Baizhanova, Aizhan

Reasoning, argumentative interaction and idea life cycles during group product ideation in higher education

Master's Thesis in Education

FACULTY OF EDUCATION AND PSYCHOLOGY

Master's Degree Program in Learning, Education, and Technology

April 2023

University of Oulu (Oulun Yliopisto)
Faculty of Education and Psychology (Kasvatustieteiden ja psykologian tiedekunta)
Aizhan Baizhanova
Master's thesis, 58 pages, 3 appendices
April 2023

Abstract

This study presents the analysis of the use of argument in group ideation process in higher education settings. The need for such analysis is dictated by the fact that students in higher education are one step away from joining wider professional communities, where the ability to engage in joint brainstorming and evaluating new products is in high demand. The study data consists of transcripts of ideation discussions of two groups of master's degree students. The task for both groups was to imagine and formulate a future AI-based teaching/learning assistant, prepare a short verbal presentation of the product, and present it to the whole class. The analysis is arranged in three steps. First, frequency and quality of grounded claims is evaluated using Toulmin's Argumentation Pattern. Then, the type of talk is determined using the indicators of exploratory, cumulative and disputational talk (Mercer, 1996), the interplay between types of talk is examined. Finally, idea life cycles and reasoning behind idea demotion is investigated. The results indicate that 1) arguments are provided rarely, but when provided, most of them (2/3) are complete; 2) exploratory talk manifests mostly in elaborations on peers' ideas, whereas reasoning (justifications) to own ideas and critical evaluation is less frequent; these factors characterise the discussions more as co-constructive interaction rather than exploratory talk; 3) dominance of elaborative comments on an idea leads to inclusion the idea in group solution; reasoning for idea demotion varies remarkably between the two groups (56% vs. 80%). These outcomes indicate that students might benefit from enhancing their reasoning to be ready for workplace ideation in groups. From task design view, clear product metrics should be set, and a line drawn between brainstorming and evaluation phase, to prevent unreasoned idea demoting in brainstorming and stimulate questioning and reasoning in evaluation.

Key words: reasoning, argument, exploratory talk, ideation

Contents

List of figures

List of tables

1. Introduction

Decades of research on collaborative learning equipped teachers with a powerful instrument for increasing students' learning outcomes (Macgregor, 1990; Dillenbourg, 1999; Kirschner et al., 2009). But in addition to "learning through the talk" (Gerlach, 1994), the use of this approach in higher education pursues the goal of better preparing students for working life (Foster & Yaoyuneyong, 2016; Mohan et al., 2018). Knowing that in many industries, the success of companies depends on creativity and productivity of cross-functional teams in developing innovative, customer-oriented solutions, universities strive to help students develop the ability to solve complex tasks in diverse, interdisciplinary, and cross-functional teams (Lovelace et al., 2001; Tancig, 2009). Cooper and Edgett (2008) pointed out that most firms seek proven approaches for their teams to generate "game-changing new product ideas". This call permeates the instruction in graduate programs and influences the design of course tasks and projects, by connecting these to working life and providing a space for students to jointly ideate, prototype, and then conduct feasibility studies of their solutions. Ideation is an important phase of such group projects. Depending on which ideation efficiency metrics is applied to measure the quality of ideation session (López-Forniés et al., 2017), the questions to group's solutions may include: how novel is the product, how useful is it for potential users, how does it compare to those existing on the market and how does it take in account ethical considerations. In this sense, course-related ideation discussion in higher education is not only a collaborative learning episode, but an event that is comparable to a workplace brainstorming session. To produce a novel and well-grounded solution, students, as future engineers, and designers, need to explore the available evidence and demonstrate divergent thinking (Runco, 2010; Cropley, 2015). They need to evaluate all ideas jointly and combine them, and if the latter is not possible, negotiate which ones to include in their product, maintaining transparent decision-making. Since this process is rarely facilitated by teachers, students themselves become responsible for their talk to be productive (Brown, 2002), for each other's arguments to be heard (Baker, 1999), and for solutions to be creative and purposeful (Liu & Lu, 2013). In this study, I suggest examining ideation talk in higher education through the lens of preparedness of students to act as

collaborative designers in their soon to come working life. In Section 2 below I set out the theories that informed this study.

2. Theoretical framework

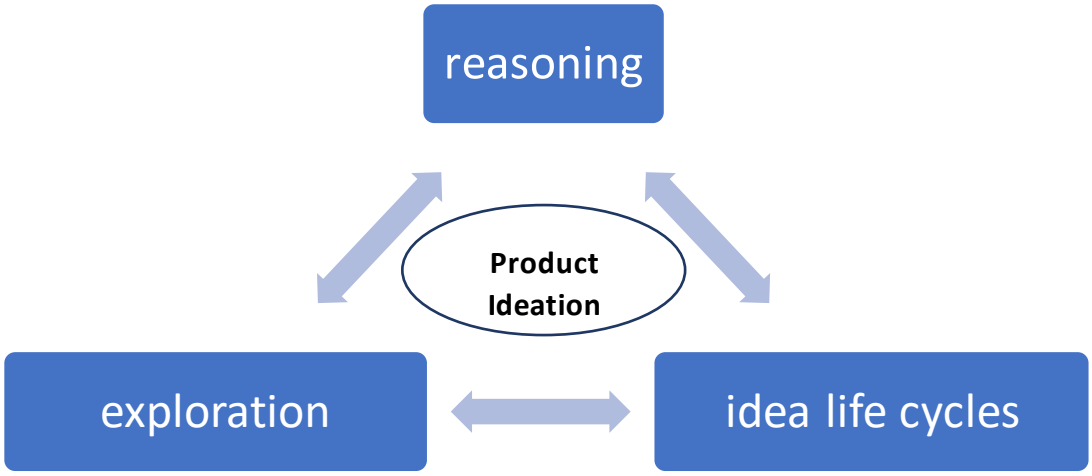
Studies on student talk are rooted in two major branches of thought, the one pioneered by Piaget (1971) focusing on inner cognitive development process, and the one initiated by Vygotsky (1978), focusing on co-operational interaction with society and culture. These were followed by studies of socio-cognitive conflict a learner's mind from interaction an alternative understanding of the matter (e.g. Doise and Mugny, 1984) and studies of how child's thinking and understanding improves through talking to the "more knowledgeable other" (Wells, 1999; Gibbons, 2002; Todd & Todd, 1979). In higher education group talk, when teachers do not intervene as "more knowledgeable ones", there is often a "symmetrical" talk, which flows in a variety of ways: some students explain and reason their standpoints and some choose to insist bluntly on their opinion, some agree to peers' suppositions to avoid conflict, and some participate very little and keep their ideas to themselves (Messer et al., 1993; Binkley, 1995; Light & Glachan, 1985). Barbieri and Light (1992) observed students talking in pairs while solving an adventure quest and concluded that pairs which displayed reasoning and collective decision making solved the problems more successfully compared to other pairs. Barnes and Todd (1977, 1995) suggested that the talk in which students justify their ideas and constructively criticize others' ideas, carries an exploratory nature.

However, in joint ideation, reasoning alone is not enough. Generation and development of ideas requires joint facilitation and ability to engage constructively in product ideation. By "facilitation" most authors in education field mean teacher facilitation (Hmelo-Silver & Barrows, 2008; Clifton, 2006), but in student-led ideation students themselves are viewed as facilitators (Brown, 2002). According to Kramer et al. (2001), good facilitation is the main condition of effective brainstorming. The last decade's developments in the field of computer-supported collaborative learning (Royle et al., 2010) may lead in the future to mass use of "intelligent conversational agents" (Adamson et al., 2014; Strohmann et al., 2017; Kumar & Rosé, 2011) for group brainstorming and knowledge co-construction. However, in current realities, the success of group work in higher education greatly depends on how active the students are in eliciting each

other's views and ideas when co-creating subject-related products. I suggest the combination of different lens to examine student ideation process, as shown in Figure 1.

Figure 1.

The lens used in this study to examine product ideation talk in higher education



In the next subsections, I outlay the theories on reasoning and argumentation, then overview theories on types of talks and their interplay and finish with synthesis of studies on brainstorming and ideas life cycles.

2.1. Reasoning and arguing. Argumentative interaction.

In broad sense, reasoning is "logical, thoughtful thinking" or a process of organising one's thoughts into a structure. To reason, or to claim something, one infers assumptions as building bricks for the claim; if the assumptions are incorrect, then the claim is not true (Walton, 1990). The smallest elements of reasoning are inferences, defined as new cognitions from previous information (Moshman, 1998). To explain development of reasoning, Moshman suggested to consider reasoning in three levels: i) beginning level is uncoordinated cognition, when inferences occur spontaneously in individual's mind to interpret social information and form new concepts, ii) next level is thinking, when individual purposefully coordinates the inferences with the same

aim, iii) upper level is reasoning, when individual verifies his or her thoughts through the prism of "self-imposed standards of evaluation". Individuals with more developed higher-order thinking are likely to display higher quality of reasoning, as they tend to apply the reasons correctly, use the inferences that are relevant, and think consistently (Walton, 2006). Exploring student reasoning, other researchers observed students solving fraction tasks and found out that when students were encouraged to share and discuss incorrect inferences, they displayed more varied reasoning in mathematical discourse and managed to clear existing misconceptions in the discourse (Mueller et al., 2012; Andriessen et al., 2003).

Reasoning is used in both explanation and argumentation, both processes are transactive, both aim to modify other person's understanding and often contain words indicating reasoning, such as "therefore", "since", "because". The difference is that in explanation individual uses reasoning to explain certain events or concepts by providing inferences, while in argumentation individual uses reasoning to increase the acceptability of his/her viewpoint by the conversational partner (van Eemeren et al., 1996; Baker, 2002; Leitão, 2000). Arguing is a process that is wider than reasoning. Binkey (1995) defines arguing as an attempt of verbal (linguistical) influence on others' judgement by giving reasons, where "giving reasons" is making statements aiming for the listeners to accept these statements "as a premise in a reckoning which the audience itself is constructing". In other words, if a person picks a side and intends to persuade the audience that that side is the correct choice for that audience then he or she provides reasons for it; in this context reason and argument mean the same.

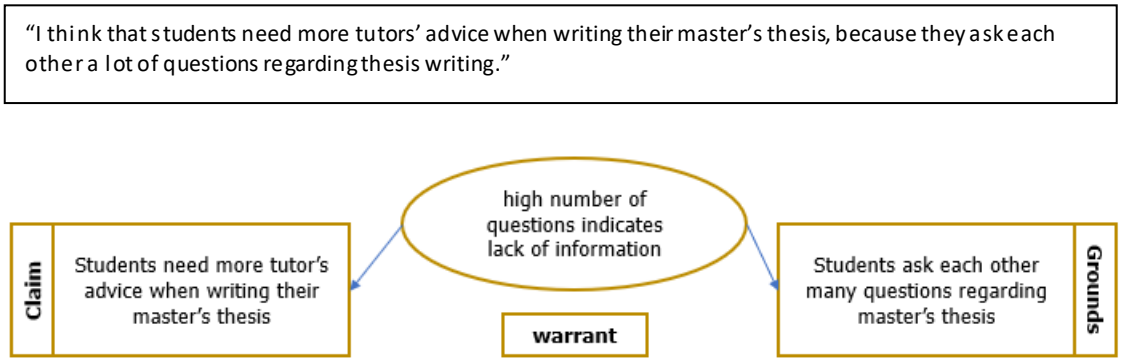
Toulmin et al. (1979) suggested the model that shows argument as pattern consisting of several interrelating components. In Toulmin's Argument Pattern (TAP) a "claim" is an assertion that needs to be supported by "grounds", the two together constituting an argument. Between the two, a "warrant" required, answering the question "how do these grounds support this claim?". If the warrant is not accepted, then the whole argument may be rejected. In addition to these components, TAP also includes backing (support to warrant), qualifier (the degree of certainty), and rebuttal (alternative valid point).An example of an argument is shown in Figure 2.

Nielsen criticized Toulmin's argumentation model, one of his main concerns being that education researchers view argumentation as dialectical and empathetic in nature, aimed at finding the truth through consensus, whereas Toulmin's method, in Nielsen's view tends to narrow the discourse analysis down to a scheme of components (Nielsen, 2013). Nielsen argued

that anyone applying TAP for analyzing classroom argumentation was forced to supplement and adjust their analysis to reflect its dialogical features. This view is supported by the conclusions of Andrews (1994), who recommends adjusting the expectations from "ideal argument" towards an argument relevant to the situation it is observed in.

Figure 2.

Toulmin's model of argument



I partially share Nielsen's critics, because ideally constructed arguments are not always the main purpose of dialogic activities. On the other hand, if the requirement for argumentation is entirely dismissed, it might impact the relevance and purposefulness of group solutions. Therefore, in this study I test the applicability of TAP to arguments put forward by higher education students in their ideation talk.

Considering multiple definitions of argumentation, to avoid cluttering of concepts Baker (1999) suggested separating the dimensions of interactions in educational settings into dialectical, epistemological, conceptual, rhetorical, and interactive and to name verbal interaction that manifests dialectical dimension as "argumentative interaction". Discussing argumentative interaction in educational settings, Baker (1999) states that reaching an agreement in joint problem-solving requires students to modify their initial suggestions to make them acceptable to the peers in the process through negotiation. Similar conclusions are set out by Roschelle and Teasley (1995), who stated that when students argue and persuade their peers, they "experience an intersubjective dissonance and socio-cognitive conflict" which may result in reframing and expanding the concepts they previously internalized. Mevarech and Light (1992) suggested that

cooperative resolution of reasoning in argumentative interactions increases students' cognitive ability. While arguing academically, students tend to employ more of their cognitive and linguistic abilities to defend their own views or challenge others' opinions (Baker et al., 2019). Argumentation practices foster development of students' scientific debate skills (Lehrer & Schauble, 2006, Lave & Wenger, 1991). The thought that seems to connect all these studies is that value of clear argumentation for educational purposes is based in its capacity to transform students' thinking.

Argumentation in educational settings is also conditioned with the sense of purpose. Andrews (1994) puts the question straightforward: "Why do we argue?" and enlists the following functions of argument: clarify (to determine issues or people's positions), persuade (to change others' feelings or thinking), win (to defeat others' arguments), entertain (to make people feel and think, as in comical stand-ups), unload (to release interpersonal tension between individuals), resolve (to come to consensus, to end conflict), find identity (as when teenagers oppose their views to parents' ones to mature). Of all these, *clarify, persuade,* and *resolve* seem to be the most applicable for small group work in higher education settings, as students often have different ideas, yet they are expected to produce one solution from the group.

