutes of the problem resolution.

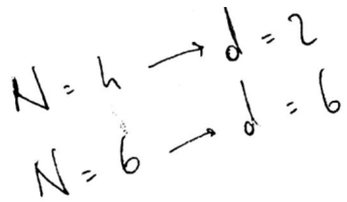
	Who	Speech	Gesture and inscription	Level
1	**T**	Diagonals… every diagonal is… I mean if I have five vertices, for instance…	*He draws 5 points on the sheet representing the vertices of a pentagon.*	*S: Ground* *I: Ground*

2	**T**	The diagonal starting from a vertex is this one and this one, right?	*With the pen, ha traces without marking two diagonals starting from one vertex.*	*S: Ground* *G: Ground*
3	**I**	Yes. Sides are not diagonals.		
4	**T**	Mhm. So, for every vertex, let's say, I can link a number of vertexes equal to the total number of vertices minus these three.	*He keeps the pointing finger on one of the vertices and then he points to the two consecutive vertices.*	*S: Ground* *G: Ground*
5	**T**	That is, the vertex and the adjacent ones… For every vertex.	*With the pen he points to the vertex he was pointing in line 4 and then he points to the two consecutive vertices*	*S: Ground* *G: Ground*
6	**I**	Ok.		
7	**T**	Therefore, for every vertex… every vertex is linked to the total number of vertices minus three.	*He writes the following inscription: '1 vertex' then three lines and then 'N-3':*  *Fig.11.32*	*S: Ground* *I: Ground*
8	**I**	Ok		
9	**T**	Wait… there are the repetitions...	*He points to the drawing of the five vertices of line 1.*	*S: Ground* *G: Ground*
10	**T**	… [10 seconds]		
11	**T**	I could also do, with this I do this, with this I do this, with this I do this, with this I do this and with this I do this… done	*Starting from every vertex he iteratively draws a diagonal obtaining the following inscription.*  *Fig.11.33*	*S: Ground* *I: Ground*
12	**T**	No, this has five… if it has six… give me a moment	*He draws six points representing the vertices of a hexagon.*	*S: Ground* *I: Ground*

13	**I**	You have all the time you want.		
14	**T**	I need to visualize it, I do not… [5 seconds].		*S: Meta*
15	**T**	Then I do like this: one, two, tree, four, five, six. Done.	*For each of the six vertices drawn at line 12, he draws a single diagonal.*	*S: Ground* *I: Ground*
16	**T**	So it is n plus…	*He writes 'N+' at the right of the just made inscription.*	*S: Ground* *I: Ground*
17	**T**	For every vertex there is surely one diagonal. This particular diagonal.	*With the pen he points to one of the just drawn lines.*	*S: Ground* *G: Ground*
18	**T**	And then there is… two, three.	*He draws other three diagonals*	*S: Ground* *I: Ground*
19	**T**	So I do… ta, ta, ta…	*He moves the pen on three diagonals already drawn.*	*S: Ground* *I: Ground*
20	**T**	Let's try with seven…	*He draws seven new points on the sheet representing the vertices of a heptagon.*	*S: Ground* *I: Ground*
21	**T**	I can do… one, two, three, four… one, two… help… there are more here. Wait.	*Starting from one of the just drawn points he draws four lines corresponding from four diagonals, and then starting from a second point other two lines.*	*S: Ground* *I: Ground*
22	**T**	Ok, so, if I take a vertex, I trace… n minus three… n minus three?... For a vertex I trace n minus three, for the second vertex n minus three again, for the third N minus two…	*He writes the following inscription.*  *Fig.11.34*	*S: Ground* *I: Ground*
23	**T**	No, no… there must be a faster way…		*S: Meta*
24	**T**		*With the pen he marks again for several times the diagonals of the inscription for the pentagon made at line 11.*	*I: Ground*
25	**T**	One, two … five.	*He points to the five just marked diagonals.*	*S: Ground* *G: Ground*

26	**T**	One, two, three, four, five… One, two…	*He marks again and draws other lines on the inscription for the hexagon made at line 15.*	*S: Ground* *I: Ground*
27	**T**		*With the pen he follows some of the lines of the inscription for the heptagon.*	*G: Ground*
28	**T**	Mhm. If n is even… if n is even…	*He draws other lines on the inscription for the hexagon, obtaining the following drawing.*  *Fig.11.35*	*S: Ground* *I: Ground*
29	**T**	One, two…	*With the pen, he touches two of the just drawn lines.*	*S: Ground* *G: Ground*
30	**T**	No…What have I done?		*S: Meta*
31	**I**	Wait in this drawing I don't understand which one are the diagonals and which one the sides.		
32	**T**	Yes, me too…		*S: Meta*
33	**T**	[5 seconds]. Let's see… if I decompose it in triangles.	*Following some lines of the inscription made at line 28, he follows the perimeter of a triangle whose vertices are two vertices of the hexagon and the intersection point of two diagonals.*  *Fig.11.36*	*S: Ground* *G: Ground*

34	**I**	Ok.		
35	**T**	So, we say that if n is even.	*He draws six new points on the sheet corresponding to the vertices of a hexagon.*	*S: Ground* *I: Ground*
36	**I**	Those are six, right?		
37	**T**	Yes, they are six. So… one, two, three… plus three, plus two, Plus three again. So… one, two, three, then… this is done, this is missing, this is missing, so they are two… plus this one.	*Starting from one of the vertices he draws three diagonals. Then from a second vertex he draws three new diagonals. From a third vertex he draws two diagonals. From a fourth vertex he draws a new diagonal.*  *Fig.11.37*	*S: Ground* *I: Ground*
38	**T**	So again, these are six.		*S: Ground*
39	**T**	But for a square they are two…	He draws a square with the two diagonals.  *Fig.11.38*	*S: Ground* *I: Ground*
40	**I**	Yes, for a square they are two.		
41	**T**		He writes ‘N=4 →’	*I: Ground*
42	**I**	You can call d the number of diagonals.		
43	**T**		He continues the inscription writing ‘d=2’	

44	**T**	… We said six.	He writes ‘N=6 → d=6’ underneath the previous inscription obtaining the following one.  *Fig.11.39*	*S: Ground* *I: Ground*
45	**T**	Oh mother… No, I cannot make it… I am one hundred percent sure… Never.		*S: Meta*
46	**I**	Why do you say this?		
47	**T**	Because I am getting stuck in things which are too difficult…		*S: Meta*

The interviewer decides to intervene and to guide the student toward a possible solution for the problem, not presented here, that is obtained starting from what Tommaso wrote at line 22.

<u>Analysis</u>

In this excerpt Tommaso explores the problem drawing a series of different polygons and counting their diagonals in different ways. None of them, however, is further developed by Tommaso to solve the problem. Firstly (line 1-7) he explores the case of a pentagon and generalises the example concluding that every vertex can be linked with N-3 other vertices to form a diagonal. Then (line 9) he realises that counting the diagonals in this way there will include some repetitions. Then he seems to abandon this strategy and he starts exploring again the problem in the case of a pentagon counting the diagonals in a different way, obtaining the star-shaped inscription of line 11 (note that in this inscription not all the diagonals have been considered). Then he does the same with a drawing of a hexagon (lines 12-19) and of a heptagon (lines 20-21). In line 22 he observes that counting the diagonals starting from a first vertex, one obtains N-3 diagonals for the first two vertices and N-2 for the third one. However, Tommaso decides to abandon this counting strategy (“There must be a faster way”). In lines 24-29 he explores again the drawings of the pentagon, hexagon, and heptagon drawing other lines and counting the diagonals several times. At this point (line 30) Tommaso seems to have lost the trace of what he was doing, as highlighted by the question he poses to himself (“What have I done?”) and by the answer he gives to the interviewer who says not understand what the lines made in the drawings represent (“Yes, mee too”). In line 33 Tommaso counts again the diagonals of the hexagon, but this time with reference to a possible division of the polygon in triangles. Finally in lines 37-44 he tries to find the relation between the number n of vertexes and number d of diagonal when n is even. Firstly, he draws new hexagon with all the diagonals. He correctly draws all of its 9 diagonals, but when counting them he stops at 6. Then he draws a square with the two diagonals. In line 44 he completes the inscription summarising the results he just found: if n=4 then d=2, if n=6 then d=6 (as noticed, the number of diagonals should be 9 for the hexagon instead).

Finally (lines 45-47), Tommaso interrupts the exploration of the problem claiming to be "stuck" and not know how to solve the problem ("I cannot make it… I am one hundred percent sure… never").

In the whole excerpt, Tommaso's semiotic production is almost uniquely ground level. Tommaso makes several drawings representing polygons and diagonals and counts them in different ways, without finding a way that he feels it could help to solve the problem. There are only a few meta level interventions, which, moreover, do not seem to be used by Tommaso to describe and reflect on the problem or on some possible heuristics, but only to introduce a drawing or to state that he has not found a solution for the problem yet. In line 14, when considering the case of the hexagon, his words "I need to visualize it" anticipate the drawing of a hexagon of the sheet. In line 23, he says "there must be a faster way" when deciding not to further develop his reasoning of counting the diagonals vertex by vertex without considering the already counted diagonals. In line 30 and 32 he reflects on some drawings he did ("what have I done?") and admits being a bit lost in those. Finally in lines 45 and 47, Tommaso states to be "stuck" and to be "one hundred percent sure" that he is not going to find a solution for the problem. No meta level gestures or inscriptions have been registered. All the rest of the discourse is ground level, referring to polygons and to the count of diagonals. Tommaso's discourse seems to be missing a reflection on the problem itself, and on the different heuristics he tried to use for solving it, as the absence of such a meta-level semiotic production seems to highlight. By reflecting on his explorations, in fact, Tommaso could have recognised that he already found a solution for the problem at line 22 which he did not further explore. This awareness, however, seems to be missing, and the episode concludes with Tommaso sure of not being even close to a solution for the problem.

Let us notice that in this excerpt, some of Tommaso's meta level sentences are expressions of disappointment and emotional disposition. This aspect is related to another classic research theme in mathematics education which is the study on 'Affect' (for a summary, see Zan et al., 2006; Hannula M.S., 2014; Hannula, M. S. et al., 2019), which was not addressed in this thesis. This observation, however, opens to a possible direction for future research. The analysis of meta-level semiotic production could be enriched by introducing the affective dimension to further differentiate the meta-level semiotic production. This could provide interesting elements to further develop the study.

The graph reported in figure 11.40 models the analysis of Tommaso's excerpt and further highlights how for the whole excerpt his semiotic production has been almost uniquely ground level, with an absence of liftings or projections and of merging points or level's mismatch.

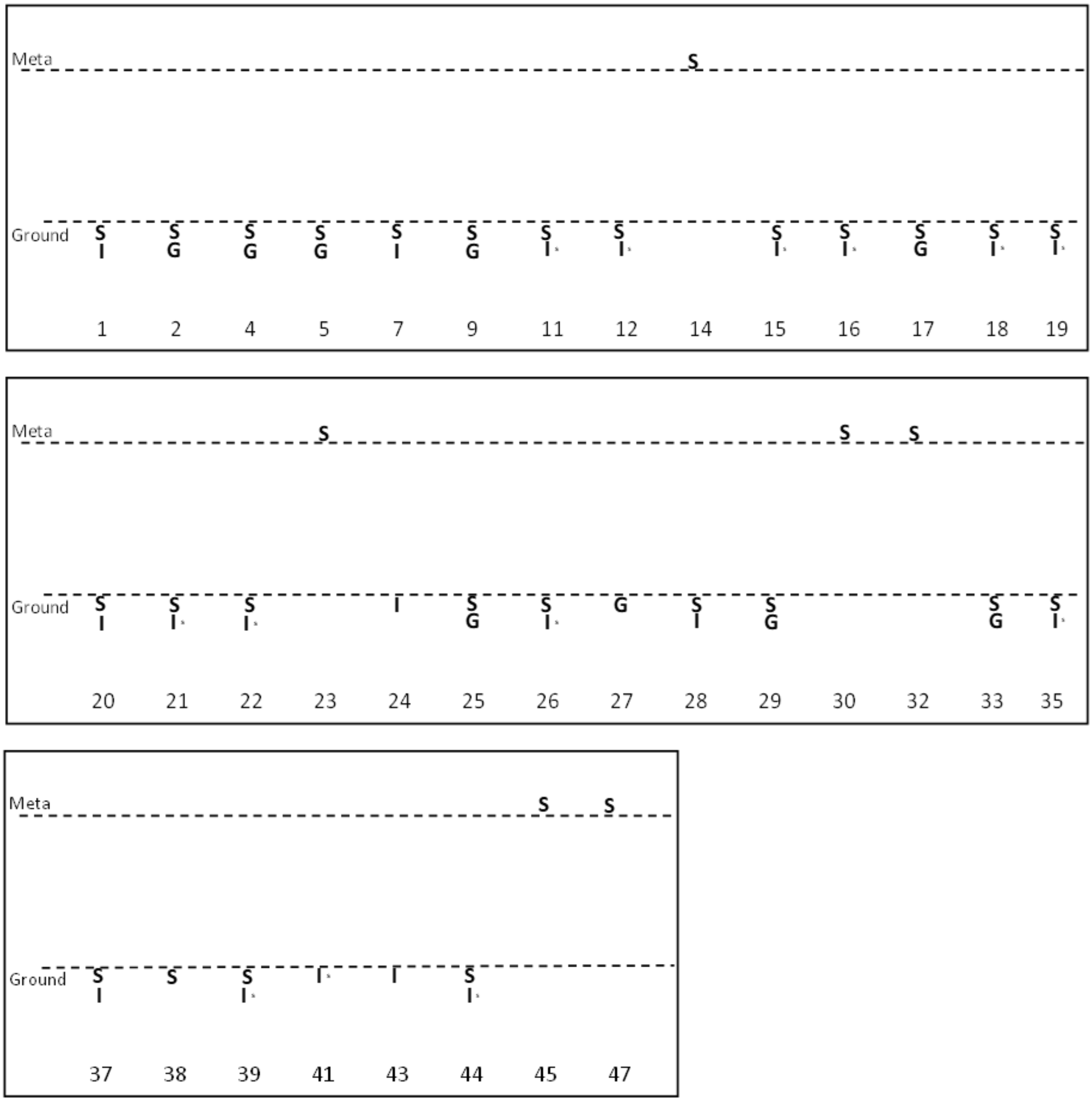


Figure 11.40 Diagram for the analysis for the excerpt 11.5.

11.1.6 The dialectic between ground and meta level semiotic production as a resourse.

In the previous examples of Giuditta (11.3) and Silvio (11.4) I registered a series of continuous transitions between ground and meta level semiotic production during their problem resolution. In particular, the presence of several liftings and projections, as well as of level's mismatches and merging points highlighted how their semiotic production was composed by a series of extremely intertwined meta and ground level signs. In other terms their semiotic production seems to be characterised by a sort of 'permeability' between the two levels. As we observed this dialectic between the two levels seems to have supported the two students in developing their recursive argumentation. On the other side, Tommaso's examples (11.5) showed a completely different situation. The student, after a series of different calculations and drawings states to be stuck and abandon the problem resolution. In this episode we registered a semiotic production which was, almost uniquely, ground level and without the continuous transitions which characterised the previous examples. As we noticed, as the absences of meta level signs might highlight, Tommaso misses a reflection on what he previously did, through which he could have recognised a possible solution for the problem.

The dialectic between the two levels seems to be crucial during the problem resolution. In particular, by lifting an argumentation from ground to meta level a subject might reflect on the structure of the argumentation itself and recognise a known structure for it (in our cases, for instance, a recursive argumentation). On the other side by projecting an argumentation from meta level to ground level a subject might follow the structure of known argumentative schema to construct a specific argumentation on the problem involved. Both the transitions, as the examples of Giuditta and Silvio of above have shown, seems to be important and effective for the resolution of the problem and for the construction of the recursive argumentation and of a proof by MI.

To further deepen this point, it is worth it to focus on the specific roles that these transitions have for a student during a problem resolution and, in particular, during the construction of a recursive argumentation or of a proof by induction.

First of all, they have a communicative role. We should, in fact, remember that during the interviews the students were asked to 'think aloud' and to try to describe their reasonings to the interviewer. Thus, it is possible that this request might have prompted the subjects to reflect on their reasonings, to describe them, and thus to produce meta level signs intertwined with ground level signs.

In addition to this, moreover, we can observe how these transitions between levels have also had an active role in supporting the students' resolution of the problem. In other terms, they can become resources for the subjects producing them.

In particular, how some of the previous examples showed, some crucial moments during the problem resolution seem to take place when the subject reflects on the exploration of the problem exploration that s/he has done (or that s/he is planning to do). In this reflection s/he might recognise that a recursive argumentation could be helpful to reach a problem solution. The transition from ground to meta level might play an important role in this process of 'recognising' since meta level signs can have a sort of potential generality. Indeed, they refer to the specific problem under consideration, but they could also be meaningful if referring to a different problem. Let us consider, for instance, the words "every case is reconducted to the previous one", "moving forward to the next case", "I could do something like by induction", or the arc-shaped gestures corresponding to meta level linking or IS. These are signs which were produced during the resolution of different problems. Therefore, the potential generality of these signs could support a student in recognising that the problem under consideration can be solved with the use of a heuristics already interiorized and used for other problems, such as a recursive argumentation. A similar aspect can be observed also in a different phase of the problem resolution, corresponding to the construction of a proof to formalise a previous argumentation. In this process, the meta level semiotic production could support the student in the organisation and the structuration of the proof, providing a structure (on the meta level) to follow and to project on the ground level with the objects of the problem.

To summarise, thus, a semiotic production rich of transitions between the two levels could support a student in the different phases of the problem resolution:

- In the exploration of the problem. In particular, the production and use of meta level signs could let the student reflect on the problem itself and on the resolution process. Through this the student might see in them some features ("Every case is reconducted to a previous one", to quote Silvio) which let her/him recognise that the problem could be solved using a recursive argumentation.

- In the construction of the recursive argumentation (or of a proof by induction). In particular a student might be aware of the structure that a recursive argumentation should have. In this case the production and use of meta level signs could support the student in organising the construction of the argumentation or of the proof which can be obtain as a projection from the meta in which the structure of a generic argumentation is described to the ground level of the specific problem under resolution.

Following what just said, thus, the presence of liftings and projections in a student's semiotic production could highlight some crucial moments in the resolution of a problem:

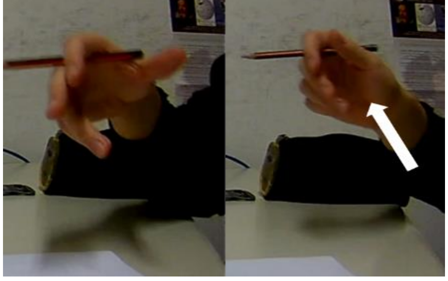
- Liftings could highlight the student's identification of a heuristic or of an argumentative structure, in our case a recursive argumentation, as resolutive for the specific problem (as observed in the case of Giuditta in the excerpt 11.3).
- Projections could highlight the student's application of an argumentative structure, in our case a recursive argumentation, to the specific problem under resolution (as observed in the case of Silvio in the excerpt 11.4).

It is important to specify that the exploration of the problem and the construction of an argumentation, as well as the identification and the application of heuristics or argumentative structures are not, generally speaking, separate moments, and thus in general they require different transitions between levels. In fact, in effective resolution processes, as the previous examples of Giuditta and Silvio have shown (excerpts 11.3-11.4), the transitions between levels turn out to be very complex and intertwined. For this reason, I have spoken of a dialectic between the two levels.

In the next paragraph I will present an example showing a complete resolution of a problem, both in its explorative phase and in the construction of an argumentation and then of proof phase. In particular the episode will highlight, as just discussed, how the dialectic between the two level could support the student firstly in the exploration of the problem and in the generation of a recursive argumentation and then in the successive construction of a proof by induction.

