BENIGN MANIPULATION IN EDUCATION

by

Magen Inon

A thesis submitted in partial fulfilment

of the requirements for the degree of

Doctor of Philosophy

at the

University College London

2022

2

Declaration

I, Magen Inon, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Benign manipulation is an integral part of educational relationships and processes, yet it has received little scholarly attention to date. The goal of this thesis is to provide educators with a conceptual framework by which to assess instances and practices of benign manipulation in education. To do so, I provide a theoretical analysis of the meaning and significance of benign manipulation in education, as well as multi-faceted criteria by which to assess manipulation. I maintain that although there is no consensus regarding the nature of manipulation, for the practical purpose of assessing manipulation, it would be reasonable to consider as manipulation any attempt to influence others which involves deception, trickery, subversion of rationality, pressure or a failure to track reasons. If the motivation for such influence is to advance the interests of the manipulee, I term it benign manipulation. Assuming that manipulation, at least in the case of benign manipulation, should be understood as a non-moralised concept, I establish a framework for the assessment of benign manipulation in education. The framework consists of three criteria: A) the consequences of the manipulation, B) the effect the manipulation has on both the manipulator's and the manipulee's understanding of the situation, and C) whether the manipulation breaches the trust that exists between the manipulator and the manipulee. To demonstrate the practicality of the framework, in the second part of

4

the thesis I apply it to three case studies, all of which have educational implications: benign deception, pharmacological cognitive enhancement and gamification.

5

Impact Statement

The analysis and insights presented in the thesis below could be put to beneficial use in various ways. Firstly, as the thesis aims to provide educators with a framework to assess the use of manipulation in education, it may be used to improve educators' awareness and understanding of the desirability and possible effects of many routine practices used in schools and in other educational contexts. Provided with the conceptual analysis presented in the thesis, educators would be able to have a greater understanding of their role and commitments as educators, as well as the possible efficacy of various educational techniques. To use an analogy, the analysis aims to provide educators with a map of the possible advantages and pitfalls of many various means used in educational contexts. To increase the impact of the thesis, I aim to share the analysis and insights with as many educators as possible. Secondly, the thesis highlights the prevalent use of manipulation in education, and aims to draw scholarly attention to these phenomena. Future research will be able to use the conceptual framework presented here and apply it to various other uses of benign manipulation in education. In this, the thesis makes a valuable contribution to the discipline of Philosophy of Education, as well as to other areas of educational research. Thirdly, some suggested uses of benign manipulation, such as the use of pharmacological cognitive enhancement, raise

6

concerns for many outside academia as well as within it, and the present work contributes to the ongoing debates about the regulation of such techniques. Specifically, a version of Chapter Six which deals with pharmacological cognitive enhancement, has been published in an academic journal (*Ethics and Education*) and drew interest from scholars as well as journalists and the general public. I intend to publish other parts of the thesis as well. Lastly, and perhaps most significantly, my interest in the phenomenon of benign manipulation in education started from a sense of personal confusion and lack of understanding. Through working towards the completion of the thesis, I have gained greater clarity and managed to shed some light on an issue of personal interest. I have already noticed the benefits of this understanding in my daily life, most significantly as a teacher and a father. I strongly believe that others may benefit from greater understanding of benign manipulation, and aim to share it with many people who routinely use manipulation to advance the interests of others.

Table of Contents

Chapter 1 Introduction

1.1 Why does benign manipulation merit attention?

A short while after I started my first teaching job, I noticed that I was using quite a few dubious techniques in order to encourage my students to do what I wanted. For example, I would motivate a student by saying 'I am sure you can do this', even though I was not sure they were capable of doing what I was asking them to do. Or, I would devise a game or some kind of competition in order to have my students focus on some curricular content. Sometimes, I would act as if the coming exam was significantly more important than I thought it was. Other times, I would teach them something that I knew was inaccurate, false or true only in specific cases, like: 'in a triangle, the three interior angles always add up to 180°'. After one of my first days of teaching, I distinctly remember thinking to myself that I had never lied so much on any other day in my life. These might seem minor incidents, almost insignificant, but saying or doing something less than what I considered honest, genuine and true many times a day, with many different students, got me to question myself: am I teaching my students by manipulating and deceiving them? Is this OK? Personally, I felt that I was behaving more like a sleek salesman than an educator.

11

Though these worries were the catalyst to start my research into the use of manipulation in education, I am confident that the issue merits attention far beyond my own personal curiosity. Often when I mentioned the topic of my thesis in a conversation, I would hear a reply along the lines of: 'well, isn't all education a form of manipulation?'. Indeed, it appears that the use of manipulation is very common in healthy and caring educational relationships. For example, a parent might hide candy from her child to avoid him eating too much sugar, or a teacher would set up a game in order to motivate children to learn. The question arises though, is every well-intended manipulation in education appropriate? In order to answer this question, we require an understanding of what benign manipulation is, and what are the relevant considerations when we decide whether it is an appropriate course of action. A significant part of this work aims to answer these questions.

A further reason to focus on the topic is the growing body of scholarship in analytical philosophy on manipulation. This partly stems from a reaction to a wave of behavioural science, such as the work of Richard Thaler and Cass Sunstein documented in their popular book *Nudge*. In brief, Thaler and Sunstein claimed that predictable cognitive processes can be harnessed by the state and other benign actors to better the lives of its citizens. For example, in order to promote a smarter saving strategy, savers are automatically enrolled into a saving scheme, and can opt

12

out if they wish to. This is based on research that shows people are more likely to stay with the default option (Sunstein and Thaler 2008). In reaction to these policy recommendations, philosophers have responded in various ways. Some of the normative issues raised to challenge the use of nudges include: lack of transparency, problems involved in the harnessing of bad reasoning, and manipulation (Bovens 2009; Dworkin 2019). While the work of Thaler and Sunstein on nudges was the catalyst for renewed interest in the concept of manipulation, the philosophical discussion went beyond 'nudges', and involved an in-depth analysis of the concept and the normative issues involved in manipulation. To the best of my knowledge, this body of work has not been used to inform a systematic analysis of educational processes and relationships. As the concept of manipulation has not received much attention from philosophers of education, I believe it would be a valuable contribution to the field.

Similar to the ways in which behavioural science research conducted in disciplines such as psychology and economics has been used to inform policies and practices, in the educational arena significant efforts were made in last few decades in order to discover 'what works'. I believe this is a characteristic of a larger aspect of modern education - that of thinking of education as a planned instrumental project with concrete and well-defined aims, specifically performance-oriented aims; a

13

project that is administered from the top down because someone 'wants to see results'. This view is not always bad or undesirable. However, this simplistic view of education risks promoting techniques which are effective, but might be inappropriate for other reasons.

One notable example is John Hattie's book *Visible Learning*, in which he produced a meta-analysis of educational research that rates the effect size different interventions in education have on students' achievement (Hattie 2009). There is no mention in his work, however, of reasons to be critical towards some of these methods, even when they involve deceiving students, for example. Some of the techniques mentioned by Hattie and others can usefully be analysed as manipulation. Therefore, the present work uses insights from philosophical scholarship on manipulation and applies them to the educational arena, focusing on some suggested effective techniques. Specifically, I look into false encouragement and the use of so called 'smart drugs' and gamification, as instances of manipulation in education. In doing that, the present work offers a productive bridge over the gap between analytic philosophy and educational research.

As I find manipulative techniques to be part of many educational relationships and processes, it should be noted that I use the term "education" in the broadest

14

sense, which includes informal forms of education, pedagogy and upbringing, as well as formal schooling.

1.2 Benign Manipulation

Here are a few examples of the use of benign manipulation in education. The first comes from a parents' website on the internet, where a parent describes an act that they themselves considered to be manipulation:

> **Long way home**: My in-laws live a mile away. When my son was an infant/toddler we would "take the long way home". He would fall asleep and we would go home. Sometimes it would be a good 30/40 minutes before he would be out. When he was 2 1/2 years old he asks us while on a ride "Why is it so quick to get to Nana and Gramp's house but it takes soooooooooooooooooooo loooooooooooooooooong to get home?" (Pinola 2015)

In this example, the parents influenced the behaviour of their child by taking a longer path home then they would have otherwise. They are not influencing their child by strictly coercing him, nor are they reasoning with him. The nature of sleep as a mental state is such that it is very difficult to coerce someone into sleep (excluding

15

drugs, etc.), and it would not make much sense to reason with the child until he understands that he should go to sleep. In addition, it appears that the motivation for their action is the well-being of their child, and that the parents have good reasons to believe that going to sleep would be good for their child. Hence, there does not seem to be anything objectionable about their use of 'taking the long way home'. That said, I still think there is a good reason to describe their action as manipulation, as is also apparent in the child's challenging reaction.

The child is manipulated into falling asleep, by a method that might be described as 'environmental manipulation', in which the manipulator does not target beliefs or emotions but sets up the environment to bring about the desired behaviour, in this case – sleep. Even though the manipulation here is clearly done in the interests of the child, it is insightful to notice that the young boy learns at a certain point that something odd is going on, and in his own way he questions his parents about it. It seems that what the boy is trying to get at is that he does not understand the reasons for his parents' action of taking the long way home.

While some acts of manipulation in education are intuitively unproblematic, others are much more questionable. A famous example can be found in Plato's *Republic*. In the dialogue, Socrates proposes to tell the future citizens of the Republic a myth that is supposed to make them more inclined to care for one

16

another. He refers to this myth as "…one of those lies that come into being in case of need, of which we were just now talking, some noble lie to persuade...". The noble lie Socrates proposes is the following:

> **Noble lie:** When the job had been completely finished, then the earth, which is their mother, sent them up. And now, as though the land they are in were a mother and nurse, they must plan for and defend it, if anyone attacks, and they must think of the other citizens as brothers and born of the earth. (Plato et al. 2016, pp. 309-11)

Socrates recommends deceiving the youth of the Republic about their origins by telling them a myth - they were all born from the ground and therefore, as brothers and sisters, they are to protect the state and each other. In this case Plato recommends deceiving the youth of the state for the greater good of safeguarding the country and its citizens. The myth is a case of manipulating individuals to make them cooperate better with one another. If we assume Socrates has good reasons to believe this myth would promote the well-being and interests of the citizens (a questionable assumption indeed), then it is an example of benign manipulation.

Other famous examples of manipulations in education can be found in Rousseau's *Emile, or on Education*. The book is a treaty on education, in which Rousseau employs two main characters; the child Emile and his tutor Jean-Jacques.

17

At times, the tutor uses techniques which involve deception and trickery in order to bring about the desired behaviour or understanding on part of Emile. One such example occurs when Emile questions Jean-Jacques on the value of learning astronomy, and the next day Jean-Jacques orchestrates them getting lost in order for Emile to understand the value of astronomical knowledge for navigation:

> **Lost**: The next morning, I suggest to him a walk before lunch. He does not ask for better. Children are always ready to take a run, and this one has good legs. We go up to the forest; we roam the fields; we get lost; we no longer know where we are; and when we have to go back, we can no longer find our path again. Time passes; it gets hot; we are hungry. We hurry; we wander in vain in one direction and another. We find everywhere only woods, quarries, plains, and no sign by which to locate ourselves. Very hot, very tired, and very hungry, we accomplish nothing by our racing around other than to get more lost. Finally, we sit down to rest and deliberate. Emile, who I am supposing has been raised like any child, does not deliberate; he cries. He does not know we are at the gate of Montmorency and that a simple copse hides it from us. But this copse is a forest for him; a man of his stature is buried in bushes.

18

After some moments of silence, I say to him with a worried air, "My dear Emile, what shall we do to get out of here?"

EMILE (all in a sweat and crying hot tears) I don't know. I'm tired, I'm hungry, I'm thirsty. I can't go on.

JEAN-JACQUES Do you believe I am in a better condition than you, and do you think I would blame myself for crying if my tears would do for my lunch? Crying isn't what has to be done. What we have to do is find ourselves. Let's see your watch. What time is it?

EMILE It's noon, and I haven't eaten.

JEAN-JACQUES That's true. It is noon, and I haven't eaten.

EMILE Oh, how hungry you must be!

JEAN-JACQUES The misfortune is that my dinner won't come and look for me here. It is noon? That is exactly the time yesterday when we were observing the position of the forest from Montmorency. If we could observe the position of Montmorency from the forest in the same way...."

19

EMILE Yes, but yesterday we saw the forest, and from here we don't see the city.

JEAN-JACQUES That's the difficulty .... If we could find its position without seeing it. ...

EMILE Oh, my good friend!

JEAN-JACQUES Did we not say that the forest was ...

EMILE North of Montmorency ....

JEAN-JACQUES Consequently Montmorency ought to be ...

EMILE South of the forest.

JEAN-JACQUES We have a means of finding the north at noon.

EMILE Yes, by the direction of the shadow.

JEAN-JACQUES But the south?

EMILE What's to be done?

JEAN-JACQUES South is the opposite of north.

20

EMILE That's true. We have only to look for the opposite of the shadow. Oh, there is the south! There is the south! Surely Montmorency is in that direction. Let's look in that direction.

JEAN-JACQUES You might be right. Let's take this path through the woods.

EMILE (clapping his hands and letting out a cry of joy) Oh, I see Montmorency! There it is straight ahead of us in full view. Let's have lunch! Let's dine! Let's run fast! Astronomy is good for something. (Rousseau 1979, pp. 180-81)

In this example, Jean-Jacques uses various means in order to achieve his goal. He sets up the environment to encourage Emile to make certain references, he lies about the fact that he himself is unaware that they are not truly lost, he nudges Emile to think of certain things, and more. In all of those, he is manipulating Emile as part of the educational process.

The elaborate manipulations Plato and Rousseau describe are somewhat far-fetched, but it is not at all difficult to think of daily routines in which educators make use of manipulation. One such example of manipulation by deception is giving students false encouragement:

False Encouragement: A teacher judges her student to be of limited ability in her specific subject, she believes he is not likely to perform well on the coming exam. Nonetheless, when the student asks her if she thinks he'll do well on the exam, she answers: "Yes, I am sure you'll do very well". She does this because she believes that being truthful in this instance will cause the student to put very little effort into his studies.

If we assume that the teacher deceives the student because she believes that it is in the interests of the student to learn more about her subject, then she is acting with good intentions, having the student's interests at heart, and she is also deceiving him regarding her beliefs (she does not believe he will do well on the exam). This is a case of benign manipulation which involves deception. It is worth noting that deceiving someone does not necessarily have to involve lying to them. For example, in False Encouragement, the teacher might truthfully reply to the student's question: "You *did* achieve a high grade for the essay on X" thus highlighting the only time in which the student did achieve a good grade. In this case, she is not lying to the student, but she is encouraging him to make the inference that she truly believes he will perform well on the exam, even though she believes the opposite. She deceives him by telling him the truth.

22

In the examples of manipulation above there is use of some deception or hiding of the truth in order to achieve the manipulator's desired result. However, educators use techniques other than deception in their attempts to manipulate the educated for their own good. Blumenthal-Barby provides a long list of what may be seen as general acts of manipulation:

> …incentivizing, offering, increasing options, decreasing options, tricking, using (resistible) threats of punishments, managing information, presenting information in a way that leads to predictable inferences, deceiving, lying, making a false promise, withholding information or options, slanting information, providing irrelevant inputs or crowding out relevant inputs, exaggerating information in a misleading way, using misleading packaging or misleading images, creating impressions by imagery, using loaded language, trading on fear, subliminal suggestion, insinuations, flattery, guilt, appealing to emotional weaknesses or needs, initiating psychological processes that are difficult to reverse or that lead to predictable behaviours or decisions (e.g., the tendency to continue with an active decision even after becoming aware that it is more costly than originally thought, or the tendency to view an option as more desirable when shown its

23

contrast), browbeating or otherwise wearing the person down, reverse psychology, and seduction (Blumenthal-Barby 2014).

It should be apparent to parents and teachers that many of these techniques are commonly used within what we consider educational relationships and processes. Often, these are not motivated by a concern for the educated, but when they are – it is what I term benign manipulation.

As a more general account, I take benign manipulation to be 1) an intentional influence of one person on another person's beliefs, desires, emotions or behaviour. This influence is 2) not coercive, nor is it persuasion through ideal reasoning. And, 3) the manipulator has good reasons to believe that the action will benefit the manipulee. Chapters Two and Three will explain further what is meant by manipulation as an intentional influence which is not coercive nor ideal reasoning. Here, I wish to make some comments to support my choice of (3) – that benign manipulation is a form of influence in which the manipulator has good reasons to believe the manipulation will benefit the manipulee. This will also explain my choice to use the term 'benign manipulation' as opposed to 'manipulation' alone or other terms that appear in the literature.

The word 'manipulation' has two primary meanings in ordinary language. First, it is more commonly used to describe an interpersonal influence which is

24

morally problematic and objectionable, as in: "The shoe seller manipulated me into buying these ugly shoes". Second, and less regularly, it is used in scientific jargon to describe operations performed on objects or things, such as genes, functions, robots, etc. as in: 'the effects of gene manipulation in rats'. These two meanings of the word manipulation are quite distinct from one another, but both capture important aspects of the phenomenon that might be helpful to understand benign manipulation as a form of influence.

The first meaning of manipulation is value laden, while the second is not. But, which one fits better the process in which one person manipulates another? It appears we need to decide whether manipulation, when we refer to one person intentionally influencing another, should be understood as a moralised concept or not. If it is to be understood as a moralised concept, such as 'murder' for example, that would mean that every instance of manipulation is at least prima facie wrong and requires some form of justification. If it is to be understood as a non-moralised concept, that would mean that manipulation is not prima facie wrong and that it does not necessarily require justification.

The philosophical literature discussing manipulation diverges on the question of whether manipulation should be understood as a moralised concept or not. Although this is in contrast to the common use of the word in ordinary language, I

25

follow scholars such as Sarah Buss (2005), Marcia Baron (2014), Alan Wood (2014) and Moti Gorin (2014b), who claim that manipulation should be understood as a non-moralised term. To illustrate, if a certain action is rightly characterized as an instance of manipulation, understanding manipulation as a non-moralised term would mean that it should not be assumed to be morally suspect just because it is an instance of manipulation. This is not the same as claiming that manipulation is wrong but might be justified consequentially or otherwise. It is claiming that often there is nothing morally objectionable about it to begin with. For example, I believe Marcia Baron expresses an accurate intuition in writing that:

> …benign manipulation is more nearly continuous with perfectly appropriate behavior, even valuable behavior, than is benign deception … in some cases [of manipulation] there is nothing to justify; the manipulation … not only is not wrong but is not even regrettable' (Baron 2014, p. 115)

If this is true, then even though the second, more scientific, meaning of the word 'manipulation' is less commonly used, it might be the more central to our understanding of the concept, even when applied to interpersonal influences. While regrettably distant from the meaning of the word 'manipulation' in everyday language, the analysis below will show that assuming a non-moralised concept of

26

manipulation allows for greater clarity when assessing actual acts, methods and techniques. Indeed, one of the major problems of understanding manipulation as a moralised concept is that the same technique will sometimes be classified as manipulation when it is wrong but would not be classified as manipulation when it is not wrong.

Ultimately, it is a matter of choosing which concept is at the centre of the analysis. Those who wish to remain closer to the most common use of the word 'manipulation' in ordinary language analyse a moralised concept and are required to explain why manipulation is morally suspect. Those who wish to analyse the use of methods of influence which are not coercive nor involve ideal reasoning, but are not necessarily morally suspect, are analysing a different concept. Though there are similarities between the two concepts, they are not the same. Therefore, it might have made more sense to use a neutral term to describe these acts of manipulation, thus avoiding the negative connotations of the word 'manipulation'. This has been the approach of several scholars. For example, Richard Thaler and Cass Sunstein have coined the term 'nudge' to refer to such influences, and J. S. Blumenthal-Barby refers to these kinds of influences as "non-argumentative influences" (Blumenthal-Barby 2014, p. 123). In Psychology, academic centres for 'Behaviour Change' are part of the interest in using different forms of manipulation. And in the field of

27

Economics these types of influence sometimes fall under the category of 'incentives'. All of these terms are lacking for various reasons in capturing the different actions under discussion here. To explain briefly, 'nudge' does not capture many of the paradigmatic examples of manipulative influences, such as those involving deception. The term 'non-argumentative influences' is problematic because arguments can be used in acts of manipulation, 'behaviour change' can be reached by coercion and ideal reasoning, and the term 'incentives' is a bad fit to describe actions that involve lying and deception, which are part of at least some acts of manipulation.

For this reason, I have chosen the term 'benign manipulation' to denote actions which are reasonably believed to promote the interests of the manipulee. In this, the term is closer to the term 'white lies', though there are some differences. 'White lies' usually assumes that lies are prima facie wrong, but there is a justifying reason that makes these lies overall not wrong. 'Benign manipulation' does not assume that manipulation is prima facie wrong and requires justification. Many instances of benign manipulation are not intuitively objectionable on moral grounds. It is often the case that these acts of benign manipulation do not even require justification in the form of 'all things considered it was the right thing to do, but prima

28

facie it is wrong'. Often, benign manipulation clearly promotes the interests of the manipulee, without it being regrettable that these methods are used.

The philosophical literature covering these topics uses two concepts that are relevant to this discussion: paternalism and manipulation. Paternalism I take roughly to mean: "an intervention by one person (the 'intervener') in the activities of another (the 'intervened') that is aimed at achieving significant good for or avoiding significant harm to the intervened" (Mullin 2014). It is important to note that there are paternalistic actions which are not manipulative. These might be coercive, for example holding a child forcefully when he attempts to play with something dangerous, or rational persuasion, such as the case when educators adequately challenge a course of action chosen by the educated. Also, and quite clearly, there are acts of manipulation which are not paternalistic, such as subliminal advertising, or a political spin. The scope of this work is to discuss acts of manipulation which are other regarding, or at least reasonably claim to be so. For this reason, acts of manipulation which are meant to serve the interests of the manipulator at the expense of the interests of the manipulee, are not discussed in detail. For instance, school management who intentionally exclude under-performing students in order to raise a school's ranking in league tables.

29

In Noble Lie, Plato recommends manipulating the future citizens of the Republic because he believes this will benefit them. However, there is a distinction between 'believing that' and 'having good reasons to believe that'. For example, the parents in Long Way Home have good reasons to believe that the child's well-being will be served by him getting more sleep. However, contrast this with the Jehovah's Witness parents of a child denied a blood transfusion who offer him a false story about the process to prevent him from receiving a potentially life-saving transfusion, in the mistaken belief that this will be good for him. One approach to deal with this difficulty would be to differentiate between objective and subjective accounts of well-being. Using such a distinction would make it possible to draw the line between 'believing that' in the subjective account of well-being and 'having good reasons to believe that' in the objective account of well-being. However, this approach is not practical for my aims, as I wish to analyse methods and techniques that are agreed to benefit the manipulee at least in some sense. In other words, my discussion is within the realm of 'having good reasons to believe that' a certain method is beneficial to the manipulee. Therefore, the central question I wish to answer is: in cases where educators have a good reason to believe that the manipulation will promote the interests of the manipulee, is it appropriate to act on that belief? In order to answer this question, I shall analyse in the first part of the present work the concept of manipulation and offer a conceptual framework to assess acts of benign

30

manipulation. In the second part I will assess three different benign manipulations in education: false encouragement, pharmacological cognitive enhancement and gamification.

1.3   Research approach

My research approach uses conventional methods of analytic philosophy, such as conceptual analysis, paradigmatic examples, thought experiments and critical appraisals of intuitions. Mainly, in chapters Two and Three I assess others' definitions of the nature and ethics of manipulation, in order to suggest a framework by which to assess instances of manipulation in education. Furthermore, my analysis relies on my own experiences as an educator, which serves as a source for reflections, examples, intuitions and thought experiments, as well as a venue in which to experiment and test ideas. I find the combination of conceptual analysis alongside the reliance on my own professional practice key to developing true and useful theories in the educational sphere. In this, I echo a long-held belief by those who were nicknamed 'pragmatists'.

Pragmatism is a broad family of philosophical theories that "understands knowing the world as inseparable from agency within it" (Legg and Hookway 2019).

31

Under this heading, there are various theories and interpretations, in many different sub-parts of philosophy. John Dewey, who was one of the main figures in the development of early Pragmatist thinking is especially relevant for educational philosophy as much of his work concentrates on education. Although I would not attempt to articulate a full and complete picture of Dewey's thinking about education, I would like to use some of his work to present key aspects of my research approach. To do that, I will use Dewey's essay 'The Sources of a Science of Education'. Indeed, I find this essay, first published in 1928, so relevant to our current educational challenges, that it strikes me with awe and a deep feeling of frustration at the same time.

To begin, I follow Dewey in assuming that education is an extremely complex process, and this often-overlooked fact is of immense importance for education as a practice. If we assume that educational research is a science, it is claimed that it is the hardest science of all. David C. Berliner provides the following reasons to support this view: 1) context is extremely important in education which makes it hard to generalise findings, 2) the number of interactions in education is significantly larger compared to the physical sciences, and 3) research findings often have a short life span, due to social changes and other factors:

32

Doing science and implementing scientific findings are so difficult in education because humans in schools are embedded in complex and changing networks of social interaction. The participants in those networks have variable power to affect each other from day to day, and the ordinary events of life (a sick child, a messy divorce, a passionate love affair, migraine headaches, hot flashes, a birthday party, alcohol abuse, a new principal, a new child in the classroom, rain that keeps the children from a recess outside the school building) all affect doing science in school settings by limiting the generalizability of educational research findings. Compared to designing bridges and circuits or splitting either atoms or genes, the science to help change schools and classrooms is harder to do because context cannot be controlled. (Berliner 2002, p. 19)

This echoes a wider debate about educational research. For the reasons mentioned above, as well as various other reasons, many theorists claim that educational practice should not be based on research that is modelled on the natural sciences (Biesta 2007; Fendler 2006; Smeyers 2006).

Because educational settings are so complex, it is not clear if and to what extent it is possible to codify and transfer knowledge across contexts. To say the

33

least, the transfer of knowledge should always incorporate a sensitivity to contextual factors. Secondly, because education is always embedded in a specific contextual setting, there is a good reason to believe that educational processes should be understood to have multiple aims and objectives. This is probably true in practice, but in addition, this is likely to be true in principle. Within the broad liberal tradition, the ongoing decades long discussion regarding the aims of the practice which did not result in consensus suggests that it is wise to hold a pluralistic approach regarding the aims of education (Marples 1999). Thirdly, if we accept the assumption that education is a complex process, influencing others also involves a plurality of means. When interpersonal influences are part of an intentional educational process, they might take the form of very different techniques and methods. Even if one has a clear understanding of the aims of the practice, they are still required to choose the appropriate means to achieve those aims. Furthermore, in education means and ends are often inseparable, in the sense that the means chosen have an educational value, and this again is evidence of the complexity of the field (Biesta 2007). Therefore, because education is highly contextual in nature, it is likely that it should be understood as having a plurality of aims and means, which are also bound up with each other. The complexity of education should not discourage us completely in our attempt to understand certain aspects of it, but it does serve as a warning that one approach, or one theoretical framework, is always going to be limited.

34

The question is how to proceed from here, with the understanding that education is extremely complex? Dewey, confronted with the complexity of educational processes, raises the question of how it 'can be conducted with systematic increase of intelligent control and understanding?... From what sources shall we draw so that there shall be steady and cumulative growth of intelligent, communicable insight and power of direction?' (Dewey 1929).

In his answer to this question, Dewey maintains that education as a practice should be informed by scientific findings (i.e. empirical, generalizable findings), but not solely determined by these. There is danger in converting empirical findings to rules and recipes to be uniformly followed by educationalists:

> When, in education, the psychologist or observer and experimentalist in any field reduces his findings to a rule which is to be uniformly adopted, then, only, is there a result which is objectionable and destructive of the free play of education as an art.

> The parent and educator deal with situations that never repeat one another. Exact quantitative determinations are far from meeting the demands of such situations, for they presuppose repetitions and exact uniformities. Exaggeration of their importance tends to cramp judgment, to substitute uniform rules for the free play of thought, and

35

to emphasize the mechanical factors that also exist in schools.

(Dewey 1929)

As we can see, Dewey warns against an exaggeration of the applicability and generalizability of empirical findings. However, this century old warning did not stop many of those working in education from viewing science as a source of authority, rather than as a far less demanding guide to practice. But, the complexity of education does not allow for such a reduction of education to mechanics. Therefore, my research approach assumes that empirical findings should inform educational practice but should not be converted into rules and recipes to be uniformly followed. One of my aims is to discover some of the problems involved in suggesting that certain techniques should be followed as rules and recipes.

Furthermore, Dewey writes that educational theorising should detach itself from short-term usefulness. Science requires abstraction, and that means that certain aspects are removed from practical experience to the realm of theoretical inquiry. It allows us to widen our range of attention beyond the immediate concerns and open up new possible purposes and means to be used. For that, theorizing should be detached from the pressure to demonstrate 'a quick, short-time span of usefulness in school'. As Dewey puts it in 1929: 'There is a tendency to convert the results of statistical inquiries and laboratory experiments into directions and rules for

36

the conduct of school administration and instruction.' (Dewey 1929) When this happens, Dewey warns, 'there is not the leisure for that slow and gradual independent growth of theories that is a necessary condition of the formation of a true science'. Empirical research tends to focus on one relevant factor among many. But in education, we need to have a wider perspective:

> The significance of one factor for educational practice can be determined only as it is balanced with many other factors… No conclusion of scientific research can be converted into an immediate rule of educational art. For there is no educational practice whatever which is not highly complex; that is to say, which does not contain many other conditions and factors than are included in the scientific finding. (Dewey 1929)

For this reason, I aim to theorize in a way that looks beyond short-term usefulness, having a wider perspective that takes account of many factors and conditions.

Moreover, my research, while directed partially at philosophers of education, is directed primarily at educational practitioners, and is meant to inform educational practice. I aim to provide a framework for the understanding and application of scientific results that may assist educators with a useful intellectual instrument in deciding the significance of these findings for their specific situations. This is in line

with Dewey's claim that '…laws and facts, even when they are arrived at in genuinely scientific shape, do not yield rules of practice. Their value for educational practice… consists in provision of intellectual instrumentalities to be used by the educator.' In other words, scientific results should inform educators' judgment, and not be followed blindly. For these reasons, I aim to offer a theory that improves the knowledge and understanding of educators. A theory that is an intellectual instrument.

This approach is aligned with how Dewey views the role of philosophy of education, which is to provide 'working hypotheses of comprehensive application'. In other words, philosophy of education should provide theories that are seen as works in progress that can be tested and modified in practice. Accordingly, my research aims to provide educators with intellectual instruments to inform their practice, but it should also be modified and tested by those practitioners. Because educational practice is both the source of the data for inquiry, and the ultimate final test of the value of the theories, I aim to theorize with some distance from the practice that allows for a wider scope of understanding, but also remain in relative proximity in order to maintain the relevance of the findings. More specifically, I will use case studies and my own experience as an educator to achieve this.

38

Though my research focuses on means used in education, I am in full agreement with Dewey's remarks that in real educational processes means are never fully separate from ends:

> When we make a sharp distinction between what is learned and how we learn it, and assign the determination of the process of learning to psychology and of subject-matter to social science, the inevitable outcome is that the reaction of what is studied and learned upon the development of the person learning, upon the tastes, interests, and habits that control his future mental attitudes and responses is overlooked. In that degree the psychological account of the process of personal learning and growth is deficient and distorted. It then deals with a short segment of the learning process instead of with its continuities. (Dewey 1929)

Hence, although I focus my attention on the means used in education, specifically manipulation, I do that in part to criticize the view that means and ends can be separated from each other. In education, means and ends are bound up with each other.

In sum, the insights I find most appealing in Dewey's approach are the following: first, educational practice can be conducted in a more intelligent and

39

systematic way than it would be by tradition and intuition alone. Second, educators should distrust empirical findings when these are presented as rules or recipes to be followed. Third, educational theory should aim to improve the knowledge and understanding of educators, i.e. theories should be seen as intellectual instruments that inform the judgement of practitioners. Fourth, educational theory should look beyond short-term usefulness and include reflection on the broader context beyond the immediate time and place. Fifth, philosophy of education should deal with concrete educational experience as a source for inquiry and as a space in which to test intellectual investigations. Sixth, in education, means and ends are ultimately inseparable.

The research approach I follow in this work is informed by these insights. Throughout the thesis, I work on the basis that educational practice should be conducted with the understanding that education goes beyond short-term usefulness and a narrow result driven motivation. My work will demonstrate the shortcomings of such a narrow way of thinking about education. Furthermore, I aim to provide educators with a conceptual instrument to better judge suggested rules and recipes that stem from empirical findings.

To do that, I use conventional methods of analytic philosophy as well as my own experience as an educator. First, I look at various conceptual understandings of

40

manipulation, focusing on the nature of manipulation as a non-moralised concept. To better understand manipulation as a non-moralised concept I make use of and assess others' definitions, thought experiments, examples and intuitions of and about manipulation. Second, even when this is not explicitly mentioned, I rely on my own experiences as a teacher and educator in very different settings. My own professional practice serves as a source of examples, thought experiments and intuitions, which I make use of in the conceptual analysis of manipulation. I have had the privilege to work as a secondary teacher in two very different schools. First, in a state school in Israel which caters for underprivileged students, and secondly in an International Baccalaureate school in London which caters for the international elite. I have also worked as a curriculum developer in the Israeli Army, mentored young offenders while they were in prison, worked on a leadership development programme and taught undergraduate students in two universities. Perhaps more importantly, I became a father just over five years ago. Throughout my work, I use these experiences both as sources for reflection, and as a place to experiment with ideas. Lastly, in the second part of the thesis, and in line with Dewey's thinking described above, I look into case-studies that both inform the analysis and serve as a testing ground for it. In this way, I maintain contact with educational practices. Moreover, while acknowledging that means and ends are never fully separate in

education, my focus on the means used in education also ensures that I remain 'in close proximity' to actual educational processes which use these methods.