To conclude on absence or presence of argumentation, Osborne et al. (2004) concluded that one of the signs of it taking place is the frequency of rebuttals, when students challenge the evidence provided by their peers. Other authors suggested that key indicator quality argumentation is correct use of scientific content as evidence supporting the claims (Sampson and Clark, 2008; Maloney & Simon, 2006). To a certain extent these parameters of quality argumentation overlap with indicators of exploratory talk, which will be discussed in the next section.

2.2. Mercer's typology of talk, indicators of exploratory talk, and ground rules

The idea of exploratory talk that uses language for reasoning belongs to Douglas Barnes (Barnes & Todd, 1995) but the first researcher who identified and described three distinct types of groups talk in joint problem-solving activities was Neil Mercer (1994, 1996). In this section, I will first set out the notions of symmetry and productivity of talk and touch upon ground rules as pre-requisite of talk productivity, then outline three types of talk by Mercer (1996) and the indicators

by which these three types of talk can be determined, as suggested by Polo and colleagues (Polo et al., 2015).

Productivity of classroom talk has been discussed in different dimensions, such as academic achievement and high-level discourse (Fisher, 1993; Gaskins et al., 1994), variance in rates of participation in a classroom (Sedova et al., 2019) and desirable prosocial behaviors, such positive attitude towards student of other races and ethnicities (Cohen, 1994). The studies on how classroom talk can support learning discuss the factors that lead to difference in talk productivity in group problem solving activities.

By symmetry the equal knowledge of the subject matter among participants is meant (student-student talk), whereas asymmetry in talk means that there is more knowledgeable other leads the conversation towards the solution that he or she is confident in (teacher-student talk). Barbieri and Light studied interactions in dyads and discovered that symmetrical talk is a way in which the learners can construct knowledge (Barbieri & Light, 1992). Building on these findings, Mercer observed triads of students working on problem-solving tasks without teacher's interference and noticed that some talk sequences were richer in alternative ideas and contained more elaborations than others (Mercer et al., 1999). Following these observations, Mercer and colleagues outlined the conditions for talk productivity. The first condition is that thinking aloud should be encouraged and it should be purposeful, because it allows peers to hear and react to each other's ideas and reasoning. The second condition is that the task should be designed carefully to promote cooperation rather than forcing students to compete. When the task is competitive, the battle of opinions is more likely to occur. The third condition is that the goal of the activity should be clearly explained by teachers and well understood by students. Mercer (1999) also pointed to the properties of productive talk: thinking is made visible to others and decisions are made collectively. The latter means that students persuade one another, analyze, and compare their arguments before a decision is made. Having set out these conditions and properties, Mercer draws readers' attention to the contextual factors of the pedagogical situation in which he collected the data: the principles the teachers were guided by while designing the activities, the learning objectives of the tasks and the ways in which computer programs may constrain student talk.

With all these contextual factors considered, Mercer (1996) sorted talk sequences by pattern of progression and revealed three distinct types of talk, which can be described as follows.

Disputational talk is characterized by short assertions, often aimed at showing the speaker's own knowledge without an intention to seek evaluation or alternative opinion. Such talk is essentially an exchange of opinions which rarely leads to expansion of understanding of the matter by conversational partners. Disputational talk resembles playing a tennis game. Speakers in disputational talk express disagreement in short messages, without explaining the reason for such disagreement. Disputational talk is characterized by unsupported assertions and sole decisions, where the participants do not seek each other's input or advice, do not create a shared pool of existing resources. Disputational nature of talk may also be caused by the design of the task. If the task is very competitive or too timing-restricted, for example if the scores are assigned for being the first group to give the right answer, that reduces the space for negotiations to take place.

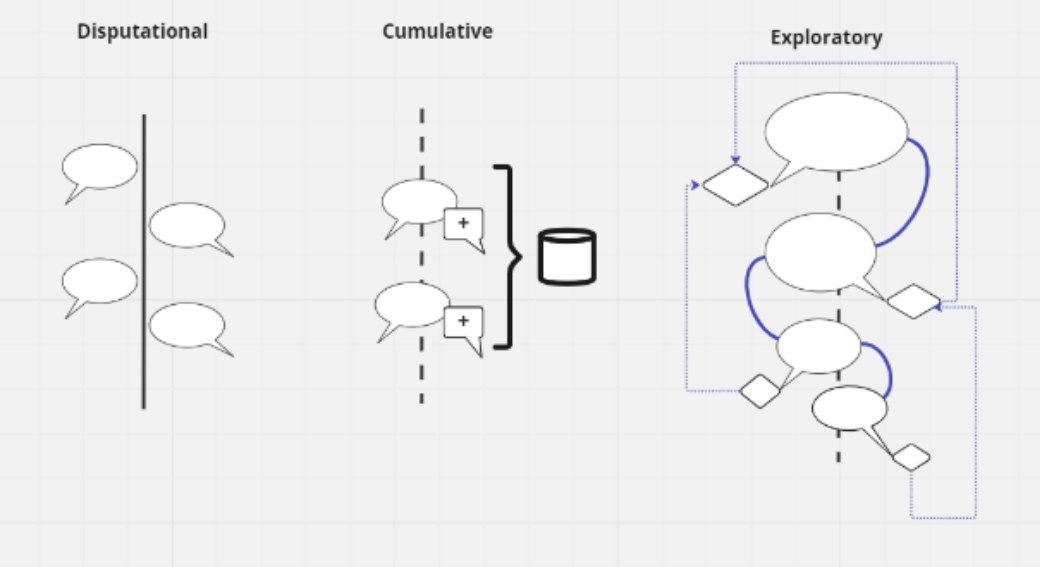
Cumulative talk is characterized by smooth and peaceful flow, in which speakers put forward their ideas and propositions without much reasoning and justifying, and these propositions are confirmed or positively validated by peers without challenging them. Common view is built through piling up resources without testing if they are valid, frequent confirmations and repetitions. This kind of talk is likely to occur between students who know each other well, or when conversations partners for any reason try to avoid any disputes. This type of talk can often be observed in informal conversations.

Exploratory talk can be distinguished from the other two types by participants evaluating each other's proposal and using reasoning to build up shared understanding of the matter. In this type of talk, participants suggest options with grounding attached to them and then jointly discuss them to resolve unclarities and doubts that may arise. In problem solving activities, exploratory talk means that students justify their suggestions and question other's suggestions constructively. Whenever it is claimed that an idea is wrong, justification is provided as to why it is wrong and alternative ideas are put forward. Once agreement is reached on a certain point, the group strives for all members to recognize and accept the progress before proceeding to the next steps, thus ensuring collective ownership of the solution.

These three types of talk do not always appear in pure form, a group discussion may sometimes include all three of them intertwined. For example, cumulative talk may be way to share information and a chance to synthesize a new, shared opinion as a group based on the

combination of the shared information (Mercer, 1994). Schematically the three types of talk can be presented as shown on Figure 3.

Figure 3.
Schematical representation of Mercer's types of talk



According to Mercer (1996), in both exploratory talk and disputational talk students criticize and challenge each other's ideas occurs, but the main difference is that in exploratory talk, when refuting a peer's suggestion or positions, students provide grounds to their refutation, for example "I don't think this is right, because…", whereas in disputations talk refutations are not grounded, for example "No, you are wrong". In exploratory talk, if two reasoned ideas collide, students put forward alternatives to resolve the controversy, whereas in disputational talk the conversational partners keep their standpoints not willing to rethink their understanding. Exploratory patterns of talk are observed more often in spaces where ground rules for exploratory talk are instilled and sustained by teachers (Mercer et al., 1999). This thought underpins the large-scale intervention course called "Talk Lessons" as part of "Thinking Together" project involving several primary schools in Milton Keynes, UK. To design this intervention, including the training for teachers and tasks for lessons, the key findings from previous phases of research were incorporated. One of the key documents published on the project's website is named "Ground Rules" which lists the set of rules to be explicitly taught and

emphasized during group activities, as shown in Figure 4. Child friendly version of the ground rules was used in Talk Lessons project, where each of ground rules began with "we will", for example "we will give reasons for what we say" or "we will try to agree about what we think".

Figure 4.

Ground rules for exploratory talk

Everyone in the group is encouraged to contribute.
Contributions are treated with respect.
Reasons are asked for.
Everyone is prepared to accept challenges.
Alternatives are discussed before a decision is taken.
All relevant information is shared.
The group seeks to reach agreement if views are different.

The concept of ground rules was criticized by Lambirth (2006), who questioned the idea of dividing classroom talk into valuable and less valuable. Lambirth viewed Mercer's proposition of explicit teaching of ground rules as debatable. Firstly, referring to O'Neill (1970), Lambirth explains that when children are taught to talk in certain way, their own inner 'proper literacy", ability to draw on their own culture and knowledge begin falling apart and the children become less confident in valuing and utilizing own experiences. Further, Lambirth argues that when children talk to make meaning, such talk does not need to follow any ideal pattern. He asserts that various ways in which talk occurs deserve to happen in classrooms, as these are natural ways for children to express their thoughts. Lambirth further explains that cultural, ethnic and language peculiarities influence speaker's way of conversing, therefore an attempt to break these peculiarities may lead to devaluating speaker's language and culture; imposing certain "technology" of productive talk may result in the feeling of alienation in those who do not fully align to that technology due to their social and ethnical origins.

Here are some propositions that can be set out here to partially resolve Lambirth's critics. Firstly, Mercer (1999) analyzes the talk at three levels: linguistic (content and function), psychological (students' interests and concerns, opposition, competition, defensive stance even when doubting

own assertions or the contrary, trust and willingness to reach an agreement), and cultural (signs of wider society's influence on how students talk). With all these aspects considered, Mercer (1999) suggests that usage of language as sociocultural medium for sharing and processing information and observance of ground rules lead to togetherness of learners rather than separation (Mercer, 1996). He also explains that types of talk are three recognizable "social modes of thinking" rather than strictly defined categories as they are influenced by various emotional and socio-cultural factors.

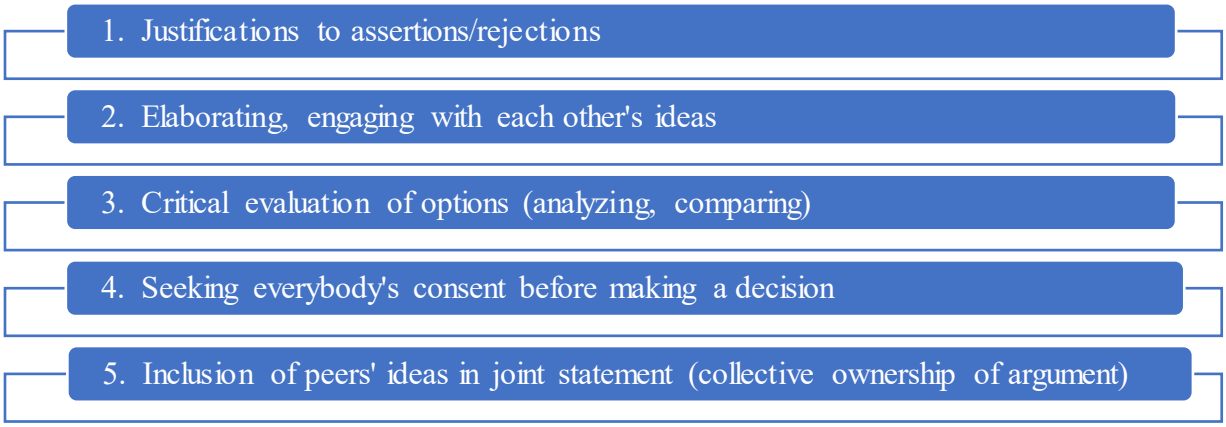
Secondly, transferring this paradigm to higher education, some researchers examine the challenges that students from non-western cultures confront when they start their studies at Western universities (e.g. Jones, 1999). Jones points out that that possession of academic discourse skills is necessary for these students to benefit from their studies, but acquiring such skills should not compromise their own cultural values. Examining cross-cultural aspects of academic group discussions, Jones (1999) states that the objective in such settings is not to transform the foreign students, but to help them adapt to the interactional manner of classroom discussions. The same careful and reflexive approach to this problem can be noticed in the work of Ryan and Viete (2009), they explain that in the era of globalization of education, it is time for universities in Australia and other Western countries to update their pedagogical practices to make it easier for students from other cultures to blend into "new discursive environment". Overall, these debates revert to the importance of academic debate skills, in which reasoning, elaboration and critical evaluation play pivotal roles. For example, the two of the ground rules "reasons are asked for" and "alternatives are discussed" are equally appropriate for problem solving talk and product ideation discussion, because at the end of both processes, groups should be able to explain all pros and cons of their solution,

In this study, the observance or non-observance of ground rules for exploratory talk will be evaluated through presence or absence of indicators of exploratory talk, which are discussed in more detail in the next subsection.

Elaborating on Mercer's typology of talk, Polo and colleagues (2015) operationalized the types of talk through how they manifested in high school students' discussions on societal equity issues. The observations took place in the US and France during "scientific café" after-school club sessions. The groups of students were requested to discuss and agree on one of the choices enlisted in the task related to economic aspects of drinking water distribution. To classify a talk

sequence as exploratory, Polo and her colleagues worked out five indicators of exploratory talk based on Mercer's definition, as shown in Figure 5.

Figure 5.
Indicators of exploratory talk by Polo et al.



The study demonstrated that while working on an open-ended task, high school students are not "locked" within one type of talk throughout their discussion. The higher the relevance of an issue to students' own experiences was, the more exploratory talk was present in a sequence. Moreover, it became clear that a cumulative talk sequence may serve as warming up phase for an exploratory sequence that follows it. Some degree of agreement to these findings is found in works of Wegerif, who pointed to the risk that requirement to justify each contribution may impact negatively on the discourse which aims co-creation rather than finding the right answer (Wegerif, 2005). Also, Barron (2003) found in triads' work observation that "most successful group had high rates of affirming, agreeing, and accepting remarks", which if converted to Mercer's (1996) framework, are the signs of cumulative talk.