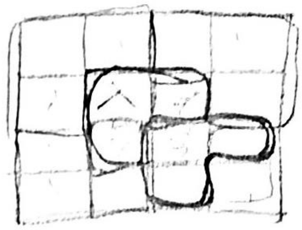
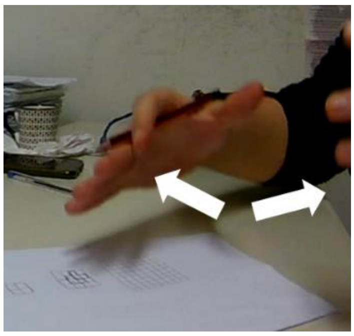
Excerpt 11.6

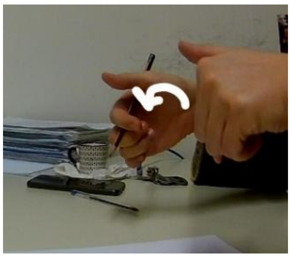
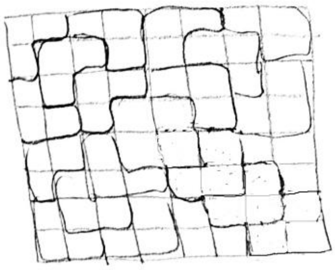
Guido is dealing with the chessboard problem. In the previous minutes, after having read the text of the problem, he has drawn a 2x2 chessboard and he has concluded that three of the four squares can be covered with a tile. Then he has drawn a 4x4 chessboard, finding a tessellation which covers all the squares except for one in the centre. After having considered these two cases, one separated from the other, without any external intervention, Guido refers to a possible proving strategy:

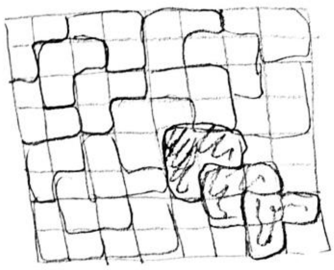
	Who	Speech	Gesture and inscription	Level
1	**G**	Ok, so, we could try… no wait, too soon. I wanted to do kind of an inductive thing, but I went too fast.	*[inductive thing]. He moves his right hand in the air in a straight line from left to right.*  *Fig.11.41*	*S: Meta* *I: Meta*

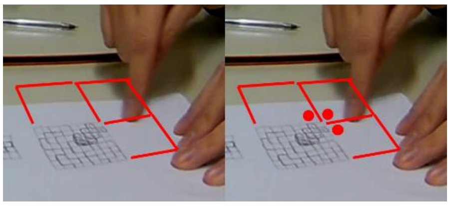
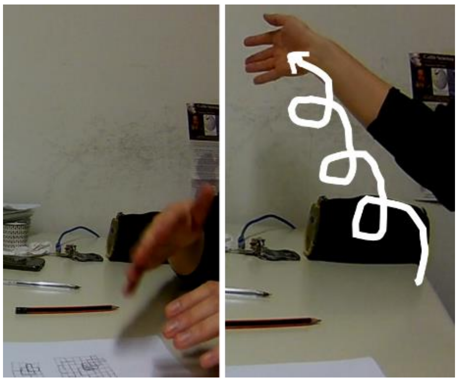
2	**I**	What is your conjecture?		
3	**G**	No, no, I don't have any conjecture yet. Let's do another case and then we do the conjecture.		*S: Meta*

First Guido, generalising from the cases of the 2x2 and 4x4 chessboard, observes that for reasons of divisibility in every chessboard with dimension $2^n x 2^n$, in the best case, a tessellation would leave out one little square. Then he draws the inscription of an 8x8 chessboard, and he observes it for almost thirty seconds moving the pen over the inscription as shaping some tiles on it.

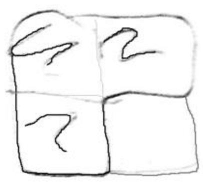
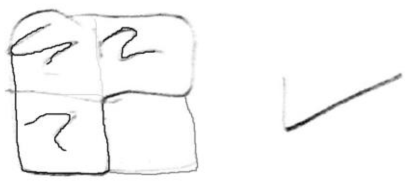
[…]				
4	**G**	Ah! If I could do… if I put this in this way, here I can do like that… Then the one left out is here on this corner, which is better.	*He changes the inscription representing a tessellation of the 4x4 chessboard so that the left out little square can occupy one of the corners of the chessboard. The obtained inscription is the following.*  *Fig.11.42*	*S: Ground* *I: Ground*
5	**I**	Tell me again this, please.		
6	**G**	Because I noticed that… widening…	*He keeps the two hands open and close to each other over the sheet; then he moves the right hand to the right and the left hand to the left.*  *Fig.11.43*	*S: Ground*

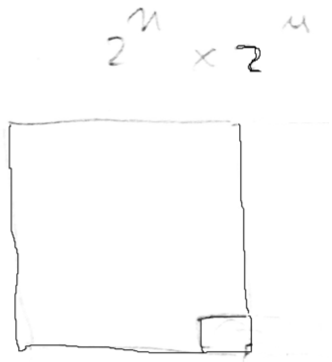
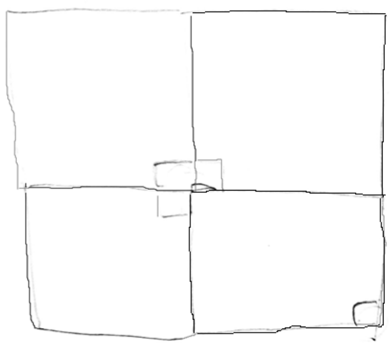
7	**G**	I would need to change a bit how I drew it here. So that I can do the successive step in a better way.	*[successive step]. He rotates forwards both the hands in the air.*  *Fig.11.44*	*S: Meta* *G: Meta*
8	**I**	Ok		
9	**G**	So, let's see… First, I'll do and then I'll try to say it better.		*S: Meta*
10	**I**	Yes.		
11	**G**		*He starts the drawing of a tessellation for the 8x8 chessboard. In doing it he seems to copy the same pattern of the tessellation of the 4x4 chessboard.*	*I: Ground*
12	**I**	What are you trying to do?		
13	**G**	All of them except for one		*S: Ground*
14	**I**	Are you using the same… method, let's say?		
15	**G**	Yes, I'm trying to generalise, let's say. So maybe from the next step I can do it for all of them…. which is absolutely not granted, I am not sure to have understood…	*He completes the drawing of the tiles on the 8x8 chessboard obtaining a tessellation which leaves out the four central little squares.*  *Fig.11.45*	*S: Meta* *I: Ground* *Level Mismatch*
16	**G**	And then we put one of any kind, like this one.	*He draws tile covering three of the four little squares in the centre of the 8x8 chessboard.*	*S: Ground* *I: Ground*

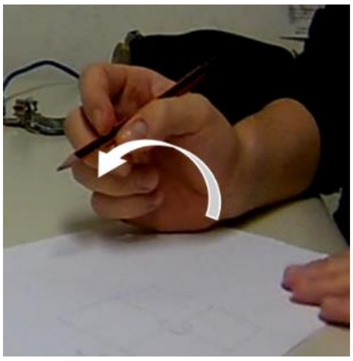
17	**G**	Ah, but I should have…		*S: Meta*
18	**I**	What?		
19	**G**	Eh, I need to do again what I have done again at the previous step…	*[previous step] He rotates his right hand backwards in the air.*	*S: Meta* *G: Meta*
20	**I**	That is?		
21	**G**	To try to flip it another time.		*S: Ground*
22	**I**	What do you mean with flip it?		
23	**G**	Yes, yes. Because I keep leaving this, but instead no, I want to leave a corner.	*He points to the centre of the inscription of the 8x8 chessboard and then to its bottom right corner.*	*S: Ground* *G: Ground*
24	**I**	Ok.		
25	**G**	So, like this, I'll do it well, like this and this, and these others…	*He changes the drawing of some tiles on the 8x8 chessboard so that the left out little square is occupies the bottom right corner of the chessboard.*  *Fig.11.46*	*S: Ground* *I: Ground*
26	**I**	Ok.		
27	**G**	At the end, now, yes, I am convinced that it is possible to cover it entirely except for one.		*S: Ground*

28	**I**	Ok		
29	**G**	Why? Because I have seen that there is a sort of rule to put them in a way that the corner here at the bottom right always remains out, or where you prefer.	*He points to the bottom right corner of the drawing of the 8x8 chessboard.*	*S: Ground* *G: Ground*
30	**G**	Then you do... Here you leave it at the bottom left, here top right, here top left...	With his right pointing finger, he draws on the sheet the perimeter of three squares attached to the inscription of the 8x8 chessboard. After drawing each of these squares he points to a position corresponding to a corner as shown in the following figure.  *Fig.11.47. The red lines indicate the finger trajectory. The red dots where he points to after the drawing of each square.*	*S: Ground* *G: Ground*
31	**G**	And you go forward in an inductive way, let's say.	He moves his right hand in the air from left to right while shaping some circles.  *Fig.11.48*	*S: Meta* *G: Meta*

After this, the interviewer asks Guido to construct a proof for the statement emerged from the previous argumentation: "It is possible to completely tile every 2^nx2^n chessboard except for one little square in one of the corners".

		[…]		
32	**I**	Ok. I ask you to try to write the proof.		
33	**G**	Oh, it’s not easy. Can I take a sheet?		*S: Meta*
34	**I**	Yes, or you can turn this one.		
35	**G**	No, no, I prefer another sheet, so I can look at this one.		*S: Meta*
36	**I**	Ok, no problem.		
37	**G**	So, how can I do? It is no easy at all to explain…		*S: Meta*
38	**I**	It can be… it can refer to some figures, it doesn’t have to be…		
39	**G**	Yes, yes… All right, so, by induction I would do like this…		*S: Meta*
40	**G**	I’ll create the base step…	*He draws a 4x4 chessboard and a tile on it covering three squares.*  *Fig.11.49*	*S: Meta* *I: Ground* *Level Mismatch*
41	**G**	Which is this one here. I have… one vertex which is empty.	*He points to the bottom right corner of the drawing.*	*S: Ground* *G: Ground*
42	**G**	And this is verified by this figure.	*He marks with a tick the just made drawing.*  *Fig.11.50*	*S: Ground/Meta* *I: Meta* *Merging Point (Speech)*

43	**G**	Then I say… so, let’s suppose it is true for n.		*S: Meta*
44	**I**	Ok		
45	**G**	So, I have a n by n square.	*He draws a square.*	*S: Ground* *I: Ground*
46	**I**	It should be two to the n by two to the n.		
47	**G**	Yes, sorry, two to the n by two to the n.	*He writes ‘$2^n x2^n$’ above the inscription of the square.*	*S: Ground* *I: Ground*
48	**G**	Ok? With this one which is missing.	*He draws a little square corresponding to the bottom right corner of the big square.*  *Fig.11.51*	*S: Ground* *I: Ground*
49	**I**	Ok		
50	**G**	And then I construct the successive one in this way.		*S: Ground/Meta* *Merging Point (Speech)*
51	**G**	It is not important how it is oriented…	*He draws three other squares extending the previous inscription and for each of them he draws a little square in one their corners.*  *Fig.11.52*	*S: Ground* *I: Ground*

52	**I**	Ok.		
53	**G**	These three, I have a tile… and here I have an empty one.	*He points to the three little squares in the centre of the previous inscription and then to the bottom right squares.*	*S: Ground* *G: Ground*
54	**G**	And so…	*He makes a rapid arc-shaped gesture in the air with the pen on the right hand, from left to right.*  *Fig.11.53*	*S: Meta* *G: Meta*
55	**G**	I constructed, starting from a two to the n by two to the n, a two to the n plus one by two to the n plus one, with a missing corner.	*He writes '$2^{n+1}x2^{n+1}$' under the inscription of line 51.*	*S: Ground* *I: Ground*
56	**I**	Ok, good. Let's stop here.		

<u>Analysis</u>

The episode starts after Guido's exploration of a 2x2 and of a 4x4 chessboard. For both cases he has found a tessellation which covers all the squares except for one. The two cases were addressed separately. This part, not reported above, was characterised by a ground level semiotic production. At the beginning of the excerpt (line 1), Guido seems to want to start the construction of a recursive argumentation, but he decides that it is too soon: "I wanted to do kind of an inductive thing, but I went too fast." This sentence and the left to right path gesture in the air co-timed with the words "an inductive thing" are meta level. When the interviewer asks Guido to explicit his conjecture, he answers not to have it yet but that he needs to explore "another case" (line 3).

After this, the semiotic production comes back to the ground level, while Guido explores the case corresponding to the 8x8 chessboard. From this exploration, Guido observes something that will be fundamental for the resolution of the problem: the 8x8 chessboard can be tiled starting from four 4x4 chessboard with an appropriate tessellation. This observation and its justification occupy several lines of the episode (4-16) in which we can register a continuous intertwining between ground and meta level semiotic production: the ground level signs, produced and used whilst the tessellation of the 8x8 and of the 4x4 chessboard is drawn and described, alternate with a series of meta level signs. Firstly, in lines 6 and 7, we can register a lifting: the words “widening” together with the gesture miming the enlargement of the 4x4 chessboard, become, in the successive line, the words “the successive step” together with a rotating forward gesture in the air. In both cases Guido refers to the transition between the 4x4 and the 8x8 chessboard, but he does it firstly with ground level signs and then with meta level signs. Secondly, in line 15, we register the presence of a level’s mismatch. The meta level spoken utterance “I’m trying to generalise, let’s say. So maybe from the next step I can do it for all of them” is produced while Guido is copying the tessellation of the 4x4 chessboard on the drawing of the 8x8 chessboard, thus while producing a ground level inscription.

In the successive lines (17-26) Guido observes that in order to iterate the reasoning to the successive chessboards after the 8x8 it is necessary that the left out square occupies one of the corners of the chessboard, exactly how it was for the 4x4 chessboard. Also in these lines we can register a dialectic between ground and meta level signs: at the beginning (lines 17-19), Guido’s discourse is in fact meta level (“I need to do again what I have done again at the previous step”) and then (lines 21-25) it is projected on the ground level when constructing and describing a new tessellation for the 8x8 chessboard leaving out one of its corners, represented by the inscription of line 25.

Guido thus has constructed a tessellation for the 8x8 chessboard with one little square in one of the corners left out by using four 4x4 chessboard tessellated so that only one little square in one of their corners was left out. In line 27 Guido states to be convinced to have solved the problem: “At the end, now, yes, I am convinced that it is possible to cover it entirely except for one”. He then justifies this statement by summarising in three lines (29-31) what observed until now. These lines contain some extremely interesting elements for our analysis. At line 29 and 30 he produces a ground level LS representing the connection between the tessellation of a chessboard (8x8 in the drawing) and the successive one. Then at line 31 Guido says that it is possible to “go forward in an inductive way” and while saying this he makes a gesture in the air, shaping some circles while his hand moves from left to right. This line, thus, contains an IS which is meta level since it refers to the recursive structure of the argumentation and not to the chessboards or to tessellations. If we focus on the development of these three last lines, we can register the presence of a lifting: Guido’s recursive argumentation is firstly constructed with a ground level semiotic production, when referring to a step between two chessboards, and then it is concluded with a meta level semiotic production, when referring to the successive iteration of steps. In other terms Guido has lifted his recursive argumentation from ground to meta level, firstly referring to some specific objects of the problem (the chessboards) and then to the structure of the argumentation itself.

This concludes the first part of the episode, corresponding to Guido’s exploration of the problem leading to a conjecture and to the generation of a recursive argumentation for it. As observed, all this part (lines 1-31) was characterised by a rich dialectic between the ground and meta level semiotic production, with a series of different level’s transitions. Before analysing the second part of the episode, corresponding to Guido’s construction of a proof by induction to formalise his solution, we can observe that until this moment, he has produced a series of meta level signs (such as the arch-shaped gestures, or the utterances like “the successive step” or “go forward in an inductive way). Some of these signs, as we will see, will be repeated in the successive part of the episode, generating a sort of ‘semiotic echo’ which will contribute to the cognitive unity between argumentation and proof.

The successive phase of the episode is opened by the interviewer’s request to Guido to write a proof for what just stated. At line 33 an interesting passage occurs: Guido asks a new blank page where to the write the proof; the interviewer answers that he can use the back of the page used until now, but Guido says that he prefers to take a new sheet so he can “look at this one” (i.e., what already written). These lines highlight that Guido intends to construct the proof following what already written during his argumentation, thus obtaining what we can interpret as a cognitive unity between the two phases of the episode.

At line 39 Guido starts constructing a proof by induction. This construction is characterised by a series of projections from the meta to the ground level.

At line 40, Guido says “I create the base step” (meta level) while drawing a 2x2 chessboard with a tile put on it (a ground level inscription) and at line 41 he further describes in ground level terms the just created base case (“I have… one vertex which is empty”). The verb to “create” used by Guido in line 40 seems to highlight that he is following a sort of a structured path (the one of a proof by induction) which has some fixed elements which Guido is trying to reconstruct with the problem’s objects.

At line 43 Guido starts the proof of the inductive step. His discourse is, initially, meta level “let’s suppose it is true for n”. Then, in lines 45-48, it becomes ground level: “So, I have a n by n square […] With this one which is missing”, together with the drawing representing the $2^n x 2^n$ chessboard completely tiled except for one little square in one of the corners (line 48). As previously, in this passage we can register a projection: firstly, Guido refers to a generic inductive hypothesis (“let’s suppose it is true for n”) which is then described, introduced by “so”, in ground terms: a $2^n x 2^n$ chessboard can be completely tiled except for one little square in one of the corners. In this perspective it is interesting the mistake made by Guido who initially refers to the chessboard n x n instead of $2^n x 2^n$, then corrected at line 47. It seems like, in the projection of the inductive step, Guido remained anchored to the sign ‘n’ of the meta level without transforming it in 2^n when referring to the chessboard’s dimension. With this mistake, however, Guido can take into consideration, simultaneously, two aspects: the fact that the chessboard is a square and the fact that it represents the ‘case n’ of a recursive argumentation.

At line 50, the sentence “And then I construct the successive one in this way”, introduces the drawing of the $2^{n+1} x 2^{n+1}$ chessboard with a tessellation obtained with four $2^n x 2^n$ chessboards obtained and described at line 51. The sentence at line 50 can be interpreted as both ground and meta level. It refers to a ‘construction’ which can be seen as a ground level expression referring to the construction of a chessboard and of a tessellation for it. However, the adjective “successive” seems to refer to the “successive step” of the iteration (to use Guido’s words of line 7) and thus can be read as meta level. Due to this duality of levels, this line can be interpreted as a merging point. In the successive lines (51-53) the semiotic production is projected on the ground level and the inductive step is explicitly constructed by extending the drawing of the $2^n x 2^n$ chessboard (line 51) and describing how to complete the tessellation for the $2^{n+1} x 2^{n+1}$ chessboard (line 53). We can observe that Guido’s construction of the inductive step is obtained by producing a LS (the inscription of line 51) which is analogous to the LS he produced at line 30 during the construction of the recursive argumentation to link the tessellation of the 8x8 chessboard with the one of the 16x16 chessboard.

In the last two lines (54-55) Guido describes a second time what just done to highlight that he has just proved the inductive step. In doing this he firstly produces a meta level LS: at line 54 he says “and so…” while making an arc-shaped gesture in the air. In this line Guido seems to describe in a very condensed way the inductive step itself and in fact the gesture seems to seems to partially repeat the meta level gestures of lines 7 and 31. Finally, in the successive line Guido projects his discourse on the ground level and he describes the just constructed proof for the inductive step, this time referring to the chessboards and tessellations (“I constructed, starting from a two to the n by two to the n, a two to the n plus one by two to the n plus one, with a missing corner”). This concludes Guido’s proof.

Figure 11.54 contains the diagram corresponding to the analyses of Guido’s episode.

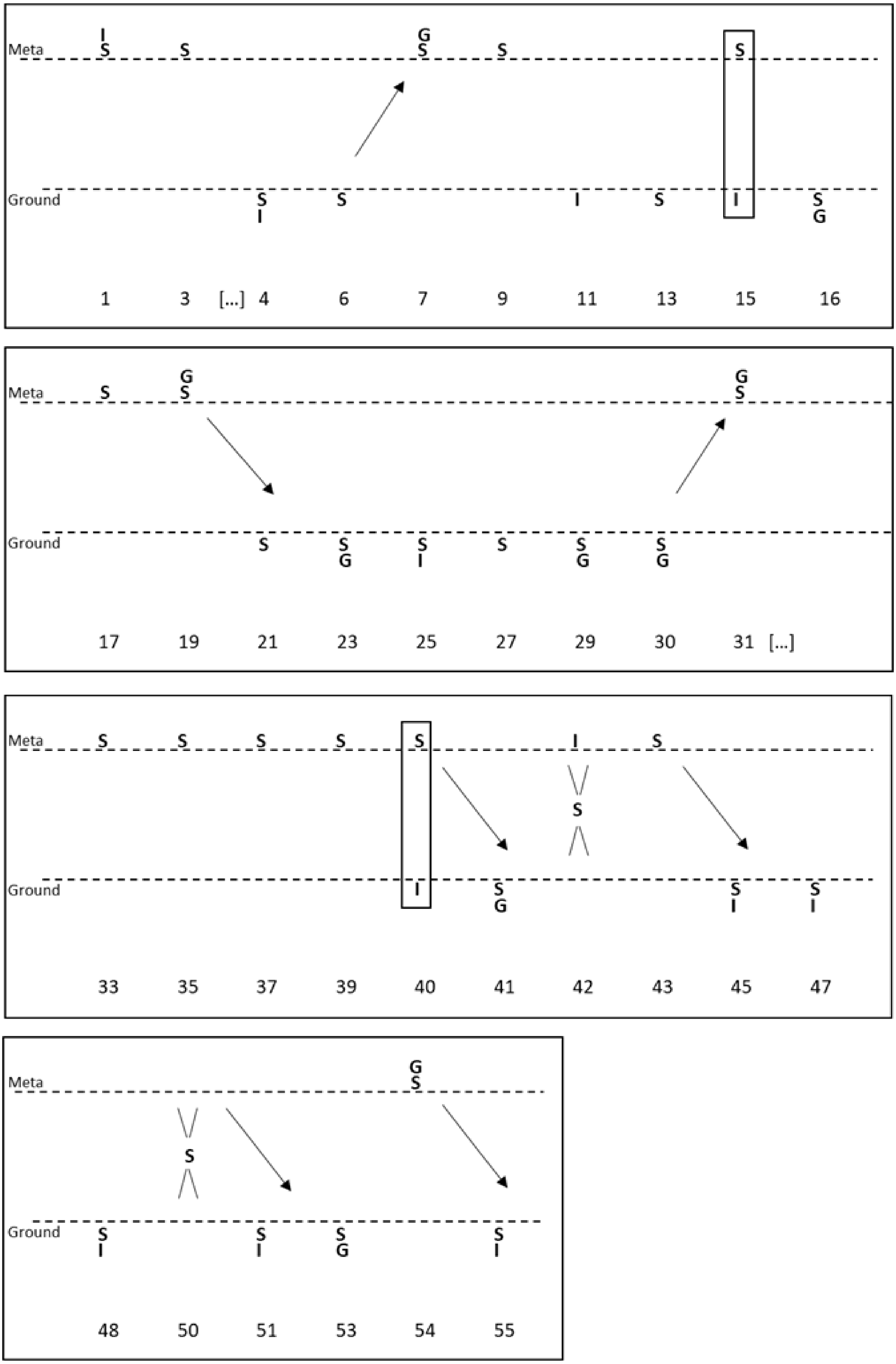


Figure 11.54. Diagram for the analysis for the excerpt 11.6.