As part of remaining 'in close proximity' to actual educational processes, I engage with empirical research throughout the thesis. Each of the case studies: false encouragement, pharmacological cognitive enhancement and gamification, is based on empirical research which, so some believe, supports the use of such interventions in education. Indeed, these empirical findings are what makes it reasonable to claim that these interventions are benign. One aim of the thesis is to criticise the interpretation of these empirical results, as calling on educators to employ these interventions with little further consideration. To do so I offer different interpretations of the meaning of those empirical results. However, as part of my criticism, and building on my own experience, I engage in a fair bit of empirical speculation regarding how students, teachers and policy makers might be affected by such manipulation. To make it clear, such empirical speculation is not decisive evidence against the use of such interventions. Rather, it is meant as suggestion of what might be some hidden features of those interventions.

42

1.4 Another map for educators

To further explain what I aim to accomplish in this work I would like to suggest the metaphor of a map. Put simply, I wish to provide educational practitioners with a conceptual map of the nature of benign manipulation in education and a framework to determine the appropriate use of such techniques.

A map is a useful tool for simplifying complex realities in a way that serves human needs. Any map makes assumptions about what is important in the current context and leaves out an infinite amount of detail. The London Tube Map, for example, represents London as a series of colourful lines intersecting each other through dots on a light background. The lines represent the underground train tunnels, and the dots represent the stations. The London Tube Map is purposefully distorted and inaccurate in many ways, meant to make it a simple and useful tool for those who wish to navigate the complex network of the underground train.

In education, a prevalent way of conceptualizing the complexity of the field is thinking about it through the data that is most readily available in the form of measured performances and other numerical information. In the words of David Labaree, to whom I owe thinking about education using the metaphor of the map:

43

In order to make education visible for their purposes, school reformers at the district, state, or national level need to construct a map; and like any map, it necessarily represents its subject in a radically simplified form. It draws on data that are easily gathered, suitable for the task at hand, and amenable to a statistical summary. This means data like student social characteristics and test scores, teacher experience and qualifications, and state and local funding. Only one thing is certain about the map that reformers create in their effort to see schooling: it leaves out almost everything. The complex ecology of the classroom disappears into the simplified columns of summary statistics. (Labaree 2010, p. 158)

I agree with Labaree that the map educational reformers construct has significant limitations. However, I also think that having a map is important for educators. Approaching the complexity of education without some kind of a simplified representation is extremely difficult and would not be productive. That said, educators would be wise to carry more than one map, together with the understanding that any map has limitations. Therefore, what can we expect from a map of education and how should we use it?

44

The Tube Map is a representation of London. It is a very useful representation if one would like to navigate her way through the complicated underground train network of the city. What makes the map so useful, however, is partly the effect of leaving many things out of the map. For one, the map is static and does not change over time, unless a new version is printed. That makes it easier to read the map if one is familiar with it already, but also makes it harder to know which stations and lines are busiest during rush hour. It also leaves out all the roads, landmarks, shops, attractions etc. Again, this 'leaving out' of things makes the map simple and easy to use. It is important to notice that we tend not to think about what is not in the map. The map appears to us as an accurate representation of real life. In this, the map might hinder our understanding of reality.

A similar occurrence has happened in educational settings over the past few decades. Simplifying education to learning, and especially to learning as measured by performance on exams – is a representation that simplifies educational reality extensively. An education system is understood by the way it is represented in statistics, exam results and league tables. What is important is learning, and this is measured mostly by performance on set tasks. As Labaree points out, this way of understanding education 'leaves out almost everything' (Labaree 2010, p. 158). However, what it leaves out is highly important, even for those who regularly use this

45

way of understanding education. By focusing on benign manipulation, I aim to offer

educators with a supplementary, and sometimes competing, map of educational

reality, one that highlights other aspects of significance in the educational reality,

such as educators' commitments to their students, and the wider, long-term,

consequences of their actions.

Any map is a simplified representation of reality, and so also includes some

systematic distortion of geographic space. This is taken to the extreme in the design

of the London Tube Map. Based on the design of electric circuits, it represents all the

Tube lines as straight lines which turn by 45 or 90 degrees only; even the river

Thames follows this rule. The representations of the lines on the map are very

different from how the tunnels travel geographically underground. The map distorts

the geographic location of stations and lines. It also has a varying scale, so that the

centre of the map is drawn with a larger scale than the outer parts of it. In fact, the

Tube Map is very different from many other maps we are used to. It is these

inaccuracies that make the map as useful as it is for commuters, to the extent that

many other underground systems around the world use a similar design. On the

other hand, there are limitations involved in the use of the map; if one would like to

travel between relatively close stations, in some cases it will be faster on foot,

although the map might suggest otherwise.

46

In education, measured performance is likewise a simplification of educational reality. It is inaccurate for at least two reasons: for one, it measures only what is measurable, and - more importantly - easily measurable, and it simplifies these even further – to the realm of numbers and statistics. Furthermore, it treats these measurements as a good indication of what is really going on. This would be akin to thinking that the lines of the underground train follow their representation on the Tube Map. If one uses the Tube Map to navigate one's way through London, the map strikes a good balance between simplicity and accuracy. However, if one wanted to use the Tube Map to plan underground engineering works, it is quite useless. The conceptual map I aim to provide in this research also simplifies reality, and for this reason it should not be seen as providing a recipe for action. What it should provide is a tool for educators to better understand their situation and navigate it. Therefore, the theory offered here, much like a map, requires its user to realize that reality, in all its contextual complexity, is not as it is represented in the theory. Educators should use their own judgement when using the theory, the same way one would use judgement when using a map.

Not only is the Tube Map (purposefully) inaccurate, the representation it offers of the city influences the lived reality of the city. The map becomes a standard way of thinking about the city – so that inhabitants will regularly name the area where they

47

live according to station names and lines, even though geographic location might be more relevant. This might influence communities, house prices, etc. London as an imagined community is visualised and understood through the Tube Map. In a similar form, and perhaps to a large extent, the educational map of exam results, statistics and league tables influences the way people live and behave. It influences university and job allocation, the places people choose to live in, etc. Importantly, it influences the flow of research funds. It also has a significant influence on how people are perceived by others, and how they understand themselves. Within the narrow scope of this work and through the lens of the concept of benign manipulation, I aim to provide educators with a parallel way of thinking about educational reality, that might serve to critically assess the usefulness and accuracy of the prevalent way of representing education.

If we take the map to be reality itself, or even a very accurate representation of reality, we will end up making many mistakes. This is true for users of the Tube, but even more so for those who oversee constructing and running the Tube. The map was designed to be used by commuters, not by engineers of the systems. In education, the model through which we understand the system is often used as the key source of information through which to influence and change it. While Tube engineers have multiple maps, and a good understanding of reality itself, it seems

that educational policy makers and practitioners sometimes lack other perspectives than those shown on the 'standard model'.

In light of the above, perhaps we should aim to educate without having any map, any conceptual scheme of education, whatsoever. However, just as one would feel utterly lost trying to travel the London Underground system without a map, so someone who wishes to navigate the complexities of education requires a conceptual scheme. A map makes it easier to navigate the Tube, but it also gives us a sense of security, and the safety feeling involved in knowing 'the lie of the land'. We are aware that something might go amiss while going from place to place; you might miss your stop, a train might be delayed, cancelled etc. Having a map makes one feel more secure in dealing with the uncertainties and complexities of life. The standard model of education gives a similar sense of security for many of those involved in the education system. For students, it makes clear what success is and what one needs to do to achieve it, it makes clear why some people get certain things and others do not, it gives us a theory that we can use to understand the world around us and feel better for that. It tells us simply if people are improving in what they do, if they are being more or less effective, etc. It is difficult to say whether we would be better off without any map whatsoever. However, we also need to resist the temptation to think that the map is an accurate representation of reality. I am

convinced that a good way to resist this temptation is to offer a different map of educational reality, one which highlights different aspects of that reality.

My work here often remains neutral in regard to the products of my analysis. For example, in the discussion of what is manipulation and what is significant about manipulation in education, I end without providing definite answers. Though remaining uncommitted may seem like a way to avoid criticism, and probably it is also that, the main reason to avoid definitive answers is that I am not trying to win an argument nor solve an enigma. I am trying to help others navigate the complex nature of educational processes. In the same way that there is more than one useful map of London, it appears obvious to me that there is more than one useful conceptual scheme of education.

To summarize, my research approach assumes that education is a complex process. This complexity means there is a multiplicity of aims and means in education, and that these are not wholly separable from one another. For this reason, empirical research findings that appear generalizable are not necessarily so. Moreover, due to the complexity of education, it appears useful to represent education in a simplified form, but a single conceptual framework is limited. In order to provide a better understanding of education, I wish to provide a conceptual

50

framework of benign manipulation in education that would be useful for educators, when they decide whether or not to employ such techniques.

1.5 Chapter description

In the second chapter, I provide a simple categorization of the methods by which we might intentionally influence others, and I discuss the nature of manipulation. Broadly speaking, the different methods of intentional influence can be gathered into three groups: coercion, reasoning and manipulation. Benign manipulation should be understood as 1) an intentional influence of one person on another person's beliefs, desires, emotions or behaviour. This influence is 2) not coercive, nor is it persuasion through ideal reasoning. And, 3) the manipulator has good reasons to believe that the action will benefit the manipulee. Building on recent work in analytic philosophy I explain different theories of manipulation: manipulation as deception, manipulation as subversion of rationality, manipulation as pressure, manipulation as failure to track reasons, and hybrid theories. I conclude the identification question by stating that an attempt to influence, which exhibits one or more of the attributes discussed, can usefully be analysed as manipulation.

51

Chapter Three establishes a framework that allows for an assessment of benign manipulation. As an analogy, just as one would like to choose a certain form of transport over another (train vs. car, etc), I claim that in order to assess the use of benign manipulation we require an understanding of the significant features of such use. A good place to look for those, is the ongoing discussion of the ethics of manipulation. I conclude by establishing the following framework for assessing benign manipulation:

A) Consequences: 1) intended consequences – to what extent is the suggested benign manipulation effective in achieving the desired result in the short term, long term and across contexts? 2) unintended consequences – to what extent does the suggested benign manipulation have unintended consequences?

B) Understanding: 1) to what extent does the manipulation prevent the manipulee from having an accurate understanding of the situation? 2) to what extent does the manipulation prevent the manipulator from having an accurate understanding of the situation?

C) Trust: is the manipulation appropriate according to the norms which guide behaviour in the relationship between the manipulator and manipulee?

Chapter Four deals with the challenge to the common belief that there is a difference between manipulating children and manipulating adults. Often, it is assumed that children are less autonomous, and because it is further assumed that manipulation harms autonomy, it is claimed that manipulation does not harm children in the same way that it harms adults. My reasons for opposing this argument are that manipulation does not necessarily harm autonomy, that adults are not fully autonomous, that children exhibit at least some degree of autonomy, and that for both children and adults we have a good reason to treat them as if they are autonomous. Furthermore, I suggest a conceptualization of childhood that is not based on 'autonomy' and can explain why children and adults should be treated differently when it comes to manipulating them. This conceptualization is based on the epistemic difference between children and adults, in simple words – that most of the time adults have a better grasp of reality than children. Because adults know better, it makes sense that they would attempt to influence children on more occasions. However, we should not assume that adults always know better than children. For this reason, we should use the framework that was established in the previous chapter to assess whether a specific act of manipulation is appropriate within an educational process that involves children.

The next three chapters delve into three case studies of manipulation in education. The first is the use of the so called 'Pygmalion Effect', the idea that teachers' expectations of students' abilities serve as self-fulfilling prophecies. I show that a recommendation to implement the 'Pygmalion Effect' necessarily involves deception and analyse its flaws and merits by using the manipulation assessment framework. My analysis shows that although in particular instances teachers would be wise to communicate false expectations, it should not be their default option.

The second case study is the use of so called 'smart drugs'. There is a growing interest in recent years in the use of pharmacological cognitive enhancement for those who are not diagnosed with any physical or cognitive abnormal condition. The liberal use of pharmacological cognitive enhancement has been criticized by some for principled moral reasons, but these are mostly unconvincing. I use the framework for assessing manipulation to draw attention to previously unacknowledged problems with the liberal use of pharmacological cognitive enhancement. Mostly, I claim that such liberal use will prevent different stakeholders from understanding the current situation and desired goals of the educational process.

The third case study involves the use of gamification in education. Gamification is the growing popular trend to use game mechanics in non-game

situations, such as training and education. I use the framework for assessing benign

manipulation in education to demonstrate that often the use of gamification

techniques in education is appropriate, but also that it can be misused and work

against the aims of well-intentioned educators.

Chapter 2   The Nature of Manipulation

2.1   Introduction

I stated that benign manipulation is 1) an intentional influence of one person on another person's beliefs, desires, emotions or behaviour. This influence is 2) not coercive, nor is it persuasion through ideal reasoning. And, 3) the manipulator has good reasons to believe that the action will benefit the manipulee. This chapter will explain what is meant by benign manipulation being an intentional influence which is not coercive nor ideal reasoning.

People influence each other's thoughts, feelings and behaviours in different ways. Many of these influences are unintentional. For example, a colleague might show up to work wearing a yellow shirt, which will remind me of a childhood incident involving a yellow shirt, and that memory would make me sad. Perhaps most of the times we influence or are influenced by others are of this kind, but unintentional influence is not what I shall be focusing on. We attempt to intentionally influence someone when we do something in order to bring about or raise the likelihood of a desired behaviour or state; such as character traits, beliefs, desires and emotions.

56

The range and diversity of the means by which we do so is extremely wide.

However, it is useful to classify these into three major categories: reasoning,

coercion and manipulation. These categories should not be taken at face value – the

boundaries between the categories are not always clear and I am not sure that

beyond certain paradigmatic cases, real life instances of interpersonal influence can

be neatly classified as falling strictly under one of these types. However, although

such classification necessarily simplifies the contextualized situations of real life, it is

useful in advancing our understanding of the different methods used to intentionally

influence others. The aim of such classification is to assist our thinking about the

means by which we influence others, without falling into oversimplification - if

possible.

The above definition and the categories of intentional influence I will discuss

below are general conceptual categories, not necessarily to do with education. This

is because much of the analytical discussion of these categories has been done by

general philosophers, not those who specialize in educational theory. However, as

my main aim in this work is to understand the use of manipulation in education, I will

use examples from educational contexts in order to explain their thinking, and

wherever possible, will bring in references these theorists make to education.

57

Moreover, when relevant I will refer to the ideas of philosophers of education and other theorists who focus on the field.

2.2 Coercion

Alan Wood defines intentional coercion as influencing the behaviour of others by threatening them or making them expect consequences in a way that makes certain options (the ones the coercer wishes the coerced to choose) the only ones reasonably acceptable (Wood 2014, p. 21). For example, threatening a student with physical punishment if she will not keep quiet is a paradigmatic case of coercing her into being silent in class. It appears that Wood's definition would not apply to less extreme punishments, such as demanding that the student leave the classroom, detention, etc.

Quite differently, Horner and McIntosh provide a definition that encompasses a significantly wider array of actions: 'Coercion is the contingent use (or threat of use) of aversive events (pain, criticism, reprimands, aggression). Coercion occurs in many forms in schools…'(Horner and McIntosh 2016, p. 330). For the discussion of coercion in schools, the definition by Horner and McIntosh appears more suitable. At least, these are indeed instances that teachers and students will experience and

58

describe as coercive. That said, their definition is not without problems. For example, would we include a bad grade as an aversive event? If the answer is yes, then it follows that every grade which is experienced as a 'bad grade' by the student is an act of coercion. If not, then the definition needs to be changed so that it would not include many aversive events. More generally, the question that Horner and McIntosh fail to answer is whether every act experienced as coercion should indeed be classified as coercion. Relatedly, the line between what is coercive and what is not is hard to establish according to their definition. I will not attempt to answer these difficult questions here. For my purposes, it is enough to draw some insights about coercion that are common to many, if not all, coercive actions.

There are a few characteristics of coercion that emerge from the definitions above. First, when coercing someone the actions and threats are visible. If not explicit, they should at least be known implicitly or commonly understood. Punishing students after an undesirable act has been committed, without telling them in advance that they will be punished for it, might coerce them only in regard to their future actions, because they will expect to be punished from now on for these kinds of actions. Importantly, because coercion is visible, it invokes reaction on the part of the coerced. Second, coercive actions are directed mostly towards behaviour and action, and not towards beliefs, desires and emotions. Suppose someone tries to

59

coerce me into believing in God by threatening to kill me. This gives me a very good reason to act as if I believe in God, but not to actually believe in him. To use another example, the threat of a fine might make me wear a seat belt, without necessarily influencing relevant beliefs about safety etc. In this example, coercion gives reasons to act in a certain way (seatbelt), and those reasons are explicit (fine), but are not necessarily the reasons that we would ideally prefer the agent to act upon (safety). Third, coercion is a method of influence in which the coercer has relative control over the reasons that influence the coerced. In the classroom this might be phrased thus: "If you do not X, I will Y! So, you had better X". The teacher has control over whether to punish the student or not and can choose which actions will be punished and which will not. Thus, on the one hand Coercion resembles rational reasoning in the sense that reasons to behave in a certain way are made visible, the coerced is recognized as responsive to reasons, and she is presented with reasons to behave in a certain way. On the other hand, and unlike ideal rational reasoning, in coercion the reasons to behave in a certain way originate in the coercer.

It might be intuitive to assume that there is something inherently wrong in coercing someone, possibly because that would be to treat that person just as a means to an end, or because it would be a violation of that person's personal autonomy and freedom, or for other reasons. The question is beyond the scope of

60

this work. To keep our intuitions balanced, it should be kept in mind that in many cases even if coercion is *prima facie* wrong, all things considered it might be beneficial to everyone involved – for example in the case of the law that threatens those who do not wear seat belts with fines. Alan Wood comments that "there are important kinds of freedom available to you by way of your being forced to be free" (Wood 2014),[1] and this might be especially true in educational relations.

The use of coercion in education, or at least some forms of it, has been criticized as being not only immoral but also ineffective (Glasser 1992; Horner and McIntosh 2016). It is questionable whether these criticisms are correct, but it is safe to assume that in the last one hundred years the use of 'hard' coercion in education, and especially corporal punishment in schools, has declined considerably, at least in liberal democracies.[2] Possible causes might be growing ethical concerns regarding children's status and rights (expressed also in the 1989 UN Convention on the

---

[1] In this he probably echoes Rousseau, who had a similar idea in mind: "whoever refuses to obey the general will shall be constrained to do so by the entire body; which means only that one will force him to be free" Affeldt, S.G. 1999. 'The Force of Freedom: Rousseau on Forcing to be Free', *Political Theory* 27(3): 299-333.

[2] Corporal punishment was banned by Parliament in state funded schools in Britain in 1987, in private schools it was banned as late as 1999 in England and Wales, 2000 in Scotland, and 2003 in Northern Ireland. In the United States the Supreme Court ruled in 1977 that the spanking or paddling of students by school officials or teachers is lawful, where it has not been explicitly outlawed by local authorities. Only 31 (out of 50) states have abolished it, either de facto or de jure. It remains a widespread practice in Alabama, Arkansas and Mississippi. In most European countries and in Canada it is illegal. For more details, see: Rules and procedures for the administration of corporal punishment. https://www.corpun.com/rules.htm.

61

Rights of the Child), or perhaps a realization that it is not an effective method of motivating students, or a diffusion of ethical concerns from adult society into institutionalized and familial children's education. Other kinds of coercion are very much still in use in education and take various different forms (Noddings 2001, p. 35), such as the threat of punishment in schools (Horner and McIntosh 2016, p. 330).

Quite recently, a few instances of the use of coercive techniques have gained much publicity and many adherents. A notable example is the bestselling book 'Battle Hymn of the Tiger Mother' by Amy Chua (Chua 2011). In the book, Chua describes the coercive methods she routinely used in the education of her two daughters. Following in her footsteps, teachers in a school in North London have contributed essays to a book by the name of 'Battle Hymn of the Tiger Teachers: The Michaela Way' in which they advocate the use of coercive techniques in the education of under privileged children (Birbalsingh 2016). Another London school, the Mossborne Academy, similarly uses coercive techniques in the education of its pupils (Fowler 2011). I believe these examples gained much attention and publicity because they represent a use of coercion that is highlighted in contrast with the spirit of the time. Also, some of the techniques used in these coercive educational processes are closer to what I mean by manipulation. One example is the panoptic

62

architecture of the Mossborne Academy, aimed at controlling student and teacher behaviour by designing the physical environment of the school in a way that promotes certain kinds of behaviours.

In sum, coercion as an intentional form of influence is characterised by the visibility of the reasons that encourage the relevant behaviour, it is directed towards actions and behaviour (rather than emotions, desires and beliefs) and the reasons to behave in a certain way originate from the coercer. Moreover, although it is intuitive to assume that coercion is *prima facie* wrong, it is not necessarily the case. And, even if coercion is *prima facie* wrong, it is often justified all things considered. That said, the trend of the last hundred years of formal and informal education appears to be characterised by diminishing use of 'hard' forms of coercion, such as corporal punishment.

2.3 Ideal Reasoning

Another category of intentional influence is ideal reasoning. First, reasoning means recognizing and respecting the other as an agent who is responsive to reasons. Second, it means providing reasons, in the form of facts, evidence, theories and arguments that justify a certain belief, state, or behaviour. It is important to note

63

that an influencer might provide reasons in the form of threats, or provide reasons that are false, fake or misguided. In these cases, even if the intentional influence is targeting the agent's reasoning capabilities, it is more suitable to refer to these as coercion or manipulation. Due to that, I refer to this category of influence as ideal reasoning, to indicate that the reasons offered do not arise out of the influencer's own actions, and that the influencer's reasons are good reasons, in the sense that they justify the belief, state or behaviour they are supposed to promote.

For example, if someone does not think global warming is real, and as a result behaves in a way that is not at all considerate of that truth, you might give that person an argument involving evidence, theories and inferences that would convince that person to change his or her behaviour because he or she will come to realize that global warming is real. This is the kind of influence that scientists are at least ideally supposed to have on one another and the public regarding their field of research. However, if the state legislators pass a bill that forbids the use of certain chemicals due to their harmful environmental effects, threatening to punish culprits, that would be influencing by coercion. In sum, ideal reasoning is similar to coercion in that the influence is visible and it involves recognising the other as responsive to reasons. However, it differs from coercion in two main aspects. First, it is normally aimed at beliefs rather than behaviour. Second, the reasons to behave are

64

dependent on the actions of the coercer, i.e. the coercer has at least some control over the reasons that influence the coerced. For example, the coercer might use threats to influence the behaviour of the coerced.

As an educator, I might want to convince others that it would be good for them to think, behave or act in a certain way, or that this is the right and correct way to think or behave. For example, I might motivate my students to put more effort into their studies by telling them that this will raise their chances of getting into the university they want. If I have good reasons to believe that this is the case (that if they study hard, they are more likely to be admitted to the university of their choice), that would be attempting to influence them by ideal reasoning. Because it seems to be a wonderful way of influencing others while at the same time respecting them as rational agents, one would expect that most instances of intentionally influencing others would follow or aim at the characteristics of ideal reasoning. But this is not the case, for various reasons which are briefly listed below.

Reasoning ability: convincing someone who is capable of understanding the truth is hard enough but convincing someone who at that specific moment is not capable of such understanding - is just impossible. Young children come to mind as an example. If a child is incapable of understanding that too much candy harms their health, we can try and convince her all we want, it would not work. Similarly, children

65

might not be capable of understanding 'advanced' truths, such as complex mathematical ideas (that the sum of all angles of a spherical triangle is larger than 180 degrees) and this gives a reason to simplify or bend the truth for them, which takes us away from ideal reasoning. Adults who are non-experts in a certain field would be in a similar situation to that of small children in that regard, as they might not be able to understand the reasons that should guide certain behaviours.

Cost to the influencer: as participants in political debates can testify, it is very hard to convince someone who holds the opposite position. This is true even when the truth is relatively straightforward, such as in the case of global warming. The production of reliable evidence, theories, reasons and arguments takes much effort and time. It is hard work. Reasoning might be a great ideal, but it would be very costly if we want to put it into practice in many instances in which we wish to influence others.

Cost to the reasoner: not only is reasoning costly for the influencer, it is also costly for the reasoner. In recent years there has been an explosion of research about cognitive biases, which are systematic flaws in human reasoning. In his book 'Thinking Fast and Slow' Daniel Kahneman, a pioneer of behavioural economics, summarizes the research findings of the past half century (Kahneman 2011). Kahneman and other scholars in psychology, behavioural economics and the

66

cognitive sciences are pointing towards the fact that actual reasoning is prone to systematic mistakes and a range of cognitive biases. Human adults tend not to be the rational reasoners we might want them to be. On the contrary, they fit more closely with our conception of children, in the sense that they constantly deal with a variety of reasoning hurdles. It is possible for someone to come closer to ideal reasoning, but that involves significant effort on the part of the reasoner herself. As a result of these findings, there is an ongoing attempt to use the knowledge produced about the actual reasoning of human beings in policy in areas as diverse as health, crime, economics, education and more. Projects such as the 'Nudge Unit' in the UK[3], and centres for behavioural economics and behavioural change have sprung up in many universities around the globe. While the efficacy of many of the suggested interventions is very much debatable, the basic insight that a person's actual reasoning is quite distinct from ideal reasoning appears to hold. To come closer to the demands of ideal reasoning the reasoner needs to engage and make a significant effort on her part. Often the reasoner lacks the motivation to do so.

---

[3] The Nudge Unit is more formally known as the 'Behavioural Insight Team'. It was founded in 2010 as a small unit within the UK Cabinet Office, and has since grown to become a multinational consulting firm, which aims to apply behavioural science in real world contexts. For more information see: https://www.bi.team/about-us/

Moreover, it is impossible to have both the motivation and the capacity to engage with all of the different issues others are attempting to influence us towards.

Weakness of the will: another problem with trying to influence someone's behaviour by reasoning with her is that even if you convince that person of the truth of the action guiding belief, it still does not guarantee that that person will behave according to that belief. So, if you have convinced a friend that smoking harms her health, that does not necessarily mean she will stop smoking. Her other beliefs, desires and emotions come into play, and she may smoke while knowing that it is unhealthy. Philosophers sometimes call this phenomenon *akrasia* or weakness of the will, a discussion that goes back to Plato and Aristotle (Aristotle 1954). There is much more to say about the nature and significance of weakness of the will,[4] but it is one more clue that ideal rational reasoning is limited in guiding actions. It is not enough to convince someone of a certain belief, because holding that belief does not guarantee they will behave in a way compatible with it. Moreover, we might actually

---

[4] The nature of the phenomenon is difficult to explain, as is also its moral significance. While some would argue that it is undesirable for someone to act against what they reason to be the right action, some might argue that sometimes acting against what one thinks is the right thing to do is actually a good thing. A case in mind is the conscience of Huckleberry Finn, who does not betray Jim, the runaway slave, even though Huckleberry Finn strongly believes that he is doing the wrong thing. See Bennett, J. 1974. 'The Conscience of Huckleberry Finn', *Philosophy* 49: 123--34.

68

want people to not always act on their beliefs, since their beliefs might be wrong, and even immoral in some cases.

The norms of reasoning: another problem with ideal reasoning is that it is not always obvious that this is what is going on. It is common for people to appear as if they are engaging in ideal rational reasoning while they are really doing something different. Politicians and others in leadership roles might appear as if they are trying to influence others by rational reasoning while in truth they are using other kinds of influence.[5] As a reasoner, one is not only required to put effort into the reasoning process itself, but also to judge whether the influencer is genuinely assisting one's reasoning process or just faking it for whatever reason. Is the shoe seller recommending these shoes because she genuinely thinks they are the best shoes for me, or is she attempting to sell me the shoes that she gets the most profit out of? How would I know if what is going on is a genuine reasoning process or a fake one? This is especially hard if we consider Aristotle's claim that different areas of knowledge in human life are not subject to the same degree of precision (Aristotle 1998). Therefore, even if I gained enough knowledge about shoes and experience in shoe stores to know whether the seller is genuinely reasoning with me or just faking

---

[5] The concepts of 'Spin' and 'Bullshit' are relevant here. See for example: Frankfurt, H.G. 2005. *On bullshit*. Oxford: Princeton University Press.

69

such a process, I might not have a clue as to whether my doctor is engaged in ideal reasoning with me or just trying to sell off the drug that was promoted by the drug company. There are two related issues here; first, the influencer might hide some reasons in order to better convince the reasoner, and second, the norms of ideal reasoning vary. What might be considered good and genuine reasoning in shoe selection, might not be good and genuine reasoning in health treatment, and as a reasoner, it is hard to judge the situation correctly. The same goes for different academic fields: what might be considered ideal reasoning in economics, might not be considered ideal reasoning in mathematics, or the arts, or any other discipline. The norms of reasoning vary depending on the nature of the knowledge the process is aiming for.

To conclude, while in theory we might hold ideal reasoning as the best way to influence others' behaviour, there are many different difficulties in putting this ideal into practice. Some people might not be capable or motivated to engage in rational reasoning at the level required; it is very costly for both the influencer and the reasoner to engage in rational reasoning; reasoning is meant to produce beliefs, but holding a belief does not guarantee behaviour that is compatible with that belief; and it is challenging for the reasoner to correctly judge the situation as one in which rational reasoning is taking place, and also to apply the appropriate norms of

70

reasoning to the topic at hand. Since ideal reasoning is very difficult to put into practice in many cases in which we wish to intentionally influence others, and there is a growing suspicion towards coercion, manipulation appears to be an obvious way forward.

2.4 Manipulation

A few examples of manipulation were mentioned earlier. However, is it possible to give a positive account of what manipulation is, which goes beyond providing examples and differentiating it from coercion and ideal reasoning? In the next few sections, I will present different theories which explain what is characteristic about manipulation, and how it might be identified. As the analysis will show, there is currently no consensus on the matter.

2.4.1 Manipulation as Deception or Trickery

If we think of the Long Way Home example, the parents appear to be deceiving their child into believing that the way back home is longer, deceiving him in regard to their intentions by withholding important information from him (that they want him to fall asleep) or tricking him into sleep. In line with this account, one theory would be that manipulation involves deception or trickery, and that this is what

71

makes it different from coercion and ideal reasoning. This view is motivated by a consideration of many seemingly paradigmatic examples of manipulation which do appear to involve some form of deception.

An example of this view can be seen in Vance Kasten writing about manipulation in education. He claims that being manipulated involves being misled:

> … manipulation occurs when there is a difference in kind between what one intends to do and what one actually does, when that difference is traceable to another in such a way that the victim may be said to have been misled. (Kasten 1980, p.54)

This is certainly the case with many examples of manipulation. However, even in the Long Way Home example, what causes the child to fall asleep does not appear to be necessarily the deceit, but rather the setting up of the situation in a way that encourages sleep. So, even if often manipulation does include deceit, even if it is present, it is not necessarily the cause of the desired influence. This can be generalised to many cases in which the influencer raises the chances of the preferable action happening by making it easier for the person influenced to do it. A health official who would like to get more people vaccinated would influence people by opening many vaccination clinics with no queues. In such a case she would be influencing the behaviour of others, without coercing them, reasoning with them, or

72

deceiving them. The concept of manipulation under analysis here, which is a non-moralised one, includes such instances.

Moreover, not all instances of what we might consider morally suspect manipulation involve deception at all. Moti Gorin provides the following counterexample to the deception account of manipulation:

> **Off the Wagon**: Wilson and Adams are up for promotion, though only one of them will get the job. Adams is a recovering alcoholic and Wilson sets out to encourage a relapse, intending this to disqualify Adams for the promotion. Wilson consistently drinks alcohol in front of Adams, offers her alcoholic beverages, vividly describes to her whatever benefits there are to drinking and to drunkenness, and so on, all the while making no secret of his intentions. During a moment of weakness brought on by a particularly difficult and stressful event, Adams takes a drink, which leads to more drinks, missed days at work, and an overall decreased ability to meet the demands of her job. When the time comes to announce who will be promoted, Adams is told by her managers that her recent poor performance has made it impossible for them to give her the new job

73

and that they have selected Wilson for the promotion. (Gorin 2014b,

pp. 80-81)

In this example, Wilson has manipulated Adams by encouraging her desire to drink alcohol. Adams knows the true intentions of Wilson, and yet this does seem to be an instance of manipulation. In other words, this is a case of manipulation which does not involve deception.

In the field of education, it is possible to think of setting up a competition in class as a form of manipulation which does not involve deception:

Competition: A teacher who is being financially rewarded if her students perform above a certain threshold decides to motivate her students to improve their grades in the coming exam, by setting up a competition between them. She tells them that the list of students ranked by their score on the exam will be placed on a public message board at the school for all to see, including parents. Students at the top of the list will be hailed as successful, capable and praiseworthy, while students at the bottom of the list will be shamed as dumb and lazy. She explains to them her reasons for setting the competition without hiding her intentions and motives.

74

Let us assume that at least some of the students do learn more about the topic than they would have done without the competition. In this case, the teacher influences those students, but she is not deceiving them. To be clear, perhaps some of the students will experience the teachers' actions as coercive and feel (rightly or not) that they are being coerced into studying harder, but some students might not consider this a coercive intervention – they will be motivated to win, without resenting the teacher for what she is doing. For these students, the teacher has influenced them to behave in the way she wanted them to behave without misleading them. This example shows that manipulation does not necessarily involve deception, even if often – it does.