When applying Mercer's framework to non-mathematical group work situations, Rojas and colleagues (2006) also questioned to what extent explicit reasoning is necessary if the task is not about finding the correct answer but about co-creating a text summary. In their study Rojas et al. suggested taking the need for explicit reasoning out of the equation and measuring the transferability of the rest of the features of exploratory talk in creative tasks. They discovered that children from experimental group who received the training on productive talk applied

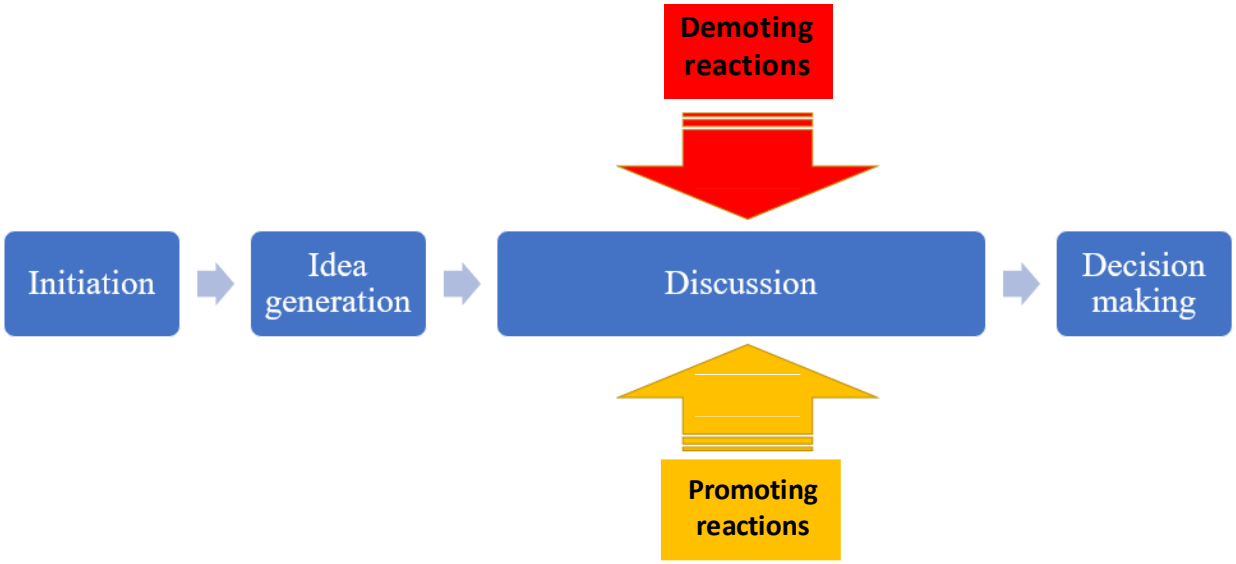
ground rules, such as turn taking, offering alternative ideas, considering each other's suggestions, seeking everyone's acceptance, in their work on creative tasks. Reporting this result, Rojas et al. bring about the concept of "co-constructive talk", the talk that does not require explicit reasoning for interim choices but appears to be fruitful for co-creation tasks. Rojas et al.'s conclude that exploratory talk is a form of co-constructive interaction that is best suited for mathematical, logical, and problem-solving tasks. In the analysis part of this study, I apply the indicators of exploratory talk developed by Polo and colleagues (2015), indicators of cumulative and disputational talk based on Mercer's definitions (1996) and the interplay between the three types of talk.

2.3. Idea life cycles: promoting and demoting reactions

The question of how ideas evolve during group work in educational settings was studied mostly in problem-solving situations. Brigid Barron in her article "When Smart Groups Fail" observed and compared interactions in triads of students to figure out why in some triads the correct ideas were ignored and in others the correct ideas were discussed and documented (Barron, 2003). Barron classified the reactions of peers to correct proposals into three categories: "accept", "discuss", "reject/ignore" and found that more successful groups accepted and discussed correct proposals more often than less successful groups. Barron's approach provides suitable lens to evaluate the effect of peers' reactions to ideas evolution in well-structured problem, where a single correct answer is sought for by students. However, in this study, the students are not looking for a single correct answer, they co-create a product, negotiating what it should be and what functions it should have. They are expected to think about customer needs, the degree of product utility for customer, and ground their viewpoints. Therefore, I felt that Barron's approach should be tailored to ideation activity by findings from studies on brainstorming (Osborn, 1953) and productivity of idea-generating groups (Diehl & Stroebe, 1987). Isaksen (1998) listed the following rules of productive brainstorming: i) judgement of ideas should be postponed until a later stage; ii) freewheeling is allowed: suggestions can be wild and unusual; iii) blocking is not welcomed: the more ideas the better, iv) improving and combining: participants can suggest how other's ideas can be improved or combined with other ideas to result in more ideas. Ideation is a process of forming ideas that includes brainstorming phase.

Strohmann (2017) studied AI-based intelligent moderators (IMO) and their influence on ideas in brainstorming (Strohmann, 2017). When programming the IMO, Strohmann considered various idea-demoting factors, such as "Fifty phrases that kill creativity". Among these 50 phrases, some appear more relevant to business settings, such as "It's against company policy" or "It isn't in the budget", but other phrases may well relate to group ideation in educational settings, for example: "I don't like the idea" or "It's impossible", "Good thought, but impractical". The intelligent moderator in Strohmann's study (2017) was able to detect killer phrases and intervene when conversation drift away from the topic. In Figure 6 illustrates a simplified representation of idea life cycle through peers' reactions.

Figure 6.
Reactions from group members to idea



In the analysis part of this study, I apply this simple approach to analysis of ideas evolution, by classifying students' reactions to ideas as "promoting" or "demoting" within idea thread and tracking inclusion of ideas in the group solution. I then evaluate whether demotion of ideas is reasoned or not.

3. Aim and research questions

The frameworks outlined in the theoretical background part of this study were traditionally tested in school settings in problem-solving activities (reasoning, exploratory talk) or in workplace brainstorming. However, these frameworks have not been applied in the context of ideating activities in higher education. Knowing whether students use arguments in this context and how their arguments are constructed might reveal the need for the tasks to be requiring argumentation. I believe that understanding of what kind of talk prevails may help to assess the preparedness of students to learn from engagement in scientific debate and participate in design taskforce in their fore coming careers. I also believe that tracing the ways how ideas are promoted or demoted in ideation in the absence of external facilitators may bring insights on how the self-facilitative discussions can be scaffolded in groups where such skills are weaker than in others. This study aims to fill these gaps by answering the following research questions:

RQ 1: How often do students argument their views? Are their arguments complete?
RQ 2: Which type of talk prevails in ideation and how do types of talk interplay?
RQ 3: How peers' reactions influence the idea life cycles?

4. Method

In this section I set out the details of the study design and the analytical tools that were used to approach the above listed research questions. Count and detailed analysis of arguments put forward by students was conducted to answer RQ1. Further, to answer RQ2, I conducted deductive discourse analysis using the indicators of exploratory talk tested by Polo and colleagues in "scientific café" settings (2015) and described the most frequent patterns of interplay between different types of talk. Exploring RQ3 required detecting all idea threads in both groups and presenting them as summary showing promoting or demoting reactions. Supposedly, more promoting reactions should lead to inclusion of an idea in group solution. For cases of idea-demotion, I expect to find grounding provided to these demoting reactions.

4.1. Participants and data collection

A total of 10 higher education students participated in this study, of them 1 male student and 9 female students. The mean age of participants is M = 30.02 (SD=5.16). Of the 10 participants, 8 represented the whole cohort in an interdisciplinary master's program and 2 participants were exchange students from another university joining this cohort for one semester. The participants come from 8 countries, with professional backgrounds in elementary school teaching, software development, marketing, management, and other fields. All participants were given detailed privacy notes and gave consents for the data collected during the session to be used for this study. The study was conducted in full accordance with has research ethics guidelines of the Finnish Advisory Board on Research Integrity (TENK, 2019). Participants names were changed for the purpose of the study, to exclude any risk of personal identification.

Videotaping was arranged, with participants' consent, with the only purpose for the author to identify who the utterances are pronounced by. Videotaping was also used to figure out which objects participants point at when they pronounce phrases like "I mean like this one". As soon as transcriptions were finalized, the videos of the conversations were deleted.

4.2. Experiment Task Design

The task was identical for both groups. Conditions for group product were made to mimic the nature of tasks that are given in interdisciplinary higher education in courses like "Educational Technology Project", "Entrepreneurship Mindset in Education", "Emergent Technologies in Education". Since this cohort of degree students were acquainted with fundamentals of learning sciences and were introduced to current trends in educational technologies, the task was expected to be of interest for them by activating their prior knowledge. The main goal was to provoke problematizing and debating on various features of the join product: its name, functions, usefulness, and consequences of its use, because according to Cheon et al., exchange a lot more ideas in the process than can be seen in the final design (Cheon et al., 2019). Cheon et al. also indicated that a closer look at co-designing process allows to see "soft conflicts" arising when

societal and cultural implications of the design are discussed. I also refer to Osborne et al., (2004) who suggested that tasks asking participants to take a particular position (for or against) trigger argumentation. Dunne and Raby enlist the following forms of design fictions: speculative design, radical design, design for debate, discursive design, critical design, future-scaping, which all require active dialogue between participants (Dunne and Raby, 2013). The fictional future objects in design fictions are *believable and plausible*, so it is often used to elicit participants' hopes, concerns and underlying values while they discuss the implications of the fiction objects coming to being (Baumer et al., 2020). Tanenbaum (2014) also mentioned that participants of design fiction often talk about ethical and social concerns around the design objects even without being prompted to do so, as they connect the concepts and future technologies into their own life experiences and emotions.

When used in pedagogical situations, tasks related to creating future objects increase student engagement due to their speculative and entrepreneurial orientation (Maxwell et al., 2019). Barnes suggested that if students are expected to engage in dialectic and truly exploratory talk, they need to feel free and not be afraid of saying wrong things (Barnes, 2008). I assumed that futuristic nature of group tasks reduces the fear of voicing a silly idea, as there is a tacit assumption that everything is possible in future. While co-creating a non-existing object, creative power of the group can be unleashed in search of fresh ideas (Lavonen et al., 2004). Referring to these studies, the task for participants of this study was designed to allow them space for imagining, speculating, and drawing on their previous experiences as teachers and students. The purpose of the workshop was announced as "Ideating future AI-based educational technology". Task Description was read aloud, and printed versions handed out (Appendix A). The task required the participants to imagine themselves in the year 2042 in a country called "Sunland". It was explained to participants, that in that imaginary future they are no longer master's degree students, but two competing R&D teams who both wish to win funding announced by Sunland government for designing the best product which would assists students in higher education who are overwhelmed by their studies or the best product to assist university teachers who complain that their wellbeing is endangered by the workload. The debates in the groups were expected to be about: category of the object (only 2 options: learning assistant or teaching assistant), name of the object (no restrictions), functions of the object and appearance of the object (no restrictions).

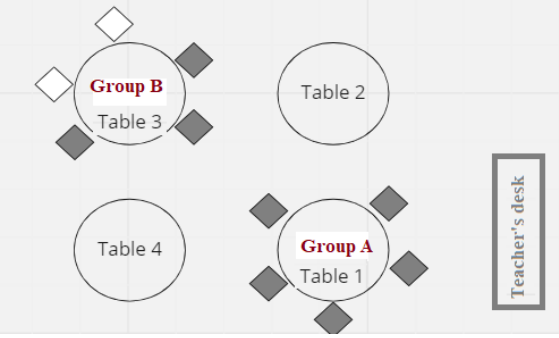
4.3. Design of the study

Data for this study was collected during a purposefully designed workshop that took place at the university campus, in one of the rooms equipped for classes about educational technologies. For all participants this was one of their normal locations, to avoid any discomfort. The participants were aware of the workshop time and venue, but not aware of the topic of the workshop. No preparation was required from the participants, no materials were sent in advance, except the privacy notes. It is worth mentioning that prior to this experiment workshop, the participants attended their program courses which contained topics related to the topic of the experiment, for example, "Learning Environments and Technologies" and "Entrepreneurship mindset in education". Such shared prior experience could make it easier for the participants to begin their ideation during the experiment workshop.

The room had 4 round desks, each housing up to 8 people. Upon arrival at the class, the students took places on their discretion. Group seating is shown in Figure 7, in which grey rhombuses stand for students from the same cohort and white rhombuses stand for exchange students from another university.

Figure 7.

Group seating plan



The seating details are provided for possible connection between the groups structure and argumentation dynamics during ideation discussions in these groups. The length of the workshop was 40 minutes, divided into four phases as shown in Table 1.

Table 1.

Workshop phases

Workshop phase	Duration
1. Greeting, explanation the purpose of the workshop and the group task	5 minutes
2. Ideation discussions in groups	25 minutes
3. Statement to whole class from each group	6 minutes
4. Answering questions from opponent groups and wrap up	4 minutes

4.4. Instrumentation

Transcripts of group discussions and group statements were prepared by listening to audio files and typing all the utterances in a table in WORD. Video files were used to identify who the utterances belong to. In the case of two or more group members speaking at a time, the utterances of individuals were not audible, therefore these were not transcribed. Where speakers used informal versions of words such as "gonna" instead of "going to", these were preserved without editing. The transcription was checked for typographical errors and all turns were numbered. Ideation discussions included a total of over 603 speaking turns in both groups (368 speaking turns in Group A and 236 speaking turns in Group B) and statement to whole class made by 1 representative from each group. At the end of each group presentation, the other group was given an opportunity to ask a few questions regarding the product described by the presenting group. The answers to these questions given by the presenting group were treated as part of group's presentation during the analysis related to RQ1.

Following the meaning-making check of text, the transcripts of group talks were uploaded to NVIVO software and coded for the study section. The first layer contained 9 codes for various types of talk, including 1 code overlapping between RQ1 and RQ2 ("grounded claims"). The second layer of coding, idea threads, resulted in 36 codes: 19 idea threads in Group A and 17 idea threads in Group B talk. To ensure coding reliability, a second coder was invited, who is a master's degree student in the same program but from a different cohort. One round of

negotiations with the second coder was conducted to achieve common understanding of formulations and meanings of codes related to RQ1 and RQ2. The codes for RQ3 were not tested with interrater reliability check, because i) I assumed that detecting all utterances related to an idea by one coder is less subject to subjectivity and ii) coding idea threads would involve coding of the entire transcription, i.e. too large scope to assign on volunteer basis. The fragment of transcript sent to the second coder contained 96 out of 603 speaking turns (15.9%). For interpretation of Kappa values, I used ranges suggested by Landis and Koch (1977): 9 tested codes fell into the following ranges: not present in the test fragment of talk - 1 code, moderate agreement - 4 codes (0.41-0.60), substantial agreement - 2 codes (0.61-0.80) and perfect agreement - 2 codes (0.81-1), these values are presented in Tables 2 and 3 below.