The just presented analysis of Guido's excerpt has allowed us to observe some interesting aspects which were already partially described when introducing the excerpt.

First of all, we registered how the whole episode is characterised by several liftings and projections. Some differences in these level's transitions can be tracked if we divide the episode in two parts, respectively the one involving the generation and the construction of the argumentation (1-31) and the one involving the construction of the proof by induction (32-56). In the argumentation phase, even if not in a linear way, Guido's semiotic production seems to start from the ground level and then to be lifted to the meta level. The connection between two chessboards and their tessellation are initially represented by ground level LS (as the gesture-speech-inscription bundle of line 30) and then, later, by meta level signs (as in the meta level IS of line 31). In this first phase, thus, the main direction of the transitions is from ground to meta level: Guido, exploring the connections between two chessboards, becomes aware that a recursive argumentation could be used to solve the problem. In the second phase of the episode, instead, when a proof by induction is constructed, the level's transitions in Guido's semiotic production have the opposite direction, from meta level to ground level. Both the base of the induction (lines 40-42) and the inductive step (lines 43 -53 the first time, and 54-55 its repetition) are firstly described in meta level terms and then in ground level terms, forming, as a result, a series of different projections.

Secondly, it was possible to register a cognitive unity between Guido's argumentation and his successive proof. This unity can be seen in the parallelism between the logical structure of the argumentation (a recursive one) and of the proof (by induction) and in Guido's semiotic production as well. In particular some signs produced during the construction of the argumentation have been repeated during the construction of the inductive step. It interesting that some of these signs are meta level signs (such as the meta level linking and iteration gestures). The fact that Guido uses similar meta level signs when describing the structure of the recursive argumentation and of the proof by induction might suggest that Guido is perceiving them as argumentations which are structurally analogous in general terms and not only for this specific problem. In other terms, meta level signs might contribute to create a *structural continuity* (Boero et al. 2010, see section 3.2) between recursive argumentation and proof by induction.

In the introduction to this last episode, I observed that the dialectic between ground and meta level semiotic production, in addition of having a communicative role, could become a resource for students themselves. In particular, driven by previously analysed examples, I observed that producing and using meta level signs could support a student in recognising that a recursive argumentation could be used to solve the problem, and in constructing such an argumentation, or, more specifically a proof by induction. Guido's episode confirms these observations, showing an example, that could become paradigmatic, of a complete and successful transition from exploration of the problem to the construction of a recursive argumentation, first, and of proof by MI then. This transition, as observed, is characterised by an extremely rich dialectic between ground and meta level semiotic production. The result of this dialectic process is that Guido's argumentation and his successive proof are structured and described both with meta level signs referring to their general logical structure (recursive or by MI) and with (ground level) signs referring to the specific problem under resolution.

11.2 FIRST CONCLUSIONS ON GROUND AND META LEVEL SEMIOTIC PRODUCTION

In this chapter, I presented how, analysing the interviews with a focus on the students' semiotic production, it was possible to distinguish signs which can be interpreted as belonging to two different categories: ground-level signs and meta-level signs. I then presented a methodological tool, introducing the notions 'lifting', 'projection', 'merging point', and 'level's mismatch', to observe, describe and analyse the dialectic process between these two levels of semiotic production. Analysing the interviews with this perspective some aspects have been observed.

Firstly, I registered how **a subject's meta level semiotic production highlights some crucial moments in the problem resolution in which the subjects reflect on the problem itself or on the resolution process**. In those moment the subject seems to pause calculations, symbolic manipulations, geometrical transformations, etc., in order to observe, reflect on, the structure of the problem or of her/his own argumentation or to describe what s/he has done, what s/he is planning to do. These reflective moments, highlighted by a meta level semiotic production, have been registered continuously in all the interviews with expert students, as presented in several examples in this chapter. However, as shown in the Tommaso's excerpt 11.5 it was also possible to register episodes in which these reflective moments seemed to be missing and the semiotic production was almost uniquely ground level. Within this perspective, thus, meta level signs can become sort of indicators for a researcher to investigate the reflective moments during the resolution of a problem presented above. This result support the methodological decision of interviewing expert students, for which a rich dialectic between the two levels was registered. If the study had been restricted only to less expert students it is possible that this aspect would have not been registered. When this methodological choice was presented in section 8.4.2, I observed that the observation of effective process in expert students could provide elements for interpreting students' difficulties in terms of non-activated processes. This point was confirmed in the analyses of above. In particular it was possible to interpret Tommaso's difficulties in 'navigating' in the problem resolution without 'getting stuck' (to use his expression) in terms of a semiotic production almost uniquely ground level without a dialectic with a meta level semiotic production.

With a focus on the previous aspect, I noticed that meta level signs can become resources for the subject involved in the construction of a recursive argumentation or of a proof by induction. In fact, we observed that **meta level signs could play an important role in supporting a subject in the generation of a recursive argumentation**. Meta level signs, in fact, could let the student recognise in the problem some characteristic for which it can be solved with a recursive argumentation. The use and production of **meta-level signs could also support a student in the construction of a proof by induction**, which can be obtained as a projection of the structure of a generic proof by induction, expressed on the meta level, on to the ground level in which it is reconstructed with the problems' objects. In this case, thus, meta level signs can guide the construction of a specific proof by induction.

As a further aspect, the distinction between ground and meta level semiotic production allowed me to investigate the students' cognitive unity between recursive argumentation and proof. In particular, as shown in the last excerpt, we observed that similar meta level signs, in particular LS and IS, can be used and produced to refer to the structure of a recursive argumentation and of a proof by induction. **This continuity on meta level signs on a subject's semiotic production can highlight the presence of a structural continuity between recursive argumentation and the successive proof by induction**. In other terms, a continuity on meta level signs could be seen as a trace of the fact that the student is recognising a similar structure between recursive argumentation and proof by MI.

The fact that meta level signs can be important resources for students brings to a direct didactical implication: the construction of such a meta level semiotic production in students could be a didactical goal for teachers. As the analysis of the survey in the previous chapter has highlighted, producing, and giving meaning to a meta level discourse on MI can be problematic for students. Thus, it could be helpful, during the teaching design, to plan activities involving the use and production of meta level signs. In particular, referring to the recursive argumentation and the proof by induction, students should be able to construct a proof by induction for specific problems (ground level), however they should also be able to talk about a generic proof by induction, describing its structure and recognising in which cases it is possible to prove by induction (meta level). This dual goal seems to be didactically problematic: to talk about induction on the meta level seems to require the experience of proving by induction in some specific problems, that is to have experience of a ground level semiotic production related to MI. On the other hand, the construction of a proof by induction on a specific problem seems to require for the subject to know how the structure of a generic proof by induction should be and when it can be used. In other terms it seems like the two levels of semiotic production related to MI are prerequisite one to the other. This sems to be analogous to the “vicious circle” described by Sfard (1991) referring to the dual nature of mathematical entities as processes or objects:

> The thesis of the "vicious circle" implies that one ability cannot be fully developed without the other: on one hand, a person must be quite skilful at performing algorithms in order to attain a good idea of the "objects" involved in these algorithms; on the other hand, to gain full technical mastery, one must already have these objects, since without them the processes would seem meaningless and thus difficult to perform and to remember. (Sfard, 1991, p. 32).

Following this perspective, thus, the two levels (ground and meta) should be continuously and reciprocally constructed, without any prevalent direction. With this aim, therefore, it can be useful to design activities with recursive argumentation or proofs by induction, involving the transitions between the two levels: lifting on the meta level when reflecting on the structure of a recursive argumentation or projecting on the ground level when shaping such an argumentation on a specific problem.

12 OTHER RESULTS AND CONCLUSIONS OF THE STUDY WITH THE INTERVIEWS

This last chapter dedicated to the interviews is divided into two parts. In the first one, other results registered form the analysis of the interviews and related to other research questions of this thesis will be presented. In the second part, I will summarise what observed in this and the previous two chapters, discussing the overall conclusions of the whole empirical study with the task-based interviews

12.1 OTHER RESULTS

In the previous two chapters I focused on the analysis of the interviews investigating the research questions RQ4, addressing the use and production of signs by students involved in the construction of recursive argumentations and proofs by induction. During the interviews, however, it was possible to register elements which contribute to investigate other research questions, specifically the question on the reconstruction of the explain induction process (RQ2) and the question on the intuitive acceptance of a theorem by MI as a triplet (RQ3). In the first part of this chapter I will briefly present some aspects on these two points emerging from the interviews. All the presented excerpts will refer to the last part of the interviews when questions on MI itself were posed to students, in particular on the justification of its validity. The excerpts will be analysed with reference to the theoretical elements introduced in this chapter, in particular the notions of LS and IS. Since the excerpts refer to questions involving the description of the functioning of MI as a proving scheme and not the resolution of a specific problem, the analyses that I will present will not involve the notions related to ground and meta level which focus on a subject's semiotic production during the resolution of a mathematical problem.

12.1.1 Explain induction processes

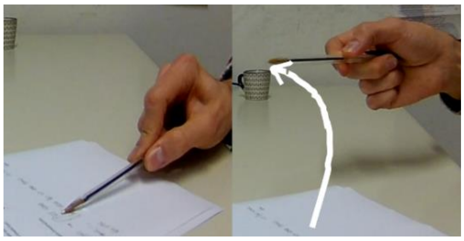
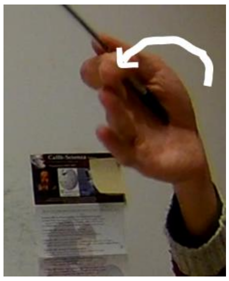
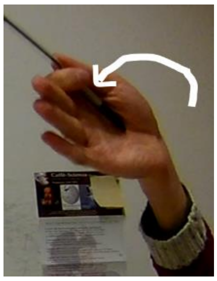
The semiotic perspective allowed us to register the reconstruction of the explain induction process in several of the interviewed students. In particular, the justification of the validity of MI was described by students with the production of several LS and IS, used by them to represent the chain of logical inferences involved in the explain induction process. In the GD of MI, two forms of the explain induction process were proposed: the direct and the indirect form, respectively involving the construction of a chain of MP or of MT. In the interviews, it was possible to register traces of both the processes.

I will present here two examples, involving the same person, Silvio, who in two different moments of the interview, constructs the two explain induction processes (direct and indirect). The examples will also show how the semiotic perspective, focusing on LS and IS, can offer a lens through which observe and analyse these processes.

Excerpt 12.1

Silvio has just found the flaw in the (false) proof by induction for the false coin problem (presented as 'A false proof by induction', in the methods section). The interviewer asks to Silvio what justification he gives to himself of the validity of a proof by MI.

	Who	Speech	Gesture and inscription
1	**S**	You can do this like a chain. It makes sense, in reality…	*He points to the sheet where the following inscription is written:* *'P(0) True; if P(n) is true → P(n+1) is true. Therefore P(n) is true ∀n'.*

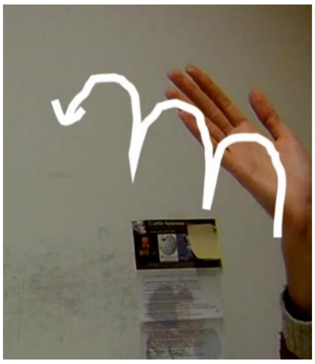
2	**S**	Because you say: In zero it holds, then it holds for the consecutive one, which is one.	*Starting with the pen touching the sheet were 'P(0) True' is written, he makes an arc-shaped gesture in the air, from left to right.*  *Fig.12.1*
3	**S**	Then, if it holds for one it holds for two,	*With the pen air in the air, he makes an arc-shaped gesture from left to right.*  *Fig.12.2*
4	**S**	if it holds for two it holds for three,	*He makes the same arc-shaped gesture in the air of line 3.*  *Fig.12.3*
5	**S**	Etcetera, etcetera.	*He moves the hand from left to right, whilst drawing some circles high in the air.*  *Fig.12.4*

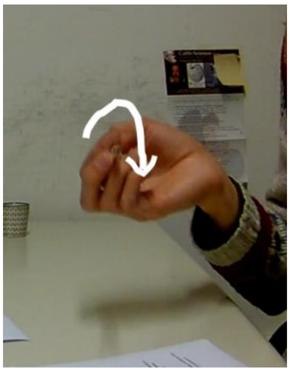
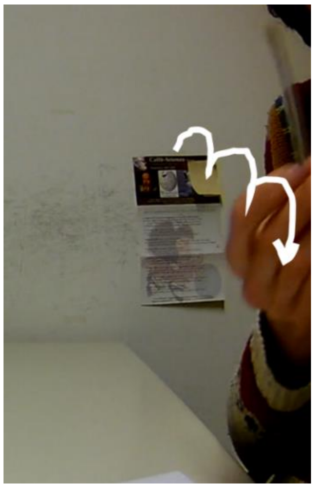
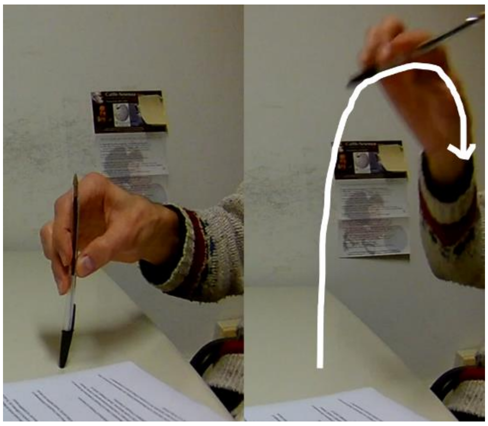
6	**S**	And therefore, practically you have all of them.	

The excerpt shows Silvio's construction of the explain induction process in direct form. Silvio describes a series of application of modus ponens starting from the base of the induction and the inductive step. At line 2 he produces a LS connecting P(0) with P(1): with his speech he refers to one application of the modus ponens, while with his gesture he represents it as a metaphorical path from in the air. The same thing happens at line 3 and 4 with the second and third application of the modus ponens. Then, at line 5 he produces an IS: the bundle composed by the words "etcetera, etcetera" and the gesture in the air. In this IS, we can see how the gesture enriches Silvio's words which alone only refer to a repetition. The gesture, instead, represents a series of step (modus pones) each ending where the successive one starts. In other terms the gesture itself is representing the chaining process involving several modus ponens.

Excerpt 12.2

Silvio is answering to the questions called 'meta questions on MI' in the methods section. In particular after having read a (correct) proof by induction made by Matteo, an invented student, he reads the affirmation of another student, Marco, claiming that a n=2357 could be a counterexample for the statement.

	Who	Speech	Gesture and inscription
1	**S**	It cannot be something half. In the moment in which it is true for one, then it holds for the second, and the third, and the fourth, etcetera, etcetera.	*Keeping the right hand perpendicular to the sheet, he moves in high in the air from left to right shaping some arcs.*  *Fig.12.5*
2	**S**	But there cannot be an interruption. There cannot be…	
3	**I**	Why?	
4	**S**	Eh, because if for n it is not true, if for n… what is it here? I mean…	

5	**S**	If for k it is not true, it would not be true for the n minus one either,	*He draws an arc in the air moving the hand from the right to the left.*  *Fig.12.6*
6	**S**	and so on… backwards.	*He rapidly draws a series of arcs in the air moving the hand from the right to the left.*  *Fig.12.7*
7	**S**	And so, this would contradict the first one.	
8	**S**	Thus, if it was not true in a certain point, consequently it would not be true for the first one.	*He touches the table on a point with the pen and then, starting from this position he draws, an arc in the air from the right to the left.*  *Fig.12.8*

Silvio explains why, after having proved the base of the induction and the inductive step, there cannot be a number greater than the base for which the proposition is false. In line 1, Silvio justifies this by constructing an explain induction process with direct form. As in the previous excerpt Silvio represents the chain of MP starting from the base and reaching the second number after the base, then the third, the fourth, and "etcetera, etcetera". In the successive lines, however, Silvio justifies again why the proposition cannot be false for a number greater than the base, but this time he constructs an explain induction process in indirect form. Firstly, at lines 4-5, he describes an application of the MT: "if for k it is not true, it would not be true for the n minus one either". With his words Guido is referring to a generic number, as indicated by his use of the variable n (which was k in the first part of the sentence). The arc-shaped gesture is analogous to the ones used when referring to the MP, but this time the hand moves from the right to the left, indicating the transition from n to n-1, coherently with the logical structure of MT. At line 6-7, Silvio describes the successive application of MT each chained to the following one to conclude that the existence of a number greater that the base for which the proposition is false would contradict the base of the induction: "this would contradict the first one". He produces an IS at line 6. With his words he refers to a backward repetition ("and so on… backwards"), while he is drawing some arcs in the air, a gesture which iteratively repeats the arc-shaped gesture produced when describing the first application of MT. Thus, the bundle made of his speech and his gesture contains Silvio's description of the application of a series of MT. As observed also in the previous excerpt, the gesture contributes to represent the chaining of these logical inferences: the fact that the conclusion of one inference corresponds to the premise of the successive one is metaphorically represented by the fact that the ending point of an arc corresponds to the starting point of the following one. At line 8, Silvio summarises the whole argument: "Thus, if it was not true in a certain point, consequently it would not be true for the first one". Whilst saying this, he performs another arc-shaped gesture in the air, still from the right to the left. With this last gesture, Silvio seems to metaphorically represent the chain of MT itself as a unique path in the air. In this line, the IS used to represent the chaining of modus tollens in the previous line has evolved into a LS connecting the start and the end of the chain of MT (respectively a generic number k for which the proposition is false and the base of the induction).

Both the just analysed excerpts have shown how a semiotic lens could be used to observe and analyse the explain induction processes of both forms. In particular, the gestures themselves contribute to represent the logical structure of the student's argumentation. As observed the iterative chaining of syllogisms were metaphorically represented by Silvio's gesture. Moreover, the direction of the gesture (from the left to the right or the opposite one) was used by Silvio to represent the different direction of the iterations (from the base of the induction to greater numbers, or from a generic number to the base).

12.1.2 Traces of difficulties in intuitively accepting the inductive step.

The interviews were also analysed focusing on the student's intuitive acceptance of the proof by induction. In particular, in the last part of the interviews questions asking to describe if there were any unclear aspects of MI or asking if a proof by MI was considered as a convincing proof. All of the interviewed students stated to be convinced by a proof by MI and that there were not any unclear aspects of it. However, in some cases the students state to remember some difficulties to accept MI the first times it was presented to them, in particular in relation to the inductive step. In the following two experts I will present two examples in which two students, when remembering their first experiences in proving by induction, refer to some difficulties which can be interpreted as a not complete intuitive acceptance of the inductive step.

Excerpt 12.3

Valentina, when asked by the interviewer, stated that the first times she encountered MI she did not understand it ("At the beginning I didn't understand"). The interviewer, thus, asks Valentina to try to remember if there were any specific aspects that she could not understand.

	Who	**Speech**	**Gesture and inscription**
1	**V**	I remember it was not clear why it was true…	
2	**V**	I mean… why one could… one could suppose…	*She moves the right hand closing it towards her body as miming the action of bringing an object close to her body.*  *Fig.12.9*
3	**V**	why it was enough to prove something for an n….	*She makes a rapid arc-shaped gesture in the air rotating the right hand from the left to the right keeping the pointing finger and the thumb at a constant distance.*  *Fig.12.10*
4	**V**	Let's say the inductive step… it was not very clear to me.	*She repeats the gesture of line 3 rapidly two more times.*
5	**I**	Ok, Ok. It is a bit critical thing, indeed. And, in your opinion which are the main difficulties? I mean for someone who doesn't know it.	

6	**V**	Eh, I think the fact that they explain it during the first lessons…. During the first lessons of mathematics and so you are not used to this abstract way of reasoning yet, and you have this obstacle…	
7	**I**	Mhm…	
8	**V**	I mean, it is a very mathematical thing, this thing of to suppose something true and then to make some reasonings starting from that supposition.	*[suppose something]. She puts the two hands as for holding something in from of her to the left.* *[make some reasonings]. Starting from the previous position, she moves both hands from the left to the right shaping an arc in the air.*  *Fig.12.11*
9	**I**	Ok.	
10	**V**	If you are not used to this way of reasoning, at the beginning it can look a bit weird.	