2.4.2 Manipulation as Non-Rational Influence

Another theory maintains that manipulation is a kind of influence that 'subverts' or 'bypasses' the rational capacities. For example, Wood identifies manipulation as a method that "influences people's choices in ways that circumvent or subvert their rational decision-making processes, and that undermine or disrupt the ways of choosing that they themselves would critically endorse if they considered the matter in a way that is lucid and free of error." (Wood 2014, p. 35) According to this view, manipulation is a form of influence that either fails to engage the rational

75

capacities of the manipulee, or else engages these capacities in some way that undermines their function (Gorin 2014b, p. 90).

The idea that manipulation is a form of influence that subverts or bypasses the rational capacities of the manipulee is appealing for various reasons. Firstly, because manipulation is distinct from ideal reasoning, then it would make sense that what is unique about manipulation is related to its divergence from engaging the rational capacities of the manipulee. Secondly, examples in which the form of influence clearly bypasses or subverts the manipulee's rational capacities seem intuitively to be forms of manipulation. For example, if subliminal advertising is effective, this would be a way of influencing someone by bypassing her rational capacities, and this intuitively appears to be manipulative. Therefore, it appears reasonable to understand manipulation as an exploitation of psychological mechanisms and techniques that can promote desired behaviour, without the need to engage with the rational capacities (Noggle 2018).

However, not many examples of manipulation seem to fall under the category of 'bypassing' the rational capacities completely. In educational contexts, one such example might be the Long Way Home, in which the parents appear to bypass the rational capacities of their child as they attempt the induce a non-rational state such as sleep. However, in the examples mentioned earlier of False Encouragement and

76

Competition, the rational capacities of the manipulees are engaged. The students manipulated in Competition, are given reasons to behave a certain way. They are encouraged to compete with each other because winning will bring about good consequences for them. Likewise, in False Encouragement, students are given reasons, although deceptive, to have certain beliefs about themselves. Therefore, their rational capacities are engaged in the process of manipulation. Indeed, at times the rational capacities are the ones being targeted by the manipulation, such as in the case of Competition, in which the students are encouraged to deliberate on the consequences of their actions, and realize they have reasons to compete against their peers. Therefore, it appears to make more sense to think of manipulation as a subversion of the rational capacities, and not a complete bypassing of them. So perhaps manipulation should be understood as "bypassing rational deliberation" in the sense that it brings into the deliberation non-rational influences (Noggle 2018). But, many non-rational influences do not appear to be manipulative. Some influences appeal to feelings rather than the facts, and yet we would not intuitively think these are forms of manipulation. For example, appeals to the Golden Rule invite the person to put herself in someone else's position, to feel as if she were that someone. It does not appear to be necessarily manipulative to do so. Therefore, not every non-rational influence is necessarily manipulative.

77

Perhaps the most serious challenge to the claim that manipulation bypasses or subverts reason comes from Moti Gorin (Gorin 2014a). He argues that manipulation can happen even if the manipulee is offered only good reasons. Here is a similar example to the one Gorin presents in support of this claim:

> **No God**: James wishes for Jacques's death, since this would enable James to inherit a large fortune. James knows that Jacques believes that (1) God exists, and that (2) if God did not exist, life would be meaningless, and he would have no reason to go on living. James provides Jacques with rational arguments against the existence of God. These arguments fully engage Jacques's rational faculties, and consequently Jacques concludes that God does not exist. Jacques promptly commits suicide—just as James had hoped he would. (Noggle 2018)

In this example, James does nothing to undermine the rational capacities of Jacques. If we accept that this is an act of manipulation by James, then it is a serious challenge to the explanation that manipulation always subverts the manipulee's capacity for rational deliberation. Furthermore, this example shows that it is possible for manipulation to be compatible with ideal reasoning in some cases, as James manipulates Jacques by using ideal reasoning.

78

A distinction is sometimes made between rational and non-rational methods of influence when it comes to educating children. John Tillson claims that 'Rational means affect influence by providing sound reasons for instantiating attributes… Non-rational means cover all other methods' (Tillson 2021). According to Tillson, non-rational means include: lies, deception, bullshit, coercion, manipulation and indoctrination. My view aligns more closely with that of Gorin, in which a manipulator can make use of sound reasons, and appeal to the rational capacities of the manipulated. That is, manipulation is not necessarily non-rational. Moreover, Tillson's distinction is problematic because he assumes that rational methods are unobjectionable while non-rational methods can be wrongful, and this is because of their different nature. Specifically, Tillson claims that non-rational methods can be wrongful by offending children's interests in wellbeing, autonomy, and equal regard (Tillson 2021). However, as the No God example above shows, good reasons can also be used against the interests of the person influenced. The opposite is also correct, i.e. that non-rational methods can be used to promote the interests of the person influenced. The question of whether and when manipulation is wrong must be related to something other than its rational or non-rational nature. I will discuss this further below.

79

2.4.3 Manipulation as Pressure

Other theories maintain that manipulation is best understood as an influence in which the manipulator pressures the manipulee to do as she wishes (Noggle 2018, p. 18). Examples of manipulation such as emotional blackmail and peer pressure seem to be paradigmatic of this kind of influence. One reason to think of manipulation as pressure, is that it can be placed on a spectrum between rational persuasion at one end, in which there is no pressure, and coercion at the other, in which there is very significant pressure. The example of Competition discussed above can be viewed as an instance of such pressure manipulation. Manipulation will then account for all instances of influence that fall below the coercion threshold of pressure (Wood 2014). Michael Kligman and Charles Culver offer such an account:

> The attempt to influence B's behavior takes on a manipulative character when … A's primary intent is no longer to convince B, in a good faith manner, that acting as desired by A would be in keeping with B's rational assessments of outcome; [but rather] to procure or engineer the needed assent by bringing pressure to bear, in a deliberate and calculated way, on what he presumes to be the

manipulable features of B's motivational system. (Kligman & Culver 1992: 186–187)

One challenge for such a theory is to explain where to set the threshold which differentiates coercion from manipulation. Kligman and Culver go on to distinguish manipulative pressure from coercion by claiming that the latter, unlike the former, involves "sufficiently strong incentives … that it would be unreasonable to expect any rational person not to so act" (Kligman & Culver 1992: 187). Wood offers a similar understanding of coercion: "I am coerced to do something when I either do not choose to do it or if, when I do choose to do it, I do it because I have no acceptable alternative" (Wood 2014, p. 21). However, even if it is possible to set the threshold between coercion and manipulation, the manipulation as pressure account does not fit well with cases of manipulation which use deception – in which there does not appear to be much pressure.

The theory that explains manipulation as a form of pressure is probably not a complete theory of manipulation. Most authors who discuss the pressure theory think that there are other forms of manipulation, specifically – that these are related to deception. Or, they maintain that pressure is still related to one of the other theories of manipulation, such as the subversion of rationality one (Wood 2018, p. 35). It is more accurate to speak of the pressure model as claiming that exerting non-coercive

81

pressure is sufficient (but perhaps not necessary) for an influence to count as manipulative (Noggle 2018, p. 20-21).

2.4.4 Manipulation as a failure to track reasons

A novel account of manipulation comes from the work of Moti Gorin. He claims that manipulation is a form of influence that fails to track reasons, and this characteristic is also what makes it morally suspect (Gorin 2014, p. 92). Gorin makes the very illuminating suggestion that for manipulators, who only care about the causal efficacy of reasons and arguments, it does not matter whether they were to use good reasons and arguments or bad ones. This is related to the attitude of the bullshitter, who does not care about the truth value of his statements, only the possible effect his message might have on his listeners (Frankfurt 2005).

The example Gorin provides is this one:

> **Election**: Jones is running for president. He has never taken the time to educate himself about matters of public concern and in fact could not care less about politics. Jones is attracted to the presidency because he intensely craves attention and knows that those who occupy this office are constantly in the spotlight. In order to make himself politically saleable Jones hires a stable of pollsters,

82

speech writers, and acting coaches. He constructs a political platform with the maximal vote-getting potential, as determined by his polling experts. He also commits to following through on his campaign promises once elected, since he calculates that this may help him win a second term and hence to extend his time as the centre of attention. Though personally indifferent to the ideals reflected in the policies he publicly defends, Jones—with the help of his speechwriters and coaches—convincingly advocates for his platform in speeches and debates and garners enough support to win the election. (Gorin 2014b)

In this example, Jones is manipulating his voters and followers by using good reasons and arguments, only he does not care whether these are good reasons and arguments. His choice of reasons and arguments is based on their efficacy and not on their merit as real attitudes and the bases of real policies. This example illustrates the idea that even good reasons and arguments can be used manipulatively. Gorin believes this is a form of manipulation because Jones's voters and followers, though he supplies them with good reasons, fail to realize that he is using these arguments and reasons not because he truly believes in them, but because he knows they are the most effective means of promoting his own desire for attention.

83

Gorin distinguishes between three forms of manipulation in relation to their use (or abuse) of reasons. The first he calls "non-paternalistic reasonable manipulation". In this kind, the manipulee acts in a way that is supported by reasons, such as in the case of Jones's voters in Election. They are moved to vote for Jones for good reasons. This is an act of manipulation because Jones does not hold his policy suggestions on the basis of their merit to do good, but on the basis of their efficacy to motivate his potential voters. The second kind of manipulation Gorin names "paternalistic manipulation". In these, the manipulee is motivated to act in ways which the manipulator believes are supported by good reasons, from the perspective of the manipulee. Many examples discussed under the heading of 'nudge' are instances of paternalistic manipulation. In these, the manipulator does think that there are good reasons for the manipulee to act in the way she is manipulated to act. However, even though paternalistic manipulators aim for behaviour that is supported by good reasons (such as the well-being of the manipulee), and not only because it is effective, they choose methods that fail to track reasons. In other words, the manipulee is likely not to fully understand why she is acting in the way she does. In the words of Gorin: "Paternalistic manipulators typically choose manipulation only when reason-tracking forms of influence are unlikely to work and when they believe the reasons that support the behavior they seek to bring about are sufficiently weighty to justify non-reason-tracking means of

84

influence" (Gorin 2014, p. 95). The example of Long Way Home fits well in the category of paternalistic manipulation. The parents have good reasons to choose a technique that does not track reasons, because a) the child cannot be reasoned into sleep, and b) it would be good for the child to fall asleep. Moreover, the child's challenging question is also explained by this account, since the child does not understand the reasons why the way back home is longer than the way towards his grandparents. In his own way, he is asking for a reasoned explanation. The third kind of manipulation according to Gorin is "unreasonable manipulation", in which the manipulator does not believe the aimed behaviour of the manipulee can be supported by good reasons from the viewpoint of the manipulee. Hence, the manipulator resorts to other kinds of influence such as deception, pressure, or subversion of her rational capabilities (Gorin 2014, pp. 95-6).

Although I find Gorin's suggestion that manipulation is a form of influence that fails to track reasons insightful, it does have some limitations. In refuting the theory of manipulation which focuses on deception, Gorin discusses an example that I believe refutes the failure to track reasons account of manipulation. I have provided this example above, but let me add it here again:

> **Off the Wagon**: Wilson and Adams are up for promotion,
> though only one of them will get the job. Adams is a recovering

85

alcoholic and Wilson sets out to encourage a relapse, intending this

to disqualify Adams for the promotion. Wilson consistently drinks

alcohol in front of Adams, offers her alcoholic beverages, vividly

describes to her whatever benefits there are to drinking and to

drunkenness, and so on, all the while making no secret of his

intentions. During a moment of weakness brought on by a particularly

difficult and stressful event, Adams takes a drink, which leads to more

drinks, missed days at work, and an overall decreased ability to meet

the demands of her job. When the time comes to announce who will

be promoted, Adams is told by her managers that her recent poor

performance has made it impossible for them to give her the new job

and that they have selected Wilson for the promotion. (Gorin 2014b,

pp. 80-81)

Gorin maintains that in this example Wilson manipulated Adams by engaging her

compulsion to drink alcohol, and that Adams is aware of the intentions of Wilson.

Therefore, he concludes, this is a counterexample to the manipulation as deception

account, as there is no deception going on, yet intuitively this seems to be a case of

manipulation. If this is true, this shows that the failure to track reasons accounts also

fails to capture all cases of manipulation, as Adams is aware of Wilson's reasons to

86

behave the way he does (he wants the promotion) as well as her reasons to behave the way she does (she has a compulsion to drink alcohol). Therefore, it is not the case that every manipulation necessarily involves a failure to track reasons.

It is possible to maintain that Adams fails to track the reasons that cause her to desire alcohol, apart from the soliciting behaviour of Wilson. However, we are hardly ever aware of many of the different reasons which play a part in our mental and physical lives. Moreover, many interpersonal influences fail to track reasons, yet we would not consider them instances of manipulation. When I am moved to tears by a dramatic movie or a play, I can give as the reason that I was moved by the fantastic acting, but why good acting causes me to tear up is a much more complex question and I do not have immediate access to the reasons that explain it. Just like Adams, when I am moved to tears, I fail to track some of the reasons that play a part in causing my behaviour, but it would be odd to claim that the actor manipulated me. Hence, not every interpersonal influence which fails to track reasons is an instance of manipulation, and not every manipulation necessarily involves a failure to track reasons. That said, a failure to track reasons does appear to be a characteristic of many instances of manipulation.

87

2.4.5 Other Views of Manipulation

We have seen that even though there are reasonable accounts of manipulation, none of them manages to capture all examples of manipulation which we intuitively judge to be instances of the phenomenon. The theories of manipulation presented above are either too narrow so that they do not capture all cases of manipulation, or too wide so that they fit actions which we do not intuitively consider as manipulation. Different instances of manipulation appear to incorporate features such as deceptiveness, hinderance of rational deliberation, pressure, and a failure to track reasons, but none of these features is a sufficient condition in identifying manipulation, and no single condition of the above is necessary for some action to count as manipulation. The fact that manipulation operates by different mechanisms requires us to question the assumption that the concept 'manipulation' refers to one type of actions. There are a few possible ways to proceed from this point.

Virtue Ethics Perspective

In her earlier work on the concept Marcia Baron focuses her explanation of manipulation on the character trait of manipulativeness. She explains manipulativeness through a virtue ethics framework, drawing on Aristotelian ideas.

Baron claims that manipulation is a vice, and what this vice gets wrong is fundamentally:

> How much to steer others – and which others, and how, and when, and towards what ends; and more generally, to what extent - and how and when and to whom and for what sorts of ends – to seek to influence others' conduct. One errs in either or both of the following ways: seeking to change another's conduct when one should not; or using means one should not use, in those circumstances and towards that person. The manipulative person is too ready to think it appropriate – or appropriate for him – to orchestrate things so as to lead others to act as he wants them to; and in those instances where it is not inappropriate to try to engineer such a change, the manipulative person is too ready to employ means that should not be employed. By contrast, the person who has the virtue corresponding to manipulativeness - a virtue for which we do not, I believe, have a name - knows when it is appropriate to try to bring about a change in another's conduct and does this for the right reasons, for the right ends, and only in instances where it is warranted (and worth the risks), and only using acceptable means. The virtuous person tries to reason

89

with the other, not to cajole or trick him into acting differently - with some exceptions…" (Baron 2003, p. 48)

As a general description of manipulativeness I find Baron's explanation illuminating. The idea of manipulation as a vice related to the attempt to steer and influence others captures what I believe we mean when we say that someone is a manipulator or that someone acted manipulatively. The issue becomes more complicated when we try to provide details that might assist in deciding whether a certain act, behaviour or practice is in fact manipulative. That is, thinking in terms of virtue ethics places the emphasis on the character of the person, and not on the action itself. As my main aim in this work is to discuss methods used in education, and not the character traits of the agents involved in the process, the explanation of manipulation within a virtue ethics frame is of limited use.

Further problems arise from the explanation of manipulation within the framework of virtue ethics. For one, sometimes manipulation is indeed done in the interests of the manipulated (for example, in Long Way Home). So, if the manipulator mostly influences others in a manipulative fashion in order to contribute to their well-being, does that mean they should be regarded as having the vice of manipulativeness? Also, the description of manipulation above does not explain well

the difference between coercing someone and manipulating them. It does not tell us what is unique about manipulation as a form of influence.

Disjunctive and Hybrid views

Another approach would be to theorize manipulation as a disjunctive or hybrid term. A disjunctive view of manipulation will maintain that we can identify manipulation as a form of influence which is either deceptive or subverting rationality, or pressuring the manipulee, or fails to track reasons, etc. A hybrid view of the term would allow for combinations of the above.

Manipulation might best be explained, then, using the idea of 'Family Resemblance' (Wittgenstein 1967). 'Family Resemblance' is the idea that concepts refer to particular instances without there being a set of sufficient and necessary conditions that may provide a common definition for all those particular instances. Rather, what links the different instances is a looser 'family resemblance' where one instance would be somewhat similar to another in some regards, and different in others, and then similar to another instance in other aspects, etc. This would be akin to family members who might not share one common feature among all of them, but each shares some common features with some of them. The analysis of such concepts should not aim for a strict definition, rather it should aim for a looser drawing out of the similarities and differences between the various family members.

To explain, consider the categorization of games, which is the example that Wittgenstein employs in his discussion of ordinary language (Wittgenstein 1967): one can assign games to different categories (ball games, board games, juggling etc.), in order to draw some general insights about different types of games, for instance that some games require a ball, others are based on a board. This may be useful for advancing our understanding even if every single act of playing a game is inherently contextualized to such a degree that to assign it to a category will be to oversimplify the reality of that specific instance of playing a game. Just like we would find it impossible to provide a definition of the word 'game' in the form of sufficient and necessary conditions, so are we unable to provide a single theory to explain the term 'manipulation'. If this is correct, then the classification of the means of intentional influence is not meant to provide definitions in the form of sufficient and necessary conditions that would enable us to recognize a certain act or process as falling under one of these types of influence. Instead, it aims to clarify some conceptual confusion regarding different means used to influence others, to better enable us to understand what is involved in the different means we employ in doing so. This should put us in a better position to reflect upon such uses.

For the purpose of this work, an understanding of manipulation in the form of a family resemblance is the best way forward. As my aim is to analyse practices

92

used in education, I am less concerned with supporting one theory or definition of manipulation over another just for the sake of getting the theory right. The purpose of this chapter is to use others' definitions and explanations of manipulation to draw out some characteristics of manipulation, which are relevant to the educational process. It is not necessary that each and every act of manipulating someone will demonstrate all or even most of these characteristics, but these are things we should bear in mind while looking at specific actions, routines and techniques used in educational settings. Therefore, the discussion above may be concluded thus: actions which are either deceptive, pressuring, perverting choice or fail to track reasons, are candidates to be considered acts of manipulation. Hence, whenever an educator chooses a method of influence which involves one or more of these features, then it should be considered an act of manipulation. As my focus is benign manipulation and I assume a non-moralised understanding of manipulation, in the sense that the manipulator has good reasons to believe that the action will benefit the manipulee and manipulation is not *prima facie* wrong, to claim that an instance of interpersonal influence is indeed an instance of manipulation does not yet tell us whether such action is desirable or not. This will be the focus of the next chapter.

2.5 Indoctrination

It is useful to supplement the above discussion about the nature of manipulation with reference to indoctrination, a concept that is often engaged with in the philosophy of education literature. Early discussion of it appears in R.S. Peters' 'Reason and Compassion' in which he writes that "'indoctrination' involves the passing on of fixed beliefs in a way which discourages questions about their validity". (Peters 2015, p. 47). More recently John Tillson characterizes indoctrination as 'the sustained engagement in practices that increases the likelihood of subjects' formative and behavioural modes becoming resistant to revision in the face of compelling countervailing reasons' (Tillson 2021). While Tillson does not make much of the distinction between beliefs and actions, other theorists maintain that the concept of indoctrination should be reserved for educational processes in which people are made to believe certain things, while the concept of conditioning should be reserved for feelings and actions (Wilson 2010). In short, indoctrination, as it is mostly discussed in the literature, is an educational process in which at least some of the following: the intention of the educator, the methods used, the content delivered and the outcomes of the process, aim for or result in 'closed-minded' individuals who are 'unable or unwilling to give due regard to reasons that are available for revising their current beliefs' (Callan and Arena 2009, p. 111).

94

Manipulation, as used in the current work, diverges from indoctrination in a few important ways. First, unlike the common assumption that indoctrination is morally wrong, or at least morally suspect, I favour a view of manipulation that does not assume it to be objectionable from the start. This point will be further developed below. Second, manipulation is a type of interpersonal influence, and is not necessarily related to education or teaching. Indoctrination, on the other hand, is commonly seen as a type of 'bad teaching' (Taylor 2017, p.40; Wareham 2018, p. 43). Although my interest in manipulation stems from my work as a teacher, and this work is partly aimed at informing teachers' practices, the concept of manipulation is relevant to many other spheres in which interpersonal influence takes place. Third, although it is possible that indoctrination will involve methods and techniques that can be categorized as manipulation, indoctrination can also be done by coercive techniques as well as rational persuasion. Thus, my focus on methods raises questions that are unlikely to be answered in the discussion of indoctrination. For instance, the example of giving false encouragement, 'fixing' the environment, setting up a game in class or providing pharmacological enhancers, appears to be relevant to educational processes whether these are indoctrinating or not. Lastly, benign manipulation does not necessarily aim for or result in 'closed-minded' individuals. In fact, the interest I find in benign manipulation stems from it being a part of many, if not most, educational processes, including the best of these. As will

95

be discussed later, benign manipulation in education often aims and results in the opposite of 'closed-mindedness'.

Although my focus is on benign manipulation, the discussion does bear on the concept of indoctrination. For example, as the example of No God above shows, one is able to manipulate another by using true statements, good reasons, and appeal to the manipulee's rational capacities, and I believe this would apply to indoctrination as well. That is, one is able to indoctrinate by using true statements, good reasons and by appeal to the rational capacities of the target of indoctrination. One non-hypothetical example that I've personally experienced is the focus on Holocaust studies in Israeli schools, which, at least in principle, may result in 'closed minded' individuals who are less likely to consider the suffering of people whose identity is different than theirs. This specific example as well as the possible connections between manipulation and indoctrination merit far more attention than I give it here, and I hope to pursue the issue in the future.

Chapter 3   A framework to assess benign manipulation

3.1   Introduction

As was discussed in the previous chapter, it is not clear whether we can define manipulation using a set of necessary and sufficient conditions. Therefore, I suggest we proceed with a 'family resemblance' understanding of manipulation. Interpersonal influences which do exhibit one or more of the characteristics mentioned above: deception, subversion of rational capacities, pressure, and failure to track reasons, should be thought of and analysed as instances of manipulation.

When an action is an instance of benign manipulation, in the sense that the manipulator has good reasons to believe that the influence would benefit the manipulee, it still does not tell us whether it would be an appropriate course of action. To decide whether it is, we need to consider the relevant reasons that would make it appropriate or not in the specific context. To explain, consider one wishes to travel from point A to B. In order to choose which method to travel by, we would require a description of the facts of the matter, such as how far A is from B, what

97

possible routes and methods of transport are available to get from A to B, what

capabilities our traveller has (can she walk? drive? cycle?) and so forth. Next, we

might want to consider the benefits and drawbacks of using each method. For

example, taking the bus from A to B might be quicker, but travelling by bicycle has

the added benefit of exercise. We could, if we so wished, examine multiple cases

and list the relevant aspects in assessing which method to use in general terms.

However, this would not be the end of our decision process in any particular case, as

we also need to consider the relevant features of the specific context. In other words,

even if we are aware of the different aspects that should be considered, we would

still be required to judge how these aspects play out in the specific situation.

Similarly, in what follows I wish to offer a framework that would assist our

thinking about manipulation as a method of influencing others.[6] Suppose that an

educator, parent or teacher, is faced with the choice of whether to use a technique

that involves one or more of the features mentioned in the previous chapter

(deception, subversion of rational capacities, pressure, and failure to track reasons).

As in the case of travelling from A to B, she would need to assess the benefits and

---

[6] My framework is based on the work of J. S. Blumenthal Barby. While there are significant differences in the details of the framework itself, I found her general methodology and the majority of her analysis extremely useful. For more information see: Blumenthal-Barby, J. 2014. 'A Framework for Assessing the Moral Status of Manipulation', in Manipulation: Theory and Practice, Coons, C. & Weber, M. (eds.), Oxford University Press, pp. 98-120.

drawbacks of the method. In order to do that, we need to understand what the relevant aspects of benign manipulation are that should be considered. Therefore, the overall aim of this chapter is to discuss the salient aspects of an act of manipulation we should be aware of when deciding whether to use manipulation or not. However, once we have a better understanding of the different features of manipulation, it should still need to be considered in the specific context. To exemplify an assessment of specific cases I discuss three case studies in detail in the later chapters.

Although I am concerned with benign manipulation, in which the manipulator has good reasons to believe that the manipulation will benefit the manipulee, in the next few sections I will look into some of the accounts that scholars who judge manipulation to be morally suspect have given in their explanation of what is problematic about manipulation. These will provide us with insight into what we should consider when assessing the use of benign manipulation. This inquiry will show that there are three main aspects we should look into when assessing the use of benign manipulations in education. First, we need to ask whether the technique is likely to be effective in achieving the desired result, and whether there are possible unintended consequences of using it. The second relevant aspect might be called the epistemic one. This aspect deals with the epistemic burden that is placed on both

the manipulee and manipulator due to the use of benign manipulation. Third, as the norms that govern social interaction vary from one social context to another, we should inquire whether in the specific context an act of benign manipulation breaches the norms of the relationship. In such instances an act of benign manipulation, even if effective, safe and epistemically unproblematic, might constitute a betrayal of trust. I will consider other suggested harms of manipulation, especially manipulation's supposed harm to personal autonomy, to show that these are not useful in assessing benign manipulation.

3.2 Criterion A) Consequences

Some accounts of the wrong involved in manipulation rest on the idea that manipulation is often done by the manipulator to gain some advantage at the expense of the manipulee. Following the act, the manipulee is worse off in some way, and this consequence of the action has something to do with the kind of influence manipulation is. These theories view manipulation as a moralised concept, which agrees with the use of the term in everyday language. In other words, whenever manipulation is done, there is something objectionable about it, at least *prima facie*. As discussed above, because the focus of my work is benign manipulation, I am less concerned with manipulative acts which are intended to

100

exploit the person being manipulated. Assuming that the term manipulation may refer to acts which are motivated by a concern for the manipulee, examples of manipulation which are beneficial to the manipulee are easy to find. Thaler and Sunstein discuss many examples in their book 'Nudge', such as this one:

Cafeteria: A cafeteria manager at some organization notices that customers tend to choose more frequently items that are presented earlier in the line. Therefore, she is presented with a choice – she may influence which food the customers choose to eat. Her choice might be to maximize customers' well-being by presenting healthy food before less healthy food, she might maximize profits by presenting the most profitable food first, she might take revenge on her unthankful customers by presenting the less healthy food first, she could present the food in the way she thinks her customers would want it to be presented, or else she might make a random choice. Whichever choice she makes, at least some of the customers will be nudged to choose food they would not have chosen otherwise. In this case, they will be manipulated, but this will not jeopardize their autonomy – as they can still choose the food further on down the line. (Sunstein and Thaler 2003, p. 1162)

101

In the case of Cafeteria, if the manager decides to present the food in a way that would influence the customers to eat a healthier diet, she would be influencing their conduct, in a way that operates on their choice selection. This appears to be a benign act of manipulation, which does leave the customers better off, at least in the important aspect of their health.

Some might claim that any interference in another person's conduct is morally suspect. This is related to the issue of autonomy that will be discussed in more detail below. However, putting aside for the moment the possible infringement on one's autonomy, it is clear that in some instances a person's judgement of their own interests is not beneficial to them. This might be due to non-rational tendencies, such as the ones Thaler and Sunstein focus on, or it might be due to other factors such as being tired, drunk or just mistaken. This does not mean that intervening in their conduct is always justified, but it seems unreasonable to think that it is always morally suspect. Indeed, as Baron suggests, avoiding any sort of intervention in others' conduct can also be thought of as a vice, which she names 'isolationism' – an 'it is non-of my business' attitude. Consider a drunk friend who is about to enter her car to drive home. If she will not listen to reason, manipulating her to hand over the keys would be appropriate, and abstaining from doing so would be the morally problematic decision (Baron 2003, p. 48). If an effective and benign way to intervene

102

in someone else's conduct is manipulating them, it is unreasonable to claim that any form of influence which involves manipulation is wrong. Therefore, we may conclude that at least some instances of manipulation are benign or at least do not cause harm to the manipulee. Many educational techniques which are intended to be beneficial to the manipulee appear to fall under this category.

Because manipulation does not necessarily harm the manipulee, a good place to start assessing the use of benign manipulation would be the consequences of the action. Discussions of the appropriate use of benign manipulation often start with the assumption that the manipulation in question is effective in achieving the manipulator's desired result. I will follow this assumption, although when we assess suggested techniques or methods as manipulation, it should be stressed that the empirical research into these techniques is often overstretched, overstated or just poorly conducted. In my analysis of the case studies in chapters five, six and seven I will support this claim by dedicating more attention to the actual effectiveness of the methods under discussion. For now, it is certainly the case that effectiveness in achieving the manipulator's desired result is a key consideration in choosing whether to use the manipulation or not.

That said, benign manipulation might operate differently over the long term than it would in the short term. For instance, in the Long Way Home example, the

parents' manipulation is effective for a while, but it will probably lose its effectiveness over time, as the child gains a better understanding of what is happening. Very often, pedagogical manipulative techniques are claimed to be effective through research conducted over the short term, but their effectiveness in the long term is not assessed. However, when put into practice, educators should consider whether suggested techniques are effective over the long run, and not just in the present moment. Furthermore, benign manipulation should not be assumed to be similarly effective across contexts. For example, we might find, through research or personal experience, that setting up a competition promotes learning for a group of students. This does not yet assure us that the same method would be similarly effective for a different group of students.

Moreover, recommendations to use a certain technique in education are sometimes blind to possible unintended consequences of that technique. For example, it is often the case that the recommendation to use a certain technique in the classroom is based on statistical findings that show it was effective in raising academic performance. We might find that using method X improved the results of 80% of the students. But, what about the other 20%? Has method X had no effect on those students, or maybe it had a detrimental effect on them? In medical research, the undesired effects of a certain intervention are often discussed under 'safety

104

concerns' and 'side effects', so that even if an intervention is very beneficial in one aspect or for some patients, it will be scrutinized in order to discover its effects on all patients. In educational research, this does not appear to be the case, as there is no regulatory educational equivalent to 'safety' concerns. But, if an intervention has unintended consequences this is very important in practice. Therefore, in assessing an act of manipulation, we should ask what the overall effects of it are, not just have statistical evidence that it is useful for some.

In sum, the discussion above shows that Criterion A) Consequences should consist of the following considerations: 1) Intended consequences – to what extent is the suggested benign manipulation effective in achieving the desired result in the short term, long term and across contexts? 2) Unintended consequences – to what extent does the suggested benign manipulation have unintended consequences in the short term, long term and across contexts?

3.3 Criterion B) Understanding

I have mentioned before that there are two common uses of the term manipulation. One, when it is applied to acts of interpersonal influence which are morally questionable, and the second when it is applied to the intentional handling of

things, such as the operations performed on objects of scientific inquiry. Some accounts of the moral wrong of manipulation are based on the idea that there is something problematic in applying the kind of engineering approach to actual persons. In other words, that it is morally wrong to treat persons as things (Noggle 2018, p.33). This is in line with Kant's claim that we have a duty to treat others always as an end in themselves: 'Act in such a way that you always treat humanity, whether in your own person or in the person of any other, never simply as a means but always at the same time as an end' (Kant and Paton 1967). Kant's claim may be read as a prohibition on treating human beings solely as objects we can operate on.

However, in cases where treating someone as an object increases their overall well-being, is it necessarily wrong? To keep our intuitions balanced, consider Cafeteria again. The cafeteria manager treats the customers as less than fully rational agents, and if she builds on the general tendency to choose food that is presented earlier in the line in order to improve the health of her customers, she is treating them as clever machines that can be operated upon for their own benefit. In other words, they are being treated as things for their own good. Similarly, in the examples of Long Way Home and Competition, the parents and teacher are trying to get the children to do something by influencing their conduct in a non-rational fashion. However, as discussed above, this does not necessarily involve lack of

106

respect. To argue that it does assumes that the only way to influence others respectfully is to engage in rational persuasion with them. This would exclude important and desirable ways of influencing others, such as appeals to emotions, imagination, etc. The assumption that every non-rational influence necessarily involves lack of respect for the person as an agent, assumes an over cognitive and rational human agency (Noggle 2018).

Nonetheless, I believe there is a sense in which not treating the other with respect does point towards an aspect of manipulation that is very significant. The problem with treating persons as things can be understood primarily as epistemic, and only secondarily as normative. Manipulating someone often disregards them as a knower. The manipulator is interested in influencing the manipulee without having to deal with the possibility that the manipulee might object to such influence. In the most extreme examples of benign manipulation, the manipulee's point of view is not given much importance, as the manipulator assumes that her understanding of the situation is correct, as well as her understanding of the desired state of affairs. The manipulee does not get a say on the matter. In this, the manipulee is not being respected as a source of valuable information, even if she is left with a choice. She is respected as a chooser but not as a knower.