To answer RQ1, the transcription of both groups discourse was analyzed through the prism of Toulmin's Argumentation Pattern. Firstly, all utterances containing argument were labeled as "grounded claims" and their share in total number of utterances determined. Utterances containing argument are mainly those in which words indicating reasoning are found: "because", "since", "otherwise", "therefore", but also where such connection is not expressed with a reasoning but is clear from the context, for example "I think that it should be a name that says that it is a teacher". The cases where a grounding to an assertion of refutation was provided in a later turn were not classified as grounded claims. To exclude the cases where the use of reasoning words is nominal and to establish whether the argument is full, all grounded claims were decomposed using simplified version of TAP: claim, grounds, and warrant. The objective of this part of the analysis was to understand which of the grounded claims has strong, difficult-to-dispute warrant within their structure.

To answer RQ2 the transcripts of group discussions were coded using pre-determined codes (Hsieh & Shannon, 2005) corresponding to indicators No 1, 2, 3, 4 of exploratory talk by Polo and colleagues (2015). The codes for indicators 1-4, examples of utterances matching these codes and Cohen's Kappa values for the codes are shown in Table 2. The presence of indicator No 5 was examined through analysis of group statement transcripts (Appendix C). On every instance that that an idea expressed during ideation was found in group statement, I indicated the code of relevant idea thread and treated it as manifestation of indicator No 5 of exploratory talk, based on definition given by Polo and colleagues.

Table 2.

Indicators of exploratory talk with relevant codes and examples from transcripts

No	Indicator by Polo et al	Code in NVIVO	Examples of utterances	Cohen's kappa
1	Justifying assertions and refutations	Grounded claim	*It should take notes of student's mental health, you know there are students with OCD and you have this robot*	0.73
2	Elaborating on argumentative content of previous turns	Building on an idea	*So, it will be like a teacher assistant "Good morning*	0.62
3	Critical evaluation of others' arguments	Criticize, evaluate	*And I think it is a good function – the teacher won't sit and assess the work*	0.46
4	Seeking everybody's consent	Calling for votes	*Does everybody agree with this?*	0.94
5	Inclusion of peers' ideas in joint statement	Separate detailed analysis of each group's presentation was conducted to track this indicator		n/a

To understand the shares of different types of talk in ideation process, I deepened the analysis by creating 5 codes for utterances related to cumulative and disputational nature based on how Mercer (1996) defined them: disputational is short messages of disagreement, non-grounded assertions, sole decisions, cumulative is peaceful, non-critical build up, checking the task and mere repetition of peer's ideas. The utterances as "yeah" and "hm-mm" were not counted as cumulative utterances because their meaning cannot always be determined with high degree of confidence. For example, in some cases "hm-mm" carries a shade of doubt in it, a message like "This sounds right, but for me to agree to it, please elaborate and give some reasons". Therefore, I deemed that "yeah" and "hm-mm" utterances constitute a "positive backchannel" encouraging the speaker to continue (Sannomiya et al., 2003).

The quantification of cumulative and disputational talk instances was necessary to determine the proportions between the types of talk in both group discussions. It was also necessary to investigate what are the most frequently occurring patterns of interplay between the types of talk.

Table 3 shows 5 codes, that were used to highlight the utterances of cumulative and disputational talk, with examples and respective Cohen's Kappa values.

Table 3.
Codes for utterances indicating cumulative and disputational talk.

Type of talk	Code in NVIVO	Examples of utterances	Cohen's Kappa
Cumulative	Repetition of peer's words	*"So, we want it to be less distracting"*	n/a
	Checking the task	*"Here it says that ..."*	0.87
	Sharing information	*"So, nowadays we also have AI in LMS, so.. It is already collecting data for everything"*	0.58
Disputational	Ungrounded assertion or refutation	*"Record keeping of the students"* *"No-o-o"*	0.52
	Pushing own idea (repeating own suggestion after it was ignored or rejected)	*"What does it look like? Like this camera"*	0.42

Finally, to answer RQ3, I used the third layer of coding, devoted to idea production and idea evolution. For this part of the analysis, "idea" means any suggestion regarding the category of the product, its name, appearance, or its functions. If an idea did not cause any reaction from peers, it was labeled as "single utterance idea" and set aside. If an idea did entail some discussion, it was labeled as idea thread which includes all utterances related to such idea. Further, within each idea thread, I marked all reactions as either "promoting" (elaborating, clarifying, supporting) or "demoting" (expressing doubt, ironical comment, direct cancelling, or rebuttal) and tracked whether dominance of promoting reactions leads to inclusion of the idea in group solution or not. For each instance of demoting reaction, I also sought if such demoting contains any reasoning in it. This inquiry links the results of analysis in RQ3 to in RQ1 and allows to conclude on how reasoning manifests in idea demoting.

5. Results of the analysis

In this section, the results of the analysis are presented. The section begins with thorough consideration of use of argument: number of argument-containing speaking turns and their weight in total number of speaking turns is illustrated, as well as completeness of arguments according to TAP (5.1.). Further, quantity and comparative share of exploratory talk utterances is evaluated and their distribution across the discussions shown on conversational. Several talk sequences illustrating the most frequent patterns of interplay are discussed (5.2.). Further, summary of all idea threads is shown with indicated types of peers' reactions (promote/demote) to such ideas (5.3). The influence of these reactions to idea inclusion or non-inclusion in group solution is illustrated and discussed.

5.1. Frequency and completeness of arguments

Arguments were found in 34 out of 603 speaking turns (5.6%), both groups taken as one. The difference in frequency between the groups is in the same range: 21 of 368 speaking turns in Group A (5.7%) and 13 of 235 speaking turns in Group B (5.5%).

As was explained in p. 4.4., where the instruments of this study are described, the code "Grounded Claim" was used for two purposes: 1) within RQ1 analysis, to code all instances of use of argument and 2) within RQ2 analysis, to code all instances of justifications to own ideas as the first indicator of exploratory talk. Depending on how a talk sequence flows, grounded claims were found either as grounding to ideas related to product features or grounding to refutations against the ideas; in rare cases argument was found in "thinking aloud" utterances.

Some of the arguments are complete, while in others the grounds and warrant need eliciting. For example, a short utterance from John "It would look better" does not indicate whether the product would look better for potential users of the product, or it is according to his personal taste. Same concerns Anna's "…because it has to help the teacher", which is a grounding detached from the claim, so could be related to one of the suggestions expressed in previous

turns. In such cases, the whole argument requires re-assembling to be evaluated by conversational partners.

The completeness of arguments, which was evaluated as presence of claim, grounds, and warrant, varies significantly depending on which participant expressed it; some participants consistently provide full arguments, while others limit themselves with shortened arguments. The rigidity of warrants themselves in arguments also varies significantly: some students provide grounding that is difficult to argue against, while others provide grounding that is easily refutable due to less rigid warrant.

Figures 8 presents a complete argument with a rigid warrant, put forward by Toney in Group B. In this argument by Toney, the grounding is "sustainability" and the warrant is s "generations of students can use the platform". The expression "I'm thinking" leaves a space for rebuttals and elaborations by others

Figure 8.

Example of complete argument (and more rigid warrant)

'My opinion is that we would focus on students, to see if there is… I don't know. Why? I am thinking about sustainability. Because teachers… hm.. now I'm thinking that teachers are going to retire, if there is a learning assistant. All the students when they come there, they will learn in the platform how to do their work. And every generation can use that platform

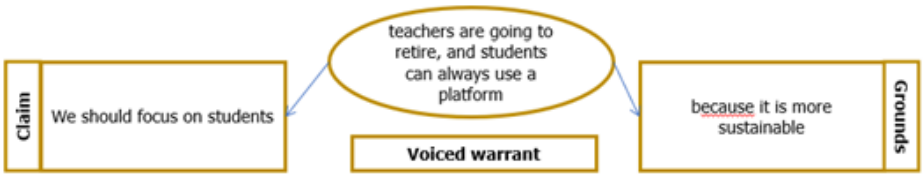


Figure 9 below shows a complete argument, but with less rigid warrant, put forward by Jonna in Group. Jonna asserts that the product must have the function of highlighting assignments because "this is a teaching assistant". The warrant in this argument is that "highlighting students' assessments" is a teaching assistant's duty. This warrant could be challenged by peers, as this statement might not be true for all educational systems. However, the expression "of course" makes the whole argument sound assertive and not welcoming objections.

Figure 9.

Example of complete argument with less rigid warrant

"So this is a teaching assistant, so of course, it should be highlighting the assignments that were given".

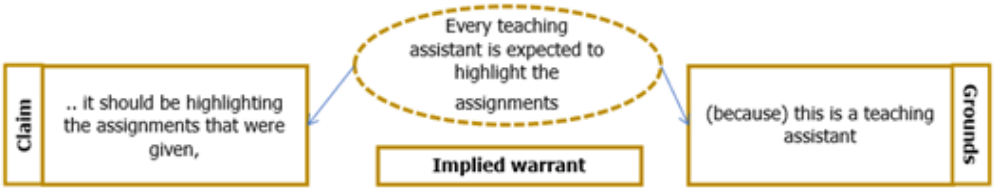
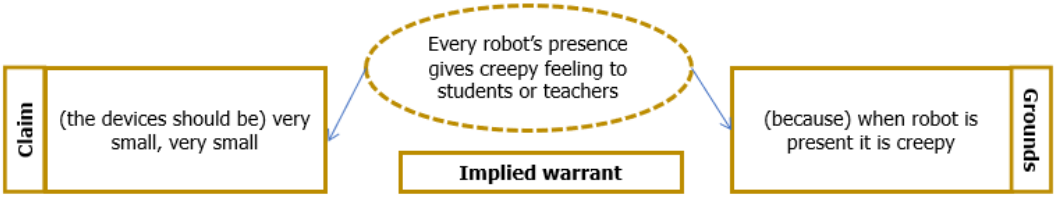


Figure 10 shows an incomplete argument with hardly extractable warrant. Felicia puts forward an argument against a robot (or flying robot) because it would be "creepy", but the reasoning is not quite clear: whether robot itself is creepy or the fact that it would be flying around the students. If Felicia voiced a warrant to her claim (e.g. "students might feel uncomfortable about a machine sitting among them"), this would open a space for grounded rebuttals (e.g. "students in future might be used to interacting with robots"). However, since the reasoning was not made explicit, this statement did not lead to either dispute or elaboration.

Figure 10.

Example of incomplete argument (difficult-to-extract warrant)

Felicia: "Very small, very small. Because if it's like a robot it's creepy then… flying around…".



In both groups, the quality of argument on many instances is questionable, for example, in Group A, Anna pushes her peers towards a decision to create a teaching assistant by building the

argument into a prompting question: "*Which one would you want to go with (teaching assistant or learning assistant)? I mean, we are all teachers?*"; in this format, the argument is less visible and less open to critical evaluation by peers.

Overall, almost 2/3 of the grounded claims (21 out of 34) put forward by participants in both groups were complete arguments with either explicitly formulated or easily extractable warrants.

5.2. Prevailing type of talk and interplay between the types of talk

Indicators of exploratory talk were found in ideation and group statements of both groups; summary is presented in Table 4. Concerning Indicator 4 "Seeking everybody's consent", utterances relating to voting occurred very rarely. I found only two instances of explicit voting Group A talk, which as voting for Product Category (teaching assistant or learning assistant) and votes for Product Look ("should it be a screen or non a screen") and only instance of explicit voting in Group B talk, which was voting for Product Category.

The rest of voting in both groups was not explicit. It appears that absence of objections is treated by group members as general consent. Another non-explicit way of voting is when student taking the notes announces what he or she is writing, tacitly inviting anyone who does not support the idea to speak up.

Table 4.

Summary of indicators of exploratory talk in ideation discussions

	Group A		**Group B**	
Justifications (grounded claims)	20	22%	13	15%
Elaborations	50	54%	38	45%
Critical evaluation	13	14%	25	30%
Seeking general consent	4	4%	3	4%
Inclusion of peer's ideas	6	6%	5	6%
Total	**93**	**100%**	**84**	**100%**

The presence of Indicator 5 "Inclusion of peers' ideas in joint statement" was detected through analysis of group statements. For this purpose, I must introduce the term "presenter". Presenter is

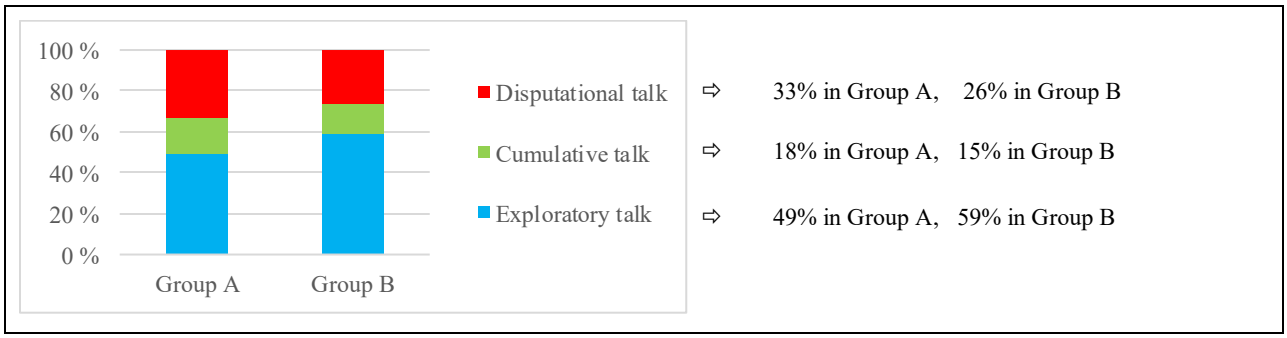
a member of a group who is elected by a group to present that group's statement to whole class. When presenting the group's solution, the presenter is expected to tell the whole class all decisions made by the group in narrative style. Overall, Group A group statement contains 7 ideas, of which 1 was initiated by presenter herself and 6 were initiated by her peers. Group B group statement contains 12 ideas, of which 8 were initiated by presenter herself and 5 were initiated by her peers.

Total number of exploratory talk instances in 93 in Group A and 84 in Group B. The next step is to compare these totals with totals of other types of talk. In Figure 11 below, proportions of each type of talk are shown for both groups of participants. Based on this diagram, it can be noted that Group B is engaged in exploratory manner of discussion more than Group A which is compensated by higher number of disputational type utterances.