Valentina says to remember that at the beginning it was not clear to her why a proof MI "was true". In particular, as described in lines 2-4, Valentina remembers to have had some difficulties in accepting the validity of the inductive step. She initially describes the inductive step by referring to the assumption of the inductive hypothesis, metaphorically represented by the gesture of line 2 miming the action of taking an object and hold it with one hand. Then, in line 3 and 4, Valentina describes again the inductive step, this time with some arc-shaped gestures in the air. The interviewer asks Valentina what specific aspects of the inductive step could be the most difficult to understand. Valentina firstly (line 6) answers by saying that the proof of the inductive step (called "this abstract way of reasoning") is something new and not familiar for students and therefore could be problematic to understand. Then (line 8), Valentina better clarifies what she meant in the previous line: she says that the inductive step requires "to suppose something true and then to make some reasonings starting from this supposition". This utterance is co-time with a metaphorical gesture: firstly, the two hands are put as to hold an object, a gesture which seems to metaphorically represent the initial supposition (i.e., the inductive hypothesis), and then they are moved to her right side shaping a path in the air, metaphorically representing "some reasonings starting from the supposition" (i.e., deducing P(n+1) from P(n)). In this line, therefore, Valentina is using some multimodal signs to describe and represent the process of proving an inductive step. In line 8, Valentina concludes that "this way of reasoning" is something that "can look a bit weird".

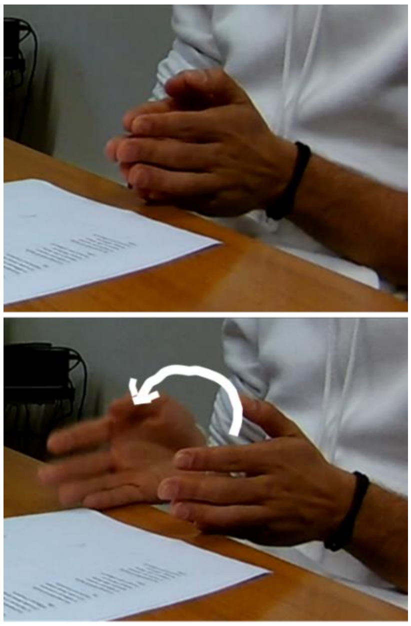
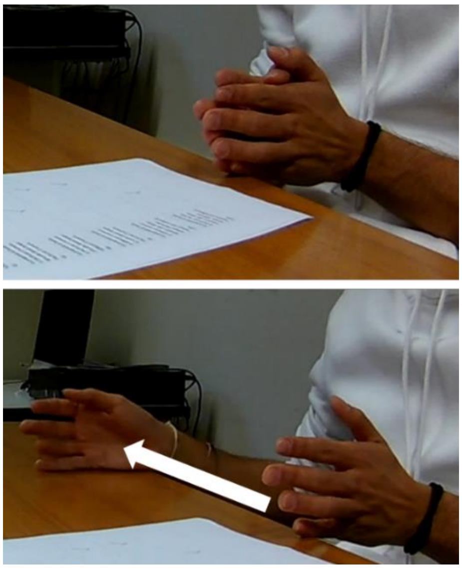
As said when analysing the intuitive acceptance of the inductive step, one of the crucial aspects is to accept its validity independently from knowing the truth value of P(n) and P(n+1). In this excerpt, Valentina seems to refer exactly to this point. Indeed, she seems to say that a difficult part of a proof by induction is that in the inductive step we need to prove P(n+1) assuming P(n) without knowing, within the proof of the inductive step itself, if P(n) is actually true or not. Valentina says that the inductive step is an "abstract way of reasoning". The word 'abstract' seems to refer exactly to what just said: the inductive step should be accepted as valid independently from the truth value of P(n). In other terms, the fact that the proof of the inductive step is valid even if we do not know if P(n) is true or not, makes appear this proof, for Valentina, as an "abstract way of reasoning".

In the following example a similar situation is shown. As we will see, another student, Claudio, as it happened with Valentina, will refer to his initial difficulties in accepting the proof of the inductive step. However, if Valentina seemed to have intuitively accepted the validity of the inductive step as an independent theorem, for Claudio the situation seems to be a bit different. Claudio, in fact, will justify the validity of the proof of the inductive step, in particular of the assumption of the inductive hypothesis, by saying that the base of the induction guarantees the validity of this assumption. Claudio's excerpt, therefore, will show an example of what I have called an interference between the two theorems corresponding to the base case and the inductive step (See 3.3.3). In particular the validity of the proof of the inductive step is accepted if together with the base of the induction.

Excerpt 12.4

In the last part of the interview Claudio is asked to describe "why a proof by induction assures the truth of the proposition for all the natural numbers".

	Who	**Speech**	**Gesture and inscription**
1	**C**	Because…. What does it do? It says to you: if this proposition is true for n minus	*He puts the two hands together perpendicularly to the table and then he moves the right one to the left shaping an arc in the air.*

		one, then… let's assume for a natural number… then it holds for the successive one.	 *Fig.12.12*
2	**C**	Therefore, as a consequence, from this point on, it will hold for all of them.	*He puts again the two hands together perpendicular to the table. Then he moves the right hand to the right, dragging it on the table.*  *Fig.12.13*
3	**C**	And if I prove that it holds for one… as I said before for Evelin[94], it will hold for all the others as well.	*He repeats twice the same gesture of line 2.*

[94] In the previous part of the interview, Claudio was involved in solving the task on Evelin's property taken from the survey, for which he answered correctly to every question.

After this, the interviewer asks to Claudio if there are any unclear or not convincing aspects of a proof by MI.

[…]			
4	**C**	Ok…let's say that at the beginning, the first time I approached to induction it looks like… That formula that you, as a student, want to deny because you say: no, it is not possible…	
5	**C**	I mean… it cannot be valid, it is impossible. But then, after some time, by doing it and by reasoning on it…. Surely it avoids a lot of proofs, I mean it helps a lot.	
6	**I**	Ok. In your opinion, this is a difficult question, why the first time that you saw it, or the first times, did you have this feeling of no, no I do not want it?	
7	**G**	Eh, because… let's say… you say: it holds for n minus one, then it holds for n. Yes, but who says to me that it holds for n minus one?	
8	**I**	Ok	
9	**G**	But then you say: eh, no, there is the inductive base! At this point, more or less, you understand.	

In the first three lines of the excerpt, Claudio describes why a proof by MI 'works'. Firstly (line 1) he refers to the inductive step, described with the utterance "if this proposition is true for n minus one, […] then it holds for the successive one", where n is a natural number, together with an arc-shaped gesture over the table, namely a LS. Then (line 2), Claudio says that "as a consequence, from this point on" the proposition will hold for all the natural numbers. In saying this he performs an interesting gesture: starting with the two hands in the same position of when describing the inductive step, he moves the right hand to the right dragging it on the table. With this gesture Claudio seems to refer to the iterative application of the inductive step to reach every natural number greater than the starting point, therefore this is a LS. Interestingly, at this line he has not mentioned the case base yet. He seems to anticipate the consequence of the inductive step, once the base case will be proved. In the successive line, in fact, he says that "if I prove that it holds for one […] it will hold for all the others as well", repeating twice the gestural IS of the previous line. Claudio's description of the functioning of MI is complete. He has constructed an explain induction process, as his gestures of lines 2 and 3 highlight. Moreover, from what said in the last line, he seems to be aware of the necessity of the case base for starting the iterative application of the inductive step. In conclusion, Claudio's description of the functioning of a proof by Induction seems to be correct.

In the second part of the excerpt, however, when asked by the interviewer about some unclear or not convincing aspects of MI, Claudio admits that the first times he experienced it, he was not convinced of its validity: "it looks like… That formula that you, as a student, want to deny because you say: no, it is not possible […] It cannot be valid, it is impossible" (lines 4-5). By saying this, Claudio is revealing that MI was not intuitively accepted by him. The interviewer, thus, invites Claudio to clarify this point asking the reason why he was having this feeling of "it cannot be valid". Claudio's answer (line 7) explicitly refers to the inductive step, in particular to the assumption of the inductive hypothesis: "you say: it holds for n minus one, then it holds for n. Yes, but who says to me that it holds for n minus one?". Claudio is remembering that he was not intuitively accepting the proof of the inductive step since it was not clear to him why one could assume the inductive hypothesis. In the successive line, Claudio explains the justification that he, after some time, gave to himself for the validity of assumption of the inductive hypothesis: "but then you say: eh, no, there is the inductive base!". In other terms Claudio seems to say that the inductive base guarantees the fact that one can assume the inductive hypothesis. We can assume P(n) true for a fixed generic natural number since we know, by the inductive base, that $P(n_0)$ is true. In a certain sense, Claudio is saying that the proof of the inductive step requires the case base. This is an example of an interference between the case base and the inductive step. After having noticed this, it is interesting to come back to the first part of the excerpt, when Claudio was describing the functioning of a proof by induction. In particular, in line 1 and 2 he has described it referring to the inductive step (line 1) and saying that "from this point on, it will hold for all of them", without any reference to the base case. When saying "from this point on" therefore, he seems to refer to the number n-1 (involved in the inductive hypothesis) as the starting point of the iteration. This is also highlighted by the fact that the gesture of line 1 (the LS referring to the inductive step involving n-1 and n) and the gesture of line 2 (the IS referring to the iterative application of the inductive step) start from exactly the same point on the table with the hand in the same position. In other terms, in this description, Claudio seems to be overlapping the number n-1 of the inductive hypothesis with the base of the induction, which is coherent with the fact that later he says that the inductive hypothesis can be assumed since "there is the inductive case". It seems like Claudio is saying that, when proving the inductive step, the inductive hypothesis corresponds to the base case so that we assume something which we know to be true. To conclude, Claudio seems to have overcome his difficulties in intuitively accepting the validity of the proof of the inductive step as independent from knowing the truth value of P(n-1) and P(n), by overlapping the number involved in the base case with the 'n-1' of the inductive step for which, then, a truth value of the proposition is indeed known. To say it with different words, with the interference between the base case and the inductive step, Claudio seems to have accepted as legitimate the assumption of the inductive hypothesis. It seems like for him the base of the induction, necessary for the meta theorem, has become necessary for proof of the inductive step as well.

12.2 OVERALL CONCLUSIONS OF THE SECOND EMPIRICAL STUDY

In this and the previous two chapters, I have presented the results emerging from the analyses of the interviews, carried out within a multimodal semiotic perspective. To better describe the findings of this empirical study, they presentation has been divided in three chapters.

Chapter 10 focused on two particular categories of signs that have been identified and studied: the LS and IS. The analysis of the use and production of these signs by students involved in the generation of recursive argumentations and in the construction of proofs by induction allowed us to observe a series of interesting aspects.

First of all, these signs seem to characterise the generation of a recursive argumentation during the student's exploration of a problem. In other terms, the production of LS and IS highlights those crucial moments in the exploration of the problem for the generation of a recursive argumentation to support a conjecture. Thus, within this perspective, LS and IS can be read by an external observer (a researcher or a teacher, for instance) as signs of a recursive argumentation emerging from the student's exploration of a problem.

Secondly, with the analysis of the interviews, it was also possible to observe that LS and IS are not just signs which provide a "window" for the researcher on students' reasonings, but they can also become resources for the students themselves. The use and production of these signs, in fact, seem to have an active role in supporting students in the construction of a recursive argumentation or of a proof by induction. Sometimes a student recognises in a previously produced LS the structure of an inductive step and, starting from it, s/he succeeds in constructing a recursive argumentation or the inductive step itself within a proof by induction. Some other times, instead, a student anticipates, through the production of LS and IS, the structure that a recursive argumentation should have and then s/he tries to reconstruct such an argumentation for the problem under investigation. Moreover, I have also registered that the suggestion of a LS made by an external intervention, in some cases, might help a student in developing a recursive argumentation or a proof by induction. For this last point, we observed that a crucial aspect for the effectiveness of the intervention is the interpretation that the student gives to the suggested sign.

Moreover, I showed that the observation of LS and IS can become a methodological instrument with which to register the possible students' cognitive unity between argumentation and proof, with a focus on MI. In particular, in fact, the production of those signs during the resolution of a problem could occur both during the phase of construction of the argumentation and the phase of construction of a proof. This can happen, for instance, when a student produces an analogous LS when developing one step of the recursive argumentation and when constructing the inductive step of the proof by induction. Otherwise, it is also possible that a student, when describing the structure of the argumentation s/he is constructing, produces (linking and iteration) signs which then are repeated when describing the structure of the proof by induction used to formalise her/his argumentation. In both cases, the observation of this semiotic continuity might reveal a student's cognitive unity between her/his argumentation and the subsequent proof by induction.

In chapter 11, a different kind of analysis of the interview was presented. In particular, I showed how, focusing on students' semiotic production, it was possible to distinguish signs which can be interpreted as belonging to two different categories: ground level signs and meta level signs.

This distinction, first of all, allowed me to identify some crucial moments in the problem resolution in which the subjects seem to reflect on the problem itself or on the resolution process. In those moments the subjects seem to pause calculations, symbolic manipulations, geometrical transformations, etc., (highlighted by a ground level semiotic production), in order to observe and reflect on the structure of the problem or of their own argumentation, or to describe what they have done, or to describe what they are planning to do (highlighted by a meta level semiotic production). These moments seem to be characterised by a rich dialectic between ground and meta level signs. In order to better describe and analyse this dialectic process, within the theoretical lens of the semiotic bundle, a series of notions have been introduced (liftings, projections, merging points and level's mismatches), together with a descriptive tool (the two lines diagram) used to model the semiotic analysis of an excerpt. This development of the theoretical lens allowed me to register how, for the expert students involved in the study, the two levels of semiotic production seem to be extremely permeable. With this I mean that during the problems' resolution the students seem to move their discourse smoothly on the two levels, with a series of liftings or projections. However, it was also possible to observe that some cases of less expert students in which these reflective moments seemed to be missing, were characterised by an almost uniquely ground level semiotic production. Within this perspective, thus, meta level signs can become sort of indicators for an external observer to investigate the reflective moments during the resolution of a problem.

Analogously to what already said for LS and IS, the dialectic between ground and meta level semiotic production does not have only a communicative role for the subjects involved, but it can become a resource for the subjects themselves supporting them during the problem resolution. In particular, I observed that this dialectic could play an important role in supporting a subject in the generation of a recursive argumentation and the construction of a proof by MI. Firstly, meta level signs, in fact, seem to guide students in recognising that the problem can be solved with a recursive argumentation. Specifically, the fact that meta level signs can have a potential generality, since they refer to the specific problem under resolution and they could also be used to refer to other problems, might support students in recognising that a specific problem has some common aspects with other problems that were already solved with a recursive argumentation. Moreover, the use and production of meta-level signs could also support students in the construction of a proof by induction. These signs, in fact, can be used by the subject to describe and plan the structure of a such a proof before constructing it with reference to the specific objects of the problem. In this case, thus, the construction of a proof by induction can be obtained as a projection of the structure of a generic proof by induction, expressed on the meta level, onto the ground level in which it is reconstructed with the problems' objects.

Furthermore, focusing on the use and production of meta level signs allowed me to deeper investigate the students' cognitive unity between recursive argumentation and proof by MI. As written above the use and production of LS and IS in different phases of the problem resolution could contribute to highlight the possible cognitive unity between argumentation and proof. This aspect can be further enriched if we focus specifically on meta level signs. We observed that similar meta level LS and IS are used and produced by students to refer to the structure of a recursive argumentation and to the structure of a proof by induction. This continuity on meta level signs can therefore highlight the presence of a structural continuity between recursive argumentation and the successive proof by induction. In other terms, a continuity on meta level LS and IS could be seen as a trace of the fact that the student is recognising a similar structure between recursive argumentations and proofs by MI.

Finally, both the analyses focusing on LS and IS, and on ground and meta level signs allowed me to register a further aspect which relates specifically to the use of gestures in constructing recursive argumentations. As said in the presentation of the conceptual framework, recent research (Arzarello & Sabena, 2014; Krause, 2015; Sabena 2018) highlighted how gestures could actively contribute to support students in structuring their argumentations. The analyses presented in this chapter offer a confirmation to these results. In particular, we observed that the use and production of LS and IS, both ground and meta level, as well as other meta level signs, seem to actively support students in the generation of a recursive argumentation and in the construction on a proof by induction. Several of these signs, as it was shown, are gestures or a bundle of them with other semiotic sets. To summarise, thus, in this study it was possible to register how the production of gestures does not seem to have only a communicative role for students (which is, of course, present), but it also seems to support them in planning, organising, and structuring a recursive argumentation or a proof by induction, both on the level of the argumentative or proving scheme (meta level) and on the level of an argumentation or proof for a specific problem (ground level).

Finally, in the first part of this chapter, I presented other results emerging from the interviews which offered elements to investigate the other research questions formulated outside of the theoretical semiotic perspective.

Firstly, I presented two examples of a student's reconstruction of the explain induction process both in direct and in indirect form. In the development of the conceptual framework, when referring to APOS theory, I proposed modification of the GD of MI. In particular, I observed that within student's Schema of MI, together with the explain induction process in direct form, there could also be the explain induction process in indirect form. The presented examples offer an empirical confirmation of this point, showing how the first process was constructed by a student to justify why a proposition proved by induction is true for every number greater than the base, whilst the second one was constructed to justify why the same proposition cannot be false for a given number n* greater than the base. The semiotic perspective, with the development that it had in this chapter with the definition of LS and IS, provided a lens through which to observe these two processes. In particular, as we noticed, the analysis of some gestures allowed us to observe the subject's processes of chaining of inferences (both modus ponens and modus tollens) within the explain induction process.

Lastly, two other episodes were presented, referring to the issue of the intuitive acceptance of the proof by MI. Even if all of the interviewed students stated to be fully convinced by a proof by MI and that for them there were not any unclear aspects about it, it was still possible to register answers referring to some sort of difficulties in accepting the validity of MI, in relation to the inductive step, as the two presented. examples have shown. In particular, in the first one, the interviewed student described the proof of the inductive step as an "abstract way of reasoning", remembering some difficulties in accepting it the first times. In the second episode it was shown the presence of an interference between the inductive step and the base of the induction, with the latter which seems to legitimate for the student the assumption of the inductive hypothesis within the proof of the inductive step. The two episodes, even if different, confirm the results of the survey and of other research (Fischbein & Engel 1989) showing that the acceptance of the validity of the inductive step as independent from knowing the truth value of its antecedent and consequent could be rather problematic for students.

13 GENERAL CONCLUSIONS

In this final chapter I will discuss the conclusions of the thesis, explaining to what extent the study conducted led to findings in relation to the research questions. In the following sections, I will answer them, highlighting the main contributions that this study offers, I will describe some didactical implications and, finally, I will outline directions for further research that could stem from it.

13.1 ANSWERS TO THE RESEARCH QUESTIONS

13.1.1 Answer to R.Q.1

R.Q.1. What mathematical aspects characterised the historical genesis of the proof by MI? In particular, what characteristics and turning points emerge from the traces of proofs by MI that the historiographic research has identified?

The answer for this question has already been discussed in section 2.9. This was necessary since the results emerging from the historical-epistemological analysis answering the question were used to structure some elements of the successive conceptual framework. Here I will just report the following points summarising the series of characteristics highlighted by the analysis of the historical traces of MI:

(A1) <u>Natural numbers as aggregations or as progressions</u>

(A2) <u>Potential infinity or actual infinity</u>

(A3) <u>Iteration</u>

(B1) <u>Absence or presence of the parameter</u>

(B2) <u>Generic example as a general case</u>

(B3) <u>Infinite proof or finite proof</u>

(B4) <u>Inductive step as an independent proposition</u>

(B5) <u>Increasing or decreasing iteration</u>

(C1) <u>Statement with an implicit or explicit recursion</u>

(C2) <u>Not necessity for trivial cases</u>

13.1.2 Answer to R.Q.2

R.Q.2. To what extent do students construct the Explain Induction Process, both direct and indirect form?

This research question has been principally investigated with the first empirical study presented in this thesis (Chapter 9), the online survey with university students, specifically in the tasks concerning situations in which, the inductive step is said to be valid and some truth or false values relating to certain natural numbers different from 0 and from 1 are known. The tasks were formulated on a ‘meta’ level, i.e. involving some generic and unknown numerical properties, coherently with the aim of investigating students’ construction of the explain induction process when involving a generic predicate P(n).

Before discussing the results of the survey related to this R.Q., it is important to observe that the presence of what can be interpreted as a trace of explain induction process (both in direct and in indirect form) has been registered also in the other studies composing this thesis (the historical-epistemological analysis of MI and the interviews with expert students) when a justification of the validity of MI or of a recursive argumentation is constructed. In particular, as I have described in the analysis of the interview related to this point (section 12.1.1), expert students used autonomously argumentations which highlight the construction of the Explain induction process (both direct and indirect) to justify the validity of MI and to contradict the presence of possible counterexamples for it. In other terms, it was possible to register empirical evidence of the fact that this process (the indirect one, as well) is an effective process for students for resolution of problems and for constructing argumentations. This aspect supports the assumption that the interiorization of the Explain induction process in both forms is an important step for the construction of the MI Schema in students and, consequently, my proposal to also include the presence of the explain induction process in indirect form in the Genetic Decomposition of MI within the APOS Theory.