107

There are some cases where this is not a serious problem. For example, in the Long Way Home, the parents are manipulating their child, and it is not problematic that their child's perspective is not being respected. The parents are well placed to judge what the current state of affairs is (their child is tired) and they are well placed to judge the desired state of affairs (sleep). And, of course, they have their child's best interests at heart. Therefore, whether the child has a different opinion or view of either the current or desired state of affairs is less important. However, in other cases, for example in False Encouragement, the manipulation might prevent the student from having an accurate understanding of the situation and she would not challenge the teacher over the teacher's assessment of her abilities, or maybe even disagree that the aim of performing better on the exam calls for deceiving her. In manipulating her, the teacher does not respect her as a knower that might challenge the teacher's understanding of the current and desired state of affairs.

The conjoined effect is that by using manipulation the manipulator might be prevented from acquiring valuable information as well, namely – the information of what the manipulee feels or thinks about what is going on. As an example, consider an online dating site in which a user takes on a completely different persona in order to appear more attractive to a possible partner. In such a case, not only is the

manipulee deceived in regard to the real situation, but the manipulator will also lose important information – the feelings or thoughts of the possible partner about who they genuinely are. In Competition, if a competition is set in order to motivate students to learn more about a certain subject, the teacher might lose important information such as what is the internal motivation the students have for learning this topic. Therefore, at times benign manipulation will cause both the manipulator and the manipulee to have a less accurate understanding of the situation.

The manipulator's fallibility is a significant issue in the assessment of benign manipulation. In her discussion of paternalistic interventions, Mullin maintains that we should be aware of the possible pitfalls for the intervening actor as well as the dangers to the target of the intervention (Mullin 2014, p. 423). Educators should consider that they might be mistaken in their understanding of the situation at hand, in their assessment of what the interests of the child under their care are, etc. Because manipulation often makes it harder for the manipulee to know what is going on and to challenge the power exercised over her, this makes manipulation especially problematic from the epistemic point of view of the manipulator. Successful manipulation will make it harder for the manipulator to receive good and accurate feedback, which is required in order to make the manipulator at least sometimes reflect upon her actions and beliefs.

109

Therefore, I believe the lack of respect involved in manipulation is best understood in epistemic terms and not in normative terms, though these two aspects are intertwined. Epistemically, manipulation is more likely to hide the true state of affairs from both the manipulee and manipulator, more than ideal reasoning and coercion. For this reason, we should add the following to the framework: Criterion B) Understanding, which contains two important elements. First, to what extent does the manipulation prevent the manipulee from having an accurate understanding of the situation? And second, to what extent does the manipulation prevent the manipulator from having an accurate understanding of the situation?

3.4   Criterion C) Trust and Trustworthiness

I would like to return to the analogy between the choice regarding how to influence someone and the choice about how to travel from point A to B. I have claimed that broadly we have two sets of questions, one about the specific context (such as, the distance between A and B) and another about the methods we might choose to travel by. So far, I have focused on the latter, and presented some of the characteristics of manipulation as a method of influence that we should be mindful of when debating whether to use it or not. Next, I wish to discuss an important aspect of the context manipulation might take place in, which is relevant in assessing it as a

method of influence - the nature of the relationship between the manipulator and the manipulee.

Patricia Greenspan suggests that when manipulation is wrong, it is because it violates the terms of the relationship between the manipulator and his target—terms that will vary according to the nature of the relationship between them (Greenspan 2003). In other words, at least *prima facie,* 'the objection to manipulation would be that it betrays the sort of trust taken for granted in certain interpersonal interactions' (Greenspan 2003, p. 163). Such a view suggests that the moral status of a given instance of manipulation will depend at least in part on the nature of the relationship between the influencer and the target of the influence. Greenspan provides these two examples:

> Consider generating interest in one's academic field among students by ignoring problems with it that might deter a beginner, feigning a kind of uncritical enthusiasm, presumably for the manipulee's ultimate benefit rather than one's own. Or consider cheering up a friend by drawing attention to various positive features of a life that lately has been depressing. If someone in a state of melancholy may be said to want to dwell on his misfortunes, spelling out the end here might undermine the effort. These are examples of

111

a kind of paternalism, but far outside the legal contexts in which we object to paternalism. In these cases the manoeuvres in question seem to be well within the scope of "caretaking" authority accorded to a teacher or a friend. (Greenspan 2003, pp. 155-6)

According to this view, acts of manipulation are a part of everyday relationships. A relationship here does not necessarily mean intimacy - it is a relationship in the sense that both parties expect certain forms of behaviour of one another. In other words, there are norms guiding the behaviour of the two parties. Different types of relationship are guided by different expectations and norms of behaviour, and in certain relationships the parties trust one another, either explicitly or implicitly.

Although there are many different ways to understand trust, most accounts of interpersonal trust agree with this general articulation:

Trust is an attitude that we have towards people whom we hope will be trustworthy, where trustworthiness is a property, not an attitude. Trust and trustworthiness are therefore distinct although, ideally, those whom we trust will be trustworthy, and those who are trustworthy will be trusted. For trust to be warranted (i.e. plausible) in a relationship, the parties to that relationship must have attitudes

towards one another that permit trust. Moreover, for trust to be warranted (i.e. well grounded), both parties must be trustworthy.

Trusting requires that we can, 1) be vulnerable to others (vulnerable to betrayal in particular); 2) think well of others, at least in certain domains; and 3) be optimistic that they are, or at least will be, competent in certain respects. Each of these conditions for trust is relatively uncontroversial. (McLeod 2015)

To highlight and explain some of the details in this general account, consider the example of trust between two friends, Amy and Beth. For trust to be possible, the two friends should be somewhat vulnerable to one another. If Amy does not allow herself to be vulnerable to Beth, then she would not be trusting her, even if Beth is trustworthy and deserves to be trusted. Also, Amy should expect goodwill and competence of Beth, at least in certain aspects. If she believes that Beth is malicious, or well-intentioned but incompetent, she would not be able to trust Beth in that regard. Therefore, for mutual trust to be possible, both friends need to have these attitudes towards one another. However, they might have these 'trust facilitating' attitudes, but be mistaken in having those, because their friend is not trustworthy. Amy might believe that Beth is well-intentioned, competent, and would make herself vulnerable to her, but she might be mistaken. Trust is not just the

attitude or feeling of being trustful, it also needs to be grounded in the trustworthiness of the person being trusted.

It would have been useful if we were able to provide a list of attributes which makes one trustworthy, and so explain when it would be justified to trust that person. Without such substance, the definition of trustworthiness as the property of the person whom it would be justified to trust appears circular. However, even if we think of some obvious possible attributes, such as honesty, for example, it is not the case that an honest person is necessarily trustworthy in that regard, or that a dishonest person would be untrustworthy. What we expect from one another can be quite complex, and involve a mixture of contrasting behaviours and attributes, such as both deception and honesty. To take a simple example, two poker players expect each other to do their best to deceive one another in regard to the strength of their 'hand' of cards, but also not to cheat by various means. Their trust in one another requires them to be deceitful in certain aspects. Many other interactions involve expectations of a mixture of behaviours. Romantic partners might expect each other to be honest in various domains, and yet dishonest in others. Shakespeare probably says it best in Sonnet 138:

> When my love swears that she is made of truth,
>
> I do believe her, though I know she lies,

114

That she might think me some untutored youth,

Unlearned in the world's false subtleties.

Thus vainly thinking that she thinks me young,

Although she knows my days are past the best,

Simply I credit her false-speaking tongue:

On both sides thus is simple truth suppressed.

But wherefore says she not she is unjust?

And wherefore say not I that I am old?

Oh, love's best habit is in seeming trust,

And age in love loves not to have years told.

    Therefore I lie with her and she with me,

    And in our faults by lies we flattered be.

Both the example of the poker players and Shakespeare's romantic relationship highlight that a relationship can be trustful, and yet incorporate and indeed flourish due to lies, deceptions and manipulation. If the manipulee expects, would not object and even prefers to be manipulated, the manipulator might not be doing anything regrettable when she manipulates in order to promote the interests of the manipulee. In such cases, the mutual expectations that guide behaviour in the relationship make some manipulations unproblematic. Therefore, in order to better understand what

trustworthiness consists in we need to look closer at the context in which trust takes place. In other words, the substance of what trustworthiness consists in is different in different domains (poker game, romance), and so in later sections, I will explain what generalisations about trustworthiness we might make in relation to education.

Finally, I wish to add that one's trusting attitude in another should not be seen as a license to treat her in whichever way the trusted party wishes. This is true even if the trusted has consciously chooses to make herself vulnerable to the other person. Greenspan provides the following example:

> Imagine Ulysses tying himself to the mast to withstand the Sirens - but then betrayed by a disloyal crew that takes advantage of the occasion to stage a mutiny. Autonomy has been violated here, though the violation depended on a voluntary transfer. (Greenspan 2003, p. 161)

The example illustrates that even in the case where one voluntarily sacrifices some of her autonomy, her self-government, it does not give license to the other person to manipulate her in whichever way she pleases. In the example above the problem is that the sailors exploit the trusting attitude of Ulysses to advance their own interests. Therefore, we might wish to add the condition that if someone voluntarily chooses to trust another person, to make herself vulnerable, this should not be used against

them. In the case of manipulation, the condition would be that within a trusting relationship one party should not abuse the other party's trust in order to exploit her. However, how does this play out in cases of benign manipulation, in which the manipulation is aiming to advance the interests of the manipulee? In other words, can benign manipulation ever constitute a breach of trust? The answer appears to be yes, but in order to understand why we need to look further into the issue of trust.

Some trust relationships involve allowing oneself to be vulnerable in a voluntary and conscious way. For instance, when a patient decides to get an operation in which she will be sedated and operated on by a physician, she may consciously give her consent to that. However, other trust relationships do not involve conscious consent. In some we have something like unconscious or implied consent, such as the trust one might have in a bus driver if she has decided to travel on the bus. In these examples, whether conscious or unconscious, trust is focused on competency in a specific skill or attitude, the operating skills of the surgeon and the driving skills of the bus driver. The patient might trust the physician about her health, but not with her life's secrets. This is often the nature of trust in short-term social interactions.

One way in which benign manipulation might breach trust is if the manipulator steps beyond the focused trust the trusting party placed in them. Placebo treatment

117

comes to mind. Placebo treatment is the use of sugar-pills or some other safe medical intervention with no empirically proven active ingredients, that at times has positive healthy outcomes for patients. If a physician decides to use a placebo without letting the patient know that the drug they are being prescribed is indeed a placebo, then although she might have the patient's best interests in mind and aim for improving the patient's health, she might be breaching the patient's trust. At least some patients would rather not receive placebo treatment without consenting to it, even if it will improve their health. For them, being manipulated in this case is a betrayal of the terms of their relationship with their medical caretaker, which they expect not only to improve their health, but also to do so by using certain means (real drugs) and not others (fake drugs). Hence, there are definitely examples of benign manipulation that do breach the focused trust the manipulee has placed in the manipulator.

However, in encompassing long-term relationships, such as we have with family members, the norms that govern the behaviour in the relationship might be incredibly hard to formalise, unlike in the example of the poker game. In such close relationships, it makes more sense to think of trust as a mindset or attitude, that is not focused on specific skills. Furthermore, in many of these long-term relationships where trust is of utmost importance, it is not based on consciously or unconsciously

118

putting oneself in a vulnerable position towards the other. The vulnerability is a constituent part of the relationship, such as is the case of children's trust in their parents, or romantic partners. In such relationships, some extreme breaches of trust, such as cheating or abuse, will obviously be a breach of general commitments that are part of the relationship. However, other breaches of trust might be subtler, and require further attention to understand. In such relationships whether benign manipulation is appropriate or not depends on the specifics of the relationship and the mutual expectations that guide the social interaction. Below, I look more closely at the nature of trust in teacher-student relationships in order to judge if and when benign manipulations are appropriate in some educational processes.

To conclude, we now see that the norms of the relationship in which manipulation is taking place are also relevant in judging the merits and flaws of such a method of influence. Therefore, we need to add the following to the framework: Criterion C) Trust. This criterion focuses on whether the manipulation is appropriate within the norms of behaviour which constitute the relationship between the manipulator and manipulee. In other words, does the benign manipulation breach the trust the manipulee has in the manipulator?

119

3.4.1 Teachers' trustworthiness

In what follows I wish to investigate further what trust means in teacher-student relationships. I explained that trust should be understood as an attitude we have towards people whom we hope are trustworthy (in the aspect in which we trust them). The trustor is vulnerable towards the trustee, should think well of the trustee (at least in the domain in which they are trusted), and believe that they are competent in the relevant aspects (McLeod 2015). For trust to be justified, the trustee must be trustworthy, and be trusted for the right reasons, i.e. because they are trustworthy. If we apply this to a student's trust in their teacher, then the student who trusts their teacher is vulnerable to the teacher, should think well of their teacher and believe that the teacher is competent (at least in the aspects in which they trust them). For the student's trust to be justified, the teacher should be trustworthy. The question I focus on in this part, is what does it mean to be a trustworthy teacher?

First, it is important to note that although one might present attributes that make her seem trustworthy, she might not be genuinely trustworthy. It is possible that someone acts as if they are trustworthy, but even if they are being trusted, that should not count as a trustful relationship, as they are not truly trustworthy and the trust in them is not justified. For example, someone might fake a caring attitude, which will make them less trustworthy than genuinely having that attitude. Faking a

120

caring attitude, or acting as if one is trustworthy, may be considered a form of manipulation, as it involves deception. Again, as I assume an un-moralised understanding of manipulation, faking trustworthiness might be entirely appropriate. An important point here is that empirical data about what members of the school community perceive as trustworthy, has limited value, as it does not tell us what being a trustworthy teacher is, rather – it tells us what people perceive as trustworthy. For example, in a summary of empirical research on trust in schools, Tschannen-Moran and Hoy combine a list of different definitions of trust to the following formulation of the meaning of trust in schools:

> Trust is one party's willingness to be vulnerable to another party based on the confidence that the latter party is (a) benevolent, (b) reliable, (c) competent, (d) honest, and (e) open… Moreover, there is empirical evidence that all of these facets are important aspects of trust relations in schools. (Tschannen-Moran and Hoy 2000, p. 556)

This definition does not tell us what being trustworthy is, but what is perceived as trustworthy. Indeed, some aspects of what is perceived as trustworthy may contradict one another. Specifically, benevolence may be contradictory to being honest in various cases. Therefore, the empirical evidence of what members of a school community believe makes one trustworthy is of limited value if we wish to

121

understand what makes one trustworthy, rather than what makes one appear

trustworthy.

Hence, is it possible to provide a general account of what attributes make one

a trustworthy teacher? I shall follow the work of various theorists in assuming that

these attributes are truthfulness[7], goodwill and competence (Baier 1986, p. 251).[8]

That is, for a teacher to be trustworthy they should be motivated by goodwill towards

the student, be truthful and competent. For the student's trust to be justified, she

should trust the teacher because the teacher is trustworthy, i.e. truthful, caring and

competent. But again, the question arises, what happens when two of these

attributes contradict one another? If the teacher thinks that what is good for the

student is to deceive her, these two aspects of trust conflict. The conflict between

truthfulness and goodwill is characteristic of asymmetrical trust relationships, and

---

[7] See Cooper, D.E. 2008. 'Teaching and Truthfulness', Studies in Philosophy and Education 27(2-3): p. 81. For a discussion of truthfulness in relation to trustworthiness see Williams 2002, p. 92; Strudler, A., 'Deception and trust', in *The Philosophy of Deception*, Martin C. W. (ed.), Oxford University Press, p. 139.

[8] On trust as truthfulness and goodwill see Cooper 2008, p. 85. On trust as goodwill see Baier 1986, p. 259; Strudler 2009, p.141. The concept of trust does not have to be understood as containing these two aspects; or might even be understood as two different concepts. This is not very important for my argument - if one does not agree with this analysis of trust, she may still consider both truthfulness and goodwill to be central to educational relations. What interests me here is that though both truthfulness and goodwill are pivotal in education, they are sometimes in conflict.

sometimes referred to as 'Paternalistic Lies'.[9] I discuss these more fully in Chapter Five.

The attributes of truthfulness, goodwill and competence are central aspects of trustworthiness in other role derived practices. Suppose a patient trusts her physician. This means the patient believes what the physician is telling her is what the physician believes to be true, and also that the patient has confidence in the good intentions and professional competence of the physician. The ethical discussion about placebo treatment stems from the conflict between these commitments of the physician. If the patient receives a prescription from the physician, she believes 1) the physician is truthful – what is written in the prescription is not deceptive, and 2) the physician acts out of goodwill towards the patient – the patient believes the physician gave her the right treatment to improve her health. Now suppose the physician prescribes a placebo to the patient by writing a code name on the prescription, which the pharmacist will understand but the patient will not. Suppose also that placebo treatment is the right treatment for the patient in this situation. In this case, the second aspect of trust overrides the first - the physician deceives the patient but she does so in order to maintain and promote the second aspect of trustworthiness; the patient's belief that the physician is doing her best to

---

[9] For a more elaborate though inconclusive discussion of 'Paternalistic Lies' see Bok 1980.

improve her health. In this case, is the physician trustworthy? This is a very difficult question to answer.[10] However, we can find some clues in the fact that placebo treatment is not without problems, even if it is medically safe and improves the well-being of the patient (Criterion A). Firstly, the use of a placebo might hinder the accurate understanding of the situation for both the physician and the patient (Criterion B). Secondly, if prescribed often, such practice might jeopardise their trustworthiness and the extent to which they are trusted by patients (Criterion C). This is not to mean that placebos should never be prescribed, but neither is it the case that they should be used without careful consideration of non-deceptive alternatives (Bok 1980, p. 72).

Although there are significant differences between physician trustworthiness and teacher trustworthiness, some similarities exist. For teachers, like physicians, sometimes the attributes of truthfulness and goodwill might contradict one another. To better understand how these might manifest in educational relationships, it is worth considering some of the characteristics of these relations. First, within the institutional settings of compulsory education these relations are asymmetrical and

---

[10] See discussion in Gold, A. and P. Lichtenberg. 2012. 'The Moral Case for the Clinical Placebo', *Journal of Medical Ethics*; and Lichtenberg, P., U. Heresco-Levy, and U. Nitzan. 2004. 'The Ethics of the Placebo in Clinical Practice', *Journal of Medical Ethics* 30(6): 551–554.

non-contractual.[11] Normally, the student enters the student-teacher relations not as a free agent and is the weaker part in these asymmetrical relations. The teacher in her role as a teacher has formal command over the student and the relations are asymmetrical also because she is older, has a better understanding of the long-term goals of the educational process, a better understanding of the means to be used, more knowledge of the subject matter, skills or values to be learned, etc. The hierarchical structure of these relations makes the teacher responsible more than the student at least for the means used in the educational process if not for the results of the process. The hierarchical structure, because it is usually accompanied by an unequal distribution of power, also makes it possible for the teacher to deceive students without suffering very much when the deception fails. That is, teachers are in a good position to use deceptive means and justify them by appeal to their goodwill. This makes deception a very handy tool for teachers who prefer not to use visible means such as coercion or reasoning. Of course, it being easy to deceive students does not make it right to do so.

Another important consideration is that student-teacher relations are ongoing, and so an abuse of trust may have lasting implications. If a teacher deceives a

---

[11] Baier discusses trust in asymmetrical relationships and the importance of a moral theory to guide us in these, as compared to and distinct from contractual relations. Baier 1968, pp. 244-253.

125

student and is not believed, this might influence how the student trusts the teacher in the future (Tschannen-Moran and Hoy 2000, p. 563). Also, because it is a complex interaction, a student may find the teacher trustworthy in certain things – such as the curricular content knowledge of that teacher - and untrustworthy in others – such as her expectations or praise (Tschannen-Moran and Hoy 2000, p. 570). Lastly, as the social roles of 'being a teacher' and 'being a student' are shared among many different individuals, the actions of individual teachers and students have an effect on how others in the same role are viewed. For example, if a teacher or student is caught lying, that might harm the way others in that role are perceived.

I believe that the mistake those who recommend communicating false encouragement as a policy make is viewing goodwill as a means to achieve an aim, rather than as a commitment the teacher has. To explain and develop this point, I wish to consider three possible views of the importance and meaning of trustworthiness in education, understood as truthfulness, goodwill and competence. The first possibility is that the attributes of trustworthiness should be viewed as means used to achieve other educational goals; the second is that they should be viewed as commitments, and the third is that they should be viewed as virtues. There is some truth in all these options, but the core understanding of teacher's trustworthiness is the second one – that it is a commitment.

126

Truthfulness, goodwill and competence as means

One possible explanation of the importance of trustworthiness in education is to view it as a means to reach desired educational goals (Tschannen-Moran and Hoy 2000, p. 549). This implies education has some aims, and the teacher should promote her trustworthiness in order to achieve the ends of the practice. However, unlike medicine, in which most would agree that the goal of the physician-patient relationship is the health of the patient, in education there are diverse theories regarding the aims of the practice. To name a few, here is a list out of the contents of a collection of essays on the issue of educational aims: Citizenship, Autonomy, Critical Thinking, Self-Determination, Well-Being, National Identity, Fairness and more (Marples 1999). Each of these educational aims usually gives a different set of recommendations to show how the practice of teaching under this aim is to be carried out.

The problem with viewing truthfulness and goodwill as means to achieve the aims of education is that different educational aims will give us a different account of whether the teacher may betray that trust in favour of achieving those aims. If the physician can justify her untruthfulness by appeal to her goodwill regarding the health of the patient, the lack of general agreement in regard to the ultimate goal of education and the means required to reach this goal makes this kind of justification a

127

harder task for educators. It is possible that a certain community will contend that truthfulness may be jeopardized in order to promote what that community believes is the goal of education. For example, in order to promote civil unity, Plato famously proposes to deceive all future citizens of the republic about their origins.[12] However, this will be true only within that specific community which agrees that this goal justifies such means. In most liberal democracies there is no consensus about the aims of education, and this disagreement is seen as part of the nature of a democratic society.

Moreover, appeal to the aims of education is not helpful because it is either too narrow or too wide to provide an answer to the concerns about benign manipulation. Suppose the aim of education is student learning, or growth, or some similar thick though unspecific concept.[13] This is too wide since learning might be 'mathematics' and then manipulation would be appropriate if it promotes learning mathematics, but learning might also be 'self-awareness' and then manipulation may not be appropriate. If we were to narrow down the goal of the educational process, perhaps just to 'mathematics', there will be less agreement on this goal, and the

---

[12] Or, to promote initiation into 'the civilized inheritance of mankind' as Cooper quotes Michael Oakeshott, a teacher might use deceptive means. Thus I do not see how Cooper's argument ensures truthfulness as a central teacher virtue in that tradition. Cooper 2008, p. 86-7.

[13] See Strike, Kenneth A. 1988. 'The Ethics of Teaching', *The Phi Delta Kappan* 70(2), p. 522.

community that would agree that this goal calls for manipulation will be smaller. If there is a variety of aims, then again by trying to appeal to the ends of the practice we would still have different answers to the question regarding teachers' use of manipulation.

I suggest it is worth considering that we may not get to definite answers to the question of the aims of education. However, we might still like to know what the right conduct is in regard to teachers' use of manipulation. If this is the case, viewing trustworthiness as a means to reach educational goals does not help in solving the issue. Nonetheless, realizing that a certain degree of trustworthiness is important in reaching any educational goal should be a clue that trustworthiness is more fundamental in education than other educational means.

To conclude, if students should trust teachers for the right reasons. i.e. because the teachers are trustworthy (goodwill, truthfulness, competence), then trustworthiness should not be seen as a method of achieving educational aims. As an analogy, consider the trust that Ulysses places in his crew when he asks them to tie him to the mast so that he will be able to listen to the Sirens. He makes himself vulnerable to the crew, but he has good reasons to think that they will not exploit the situation and start a mutiny, perhaps because their relationship has been strengthened by shared adventures and misfortunes. Also, he appears to trust them

129

to the right degree, at least in the sense that the episode is time-limited and meant to serve his own interests. He might be wrong, either by trusting the crew for the wrong reasons, such as if they only act as if they care for him, when they do not. Or else, he might be trusting them to the wrong degree, for instance by not limiting the time he is tied to the mast. Similarly, students should not be encouraged to trust teachers for the wrong reasons or to the wrong degree. Those who view trust as means to achieve other educational aims risk having students who place too much trust in their teachers or trust their teachers for the wrong reasons.

Truthfulness, goodwill and competence as part of being a teacher

My suggestion is to view trustworthiness first as a precondition of any intentional educational process. In other words, when it is carried out as an intentional process, education requires some minimal threshold of trust, in its truthfulness, goodwill and competence aspects. Teaching is based on and legitimized by the trustworthiness of the teacher. Because she is trustworthy,

130

parents, school management and the state entrust the students to her care and the students are willing to put their confidence in her, to learn with and from her.[14]

Therefore, some threshold level of trustworthiness is part of being a teacher. A student who does not believe the teacher is truthful (in general or in a specific area) or does not have faith in the good intentions of the teacher, would find it very hard to learn from that teacher what the teacher intends to teach. Hence, the three aspects of trust are not only important to good teaching, they are what makes teaching possible in the first place. Truthfulness, goodwill and professional competence are common day practices, role derived and institutionalized, that are part of student-teacher relations in schools. Teachers have a commitment to keep at least the threshold level of them since these are preconditions to any teaching. While we might want to maximize teacher trustworthiness in order to achieve educational aims, we should not forget that being a teacher is already being trustworthy at least to some minimal degree. In cases where this minimal threshold does not exist, there is no teacher-student relationship.

---

[14] Williams remarks about primitive trust of the learner of a mother-tongue language as primitively trusting the assertions of surrounding members of the family. Similarly, I maintain that a certain level of trust is a requirement of any intentional education.

131

It is hard to pin down the exact threshold of trustworthiness that legitimizes and makes education possible. However, even a narrow formulation of the commitments of trustworthiness can help in addressing the issue of teacher deceptions. In the truthfulness aspect, a minimal understanding of this commitment is not to deceive without good cause, in the goodwill aspect it is not to intentionally harm or allow harm to happen to the students, and professional competence is at least to know how, and be able to, uphold the minimal requirements of the first two.[15] This means that although benign manipulation should be understood as a non-moralised concept which does not require justification, within an educational process, teachers do have commitments of goodwill, truthfulness and competence, and therefore manipulations that breach the threshold need to be justified.

Truthfulness, goodwill and competence as virtues of the ideal teacher

Trustworthiness is more than the above minimal commitments. We have seen that in order to achieve educational goals, a higher level of trustworthiness is

---

[15] Teachers' professional competence probably includes other components, such as a basic understanding of the relevant subject knowledge, an understanding of the institutional requirements of the job and more. That said, even though I will not attempt to support this claim here, I believe the commitments of truthfulness and goodwill, understood in their minimal form, take precedence over other commitments that form teachers' professional competence,

132

probably required, and so teachers might wish not only to keep to the minimal

commitments of trustworthiness but also to promote the degree of their

trustworthiness. Perhaps trustworthiness should be seen as even more than that - as

a virtue of the ideal teacher or a characteristic of ideal education. As an ideal it gives

teachers the incentive to promote trust to the highest degree possible. However, it is

not so easy given the multi-faceted nature of trustworthiness.

Cooper, who follows Bernard Williams' account in 'Truth and Truthfulness',

views trustworthiness as calling on the teacher to sincerely convey accurate

information in regard to her subject matter (Cooper 2008, pp. 79-87). By doing so,

the teacher exemplifies the virtues of truthfulness - sincerity and accuracy. Under

this view truthfulness is part of the ideal education or a virtue of the ideal teacher.

However, as discussed above, truthfulness is only one aspect of teachers'

trustworthiness. In some cases, if a teacher tells the truth, the whole truth and

nothing but the truth, she might end up harming the relationship with the student,

since good relations require at least some withholding of information (which we like

to call politeness or tact).[16] Therefore, truthfulness should not be seen as an ideal we

---

[16] "Communicative desert calls for a fine-tuned sensitivity to situations where the bald truth
would defeat the ends of practice." Pendlebury, S. 2008, 'Accuracy, Sincerity and Capabilities in the
Practice of Teaching', *Studies in Philosophy and Education* 27(2–3), pp. 179-180. I agree on the need
to be sensitive to context, but think the guide to the right conduct should not be the 'ends of the
practice' but the preconditions of trust that underlie teaching as a practice.

should inspire to fully implement in all elements of teacher-student relations. Moreover, sometimes it makes sense to communicate inaccurate, false and simplified information about the subject matter, for various reasons, such as making it easier for the students to learn something, or censoring information that might be too disturbing. Just like in a poker game, some deceptions are expected and required, while others do constitute a breach of trust.

In a similar fashion, if we consider goodwill to be a virtue in the ideal teacher, this might give reason for the teacher to act in an undesired way just because she believes she is doing it out of goodwill towards the student, or because the student would perceive her actions as caring. Nel Noddings warns about this kind of understanding of care as a virtue of the individual: "the good intentions captured in the slogan 'all children can learn' can lead to highly manipulative and dictatorial methods that disregard the interests and purposes of students" (Noddings 1992, p. 19). Goodwill as a virtue is not supposed to override truthfulness on such occasions. Therefore, both truthfulness and goodwill are not fully desired each by itself if we wish to keep our commitments as educators. Moreover, as mentioned above, we would not want a too high degree of trust, which can lead to abuses of trust by teachers.

134

We might trust someone to a certain degree in some matters, trust her less in other matters and trust others more or less in different levels. Thus, one can trust another in the goodwill aspect of trust but not in the truthfulness aspect (like a child might trust an always supportive father) and vice versa (like one might trust a very harsh rival). For teachers it does make sense to want a high level of both. What is required from teachers is a prudential balance between truthfulness and goodwill, which considers the specifics of the situation and the commitments they have as teachers. I suggest to think of the ability of teachers to achieve this balance as part of their professional competence.

Unlike the first two criteria of the framework for assessing the use of benign manipulation (consequences and understanding), the third criterion deals with the norms that govern the relationship in which manipulation takes place. In this section, I gave a brief explanation of what should be the norms that govern teacher-student relationships, and how we need to understand the meaning of these norms from the teacher's perspective. In brief, teachers should be caring (have goodwill), truthful and competent. I've suggested that these teacher attributes mean that teachers have at least some minimal commitments to be caring, truthful and competent. In later chapters, I will use the analysis of teacher's trustworthiness as part of the assessment of benign manipulation in education.

135

3.5 Personal autonomy

I move on to discuss manipulation's infringement on personal autonomy, which is often mentioned as a possible wrong of manipulation. It is my view that what theorists find problematic about manipulation's harm to personal autonomy can be captured by a different conceptual scheme, specifically the three criteria discussed above (consequences, understanding and trust). However, because autonomy is commonly raised as a harm of manipulation, I look into this issue in some detail to show why it should not be part of the framework for assessing the use of benign manipulation.

Various theorists claim that even if manipulation is sometimes justified all things considered, it always involves an infringement of the manipulee's personal autonomy. Indeed, many accounts of the moral wrong involved in manipulation maintain that manipulation violates, jeopardises or is opposed to the manipulee's personal autonomy. As we have seen in the chapter regarding the definition of manipulation, it is a form of influence which, unlike ideal reasoning, is not obviously autonomy preserving. Hence, it appears there are good reasons to think that manipulation jeopardises personal autonomy (Noggle 2018). Baron, for example, claims that: "Manipulation generally involves a recklessness, a disregard for the

other qua agent (though not necessarily a disregard for the other's welfare), a determination to bring about a particular result and a willingness to be very pushy or somewhat deceptive to reach that result" (Baron 2014, p. 104). Though plausible at first glance, the extent to which autonomy is jeopardized is not a useful measure to assess the use of benign manipulation. My claim is based on two lines of argument. One, that autonomy is a vague and convoluted concept and so limited in its usefulness for practical considerations. Two, that if we understand autonomy as 'self-government', manipulation is often compatible with preserving and even promoting the manipulee's autonomy. Therefore, it is does not add a useful measure to the criteria established above.

There are many competing accounts of personal autonomy. In broad terms, theorists who discuss manipulation mostly understand autonomy as self-governance: "To be autonomous is, in essence, to do things for one's own reasons – to act according to principles, or laws, which are "'self-imposed' in the sense that one endorses and applies them to oneself" (Buss 2005, p. 199).[17] However, beyond the relative agreement about autonomy being a form of personal self-governance, there is a wide array of different accounts of personal autonomy as well as disagreement

[17] Buss mentions that this is a 'very broad, rather indeterminate, conception of autonomy', but that it is exactly this unclear conception that is often used in political and philosophical discourse. See Buss 2005, note 5, p. 199.

137

regarding the minimal conditions required for personal autonomy. The fact that 'autonomy has acquired almost as many definitions as it has defenders' as Michael Hand puts it (Hand 2006, p. 536), is one reason to be reluctant to use it as a meaningful measure for the assessment of actual actions and techniques.

Even if the suggestion that manipulation violates or undermines personal autonomy appears plausible, there are good reasons to be critical towards it. There are instances where manipulation does not undermine autonomy, and even cases where manipulation might enhance the overall autonomy of the manipulee. Noggle offers the example of a teacher who: "might manipulate a student into taking a course of study which will ultimately enhance her autonomy by opening new career options, improving her skills of critical self-reflection, etc" (Noggle 2018). Another example is manipulating an ex-drug addict not to relapse into abusing drugs again, thus enhancing her personal autonomy. These examples show that on balance, certain manipulative techniques might promote the overall autonomy of the manipulee. However, it may still be the case that this manipulation is morally problematic because it jeopardizes the manipulee's personal autonomy in the short term, even if it promotes her long-term personal autonomy. But even this qualified view appears to be wrong.