Most disputational utterances carry the code "pushing own idea" which means repeating own idea even if it was ignored or demoted in an earlier turn. The "records holders" on pushing their ideas were Felicia in Group A and Nadine in Group B, each of them repeated their suggestion more than 3 times, but the suggestion was ignored and not written down by the person who volunteered taking the notes, and therefore was not included in group statement. This fact draws attention and will be more closely examined in p. 5.4. of this study.

Figure 11.

Share of types of talk within utterances coded as specific type of talk



Text analysis shows these disputational utterances are mostly those coded as "pushing own idea" than "ungrounded assertions". Throughout both group discussions, exploratory, cumulative, and exploratory utterances are tightly intertwined.

Most frequently occurring talk alternation pattern is that first group elaborates and agrees on a suggestion, but when a doubt is voiced, then the reason for that is sought by another group member and it triggers a constructive debate. Table 5 shows one of such sequences.

Table 5.

Fragment of talk where a dispute is triggered by expression of doubt/evaluation.

317	Felicia	Have we written the.. interest or no?	
318	Anna	No, what?	
319	Felicia	Like it detects their interests	Disputational (pushing own idea)
320	Anna	How?	Exploratory (seek clarification)
321	Felicia	During teaching or something or	Exploratory (give clarification)
322	Kelly	Conversation…	Exploratory (elaborating)
323	Felicia	Conversation, because I can be in this group, but I wasn't interested, I was just sitting like and agreeing with them, but I'm not interested in whatever you're saying	Exploratory (give clarification)
324	Anna	Something that detects the student's interests	Exploratory (elaborating)
325	Felicia	Yeah	
326	John	I would feel uncomfortable...	Exploratory (evaluation, critic)
327	Kelly	Why?	Exploratory (seek clarification)
328	Felicia	Whether you are an individual or collaborative learner..	
329	Jonna	Why? Why are you uncomfortable?	Exploratory (seek clarification)
330	John	Cause whatever she is saying the robot should do	Exploratory (give clarification)
331	Jonna	Ha-ha-ha	
332	Felicia	But you want the robot to have a head and human body and now you are talking about this. (shows kind of mask around her face)	Disputational (ungrounded rebuttal)
333	John	I feel that it is invading my privacy.	Exploratory (give clarification)
334	Jonna	Exactly! There is a point where technology can do something,	Exploratory

	and technology can't.	(elaborating)

In the above fragment, John expresses unwillingness for robots to detect students' interests, and Anna requests clarification by asking "Why would you feel uncomfortable?". John's reasons his doubt with incapacity of robots to act indepdently and possible privacy invasion. Jonna supports that, mentioning the line between what technology can do and what it cannot do. It might seem that the idea is rejected by the group, but it appears in the notes (see Appendix B "Copies of group notes") and it is clearly declared in group statement (line "m", Appendix C).

Another frequent exploratory sequence pattern is when a group member brings up evidence from Task Description, and that triggers a chain of comments building on one another, resulting in a new idea. One of the sequences matching such pattern is shown in Table 6.

Table 6.

Fragment of talk where checking task description leads to generation of new idea.

32	Natasha	Is it, like, important for teachers, because it says, it can do facial, like recognizes images and emotions..?	Cumulative (checking task)
33	Veronica	Maybe it could scan the classroom and see if there are children who are like that.. ha-ha.. (The group laughs) maybe questions in their eyes.	Exploratory (elaborating)
34	Natasha	Or if someone is disturbed maybe? If there are like 30- to 40 students, then it is difficult to..	Exploratory (elaborating)
35	Veronica	Yeah, to see everybody	Cumulative
36	Nadine	So, it can scan and then send some questions to students, maybe more to like to, to like change their mood	Exploratory (elaborating)
37	Natasha	Do some funny things, to change their mood?	Exploratory
38	Nadine	Yeah, changing their mood and emotions, if they are bored	Cumulative
39	Toney	If we are going to present something, let's take notes	
40	Veronica	But that would be on the learner's side, right, not on the	Exploratory

		teacher's side? If it would do something funny for pupils?	(evaluation)
43	Veronica	I'm not sure, maybe it's both (laughs)	Cumulative (compromise)
44	Natasha	It's both, it's both. It will help the teacher to certain analysis	Cumulative (compromise)

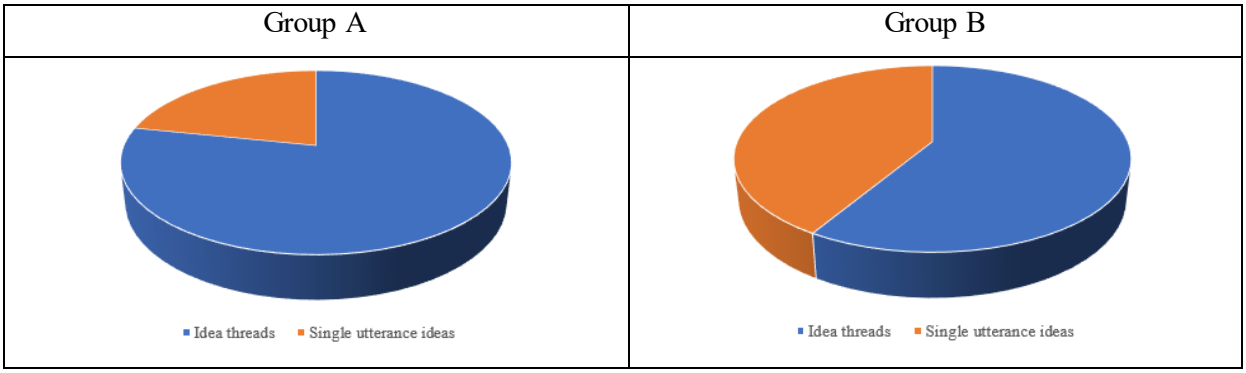
It begins with Natasha's checking the task and reading out a condition, that AI in 2042 will recognize images and basic human emotions with much higher accuracy than it does at present time. Veronica picks up that information and suggests that the AI learning assistant may could scan the classroom and detect children who have questions, and Nadine picks up on that and suggests that based on that scanning AI learning assistant could support children and change their mood. When Veronica expresses a doubt that such a function changes the category of the object, Natasha resolves that by stating that it can have both teaching assistant and learning assistant functions.

5.3. Influence of peers' reactions to idea life cycles

To approach the idea life cycles, I first classified all ideas generated in both groups in two major categories: 1) those that were followed by at least one reaction, forming an idea thread and 2) those that were not discussed, I refer to them as "single utterance idea". An example of single utterance idea is Jonna's phrase that was pronounced during discussion of AI-based teaching assistant functions: '*So it could be even 'I need a sandwich; can you go and get me a sandwich',* which was not followed by any verbal comments, except laughter. As there is no forming of an idea thread, single utterances ideas were excluded from further analysis. Figure 11 below shows proportion of idea threads to single utterance ideas in groups: 18 to 5 in Group A, 18 to 12 in Group B. Ideas were more often discussed Group A, which might indicate that members of this group are more receptive to each other's suggestions than in Group B.

Figure 12.

Shares of idea threads and single utterance ideas



Since the experiment task design implied no right or wrong ideas, even wild ideas could be considered. Whether an idea is good or bad was up to the groups to decide. During the analysis it was noticed that in both groups demoting reactions often took the form of doubting or challenging the appropriateness of suggestion, rather than direct critics. In many instances a demoting reaction was followed by an alternative proposal which was then promoted by other members. Table 7 illustrates the snowballing effect of promoting reactions (turns 60-78)

Table 7.

Promoting reactions snowballing effect

Participant, Turn	**Contents of the speaking turn**	**Type of reaction**
Jonna [turn 60]	So, if a student has issues, like it's an addition, it's maybe two months into the term, and if you are still having same issues, you..	Start of idea thread
Anna [turn 62]	You are showing the same mistakes, so repetition	Promoting
Jonna [turn 65]	It's detecting you have a problem with... from assignments, 'cause on this one question you still keep not getting it	Speaker

Anna [turn 66]	Not improving in?	Promoting
Anna [turn 68]	Analyzes errors?	Promoting
Jonna [turn 69]	It can come in between, like "Do you want me to schedule time with the teacher?"	Speaker
Jonna [turn 71]	Schedule time with the teacher, so he can do extra practice	Speaker
Jonna [turn 77]	Tutoring (offering gesture)	Speaker
Anna [turn 78]	Analyses your weaknesses and schedules your tutoring sessions?	Promoting

Table 8 illustrates another pattern, how demoting reactions lead to appearing of an alternative suggestion (turn 120-125).

Table 8.

Demoting reactions leading to alternative suggestion

Felicia [turn 120]	It should like a bag or something, it should be something you can wear, you know, these face masks	Start of idea thread
Jonna [turn 121]	Hmm.. would be intrusive, an AI?	Demoting
Anna [turn 122]	Yeah, if you are wearing it?	Demoting
Felicia [turn 123]	Not like wearing it, but you know like these face masks, they do	Speaker
Anna [turn 124]	I think if it is robot, then it should feel that there is someone sitting with us	Demoting+ alternative
Kelly [turn 125]	Yeah, someone friendly	Promote the alternative

The hypothesis for this part of analysis was that if promoting contributions prevail over demoting contributions, then such idea should be included in group solution, and vice versa. In this part of analysis, I refer to ideas showing dominance of promoting reactions as "candidate ideas".

The analysis of Group B idea threads (Table 9) shows that the logic "dominance of promoting reactions in an idea thread leads to its inclusion in group statement" proves to apply. I found 9 "candidate idea", of them 6 were found in group statement and 3 were not (A-1, A-5, A-9). There was an anomaly about idea A-1: it was found in the group notes (Appendix B) but not verbalized by the presenter in groups statement (Appendix C). Dominance of demoting reactions in an idea thread led to non-inclusion of an idea in group statement in all cases.

Table 9.

Promoting/Demoting reactions vs. idea inclusion in group statement. Group A

Idea ID	Relates to	Description of the idea	Quantity of speaker's own elaborations	Quantity of promoting / demoting reactions	Is idea found in group statement?
A-1	Function	Keeping the records of students' attendance	4	**1 / 0**	No
A-2	Function	Detecting emotions, correlating mood with academic progress, cheering	13	**6 / 0**	Yes
A-3	Function	Organizing, reminding, to-do lists for student	8	0 / 1	No
A-4	Function	Checking assignments, finding knowledge gaps, scheduling sessions based on the found gaps	14	**4 / 0**	Yes
A-5	Look	Robot-humanoid, physically present with students	2	**6 / 0**	No
A-6	Look	Goggles or mask-type device	2	0 / 2	No
A-7	Look	Compact object to be put on a desk	3	**8 / 0**	Yes
A-8	Function	Screening students' mental health and initiating student counselling if needed	1	**3 / 0**	Yes
A-9	Look	Gender-neutral device	1	**1 / 0**	No
A-10	Function	Assesses students' collaboration skills	3	0 / 0	Yes
A-11	Look	An object following humans around	1	0 / 2	No
A-12	Look	A wristband teaching assistant	1	0 / 1	No
A-13	Look	An camera-type object flying above	1	0 / 1	No
A-14	Name	Nameless device (user gives it a name)	3	0 / 6	No
A-15	Name	Gaynor	1	1 / 2	No
A-16	Name	Olga	2	**3 / 0**	Yes
A-17	Function	Detecting students' interests	6	**6 / 4**	Yes
A-18	Function	Analyzing students' behaviors	2	3 / 5	No
A-19	Type	It should be a Teaching Assistant	5	1 / 0	Yes

The fact that draws attention in Group A summary of idea threads is that certain ideas (A-2, A-3, A-4) contain high numbers of speaker's own elaborations, and two of them were finally accepted while one was not successful. This might indicate that either the speaker's reasoning is not persuasive for peers, or that the speaker is not attentive to peer's refutations.

The analysis of Group B idea threads, shown in Table 10, partially confirms the logic that dominance of promoting reactions leads to inclusion of idea in group statement: of 11 candidate ideas, 7 were found in group statements and 4 were not. In one idea thread (B-8) the demoting comment from peer did prevent the idea from inclusion in group solution.

Table 10.

Influence of promoting and demoting contributions on ideas inclusion in Group B.

Idea ID	Relates to	Description of the idea	No of speaker's own elaborations	No of promoting / demoting reactions	Present in group solution?
B-1	Type	It should be a product for learners	2	0 / 0	No
B-2	Function	Grading students' assignments, providing feedback on essays	1	**1 / 0**	Yes
B-3	Function	Spot sleepy or distracted and notify teacher	1	**2 / 0**	No
B-4	Function	Detecting and changing students' mood, cheering them with sending a voice message	2	**4 / 1**	Yes
B-5	Function	Voice message as feedback from teacher	1	0 / 1	No
B-6	Function	Creating personalized assignments	1	0 / 0	Yes
B-7	Function	Registering attendance by scanning faces	1	**1 / 0**	Yes
B-8	Function	Finding out a reason for absence via email	1	0 / 1	Yes
B-9	Function	Sending extra materials, helping with content	2	**4 / 1**	Yes
B-10	Function	Clone of teacher for every student	9	1 / 5	No
B-11	Function	Assisting with class management, substituting a teacher, repeating a material	1	**1 / 0**	No
B-12	Name	Sun-Master	1	**1 / 0**	No
B-13	Function	Recognizing emotional problems and redirecting them to the human teacher	3	**1 / 0**	Yes

B-14	Look	Hologram	2	**1 / 0**	No
B-15	Function	Back-office duties, entering grades into system	2	0 / 0	Yes
B-16	Name	Teach-Up	1	**1 / 0**	Yes
B-17	Name	SunBot	1	0 / 1	No
B-18	Look	It should be a platform or LMS	11	4 /1	Yes

Further, I bridge the analysis of idea threads with the use of reasoning (RQ1), by tracing whether reasoning was provided to demoting comments. In Group A, 4 of 8 of demoting reactions are reasoned, and the other 4 are not:

1. In thread A-3, Felicia elaborated persistently on her own idea for AI-TA to keep diaries, to-do lists for students, which was demoted by Jonna with one phrase "Don't go into to-do lists". **[unreasoned demoting]**
2. In thread A-6, Felicia suggest that AI-TA should look like goggles or face mask, to which Jonna says that it would look intrusive and Anna supports Jonna's demoting. **[reasoned demoting]**
3. In thread A-11, Felicia suggests that AI-TA could follow the students around, to which Jonna says "hm-hm" shaking her head negatively and Anna says "No, it does not follow you around" and then suggests that it should sit at the desk together with students. **[unreasoned demoting]**
4. In thread A-12, Felicia suggest that AI-TA could be something wearable, like bracelets or a bag, to which Anna says jokingly "Are you wearing your teacher assistant?" and group laughs, so the idea fades. **[unreasoned demoting]**
5. In thread A-13, Kelly suggests that AI-TA could fly above the classroom, to which Anna says jokingly "I want my TA to fly" and group laughs, so idea fades. **[unreasoned demoting**]
6. In thread A-14 Felica insists that AI-TA should not have any name so that student can name it as he or she wants, to which Anna, Jonna and John persuade her that to submit a prototype to government, the device must have some name, and Felicia agrees with peers by suggesting a name "Anu". **[reasoned demoting]**

7. In thread A-15, John suggests the name "Gaynor", but Jonna says that it sounds like a decease and Jonna supports Anna's view, so the proposal does not go through. **[reasoned demoting]**
8. In thread A-18, John suggests that AI-TA could do higher-level analysis beyond just checking the assignments, but Anna asks what teachers will do if that it done by AI, Jonna supports the demoting. **[reasoned demoting]**

Analysis of demoting reactions in Group B shows that reasoning for demotion is provided more often than in Group A, it is found in 4 of 5 cases:

- In thread B-4, Veronica suggests that AI-LA could spot students being bored or sleepy and peers elaborate on that, but then Veronica demotes her own idea by doubting that such function makes it more teaching assistant than learning assistant. **[reasoned demoting]**
- In thread B-5, Nadine suggests that AI-LA could send a voice message to student on teachers' behalf, to which Toney says that such function could lead to robot saying something that teacher is not aware about, which might lead to confusion among students **[reasoned demoting]**
- In thread B-8, Toney suggests that AI-LA could email a student who missed a class asking why he or she was absent, to which Veronica gently objects by suggesting an alternative of AI-LA only entering the absence information into the system **[unreasoned demoting]**
- In thread B-9, Nadine suggests that AI-LA could send tutorials and materials of the missed lesson to the student who was absent, to which Veronica says jokingly that it would be so easy for students that she herself would not go to lessons in such case. Group laughs and idea fades. **[reasoned demoting**]
- In thread B-10, Nadine suggests that AI-LA is a virtual clone of a teacher, so that each student could have one "copy", to which Natasha objects that such function would make teachers' job redundant, and when Nadine keeps pushing this idea, Toney adds personal reason, that if she was a student, she wouldn't something like that. Finally, Toney suggests a compromise for students to have choice to opt for it. **[reasoned demoting]**

When grounding is provided to a demoting reaction, it "closes" idea thread logically. Perhaps if sticky notes were used in the brainstorming phase, it would help the participants to summarize and see the reasons of why some ideas were not included in group solution.

6. Discussion

The aim of this study was to investigate how reasoning occurs in product ideation by students, what type of talk prevails and how idea life cycles are influenced by peer's reactions. The results indicate that the use of reasoning is tightly connected to determining the type of student talk. There would be greater presence of exploratory talk if students were willing to reason their suggestions and refutations more often. The idea threads analysis demonstrated that demotion of ideas varies significantly between the groups. In this section, I discuss these outcomes in detail.

6.1. Reasoning is rare, but most of students' arguments are complete

Despite the rare occurrence of argument (grounded claims), I feel that interpretation of RQ1 results should include a few more factors, namely by connecting these occurrences with the results obtained in exploring RQ2. I suggest viewing justifications, elaborations, and evaluations (i.e. exploratory talk indicators 1-3) as utterances which should ideally be reasoned, then the sum of these utterances is the expected (ideal) quantity of arguments. The relation of actual quantity of the arguments found in ideation talk to the expected quantity of arguments is 0.25 for Group 5 and 0.17 for Group B. The relation is still low but taking in account that experiment task required participants to brainstorm and evaluate within one session, this can be explained by the desire of participants to capture as many functions in their product as possible, leaving less time for justifying and comparing. While examining argumentative sequences in the frame "argument-counterargument-integration", Stegman et al. (2007) concluded that external scripts can be effective in improving the quality of argumentation. In this study context external collaborative scripts could help eliciting the grounds to claims, making these more visible to peers. In this study, the participants were asked to imagine themselves as members of R&D teams and the task included competitiveness: "the best product wins the funding". The aim behind these conditions

was to trigger argumentation between group members. Kuhn et al. (2008) pointed out that engagement in argumentative interaction leads to build of domain-general knowledge on argumentation among students.

During the experiment, some participants tried to retrieve grounding of others' ideas by asking "why do you think so?". Could this function be automated to develop higher quality discussions? In computer-mediated discussions, a trigger to reason can be included as programmed prompts upon detecting "candidate statements" or assertions without grounding as suggested by Adamson et al. (2014). Lastly, from the skill development perspective, referring to Vygotsky (1978) 's views, students' argumentation can be enhanced by arranging their participation in ideation in real-life projects with professionals from the industry. The professionals would act as "more knowledgeable others" to model scientific argumentation, increasing students preparedness to take part in product ideation, reason their ideas and requesting reasons for others' ideas.

6.2. The talk on product ideation is co-constructive rather than exploratory

The results revealed that although exploratory talk prevails over disputational and cumulative talk, this prevalence is achieved mostly due to elaborative utterances (indicator 2), whereas justifications and critical evaluations (indicators 1 and 3) manifest significantly less. That allows to conclude that product ideation in the chosen context is co-construction interaction rather than exploratory talk (Rojas et al., 2006). The observed frequent pattern of idea generating sequences after sharing of information and checking the task prove the conclusions by Polo and colleagues (2015), who said that that cumulative phase is sometimes necessary for new ideas to emerge. In many instances, ideas are put forward in the form of thinking aloud, in form of self-questioning, followed by "I don't know.." tacitly inviting other opinions. This resonates with a statement by Wegerif and Mercer (1997) that in exploratory talk, participant identifies him or herself with the dialogue rather than with his or her own self or the group. Long elaborations on certain suggestions were found in both groups discussions, which ensures one of the ground rules formulated by Mercer, 1999: "alternatives are discussed before a decision is made". Although Wegerif (2005) argued to what extent reasoning is needed in creative tasks, this study showed that in many cases proper grounding to ideas does not block idea production, but instead opens a space to evaluate, compare, and integrate different ideas, which in its turns triggers new

alternative ideas. Overall, it appears that the combination of cumulative and exploratory talk provides favorable ground for diverse ideas to be brought about by group members.

6.3. Idea life cycles are shaped by peers' reactions, demoting is mostly reasoned

The analysis demonstrated that there is a strong link between the nature of peer reactions and the inclusion of an idea in group solution. Most of idea threads with higher quantity of promoting reactions were found in the group statements. Idea threads in which demoting reactions prevailed, did not appear in group statement, with a few exceptions. Since this study did not aim to rank the ideas themselves, I based on assumption that all ideas are worth considering, which is one of the ground rules for exploratory talk. From this perspective, reasons for demoting reactions were examined. Reasons for demoting were provided more often in one group than in the other (50% in Group A vs. 80% in Group B). This observation resonates with findings of Barron (2003) where the lack of attention to others' ideas might lead groups to fail in solving their task. In ideation activities not paying enough attention to new ideas can lead to narrowing of spectrum of ideas, filtering out the ones that deserve consideration. Lastly, the results of RQ3 indicate the role that students own facilitation of their talk influences the ideation process. Offner et al. (2006) stated that facilitated groups produce significantly more ideas than groups without facilitators, however, in this study active self-facilitation by students can be observed, with both groups generating 18 ideas within 20 minutes.

7. Conclusion

The overarching conclusion of this study is that student reasoning and exploratory approach is present in product ideation in higher education. Practicing group ideation, with "light", not too invasive scaffolds, such as "why do you think this function should be included?" should be tested to enhance students' preparedness to participate in reflective and reasoned co-designing of services or products in their future jobs. In this study, the content of students' ideas indicates their awareness of product metrics (novelty, ease of use, usefulness) and the discussions were

mostly guided by those metrics, although the word "because" is not the most popular in these discussions. Despite rare reasoning, it is promising that even without external moderation, students directed their discussion in cumulation-evaluation sequences, creating pools of ideas and filtering those pools asking themselves "what is needed for teachers, for learners?" and attempted to combine the best of their ideas.

8. Implications

This study has implications for research in the field of product ideation in higher education, considering that co-construction and co-creation skills are in high demand in many companies. Whether these students decide to become researchers or join the industrial markets, the habit of using ground rules – justifying, building on, critically evaluating, transparent decision-making – adds positively to their skillset for both spheres. The ability to reason is especially needed in conceptual design in every industry (Liu & Lu, 2013). The study may also be of interest for research in the development of intelligent conversational agents or AI-based support for brainstorming.

9. Limitations

This study has several limitations. The first limitation is the single mode of observation: only transcribed voice recordings were used. Addition video data might help to evaluate non-verbal interaction, observe non-verbal reactions to ideas, etc. Second, it is small number of participants (10 students) and timing: knowing that they only have 20 minutes to produce a description of the product could cause a rush in groups, stopping them from justifying or evaluative comments. Also, not time was given to particpants to re-write the notes made during group discussion, which could impact the completeness of group statement. Third, it is the reliability of measurement: the average Cohen's Kappa value for 8 codes found in the second-rater fragment is 0.64, but certain codes are in range of 0.41-0.60. Especially challenging in reproducing the results could be differentiating certain codes, such as "elaborations" and "critical evaluations" due to linguistic nuances of talk of students for whom English is a second language. Same

concerns differentiating disputational utterances, such as “ungrounded assertions”. Fourth, the results could be distorted by the absence of physical or online whiteboard that participants could use to keep their ideas in groups’ visual field for categorizing, identifying relationships, framing problems (Peterson & Barron, 2007). Fifth, the investigation of brainstorming phase did not consider the studies on production blocking in idea-generating processes (Nijstad et al., 2003). Taking in account the number of interruptions omitted by the participants, not observing the turn taking, the inclusion of production blocking analysis could significantly supplement the analysis and reveal more argumentation patterns within idea threads.

However, the results illuminate important aspects of group ideation that are not often visible to instructors but potentially useful from the perspective of task design, scripting, and evaluation of quality of argumentation and level of participation with available instruments.

10.Ethical considerations

Participants of the study were provided the following information about the study: running title, researcher name, context, and purpose (master’s thesis). The privacy note handed to participants enlisted their rights according to GDPR regulations. Names of participants in transcripts were changed to exclude identification. Video-files were deleted as soon as the transcripts were finalized. It was made clear to participants that they may exit the experiment at any point of time. After 6 months from the moment of completion of the study the data source of this study will be deleted from the university servers. These measures ensure the ethical conduct of the research.

References

Adamson, D., Ashe, C., Jang, H., Yaron, D., & Rosé, C. P. (2014). Intensification of group knowledge exchange with academically productive talk agents. In *Proceedings of the 10th International Conference on Computer Supported Collaborative Learning* (pp. 10-17).

Andrews, R. (1994). Democracy and the Teaching of Argument. *The English Journal*, *83*(6), 62. https://doi.org/10.2307/820268

Andriessen, J., Baker, M., & Suthers, D. (2003). Argumentation, computer support, and the educational context of confronting cognitions. *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments*, 1-25. https://doi.org/10.1007/978-94-017-0781-7

Baker, M. J. (1999). Argumentation and constructive interaction. *Foundations of argumentative text processing, 5,* 179-202.

Baker, M.J. (2002). Argumentative interactions, discursive operations and learning to model in science. In P. Brna, M. Baker, K. Stenning & A. Tiberghien (Eds.), The Role of Communication in Learning to Model, pp. 303-324. Mahwah New Jersey: Lawrence Erlbaum Associates. https://doi.org/10.4324/9781410606280-20

Baker, M. J., Andriessen, J., & Schwarz, B. B. (2019). Collaborative argumentation-based learning. In N. Mercer, R. Wegerif, & L. Major (Eds.), *The Routledge International Handbook of Research on Dialogic Education* (1st ed., pp. 76–88). Routledge. https://doi.org/10.4324/9780429441677-8

Barbieri, M. S., & Light, P. H. (1992). Interaction, gender, and performance on a computer-based problem solving task. *Learning and instruction, 2*(3), 199-213.

Barnes, D. (2008). Exploratory talk for learning. *Exploring talk in school*, 1-15.

Barnes, D., & Todd, F. (1977). Communication and learning in small groups. Routledge & Kegan Paul.

Barnes, D., & Todd, F. (1995). Communication and learning revisited: Making meaning through talk. Portsmouth, NH: Boynton. *Cook Publishers. Brooks, F.(1993). Some problems and caveats in communicative discourse: Toward a conceptualization of the foreign language classroom. Foreign Language Annals*, *26*, 233242.

Barron, B. (2003). When Smart Groups Fail. *Journal of the Learning Sciences*, *12*(3), 307–359. https://doi.org/10.1207/S15327809JLS1203_1

Baumer, E. P. S., Blythe, M., & Tanenbaum, T. J. (2020). Evaluating Design Fiction: The Right Tool for the Job. *Proceedings of the 2020 ACM Designing Interactive Systems Conference*, 1901–1913. https://doi.org/10.1145/3357236.3395464

Binkley, R. (1995). Argumentation, Education and Reasoning. *Informal Logic, 17*(2). https://doi.org/10.22329/il.v17i2.2403

Brown, M. (2002). The facilitator as gatekeeper: A critical analysis of social order in facilitation sessions. *Journal of Adventure Education and Outdoor Learning*, *2*(2), 101–112. https://doi.org/10.1080/14729670285200211

Cheon, E., Sher, S. T.-H., Sabanović, Š., & Su, N. M. (2019). I Beg to Differ: Soft Conflicts in Collaborative Design Using Design Fictions. *Proceedings of the 2019 on Designing Interactive Systems Conference*, 201–214. https://doi.org/10.1145/3322276.3322350

Clifton, J. (2006). Facilitator talk. *ELT journal, 60*(2), 142-150.

Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of educational research*, *64*(1), 1-35.

Cooper, R. G., & Edgett, S. (2008). Ideation for product innovation. *What are the best methods*, 12-17

Cropley, D. H. (2015). Promoting creativity and innovation in engineering education. *Psychology of Aesthetics, Creativity, and the Arts*, *9*(2), 161. https://doi.org/10.1037/aca0000008

Diehl, M., & Stroebe, W. (1987). Productivity loss in brainstorming groups: Toward the solution of a riddle. *Journal of personality and social psychology*, *53*(3), 497

Dillenbourg P. (1999). What do you mean by collaborative leraning?. In P. Dillenbourg (Ed) Collaborative-learning: Cognitive and Computational Approaches. (pp.1-19). Oxford: Elsevier

Doise, W., & Mugny, G. (1984). The social development of the intellect. Oxford, England: Pergamon Press.