In the GD of MI, the explain induction process is central and crucial for the construction of the Schema of MI; I recall indeed that, in APOS terms, the Object 'Proof by MI' can be constructed by a subject as an encapsulation of the explain induction process. On one side, with the study conducted in this thesis, it was possible to register the presence of some answers highlighting a complete interiorization of the explain induction process in both forms, offering an empirical confirmation for the presence of these processes in the GD of MI. On the other side, however, the study highlighted how the construction of explain induction process (in both form) could be rather problematic for students, even among those who have already encountered the MI during their previous studies.

In relation to this last point, the analysis of the results of the survey allowed me to identify and focus on some of these problematics aspects. I recall that all of them were registered transversely in every considered group of students, from the ones with no experience with MI to masters' students in mathematics, an aspect which suggests how these problematics could be deep and persistent in in students.

13.1.2.1 Explain induction as a process of Chaining of logical inferences

The tasks A, B and C of the survey were formulated in a way that a theorem corresponding to the inductive step for a property was said to be valid, and a number different from 0 or 1 was given for which it was known if the property was true or false. A first result that I registered is the fact that the percentages of correct answers decreased when the numbers involved in the questions were progressively more distant from the number for which a known true/false value was given. This result was registered transversely in every group considered in the analysis and in all the tasks of the survey involving questions of this kind. This aspect highlights first of all how, for some students, it seems not to be immediate to recognise that, starting from the inductive base and the inductive step, it is possible to construct a chain of MP to reach every number greater than the base. Moreover, it also shows that, with reference to the explain induction in indirect form, similar problematics seem to involve the construction of a chain of MT starting from a number n^* for which the given proposition is false. Supported by the qualitative analysis of the open questions at the end of the tasks, I interpreted this phenomenon in terms of interiorization of complex processes obtained as a coordination of other processes, that I named 'chaining of MP' and 'chaining of MT' (see section 9.2.1). A critical aspect emerging in the GD of MI presented in the framework was the transition from potential to actual infinity in the encapsulation of the explain induction process which allows the subject to recognise that the chain of MP, starting from the base n_0, reaches the whole set of natural numbers greater than n_0 (see 4.4.5). The results that I discussed, however, highlighted something more. Indeed, they showed that the construction of this chain of inferences could be problematic even when it involves only a finite number of steps.

13.1.2.2 Chaining of MP vs Chaining of MT

As said, similar problematics were registered in relation to the process of chaining of MP or of MT, i.e. when considering, respectively, the explain induction process in direct form or in indirect form. However, the data analysis allowed me also to register significant differences between the two processes. In particular, comparing the questions involving the chaining of MP with those involving the chaining of MT the second ones have always obtained a lower percentage of correct answers that the first ones. This aspect was registered in every group. A wide literature has highlighted that MT is more problematic from a cognitive point of view than MP (Fischbein, 1987, pp. 72-81; Antonini, 2004; Inglis & Simpson, 2008). First of all, thus, the results of the survey offer a confirm for these studies. Moreover, the results also show that the Chaining of MP/MT seem to present the analogous problematic. In every group considered, not all the students who have constructed the Chaining of MP, also seemed to have constructed the Chaining of MT, even if they can apply correctly one MT. This result shows that the Chaining of MT could be more problematic than the Chaining of MP. In other terms, the greater difficulty of MT in respect to MP does not end with the application of an inference but it involves the chaining process. As a further confirmation of that, we can consider that in the analysis of both the open and closed questions of all the participants, I have not found any answer in which the Chaining of MT seems to be present without the presence of the Chaining of MP as well, whilst, as said, the opposite situation was registered instead. What has been recorded, therefore, shows how the construction of a chain of MT can be problematic even for students who can construct a chain of MP and who can apply single MT steps. This aspect therefore suggests that although from a logical point of view the Chaining MT is analogous to the Chaining of MP, from a cognitive point of view, however, a subject's interiorization of the former does not automatically lead to the interiorization of the latter.

13.1.2.3 The 'meta' level of the explain induction process

A further result registered relates to the particular kind of the questions of the survey. In the a-priori analysis of the survey, I observed that the fact the tasks involved some generic and unknown properties could have added a further level of complexity since the questions were posed, in a certain sense, on a 'meta' level, whose objects were propositions and properties themselves. This choice was coherent with the aim of investigating on the explain induction process. Indeed, when I presented the GD of MI within the APOS Theory (section 4.3), I observed that the Explain Induction process is the process by which a subject constructs a justification of the functioning of the MI both for a specific predicate and for a generic predicate P(n). Thus, for a complete construction of the MI Schema, a subject must be able to construct this process even for a generic predicate P(n). With reference to this point, the results of the survey registered the difficulty (perceived or not) of some students in expressing themselves and reasoning when the objects were generic (and unknown) predicates. Indeed, we registered that some students (among the master's students in mathematics as well) answered "We cannot know" to some or even all the questions of the survey claiming that, since the involved properties were unknown, it was not possible to answer the questions.

In conclusion, it is possible to summarise what just said by observing that this thesis offers two main results in relation to the R.Q.2:

First of all, it was possible to register that some students, when providing a justification of the functioning of a proof by induction construct the explain induction process which sometimes is in direct forms and some other times in indirect form. This aspect was registered both in some answers in the survey and in the task-based interviews. This aspect supports the assumption that the interiorization of the Explain induction process in both forms is an important step for the construction of the MI Schema and thus the proposal of also including the Explain induction process in indirect form in the GD of MI.

Secondly, however, it was also possible to register that the construction of the explain induction process, in both forms, could be rather problematic for students. The results of the survey, just summarised, showed how, despite a higher number of correct answers in the groups of students with more experience with MI, we have still registered a relevant number of not correct answers in every group, in master's students in mathematics as well. This suggests that some problematics aspects related to the Explain induction process (the chaining of inferences or the fact that it may involve a generic predicate P(n)) could be particularly robust and could persist even in expert students. This aspect is delicate from a didactic point of view. Indeed, generally, in teaching, when MI is introduced to students, its functioning is explained by explicitly constructing the first steps of a chain of MP inferences starting from the base of the induction to then conclude with a "and so on". How this study has shown, however, the process which underlies this "and so on" can be constructed by students in some very different ways, with chains of MP that do not reach every natural number, that stop after a few steps, or that do not even start.

13.1.3 Answer to R.Q.3

R.Q.3. To what extent a proof by MI, intended as a system of theorems (S1, P1, T), (S2, P2, T) and ((S1∧S2)→S, MP, MT), is intuitively accepted by students? Moreover, which aspects can influence its intuitive acceptance?

Students' intuitive acceptance of the proof by MI has been investigated mainly with the online survey and with the questions of the interviews. The results discussed in the previous chapters provided a clear indication on this point, highlighting that the proof by MI could be seen as not intuitive by students. Indeed, even if for some students, this proving scheme and its justification was described as obvious and self-evident, for some others this was not the case. First of all, this has been registered in the results of the survey, both in the conclusive open answers, in which some students stated themselves to have doubts about the validity of MI and opened the possibility of counterexamples for "very big numbers", and in the answers to the tasks where some students, given the validity of the inductive step and the truth of a predicate on a number n_0, did not conclude the validity of the predicate for every number greater than n_0. Secondly, it was also possible to register aspects related to a not intuitive acceptance of MI in the interviews. As shown in section 12.1.2, in fact, even if all of the interviewed students stated to be convinced by a proof by MI and that there were not any unclear aspects, however, in some cases the students state to remember some difficulties to accept MI the first times it was presented to them.

In this thesis, the registered problematics related to the intuitive acceptance of MI have been further analysed with reference to the model of proof by MI as a system of three theorems (S1,P1,T), (S2,P2,T) and (MS,MP,MT), in line with the R.Q.3.

13.1.3.1 Intuitive acceptance of the Meta-theorem

By the analyses of the results of the survey, it was possible to register some problematics for students in accepting as obvious and certain the fact that from $P(n_0)$ and $\forall n(P(n)\rightarrow P(n+1))$ it necessarily follows $\forall n\geq n_0 P(n)$, i.e. in other terms what I interpreted as intuitive acceptance of the meta-theorem. I recall that the intuitive acceptance of the meta-theorem was investigated in the survey by providing the validity of the statements S1 and S2, where S1 was modified involving numbers different from 0 or 1, and by registering if students could reconstruct the meta-theorem in these cases. The assumption was that an intuitive acceptance of the meta-theorem allowed to correctly answer the questions posed and thus the deviations from the correct answers could give indications related to its intuitive acceptance. As already observed in relation to the previous R.Q., the results of the survey clearly highlighted that the passage from $P(n_0)$ and $\forall n(P(n)\rightarrow P(n+1))$ to $\forall n\geq n_0 P(n)$ can be not intuitive for students, also for those who already encountered MI during their studies. In relation to this point, it is worth mentioning that very low percentages of correct answers to all numerical questions of the tasks were recorded in the survey (in the group formed by master students in mathematics, for instance, only 50% correctly answered all the numerical questions of the survey)[95]. The fact that several non-correct answers were registered also in students with several experiences of studying and constructing proofs by MI, can be well interpreted in terms of intuitive knowledge following Fischbein (see 5.2): these students, who possibly have a formal knowledge of MI, do not possess an intuitive knowledge of it. Indeed, in the survey, in the moment in which some aspects of MI were modified in relation to how generally it is encountered (the shifting of the base to a number distant from 0 or 1, or the absence of the case base) these students could not re-construct the meta-theorem for these different cases.

A further aspect related to the acceptance of the meta-theorem is the fact that this theorem involves a generic predicate P(n). As already observed in relation to the previous R.Q., the survey registered the problematic, for some students, in dealing with these 'meta' aspects involved in the questions. This point offers a contribution also in relation to the intuitive acceptance of the meta-theorem, highlighting that some students does not seem to accept the fact that the meta-theorem could be valid without knowing the specific predicate P(n) it involves. In other terms, the fact that the meta-theorem involves a generic and unknown predicate could influence its intuitive acceptance by students. This is a delicate aspect to take into consideration, since generally, in teaching, MI is presented to students in terms of the meta-theorem whereas the activities in which students are involved with MI are often only limited to the application of a ritual, whose difficulties are reduced to the technical level of manipulation of mathematical symbols (algebraic expressions, sums, indexes, etc.). As a result, MI risk to become a ritual procedure for proving statements which is not accompanied with the feeling that it intuitively 'works' as a proving scheme, i.e. with an intuitive knowledge of MI, in Fishbein's terms. This point will be further expanded when the didactical implications of the study will be discussed (section 13.2)

13.1.3.2 Intuitive acceptance of the inductive step

The study conducted in this thesis allowed me also to register some problematics related to intuitive accepting the inductive step. The results of the survey highlighted, firstly, the difficulty of several students in recognising the validity of the implication involved in the inductive step in itself, without any concern whatsoever for the truth of antecedent and consequent. This was registered in some answers of students stating that since it was known that for a number n*, P(n*) was false, then necessarily one between P(n*-1)→P(n*) or P(n*)→P(n*+1) must be invalid. This result offers a confirmation to the study made by Fischbein and Engel (1989) on this aspect with secondary school students, showing moreover that this problematic could resist also students with more mathematical experience (in master's students as well). Moreover, from the results of the survey a further aspect was also observed, which is the fact that some students did not recognise the inductive step as theorem itself. In particular, when an inductive step was presented alone, without a truth/false value for some n_0, as in the Task D, we registered that for some students the inductive step was interpreted as an 'empty' statement from which it is not possible to infer any conclusion regarding the involved property (indeed these students selected "we cannot know" for every statement of the task, without recognising that some of them were in contradiction with the validity of the inductive step). Furthermore, some problematic aspects related to the intuitive acceptance of the inductive step were registered in the analysis of the interviews as well. In particular, in 12.1.2, I showed two examples of undergraduate students remembering some difficulties in accepting the validity of the inductive step the first times they encountered MI. In the first example, the interviewed student described the proof of the inductive step as an "abstract way of reasoning", perceiving it as different from what it is generally done in other proofs. In the second episode an interesting aspect was registered: the student justified the validity of the proof of the inductive step, in particular of the assumption of the

[95] The complete results in relation to this point were reported in the Table 9.25 in section 9.2.5.

inductive hypothesis, by saying that the base of the induction guarantees the validity of this assumption. This example, therefore, showed the presence of what I have called an interference between the theorems corresponding to the base case and the inductive step (See 3.3.3), highlighting that it is possible that students accept the validity of the inductive step only when it is paired with the presence of a valid base case. In other terms it is possible that the intuitive acceptance of the validity of the inductive step is influenced by the presence of a number for which the proposition is known to be true, despite this is not necessary from a logical point of view. The registered difficulties in accepting the validity of the inductive step (and, specifically of an implication) without any concerns for the truth of antecedent and consequent opens an important issue from a didactical point of view. From one side, the proof by MI is based on the fact that one can prove an implication independently from knowing the truth of the statements involved. Students' intuitive knowledge of this aspect seems to be a prerequisite for their understanding of MI. On the other side MI is perhaps one of the most (and first) significant cases in which the process of proving the validity of an implication independently from knowing the truth of antecedent and consequent could be lived by students. Therefore, this results in a sort of vicious circle which could be problematics for students' construction of a productive and effective knowledge of MI.

13.1.3.3 Intuitive acceptance of the base case

Lastly, it was also possible to register some problematics related to the intuitive acceptance of theorem corresponding to the base case, specifically on its role in a proof by MI. In particular, the results of the survey highlighted two different aspects. The first one is the difficulty of some students in recognising that the number corresponding to the base case does not have to be necessarily 0 or 1. This aspect, besides being highlighted by the presence of non-correct answers in the tasks in which a base different from 0 and 1 was given, was also registered with a specific question in which several students answered "I agree" in relation statement claiming that in every task of the survey the base case was never present and thus it was not possible to conclude anything from what was stated. These answers highlight that for several students the given numbers (104 for task A, 75 for Task C) did not represent a base case. In other terms, these students do not seem to accept a statement as "the base case" if it does not involve 0 or 1. The second problematic aspect related to the base case registered in this study is the fact that for some students, the base case does not seem to be a necessary component for the validity of proof by MI. This was registered both in some open answers at the end of the survey, where the proof by MI was described only in terms of the inductive step, and in some answers of the task in which the validity of inductive step was presented alone, without a truth/false value for some n_0 (Task D) in which the students stated that as a consequence of the theorem (i.e. the inductive step) the proposition is necessary true for every natural number. These two observed aspects confirm other studies in literature reporting students' difficulties in accepting a base of induction different from 0 or 1 (Avital & Lebeskind, 1978; Stylianides et al., 2007) or registering that often proving the base is not recognised by students as necessary for a proof by MI (Ernest, 1984; Dubinsky & Lewin, 1986; Ron & Dreyfus, 2004; Garcia-Martinez & Parraguez, 2017, Larson & Petterson, 2018).

In conclusion, this study highlighted how the proof by MI could be rather problematic to be intuitively accepted by students, both in reference to the meta-theorem and to the base case and the inductive step. The model of a theorem by MI as a triplet highlighted a complexity of the proof by MI, in terms of relationships between the involved statements, proofs, theory and meta-theory. When the intuitive acceptance of MI has been described in terms of intuitive acceptance of this complex system theorems, I observed that a crucial aspect is that a student intuitively accepts them both as independent theorems and in mutual relationship (see 5.3). What this study has registered, thus, is how this complexity could be problematic for students for the intuitive acceptance of MI, showing that the construction of an intuitive knowledge of the proof by MI (as system of theorems) is far from immediate for students.

13.1.4 Answer to R.Q.4

R.Q. 4 – Is it possible to identify crucial signs involved in the generation a recursive argumentation in students' exploration of a problem? What contribution can the production and use of these signs provide to students' construction of a proof by MI?

This research question has been investigated with the task-based interviews whose analyses and discussion were presented in chapters 10, 11, and 12.

13.1.4.1 Linking and iteration signs

The adopted focus of the analysis of the interviews was on the students' production and use of signs involved in the generation and construction of recursive argumentations. As a result of this analysis, I identified two particular kinds of such signs: The Linking Signs (LS) and the Iteration Signs (IS). By investigating on the use and production of these signs in students involved in the generation of recursive argumentation and in the construction of proofs by MI, it was possible to register that LS and IS are crucial for the generation of a recursive argumentation and for the construction of a proof by MI. In particular, for these categories of signs a series of important aspects were registered.

First of all, LS and IS characterise the generation of a recursive argumentation during the student's exploration of the problem. In particular, the production and use of these signs might highlight a crucial *shift of attention* (Mason, 1989) for the student: the attention is moved from one or more cases of the problem as separated, to the link between these cases. In these moments, it is therefore possible to observe the generation of a recursive argumentation because it is precisely from this link between cases that such an argumentation can be developed. Within this perspective, LS and IS can be read by an external observer (a researcher or a teacher, for instance) as signs of a recursive argumentation emerging from the student's exploration of a problem. As observed, moreover, the presence of these signs can also be observed in the formal proofs by MI. In particular the construction of the inductive step can be interpreted as oriented to the construction of a LS connecting P(n) with P(n+1). In these cases, the LS is obtained through the use of some mathematical symbols (parentheses, nested sums, etc.). This is to say that also when considering the proof by MI, and not only the recursive argumentation, we can observe that an important part of this proof construction is an effective treatment of these (symbolic) LS.

An extremely important point to remember in relation to LS and IS is their multimodality. As shown in all the experts presented, It is through the whole semiotic bundle, with its different components and the different relations between them, that students produce and explore LS and IS. For this reason, it is from the analysis of the various components of the whole bundle with their mutual relations that an observer can register the presence of LS and IS in a student's semiotic production.

A further significant registered aspect is that LS and IS can also become resources for the students themselves during the whole problem resolution activity and, specifically, also during the proving process. The use, production, and exploration of these signs, in fact, often support students in planning, organising, and constructing a recursive argumentation or, more specifically, a proof by induction, both at the level of its general logical structure and of the specific problem involved. Sometimes a student recognises in a previously produced LS the structure of an inductive step and, starting from it, s/he succeeds in constructing a recursive argumentation or the inductive step itself within a proof by induction. Some other times, instead, a student anticipates, through the production of LS and IS, the structure that a recursive argumentation should have and then s/he tries to reconstruct such an argumentation for the problem under investigation. In addition to this, it was also possible to register that the suggestion of a LS made by an external intervention, in some cases, might help a student in developing a recursive argumentation or a proof by induction (see section 10.1.4.3). For this last point, as observed, a crucial element for the effectiveness of the intervention is the interpretation that the student gives to the suggested sign.

Moreover, it was possible to observe that of LS and IS can become a lens for the researcher with which to observe the possible students' cognitive unity, with reference to the transition between a recursive argumentation and a proof by MI. In particular the production of those signs can be registered across the problem exploration, the argumentation, and the proof construction. This can happen, for instance, when a student describes the structure of the argumentation s/he is constructing, by producing (linking and iteration) signs which then are repeated when describing the structure of the following proof by MI, or when a student produces an analogous LS when developing one step of the recursive argumentation and when constructing the inductive step. In both cases, the observation of such semiotic continuity might highlight the student's cognitive unity between her/his argumentation and the subsequent proof by induction.

A further result, related to the previous points and highlighted by the analysis of the interview, is the importance of the students' exploration of the problem for the emergence of a recursive argumentation and the successive construction of a proof by MI. As observed by Hanna (1989) proofs by MI are often perceived as not explanatory by students since they seem only to validate a statement without showing "the mathematical properties that cause the asserted theorem […] to be true" (p.51). Stylianides, Sandefur and Watson (2016) starting from Hanna's consideration observed, however, that if the problem resolution involves an exploration, then

> this exploration could, in turn, lead them [(the students)] to construct informally the inductive step and, eventually, see the utility and apply mathematical induction in their proving activity to formalize their work. In such situations, proving by mathematical induction becomes a codification of students' sense making and serves first as a method that explains and then as a proof that verifies. (Ibidem, p. 33).

The study conducted in this thesis, first of all confirms what observed by Stylianides, Sandefur and Watson. Furthermore, what we observed deepens it by highlighting that a crucial point is that this exploration could promote students' production of LS and IS and that the exploration and activity on these signs could lead students to the generation of a recursive argumentation. When this happens, the proof by MI used for formalising a previous recursive argumentation, can become explanatory for students.

13.1.4.2 Ground and Meta level semiotic production

Parallelly to the analysis that led to the identification and study of LS and IS, a second analysis of the interview has been conducted. The focus was on those moments in which the students involved in the problem resolution seemed to pause calculations, symbolic manipulations, geometrical transformations, etc., in order to observe, reflect on, or describe what they have done, what they are planning to do, the structure of the problem or of their own argumentation. Through an analysis of students' semiotic production, it was possible to identify these 'reflective' moments and to register how they play an important role in the problem resolution. In particular two different 'levels' of semiotic production have been distinguished: the ground level and the meta level.