138

First, manipulation does not have to involve the complete removal of choice. This is demonstrated by various examples of 'nudges' that change the choice environment in order to encourage certain behaviours, such as in Cafeteria (Sunstein and Thaler 2003, p. 1164). As Cafeteria shows, one might be manipulated into a certain kind of behaviour without it preventing the manipulee from making a choice, and at least in that sense she is autonomous. However, an objector might claim that just being able to choose does not constitute self-governance. What is also required is to be sufficiently free of external influence. Nonetheless, there are many external factors that influence us, such as the weather, the landscape, our bodily needs, etc. and so we are never free of all external influence whatsoever. If anyone is to be considered self-governing, this cannot mean to be free of any external influence. The objector might concede the previous point and still claim that there is a difference between external influence which no one has any control over, and external influence which is intentional and originates from another person. In this case, to self-govern would mean to be free of another person's intentional influence.

However, Buss makes a convincing argument as to why self-governance may include another person's intentional influence. Her argument rests on two key claims. First, that manipulation does not prevent the manipulee from being able to exercise

139

her own will; and second, that autonomous agents would often not object to being influenced by manipulation:

> One person can manipulate and deceive another… without preventing the manipulated person from governing herself. Moreover, when manipulation and deceit are characterized in a morally neutral way, reasonable, self-governing agents do not object to being manipulated or deceived in a wide range of circumstances. (Buss 2005, p. 197)

That the manipulee is not prevented from exercising her will can be seen in the example of Cafeteria, as discussed above. Whichever presentation the cafeteria manager chooses, it does not prevent the customers from making their own choices, i.e. exercising their will. In this case, as in other forms of manipulation, the manipulator offers or allows a choice to the manipulee, the manipulee can exercise her will. This is characteristic of many instances of manipulation; unlike coercion, manipulation often leaves the manipulee with a choice. Therefore, the manipulee's autonomy, understood as 'self-governance' is not being undermined, at least not completely. Of course, sometimes reducing the available options so that the agent is only left with a choice between actions that are desirable to the manipulator does jeopardize the autonomy of that person, and such cases are conceptually closer to

140

coercion. However, we should also be careful not to assume that only actions which are the result of a rational conscious process are autonomous actions. Many of the choices we make are based on a less than full awareness of everything that motivates us to make those choices, yet we would still consider these 'not in full awareness' choices as our choices.

Buss's second claim is that autonomous agents would often not object to being influenced by manipulation. In her view, part of a person's self-government is allowing herself to be or not regretting being moved by methods very different from rational persuasion. The example she focuses on is that of romantic seduction, which regularly involves manipulative techniques such as withholding information, deception, playing on one's feelings, setting the mood, etc. Buss claims that most of us would not want our romantic lives to be devoid of such methods of influence, where the only unobjectionable courting would be a straightforward request to reciprocate affections:

> [W]hether we know the truth about someone's intentions with respect to us is often irrelevant to whether she treats us with respect and, more generally, that, though we certainly do want to be treated with respect, we often do not really care all that much about what others have in mind when they interact with us. Sometimes, we even

141

prefer not to know—to wonder, to doubt, to be kept in suspense—to be overwhelmed, captivated, moved, seduced, played with, played upon—and yes, manipulated—by poetry, drama, music, oratory, a steady gaze, a darting side-glance. We can only pity those human beings (if any there be) whose souls vibrate to nothing but their own acts of well-informed, undistracted self-determination. (Buss 2005, p. 226)

Buss's claim is especially true in educational settings. Few learners would choose a dry, rational and self-determined learning experience over a captivating and inspirational learning experience. We are generally willing to 'hand-over' our autonomy in order to have a much more pleasurable educational experience. Sometimes, this voluntary transfer of control will result in us having overall greater autonomy. Therefore, understood as a non-moralised concept, there are good reasons to think that manipulation can be autonomy preserving, autonomy promoting, and autonomy respecting. Hence, those who claim that manipulation is objectionable because it jeopardizes the autonomy of the manipulee are mistaken.

In sum, if we understand manipulation as a non-moralised concept, the question of whether manipulation harms or jeopardizes the manipulee's personal autonomy is not useful for assessing the appropriate use of benign manipulation.

142

This is for two main reasons; one is practical in nature – autonomy is a vague and convoluted concept and so it is a difficult task to determine whether someone's autonomy has been jeopardised if we have various different ways of understanding what autonomy is. The second reason is more principled, if we understand autonomy as 'self-governing' it appears that manipulation is often compatible with preserving one's ability to govern herself. Moreover, what is believed to be significant about most accounts of manipulation's harm to autonomy can be captured by the other criteria mentioned above. If we have established the intended and unintended consequences of the action, the extent to which the manipulation prevents the manipulee and manipulator from having an accurate understanding of the situation, and whether the manipulation does not breach the norms of the relationship between the manipulee and the manipulator, thinking in terms of personal autonomy does not add to the analysis.

3.6   How do the criteria operate?

By looking at some of the reasons that others have suggested in explaining the wrong making features of manipulation, I have established that when assessing benign manipulation, we should consider the following aspects of the action:

143

A) Consequences: 1) intended consequences – to what extent is the suggested benign manipulation effective in achieving the desired result in the short term, long term and across contexts? 2) unintended consequences – to what extent does the suggested benign manipulation have unintended consequences in the short term, long term and across contexts?

B) Understanding: 1) to what extent does the manipulation prevent the manipulee from having an accurate understanding of the situation? 2) to what extent does the manipulation prevent the manipulator from having an accurate understanding of the situation?

C) Trust: is the manipulation appropriate according to the norms which guide behaviour in the relationship between the manipulator and manipulee?

An important question to answer is how the criteria should operate when assessing the appropriate use of benign manipulation. For example, we might find that an instance of benign manipulation is very effective in achieving the desired result, but that it also severely prevents the manipulee from having an accurate understanding of her situation. In order to come to an all things considered judgement about the appropriateness of an instance of manipulation we will need to

144

consider the various reasons presented in the framework. But, is there an order of importance among these?

It might be argued that the consideration of the effectiveness of the manipulation in achieving the desired result is lexically prior to all other considerations.[18] In other words, that benign manipulation must be effective in order to even become a candidate for use in practice. By analogy to transport, a method of transportation might be considered for use only if it actually enables the person to travel from A to B. But, the answer to the question of effectiveness of benign manipulation is not binary; techniques are likely to be effective to some extent, their effectiveness varies from one instance to another due to various contextual factors, etc. Moreover, techniques might need to be assessed before their effectiveness can be established in hindsight. If we realize that effectiveness is a spectrum, effectiveness should not be understood to have lexical priority over other considerations. However, it appears reasonable to assume that benign manipulation needs to be effective to some minimal degree (in the specific context) and is limited by some extreme unwanted consequences (in the specific context) before we consider the other criteria.

---

[18] An example of the lexical priority of one moral principle over another can be found in John Rawls's *Theory of Justice*, see: Rawls, J. 1972. *A Theory of Justice*. Oxford: Clarendon Press.

145

Perhaps Criterion B: Understanding is lexically prior to the other considerations? In other words, is it the case that we need to establish the extent to which an instance of manipulation prevents the manipulee and manipulator from having an accurate understanding of the situation before we assess the other criteria? Similar to Criterion A: Consequences, it appears that an instance of benign manipulation might affect the manipulee's and manipulator's understanding of the situation to various degrees. One instance of manipulation might hinder understanding to a large degree, while another will do so to a lesser degree. This again makes it difficult to argue that Criterion B: Understanding should always be lexically prior to the other criteria in all different contexts. In extreme cases, where the reality of the situation is very important for the manipulee or manipulator to know, and manipulation prevents either or both from having an accurate understanding of that reality, this criterion probably trumps other considerations. However, in most instances it should be assessed on balance with the other considerations.

Somewhat differently, Criterion C: Trust should take prominence in the assessment of benign manipulation. In some relationships the norms of behaviour and the resulting expectations of the members have a different moral weight than in others. For example, poker players will treat some acts of honesty with contempt (but do expect competitors not to cheat), while friends will value each other's honesty

146

much more. Therefore, in order to assess a specific instance of manipulation we will need to first consider the nature of the relationship in which the manipulation is taking place. For instance, we might think that the obligations of teachers to their students are different than the obligations of companies to their clients, or the state to its citizens. For this reason, we should begin the assessment of benign manipulation by thinking about the nature of the relationship between the manipulator and manipulee (Blumenthal-Barby 2014, pp. 133-4). In Chapter Five I discuss in depth the trust inherent to the teacher-student relationship. An articulation of the nature of that relationship is necessary if we want to consider what a breach of trust in such a relationship means.

3.7   Clarifying the ethical features of the assessment framework

I would like to clarify some ethical features of the assessment framework. To begin, is the framework consequentialist, non-consequentialist or a combination of the two? At least in parts, the assessment framework is certainly focused on the outcomes of the manipulation being performed. In Criterion A: Consequences, the focus is on the benefits and drawbacks the manipulation might have for the manipulee. Criteria B: Understanding and C: Trust, can also be considered

147

consequentialist, in so far as they consider the outcomes the manipulation might have on the manipulee's and manipulator's understanding of their situation, and the quality of their trust relationship. Therefore, the assessment framework certainly has consequentialist aspects. However, is it wholly consequentialist?

In Chapters One and Two I maintained that the concept under consideration in this thesis is 'benign manipulation'. It is important to clarify that 'benign' here means 'well intentioned' manipulation, rather than a manipulation that was judged to be appropriate and desirable after it was considered using the assessment framework. In this sense, the assessment framework comes into play only when the manipulator has at least some good reasons to claim that the manipulation serves the interests of the manipulee. In this sense, the framework is non-consequentialist – as only 'well-intentioned' acts of manipulation enter the mostly consequentialist assessment.

To explain the former point, consider an analogous case in medicine. Only a drug which is reasonably considered to bring benefits to patients, may be administered to a patient. Then, the physician may still want to assess the side effects of the drug alongside the positive outcomes of it. In this, only a well-intentioned medical intervention may enter the consequentialist calculation of negative and positive outcomes. This is different from assessing whether a certain

action is overall right or wrong 'from the start' solely on consequentialist grounds, such as in a 'ticking bomb' scenario where a consequentialist view would justify torturing a person who can reveal the location of a bomb that might harm many others. In the 'ticking bomb' example the torturer does not have good intentions towards the terrorist. However, in the medical intervention example, the consequentialist calculation comes into play only in cases where the physician has good intentions towards the patient. The concept of benign manipulation under discussion in this work, similarly assumes good-intentions on the part of the influencer. Only interventions which can reasonably claimed to be in the interests of the manipulee, should enter the assessment framework.

In this context, a further question arises: if Criteria B: Understanding and C: Trust are also focused on the outcomes of the manipulation, why not think of them as sub-categories of criterion A: Consequences? The answer is practical rather than principled. I find the consequences of the effect of manipulation on the manipulee's and manipulator's understanding of the situation a sufficiently different class of consequences to be considered separately. For one, consequences that fall under Criterion A: Consequences, such as learning outcomes, personal well-being, behaviour, health, etc. are somewhat easier to measure and, though still challenging, it appears simpler to weigh and balance them against each other. However, Criterion

149

B: Understanding, focuses on the epistemic consequences of manipulation, which are distinct from other outcomes to an extent that makes it harder to weigh and balance them against other consequences. To take an example from medicine again, placebo treatment demonstrates the difference between considerations of outcomes that are easier to measure and weigh, such as health benefits, as compared to harms to the patient's understanding. I believe the debate regarding the ethics of placebo treatment demonstrates that the conflict between consequences and understanding in education is of a different kind, than the conflict between health benefits and risks. That said, the distinction between Criterion A: Consequences and Criterion B: Understanding is mostly meant to assist practitioners in seeing features of benign manipulation that might be harder to consider if these were placed under the same heading. In that, the distinction is supposed to serve a practical purpose, rather than a principled one. Something similar can be said about Criterion C: Trust. First, this criterion takes some priority in considering consequences, as some consequences will not come into consideration in certain relationships. As an example, two poker players who attempt to deceive one another, will not consider it a betrayal if they are caught bluffing. In such a relationship the outcomes that accompany betrayal will not be part of the consequences. In this sense, Criterion C: Trust, has an element which is non-consequentialist, and is related to the commitments and moral obligations of the parties involved. I discuss these

150

commitments in more detail above. That said, the effects of the manipulation on the degree and quality of the trust relationship may indeed be considered under Criterion A: Consequences. However, and similar to Criterion B: Understanding, I find the effects on trust to be sufficiently different to merit a distinct and separate aspect of the assessment framework, if not solely on principled reasons, at least on practical ones – to make it easier for practitioners to see and consider the various important features of benign manipulation.

3.8   A word about meta-ethics

I have rejected the idea of manipulation as a moralised concept, yet I draw on moral intuitions and moral language to defend the criteria and how they can be applied. The easy way to reconcile this issue, is to refer again to the fact that the actions under discussion are 'benign manipulations', that is, manipulations in which the manipulator believes they act in the interests of the manipulee. Therefore, I assume that the manipulator holds substantive ethical values which are shared among most members of the relevant society. For example, that the manipulator believes that one should not harm others unless that harm is justified. In this sense, the criteria are meant to assist the manipulator in seeing different features of the

151

manipulative action which might be less visible if one were to focus on just one feature of the action, such as its effectiveness.

The more challenging answer to the above concern, is that in the process of thinking about these issues, I have come to realize that strict moral principles are limited in being a useful guide to educational practitioners. To take a simple example, if someone holds a moral principle which obliges them not to lie on any occasion, they will struggle to maintain and nourish educational relationships. It appears that a deontological approach to moral judgements is limited when it comes to educational settings. Similarly, a consequentialist or utilitarian approach, risks falling into problems, as at times the most effective measures are morally problematic, such as corporal punishment. In other words, as was also discussed above in the introduction, the context in which these judgements take place is very significant when it comes to educational relationships and processes. Deontological and consequentialist theories, both based on general principles, fail in not being sensitive enough to the situation at hand.

Moral particularism is an alternative to basing moral judgements on general principles. In its most extreme version, moral particularism is the claim that "there are no defensible moral principles, that moral thought does not consist in the application of moral principles to cases, and that the morally perfect person should not be

conceived as the person of principle" (Dancy 2017). There are more cautious versions of particularism which allow for some moral principles, but similarly downplay their importance in moral judgement in comparison with generalist theories (McNaughton and Rawling 2000). The important issue for me here is that moral particularism is sensitive to context and the specific reasons that are relevant to the case being considered. As an example, a lie would not be considered morally right or wrong before the situation is carefully assessed to consider various reasons that would make it right or wrong. In this, it is similar to how I understand manipulation as a non-moralised concept. To understand lies, or manipulation, as a non-moralised concept is not to claim that they do not have moral significance, but rather that their moral significance is dependent on the specific situation in which they occur.

If this is the case, one might think that to establish a framework for assessing manipulation is misguided. However, the moral particularist can use the experience of past cases in honing their moral sensitivity:

> …experience of similar cases can tell us what sort of thing to look out for, and the sort of relevance that a certain feature can have; in this way our judgement in a new case can be informed, though it is not forced or constrained, by our experience of similar cases in the past. There is no need to suppose that the way in which this works is

by the extraction of principles from the earlier cases, which we then impose on the new case. (Dancy 2017)

Therefore, the framework suggested above can be used to inform our moral sensitivity, and shed light on the moral relevance certain features of manipulation have.

It is beyond the scope of this work to fully assess the merits and flaws of moral particularism as a general meta-ethical theory. However, at least within the field of educational practice, this appears to be a compelling approach which seems to agree with Dewey's remarks about the science of education discussed above, as well as my own personal experience and analysis of case studies which will be presented below.[19] Therefore, I suggest that a possible understanding of the criteria for the assessment of benign manipulation is that they should be taken as an attempt to 'tell us what sort of thing to look out for, and the sort of relevance that a certain feature can have' rather than a recipe of general principles to be followed.

The discussion above also offers an explanation of whether the consequentialist aspects of the assessment framework emphasize rules (and

---

[19] For further development and exploration of the possible links and disagreements between Pragmatism and Particularism see Bakhurst, D. 2007. Pragmatism and Ethical Particularism, in New Pragmatists, Cheryl Misak (ed.), Oxford University Press, pp. 122-141.

154

classes of cases) or individual acts of manipulation. As I find the specific context in which each individual act of benign manipulation takes place to be of importance in the assessment of it, in so far as the framework offers rules, it is only as a guide to the significance a feature of the action might have. As mentioned above, the framework should be understood as a guide to assist professional judgement in each individual act, rather than a recipe to be followed.

155

Chapter 4   Manipulating Children vs. Manipulating Adults

4.1   Introduction

My main aim in this work is to provide educators, and especially those who work with children in educational institutions, a conceptual map to assist their decision making in regard to benign manipulation. Up until this point, I have not specifically focused my attention on manipulating children. The discussion above, both of the nature of manipulation and the framework by which to assess benign manipulation is, on the face of it, applicable to both children and adults. However, some theories of manipulation claim that there is a difference between manipulating children and manipulating adults. The following chapter will assess whether there is such a difference and how we should understand that difference when we wish to assess the use of benign manipulation in education.

According to Alan Wood, manipulation is less problematic when it is directed at children. He writes that: "Manipulation may be acceptable, or even admirable, when it is used to train the child in good habits, or encourage emotional responses

156

that are conductive to the acquisition of the rational and emotional capacities of a healthy adult." (Wood 2014, p. 36) Marcia Baron similarly claims: "One feature that sometimes makes what would otherwise be manipulation not be manipulation is that the person to whom our efforts are directed is a child" (Baron 2003, p. 45). Alan Ware notes that "practices that might be claimed as manipulative if experienced by adults have traditionally been a normal aspect of the raising of children" (Ware 1981, p. 164). These scholars assume that manipulation is a moralised concept. In their view, manipulating children is either not manipulation, so it is not problematic when children are the target of the 'manipulation-like' actions. Or, if these actions are an instance of manipulation, they are justified due to the fact that they are directed at children.

Often, the differential treatment of children and adults is assumed to follow from the difference in the degrees of personal autonomy adults and children possess. Because manipulation harms autonomy, so the argument goes, it cannot harm those who are not autonomous, and therefore it is not objectionable when directed at children. In other words, if manipulation jeopardises autonomy, and children are not autonomous, that would make manipulating children more acceptable than manipulating adults. I have argued in the previous chapters that whether we assume manipulation is a non-moralised concept or not, it is not the

157

case that manipulation necessarily harms personal autonomy. In this chapter I will

use that analysis to show that the supposed harm to autonomy cannot serve as a

justification for the differential treatment of children and adults when it comes to

manipulating them. I will then consider other conceptualizations of the difference

between children and adults that offer better explanations of their differential

treatment in regard to benign manipulation.

4.2 Marcia Baron's view

Marcia Baron (2003) expresses the view that children and adults should be

treated differently in regard to manipulating them. Her overall outlook is that

manipulation is a moralised concept and she offers an explanation of

manipulativeness as a vice, within a virtue ethics framework.[20] It is worth looking into

her discussion of the issue in detail, to better understand the position which sees a

difference between manipulating children and manipulating adults:

---

[20]Baron's view has since changed and in her more recent work she supports a non-moralised conceptualization of manipulation, see: Baron, Marcia. 2014. 'The Mens Rea and Moral Status of Manipulation', in Manipulation: Theory and Practice, C. Coons, & Weber, M. (eds.), Oxford University Press, pp. 98-120.

One feature that sometimes makes what would otherwise be manipulation not be manipulation is that the person to whom our efforts are directed is a child.

…The following ends all can justify doing to a child what we would usually call manipulation were it done to a fully-functioning adult: the child's long-term well-being; the child's immediate happiness; others' comfort or sanity; the future good of the world. Still, there are limits to the techniques that may be used. Filling them with terror in order to get them to behave is almost always beyond the pale, and lying to them, if not beyond the pale, is certainly more dubious than exploiting their sympathy, or withholding some information, or prodding them to view the matter in a different light (even if we know that their way of viewing it is no less accurate than the way we are proposing). (Baron 2003, pp. 45-46)

In the quote above Baron claims that some of the tactics used to influence children, which would be considered manipulative if exercised on adults, should not be considered manipulative when these are done to children. However, it is not clear what explains such differential treatment of children and adults. To illustrate, consider an instance where the same technique is used towards both children and

159

adults. For example, Jane's child and partner both like sweets and eat them in unhealthy amounts when there are sweets in the house. When she receives a bag of sweets as a present from her workplace, she does not tell them about it. Assuming this is a case of withholding information from both, according to Baron her act would be characterized as manipulative towards her partner but not manipulative towards her child. This seems odd, so let us look into the possible explanation that Baron's view might provide for this difference.

According to Baron's account there are several reasons that would make manipulating children not manipulation (or a less objectionable act of manipulation). First, she claims that the ends of the action might justify it. Therefore, if the consequences of the action are likely to bring about a significantly better state for either the manipulated or others, that would justify the action. But, why should this be any different in regard to adults? In the example above, if eating sweets would be bad for the adult partner, then manipulating him to prevent those consequences should also be justified. Indeed, the good consequences mentioned by Baron involved in manipulating children (the child's long-term well-being, the child's immediate happiness, others' comfort or sanity, and the future good of the world) apply to manipulating adults as well. Therefore, a consequentialist approach does

160

not explain well why children and adults should be treated differently when it comes to manipulating them.

However, one might argue that as part of the calculus of consequences, a child is different from an adult in two important ways: they are likely to live longer and so the consequences might be long-lasting, and they are perhaps more likely to develop habits in childhood (e.g. eating sweets) that will affect their behaviour as adults. None the less, these differences are differences in degree and not in quality. What matters in the calculus of consequences here is the long-term impact and the effectiveness (habit formation) of the intervention, and not the fact that these are direct at children. In cases where the impact will be long-lasting and effective, this should be considered, regardless whether the intervention is directed at a child or an adult. I shall elaborate further below, but for now it should be enough to see that a consequentialist approach at best offers a general 'on average' difference in degrees between manipulating children and manipulating adults, and so it is not a good explanation of why manipulating children is categorically different than manipulating adults.

Second, Baron appears to assume that children are not 'fully functioning' adults. What 'fully functioning' means in this context is somewhat obscure. I shall assume that by the phrase 'fully functioning' Baron means something like

161

'autonomous', which is in itself a convoluted concept. For my purposes here, it is enough to remain with the wide understanding of autonomy as 'self-governance' which I believe captures what Baron means. Supposedly, children are not 'self-governing' because they lack self-control, are less capable and less likely to listen to reason, and possibly other mental attributes related to a lack of self-governance. If this is the case, then using manipulative techniques to influence children, in cases where they lack autonomy, would be justified for these reasons: firstly, there might not be another way to influence them, so manipulation would be the only or significantly most effective way of influencing them. And secondly, because children are not fully autonomous, if the harm involved in manipulation is not respecting one's autonomy, then that harm is irrelevant in cases where the manipulee is not autonomous.

To illustrate this view, in the case of Jane, manipulating her adult partner would be wrong because she has other, comparatively effective, ways of influencing him, such as reasoning with him. In contrast, these methods of influence are not open to her in attempting to influence her child. Therefore, manipulating children is justified because other forms of influence are not available. Secondly, because her partner is an adult, Jane needs to respect his autonomy and let him make the choice of whether to eat the sweets or not. However, since her child lacks the capacity

162

necessary for self-governance, i.e. her child is not autonomous, manipulating her child is less objectionable because Jane does not need to respect the child's autonomy. Put simply, manipulating her child will not harm his non-existent autonomy. These two lines of argument seem plausible, but I believe they are wrong for reasons that will be discussed below.

The first argument, that manipulating children might sometimes be the only or most efficient way of influencing them and therefore it is right to do so, is a restating of the consequentialist view. Without discussing the merits or flaws of such an approach, my concern here is that it does not provide us with a good measure of differentiating between children and adults. At times, manipulation could be the only or most effective method of influencing adults. If manipulating adults is sometimes objectionable, even in cases where it is the only or most efficient way of influencing their behaviour, then the same would apply to children. Therefore, we cannot differentiate categorically between manipulating children and adults on the basis of manipulation being more effective when exercised over children.

However, perhaps there are on average far more instances where it is significantly more efficient to manipulate children compared to manipulating adults, maybe due to their innocence and not yet fully developed cognitive abilities. But this surely cannot serve as the sole justification for actually doing so. It is also much

163

easier and possibly efficient to use physical punishment on children as part of their education, but this is not reason enough to do so. It is a clear case of the naturalistic fallacy - deriving what ought to be done from what is the state of affairs. In the multi-criteria framework for the assessment of manipulation that I have suggested in the previous chapter, consequences are part of the assessment framework, but not the only relevant consideration. Though it is reasonable to assume that manipulating children might sometimes be easier, i.e. more effective, it is something to take account of, but not as an overarching reason to differentiate between manipulating children and adults.

The second argument in favour of differentiating between children and adults in regard to manipulating them, assumes that manipulation harms autonomy. Supposedly, because children are not fully autonomous, they are not harmed by manipulation like adults are. Though plausible at first, I think this argument is wrong for various reasons. First, as discussed earlier, manipulating someone does not necessarily harm their autonomy (Buss 2005). Hence autonomy, or the lack of it, cannot serve as the basis for the differential treatment of children and adults. Furthermore, as 'autonomy' is a convoluted concept, one's theory and understanding of autonomy will determine whether children are regarded as autonomous, and to

164

what extent they 'govern themselves'. At least according to some theorists, children

should be understood to have 'bounded autonomy' in some aspects of their lives:

> On my view even young children have some arenas of action
>
> where they have autonomy. Any such arena must be one where a
>
> child has volitional stability, understands the relevant options and their
>
> consequences, and is able to exercise self-control in support of his or
>
> her goals. (Mullin 2014, p. 421)

If this view of children's autonomy is correct, then manipulating children cannot be

dismissed as unproblematic on the basis that it does not harm their non-existent

autonomy.

Other thinkers, who hold a different understanding of autonomy, might claim

that children are not autonomous. However, these views would struggle to explain

the fact that many actions performed by adults may not rise to the 'autonomous'

threshold. For example, an adult might choose food for lunch, without their choice

being directed towards long-term goals or interests, or without it being in accordance

with principles, or laws, which are self-imposed. They might pick the greasy chips

option, without giving it much thought. If we hold autonomy based on the rational

capacities and a stable self, then in this case, their action is not autonomous. Would

it be appropriate to manipulate adults in these cases? The literature about

165

'Libertarian Paternalism' builds on this idea, claiming that as long as the person is left with a choice, attempts to influence her do not amount to jeopardising her autonomy (Sunstein and Thaler 2003). An adult that would pick one option over the other, because it is healthy, or because it is their favourite food, or for whatever other reason, is still respected as an agent. This is a complex issue much discussed in the literature.[21] For my purposes here, it suffices to note that on the 'self-governance' view of autonomy adults are not autonomous by default, it is only at times that they exhibit, and perhaps even care about, being fully autonomous. None the less, we tend to think that they should be respected as agents, even when not demonstrating coconscious self-governance in their actions. Therefore, it makes more sense to argue that for both children and adults, it is problematic to manipulate them in so far, and to the extent, and in the aspects in which they are, or wish to be, autonomous. This again does not amount to categorically differentiating between children and adults.

Another reason that undermines the differential treatment of children due to their supposed lack of self-governance, might be titled the pedagogical reason. As

---

[21] For the case in favor of 'Libertarian Paternalism' see: Sunstein, Cass R. 2014. *Why Nudge?: The Politics of Libertarian Paternalism*. Yale Univeristy Press. A critical assessment of this view can be found in Bovens, L. 2009. 'The Ethics of Nudge', in Preference Change, Springer, pp. 207-19.

166

Fischer and Ravizza rightly explain, treating children as responsible adults has the added value of teaching the child what it means to be a responsible adult:

> Even before children are fully responsible for their actions, we often find ourselves taking certain attitudes towards them that are in many respects similar to the full-blown attitudes of indignation and resentment (which are of course only appropriately applicable to morally responsible agents). For example, a young boy, overcome by excitement, tears open the presents belonging to the birthday girl, despite having been instructed in the proper etiquette. We might well correct him and show the customary signs of indignation, even though we are well aware that the child is not yet fully responsible for his antisocial behaviour. By adopting certain attitudes towards the child (and expressing them suitably) – by acting *as if* the child were a fully developed moral person – we begin to teach the child what it means to be such a person. Of course, this sort of training, with its characteristic set of parental attitudes and responses, is a central feature of the moral education of children. (Fischer and Ravizza 1999, p. 208)

167

Therefore, even in cases where the child demonstrates a lack of self-governance, we might still want to treat her as if she is a self-governing agent, thus assisting her to become one. If we think that manipulating a child is unproblematic because she is not self-governing, we will miss out on opportunities to advance the development of self-governance in that child. This is true for adults as well as children. When we influence an adult by benign manipulation, we might also prevent that adult from developing self-government in that aspect. In Cafeteria, if we choose to change someone's diet by carefully placing the healthy food earlier in the line, we are not changing their general preference which will be apparent in other environments. However, if we were to reason with that person, it might improve their self-governance in relation to their diet. Again, it appears that the pedagogical reason is applicable to both children and adults.

To conclude, I have listed above four reasons to be critical towards the idea that children should be treated differently than adults due to their supposed lack of autonomy. In brief, manipulation does not necessarily harm autonomy even when it is performed on adults; there are theories of autonomy claiming that children are self-governing at least to some degree; there is a good reason to think that adults are not fully self-governing; and we have a good reason to treat both children and adults as if they are self-governing, even when they are not.

168

So far, I have attempted to describe the puzzle – why the question of the differential treatment of children when it comes to manipulating them cannot be resolved by using the concept of autonomy. To exemplify, consider a case where a parent makes herself appear angry in order to influence the behaviour of her child, for the benefit of the child. And, a case where a child makes herself appear upset in order to influence the behaviour of her parent, for the benefit of the parent. If what I have claimed above is correct, we cannot judge the parent's action as less objectionable just because it is directed at a child. Likewise, we cannot judge the child's action as objectionable because it is directed at an adult. In what follows next, I would like to suggest that in order to understand when and how children should be treated differently from adults in regard to benign manipulation, we need to consider conceptualizing the difference between children and adults not in terms of their 'autonomy'.

4.3   Conceptualizing Childhood

Intuitively, it does appear to be the case that manipulating adults is different from manipulating children. If the concept of autonomy is not productive in explaining this difference, is there another reason to explain if and when children and adults should be treated differently in relation to benign manipulation? In order to answer

169

this question, we need to understand better what we mean by the concepts of 'child' and 'adult'. Evidently, we have different reasonable ways of conceptualizing the difference between children and adults, and these different conceptualizations are valuable for different reasons. In my view there is no single true conceptualization of children, and some conceptualizations of the difference would be useful for different things. In this, I am assuming a pragmatic approach to the conceptualization of childhood, in-line with the thinking of William James, which sees theories as 'instruments, not answers to enigmas' (James 1907, p. 53).

Any conceptual scheme requires us to make certain simplifications in the aim of greater clarity. This is especially true in the case of ordinary language terms, such as 'child' and 'adult'. Depending on the difference between children and adults that we wish to explain, we may conceptualize children according to different features of their lives. For example, in order to set the minimal age for riding in a car without a child seat, we would probably be most interested in children's physical features. More complex issues might be the legal responsibility age, after which children may be treated as responsible for crimes they commit. In this case we seem to conceptualize children according to a variety of features, such as cognitive ability, actual understanding of the world, and an ability to demonstrate control over their actions. Another issue might be the age of consent, which takes account of some

170

physical features, such as puberty, and some mental development features, similar to the case of the age of criminal responsibility. Other examples of differential treatment might be compulsory education, the right to vote, alcohol drinking age, etc. which will be explained by different features of the complex reality which is children's lives and development.

A useful analogy here is that of maps, which was mentioned before. Any map is a simplified version of the infinite features of a certain area. The designed layout of the map will depend on the specific need the map is meant to answer. For example, for navigating with a car, some maps would highlight roads and remove trees and green areas. Others, perhaps designed for tourists, will highlight the main attractions in an area. The London Tube Map, as another example, distorts distances and geographic locations in order to make it easier to navigate the underground train system. Clearly, there is no 'true' map that is somehow a perfect or a correct representation of reality. The same reasoning should be applied to conceptualizing children and childhood. Sometimes it makes sense to think of children in terms of their size, sometimes in terms of their cognitive abilities, sometimes in terms of their ability to suspend (or not) their reactions, sometimes in terms of their understanding of the world around them, etc. For my purposes here, the question we need to

171

answer is which features of children's lives are relevant in thinking about their differential treatment when it comes to manipulating them.

Physical features

In some cases, it makes sense to conceptualize the difference between children and adults in terms of their physical features. Differential treatment of children and adults, when regulating car seat safety, for example, would require us to treat children and adults differently based on their sizes. The differential treatment of children and adults in this case, for example that children under a certain age require a special car seat, is explained and justified due to children being smaller in size than adults. Age, in this case, would then be a simple proxy to use for physical size.

Of course, conceptualizing children in terms of their physical size does not give us a perfect representation of reality; some children will share the physical features of some adults. There would be instances where the child's size is sufficient to merit equal treatment, but age is a useful proxy even if there are some cases which it does not capture. For example, an adult with the physical properties of a child, is likely to find a suitable bicycle in the 'Kids' section of the store, because she shares the relevant physical features with children. However, the fact that there is some overlap, does not mean that it is wrong to conceptualize the difference between children and adults in terms of the different physical features of these

172

groups. For many needs, it seems like a perfectly adequate way to conceptualize the difference between children and adults.