Dunne, A., & Raby, F. (2013). *Speculative everything: design, fiction, and social dreaming*. MIT press.

Fisher, E. (1993). Distinctive features of pupil-pupil classroom talk and their relationship to learning: How discursive exploration might be encouraged. *Language and Education*, *7*(4), 239-257. https://doi.org/10.1080/09500789309541363

Foster, J., & Yaoyuneyong, G. (2016). Teaching innovation: Equipping students to overcome real-world challenges. *Higher Education Pedagogies*, 1(1), 42–56. https://doi.org/10.1080/23752696.2015.1134195

Gaskins, I. W., Satlow, E., Hyson, D., Ostertag, J., & Six, L. (1994). Classroom talk about text: Learning in science class. *Journal of Reading*, *37*(7), 558-565.

Gerlach, J. M. (1994). Is this collaboration? *New Directions for Teaching and Learning*, *1994*(59), 5–14. https://doi.org/10.1002/tl.37219945903

Gibbons, P. (2002). Scaffolding language, scaffolding learning. Portsmouth, NH: Heinemann.

Hmelo-Silver, C. E., & Barrows, H. S. (2008). Facilitating Collaborative Knowledge Building. *Cognition and Instruction*, *26*(1), 48–94. https://doi.org/10.1080/07370000701798495

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. Qualitative Health Research, 15(9), 1277–1288. https://doi.org/10.1177/1049732305276687

Isaksen, S. G. (1998). *A Review of Brainstorming Research: Six Critical Issues for Inquiry*. 29.

Jones, J. F. (1999). From Silence to Talk: Cross-Cultural Ideas on Students Participation in Academic Group Discussion. *English for Specific Purposes*, *18*(3), 243–259. https://doi.org/10.1016/S0889-4906(97)00059-8

Kirschner, F., Paas, F., & Kirschner, P. A. (2009). A Cognitive Load Approach to Collaborative Learning: United Brains for Complex Tasks. *Educational Psychology Review*, *21*(1), 31–42. https://doi.org/10.1007/s10648-008-9095-2

Kramer, T. J., Fleming, G. P., & Mannis, S. M. (2001). Improving Face-To-Face Brainstorming Through Modeling and Facilitation. *Small Group Research, 32*(5), 533–557. https://doi.org/10.1177/104649640103200502

Kuhn, D., Goh, W., Iordanou, K., & Shaenfield, D. (2008). Arguing on the computer: A microgenetic study of developing argument skills in a computer-supported environment. *Child development*, *79*(5), 1310-1328. https://doi.org/10.1111/j.1467-8624.2008.01190.x

Kumar, R., Rosé, C.P. (2011). Architecture for Building Conversational Agents that Support Collaborative Learning. *IEEE Transactions on Learning Technologies 4*(1), 21-34. https://doi.org/10.1109/tlt.2010.41

Lambirth, A. (2006). Challenging the laws of talk: Ground rules, social reproduction and the curriculum. *The Curriculum Journal*, *17*(1), 59–71. https://doi.org/10.1080/09585170600682608

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, 159-174

Lave, J., & Wenger, E. (1991). Learning in doing: Social, cognitive, and computational perspectives. *Situated learning: Legitimate peripheral participation*, *10*, 109-155.

Lavonen, J., Autio, O., & Meisalo, V. (2004). Creative and Collaborative Problem Solving in Technology Education: A Case Study in Primary School Teacher Education. *The Journal of Technology Studies*, *30*(2). https://doi.org/10.21061/jots.v30i2.a.8

Leitão, S. (2000). The potential of argument in knowledge building. *Human development*, *43*(6), 332-360.

Light, P., & Glachan, M. (1985). Facilitation of individual problem solving through peer interaction. *Educational psychology, 5*(3-4), 217-225.

Liu, A., & Lu, S. C.-Y. (2013). Impacts of Synthesis Reasoning on Ideation Effectiveness in Conceptual Design. *Journal of Mechanical Design*, *135*(6), 061009. https://doi.org/10.1115/1.4024086

Lehrer, R., & Schauble, L. (2006). *Cultivating model-based reasoning in science education*. Cambridge University Press.

Lovelace, K., Shapiro, D. L., & Weingart, L. R. (2001). Maximizing Cross-Functional New Product Teams' Innovativeness and Constraint Adherence: A Conflict Communications Perspective. *Academy of Management Journal*, 44(4), 779–793. doi:10.5465/3069415

López-Forniés, I., Sierra-Pérez, J., Boschmonart-Rives, J., & Gabarrell, X. (2017). Metric for measuring the effectiveness of an eco-ideation process. *Journal of Cleaner Production*, *162*, 865-874. https://doi.org/10.1016/j.jclepro.2017.06.138

Macgregor, J. (1990). Collaborative learning: Shared inquiry as a process of reform. *New Directions for Teaching and Learning*, *1990*(42), 19–30. https://doi.org/10.1002/tl.37219904204

Maloney, J., & Simon, S. (2006). Mapping Children's Discussions of Evidence in Science to Assess Collaboration and Argumentation. *International Journal of Science Education*, *28*(15), 1817–1841. https://doi.org/10.1080/09500690600855419

Maxwell, D., Pillatt, T., Edwards, L., & Newman, R. (2019). Applying Design Fiction in Primary Schools to Explore Environmental Challenges. *The Design Journal*, *22*(sup1), 1481–1497. https://doi.org/10.1080/14606925.2019.1594972

Messer, D. J., Joiner, R., Loveridge, N., Light, P., & Littleton, K. (1993). Influences on the effectiveness of peer interaction: Children's level of cognitive development and the relative ability of partners. *Social Development*, *2*(3), 279-294.

Mercer, N. (1994). The quality of talk in children's joint activity at the computer. *Journal of Computer Assisted Learning*, *10*(1), 24–32. https://doi.org/10.1111/j.1365-2729.1994.tb00279.x

Mercer, N. (1996). The quality of talk in children's collaborative activity in the classroom. *Learning and instruction, 6*(4), 359-377.

Mercer, N., Wegerif, R., & Dawes, L. (1999). Children's Talk and the Development of Reasoning in the Classroom. *British Educational Research Journal*, *25*(1), 95–111. https://doi.org/10.1080/0141192990250107

Mevarech, Z. R., & Light, P. H. (1992). Peer-based interaction at the computer: Looking backward, looking forward. *Learning and Instruction, 2*(3), 275–280. https://doi.org/10.1016/0959-4752(92)90013-C

Mohan, M. D., Sarfraz, I., Hewege, C., & Rajendran, D. (2018). An exploration of global employability skills: A systematic research review. *International Journal of Work Organisation and Emotion*, 9(1), 63. https://doi.org/10.1504/IJWOE.2018.10012435

Moshman, D., & Geil, M. (1998). Collaborative reasoning: Evidence for collective rationality. *Thinking & Reasoning, 4*(3), 231-248.

Mueller, M., Yankelewitz, D., & Maher, C. (2012). A framework for analyzing the collaborative construction of arguments and its interplay with agency. *Educational Studies in Mathematics, 80*(3), 369–387. https://doi.org/10.1007/s10649-011-9354-x

Nielsen, J. A. (2013). Dialectical Features of Students' Argumentation: A Critical Review of Argumentation Studies in Science Education. *Research in Science Education*, *43*(1), 371–393. https://doi.org/10.1007/s11165-011-9266-x

Nijstad, B. A., Stroebe, W., & Lodewijkx, H. F. (2003). Production blocking and idea generation: Does blocking interfere with cognitive processes?. *Journal of experimental social psychology*, *39*(6), 531-548. https://doi.org/10.1016/s0022-1031(03)00040-4

Offner, A. K., Kramer, T. J., and J. P. Winter. 1996. The effects of facilitation, recording, and pauses on group brainstorming. Small Group Research 27 (2): 283-98.

Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of Research in Science Teaching*, *41*(10), 994–1020. https://doi.org/10.1002/tea.20035

O'Neill, W. (1970). Properly Literate. *Harvard Educ Rev*, *40*(2), 260-3.

Piaget, J. (1971). The theory of stages in cognitive development. In D. R. Green, M. P. Ford, & G. B. Flamer, *Measurement and Piaget*. McGraw-Hill

Peterson, E. R., & Barron, K. A. (2007). How to Get Focus Groups Talking: New Ideas that will Stick. *International Journal of Qualitative Methods*, *6*(3), 140–144. https://doi.org/10.1177/160940690700600303

Polo, C., Lund, K., Plantin, C., & Niccolai, G. (2015). *Analyzing Exploratory Talk as a Socio-Cognitive Practice: Identity, Group Argumentation, and Class Debate Quality*. 36.

Rojas-Drummond, S., Mazón, N., Fernández, M., & Wegerif, R. (2006). Explicit reasoning, creativity and co-construction in primary school children's collaborative activities. *Thinking Skills and Creativity*, *1*(2), 84–94. https://doi.org/10.1016/j.tsc.2006.06.001

Roschelle, J., & Teasley, S. D. (1995). The Construction of Shared Knowledge in Collaborative Problem Solving. In C. O'Malley (Ed.), *Computer Supported Collaborative Learning* (pp. 69–97). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-85098-1_5

Royle, K., Jenkins, C., Nickless, J., Arrigo, M., & Ciprì, G. (2010). *Teaching kids how to hold productive learning conversations using pictochat on the Nintendo DS*. *6*(1), 22.

Runco, M. 2010. Divergent thinking, creativity and ideation. In J. Kaufman & R. Sternberg (eds), Cambridge Handbook of Creativity, Cambridge UK: Cambridge University Press: 413- 446

Ryan, J., & Viete, R. (2009). Respectful interactions: Learning with international students in the English-speaking academy. *Teaching in Higher Education*, *14*(3), 303–314. https://doi.org/10.1080/13562510902898866

Sampson, V., & Clark, D. B. (2008). Assessment of the ways students generate arguments in science education: Current perspectives and recommendations for future directions. *Science education*, *92*(3), 447-472. https://doi.org/10.1002/sce.20276

Sannomiya, M., Yamakawa, I., Kawaguchi, A., & Morita, Y. (2003). Effect of backchannel utterances on facilitating idea-generation in Japanese think-aloud tasks. *Psychological reports*, *93*(1), 41-46. https://doi.org/10.2466/pr0.2003.93.1.41

Sedova, K., Sedlacek, M., Svaricek, R., Majcik, M., Navratilova, J., Drexlerova, A., Kychler, J., & Salamounova, Z. (2019). Do those who talk more learn more? The relationship between student classroom talk and student achievement. *Learning and Instruction*, *63*, 101217. https://doi.org/10.1016/j.learninstruc.2019.101217

Stegmann, K., Weinberger, A., & Fischer, F. (2007). Facilitating argumentative knowledge construction with computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning*, *2*(4), 421–447. https://doi.org/10.1007/s11412-007-9028-y

Strohmann, T., Siemon, D., & Robra-Bissantz, S. (2017). brAInstorm: Intelligent Assistance in Group Idea Generation. In A. Maedche, J. vom Brocke, & A. Hevner (Eds.), Designing the Digital Transformation (Vol. 10243, pp. 457–461). Springer International Publishing. https://doi.org/10.1007/978-3-319-59144-5_31

Tanenbaum, T. J. (2014). Design fictional interactions: why HCI should care about stories. *interactions*, *21*(5), 22-23.

Tancig, S. (2009). Expert team decision-making and problem solving: Development and learning. *Interdisciplinary Description of Complex Systems: INDECS*, 7(2), 106-116.

TENK. (2019). The ethical principles of research with human participants and ethical review in the human sciences in Finland (Issue 3)

Todd, F., & Todd, R. (1979). Talking and Learning: Towards the Effective Structuring of Student-Directed Groups in Higher Education. *Journal of Further and Higher Education*, 3(2), 52–66. https://doi.org/10.1080/0309877790030206

Toulmin, S., Rieke, R. D., & Janik, A. (1979). An introduction to reasoning. New York: MacMillan

van Eemeren, F. H., Grootendorst, R., & Henkemans, F. S. (1996). Fundamentals of Argumentation Theory a Handbook of Historical Backgrounds and Contemporary Developments (Vol. 31, Issue 1).

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press

Walton, D. N. (1990). What is reasoning? What is an argument?. *The journal of Philosophy, 87*(8), 399-419.

Walton, D. (2006). Rules for reasoning from knowledge and lack of knowledge. *Philosophia, 34*, 355-376.

Wegerif, R. (2005). Reason and Creativity in Classroom Dialogues. *Language and Education*, *19*(3), 223–237. https://doi.org/10.1080/09500780508668676

Wegerif, R., & Mercer, N. (1997). A dialogical framework for researching peer talk. *Language and Education Library*, *12*, 49-64.

Wells, C. G. (1999). Dialogic inquiry (pp. 137-41). Cambridge: Cambridge University Press.

Appendix A. TASK FOR BOTH GROUPS

It is the year 2042…

By this time, AI has improved a lot

- It recognizes images with much higher accuracy
- It identifies basic human emotions without mistakes
- It can order things/food online within the set budget
- it can do many other things…

We are in the country called Sunland.

Higher education students have submitted a complaint to authorities that their studies have become too demanding and exhausting.

At the same time, university teachers resentful, they claim they are overloaded, their work life

So, Sunland's Government issues funding

for development of

AI-based learning assistant (for higher degree students) or **AI-based teaching assistant** (for university teachers).

You are a group of researchers in learning sciences, and you wish to get that funding. How do you get that funding? Present your prototype to take part in the bid!