The introduction of these to levels of semiotic production made it possible to observe how the problem resolution of expert students develops through a rich dialectic between different kind of signs which highlight a richness of different processes. In order to better describe and analyse this dialectic, within the theoretical lens of the semiotic bundle, a series of notions have been introduced (liftings, projections, merging points and level's mismatches), together with a descriptive tool (the two lines diagram) used to model the semiotic analysis of an excerpt. This development of the theoretical lens allowed me to register how, for the expert students involved in the study, the two levels of semiotic production seem to be extremely permeable. With this I mean that during the problem resolution, students seem to move their discourse smoothly on the two levels. By investigating on this dialectic between ground and meta level during the problem resolution it was possible to register some interesting aspects.

First of all, the dialectic between the two levels of semiotic production highlights some crucial moments in which students reflect on the problem itself, on its structure, on the structure of the argumentation. It is important to remember that the observation of these moments was possible with an analysis of the whole semiotic bundle and not only of the students' speech. These reflective moments, highlighted by a dialectic between ground and meta level semiotic production, have been registered continuously in all the interviews with expert students. it was also possible to observe the episode of a less expert student who 'got stuck' in the problem (to use his expression), in which these reflective moments were missing, and the semiotic production was almost uniquely ground level (Excerpt 11.5, in section 11.1.4). Within this perspective, thus, a rich dialectic between meta level and ground level sign can become an indicator for an external observer (a teacher or a researcher) to investigate the presence of reflective moments during the resolution of a problem.

Analogously to what already said for LS and IS, the dialectic between ground and meta level semiotic production plays an important role during students' problem solving. Indeed, it was possible to register that this dialectic supports a subject in the generation of a recursive argumentation and the construction of a proof by MI. More specifically, I identified some interesting processes with which a subject operates this dialectic. By lifting the discourse on a meta level during the exploration of the problem, for instance, a subject might reflect on the structure of the argumentation itself and recognise a known structure for it (in our cases, for instance, a recursive argumentation). The fact that meta level signs can have a potential generality, since they refer to the specific problem under resolution and they could also be used to refer to other problems, might support students in recognising that a specific problem has some common aspects with other problems that were already solved with a recursive argumentation. On the other side, projections from meta to ground level could support students in the application of a known argumentative or proving schema (such as a proof by MI) to the specific problem under resolution. In other terms thus meta level signs can be used by the subject to describe and plan the structure of a such a proof and organise its construction with reference to the specific objects of the problem. Moreover, merging points and levels mismatches allows the subjects to produce a discourse which involves both ground and meta levels signs, enabling them to take into consideration simultaneously the two levels. In this regard, it was interesting to observe how sometimes a mathematical error can be produced as a semiotic means with a specific role to make a bridge between ground and meta levels (see, for example, in the excerpt 11.6 in section 11.1.6, when Guido indicates the chessboard's dimensions with $n \times n$ instead of with $2^n \times 2^n$).

Finally, focusing on the use and production of meta level signs made it possible to deeper investigate the students' cognitive unity between recursive argumentation and proof by MI. As written above the use and production of LS and IS in different phases of the problem resolution could contribute to highlight the possible cognitive unity between argumentation and proof. This aspect can be further enriched if we focus specifically on meta level signs. In particular, it was observed that analogous meta level LS and IS are produced by students to refer to the structure of a recursive argumentation and to the structure of a proof by induction. This continuity could be seen as a trace of the fact that the student is recognising a similar structure between recursive argumentations and proofs by MI, and thus highlighting the presence of what, within the construct of the cognitive unity, is called a *structural continuity* between them (see section 3.2).

A last aspect to observe, transversal to the focus on LS/IS or on Ground/Meta level, concerns what this study has registered specifically in relation to the use of gestures in constructing recursive argumentations and proofs by MI. As said in the conceptual framework (section 6.3.2), recent research (Arzarello & Sabena, 2014; Krause, 2015; Sabena, 2018) highlighted how gestures could actively contribute to support students in structuring their argumentations. The analyses of the interviews provided a clear confirmation to these studies. In particular, several of those crucial signs registered in the interviews (LS and IS, both ground or meta level) were gestures or a bundle of them with other semiotic sets. The study presented in this thesis, therefore, registered how the production of gestures does not have only a communicative role for students (which is, of course, present), but it also become a resource, by actively supporting them in planning, organising, and structuring a recursive argumentation or a proof by induction.

13.2 Didactical implications

The results of this thesis, just discussed and presented in terms of answers to the research questions, encourage a series of observations about possible didactical implications of this study.

As widely discussed when answering the research questions, the students' intuitive knowledge of the proving scheme of MI cannot taken for granted. In this sense, it is important for teachers to be aware that, in expert students as well, deep problematics related to MI could be present. As the results of this thesis showed, for several students who already encountered MI during their studies, their formal knowledge of MI was not joined with an intuitive acceptance of this proving scheme. As a result, MI risks to become a ritual procedure for proving statements which is not accompanied by the feeling that it intuitively 'works' as a proving scheme. In tune with Fischbein (1987, pp. 72-81) to reach an intuitive knowledge it will not suffice a knowledge of the logic truth tables, or in general a formal knowledge of some propositional and predicate calculus. In relation to this, in this study it was possible to register how the implication was rarely seen by students as valid in itself, without any concern whatsoever for the truth of antecedent and consequent, also among those students who already encountered the proving scheme of MI. As observed, this aspect highlights an important issue from a didactical point of view. From one side, the proof by MI is based on the fact that one can prove an implication independently from the truth of antecedent and consequent (as it happens in the inductive step), the knowledge of which by students, therefore, seems to be a prerequisite for their understanding of MI. On the other side, it is rare that students have an intuitive knowledge of the validity of an implication in itself as independent from knowing the truth of the statements involved because it is rare that they live experiences in which this aspect of an implication is meaningfully used. They might have encountered the truth tables of the implication, but this, as well shown by Fischbein's studies (1987, pp. 72-81) is not enough to reach an intuitive knowledge. Paradoxically MI is perhaps one of the most (and first) significant cases in which the process of proving the validity of an implication as independent from the truth of antecedent and consequent could be lived by students. Therefore, we have a sort of vicious circle which could explain, in part, students' problematics in possessing an intuitive knowledge of MI. This is an important aspect of which we should be aware. Further research on how to treat, from a didactical point of view, this vicious circle is surely necessary.

In relation to the previous point, it is important to observe that often, in teaching, the introduction of MI to students is accompanied with some of images or metaphors, whose use by teachers and textbooks can be interpreted as oriented to present an intuitive meaning of MI. This is the case, for instance, of the well known image of the falling dominos associated to MI. It is important to notice that the use of metaphors and image could be an effective means to support students' construction of an intuitive knowledge of MI. However, at least two delicate aspects should also be considered.

First of all, we should be aware that such metaphors could bring with themselves some implicit features which do not align with the mathematically correct meaning of MI. This was registered, for instance, in one of the answers of the survey in which the student, by describing the functioning of MI referring to the falling dominos metaphor stated that, knowing that a domino has fallen we can be sure that all the dominos preceding it has fallen as well. This conclusion, coherent with the dominos metaphor, if transposed in mathematical terms corresponds to conclude that, as a consequence of the inductive step and that $P(n^*)$ is true for an n^*, $P(n)$ must be true also for every natural number $n \leq n^*$, which is not correct from a mathematical point of view. Therefore, we should be aware that, from one side, metaphors could be interpreted by students in ways that do not align with mathematically correct meaning of MI and, on the other side, that the fact that a student correctly describes MI with the use of metaphors such as the falling dominoes one, does not guarantee by itself that the student's conceptualisation of MI aligns with the mathematically correct one. Considering this, it is crucial from a didactic point of view, that the use of metaphors and image related to MI is accompanied by a discussion with students of those aspects of the metaphor which fit with the formal aspects of MI and those which do not fit.

Secondly, we should be aware that an intuitive knowledge of MI cannot be simply obtained with transmissive strategies in which the teacher *describes*, for instance, the falling dominos metaphor to students. Indeed, as observed by Fischbein (1982):

> in order to reach 'a basis of belief ' you must live the process. It must be the effect of some behavioral, personal (mental or external) involvement. This is what I mean by getting an intuition, an intuitive understanding of a mathematical truth (p. 14, italics in original).

In other terms, it is important for students also to 'live the process' of proving by MI in analogy with those metaphors which otherwise will remain images only used for describing MI, but detached from the students' personal experiences with the proving process by MI.

In order to reach a knowledge of MI in which a balance and a dialectic between formal and intuitive knowledge can result, it is important to 'live the process' of proving by MI, also by reflecting on the structure of the argumentation. In this sense, for making the students live the process, it is important to involve them in open problems which require an exploration, the formulation of a conjecture and the construction of an argumentation, and not only in closed problems with the form 'Prove by induction that…' as it often happens and in which students are not involved in 'discovering' that a proof by MI (or a recursive argumentation) could be appropriate for the problem. Moreover, it is also important to involve students in activities promoting their reflection on the structure of a recursive argumentation or of MI itself, on its validity, and on when it could be appropriate to be used.

With a different perspective than that one on intuitions, by analysing from a semiotic standpoint the processes of expert students involved in 'living the process' of proving by MI of above, we registered how these processes were characterised by a rich semiotic production (LS and IS) and a continuous dialectic between ground and meta level. This aspect offers as an implication in relation to teachers' role for making the students living crucial processes. Teachers should promote and support the dialectic processes between ground and meta level, as well as the production, use, and exploration of LS and IS, and, moreover, a reflection on the meaning of these signs (both ground and meta level) and thus students' awareness of these processes. The transitions between levels, for instance, could be initially promoted by the teacher who, working within learners' ZPS (Vygotsky, 1978), could take charge of the 'voice' that moves one from one level to the other, a voice which could be then interiorized by students. In tune with this, a practice that a teacher could adopt to promote crucial processes in relation to MI, is what have been called the *semiotic game* (Arzarello & Paola, 2007) between teacher and students:

> Typically, the students explain a new mathematical situation producing simultaneously gestures and speech (or other signs) within a semiotic bundle: their explanation through gestures seems

> promising but their words are very imprecise or wrong and the teacher mimics the former but pushes the latter towards the right form. […] [Students'] gestures within the semiotic bundle (included their relationships with the other signs alive in the bundle) become a powerful mediating tool between signs and thought. From a functional point of view, gestures can act as "personal signs"; while the semiotic game of the teacher starts from them to support the transition to their scientific meaning. (P.23)

In terms of LS and IS, and ground/meta level, thus, a teacher could recognise in the students' semiotic bundle the presence of a LS or IS or signs highlighting effective level's transitions (e.g., merging points or a level's mismatches) on which a 'semiotic game' could be played, aimed to foster the students' construction of a meaningful understanding of the proving scheme by MI.

As the analyses of the interview highlighted, a key aspect of the students' resolution of a problem through the construction of a recursive argumentation (or a proof by MI) was their recognition of a connection between cases of the problem. This was characterised by a crucial *shift of attention* (Mason, 1989): the attention is moved from one or more cases of the problem as separated (e.g., to obtain a given tessellation for the 4x4 chessboard and, independently, for the 8x8 chessboard) to the connection between these cases (e.g. if it is possible to obtain a given tessellation for the 4x4, then it is also possible to obtain a given tessellation for the 8x8 chessboard). This aspect then suggests a didactical objective for the teacher. Students' competencies in problem solving should be enriched with the development of what can we can call an '*Inductive eye*', which is present in the experienced problem solvers: when exploring a problem, they shift their attention from the cases of the problem to the connection between these cases, they look for this connection, they recognise it and to generalise it, if possible, to other cases of the problem, and moreover they are aware that this structural feature of the problem could lead to a recursive argumentation , or more formally to a proof by induction. The term 'Inductive eye' is used in analogy with the notion of '*mathematical eye'* introduced by Mariotti and Baccaglini-Frank (2018) to present some visual-geometrical skills crucially involved in geometrical problem solving:

> How can it happen that, looking at a scribble on a piece of paper […] the observer thinks of "a square"? Or, similarly, looking at a moving image on the screen, the observer suddenly exclaims, "it is a parallelogram!"? Though these experiences may be considered common to any student who has learned a bit of geometry, other more sophisticated experiences are common for expert mathematicians, thanks to a high level of competence in the treatment of images that supports problem solving in geometry; we can call such competence the *mathematical eye*. (Ibidem, p. 155).

With reference to the inductive eye, we can paraphrase the above quotation as follows. How can it happen that, looking at the drawing of an 8x8 chessboard divided in four 4x4 sub-chessboards, the observer exclaims "I think that one could do something like by induction", as it happened in Silvio's excerpt 10.8 (section 10.1.4)? The 'inductive eye' allowed Silvio to *see* a connection between two cases of the problem, independently from having solved a single case (indeed Silvio did not know how to tessellate the 4x4 chessboard yet), it allowed Silvio to recognise its generality and to identify a possible argumentative structure that could be constructed by this. If among the goals of teaching there is "to help high school students learn and adopt some of the ways that mathematicians *think* about problems" (Cuoco et al., 1996, p. 376, italics in original), then the development of such an effective 'inductive eye' in students should become an important objective of teaching.

13.3 Concluding remarks and directions for further research

In the introduction of this thesis, I have highlighted the research problem that moved this study, which is that despite MI has a foundational and epistemologically central role for the modern mathematics, it is a widely problematic topic from an educational perspective. As observed in the literature review, the problem is extremely complex from a cognitive and didactical point of view. On the other side, as registered for instance by the long process which preceded the genesis of the modern formulation of proofs by MI in the history, MI has an epistemological complexity also from a mathematical point of view. The main objective of this study, thus, was to investigate this complexity, shedding light on it and identifying some standpoints by which to observe it.

With this aim, a conceptual framework composed by different theoretical perspectives has been structured. Within this framework some different research questions have formulated and investigated in order to have multi-faced insight into the complexity of the research problem. In addition to the answers to the research questions discussed above, I would like to observe that the conceptual framework that I developed represents itself one of the results of this thesis. In particular some original contributions have been built to structure the framework.

- Firstly, a historical-epistemological analysis of the traces of the proofs by MI in the history have been carried out. This analysis providing a series of elements that have been used in the following development of the conceptual framework offers also an autonomous theoretical contribution to this thesis.
- With reference to the historical-epistemological analysis and to the definitions of 'Argumentation' and 'Proof' adopted in this thesis, an operative definition of Recursive Argumentation has been given. This allowed me to expand the focus of the study not only to the formal proofs by MI but also considering those less formal argumentations which can be seen as related to MI.
- The construct of Theorem as a triplet (Mariotti et al., 1997) have been used to present a model of theorem with a proof by MI as a system of theorems. An explicit reference to this model has been during the whole study, and in particular, the theoretical perspective on the intuitive acceptance of a theorem (Fishbein 1982; 1987) have been integrated with this model of proof by MI.
- A further considered perspective was the APOS Theory. Starting from the historical-epistemological analysis, I deepened and proposed some modification to the Genetic Decomposition of MI present in the literature. This proposed modification allowed me to deeper investigate a central process in the MI Schema in APOS terms, that is the explain induction process.

Moreover, some of the results presented as answers to the research questions offer other original theoretical elements which contribute in further enlarge the conceptual framework.

- The analysis of the results from the survey prompted to a deeper analysis of the Explain induction process (both direct and indirect form) which has been described and modelled in terms of the Chaining of MP/MT process. This point further expands the Genetic Decomposition of MI, in relation to the Explain induction process.
- The analysis of the interview with a semiotic perspective led me to the development of two theoretical constructs: the notion of LS and IS, and of Meta and Ground level semiotic production, together with a series of notions related to it (lifting/projections, level's mismatches, and merging points). These constructs, as discussed above, allowed me to formulate an answer for two of the RQs investigated. However, these constructs also expanded the conceptual framework originally considered, providing some research tools with which to investigate the semiotic production of students. In other words, then, the notions of LS/IS and Ground/Meta level, in addition to being stand-alone outcomes of this thesis (intended as an answer to the research questions), have also become research tools that goes to enrich the semiotic perspective in the conceptual framework.

Finally, I conclude this chapter outlining possible directions for further research that might stem from this study.

The Genetic Decomposition of MI developed in this study was used as an interpretive lens to describe and analyse students' processes and difficulties. However, as noted in the description of the framework (section 4.2.6), within APOS theory a Genetic Decomposition can also be used as a didactical tool for the design of educational interventions. Therefore, a possible direction for future research could focus on using the proposed Genetic Decomposition of MI to frame the design of educational interventions aiming to support students in the construction of a mathematically rich, efficient and coherent Schema of MI.

The semiotic perspective with the theoretical lens developed in this thesis could be effective to observe and analyse the classroom's social dynamics involving teacher and students, in relation to the teaching and learning of MI. In particular it could be interesting to carry out a long term study investigating on the evolution of the semiotic production shared in the classroom, focusing on how, eventually, the production of LS/IS and the dialectic between ground and meta level contribute in developing the classroom's discourse on MI.

As outlined when the didactical implications were discussed, further research is necessary on the role of the teacher to foster a balanced and effective dialectic between intuitive and formal knowledge of MI, to promote a dialectic between ground and meta level semiotic production, and to support student's development of the "Inductive eye". The results of the thesis suggest that the semiotic game (Arzarello & Paola, 2007) could be a suitable methodology for these aims, however further research on the teaching and learning of particular proving schemes, among which the one by MI, in relation with the semiotic game is necessary.

A further research line could involve the study related to the task design and the construction of didactic activities with specific artifacts (not necessarily, but also digital) that take into account the findings of this thesis and are in line with specific educational objectives in relation to MI. A first explorative study in this direction has been conducted in the last months. In particular, I designed and implemented a set of activities with secondary school students aiming to foster their production and use of LS and IS and to promote a dialectic between ground and meta level signs. Unfortunately, a new worsening of the pandemic situation in Italy when this explorative study was conducted slowed down its development and made its continuation problematic, which is why it was not presented this thesis.

The theoretical development within the semiotic perspective that led to the definition of LS/IS and ground/meta levels could become a tool for the a-priori analysis of problems related to MI. In particular, therefore, a possible future research direction could focus on using the developed semiotic theoretical lens to identify and describe the characteristics of "effective" problems to promote crucial processes in relation to MI, such as the production and exploration of LS/IS and the dialectical process between ground and meta level semiotic production. This possible research direction could offer new elements with which to enrich Harel's (2001) classification of problems related to MI discussed in the literature review (section 1.3.3).

In the thesis I presented examples in which the theoretical lens developed within the semiotic perspective turn out to be effective in investigating aspects proper of the APOS theory. Specifically, in Chapter 12 (section 12.1.1), I showed how the semiotic lens enriched the analysis of the students' construction of the Explain induction process both in direct and in indirect form. This point opens to a possible research direction for the future which is to investigate within a semiotic lens on other APOS elements contained in the GD of MI or in relation to the GD of other mathematical concepts. The two perspectives (semiotics and APOS) have distant foundations (Vygotskian the former, Piagetian the latter), an aspect that requires caution in using a semiotic lens as a tool for an APOS analysis. However, also the researchers who adopt an APOS perspective make use of protocols and analyse subjects' speech and writings. Extending this analysis to other semiotic sets, and in particular to gestures and inscriptions, might be a suitable and effective methodological tool for the analysis of processes within and APOS perspective, as the results obtained in this thesis suggest. In conclusion, thus, it might be interesting to develop this research line, focusing with a semiotic perspective on some key transitions in APOS terms (e.g., Interiorization, (de-)encapsulation, reversal, thematization, etc.) related to students' construction of different Schemas.

Finally, let us observe that the construct of Ground/Meta level semiotic production, together with the theoretical notions used to investigate and analyse the dialectic between the two levels (liftings, projections, merging points, level's mismatches), has a broader applicability than has been seen in this thesis. It is a general construct, which is not limited only to the proving scheme of MI. For this reason, further research could be carried out, using this construct as a theoretical lens to investigate the students' processes in relation to other specific proving schemes, and more in general in situations involving the construction of conjectures, argumentations, and proofs, and in relation to problem solving in mathematics.

APPENDIX A – DEFINITIONS AND PROPOSITIONS FROM BOOKS VII-IX OF EUCLID'S *ELEMENTS*

The list of definitions and propositions that is presented below is taken from the following website, created by David E. Joyce in 1996: https://mathcs.clarku.edu/~djoyce/java/elements/

VII Book

Definitions

1. A unit is that by virtue of which each of the things that exist is called one.

2. A number is a multitude composed of units.

3. A number is a part of a number, the less of the greater, when it measures the greater.

4. But parts when it does not measure it.

5. The greater number is a multiple of the less when it is measured by the less.

6. An even number is that which is divisible into two equal parts.

7. An odd number is that which is not divisible into two equal parts, or that which differs by a unit from an even number.