Conceptualizing the difference between children and adults in terms of their physical features, offers a good explanation of why these two groups should be treated differently in certain aspects. For example, because children are generally smaller in size than adults, it justifies requiring them to use special car seats. However, although differentiating between children and adults based on their physical features makes sense sometimes, it is clear that this way of differentiating between them does not help us in explaining or justifying the differential treatment of children and adults when it comes to manipulating them.

Autonomy

As was discussed earlier, another conceptual scheme that is used to explain the differential treatment of children and adults is based on the concept of autonomy. Because children are not fully autonomous, so the argument goes, it is justifiable to treat them differently than we would adults. There are various and competing accounts of autonomy in the philosophical literature about children (and adults). In itself, this is a reason to be somewhat suspicious about conceptualizing the difference between children and adults in terms of autonomy, if our aim is to use this conceptual scheme as a guide for decision-making regarding benign manipulation.

173

That said, as the concept is often invoked in justifying the differential treatment of children and adults, it is important to dedicate further attention to it.

First, to those who hold a moralized understanding of manipulation, the supposed lack of autonomy of children may serve as a justification of manipulating them, only if the lack of autonomy is a feature of children's inner abilities and capacities. That is, if children's lack of autonomy is mostly due to external factors such that children are not given the opportunities to demonstrate and exhibit their autonomy, then this lack of autonomy cannot serve as a justification for manipulating them. To give an extreme example, a jailed prisoner may lack autonomy due to external conditions, but that cannot serve as a justification to manipulate him, to those who believe manipulation is wrong. If the lack of autonomy is to justify manipulating him, this has to do with the internal aspects of the prisoner. For example, if the prisoner is intoxicated and lacks self-control – then this appears to be the kind of lack of autonomy that would justify manipulating him. However, in the case of children the question of when and to what extent their lack of autonomy depends on external or internal factors is not easily answered. This in itself gives reason to question whether autonomy is a useful concept in differentiating between manipulating children and adults, as it appears difficult to separate the internal and external aspects of autonomy (Frierson 2016).

174

That aside, if we focus our attention on the internal aspects of autonomy, the understanding of autonomy that might explain why children and adults should be treated differently when it comes to manipulating them can be called 'agent-autonomy', a term that was coined by Nomy Arpaly, in her discussion of different varieties of autonomy. According to Arpaly, 'agent-autonomy' "is a relationship between an agent and her motivational states that can be roughly characterized as the agent's ability to decide which of them to follow: it is a type of self-control or self-government that persons usually have and that nonhuman animals do not have" (Arpaly 2003, p. 118).

The question is then, to what extent do children have this ability for self-government? There are two broad families of answers to this question. First, what we might call the thick conception of autonomy requires that autonomous agents' self-government be based on rational faculties, abstract principles, or one's life as a whole (Schapiro 1999; White 1990; Winch 2006). To take one example, in the philosophical literature on children, Tamar Schapiro articulates a detailed Kantian account of autonomy based on the idea that humans are autonomous because, unlike animals, we are able to act on reasons that we give to ourselves. To give reasons to oneself, a person must be capable of critical self-reflection, but this is not enough. Schapiro maintains that in order to act autonomously, our decisions must be

175

driven by a relatively stable self. In the words of Schapiro, the autonomous person speaks in her own voice, "the voice of one who stands in a determinate, authoritative relation to the various motivational forces within her" (Schapiro 1999, p. 729). Autonomous agents, that is, act on principled reasons that express who they are, what they value, and what they want to accomplish, in an integrated manner across a host of domains. According to Schapiro and like-minded theorists who hold a thick concept of autonomy, young children do not possess this ability, and therefore they are not autonomous. If this is the case, the lack of autonomy of children might explain why manipulating them is different than manipulating adults. I will criticize this view below.

The second family of theories of autonomy which can be called the thin concept of autonomy, maintains that the ability of self-government does not have to involve the rational faculties, but rather can be based on 'internal attitudes [as] sources of reasons.' (Bakhurst 2011; Jaworska 2005). For example, Amy Mullin claims that autonomy should be conceived as something we can have in degrees and that it may vary between different situations. She maintains that children should be considered autonomous because they do exhibit self-governance in the sense that at times they behave with sufficient volitional stability that is in contrast with actions that are based on shifting impulses (Mullin 2014). The minimal requirements

for someone to be considered autonomous according to Mullin are the ability to care about some outcomes more than others in a relatively stable way, some self-control, the ability to imagine different courses of action and confidence (Mullin 2014, p. 416). Importantly, these do not require fully developed mental capacities or an integrated self across all life's domains.

An important point of contention between these two competing views is that those who hold the thick conception of autonomy, believe that in order to act autonomously we must inhabit a mental standpoint which is separated, at least momentarily, from our internal impulses as well as from those around us such as family members and loved ones. This solitary perspective is supposedly necessarily so that we can decide whether to endorse the motives that operate on us or not. Mullin, on the other hand, claims that our dependence on our loved ones, and the care and love we feel for them, can be a source of autonomy. This is also true for other things we care about. That is, self-governance consists in acting in accordance with what we care about, "whether these commitments are to persons, relationships, ideals, values, or even things" (Mullin 2007, p. 540). This can be done without critical self-reflection or an integrated stable self, and children do exhibit this kind of self-governance in their behaviour.

177

Back to our main purpose here, of discovering whether autonomy is a useful concept by which to differentiate manipulating children and adults, it should now be clear that the thin concept of autonomy will not do, because according to it, children are, at least sometimes, autonomous. Therefore, if the thin concept of autonomy is correct, then autonomy is not a useful concept by which to differentiate between children and adults. Moreover, if on the other hand the thick concept of autonomy is correct, this sets a high bar for acting autonomously. Even most fully functioning adults do not act at all times as a result of critical self-reflection, according to abstract principles or in accordance with one's life as a whole. Therefore, the thick concept of autonomy might be better understood as an ideal we should aspire towards, rather than a description of the normal activity of everyday life. Because not all adults appear to exhibit thick autonomy all the time, it is questionable whether there is a moral obligation to respect their autonomy when their actions are not autonomous. As an example, if someone chooses the greasy chips option at the cafeteria without self-reflection, there appears to be little reason to respect their choice, as it is not a result of critical self-reflection. If this is true, we cannot use the 'autonomy as critical reflection' to differentiate categorically between children and adults, as sometimes adults would act non-autonomously. In sum, if one accepts the thin concept of autonomy, then there is no clear divide between children and adults in regard to their autonomy. Alternatively, if one accepts the thick concept of autonomy, then it is not

178

useful to differentiate between children and adults because often adults will not act out of critical self-reflection, just like children. That is, in both accounts, autonomy is not a useful concept to explain a different treatment of children and adults when it comes to manipulating them.

To illustrate, due to the considerations above, it makes little sense to differentiate between young children and university students when it comes to manipulating them based on their level of autonomy. Consider the example of deciding what goes into the curriculum of a certain course. The choice of what gets in will certainly have an influence on the students. If we were to agree with the 'autonomy as self-reflection' account that means we would assume young children are not fully autonomous, but university students are. Does that mean a university lecturer would be wrong to craft the syllabus in a way that would be most beneficial to the students? Here is what Baron has to say about pedagogy, in which the teacher might sometimes manipulate her students:

> The role of a teacher may sometimes license rather heavy-handed attempts to shape attitudes. Indeed, some of the careful crafting of the syllabus engaged in by many professors - myself included - reflects an attitude that would in other contexts seem manipulative. I am thinking of such conduct as trying, without telling

179

our students that we are doing so, to dislodge assumptions that, we

find, get in the way of thinking philosophically; starting the term with

readings that will increase the chances that they will take the

arguments seriously, instead of lazily siding with cynics or relativists;

and not mentioning views taken by Great Philosophers that are

peripheral to our aims for the course and might alienate the students.

(Baron 2003, p. 54)

Baron's remarks about manipulation as part of an educational process have little

difference whether the educated are children or adults. Both children and adults

might be manipulated for their own good when it comes to their education and

learning. Partly, this is because of the nature of the relationship between the teacher

and the students, an issue that will be discussed later. However, it seems unintuitive

to claim that the university lecturer would be doing something wrong in carefully

crafting the syllabus to benefit the adult students while the kindergarten teacher is

not doing anything wrong by carefully crafting the syllabus to benefit the younger

children. As a reminder, I am assuming that manipulation is an un-moralised

concept, and so many acts of manipulation, and especially benign manipulation,

would be part of everyday healthy relationships, whatever age the manipulator and

manipulee are. Baron, on the other hand, assumes that manipulation is a moralised

180

concept, and for that reason she is required to explain why it would be unproblematic to use 'manipulation-like' techniques in what we intuitively see as part of common education practice. Her solution is based on the claim that children are not autonomous. This presents a difficulty because the techniques are very similar whether they are directed at children or adults. As can been seen above, if we accept the widespread understanding of autonomy as 'self-governance', contrasting accounts of autonomy are not helpful in categorically differentiating between children and adults in regard to manipulation. Furthermore, what is significant about the view that autonomy can be used to differentiate between children and adults can be captured by other conceptual schemes, which will be discussed next.

Epistemic difference

I move on to discuss two conceptual schemes that in my view are the most useful in understanding why sometimes it makes sense to treat children and adults differently when it comes to manipulating them. Anca Gheaus offers the thought experiment below in her criticism of the theory of parental rights put forth by Brighouse and Swift (Brighouse and Swift 2014). Very briefly, Brighouse and Swift's project is an attempt to justify the family in a liberal society. The problem they face, as liberals, is that the family might not be the best place for the child to be raised, and so being raised in a family might not be in the best interests of the child. If, in a

liberal society, a necessary condition for paternalism to be justified is that is serves the interests of the paternalized, then it is not clear what justifies making the choice for children to be raised in a family. Their answer is based on the idea of parental rights. In their view, adults' interests in being parents limit the child's entitlement to be raised by the best available adult or institution:

> Within certain limits, adults' interests in being a parent can trump children's interests in having the best possible parents. No child has a right to be parented by the adult(s) who would do it best, nor do children as a whole have a right to the way of matching up children and parents that would be best for children overall. Both scenarios could leave perfectly competent parents missing out on the goods of parenting (Brighouse and Swift 2014, p. 95).

In the view of Brighouse and Swift, adults' interests in being parents rest on the uniqueness of the parent-child relationship. According to their view, there are four features that make this relationship unique. First, relationships between parents and children are inherently unequal, due to children's dependency on the adults. Second, parents are responsible for their children's welfare and development to an extent far greater than that which people can take responsibility for other individuals. Third, parents shape their children's beliefs and interests. And lastly, children are capable

182

of loving their parents in an innocent and unconditional way, that is not present in other kinds of loving relationships. Due to these features, the argument goes, adults have a right to parent, as parenting is essential for the flourishing of (most) adults (Gheaus 2015, pp. 195-96).

Anca Gheaus criticises Brighouse and Swift's justification of the family. She bases her criticism on their account of the parent-child relationship. To do so, she describes a thought experiment in which a group of 'primitive' refugees share some characteristics with children so that entering a relationship with one of them would be similar to the parent-child relationship that Brighouse and Swift describe. Gheaus claims that although one might be able to have a unique relationship with one of these refugees, that does not give them the right to have such a relationship. I wish to use her thought experiment to shed light on the role of the lack of knowledge an understanding of children, as different from their lack of self-control:

> Imagine that, as a result of a natural cataclysm, a group of adult refugees reaches your country. They have nowhere else to go. You live under a just and benevolent government that automatically grants the refugees the right to stay and settle down in the country and, in due course, to become citizens. As it happens, the refugees come from a very remote culture, described by anthropologists as

183

'primitive'. They speak a language that nobody has heard of before and nobody understands, and they do not seem able to pronounce simple words in your own language. They cannot read or write, and have never been in contact with any technologically advanced civilization. They do not understand how any of the machines work, and understand complicated social rules even less. They appear scared of traffic and large crowds. Their bodies are beautiful, fragile, relatively small, and unusually agile. They quickly acquire a wonderful reputation for being uncomplicated, trusty, direct, curious, affectionate and playful. For good reason, the belief spreads that having one of these refugees around can bring into your life a kind of joy and fun that nothing else could, and hence that an intimate relationship with one of them would be a special blessing. Moreover, these people are in much need of patient introduction into your own ways of living; somebody has to take over the job of socializing them. And you are right to think that engaging in such an extraordinary task would make a significant and unique contribution to your own personal development. (Gheaus 2015, pp. 198-200)

184

In the scenario above, why would anyone think that the refugees are 'child-like' (Gheaus 2015, p. 198)? It appears that they would count as autonomous adults, in the sense that they exhibit self-control, full cognitive capacities, etc. If they are not self-governing it is not because they do not have the capacity to govern themselves, but because they lack the knowledge and understanding of the world around them. More often than not, they might end up in a situation where they are in some danger because they are not accustomed to their environment. Some danger might be physical, such as running into traffic, or it might be social, like being exploited by malicious parties. In order to protect them from such harms, it makes sense to act paternalistically towards them. In other words, what justifies a paternalistic attitude towards the refugees has little to do with their capacity for self-control, and more to do with their lack of knowledge and understanding of their environment. It is in that sense that they are 'child-like'.

Similarly, we can conceptualize the difference between children and adults in epistemic terms. This conceptual scheme offers a few advantages. Firstly, it is better suited to explain the differential treatment of individuals. If an adult or a child exhibits an understanding of the situation, it would be more problematic to treat them paternalistically. The idea of 'digital natives' vs 'digital immigrants' captures this aspect well. Children born into the modern technological world of today, are often

185

likely to exhibit expertise in the use of technology that many older adults do not. In regard to the use of technology, it would make little sense to discriminate against them because they are children. Secondly, the epistemic difference explains well the differential treatment of the same individual in different situations. Someone might behave as an 'adult' when it comes to an area in which they have experience, knowledge and understanding, but be treated as a 'child' when it comes to an area in which they lack experience, knowledge and understanding. In this case, it would make sense to trust even a young child when she is in her well-known environment, but to be much more protective when she is in an unfamiliar territory. The epistemic difference is akin to the difference between an expert and a non-expert in a given discipline. Children are sometimes non-experts in the ways of life, and for this reason we should sometimes treat them differently.

I believe that some of our social norms reflect this. A striking example is the gap between the age of criminal responsibility, which in England is 10, and the voting age, which is 18. What explains the significant eight years' gap? According to the epistemic conception that I am describing, we might explain this by realizing that criminal offences are likely easier to understand, as basic conceptions of right and wrong are well developed by the age of 10 (Killen and Smetana 2015). However, as political parties are offering competing policies and different conceptions of the good,

186

together with the fact that politicians are often not as straightforward and clear as the criminal law, it requires more experience, knowledge and understanding on the part of the individual. Therefore, it makes sense to have the age of criminal responsibility lower than the voting age. Notice that we would not get far by trying to explain this gap by using the autonomy conception of the difference between children and adults.

Differentiating between children and adults on epistemic grounds also has the advantage of explaining why sometimes we would want to listen to children. Children have unique access to an important source of information, namely their own feelings, thoughts and emotions – they are the only ones experiencing their lives. As a parent, one of the most frustrating parts of caring for my new-born was trying to figure out what exactly bothered him as he was crying; was it the teeth, or the stomach, being bored, tired or sick? As children grow, they get better and better at understanding themselves, they become experts in interpreting their own feelings and emotions. For this reason, it makes sense to ask children questions about their own lives, as they give us information that might not be available otherwise. Like adults, and maybe more so, they will make mistakes in caring for the wrong things, or in interpreting their own feelings. Hence, we should be critical about the value of the information children convey. But, because they are the only ones living their lives, we should listen to children's views and opinions, which does not seem as obvious if we

187

assume children are mostly characterized by a lack of autonomy. Importantly, this does not mean that children should always be consulted, but neither should adults.

According to this epistemic conception, the different treatment of children and adults in relation to benign manipulation will be based on their knowledge and understanding of the situation. Because adults are more likely to have an understanding of a greater range of situations, there are fewer occasions where it would be a good idea to manipulate them for their own good. In other words, on average, adults are less likely to act against their own interests due to their lack of knowledge and understanding, while children, to a greater extent the younger they are, are more likely to act against their own interests due to a lack of knowledge and understanding. But, this is not an over-arching categorical difference, and should be assessed according to the specific characteristics of the person and the situation.

Relationships

Another conceptual scheme that is important in thinking about the difference between manipulating children and adults would focus on the typical relationships' children are part of. Unlike many adults, children are continually and thoroughly taken care off by adults, namely their parents or guardians, who, at least some of the time, and ideally, have their interests at heart.

188

Parents might not be the perfect judges of what is good for their child, in the same way that adults are not always the perfect judges of what is good for themselves. However, parents care for their children and are placed in close contact with them. They are likely to have a good understanding of their child's needs, desires, weaknesses, etc. as well as an understanding of what is good for their child. Also, they are well placed to take note of feedback from their child in cases where the child wishes to challenge a course of action the parents are taking. We can see this in the example of Long Way Home. The parents know that their child needs to sleep, they know how to make their child get what he needs, and they are also there when he challenges their actions, and able to explain their actions. Because children are often a part of an unequal caring relationship, in which they are the physically and mentally weaker side, they are bound to place their trust in the more powerful side of the relationship. This enables the parents to manipulate children more, and because they care for their child and are well placed to know what is good for the child, there are more instances where it would be appropriate for them to use manipulation, when they think that would be the right course of action.

However, there are limits to the parents' knowledge of their child and what is their child's best interest. As Mullin puts it, parents should listen to other views, including those of the child:

189

When we focus on paternalistic interventions, we typically focus on whether the intervened has or lacks relevant information about his or her situation, and has or lacks the ability to pursue goals and appreciate how his or her interests may be affected by courses of action. When we focus on paternalistic interventions by parents with the liberty of their young children, it is easy to assume that the parents will be not only better judges than their children but also good and accurate judges. However, it is important to consider whether even generally competent adults should, when time permits, be expected to listen to others' views, including the views of the child in question, about the likelihood of significant harm (or much lesser good) coming to the child in the absence of intervention. (Mullin 2014, p. 423)

Therefore, although parents are often well placed to be better, good and accurate judges of what is in the interest of their child, they should also have a healthy dose of self-criticism.

Moreover, not only should parents reflect upon their knowledge of their child and the desired state of affairs, they should also consider the obligations and expectations that are part of their role as parents. Even if parents judge that a certain

190

benign manipulation would benefit their child, that manipulation might breach the trust the child has placed in the parent. For example, if the child expects her parent to be not only caring but also honest, then hiding the candy from her might involve betraying her trust. Once again, there does not appear to be a qualitative difference in this respect between children and adults. In the example mentioned above, the parent hiding the candy from both her partner and her child, might rightly be criticised for abusing the trust they have placed in her.

4.4 Back to the framework

My suggestion is that instead of assuming that children and adults are qualitatively different, we can reach a more accurate and nuanced understanding of the difference between manipulating adults and manipulating children by using the framework that was established in the previous chapter. The framework, together with insights from the different conceptual schemes mentioned above, may explain why children and adults merit a differential treatment when it comes to manipulating them.

Consequences

191

First, let us consider the intended consequences of an instance of benign manipulation. In the short term, it appears that manipulating children might be more effective than manipulating adults. This could be because children are not yet fully physically and mentally developed, as well as their lack of actual knowledge and understanding of the world. As was mentioned earlier, the use of direct physical force is also more effective when used on children, simply because they are less able to resist. While this does not give us an overarching reason to use force or manipulation on children just because it is more effective, the fact that at least sometimes manipulating children is more effective than manipulating adults should be taken into account when we consider whether to use manipulation or not.

However, sometimes there would not be a significant difference in the short-term consequences of manipulation. As is exemplified by Cafeteria, there is no difference whether the supposed cafeteria is in a kindergarten, school, university or office building. In each of these, the manipulation might be beneficial to the manipulees, without it being significant whether they are children or adults. However, in the long term, as children are expected to live longer than adults, the benefits (or harms) of manipulation to their future well-being might be greater than those of adults. For example, if children are accustomed to eating a healthy diet while they are young, they might develop a stable preference for healthy eating that would

192

continue throughout their lives. It is also possible that by using benign manipulation to influence children we are missing out on opportunities to develop stable reasoned preferences. This question would probably be best answered empirically. At least from a practical point of view, there is a very high level of uncertainty when we try to consider the long-term consequences of manipulating children. Therefore, we should consider the consequences of benign manipulation when deciding if it is the appropriate method, but also be aware of the uncertainty involved, especially when we realize the actions have long-term and wide-scope consequences, alongside the easier to measure short-term consequences.

Understanding

On the question of whether the manipulation prevents the manipulee from having an accurate understanding of her situation, it appears, again, that there is no significant difference between children and adults. True, children, whose knowledge and understanding of the world is on average not as good as that of adults, might be easier to manipulate because of that. However, in both cases, manipulating both children and adults might prevent them from having an accurate understanding of their situation, and this should be considered when we are assessing whether to manipulate them or not.

Similarly, for both children and adults, manipulating them might prevent the manipulator from having an accurate understanding of the situation. Sometimes manipulation brings about a change of behaviour, beliefs or desires in a way that is at least partially hidden from the person manipulated. In a sense, this is exactly the reason to use manipulation in the first place, because at times we would like to influence someone without having to deal with their challenge to that influence. This is not necessarily bad, especially in a relationship in which the manipulator can reliably judge the effectiveness and consequences of the manipulation. However, other times, for example if the manipulator is not present at the scene of manipulation, they might not be able to judge whether the manipulation is truly effective or beneficial, all things considered.

One reason to think that manipulating children would be more problematic than manipulating adults is that commonly children are not respected as knowers at the same level as adults. Because manipulation distorts the feedback coming from the manipulee this might be more acute when it comes to children whose voice usually carries less weight than that of adults. The point will be further developed in the case studies, and especially in the discussion of pharmacological cognitive enhancers in Chapter Six.

Trust

194

As mentioned earlier, sometimes the role of a parent or a teacher calls for the use of manipulation when it is being done within the boundaries of the obligations and expectations which constitute the nature of the relationship. However, as was noted, this does not mean that within such relationships every benign manipulation is appropriate. Sometimes, the manipulation would include a breach of the terms and norms of that relationship.

The difference between children and adults in this criterion is that children are often in more encompassing care relationships than adults. That means, they are the children of someone, they are the students of a particular teacher, etc. For this reason, as parents and teachers have a caring role in the life of the children under their care, it is likely that they are better judges of if and when to manipulate, and also are better placed to see whether the manipulation is truly effective and beneficial. That might be a significant difference from adults, who are less likely to be in such encompassing relationships.

In sum, although children and adults might be treated differently in regard to manipulation for various reasons, there is no principled argument to claim that the manipulation of children is more or less objectionable just because it is performed on children. Just like adults, the best approach would be to consider the flaws and merits of an instance of manipulation, alongside the context in which the

195

manipulation is taking place. To do that, one should use the assessment framework established in the previous chapter.

After establishing a framework to assess benign manipulation and showing that there is no principled reason to discriminate between children and adults in regard to manipulating them, I move on to discuss three case studies of techniques used in education that can be usefully analysed as benign manipulation. I have chosen these specific case studies due to the fact that they highlight aspects of the framework that are the more complex and less commonly considered when educators choose an educational technique. The three case studies are: deceiving children in the aim of raising academic performance, pharmacological cognitive enhancement, and gamifying education.

Chapter 5   Benign Deceptions

5.1   Introduction

The use of lies and falsehoods in education has received some attention in recent years. Two examples are the bestselling book *Lies My Teacher Told Me* (Loewen 1995) in which the author James Loewen criticizes American History textbooks for presenting false and inaccurate information about American history, and Jonathan Kozol's critique of lies told to underprivileged school children in the United States (Kozol 2007, pp. 189-91). In philosophical circles, Harry Frankfurt's work "On Bullshit" has drawn attention, to the extent that some universities are offering courses in bullshit detection (Bergstrom and West; Frankfurt 2005). Also, Bernard Williams' work on Truth and Truthfulness has renewed interest in this philosophical topic (Williams 2002). Not only that, recent political trends, growing use of social media and other current events have made the concepts of lies, deception, spin, falsehoods, fake news, alternative facts, etc. very prominent in our society. Often, as is exemplified in the works mentioned above, the use of falsehoods, lies and deceptions by figures of authority such as educators and politicians is criticized and even ridiculed.

197

On the other hand, there is also a growing body of scholarship that acknowledges the importance of falsehoods for practical, social, psychological and epistemic purposes. For example, philosopher Amelie Rorty lists the benefits of self-deception (Martin 2009), Peter Klein argues that falsehoods may have epistemic benefits (Klein, Peter D. in Smith 2008(ed.)), and Benjamin Rancourt (Rancourt 2015) similarly claims that our knowledge and understanding might improve through believing falsehoods. Recently, the journal *Philosophical Explorations* dedicated a special issue to the question 'How can false or irrational beliefs be useful?' (Bortolotti and Sullivan-Bissett 2017). Following these theorists, Ruth Wareham considers the links between epistemically useful delusions and indoctrination (Wareham 2019).

This complexity of lies and falsehoods being both problematic and useful plays out in educational relationships and processes. It is almost impossible to imagine a healthy and normal educational relationship which does not include at least some hiding of the truth, if not full-blown lies. Intuitively, censoring information appears to be a good idea in some instances, as well as communicating simple, yet inaccurate information, which is easier to comprehend than the complex yet more accurate information. However, there appear to be cases where deceiving children as part of their education is harmful for the child, and even if it is beneficial in some respects, the question of how to assess the use of altruistic deceptions in education

is a difficult one. I suggest that the framework for assessing benign manipulation that was established in the previous chapters is well suited to consider whether a particular lie or falsehood, which is well intentioned, is appropriate.

My analysis aims to show that although we cannot hope to achieve a comprehensive set of rules to guide educators, we can shed light on the most salient aspects educators should consider when deciding to deceive those under their care. Those aspects will be the intended and unintended consequences of the deception, the extent to which the deception hinders both the educator and student understanding of their situation, and whether the deception is appropriate within the norms that constitute the relationship between the educator and the educated.

Here are a few examples of deceiving children for their own good:

False Expectations: A student approaches her teacher after class and says: "I am really bad at Math and I will not get a good grade on the exam." The experienced teacher believes the student is probably right, as she assesses the student's chances of performing well on the exam as small. However, because the teacher has recently learned about the Pygmalion Effect, i.e. that acting as if her students are capable of achieving better results than those she genuinely believes

they are capable of leads to them performing better, she decides to say to the

student: "I have full confidence in your ability to perform well on the exam." [22]

Water the Plants: To allow an overexcited and energetic child a method to let

out some energy, her teacher asks her to water plants on the school premise using a

very heavy water bucket. The child is asked to do this every day, for a number of

days. She is told that this requires demonstrating responsibility, as the plants will die

if she forgets to water them. Unbeknownst to the student, the plants are made of

plastic (adapted from the podcast Israel Story).

Noble Lie: Socrates recommends deceiving the youth of the Republic about

their origins by telling them a myth - they were all born from the ground and

therefore, as brothers and sisters, they are to protect the state and each other.

Christopher Columbus: An American History textbook makes the following

wrong claim: Christopher Columbus attempted to reach India because there was a

dire need in Europe for spices to disguise the taste of bad meat. This shortage in

---

[22] For further discussion of the Pygmalion Effect, see below on p. 190.

spices was due to the fact that the Turks had cut off the spice trade with the East following the fall of Constantinople in 1453 (Loewen 1995, p. 34).[23]

Triangles: During a grade 7 geometry lesson, the teacher writes on the board: 'the sum of all angles in a triangle adds up to 180 degrees.' She asks the students to write the sentence in their exercise books and memorize it for their next lesson. Alongside her teaching role, the teacher is also researching towards a PhD in the area on Non-Euclidian geometry, and she is well aware of the fact that the sum of angles in triangles may be different from 180 degrees.[24]

The Hortatory Lie. Underprivileged children are repeatedly told that it is up to them whether they succeed or fail in their studies. The structural, socio-economic hurdles they are facing are not acknowledged or communicated to them. Each

---

[23] Following the common description of the life of Christopher Columbus in many American History books, Loewen writes: "Unfortunately, almost everything in this traditional account is either wrong or unverifiable. The authors of history textbooks have taken us on a trip of their own, away from the facts of history, into the realm of myth. They and we have been duped by an outrageous concoction of lies, half-truths, truths, and omissions, that is in large part traceable to the first half of the nineteenth century." (Loewen 1995, p. 32)

[24] Harvey Siegel provides a similar example in his response to Michael Hand's argument against faith schools. Briefly, Hand claims that faith schools teach for belief in religious propositions which are not known to be true, hence they are indoctrinatory. However, Siegel provides the example of a physics teacher who teaches her class the Newtonian law that F=ma, which is not known to be true, and yet she is not indoctrinating. Supportive of Hand's argument, Siegel suggests to restate Hand's argument by adding that the non-indoctrinating teacher teaches so that: "the proposition be supported by good reasons/strong evidence which confer(s) upon it positive justificatory status, and moreover that the student be encouraged to believe it on the basis of those reasons/that evidence." (Hand 2003, Siegel 2004)

201

morning they are asked to repeat the following mantra 30 times: "I have confidence that I can learn" (Kozol 2007, pp. 273-4).

As discussed in Chapter 2, when deception is used to influence others, it is often regarded as a sub-set of manipulation. The examples above are all manipulations which use deception in order to influence others with the intention to benefit the deceived. In Chapter 2 I presented the example of a teacher setting up a competition among her students as an example of manipulation which does not include deception. That example was meant to show that not every manipulation includes deception. However, many of them do, and the variety of examples presented here demonstrates that. For the remainder of this chapter, the discussion assumes that benign deception is a sub-set of benign manipulation and so can be assessed by the framework established in the previous chapters. For the majority of the chapter, I focus on the first example, that of communicating false expectations, but will address the other examples when relevant.

5.2   What is deception?

There are various definitions of deception. Sissela Bok, whose 1978 book *Lying* has revived the age-old ethical concern about the ethics of deception, formulates what seems to be the common-sense view:

> When we undertake to deceive others intentionally, we communicate messages meant to mislead them, meant to make them believe what we ourselves do not believe. We can do so through gesture, through disguise, by means of action or inaction, even through silence… I shall define as a lie any intentionally deceptive message which is stated. (Bok 1980, p. 13)

A different definition of deception can be found in Thomas Carson's 2010 book 'Lying and Deception'. Carson's definition and discussion of it sums up much of the debate about deception that was initiated by Bok's work. He claims that an attempt to deceive may be considered deception only if it is intentional, causal and successful, and that one may deceive someone else only into false belief:

> A person S deceives another person S1 if, and only if, S intentionally causes S1 to believe X (or persist in believing X), where X is false and S does not believe that X is true. (Carson 2010, p. 50)

A short comparison of these two definitions will explain how I view deception for the purpose of assessing altruistic deceptions in education. For the most part, I follow Bok's definition that intentional deception is the attempt to make someone believe something which the deceiver does not believe, due to practical reasons which will be explained below.

Assuming we consider only intentional deception, there are two main differences between the definitions. First, Carson believes that communicating false belief is deception only if one succeeds in causing another to believe the false belief is true (Carson 2010, pp. 55-56). However, in the context of teacher expectations, because a recommendation to communicate false expectations needs to consider successful and unsuccessful attempts, I will include 'successful deception' as a sub-category within 'attempted deception'. This is because unsuccessful attempts at deception have consequences as well, and if we wish to assess a deceptive technique, we need to assess the possible consequences of attempting to use it, not just the consequences of the successful attempts. To explain, suppose I try deceiving my child for his own good by telling him there is no candy in the house, only he does not believe me. My intentional unsuccessful attempt will not have the desired effect of preventing my child from eating candy (I will have to use another action – such as force) and it is more likely that my child will be upset with me and

204

will not trust me in similar circumstances. To be clear, I am not arguing against Carson's definition, but claiming that for the purpose of considering the use of intentional deception as an educational technique we need to account for the uncertainty involved. When teachers act as if they have high expectations from students, the students sometimes believe the teachers and sometimes they do not, but the teacher at least tried to deceive. For simplicity, throughout the rest of the chapter I will use the word deception to refer to both successful and unsuccessful attempts to deceive.

The second difference between the two definitions is Carson's claim that one deceives only if she makes another believe a false belief, while Bok does not provide this condition. This means that according to Carson, one is not deceiving if she causes someone to acquire a true belief. Suppose, while trying to prevent my child from eating candy I tell him there is no candy in the house, when I think there is. However, without my knowing, my daughter ate all the candy, and therefore I am actually telling the truth. Carson does not consider this deception, but I do. As my aim is to assess attempts to deceive, whether these are successful or not, I consider it deception when the agent intends to deceive, even if ultimately, they are telling the truth. This is because my aim is to assess deception before or when it is attempted and not in hindsight when the truth of the matter is discovered.

205

To elaborate further in relation to the case of communicating false expectations; when a teacher communicates false expectations that teacher knows what her expectations are, but at best she might only have a reasonable guess of what the true abilities of the student are. The teacher does not know if the student will live up to the expectations when she considers whether to communicate false expectations or not. The teacher in these cases communicates a belief about her own perception. It is the difference between 'I see a tree' and 'there is a tree'. If I see a tree and I am saying 'I do not see a tree', I consider this an attempt to deceive, even if there really is no tree, i.e. what I see and believe to be a tree is, in fact, a very well-made model of a tree; or a hologram; or, maybe, a bush that is not actually a tree according to the botanical definition. I was lying about my perception of a tree and not about the existence of a tree. Likewise, if a teacher believes a certain student is expected to perform p and acts as if the student is expected to perform $p+$, and the student performs $p+$ then in this case the teacher was wrong in expecting p, but she was deceiving the student about her expectations; she truly expected the student to perform p and communicated the belief that the student will perform $p+$.