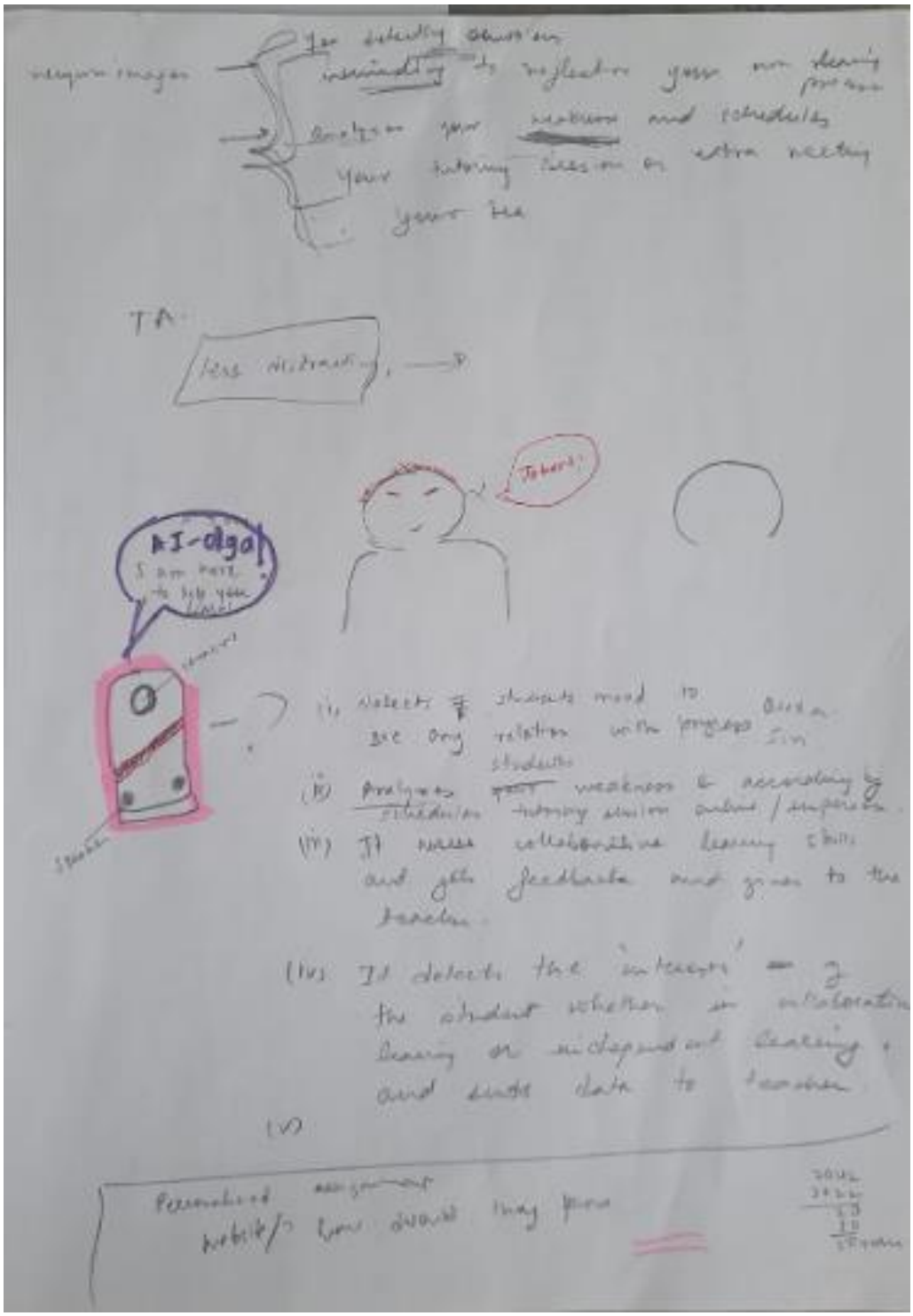
Your bid should contain the answers to:

- Is your product LA or TA?
- What is its name?
- How does it look like? (if possible, a drawing)
- What are its functions?

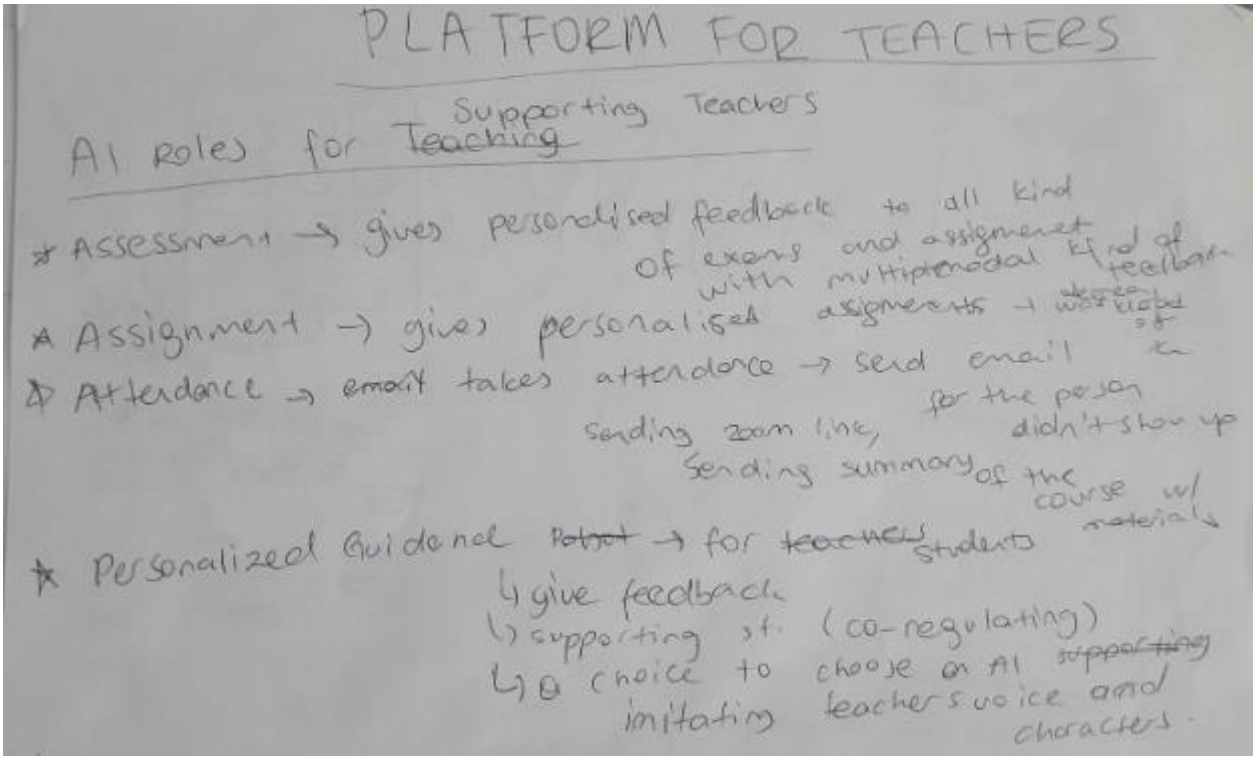
(one person from the group presents)

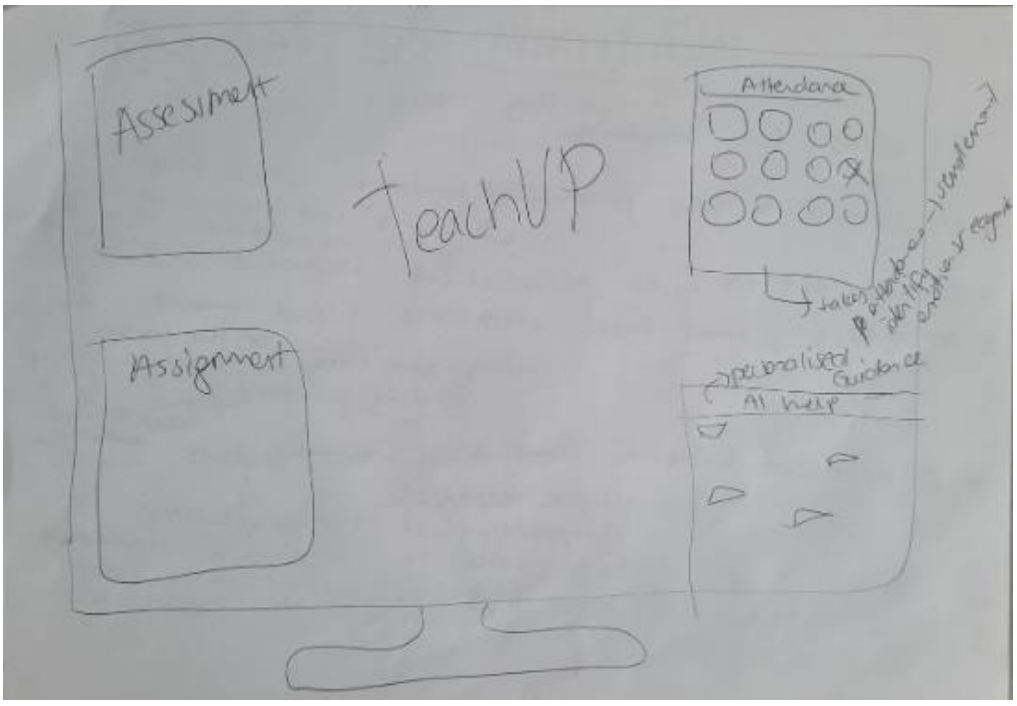
Appendix B. COPIES OF GROUP NOTES

Group A notes



Group B notes





Appendix C. ANALYSIS OF GROUPS STATEMENTS

GROUP A group statement

(Anna is presenting, her own ideas are shown in bold)

Speaker	Utterance within group statement	Idea ID
Anna	Our product is a teacher assistant, AI-based teaching assistant. And the name of our product is.. it says "hi, my **name is Olga**" and I am here to help you	A-19 and A-16
Anna	What does it look like? It looks like this (shows the sketch made during the discussion).	
Anna	**It is small, not that big**. We wanted to give a vibe like there is no, like, another person. Ok?	A-7
Veronica	It looks more like a device	
Anna	Yeah, it's actually like a device.	
Anna	There is one camera to see, **to detect your emotions** and stuff,	A-2
Anna	and **two speakers**,	was not discussed
Anna	so that if you have any questions to ask – the assistant answers	A-8
Anna	And if AI teaching assistant wants to send some information to you, it is linked with the software that you have on your screen.	
Anna	Ok, how does it look like we have showed, and what are its functions – we have chosen three functions	
Anna	It **detects students' mood** to see if there is any relationship between the child **makes progress and it looked happy**, some recent AI can do that.	A-2
Anna	And then the second – **it analyzes the student's weaknesses based on the assessments and then it sees if there is any pattern, like if the kid is making the same mistake over and again and again, it recommends specific tutoring sessions either independent learning or you can schedule a meeting with the teacher if you would like**	A-4
Anna	Then 3rd thing – it assesses collaborative learning skills and it gives.. like if a teacher has given a task where there is collaboration, it will assess how the collaboration was like, what skills different participants showed and **sends feedback to the teacher,** so that the teacher can help which students need.. No, that will help the teacher to make a decision which participants are lacking collaboration skills, because it is assumed that by 2042 it will be something.. that collaborative skills will be needed	A-10 A-18
Anna	And lastly, our AI Olga can also **detect the interests of student based on the topics and with whom, which groups is it collaborating.** So, maybe in this group I am not collaborating that enough, and based on that it detects my interest, and if I go with someone else, I am collaborating there effectively, so it *somehow analyzes your probably your personality traits*. Yeah, that's about it	A-17
Nadine	You mean a device for one student or the whole class?	
Anna	Very good question! Well **we can't afford an individual device for**	

	everyone, *so we are looking probably at 1 for 5, no 1 for 4*. So for 4 kids will be one teacher assistant.	
Toney	Knowing that it is 2042, our perspective towards AI is going to change due to ethics. So for example if someone doesn't want to be on any kind of device, how would you do that task?	
Felicia	So, that's why we are following GDPR guidelines. It depends on what guidelines will be provided in 2042. So if someone doesn't want a device, they can go with a human teacher assistant.	

GROUP B group statement

(Toney is presenting, her own ideas are shown in bold)

Speaker	Utterance within group statement	Idea ID
Toney	So, our AI-based is for teachers. It is a platform, like a website, like LMS, but it does more than we have now	
Toney	What does it do? First, we meant our product for teachers. So we first looked at possible problems for the teachers. We tried to understand what kind of work may have in 2042. So we thought like there is a **lot of assignments, assessments** and personal guidance and a lot of things such as **administration stuff**	B-15
Toney	So, first we have here (shows at the sketch) on the assessment part, this is personalized feedback to students to all kinds of assignments. It could be an essay, now giving feedback to essays might be too problematic for AI, but in 2042 it is not going to be like that, it is going to be super good.	B-2
Toney	And it is going to be not only with the multiple-choice exam grading, it will do more, **it'll rate assessments and assignments**. And, oh sorry, for the assignments it will give personalized assignments and we hope it will be also helping with the students, because the problem why the students right now have a problem it could be because the work they are being assigned it could be beyond their confidence. So it will give personalized assignments so that the kids *can first progress on their own terms*. So, here's the assignments part	B-6
Toney	Oh, by the way, the name is TeachUp. It is not that creative, just TeachUp	B-16
Toney	**And it will take attendance**. Ehm, normally it takes about 2 minutes for a teacher, but it will give us 2 minutes for personalized interactions.	B-7
Toney	It will also.. when AI is taking attendance, it will also recognize emotions of the students, it can **notify the teacher this person is feeling a little bit down** or sleepy and maybe give some recommendations.	B-4, B-3
Toney	And if somebody is not there, not in the class, **it automatically sends an email** to the person, just "*Hey, you were not in class, how is it going, like, why weren't you in the class?"*	B-8

Toney	It will also **send a summary of the class, all the materials**, so that everybody can catch up and they don't have to spend time on things like "what did they do?", like they will gain more time	B-9
Toney	And here is super good, I wish we had that, **it will be like personalized guidance, student will go there and write their problems**, if they need more information, AI will provide	
Toney	And if there is anything related to emotional, where AI cannot help, it will notify the teacher, so that the students get the chance to talk to teacher	B-13
Toney	And also, now that I'm talking it came to my mind, teachers also need to talk to parents, but in higher education maybe it's not the case, but for lower secondary school teachers can be advising with parents on that as well.	Not discussed
Toney	We have a choice for AI help, so certain students want to talk to the teacher, not talk to AI, **but with the teacher's voice and character**, it can also do that, because it's like up to your choice, we wanted to add that. This is not only for texting, **but voice messaging** and you can also send video in 2042. TeachUp guys!	B-5
Anna	I just have a doubt. It's a website, yes?	
Toney	It's LMS, platform	B-18
Anna	So, it's not a device, like how we.. How would it know that the kid is not present in the class?	
Veronica	Because it is equipped with the webcam	Not discussed
Toney	Yeah, and we think everybody is going to be in LMS, there not going to be any notebooks and pencils existing in that year, so probably all the classes are going to be done through computers.	B-18
Anna	Ok, because I think like, if everyone is in person, then how is the website going to know. Ok, thank you	