8. An even-times-even number is that which is measured by an even number according to an even number.

9. An even-times-odd number is that which is measured by an even number according to an odd number.

10. An odd-times-odd number is that which is measured by an odd number according to an odd number.

11. A prime number is that which is measured by a unit alone.

12. Numbers relatively prime are those which are measured by a unit alone as a common measure.

13. A composite number is that which is measured by some number.

14. Numbers relatively composite are those which are measured by some number as a common measure.

15. A number is said to multiply a number when the latter is added as many times as there are units in the former.

16. And, when two numbers having multiplied one another make some number, the number so produced be called plane, and its sides are the numbers which have multiplied one another.

17. And, when three numbers having multiplied one another make some number, the number so produced be called solid, and its sides are the numbers which have multiplied one another.

18. A square number is equal multiplied by equal, or a number which is contained by two equal numbers.

19. And a cube is equal multiplied by equal and again by equal, or a number which is contained by three equal numbers.

20. Numbers are proportional when the first is the same multiple, or the same part, or the same parts, of the second that the third is of the fourth.

21. Similar plane and solid numbers are those which have their sides proportional.

22. A perfect number is that which is equal to the sum its own parts.

Propositions

1. When two unequal numbers are set out, and the less is continually subtracted in turn from the greater, if the number which is left never measures the one before it until a unit is left, then the original numbers are relatively prime.

2. To find the greatest common measure of two given numbers not relatively prime.

Corollary. If a number measures two numbers, then it also measures their greatest common measure.

3. To find the greatest common measure of three given numbers not relatively prime.

4. Any number is either a part or parts of any number, the less of the greater.

5. If a number is part of a number, and another is the same part of another, then the sum is also the same part of the sum that the one is of the one.

6. If a number is parts of a number, and another is the same parts of another, then the sum is also the same parts of the sum that the one is of the one.

7. If a number is that part of a number which a subtracted number is of a subtracted number, then the remainder is also the same part of the remainder that the whole is of the whole.

8. If a number is the same parts of a number that a subtracted number is of a subtracted number, then the remainder is also the same parts of the remainder that the whole is of the whole.

9. If a number is a part of a number, and another is the same part of another, then alternately, whatever part or parts the first is of the third, the same part, or the same parts, the second is of the fourth.

10. If a number is a parts of a number, and another is the same parts of another, then alternately, whatever part of parts the first is of the third, the same part, or the same parts, the second is of the fourth.

11. If a whole is to a whole as a subtracted number is to a subtracted number, then the remainder is to the remainder as the whole is to the whole.

12. If any number of numbers are proportional, then one of the antecedents is to one of the consequents as the sum of the antecedents is to the sum of the consequents.

13. If four numbers are proportional, then they are also proportional alternately.

14. If there are any number of numbers, and others equal to them in multitude, which taken two and two together are in the same ratio, then they are also in the same ratio ex aequali.

15. If a unit number measures any number, and another number measures any other number the same number of times, then alternately, the unit measures the third number the same number of times that the second measures the fourth.

16. If two numbers multiplied by one another make certain numbers, then the numbers so produced equal one another.

17. If a number multiplied by two numbers makes certain numbers, then the numbers so produced have the same ratio as the numbers multiplied.

18. If two number multiplied by any number make certain numbers, then the numbers so produced have the same ratio as the multipliers.

19. If four numbers are proportional, then the number produced from the first and fourth equals the number produced from the second and third; and, if the number produced from the first and fourth equals that produced from the second and third, then the four numbers are proportional.

20. The least numbers of those which have the same ratio with them measure those which have the same ratio with them the same number of times; the greater the greater; and the less the less.

21. Numbers relatively prime are the least of those which have the same ratio with them.

22. The least numbers of those which have the same ratio with them are relatively prime.

23. If two numbers are relatively prime, then any number which measures one of them is relatively prime to the remaining number.

24. If two numbers are relatively prime to any number, then their product is also relatively prime to the same.

25. If two numbers are relatively prime, then the product of one of them with itself is relatively prime to the remaining one.

26. If two numbers are relatively prime to two numbers, both to each, then their products are also relatively prime.

27. If two numbers are relatively prime, and each multiplied by itself makes a certain number, then the products are relatively prime; and, if the original numbers multiplied by the products make certain numbers, then the latter are also relatively prime.

28. If two numbers are relatively prime, then their sum is also prime to each of them; and, if the sum of two numbers is relatively prime to either of them, then the original numbers are also relatively prime.

29. Any prime number is relatively prime to any number which it does not measure.

30. If two numbers, multiplied by one another make some number, and any prime number measures the product, then it also measures one of the original numbers.

31. Any composite number is measured by some prime number.

32. Any number is either prime or is measured by some prime number.

33. Given as many numbers as we please, to find the least of those which have the same ratio with them.

34. To find the least number which two given numbers measure.

35. If two numbers measure any number, then the least number measured by them also measures the same.

36. To find the least number which three given numbers measure.

37. If a number is measured by any number, then the number which is measured has a part called by the same name as the measuring number.

38. If a number has any part whatever, then it is measured by a number called by the same name as the part.

39. To find the number which is the least that has given parts.

VIII Book

Propositions

1. If there are as many numbers as we please in continued proportion, and the extremes of them are relatively prime, then the numbers are the least of those which have the same ratio with them.

2. To find as many numbers as are prescribed in continued proportion, and the least that are in a given ratio.

Corollary. If three numbers in continued proportion are the least of those which have the same ratio with them, then the extremes are squares, and, if four numbers, cubes.

3. If as many numbers as we please in continued proportion are the least of those which have the same ratio with them, then the extremes of them are relatively prime.

4. Given as many ratios as we please in least numbers, to find numbers in continued proportion which are the least in the given ratios.

5. Plane numbers have to one another the ratio compounded of the ratios of their sides.

6. If there are as many numbers as we please in continued proportion, and the first does not measure the second, then neither does any other measure any other.

7. If there are as many numbers as we please in continued proportion, and the first measures the last, then it also measures the second.

8. If between two numbers there fall numbers in continued proportion with them, then, however many numbers fall between them in continued proportion, so many also fall in continued proportion between the numbers which have the same ratios with the original numbers.

9. If two numbers are relatively prime, and numbers fall between them in continued proportion, then, however many numbers fall between them in continued proportion, so many also fall between each of them and a unit in continued proportion.

10. If numbers fall between two numbers and a unit in continued proportion, then however many numbers fall between each of them and a unit in continued proportion, so many also fall between the numbers themselves in continued proportion.

11. Between two square numbers there is one mean proportional number, and the square has to the square the duplicate ratio of that which the side has to the side.

12. Between two cubic numbers there are two mean proportional numbers, and the cube has to the cube the triplicate ratio of that which the side has to the side.

13. If there are as many numbers as we please in continued proportion, and each multiplied by itself makes some number, then the products are proportional; and, if the original numbers multiplied by the products make certain numbers, then the latter are also proportional.

14. If a square measures a square, then the side also measures the side; and, if the side measures the side, then the square also measures the square.

15. If a cubic number measures a cubic number, then the side also measures the side; and, if the side measures the side, then the cube also measures the cube.

16. If a square does not measure a square, then neither does the side measure the side; and, if the side does not measure the side, then neither does the square measure the square.

17. If a cubic number does not measure a cubic number, then neither does the side measure the side; and, if the side does not measure the side, then neither does the cube measure the cube.

18. Between two similar plane numbers there is one mean proportional number, and the plane number has to the plane number the ratio duplicate of that which the corresponding side has to the corresponding side.

19. Between two similar solid numbers there fall two mean proportional numbers, and the solid number has to the solid number the ratio triplicate of that which the corresponding side has to the corresponding side.

20. If one mean proportional number falls between two numbers, then the numbers are similar plane numbers.

21. If two mean proportional numbers fall between two numbers, then the numbers are similar solid numbers.

22. If three numbers are in continued proportion, and the first is square, then the third is also square.

23. If four numbers are in continued proportion, and the first is a cube, then the fourth is also a cube.

24. If two numbers have to one another the ratio which a square number has to a square number, and the first is square, then the second is also a square.

25. If two numbers have to one another the ratio which a cubic number has to a cubic number, and the first is a cube, then the second is also a cube.

26. Similar plane numbers have to one another the ratio which a square number has to a square number.

27. Similar solid numbers have to one another the ratio which a cubic number has to a cubic number.

IX Book

Propositions

1. If two similar plane numbers multiplied by one another make some number, then the product is square.

2. If two numbers multiplied by one another make a square number, then they are similar plane numbers.

3. If a cubic number multiplied by itself makes some number, then the product is a cube.

4. If a cubic number multiplied by a cubic number makes some number, then the product is a cube.

5. If a cubic number multiplied by any number makes a cubic number, then the multiplied number is also cubic.

6. If a number multiplied by itself makes a cubic number, then it itself is also cubic.

7. If a composite number multiplied by any number makes some number, then the product is solid.

8. If as many numbers as we please beginning from a unit are in continued proportion, then the third from the unit is square as are also those which successively leave out one, the fourth is cubic as are also all those which leave out two, and the seventh is at once cubic and square are also those which leave out five.

9. If as many numbers as we please beginning from a unit are in continued proportion, and the number after the unit is square, then all the rest are also square; and if the number after the unit is cubic, then all the rest are also cubic.

10. If as many numbers as we please beginning from a unit are in continued proportion, and the number after the unit is not square, then neither is any other square except the third from the unit and all those which leave out one; and, if the number after the unit is not cubic, then neither is any other cubic except the fourth from the unit and all those which leave out two.

11. If as many numbers as we please beginning from a unit are in continued proportion, then the less measures the greater according to some one of the numbers which appear among the proportional numbers.

Corollary. Whatever place the measuring number has, reckoned from the unit, the same place also has the number according to which it measures, reckoned from the number measured, in the direction of the number before it.

12. If as many numbers as we please beginning from a unit are in continued proportion, then by whatever prime numbers the last is measured, the next to the unit is also measured by the same.

13. If as many numbers as we please beginning from a unit are in continued proportion, and the number after the unit is prime, then the greatest is not measured by any except those which have a place among the proportional numbers.

14. If a number is the least that is measured by prime numbers, then it is not measured by any other prime number except those originally measuring it.

15. If three numbers in continued proportion are the least of those which have the same ratio with them, then the sum of any two is relatively prime to the remaining number.

16. If two numbers are relatively prime, then the second is not to any other number as the first is to the second.

17. If there are as many numbers as we please in continued proportion, and the extremes of them are relatively prime, then the last is not to any other number as the first is to the second.

18. Given two numbers, to investigate whether it is possible to find a third proportional to them.

19. Given three numbers, to investigate when it is possible to find a fourth proportional to them.

20. Prime numbers are more than any assigned multitude of prime numbers.

21. If as many even numbers as we please are added together, then the sum is even.

22. If as many odd numbers as we please are added together, and their multitude is even, then the sum is even.

23. If as many odd numbers as we please are added together, and their multitude is odd, then the sum is also odd.

24. If an even number is subtracted from an even number, then the remainder is even.

25. If an odd number is subtracted from an even number, then the remainder is odd.

26. If an odd number is subtracted from an odd number, then the remainder is even.

27. If an even number is subtracted from an odd number, then the remainder is odd.

28. If an odd number is multiplied by an even number, then the product is even.

29. If an odd number is multiplied by an odd number, then the product is odd.

30. If an odd number measures an even number, then it also measures half of it.

31. If an odd number is relatively prime to any number, then it is also relatively prime to double it.

32. Each of the numbers which are continually doubled beginning from a dyad is even-times even only.

33. If a number has its half odd, then it is even-times odd only.

34. If an [even] number neither is one of those which is continually doubled from a dyad, nor has its half odd, then it is both even-times even and even-times odd.

35. If as many numbers as we please are in continued proportion, and there is subtracted from the second and the last numbers equal to the first, then the excess of the second is to the first as the excess of the last is to the sum of all those before it.

36. If as many numbers as we please beginning from a unit are set out continuously in double proportion until the sum of all becomes prime, and if the sum multiplied into the last makes some number, then the product is perfect.

APPENDIX B – CONSENT FORM FOR THE INTERVIEW

This appendix contains the consent form that the subjects involved sin the interviews signed. It is presented in Italian, as in original.

CONSENT FORM FOR UNIVERSITY STUDENTS

Documentazione Video Fotografica
Lettera Liberatoria

Io sottoscritto/a ____

autorizzo

il dott. Bernardo Nannini, dottorando presso l'Università degli Studi di Firenze, sotto la direzione scientifica del prof. Samuele Antonini (Università degli Studi di Pavia) e della prof. Veronica Gavagna (Università degli studi di Firenze), all'uso di foto e video raccolti nel corso di interviste svolte tra ottobre 2019 e dicembre 2020, nell'ambito di un progetto di ricerca riguardante i processi argomentativi coinvolti nella risoluzione di problemi matematici per realizzare materiale documentativo mediante, anche, riduzioni o adattamenti. Nel caso in cui questo materiale dovesse essere reso pubblico, l'identità degli intervistati sarà protetta dall'anonimato o dall'uso d uno pseudonimo.

Il materiale documentativo sarà utilizzato senza limite di spazio e tempo per scopi:

- Tecnico – scientifici
- Informativi – formativi
- Didattici

Non è richiesto, in proposito, alcun tipo di compenso.

Data ____________________

Firma ____________________________

BIBLIOGRAPHY

Acerbi, F. (2000). Plato: Parmenides 149a7-c3. A Proof by Complete Induction? *Archive for History of Exact Sciences*, *55*(1), 57–76.

Antonini, S. (2004). A statement, the contrapositive and the inverse: intuition and argumentation. In M. J. Høines & A. B. Fuglestad (Eds.), *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 47–54). Bergen, Norway: PME.

Antonini, S. (2011). Generating examples: Focus on processes. *ZDM*, *43*(2), 205–217.

Antonini, S. (2019). Intuitive acceptance of proof by contradiction. *ZDM*, *51*(5), 793–806.

Antonini, S., & Mariotti, M. A. (2008). Indirect proof: What is specific to this way of proving? *ZDM*, *40*(3), 401–412.

Antonini, S., & Nannini, B. (2021). Linking and iteration signs in proving by induction. In M. Inprasitha, N. Changsri, & N. Boonsena (Eds.), *Proceedings of the 44th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 19–28). Khon Kaen, Thailand: PME.

Arnon, I., Cottrill, J., Dubinsky, E., Oktaç, A., Roa Fuentes, S., Trigeros, M., & Weller, K. (2014). *APOS theory: A framework for research and curriculum development in mathematics education*. Springer-Verlag.

Arzarello, F. (2006). Semiosis as a multimodal process. *RELIME. Revista Latinoamericana de Investigación En Matemática Educativa*, *9*(Extra 1), 267–300.

Arzarello, F., & Edwards, L. (2005). RF02: Gesture and the construction of mathematical meaning. *Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education* (Vol.1, pp. 123–154). Melbourne, Australia: University of Melbourne.

Arzarello, F., & Paola, D. (2007). Semiotic games: The role of the teacher. In J. H. Woo, H. C. Lew, K. S. Park, & D. Y. Seo (Eds.), *Proceedings of the 31st Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 17–24). Seoul, South Korea: PME.

Arzarello, F., & Sabena, C. (2014). Analytic-Structural Functions of Gestures in Mathematical Argumentation Processes. In L. D. Edwards, F. Ferrara, & D. Moore-Russo (Eds.), *Emerging perspectives on gesture and embodiment* (pp. 75–103). Information Age Publishing.

Austin, K. (1988). A paradox—Four weighings suffice. *The Mathematical Gazette, Note No. 72.15*, *72*(460), 113.

Avital, S., & Libeskind, S. (1978). Mathematical induction in the classroom: Didactical and mathematical issues. *Educational Studies in Mathematics*, *9*(4), 429–438.

Balacheff, N. (1988). Aspects of proof in pupils' practice of school mathematics. *Mathematics, teachers and children*, *216*, 235.

Balacheff, N. (2008). The role of the researcher's epistemology in mathematics education: An essay on the case of proof. *ZDM*, *40*(3), 501–512.

Bikner-Ahsbahs, A., & Prediger, S. (Eds.). (2014). *Networking of Theories as a Research Practice in Mathematics Education*. Springer International Publishing.

Boero, P. (1999). *Argumentation and mathematical proof: A complex, productive, unavoidable relationship in mathematics and mathematics education*. International Newsletter on the Teaching and Learning of Mathematical Proof. http://lettredelapreuve.org/OldPreuve/Newsletter/990708Theme/990708ThemeUK.html

Boero, P., Douek, N., Morselli, F., & Pedemonte, B. (2010). Argumentation and proof: A contribution to theoretical perspectives and their classroom implementation. In M. M. F. Pinto & T. F. Kawasaki (Eds.), *Proceedings of the 34th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 179–209). Belo Horizonte, Brazil: PME.

Boero, P., Garuti, R., & Mariotti, M. A. (1996). Challenging the traditional school approach to theorems: A hypothesis about the cognitive unity of theorems. In L. Puig & A. Gutierrez (Eds.), *Proceedings of the 20th Conference of the International Group for the Psychology of Mathematics Education* (Vol.2, pp. 113–120). Valencia, Spain: Universitat de Valencia.

Bussey, W. H. (1917). The Origin of Mathematical Induction. *The American Mathematical Monthly*, *24*(5), 199–207.

Bussey, W. H. (1918). Fermat's Method of Infinite Descent. *The American Mathematical Monthly*, *25*(8), 333–337.

Cajori, F. (1918). Origin of the Name «Mathematical Induction». *The American Mathematical Monthly*, *25*(5), 197.

Cantor, M. (1880). *Vorlesungen über Geschichte der Mathematik*. B.G. Teubner.

Carlson, M. P., & Bloom, I. (2005). The Cyclic Nature of Problem Solving: An Emergent Multidimensional Problem-Solving Framework. *Educational Studies in Mathematics*, *58*(1), 45–75.

Carotenuto, G., Coppola, C., & Di Martino, P. (2018). Mathematical induction at the tertiary level: Looking behind appearances. In E. Bergqvist, M. Österholm, C. Granberg, & L. Sumpter (Eds.) *Proceedings of the 42nd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 219–226). Umeå, Sweden: PME

Cassinet, J. (1988). The first arithmetic book of Francisco Maurolico, written in 1557 and printed in 1575: A step towards a theory of numbers. In C. Hay, *Mathematics from manuscript to print, 1300-1600* (pp. 162–179). Oxford University Press.

Cavaillès, J. (1962). *Philosophie mathématique*. Hermann.

Cornu, B. (1991). Limits. In D. Tall (Eds.), *Advanced Mathematical Thinking* (pp. 153–166). Kluwer.

Cuoco, A., Goldenberg, E. P., & Mark, J. (1996). Habits of mind: An organizing principle for mathematics curricula. *The Journal of Mathematical Behavior*, *15*(4), 375–402.

Cusi, A., & Malara, N. (2009). Improving awareness about the meaning of the principle of mathematical induction. In O. Figueras, J. L. Cortina, S. Alatorre, T. Rojano & A. Sepulveda (Eds.), *Proceedings of the Joint Meeting of PME-32 and PME-NAXXX*, (Vol.2, pp. 392–398). Morelia, Mexico: PME.

Di Martino, P. (2018). Matematica strumentale vs matematica relazionale: Il caso del principio d'induzione. *L'insegnamento della matematica e delle scienze integrate, 41B*(1), 29–52.

diSessa, A. A. (2007). An Interactional Analysis of Clinical Interviewing. *Cognition and Instruction*, *25*(4), 523–565.

Dörfler, W. (2016). Signs and Their Use: Peirce and Wittgenstein. In A. Bikner-Ahsbahs, A. Vohns, O. Schmitt, R. Bruder, & W. Dörfler (Eds.), *Theories in and of Mathematics Education: Theory Strands in German Speaking Countries* (pp. 21–31). Springer International Publishing.

Douek, N. (1999a). Some remarks about argumentation and mathematical proof and their educational implications. In I. Schwank (Ed.), *Proceedings of the 1st Conference of the European Society for Research in Mathematics Education* (Vol. 1, pp. 125–139). Osnabrück, Germany: ERME.

Douek, N. (1999b). Argumentative aspects of proving: Analysis of some undergraduate mathematics students' performances. In F. Hitt & M. Santos (Eds.), *Proceedings of the 23rd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 273-280). Haifa, Israel: PME.

Douek, N. (2002). Context complexity and argumentation. In A. Cockburn & E. Nardi (Eds.), *Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 297-304). Norwich, England: PME

Douek, N. (2007). Some remarks about argumentation and proof. In P. Boero (Ed.), *Theorems in school* (pp. 163–181). Sense Publisher.

Dubinsky, E. (1986). Teaching Mathematical Induction I. *Journal of Mathematical Behavior*, *5*(3), 305–317.

Dubinsky, E. (1989). Teaching mathematical induction II. *Journal of Mathematical Behavior*, *8*(3), 285–304.

Dubinsky, E. (1991). Reflective Abstraction in Advanced Mathematical Thinking. In D. Tall (Eds.), *Advanced Mathematical Thinking* (pp. 95–126). Kluwer.

Dubinsky, E., & Lewin, P. (1986). Reflective abstraction and mathematics education: The genetic decomposition of induction and compactness. *The Journal of Mathematical Behavior*, *5*(1), 55–92.