I wish to make two more remarks about how I understand deception as opposed to lying in the context of student-teacher relations. Lying I take from Bok to mean an intentional deception that is stated. When thinking about intentional

206

deception there is no reason to think that there is a normative difference between lying to someone and deceiving someone. For example, I ask a student who did not show up for school why she did not come. The student says: "I was at the hospital". Now, if she was not at the hospital she is lying. But what if she deceives me by going to the hospital for no other reason than not to lie when she says: "I was at the hospital"? Many lies can be made into not being a clear-cut lie, but that does not change their fundamental qualities as deceptions. For that reason, I am concerned with the larger concept of deception and not only lying. A teacher may deceive about her expectations without lying but that does not mean she is acting appropriately just because her deception is not stated in words.

Moreover, deception does not require the deceiver to know the truth. For example, a teacher deceives a student if she communicates high expectations and does not have any expectations at all, high or low. This type of communication, in which the deceiver has little or no concern for the truth, was aptly named 'Bullshit' by Harry Frankfurt (Frankfurt 2005). Thus, telling a newly met student that you have high expectations for her is deceiving her. Also, and from my experience that is a common practice in schools, telling students they have an unfulfilled potential, when the teacher has no reason to think so, also counts as deception. For now, I do not

207

argue against such a practice (I actually think it is reasonable in some cases, and will elaborate below), just emphasize that such practice is indeed deception.

In sum, intentional deception in the context of interpersonal expectations in teaching is any attempt by a teacher to communicate false expectations to a student, whether successful or not in making the student believe these expectations are what the teacher believes. Also, this intentional deception may be carried out by a lie, i.e. a false statement, or other means of communication without there necessarily being a moral difference between these different methods of deceiving. And, a teacher deceives if she communicates high expectations while having no expectations at all.

5.3   Is deception wrong?

As I assume that deception is a type of manipulation, and manipulation should be understood as a non-moralised concept, I want to support the view that deception should also be understood as a non-moralised concept. This is unintuitive. We feel that deception is wrong. However, it is extremely hard to explain why. Kant famously argued that we have a duty not to lie, even to a would-be murderer in order to save

an innocent life (Kant 1993). If this is really what Kant thought about lying,[25] it seems to be too extreme, and even if every lie is a crime, it does not imply that every deception is prima facie wrong. Bok takes it for granted that deception requires justification and that different lies should be justified by appeal to their consequences for individuals and institutions (Bok 1980). For example, placebo treatment would have been wrong if it had no benefits, but it might be justified if it improves the health of the patient and does little or irrelevant harm to the general trust in doctors and the medical institutions. An interesting attempt to explain why deception is wrong is Bernard Williams' claim that trustworthiness is intrinsically good, and that sincerity and accuracy are virtues, hence deception is wrong (Williams 2002). Nonetheless, I believe the issue of why deception is wrong is still unresolved.

Consider the example of Triangles above. In this case, when the mathematics teacher chooses to communicate information which she herself does not believe is true, she appears to have a good reason to do that, as students might need to learn the simpler Euclidean geometry before they can learn more complex non-Euclidean geometries. Nonetheless, she communicates information which she herself does not believe is true, and this counts as deception according to Bok's definition. Also,

---

[25] Kant's view of lies is probably more complex, see Mahon, J.E. 2009. 'The Truth about Kant on Lies', in *The Philosophy of Deception*, Martin C. W. (ed.), Oxford University Press, pp. 201-24.

209

consider the many stories we tell children that we ourselves do not believe are true. And, being polite and tactful can very often involve communicating information which we ourselves do not believe is true. Or, consider a play based on true events, which does not portray past events accurately, as these are dramatized for the enjoyment of the audience. All of these examples fit the definition of deception, but are they prima facie wrong and require justification? It does not appear that we are required to justify simplifying information, telling children made up stories, or being polite or entertaining. I believe these examples are enough to demonstrate that deception is often unproblematic. Although one might claim that the deceptions described in these examples are unproblematic because there is an accompanying reason that makes them so, it is also possible to claim that in examples where the deception is wrong there is an accompanying reason that makes them wrong. Ultimately, I do not think there is much difference between these two approaches. In both, the focus should be on what are the accompanying reasons that make deception wrong, right or unproblematic. For practical reasons, I proceed by assuming a non-moralised understanding of deception. When deception is wrong, right or unproblematic, it is because there is an accompanying reason that makes it so. These might include harm caused to the deceived, that the deception prevents the deceived from having an accurate understanding of her situation, that it constitutes a breach of trust, or other reasons.

In relation to teachers communicating false expectations, according to my approach, the teacher's action should not be considered inappropriate for the sole reason that her action involves deception. To argue whether or not the teacher is acting appropriately, we would need to consider her action in light of the framework for the assessment of benign manipulation that was established in the previous chapters. To repeat briefly, in order to decide whether the teacher acts appropriately if she decides to deceive the student, we need to ask about:

Criterion A) Consequences: 1) intended consequences – to what extent is the suggested benign manipulation effective in achieving the desired result in the short term, long term and across contexts? 2) unintended consequences – to what extent does the suggested benign manipulation have unintended consequences in the short term, long term and across contexts?

Criterion B) Understanding: 1) to what extent does the manipulation prevent the manipulee from having an accurate understanding of the situation? 2) to what extent does the manipulation prevent the manipulator from having an accurate understanding of the situation?

Criterion C) Trust: is the manipulation appropriate according to the norms which guide behaviour in the relationship between the manipulator and manipulee?

211

Before moving on to look more closely at the idea that communicating false expectations might be a good idea, I wish to make one last comment. Throughout the chapter I focus on the personal interaction between a teacher and a student, since that is the minimal interaction in which teachers deceive. However, probably the most common interaction in which teachers deceive students happens in the classroom. Classroom interaction is more complicated because it involves many different relations: teacher-student/s, student-student, etc. For example, a teacher wishing to treat her students equally may wish to make them believe she is equally fond of each of them. This brings in the question of fairness in class, and other morally complex issues. One might argue that communicating similar high expectations to all students in class is the just and fair thing to do, even if that means deceiving those students of whom the teacher has low expectations. These considerations are beyond the scope of this work, but the analysis does have implications that are relevant to broader questions of educational justice, and I will discuss some of those below.

5.4 What is the Pygmalion Effect?

The idea that false expectations might cause students to live up to those expectations has historical roots in what is known as the 'Pygmalion Effect', an idea

that came to prominence a few decades ago and has since subsided. I have chosen

to analyse this specific example because the time that has passed since it first

appeared on the educational scene can assist in achieving greater clarity regarding

the wider effects and consequences of using this technique specifically, and

deceptions in education more generally. Also, because my initial intuition was that

some instances of communicating false expectations are appropriate, but to

communicate false expectations as a rule appears inappropriate, I find it an insightful

case to analyse and draw conclusions from. Moreover, although the debate about

the 'Pygmalion Effect' has subsided, there are related incarnations of the idea that

one's beliefs about oneself have educational implications and educators need to

promote a change in those beliefs. This can be seen in the example of The Hortatory

Lie mentioned above, as well as in currently popular theories such as mind-set, brain

training, grit, deliberate practice and the bilingual advantage (Moreau *et al.* 2018).

These theories assume that easily changeable environmental factors play a

significant role in determining the success of individuals in a range of real-world

pursuits, in contrast to that person's innate talent or socio-economic background.

They share a common behavioural assumption with the Pygmalion Effect – that

abilities and other traits are very malleable, and the theories often suggest a simple

intervention as a recipe for success. Much like the research on the Pygmalion Effect,

which I will discuss at length below, these theories have been criticized for

213

overstating the evidence that provides support for the accuracy and efficacy of both the theories and their implementations in practice (Moreau *et al.* 2018). Although my focus is on the somewhat forgotten Pygmalion Effect, I believe a similar analysis can be undertaken for any theory which overstates the degree to which a simple intervention may improve individuals' performance in various fields, such as academic studies, sport or financial success.

In 1966 Robert Rosenthal, a Harvard University researcher and Lenore Jacobson, the principle of an elementary school, tested whether differences in teacher expectations led to differences in students' performance (Rosenthal and Jacobson 1966). In their research, children from first to sixth grade were tested for IQ, and teachers were made to believe the test examines the children's ability to make intellectual gains. The teachers were told that some children are expected to be "growth spurters", that is - are expected to show an 'unusual intellectual gain' in the next year, while in truth those children were chosen at random (Rosenthal and Jacobson 1966, p. 115). Eight months after the initial IQ test the children were retested, and according to the researchers the children randomly labelled as "growth spurters" 'showed a significantly greater gain in IQ score than did the control children' (Rosenthal and Jacobson 1966, p. 116). Although the research shows the

214

effect was significant mostly for the first and second grade children only, the researchers thought it provided:

> …further evidence that one person's expectations of another's behavior may come to serve as a self-fulfilling prophecy. When teachers expected that certain children would show greater intellectual development, those children did show greater intellectual development. (Rosenthal and Jacobson 1968)

The researchers' book about the subject, named *Pygmalion in the Classroom,* initiated a great interest in the phenomenon as well as significant criticism. But, if Rosenthal's and Jacobson's conclusion is valid,[26] it appears reasonable to infer from it that teachers should communicate high expectations to their students in order to improve the students' performance, regardless of their real expectations. However, there are some important differences between research and practice in this case.

In the research conducted by Rosenthal and Jacobson and similar other researches on the Pygmalion Effect, teachers are unaware of their expectations

---

[26] For a meta-analysis of Pygmalion Effect and Self-fulfilling prophecies in education see Jussim, L. and K.D. Harber. 2005. 'Teacher Expectations and Self-Fulfilling Prophecies: Knowns and Unknowns, Resolved and Unresolved Controversies. *Personality and Social Psychology Review* 9(2): 131-55.

being influenced. When teachers are made to think that some of their students are "growth spurters", the process taking place can be described generally like this:

1) X believes Y's performance will be p

2) X communicates expectations that (acts as if) Y's performance will be p

3) Because of (2), Y performs p

In the research outlined above and others similar to it, X represents teachers and Y students. However, this formulation does not represent the agent that deliberately communicates false expectations to improve the performance of the other actor, i.e. knowingly deceiving. In the case of the research above, those performing the manipulation are the researchers themselves, since they are the ones who knowingly deceive others. Hence, if we wish to consider whether teachers should deliberately communicate false expectations, a different description of the process is required:

1) T estimates S's performance to be p

2) T communicates expectations that (acts as if) S's performance will be p+

3) S performs p+ because of (2)

216

This formulation allows to see that in the original research of Rosenthal and Jacobson, this process was not performed on the students but on the teachers. The researchers estimated the teachers' performance to be p towards the control group, they communicated expectations that the teachers' performance will be p+ towards some students by deceiving the teachers to believe these students are "growth spurters", and (supposedly) teachers really performed p+ towards these students, due to that.[27]

The findings of the initial research were quickly turned into policy recommendations, calling teachers to 'act as if' they have high expectations from students even if they did not believe the students are capable of fulfilling those. Here are some fairly recent recommendations to use the 'Pygmalion Effect' in class:

> Even if a teacher does not truly feel that a particular student is capable of greater achievement or significantly improved behavior, that teacher can at least act as if he holds such heightened positive

---

[27] Note that this formulation does not describe a process in which the teacher believes the student can really perform p+. For example, if the teacher estimates that the reason for the gap between the student performance and her high expectations from the student is due to the self-image of the student. Then if the teacher acts as if the student can perform p+ he's not deceiving, since he believes the student can perform p+, knows why she does not, and acts in a way that is meant to narrow the gap – she helps improve the student's self-image. In such a case there is no deception going on.

217

expectations. Who knows, the teacher very well may be convincing to the student and, later, to himself. (Tauber 2009, p. 3)

It can be difficult to deliberately change our expectations of others. But we can consciously change our behavior. By adopting the set of behaviors above, we'll be acting like our kids, our students or our employees have great potential — potential that they'll more than likely live up to. (Paul 2013)

…practice is not research. We should harshly judge the clinician who abstains from using the placebo effect as an aid to treatment. If smiling, nodding reassuringly, and saying to a patient, "I am sure the treatment I have prescribed will make you feel a lot better" can augment the drug in promoting the healing process, then physicians should add these placebo interventions to their treatment repertoire. Refraining from creating physician expectation effects on methodological grounds is wrong both practically and ethically. The same reasoning applies to wilful creation of leader expectation effects. (Eden 1992, p. 275)

The recommendation to 'act as if' a certain student is capable of greater achievements or significantly improved performance may arise from a concern for

218

the student and so be well intentioned. Only, it does not give much thought to the fact that this involves deceiving her. As discussed above, although using the Pygmalion Effect necessarily involves deception, it does not follow that it is wrong, prima facie or overall. In order to establish whether it is an appropriate action, we should use the framework for assessing benign manipulations.

5.5 Criterion A) Consequences

5.5.1 Intended consequences

The results of the initial Pygmalion Effect research were criticized by many and generated various discussions and follow ups, questioning the validity of the methodology and conclusions of the research, the real size of the effect, its implications for education and more. In a comprehensive meta-analysis research authors Lee Jussim and Kent D. Harber summarize the findings of the last 35 years. They conclude that empirical research on the subject justifies the following conclusions:

> (a) Self-fulfilling prophecies in the classroom do occur, but these effects are typically small, they do not accumulate greatly across perceivers or over time, and they may be more likely to

219

dissipate than accumulate; (b) powerful self-fulfilling prophecies may selectively occur among students from stigmatized social groups; (c) whether self-fulfilling prophecies affect intelligence, and whether they in general do more harm than good, remains unclear, and (d) teacher expectations may predict student outcomes more because these expectations are accurate than because they are self-fulfilling. (Jussim and Harber 2005, p. 131)

From this we may infer that if communicating false expectations were to be implemented as a general policy, the benefits would be rather small. However, if it is indeed the case that 'powerful self-fulfilling prophecies' do occur among certain groups, it might be a reason to use this technique for those groups only. Furthermore, one might claim that although the benefits are small, they do exist, and since there is no obvious harm in communicating false expectations, why not do so? This is especially relevant to my own understanding of deception as a non-moralised concept. If deception is not necessarily even prima facie wrong, there appears to be no reason to avoid using it even if the benefits are small. In what follows next I will discuss possible accompanying reasons that make the communication of false expectations less desirable.

5.5.2 Unintended consequences

While at times deceiving students about what is expected of them may improve their performance, there are good reasons to be critical towards advising teachers to use this technique. First, as discussed above, implementing this technique by teachers is asking them to deliberately deceive students, which might be less effective than the empirical research shows. That is, the findings of the research summarized by Jussim and Harber relate to manipulation of teacher expectations inflicted by researchers and not to teachers knowingly acting as if they have high expectations. A recommendation to use the Pygmalion Effect in class requires teachers to wilfully communicate expectations that they do not believe are true. Hence, if they are not successful in mimicking perfectly, by verbal and non-verbal moods of communication, genuinely having high expectations, then the effect is probably smaller even than the small effect suggested by Jussim and Harber. In other words, only if the teachers succeed in their deception on all occasions and students never notice they are being deceived, will the effect equal those measured in research. But, is it possible for teachers to be such good deceivers? On this exact question Bella DePaulo writes that:

> To the question of whether teachers can communicate high
> expectations if only they try, the Feldman (1976) study, as well as the

221

body of literature on the communication of deception, suggests the answer 'sort of.' Teachers who try to convey high expectations will probably appear to have such expectations. However, they will be a tad (but a significant tad) less convincing than teachers whose expectations really are high. (DePaulo 1993)

This suggests teachers are not all-powerful deceivers, and the desired effect of 'acting as if' one has high expectations is smaller than genuinely having those expectations. Also, it means the undesired effects of using the technique are larger, since the attempt to deceive is likely to fail more often. Nonetheless, it is possible that for some teachers deceiving some students the effect of knowingly deceiving students will be similar to the effects of the research, for example if the teacher is a very good deceiver or the student very insensitive to deception. Perhaps that is the case for teachers teaching low grades or students new to the school, but these are the exceptions. Generally, those who want to recommend the use of Pygmalion Effect should be mindful that the desired effect is not only small as shown in research, but actually smaller since teachers who are aware they are deceiving will probably fail more often in convincing students they have high expectations, compared to teachers who are themselves manipulated by researchers to have high expectations.

222

A related issue is that if an attempt to deceive fails, that would harm the trust placed in the deceiver. I shall elaborate further below on the implications of deceptions for the trust which is part of the student-teacher relationship. For now, consider the moral of the story about the boy who cried wolf. If a teacher deceived about her expectations often, and on quite a few occasions those expectations do not materialize, students will be less likely to believe similar assertions in the future. Hence, from a practical perspective, the teacher should communicate false expectations quite rarely, if she wishes those utterances to be believed.

Those who recommend the use of Pygmalion Effect as a policy, assume that students generally believe that what their teachers are telling them is true. However, adopting a technique that includes deceiving students will jeopardize this belief. If the deception is not successful, that is the student does not believe the expectations the teacher is communicating to be her real expectations, that would make the student less likely to believe the teacher in the future. If the deception is successful, and the student does believe the teacher's expectations are her real expectations, but there is little connection between these expectations and what actually happens (in the next test, for example) the student learns the teacher is wrong most of the time. The student will think the teacher's assertions are not aligned with reality and this again will make the student less likely to believe the teacher in the future. Another

223

possibility is that the student will keep believing that the teacher expects her to

perform well, but conclude that the fault was her own, that she did not work hard

enough or that she is not really academically capable to the extent the teacher

believes.[28] Holding those beliefs might move the student further away from an

accurate understanding of her situation, an issue which I touch upon below.

Another way to illustrate the effect deception has on the trust in the accuracy

of the information conveyed is to think of the effect the manipulation of the

researchers of the Pygmalion Effect had on the trust teachers have in researchers.

The manipulation done by the researchers of the Pygmalion Effect was possible

because there was a basic trust of the teachers in the researchers. That trust was

---

[28] Carole Dweck's work is of relevance here. She claims that people may be divided roughly
into those with a growth mind-set and those with a fixed mind-set. Those with growth mind-set believe
that human attributes (e.g. intelligence) are malleable with effort, while those with a fixed mind-set
believe that human attributes are innate and are not malleable (Dweck 2006). If this theory is correct,
it circumvents some of the problems I find with the Pygmalion Effect, in that it places the emphasis on
the belief the student has about how she can learn and develop, rather than what she is innately
capable of. To explain, if the belief the teacher communicates to the student is that 'the brain is like a
muscle, and you will be able to develop your abilities with practice' then the teacher is not deceiving
the student if the teacher truly believes that. However, there are a few problems with the mind-set
theory and practice, and a similar argument to the one presented here in regard to the Pygmalion
Effect can be made against it. First, the theory appears to exaggerate the malleability of human
attributes. The empirical evidence that people with a 'growth mind-set' perform better than those with
a 'fixed mind-set' is far from being conclusive (Sisk et al. 2018). Second, even if the theory is correct,
it places significant emphasis on one's responsibility and control for one's performance, rather than
various other environmental factors. This emphasis, if overstated, may in itself be considered as
manipulation. To explain, a student that buys into the theory of a growth mind-set may believe that
whether she succeeds in her studies is mostly up to her. In this, the theory might serve an ideological
agenda. Especially, in the competitive zero-sum game of schooling, someone has to fail, as David
Labaree puts it, and so whether one succeeds or not depends at least in part on the rules of the
game, rather than how good one is (innately or not) at playing it (Labaree 2010).

224

strengthened by the researchers' choice to give a grandiose name to the test that was supposedly meant to check which of the students were 'growth spurters', the name was: "Harvard Test of Inflected Acquisition". Thus, the researchers reinforced the credibility of their deception not only by using technical terms, but by using the whole of Harvard's reputation. However, as the research became famous in the educational community, and even taught in teacher training, it probably became harder for researchers to gain the confidence of teachers in these kinds of research. Similarly, the more teachers adopt Pygmalion Effect as a policy, the harder it will be for them to convince students of the truthfulness of their communicated expectations. The reason for this mistrust in information is a weak link between the communicated information and the facts as perceived by the students. That is, the harms to trustworthiness seem to accumulate over time and across different situations. Even if it might be efficient to deceive in certain cases, it surely harms trustworthiness to 'act as if' on many occasions. Unsuccessful prophecies, whether unsuccessful because they are not believed or because they do not come true, harm the prophets' trustworthiness.

There are also other harms to be considered. If the deception is unsuccessful, the deceived trust in others somehow 'like' the deceiver and in other people in general might diminish. The student will not only mistrust the teacher when the

225

teacher acts as if she has high expectations from the student but possibly also other elements of the teaching profession, and other teachers. As teachers, when a student deceives us about something, this has implications for our general trust in her and our trust in other students. In a similar fashion, when deceiving in regard to her expectations the teacher loses credibility in other matters and other teachers lose credibility as well. Hence, not only are the benefits of the Pygmalion Effect as shown by empirical research small, these benefits assume the students have trustful relations with their teachers. But, because an implementation of Pygmalion Effect as a policy will harm teachers' trustworthiness it probably will not have the benefits those who recommend it hope for.

There is certainly quite a bit of empirical speculation in the reasons I list above. Although I believe these negative consequences of deceiving students are possible, if not likely, the point is that they should at least be considered, and teachers should be 'on the lookout' for them when they decide to deceive students. In general, those who recommend an extensive use of Pygmalion Effect seem to concentrate on the possible good consequences of the successful attempts without considering the harms of the unsuccessful ones. This is like recommending a basketball player to shoot the ball far from the basket because the ball might get in, without realizing that the chances of it happening are small and that he will probably

226

turn over the ball. A justification of the policy to 'act as if' must take into consideration the unsuccessful as well as the successful attempts to raise academic achievement and performance by deceiving students.

To conclude, if the recommendation to implement Pygmalion Effect were to be applied as the default option by teachers, it appears the benefits of it would be generally small, if there were any such benefits at all. However, there might still be cases in which Pygmalion Effect is effective in achieving the desired result, even if not applied as the default option. In order to continue the analysis of false expectations as manipulation, I move on to discuss Criterion B) Understanding, and Criterion C) Trust.

5.6 Criterion B) Understanding

5.6.1 The Manipulee's understanding

By definition, successful deception prevents the deceived from knowing the truth, at least in the short term. Sometimes, the truth in question will prevent the deceived from having an accurate understanding of important aspects of her situation. When a teacher deceives a student in regard to her expectations from that student, she might impede their knowledge and understanding of the situation they

227

are in. If I believe that the teacher thinks I am expected to do well on the exam, that might cause me to act against my interests, since I am not aware of the real situation. For example, I might not study before the exam, because the teacher already expects me to do well.

However, at other times believing some falsehoods can be useful for the deceived. For instance, falsely believing that the teacher has high expectations from me, might raise my confidence and motivate me to put more effort into my studies. Moreover, some falsehoods might even be epistemically useful. For example, Peter Klein discusses the example of astronomy students in the Middle Ages using the Ptolemaic model to calculate the position of Mars and Earth 800 years into the future. Assuming their calculations are correct, and the model allows for such extrapolations, they would be gaining knowledge that is based on false beliefs regarding the fixed position of the Earth and the orbits of the Sun and Mars (Klein 2008). Similarly, a teacher might communicate false expectations, which will cause the student to live up to these expectations, to the teacher's surprise. In such a case, the teacher's false expectations do serve as a self-fulfilling prophecy.

It is difficult to draw principled conclusions as to whether for epistemic reasons teachers should or should not deceive their students regarding their expectations from the student. One possible solution would be to try different

228

approaches at different times. As the relations between teachers and students are ongoing, the teacher can, when she finds it reasonable, communicate false expectations to see if these might assist in discovering unseen aspects in her student, and other times, she can communicate her real expectations, to allow the student to have a better understanding of her situation. In choosing which approach to use in a given instance, the teacher would draw on her experience and professional skills, i.e. her current knowledge of the student, the environment, their relationship, etc. This is already quite distinct from a recommendation to always or mostly 'act as if' the teacher has high expectations from the student.

Even if we assume the teacher has reasonable contextual knowledge, it may still be extremely hard for her to know how to act. For example, there may be students from disadvantaged groups, or in particular personal circumstances, for whom a teacher saying they expect them to do badly in a particular exam or subject would have the effect of making them try harder just to prove the teacher wrong. Would a teacher in such circumstances be justified in knowingly deceiving the student – i.e. deliberately communicating lower expectations than she actually has? A simple answer here would be that teachers should not act on the sole basis of the causal efficacy of their expectations but instead convey accurately whatever expectations they do have.

229

5.6.2 The Manipulator's understanding

When a teacher deceives a student in regard to her expectations, the teacher might also impede her own knowledge and understanding of the situation. Deceiving a student interrupts our perception of the real difficulties the student is facing. Often, it is a way of brushing off the real problems. One reason why deceiving another person might impede our own perception of reality is that deception, and other forms of manipulation, do not encourage the kind of feedback that telling the truth does. Deceiving a student that they are expected to perform well on the exam, when the teacher thinks otherwise, does not bring to the foreground the question: "Why?" as much as telling the student that you do not expect her to do well on the exam. In such a case, the teacher will reflect less on the reasons that might cause the student to perform at the level that she does. As was discussed earlier, in this, manipulating someone is different from coercing them. Since coercion is visible and manipulation is often not visible, at times coercion might actually be a better way of promoting knowledge and understanding. Feelings, emotions and attitudes of the manipulated are distorted when we choose to manipulate, in a way that is less likely to happen when we choose to coerce or convince someone. This is partly why manipulation works when it does, because it circumnavigates the resistance more than coercing someone or reasoning with them. However, resistance can be an important aspect of

230

human attempts to influence one another. So, there are good reasons for the influencer to want to know whether the influenced resists the attempt to influence her. In this sense, manipulation might impede our understanding of reality; the manipulator does not know what the manipulated genuinely thinks or feels about the attempt to influence her. Therefore, even if manipulation might assist in maximising certain aspects of what the manipulator believes is good for the manipulee, it might prevent the manipulator from considering other interests of the manipulee, by distorting the manipulees reaction to the influence operating on her.

To explain, consider two examples, one in which the teacher manipulates the student by communicating false expectations, and the other in which the teacher forces the student to practice extensively before the exam. In both cases, the teacher might not know, or care, what the student is capable of, whether the student will improve their overall academic performance due to performing well on the exam, or other aspects of the situation. By communicating false expectations, the teacher is less likely to be challenged over her beliefs and the course of action she chose. This is because often manipulation works in a way that is hidden from view and does not raise questions about the efficacy and legitimacy of the influence one is subjected to. In contrast, forcing a student to study hard for the exam is more likely to raise the

231

question 'why should I do that?', as the use of coercion is visible, and does raise questions about efficacy and legitimacy.

Furthermore, educators would be wise to recognize their own fallibility. Though teachers and other educators are often more knowledgeable than students and have a good understanding of the situation, they are also likely to be mistaken quite often. By communicating false expectations, we will be creating an environment that undermines the values of honesty and truthfulness. Such an environment will make it more likely for educators to turn a blind eye, or even self-deceive. While not every occasion of deceiving others necessarily involves impeding our own perception of reality, it is a possible consequence of deception that is often overlooked. We would do well to consider it before recommending and implementing a policy that involves deceiving others.

To conclude, although it is difficult to draw general practical recommendations from the discussion above, it is clear that deceiving students impacts both the student's and the teacher's understanding of their situation. By being deceived, students are less likely to have an accurate understanding of the situation, thus – less able to govern themselves with an accurate understanding of reality. The conjoined effect is that by deceiving, teachers would be less likely to look for and

receive accurate feedback from the students, and thus will also have a less accurate understanding of the situation.

5.7 Criterion C) Trust and Trustworthiness

As part of the discussion of criterion C) Trust of the framework in Chapter Three, I claimed that teachers have commitments of trustworthiness that are part of their role as teachers. These commitments are, at least in the most minimal sense, to be truthful unless there is a good reason not to tell the truth, to have an attitude of goodwill towards the students, and to be competent in various skills, including knowing how to balance truthfulness and goodwill. Here, I wish to consider whether teachers' deceptions are wrong, not because the nature of the act is wrong, but because they might constitute a breach of trust.

Consider again Ulysses and his crew. If they take over the ship, this is a betrayal of the trust Ulysses has placed in them. Similarly, because students place trust in their teachers' assessments of their capabilities (competence) and trust that the information teachers are communicate is accurate (truthfulness), communicating false expectations might constitute a betrayal of that trust. If this is correct, there is at least one reason why teachers should not communicate false expectations.

233

However, as explained above, teachers also have a commitment of goodwill. Therefore, at times, they might wish to communicate false expectations in order to hold that commitment. For example, deceiving a student in order to prevent her from physically harming herself or others would be an act in accord with the teacher's commitments of trustworthiness. In a case such as this, the goodwill aspect of trust overrides the truthfulness aspect of it because it is necessary in order to prevent harm, and it is reason enough to deceive in such a case.

That said, though deception harms the truthfulness aspect of trustworthiness, too much of the truth might work against the goodwill aspect of it. Telling the truth may well strengthen the student's trust that what the teacher says is what she truly believes, but it can also harm the student's well-being. Generally, the truth, the whole truth and nothing but the truth might harm those who are unable to understand it or would be severely pained by it. Telling a student your real expectations of her, especially if these are rather low, might harm the student and so be a violation of the commitment of goodwill.

For example, telling a student you do not care for her would be too much of the truth and an act that breaches teacher-student trust. As a reminder, in the discussion in Chapter Three I explained why truthfulness, goodwill and competence should not be understood as ideals or means to an end, but rather as commitments.

234

Understanding these as commitments assists in answering questions about what teachers should do in cases when these aspects of trust are in conflict. If telling the truth would seriously harm the student, then deceiving the student will not constitute a breach of trust. Hence, certain cases of communicating false expectations are justified if telling the truth will harm the student. In a case such as this, the commitment to keep a minimal level of goodwill makes deception the reasonable option.

Except in obvious cases, it would be hard to determine whether the student would be harmed by hearing the truth. This is especially true in real life situations in which, for example, the truth might harm the student's immediate well-being but would serve her longer-term interests. Therefore, whether a certain situation calls for deception in regard to teacher expectations depends very much on the specifics of the relationship: Who's the teacher? Who's the student? How long do they know each other? How much does each trust the other? Is the teacher's respectful and caring attitude towards the student well established? What would others do in such circumstances? Etc. Suppose the teacher is familiar with the student for some time and knows that she's capable in other academic or non-academic fields, that her self-esteem is high etc., then the teacher may well communicate her real expectations. This is because the student will not perceive this as uncaring or un-

235

respectful and will probably not be seriously harmed by it. Also, if the student has none of the good attributes above, but the teacher has established a caring and respectful relationship with the student and telling her the truth will not be perceived as demeaning, then communicating real expectations is the appropriate action.

Now suppose the student is one who the teacher is not familiar with, or one who, while she will probably not perform any better, needs encouragement to the extent that not showing it would be an act that violates the commitment of goodwill towards her. In these situations, the student will probably experience the teacher's communicated expectations as implying whether the teacher has goodwill towards her or not. These are good reasons that make it not only reasonable to communicate false expectations in these circumstances, but in some extreme cases even a must.

However, while a few instances of communicating false expectations do not seem to constitute a breach of the minimal threshold of truthfulness, communicating false expectations often, as a policy, by many teachers, will result in dissolving the trust that students have in the truthfulness of these utterances. Therefore, such deceptive techniques require careful consideration. In light of the discussion above of the various parts of the framework for assessing benign manipulation, we may conclude that as a policy, communicating false expectations is not appropriate.

236

I believe many teachers who communicate false expectations are doing so in order to show their care and goodwill towards the students. While we have to be careful not to abuse the trust the students place in us, I think this is a better reason to communicate false expectations than that of raising academic achievements and performance. Apart from the obligation to deceive in the extreme cases mentioned above, in other situations we should consider the specific characteristics of the situation before choosing to deceive. In each specific case, teachers who wish to uphold their commitments should prudentially determine whether the benefits to the goodwill aspect of trustworthiness are worth the possible harms to the truthfulness aspect of it. Lack of such careful consideration, deceiving as a default, will constitute a breach of trust.

5.8 Conclusion

Let us reconsider the opening scenario. The student says: "I am really bad at X", you expect she will not do any better in the future, and following the discussion of Pygmalion Effect and deception above, you now believe that deceiving her in regard to your expectations has only a small chance of improving her performance, might hinder her understanding of the situation as well as your own, and betray her trust. So, should you deceive her? I believe that by using the framework for assessing

237

benign manipulations teachers would be better placed to act appropriately in specific contexts.

The recommendations to use the Pygmalion Effect, to communicate false expectations, follow this line of reasoning: teacher expectations influence student achievement. High expectations by teachers raise the chances of students achieving more (the Pygmalion Effect). Hence, regardless of their real expectations, teachers should act as if they expect the students to have high achievements (Pygmalion Effect Knowingly). I have shown that this process necessarily involves deception as teachers are asked to deceive their students in order to have their students perform better. Thus, the recommendations of a general policy of deceiving students to believe teacher expectations from them are higher than they really are – aims at the good cause of improving students' performativity. This is a case of benign manipulation and can be assessed using the framework that was established in the previous chapters.