Dubinsky, E., Weller, K., McDonald, M. A., & Brown, A. (2005a). Some Historical Issues and Paradoxes Regarding the Concept of Infinity: An Apos-Based Analysis: Part 1. *Educational Studies in Mathematics*, *58*(3), 335–359.

Dubinsky, E., Weller, K., McDonald, M. A., & Brown, A. (2005b). Some Historical Issues and Paradoxes Regarding the Concept of Infinity: An Apos Analysis: Part 2. *Educational Studies in Mathematics*, *60*(2), 253.

Duval, R. (1991). Structure du raisonnement deductif et apprentissage de la demonstration. *Educational Studies in Mathematics*, *22*(3), 233–261.

Duval, R. (2006). A Cognitive Analysis of Problems of Comprehension in a Learning of Mathematics. *Educational Studies in Mathematics*, *61*(1), 103–131.

Edwards, L. D. (2010). Doctoral students, embodied discourse and proof. In M. M. F. Pinto & T. F. Kawasaki (Eds). *Proceedings of the 34th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 329–336). Belo Horizonte, Brazil: PME.

Ernest, P. (1982). Mathematical Induction: A Recurring Theme. *The Mathematical Gazette*, *66*(436), 120–125.

Ernest, P. (1984). Mathematical induction: A pedagogical discussion. *Educational Studies in Mathematics*, *15*(2), 173–189.

Ernest, P. (2006). A Semiotic Perspective of Mathematical Activity: The Case of Number. *Educational Studies in Mathematics*, *61*(1), 67–101.

Fermat, P. de. (1894). *OEuvres de Fermat. Tome 2, Correspondence* (Tannery, P. & Henry, C., eds.). Gauthier-Villars et fils.

Ferrari, P. L. (2004). *Matematica e linguaggio: Quadro teorico e idee per la didattica*. Pitagora.

Fischbein, E. (1982). Intuition and Proof. *For the Learning of Mathematics*, *3*(2), 9–24.

Fischbein, E. (1987). *Intuition in Science and Mathematics: An Educational Approach*. Kluwer.

Fischbein, E. (1999). Intuitions and Schemata in Mathematical Reasoning. *Educational Studies in Mathematics*, *38*, 11–50.

Fischbein, E., & Engel, H. (1989). Psychological difficulties in understanding the principle of mathematical induction. In G. Vergnaud, J. Rogalski, & M. Artigue (Eds.), *Proceedings of the 13th Conference of the International Group for the Psychology of Mathematics Education* (Vol.1, pp. 276–282). Paris, France: PME.

Fowler, D. (1994). Could the Greeks have used mathematical induction? Did they use it? *Physis*, *31*(1), 253–266.

Freudenthal, H. (1953). Zur Geschichte der vollständigen Induktion. *Archives Internationales d'Histoire Des Sciences*, *22*, 17–37.

Freudenthal, H. (1973). *Mathematics as an educational task*. Reidel.

García-Martínez, I., & Parraguez, M. (2017). The basis step in the construction of the principle of mathematical induction based on APOS theory. *The Journal of Mathematical Behavior*, *46*, 128–143.

Garuti, R., Boero, P., & Lemut, E. (1998). Cognitive unity of theorems and difficulty of proof. In A. Olivier, & K. Newstead (Eds.), *Proceedings of the 22th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 345–352). Stellenbosch, South Africa: PME.

Ginsburg, H. (1981). The Clinical Interview in Psychological Research on Mathematical Thinking: Aims, Rationales, Techniques. *For the Learning of Mathematics*, *1*(3), 4–11.

Goldin, G. A. (2000). A scientific perspective on structured, task-based interviews in mathematics education research. In R. Lesh, & A.E. Kelly (Eds.), *Research design in mathematics and science education*. Lawrence Erlbaum Associates, pp. 517-545.

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Belknap Press of Harvard University Press.

Goldstein, C. (1995). *Un théorème de Fermat et ses lecteurs*. Presses universitaires de Vincennes.

Gunderson, D. S. (2011). *Handbook of mathematical induction: Theory and applications*. Chapman & Hall/CRC.

Hadamard, J. (1954). *The psychology of invention in the mathematical field*. Dover.

Hanna, G. (1989). Proofs that Prove and Proofs that Explain. In G. Vergnaud, J. Rogalski, & M. Artigue (Eds.), *Proceedings of the 13th Conference of the International Group for the Psychology of Mathematics Education* (Vol.2, pp. 45–51). Paris, France: PME.

Hannula, M. S. (2014). Affect in Mathematics Education. In S. Lerman (Ed.), *Encyclopedia of mathematics education* (pp. 23–27). Springer Reference.

Hannula, M. S., Leder, G. C., Morselli, F., Vollstedt, M., & Zhang, Q. (2019). *Affect and mathematics education: Fresh perspectives on motivation, engagement, and identity*. Springer Nature.

Harel, G. (2001). The Development of Mathematical Induction as a Proof Scheme: A Model for DNR-Based. In S. Campbell & R. Zazkis (Eds.), *Learning and Teaching Number Theory: Research in Cognition and Instruction* (pp. 185-212). Kluwer.

Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. *American Mathematical Society*, *7*, 234–283.

Heath, T. L. (1956). *The thirteen books of Euclid's Elements translated from the text of Heiberg with introduction and commentary* (2nd edition revised with additions). Dover.

Hyde, D., & Raffman, D. (2018). Sorites Paradox. In E. N. Zalta (Eds.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University.

Inglis, M., & Simpson, A. (2008). Conditional inference and advanced mathematical study. *Educational Studies in Mathematics*, *67*(3), 187–204.

Kilpatrick, J. (1985). Reflection and Recursion. *Educational Studies in Mathematics*, *16*(1), 1–26.

Krause, C. M. (2016). *The mathematics in our hands: How gestures contribute to constructing mathematical knowledge*. Springer Spektrum.

Krause, C. M. (2015). Gestures as part of discourse in reasoning situations: Introducing two epistemic functions of gestures. In K. Krainer & N. Vondrová (Eds.), *Proceedings of the 9th Congress of the European Society for Research in Mathematics Education* (pp. 1427–1433). Prague, Czech Republic: Charles University in Prague, Faculty of Education and ERME.

Laertius, D. (1989). *Lives of Eminent Philosophers* (E. Capps, T. E. Page, & W. H. D. Rouse, Eds.), Vol. 2. Loeb.

Larson, N., & Pettersson, K. (2018). Proof by induction – the role of the induction basis. In E. Norén, H. Palmér, & A. Cooke (Eds.), *Nordic Research in Mathematics Education. Papers of NORMA 17* (pp. 99–107). Linköping, Sweden: SMDF.

Leron, U., & Zazkis, R. (1986). Computational Recursion and Mathematical Induction. *For the Learning of Mathematics*, *6*(2), 25–28.

Lolli, G. (2008). *Guida alla teoria degli insiemi*. Springer.

Lowenthal, F., & Eisenberg, T. (1992). Mathematical Induction in School: An Illusion of Rigor? *School Science and Mathematics*, *92*(5), 233–238.

Manvel, B. (1977). Counterfeit coin problems. *Mathematics Magazine*, *50*(2), 90–92.

Mariotti, M. A. (2006). Proof and Proving in Mathematics Education. In A. Gutiérrez, & P. Boero (Eds.), *Handbook of Research on the Psychology of Mathematics Education: Past, Present and Future* (pp. 173–2014). Sense publishers.

Mariotti, M. A., & Baccaglini-Frank, A. (2018). Developing the Mathematical Eye Through Problem-Solving in a Dynamic Geometry Environment. In N. Amado, S. Carreira, & K. Jones (Eds.), *Broadening the Scope of Research on Mathematical Problem Solving: A Focus on Technology, Creativity and Affect* (pp. 153–176). Springer International Publishing.

Mariotti, M. A., Bartolini Bussi, M. G., Boero, P., Ferri, F., & Garuti, R. (1997). Approaching geometry theorems in contexts: From history and epistemology to cognition. In E. Pehkonen (Ed.), *Proceedings of the 21st Conference of the International Group for the Psychology of Mathematics Education* (Vol.1, pp. 180–195). Lahti, Finland: PME.

Mason, J. (1989). Mathematical Abstraction as the Result of a Delicate Shift of Attention. *For the Learning of Mathematics*, *9*(2), 2–8.

Mason, J., & Pimm, D. (1984). Generic Examples: Seeing the General in the Particular. *Educational Studies in Mathematics*, *15*(3,), 277–289.

Maurolico, F. (1575). *Arithmeticorum libri duo*. Franciscum Franciscium Senensem.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press.

McNeill, D. (2005). *Gesture and Thought*. University of Chicago Press.

McNeill, D., & Duncan, S. (2000). Growth points in thinking-for-speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 141–161). Cambridge University Press.

MIUR (2010). Schema di regolamento recante "Indicazioni nazionali riguardanti gli obiettivi specifici di apprendimento concernenti le attività e gli insegnamenti compresi nei piani degli studi previsti per i percorsi liceali". https://www.indire.it/lucabas/lkmw_file/licei2010/indicazioni_nuovo_impaginato/_decreto_indicazioni_nazionali.pdf

Morin, E. (2008). *On complexity*. Hampton Press.

Moscheo, R. (2011). Matematiche e storia a Messina alla vigilia del sisma del 1908 (Note intorno ad un coevo e incompiuto progetto editoriale). *Atti della Accademia Perloritana dei Pericolanti. Classe di Scienze Fisiche, Matematiche e Naturali*, *89*(1).

Movshovitz-Hadar, N. (1993). The false coin problem, mathematical induction and knowledge fragility. *The Journal of Mathematical Behavior*, *12*(3), 253–268.

Mueller, I. (1981). *Philosophy of Mathematics and Deductive Structure of Euclid 's «Elements»*. Dover Publications.

Nannini, B. (2022). Representing mathematical induction in proving processes. *ACCEPTED FOR Proceedings of the 12th Congress of the European Society for Research in Mathematics Education*.

Nardi, E., & Iannone, P. (2003). The rough journey towards a consistent mathematical proof: The P(n+1) -> P(n+1) step in mathematical induction. In A. Gagatsis, & S. Papastavridis (Eds.), *Proceedings of the 3rd Mediterranean Conference on Mathematics Education* (pp. 621–628). Athens, Greece: Hellenic and Cyprus Mathematical Societies.

Oms, S., & Zardini, E. (2019). *The Sorites paradox*. Cambridge University Press.

Palla, M., Potari, D., & Spyrou, P. (2012). Secondary school students' understanding of mathematical induction: Structural characteristics and the process of proof construction. *International Journal of Science and Mathematics Education*, *10*(5), 1023–1045.

Pang, A., & Dindyal, J. (2012). Students' Reasoning Errors in Writing Proof by Mathematical Induction. In B. Kaur, & T. L. Toh (Eds.), *Reasoning, communication and connections in mathematics* (pp. 215–237). World Scientific Publishing.

Pascal, B. (1665). *Traité du triangle arithmétique avec quelques autres petits traitez sur la mesme matière*. Gvillavme Desprer.

Pasquotto, G. (1998). Maurolico e il principio di induzione. *Periodico di Matematiche*, *5*(2-3), 20–42.

Pedemonte, B. (2005). Quelques outils pour l'analyse cognitive du rapport entre argumentation et démonstration. *Recherche en didactique des mathématiques*, *25*(3), 313–348.

Pedemonte, B. (2007). How can the relationship between argumentation and proof be analysed? *Educational Studies in Mathematics*, *66*(1), 23–41.

Pedemonte, B., & Buchbinder, O. (2011). Examining the role of examples in proving processes through a cognitive lens: The case of triangular numbers. *ZDM*, *43*(2), 257–267.

Peirce, C. S. (1931). *Collected papers (CP, Vols. 1-8)*. Harvard University Press.

Piaget, J. (1929). *Child's conception of the world*. Harcourt.

Platone. (1981). *Dialoghi filosofici. Volume secondo* (G. Cambiano, Ed.). Unione tipografico-editrice torinese.

Poincaré, H. (1905). *Science and hypothesis*. The Walter Scott publishing co.

Polya, G. (1945). *How to solve it: A new aspect of mathematical method*. Princeton University Press.

Prediger, S., Bikner-Ahsbahs, A., & Arzarello, F. (2008). Networking strategies and methods for connecting theoretical approaches: First steps towards a conceptual framework. *ZDM*, *40*, 165–178.

Presmeg, N., Radford, L., Roth, W.-M., & Kadunz, G. (2016). *Semiotics in Mathematics Education*. Springer International Publishing.

Rabinovitch, N. L. (1970). Rabbi Levi Ben Gershon and the origins of mathematical induction. *Archive for History of Exact Sciences*, *6*(3), 237–248.

Radford, L. (1997). On Psychology, Historical Epistemology, and the Teaching of Mathematics: Towards a Socio-Cultural History of Mathematics. *For the Learning of Mathematics*, *17*(1), 26–33.

Radford, L. (2008). Iconicity and contraction: A semiotic investigation of forms of algebraic generalizations of patterns in different contexts. *ZDM*, *40*(1), 83–96.

Rashed, R. (1972). L'induction mathématique: Al-Karajī, as-Samaw'al. *Archive for History of Exact Sciences*, *9*(1), 1–21.

Reid, D. A., & Knipping, C. (2010). *Proof in mathematics education: Research, learning and teaching*. Sense Publishers.

Relaford-Doyle, J., & Núñez, R. (2018). Beyond peano: Looking into the unnaturalness of natural numbers. In S. Bangu (Ed.), *Naturalizing Logico-Mathematical Knowledge: Approaches from Philosophy, Psychology and Cognitive Science* (pp. 234–251). Cognitive Science Society.

Rips, L. J., & Asmuth, J. (2007). Mathematical Induction and Induction in Mathematics. In A. Feeney & E. Heit (Eds.), *Inductive Reasoning: Experimental, Developmental, and Computational Approaches* (pp. 248–268). Cambridge University Press.

Ron, G., & Dreyfus, T. (2004). The use of models in teaching proof by mathematical induction. In M. J. Hoines & A. B. Fuglestad (Eds.), *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, 113–120). Bergen, Norway: PME.

Sabena, C. (2008). On the semiotics of gestures. In L. Radford, G. Schubring, & F. Seeger, *Semiotics in mathematics education: Epistemology, history, classroom, and culture* (pp. 19–38). Sense Publishers.

Sabena, C. (2018). Exploring the Contribution of Gestures to Mathematical Argumentation Processes from a Semiotic Perspective. In G. Kaiser, H. Forgasz, M. Graven, A. Kuzniak, E. Simmt, & B. Xu (Eds.), *Invited Lectures from the 13th International Congress on Mathematical Education* (pp. 541–559). Springer International Publishing.

Sabena, C., Radford, L., & Bardini, C. (2005). Synchronizing gestures, words and actions in pattern generalizations. In H. L. Chick, & J. L. Vincent (Eds*.), Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 129-136). Melbourne, Australia: University of Melbourne.

Schoenfeld, A. H. (1982). Measures of Problem-Solving Performance and of Problem-Solving Instruction. *Journal for Research in Mathematics Education*, *13*(1), 31–49.

Schoenfeld, A. H. (1985). *Mathematical problem solving*. Academic Press.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics*, *22*(1), 1–36.

Sierpinska, A. (2005). On Practical and Theoretical Thinking and Other False Dichotomies in Mathematics Education. In M. H. G. Hoffmann, J. Lenhard, & F. Seeger (Eds.), *Activity and Sign* (pp. 117–135). Springer-Verlag.

Skemp, R. R. (1976). Relational understanding and instrumental understanding. *Mathematics teaching*, *77*(1), 20–26.

Soto-Andrade, J. (2014). Metaphors in Mathematics Education. In S. Lerman (Ed.), *Encyclopedia of mathematics education* (pp. 447–453). Springer Reference.

Stylianides, A., Bieda, K., & Morselli, F. (2016). Proof and Argumentation in Mathematics Education Research. In Á. Gutiérrez, G. C. Leder, & P. Boero (Eds.), *The Second Handbook of Research on the Psychology of Mathematics Education: The Journey Continues* (pp. 315–351). Sense Publishers.

Stylianides, G. J., Sandefur, J., & Watson, A. (2016). Conditions for proving by mathematical induction to be explanatory. *The Journal of Mathematical Behavior*, *43*, 20–34.

Stylianides, G. J., Stylianides, A. J., & Philippou, G. N. (2007). Preservice teachers' knowledge of proof by mathematical induction. *Journal of Mathematics Teacher Education*, *10*(3), 145–166.

Tall, D. (1977). Conflicts and catastrophes in the learning of mathematics. *Mathematical education for teaching*, *2*(4), 2–18.

Tall, D. O., & Schwarzenberger, R. L. E. (1978). Conflicts in the learning of real numbers and limits. *Mathematics Teaching*, 44–49.

Unguru, S. (1991). Greek mathematics and mathematical induction. *Physis*, *28*(2), 273–289.

Unguru, S. (1994). Fowling after induction. *Physis*, *31*(1), 267–272.

Vacca, G. (1909). Maurolycus, the first discoverer of the principle of mathematical induction. *Bulletin of the American Mathematical Society*, *16*(2), 70–74.

van Dalen, D. (2008). *Logic and structure* (4th ed.). Springer.

Van den Heuvel-Panhuizen, M., & Drijvers, P. (2014). Realistic mathematics education. In S. Lerman (Ed.), *Encyclopedia of mathematics education* (pp. 521–525). Springer Reference.

Vitrac, B. (1994). *Euclide D'Alexandrie. Les Éléments. Traduits du texte de Heiberg. Volume II: Livres 5-6: Proportions et similitude ; Livres 7-9 Arithmetique.* Presses Universitaires de France.

Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes.* Harvard University Press.

Young, J. W. A. (1908). On Mathematical Induction. *The American Mathematical Monthly*, *15*(8/9), 145–153.

Zan, R., Brown, L., Evans, J., & Hannula, M. S. (2006). Affect in mathematics education: An introduction. *Educational studies in mathematics*, *63*(2), 113–121.

Ringraziamenti

Il lavoro descritto in questa tesi e la tesi stessa non sarebbero stati possibili senza l'aiuto e il supporto di numerose persone ed istituzioni. I ringraziamenti che meritano vanno ben oltre queste poche parole.

Il primo grazie va alla Professoressa Veronica Gavagna e al Professor Samuele Antonini, miei tutori durante il percorso di dottorato. A loro va tutta la mia gratitudine per avermi guidato in maniera salda e paziente durante questo tortuoso percorso, indicandomi una direzione ma allo stesso tempo lasciandomi la libertà e il tempo di esplorare secondo i miei interessi, di perdermi, di sbagliare, di scoprire e di imparare.

In secondo luogo, vorrei ringraziare la Professoressa Francesca Morselli e il Professor Mirko Maracci, le cui attente revisioni di questo elaborato hanno permesso di migliorarlo e i cui commenti, profondi e stimolanti, potranno essere punto di partenza per ulteriore ricerca.

Un particolare ringraziamento va al Dipartimento di Matematica e Informatica "Ulisse Dini" dell'Università degli Studi di Firenze, che ha contribuito alla mia formazione come dottorando con seminari e corsi e che ha fornito materiali e spazi per effettuare la ricerca necessaria durante questi anni. Un ringraziamento sincero va anche al gruppo GNSAGA dell'Istituto Nazionale di Alta Matematica "Francesco Severi" che ha supportato economicamente diverse delle mie partecipazioni a conferenze e a scuole di dottorato.

Vorrei inoltre ringraziare l'Associazione Italiana di Ricerca in Didattica della Matematica per l'immenso sforzo e impegno con cui si prende cura della formazione di dottorandi e giovani ricercatori in didattica della matematica. Ogni incontro, scuola, conferenza, o seminario che ha organizzato e a cui ho partecipato è stato così ricco di contenuti e stimoli da renderlo fondamentale per la mia formazione.

Grazie ai protagonisti di questo studio, studenti, dottorandi e docenti che hanno deciso di dedicarmi tempo ed energie. Senza di loro questo lavoro non esisterebbe.

Grazie a Pietro, Giulio ed Iseabail per l'aiuto nella revisione della traduzione in inglese di questo elaborato.

Ci sono poi una serie di persone che colorano quotidianamente la mia vita rendendola speciale e per cui ogni ringraziamento sarebbe troppo breve. Grazie a Beatrice, Lorenzo, Andrea e Edoardo per esserci stati, sempre e veramente, che a separarci fossero cinquanta centimetri o migliaia di chilometri. Grazie agli amici del mondo del canottaggio, per i ricordi indelebili che conservo con orgoglio. Grazie al gruppo del Dini, compagni di studio negli anni passati e veri amici oggi.

Grazie alla mia famiglia, ai miei genitori Patrizia e Alessandro e a mia sorella Giulia per non avermi mai fatto mancare affetto e supporto e per il bene che ogni giorno mi danno.

Infine, grazie ad Iseabail, mia compagna di avventure. Isea, non ho parole per ringraziarti abbastanza. In questi anni sei stata la persona migliore che potessi avere accanto. La tua vicinanza, il tuo sostegno, il tuo amore mi hanno dato la forza nei momenti difficili e hanno reso ancora più unici i momenti felici. Sono contento di aver vissuto insieme anche questa avventura e, insieme, di essere arrivati in fondo.