As was discussed above, empirical findings show the general gain in performance is likely to be small (Criterion A – Intended consequences). Also, there are good reasons to think that the long term and wide-scope consequences of communicating false expectations as a general policy are undesirable and would undermine the benefits of the technique (Criterion A – Unintended consequences). In

238

regard to epistemic considerations, though there seem to be instances where communicating false expectations might enable a teacher to have a more accurate understanding of the situation, it appears that in most cases communicating false expectation might hinder both the teacher's and the student's understanding of the situation (Criterion B – Understanding). Lastly, communicating false expectation without considering the specific context, would constitute a breach of trust, even if it is well intentioned (Criterion C – Trust). Following this discussion, we may conclude that in most cases, though not all, teachers should not communicate false high expectations to students.

However, the above discussion also demonstrates that there are many occasions where benign deception is appropriate and would be the correct thing to do. Let me consider some of the examples mentioned at the start of this chapter. In **Water the Plants,** the child is asked to water plastic plants as a way of letting out some energy and demonstrating their responsibility. In this case, the consequences of the deception appear to be good for the student, the class and the teacher. The child gets exercise, and possibly learns some responsibility. The lie touches on an aspect of reality that is probably unimportant to the child, it does not hide important truths from her. Lastly, it is questionable whether the teacher is betraying the child's trust in this case. The teacher certainly deceives the child, but it is done out of

239

concern for the child's well-being, and at least for some children in some situations, would be a clever and professional choice by the teacher. To conclude, at least in some specific contexts, even such an elaborate lie appears to be appropriate.

In **Noble Lie,** Socrates recommends deceiving the youth of the Republic about their origins by telling them a myth - they were all born from the ground and therefore, as brothers and sisters, they are to protect the state and each other. In this case the consequences of such a lie may be hard to predict and measure. However, we can speculate that such a lie would have negative consequences if it is discovered to be a lie, harming the citizens' trust in the Republic. Even if it is never uncovered, it might create unity among the citizens that is problematic. How should the citizens who were 'born from the ground' treat foreigners? Again, it is extremely hard to forecast the consequences of such a lie, but we can at least imagine such a lie as having bad consequences as well as good ones. Moreover, such a lie will seriously harm the citizens' understanding of an important aspect of their lives – their identity and origins. It will mask important aspects of reality from them. Such a lie may also harm the liar's understanding of reality, as the citizens may show loyalty to the Republic in cases where they should offer criticism of the state's actions. Lastly, as mentioned before, such a lie might constitute a breach of trust between the leaders of the state and its citizens.

240

In the case of the Noble Lie, I believe the analysis through the assessment framework suggested in this work shows that such a lie is very unlikely to be appropriate. However, the Noble Lie that Socrates proposes can be generalised to many similar falsehoods or imaginary constructions which aim to promote a sense of unity and belonging among citizens of a certain nation or state. A false historical account of the 'birth of a nation' might be such a lie. In many cases, this might be an appropriate benign manipulation. My suggestion is that to assess whether such a lie is appropriate, we should move beyond the proof that it is indeed a falsehood, and consider the three criteria suggested above.

In **Triangles,** a Grade 7 geometry teacher asks the students to memorize as fact that 'the sum of all angles in a triangle adds up to 180 degrees', even though she knows this to be true only in the specific case of Euclidian geometry. In this case, the teacher intends to simplify the topic to the level that her students will be able to comprehend it. She acts in the interests of her students by teaching something that is inaccurate and even false, but does so for their own interests, which is a case of benign manipulation. As for the consequences of her actions, it seems reasonable to assume that in this case there might be some positive consequences, such as student learning, and that extreme negative consequences are unlikely. A possible reason for this is related to criterion B – Understanding. The

sum of angles in a triangle is not an important aspect of the students' life, and so they will be less likely to 'get attached' to it. When they mature and develop their mathematical skills and understanding, they will be able to let go of the 'fact' that triangles have 180 degrees, and acquire a more subtle view that only triangles on a flat surface have that quality. As the teacher in this case is deceiving in a way that is reasonable and even required by her role and professional judgment, such a lie will often be appropriate, and will not betray the students' trust. This example may be generalized to many instances in which teachers might decide to teach simplified and inaccurate information, though each specific case should be assessed against the framework. Here again, we can see that some deceptions, understood and assessed as benign manipulation, may indeed be appropriate.

Chapter 6   Fooled by 'Smart Drugs'

6.1   Introduction

The previous chapter looked into the use of deception in education, and specifically the suggested use of 'Pygmalion Effect' – the idea that teacher expectations of student's performance can serve as a self-fulfilling prophecy. I move on to discuss the liberal use of pharmacological cognitive enhancement (PCE) in education as a manipulation designed to raise academic performance. Building on empirical research that has demonstrated that PCE has moderate benefits and poses little risks for so called 'healthy' individuals, the advisability of widespread liberal use of PCE in education has gained academic traction in recent years. The debate regarding the desirability of such measures is ongoing among scientists, bioethicists, moral philosophers and educational theorists and practitioners. It is my view that this debate can be enriched from an analysis of PCE as manipulation, similar to other educational interventions. Such an analysis places the concerns regarding PCE in a wider context, which is rarely acknowledged in the ongoing debate.

243

The reasons for viewing the use of pharmacological cognitive enhancement as manipulation stem from my understanding that a technique which is not overtly coercive and does not involve rational reasoning, is at least a good candidate to be thought of as manipulation. In the case of pharmacological cognitive enhancement, the influence that is exerted on the student clearly bypasses their reasoning processes. To be precise, one might reason with the student whether or not they should take the drugs, but once they do, the drugs are meant to influence the student's cognitive functioning and behaviour without the need to reason with the student. Moreover, when considering the widespread use of PCE, it is normally assumed that students will not be coerced to use these drugs, although they might be pressured to do so for various reasons. Therefore, the use of PCE is not coercive and the influence of these drugs is not as visible as the use of other coercive techniques. For these reasons, I believe we can benefit from analysing the use of pharmacological cognitive enhancers as benign manipulation.

The discussion below aims to address the issue of widespread liberal use of PCE. This does not have to include deceptive measures by educators and institutions in administering the drugs. That is, students may well choose to take these drugs to enhance their own academic performance, and do so freely. Unintuitively, I believe this still counts as manipulation. As a reminder, in my analysis

244

of the nature of manipulation in Chapter 2, I claimed that manipulation does not have to include deception, using the example of a teacher setting up a competition among her students. The student may freely choose to participate in the competition, but setting up the competition still counts as manipulation, in that it influences the behaviour of the students without coercing them or reasoning with them. In a similar way, I believe students may make an aware decision to take PCE, but that to allow easy access to PCE would still count as manipulation by whoever offers and allows easy access to these drugs.

To be clear, as I support a non-moralised view of manipulation, the analysis of pharmacological cognitive enhancers as manipulation does not start with assuming that there is something problematic about the use of such drugs. To establish whether or not there is a problem with such use, I will analyse the use of PCE according to the framework for assessing manipulation which was established in Chapter Four.

6.2 Pharmacological Cognitive Enhancement

Research in the domains of neuroscience and pharmacology has revealed that various drugs may have modest cognition enhancing effects when used by

245

healthy individuals (Bagot and Kaminer 2014; Maslen *et al.* 2014, p. 1). The effort to find better drugs with ever stronger effects on cognition but weaker side effects is ongoing. In the future, pharmacological cognitive enhancement (PCE) might prove to be the 'silver-bullet' enhancing educational attainment immeasurably. Imagine a classroom with no disruptions, where all students are focused and on track, measured performance exceeds expectations, and that all this is the result of cheap mass-produced pills delivered by the state to every classroom. Students would excel, parents would be satisfied, and schools would achieve their targets – why then, should we not use these drugs?

Although suggestions to distribute cognitive enhancing pills to pupils can be found in the literature (Bostrom and Sandberg 2009; Porsdam Mann *et al.* 2018; Ray 2016), the proposition to do so is generally not taken seriously by many educators and educational theorists. Nevertheless, it is analysed in some depth in this chapter for several reasons. Firstly, the use of PCE is increasing, as many students, parents and teachers are taking an interest in how pharmaceuticals improve performance. Even if it is unlikely that PCE pills would ever be freely distributed to all students, it is a concern that growing numbers of students appear to be using these drugs (Maslen *et al.* 2014; Singh *et al.* 2014). Secondly, although many parents and educators intuitively vigorously oppose the liberal use of PCE in education, there is little

246

evidence to justify this feeling of unease. For example, David Bakhurst comments,

he is "amazed at the naiveté´ of scholars who ponder the possibility of 'popping a pill

to learn'", but offers no clear argument against such use (Bakhurst 2009). He hints at

ethical issues, but previous attempts to counter PCE use on moral grounds have

proven largely unsuccessful, as will be discussed briefly below. Thirdly, the

proposition to deregulate PCE use exemplifies the extremes a purely positivistic

perspective of education might prompt. Discussions of PCE and the associated

implications for educational theory and practice expand well beyond the question of

whether or not to deregulate 'smart drugs', as will be elaborated herein.

Unlike other arguments opposing the liberal use of PCE, which have focused

on the moral aspects of such use, the objections raised in this chapter can best be

described as epistemic in origin. In that, these objections align with the epistemic

criterion of the framework (Criterion B – Understanding). In brief, the concern

expressed here is twofold. First, following Charles Taylor (Taylor 1985), I argue that

one's emotions, and the related behaviour and thoughts associated with those

emotions, are linked to an interpretation of one's situation. Therefore, emotions and

behaviour can offer clues regarding the situation that a person is experiencing, to

benefit both herself and others. Second, I suggest that liberal PCE use might

influence one's emotions and behaviour in a way that would make it more

247

challenging for those involved to interpret a situation and formulate an accurate understanding of it. Although the latter claim would probably be best supported by empirical evidence, this study offers some theoretical explanations for why this is likely to be the case. I conclude that because the constant interpretation of self and others plays a key role in educational relationships, deregulation of PCE resulting in increased use of pharmaceuticals might distort the understanding of those involved. Due to the fact that an adequate understanding of the educational situation is important for the educational process to be both pedagogically effective and ethically praiseworthy, we should be very careful about using PCE, especially because interpreting behaviour is challenging anyway. We require further research into the effects of these drugs in real life contexts to be able to judge whether it is appropriate to deregulate them or not.

The discussion will focus on pharmacological cognitive enhancements, nicknamed 'smart drugs'. These drugs have attracted attention in recent years, both empirically and philosophically. However, there are other pharmacological behaviour modifiers relevant to the educational context, such as anti-depressants to enhance well-being, psychedelic drugs that enhance creativity, and MDMA to enhance empathy. With minor adjustments, the argument presented here is applicable to any

wide scale pharmacological intervention in education, with regard to how the intervention affects the emotions and behaviour of those taking the drugs.[29]

The use of PCE is already wide-spread, and there are three main PCE substances in common use today: methylphenidate (e.g. Ritalin, Concerta) and related compounds, mixed amphetamine salts (sold in the US under the name Adderal), and modafinil (Provigil). Methylphenidate (Ritalin, Concerta) and mixed amphetamine salts (Adderall) are psychostimulants and are the most widely used pharmacological treatments for Attention Deficit/Hyperactivity Disorder (ADHD). Modafinil is used as treatment for narcolepsy and sleep disorders (Singh *et al.* 2014, p. 1). According to recent data, the use of these drugs in educational settings is rising. In the US, the prevalence of children aged 4-17 taking ADHD medication has grown from 4.8% in 2007 to 6.1% in 2011 (CDC 2017). Alongside evidence that prescribing is also increasing, various estimates have been made suggesting the non-medical use of PCE by students in US universities ranges from 5%-30% (Maslen *et al.* 2014). In the UK, a survey of university students found that lifetime

---

[29] I am assuming that enhancing cognition entails a change in behaviour. This is not necessarily true. But, if it does not manifest itself in behaviour (better concentration, memory enhancement, improved exam scores, etc.) it will probably have no desirable effect whatsoever. And if that is the case, then what it the value in taking the drugs? See also Vrecko, Scott. 2013. 'Just How Cognitive Is "Cognitive Enhancement"? On the Significance of Emotions in University Students' Experiences with Study Drugs', AJOB Neuroscience 4 (1): 4-12.

prevalence of PCE use was 9.4% and past regular and current PCE users comprised 0.3%-4% of the population surveyed. Researchers have also reported that about a third of students express interest in PCE (Singh *et al.* 2014, p. 1).

University students' attitudes towards PCE align with academic interest in the effects of PCE when used by so called "healthy individuals". For example, a group of leading researchers wrote a commentary article in the journal *Nature* (Greely *et al.* 2008), calling for an evidence based approach to the use of PCE by healthy people. They claimed there is no principled argument against the use of cognitive enhancers, and that they should be treated like other technologies, maximizing benefits and minimizing scope for harm (Greely *et al.* 2008, p. 702). In the authors' view of the future, if the safety and long term effects of PCE are properly researched and evaluated, school aged children could be encouraged to use the drugs quite freely, because 'liberal use of cognitive enhancers would be expected to encourage classroom order and raise standardized measures of student achievement, both of which are in the interests of schools' (Greely *et al.* 2008, p. 704). Others have maintained that if there is potential for extensive societal benefits resulting from effective cognitive enhancement, the state would be well advised to subsidize PCE for the poor and socially disadvantaged (Bostrom and Sandberg 2009, p. 1234; Porsdam Mann *et al.* 2018; Ray 2016).

250

In what follows, I will briefly assess some of the principled objections to PCE use. If any of these is correct, it would mean that PCE use should be considered undesirable, at least prima facie. I will not provide a detailed discussion of these objections for various reasons. First, as I assume a non-moralised understanding of manipulation, I wish to extend this assumption to PCE in order to see where this assumption might lead, even if it is unintuitive. Second, proponents of liberal PCE use have counterarguments to these principled objections. Even if these counterarguments are not fully convincing, I would like to suppose that they are, in order to provide an argument that would be more compelling even for proponents of PCE who accept these counterarguments as convincing. The principled objections to PCE that will be discussed are that PCE are a form of social control and that they harm people's autonomy, that they are an unnatural intervention, that they are socially unjust, and that they undermine the value of human effort.

Social control and authenticity

Some of the principled objections assume that drugs enhancing cognitive skills are an intervention that requires the exercise of too much power by educators, parents and institutions, making them a form of social control. This is sometimes considered to harm the 'autonomous', 'authentic', and 'natural' realisation of the student's desires, feelings and beliefs, or that it involves treating students only or

251

primarily as a means, rather than as ends in themselves (Stein 2010; Stein *et al.* 2011).

The counter argument suggests the opposite, i.e. that PCE could potentially strengthen the capacities needed for autonomous agency, enabling a person to act more in line with reasoned judgment. Therefore, the use of these drugs might promote autonomy, rather than damage it (Bostrom and Sandberg 2009, pp. 326-27). Accordingly, Bostrom and Sandberg (2009) further claim that children only attend educational establishments because they are coerced by parents and other adults into developing certain attitudes, skills and knowledge. For example, contemporary society expects children to become literate and numerate. Most of us consider education a benign form of coercion, and if this is the case, PCE to develop requisite skills and traits should be deemed acceptable as an extension of that agenda.

Naturalness

Similarly, another objection originates from the idea that the use of PCE is unnatural. However, the line that divides what is natural and unnatural is unclear. We are using various technologies that might not be considered natural, like reading and writing, drinking coffee, electronic devices, etc. Even agriculture may be considered unnatural in some ways. Moreover, even if we could draw the line between what is

252

natural and what is unnatural, it is not clear if this division has any normative significance. We might produce potent poisons 'naturally', or else use very 'unnatural' life saving technologies. Whether something is natural or not does not yet tell us if it is right and appropriate to use.

Social justice

Further objections to the liberal use of PCE proceed from social justice concerns. Typically, it is argued that PCE might be a form of cheating, similar to performance enhancing drugs in sports. In addition, it is claimed that PCE might threaten equality of opportunity, because it advantages those privileged individuals who can access these drugs more readily. In response to this concern, the supporters of PCE respond that 'cheating' only becomes a problem if we ban some students from using the drugs and not others (Porsdam Mann *et al.* 2018). If all students have access and are allowed to use PCE, then those who choose to use it would not be cheating. They liken this to allowing all students to use calculators, instead of only allowing some to do so. To the charge that this approach benefits those who are already privileged, the supporters of PCE argue that this is why we should promote liberal, possibly state funded, use of such drugs to promote equality, not harm it (Porsdam Mann *et al.* 2018). Keisa Ray specifically called for exploring whether the use of PCE 'could be a means of remedying underprivileged children's

253

experiences of social inequalities that are born from inadequate schools for the sake of increasing their chances for opportunity and well-being' (Ray 2016).

The Value of Effort

Another objection to PCE rests on the idea that PCE undermines the value of human effort. Because we value effort, so the argument goes, shortcuts such as PCE undermine the value of any results achieved that were aided by these measures. Similarly, achievements in sports which are aided by drugs, would not be considered as valuable, partly because they did not require the same level of effort they would have required without use of such drugs. However, supporters of PCE use claim that what we value in effort is not just the difficulty, or else we can dig holes and fill them back in. Unlike sport, in which the nature of the task remains the same, cognitive enhancement might be used to direct a person to other tasks. What we value is not plain effort, but effort that is directed towards certain goals. For example, we do not require athletes to run barefoot, even though that might require extra effort on their part. Therefore, PCE might not undermine the value of effort, as it would assist the individual in directing their effort towards more meaningful challenges:

> In many elementary schools, calculators are disallowed in mathematics lessons, where the goal is to understand basic

254

arithmetic, but they are allowed and increasingly necessary in the higher grades. The basics have by then been mastered, and the goal becomes to understand more advanced topics. These examples illustrate that cognitive enhancement aimed at extending and completing a person's talents may promote authenticity by offloading irrelevant, repetitive, or boring tasks and enabling a person to concentrate on more complex challenges that relate in more interesting ways to his or her goals and interests. (Bostrom and Sandberg 2009, p. 326)

I believe that at least some of the ethical concerns raised should be taken more seriously by the supporters of liberal PCE.[30] However, it is important to also explore whether, even if liberal PCE use is physiologically safe, just, fair and effective, it might still come with a price. In light of the above, I will assess the suggestion for liberal use of PCE in schools by using the manipulation assessment

---

[30] That said, in a recent paper which looks more closely into the reasons that might prohibit the use of 'smart drugs' by healthy individuals in school exams, Zdenko Kodelja arrives to the conclusion that the arguments against what he calls 'intellectual doping' are not sufficient to prohibit the use of "Smart Drugs". See: Kodelja, Z. 2021. 'Intellectual Doping and Pharmaceutical Cognitive Enhancement in Education: Some Ethical Questions', *Journal of Philosophy of Education* 55(1): 167-85.

framework that was established in Chapter Four. That is, by looking into the intended and unintended consequences of such use (Criterion A – Consequences), the effect such use will have on students, teachers and policy makers understanding of their situation (Criterion B – Understanding) and whether the use of PCE is appropriate within the norms of the educational relationship (Criterion C – Trust).

6.3 Criterion A) Consequences

Even-though pharmacological cognitive enhancement has received much interest from philosophers and scientists, it is still unclear how effective these substances are: "[W]here there is evidence of enhancement effects, they often tend to be limited to improvements on specific tasks, are only seen at certain dosages and are not observed in all people" (Maslen *et al.* 2014, p. 3). While varying from substance to substance, it appears that the effectiveness of PCE is limited (especially for 'healthy' individuals) but significant (especially for those on the 'lower' side of the cognitive spectrum). Hence, one reason against using these drugs, at least for now, is a lack of empirical evidence regarding their effectiveness (Hildt 2013).

A related issue is whether there are medical safety concerns regarding PCE use. The empirical research shows that this varies from substance to substance. However, generally the drugs mentioned above are well-tolerated by users, though some side effects were discovered (Maslen *et al.* 2014, p. 4). It should be noted that there are no longitudinal studies examining the long-term effects of the drugs mentioned above (Maslen *et al.* 2014, p. 3). Also, these results are currently relevant to laboratory settings, and not to real-world contexts. In other words, issues such as psychological effects, addiction and long-term physiological effects were not researched.

The cautionary empirical results did not prevent some PCE enthusiasts from arguing in favour of liberal PCE use. They maintain that PCE use should be assessed by weighing positive aspects of it against the negative aspects of it. Therefore, if PCE is effective in improving some cognitive process, it should be used to enhance those. Furthermore, at least in their view, even if current drugs might not be effective yet, future research will bring about the discovery of such substances (Bostrom and Sandberg 2009, p. 334).

Relatedly, on the issue of negative consequences, proponents of PCE make an interesting point. They observe that other human activities are risky, including interventions meant to influence others:

257

Even education is a risky enhancement method. Education can enhance cognitive skills and capacities, but it can also create fanatics, dogmatists, sophistic arguers, skilled rationalizers, cynical manipulators, and indoctrinated, prejudiced, confused, or selfishly calculating minds. Even high-quality education that includes training in formal methods and critical thinking can have problematic effects. (Bostrom and Sandberg 2009, p. 322)

Bostrom and Sandberg are relying on a concept of education that holds no normative value. They appear to refer to education from a sociological or anthropological perspective, which would encompass many aims and means, some perhaps immoral. Many philosophers of education hold a different concept of education which does hold normative value. For example, R. S. Peters maintains that education 'implies that something worthwhile is being or has been intentionally transmitted in a morally acceptable manner. It would be a logical contradiction to say that a man had been educated but that he had in no way changed for the better, or that in educating his son a man was attempting nothing that was worthwhile' (Peters 1966, p. 25). Peters' understanding of education would exclude some of the consequences that Bostrom and Sandberg argue are risks of education. If indoctrination is conceptually different from education, i.e. it differs from 'transmitting

258

something worthwhile in a morally acceptable manner', then being indoctrinated would not be a result of education, but of indoctrination.

However, even if we hold a concept of education that has a normative value, we cannot assume in advance that a certain method is immoral, and then argue against its use because education involves influencing others in a morally acceptable manner. We are required to explain why a certain method is immoral, before we can argue that it should not be used in education. Therefore, I believe the unintuitive thought that education is a 'risky enhancement method' can be explained by the idea that PCE should not be assumed to be immoral just because it involves some risk. In other words, proponents of liberal PCE use claim there is no principled argument against the use of such substances, even if they have some risks. As unintuitive as it is, we should bear in mind that many activities are not risk free, such as dangerous sports, cosmetic surgery, or just driving. The fact that something involves risk does not necessarily make it immoral. Indeed, certain individuals might find PCE valuable to the extent that they will be willing to take the risks involved, even if there are risks (Maslen *et al.* 2014, p. 4).

6.4　Criterion B) Understanding

6.4.1 Behaviour as evidence

The arguments in favour of liberal PCE use listed above, demonstrate that if one is interested in the narrow aim of success as measured in assessments of academic performance, then there is little reason to place importance on examining emotions and behaviour, such as a lack of motivation or distraction. Behaviour is taken to be effective or ineffective, and therefore, not something to be considered as possibly valuable evidence. More generally, disregard for the importance of interpretation is part of a positivistic approach towards educational research and practice, which represents a leading view of education in many modern societies (Hyslop-Margison and Nasee 2007).

The positivistic approach assumes educational processes can be researched in a manner similar to the natural world, using comparable methods and techniques. This is because, on this view, educational processes are law-like and mechanistic. Building on those assumptions, it offers a theory of what a student is (a cognitive being) and what the desired aims and goals of education are (learning as it is measured through performance on exams). Although complex, the person is assumed to be an individual object which can be acted upon to leverage the desired

260

effect.[31] This attitude is apparent in how supporters of PCE refer to human cognition as an information processing system (Bostrom and Sandberg 2009, p. 312). Even though a positivistic approach has proven very successful in the domain of the natural sciences, ample arguments and evidence suggest it is less appropriate in the educational sphere.[32]

An alternative view to the positivist approach is the interpretative one, as articulated by Charles Taylor (1985). Taylor establishes the unique ontological status of human agency, following Heidegger's argument that our emotions (and hence our agency) are always coloured and structured by interpretation, meaning man 'cannot be understood simply as an object among objects, for his life incorporates an interpretation' (Taylor 1985, p. 75).

---

[31] It might appear as if supporters of liberal PCE use violate the Kantian maxim by treating others merely as a resource. But, at least in their view, this is done to benefit students by serving their interests, and thereby not treating them as mere means, but also as ends in themselves. In the view I wish to promote, the morally suspect attitude is that of not respecting the students as valuable (even if not equal) sources of information. This would be disregarding the behaviour, feelings and thoughts of the students when deciding their interests. PCE may well improve academic performance, but it might also affect other areas of interest, such as food availability, social justice, inspiration towards worthwhile life pursuits, etc.

[32] Cohen et al. 2007, p. 11, put it thus: "Where positivism is less successful, however, is in its application to the study of human behaviour where the immense complexity of human nature and the elusive and intangible quality of social phenomena contrast strikingly with the order and regularity of the natural world. This point is nowhere more apparent than in the contexts of classroom and school where the problems of teaching, learning and human interaction present the positivistic researcher with a mammoth challenge". See also Biesta 2007 and Hyslop-Margison and Naseem 2007.

Taylor supports his argument regarding the role emotions play in human agency by providing a phenomenological analysis. In brief, he claims that emotions are not 'merely subjective projections but cognitive modes of understanding that grant access to certain features of reality' (Blakely 2013). In other words, emotions communicate something of importance about the reality of which we are a part. Moreover, our experience of being motivated to do something, or even not to do it, is experienced as an emotional response (Blakely 2013). This requires interpretation, because even the most basic emotions we experience need to be decoded in the context of our experience of the world. For example, consider the pain we feel when pricked by a needle. We have a completely different emotional experience if we interpret the situation as arising from a moment of care, as when we are given a vaccination, or if it is one of abuse, being pricked by someone intending to inflict harm. Thus, interpretation necessarily affects our emotions, and our emotions inform our experience, motivating us to act in a certain way. In the care scenario, we try to contain our pain and be grateful to the other person, whereas in an abuse scenario we will resent or fear the other person. Hence, emotions arise from our interpretation of reality and tell us something about our understanding relative to the experience.

There are good reasons to reject at least parts of Taylor's overarching theory.[33] Moreover, it is probably wise to seek a balance between the positivistic view of education and the interpretivist one. Undoubtedly though, not attaching value to the interpretation of self and others in the context of education is unreasonable. In fact, I believe it is a common-sense way of thinking about educational relationships. Experienced educators can read emotions, behavioural and other cues proficiently, and form judgements accordingly. If they do not do so, the first line of 'criticism' and feedback they receive takes the form of disruptive and uncooperative student behaviour. This 'feedback loop' is important in educational relationships, as it might inform the educator about which are appropriate aims to set, and which are the proper means to achieve those aims. For example, the uninterested demeanour of a student might tell the teacher a number of things, e.g. that she has not understood something, that she did not have breakfast, or that she is being bullied. At least in some cases, the interruption should prompt the teacher to investigate what is going on and pay attention to it.

To clarify, I am not claiming that emotions and behaviour are always effective clues to a situation, or even that an accurate understanding of the situation is

---

[33] For example, it might be true that some animals can interpret and express what Taylor claims to be distinctly human emotions. I owe this interesting insight to an anonymous reviewer of the journal Ethics and Education.

263

necessarily of overriding importance in the context of education. In fact, sometimes the behaviour, feelings and beliefs of students are of little importance to educators. If a student pulls a face or chats with her friend in class, this can be just something to be managed, and not an important and meaningful signal of a deeper issue. Moreover, sometimes it makes sense for educators to dismiss emotions and behaviour even when they imply certain facts about a situation. For example, I can imagine a loving parent saying: "I do not care if you're upset, we are going to climb that mountain" to promote resilience and self-confidence in her child. However, we should not downplay the significance of ongoing interpretation either. Depending on context, the emotions and behaviour of a student might indicate a problem demanding further attention from either herself or others.

6.4.2 The manipulee's and manipulator's understanding

Thus far, I have provided reasons to support the claim that the interpretation of one's own and others' emotions and behaviour plays a key role in contextual understanding. In view of this, this section examines the claim that PCE will impede the interpretation of what is really going on, obscuring issues of significance in the education setting.

264

We can consider an imaginary scenario in which an educational researcher wishes to understand which factors correlate with, cause and ultimately promote academic success in a certain school, with only two classes. One class in the school regularly use PCE pills (PCE class), while the students in the other class do not (non-PCE class). The specific research focuses on four factors that might affect academic success: 1) Do students regularly eat a good healthy breakfast? 2) Do teachers in the school provide high quality instruction? 3) Do students in the school value academic knowledge? 4) Does the socio-economic background of the students influence their performance? The researcher plans to conduct surveys and interviews with both students and staff, to observe lessons in the school and analyse examinations and other assessment instruments. However, her funding only allows her to research one of the two classes. To have a better understanding of how the four different factors influence academic performance, which of the two classes would she find most appropriate to conduct her research with?

She judges that in the case of the students who use PCE pills it will be more difficult for her to discover the effects the above factors have on performance, because it would not be possible to know whether students are focused and motivated due to one or more of the factors above, or if they are focused and motivated because they took a pill, or a combination of both. Therefore, the reason

for her decision is mostly epistemic; that is, she would have a better understanding of the educational situation if she were to conduct her research on non-PCE users. Although still a resource-consuming and challenging task, it would be easier for her to study the non-users because the pills add a further variable complicating interpretation of the situation. I suggest that this would be the reality for many educators, who base their pedagogical and moral reasoning on their interpretations of the educational situation. Similar to the researcher, their resources are limited, and judgments are often made instantaneously. This is also true for the students, who will have an important source of information about their experience of the world, their own emotions, distorted.

Supporters of liberal PCE use might claim that it will not have a significant effect on students' behaviour, and that any change would not make it harder for educators to interpret nuanced cues in class. I will discuss two possible theoretical responses to this claim and detail some initial empirical findings that appear to undermine their claims. First, if the effects of the pills on emotions and behaviour are insignificant, then why use them? These drugs are often considered under the umbrella of 'behaviour modifying drugs', precisely because they modify emotions and behaviour. Second, supporters of deregulating PCE acknowledge that current PCE drugs offer only modest cognitive improvement and are basing their views on

266

an unproven prediction that drugs will be more effective in the future. More effective pills would surely create a more significant change in students' emotions and behaviour, which would correspondingly make it more difficult for teachers to interpret them properly. This is also the difference between PCE and other unproblematic forms of mild cognitive enhancement, such as good nutrition or coffee; i.e. if the degree to which the substance influences behaviour is small, it will not distort interpretation palpably. However, the greater the effect on behaviour, the more significant the interruption. Effective PCE drugs threaten to be powerful behavioural modification substances.

Notwithstanding the above, the question of whether PCE pills affect behaviour, impeding understanding of a situation by those involved, is ultimately an empirical one. To date, the majority of the empirical research available regarding the effects of PCE has been quantitative and performed in labs. Supporters of PCE extrapolate the predicted effects of PCE to the world beyond the lab (Greely *et al.* 2008). The predicted effects of these drugs (i.e. improved concentration) are what makes them desirable according to supporters of deregulation. Even the more cautious researchers recommend further neuroscientific research to deal with the concerns these drugs raise (Maslen *et al.* 2014). From a purely positivistic perspective this makes sense, as it assumes that the law-like connections and

267

predictable effects discovered in lab conditions are readily transferable to real world educational settings. However, the anticipated consequences of using the drugs are typically identified in experimentation, partly because that is the aim of this type of research. It might well be the case that other significant aspects of drug use are less predictable, however they are nevertheless important to understand. As an example, two pints of beer will most likely have a detrimental effect on my motor skills, and this is something that can be proven in a clinical experiment. However, whether I feel very good about myself or sit gloomily and depressed in the corner of the pub, is less predictable and reliant on many other factors. If I wish to choose whether to drink on a certain evening or not, the clinical effects of alcohol are of importance, but so are factors more specific to me in the current context. It might well be the case that PCE use will change behaviour in ways that are less predictable than suggested under laboratory conditions.

To date, PCE research has been largely silent on issues of interpretation. One intriguing exception is work by Ilina Singh (2005; 2014) who writes about children diagnosed with ADHD and their families. Her findings suggest that ADHD medication adds a further layer of complexity to interpreting behaviour. She finds that

> Parents of children taking stimulant medication as treatment for ADHD spontaneously deploy authenticity in reasoning about, and

268

in justifying, dosing decisions. Their reasoning about the relationship between authenticity and medication is inconsistent, however: on the weekend, withholding medication allows a child to be 'free to be who he really is'; during the week, providing medication allows a child to 'know who he really is—a successful learner.' …even young children, whose self-concept is just emerging, will say, "I do not feel myself" in describing side effects.... In some cases of highly aggressive children with ADHD, a sense of the 'true self' is bound up with very negative self-attributions ('inside I must be evil')... In these cases, stimulant medication is viewed as helping to restore a positive sense of self through a greater ability to pause, reason and act; this reduces the experience of being, "not me; like not knowing what I was doing or why I was doing it." (Singh 2014, pp. 237-38)

These findings suggest that both in terms of the self and others, the use of ADHD medication significantly changes behaviour creating confusion over whether to attribute one's behaviour to one's self, or to the medication. The complexity and contradictions involved in interpreting the behaviour of children on ADHD medication, by both the children themselves and their parents, offer initial empirical support for

269

the claim that PCE pills will have a detrimental effect on the ability to interpret the emotions and behaviour of those who use them.

To reiterate, interpretation of the self and others is a challenging task, and both children and adults sometimes mask important facts and feelings, even without the use of measures such as PCE. However, I hypothesise that PCE adds another layer of difficulty for whoever is trying to understand what is really going on. Therefore, even if the moral concerns regarding liberal PCE use are answered, there is an epistemic worry that we should be mindful of. In what follows I will discuss in greater depth how liberal use of PCE might distort students', teachers' and policy makers' interpretations of the requirements of them within the learning environment.

To illustrate this point, we can refer to George Orwell's experiences as described in his memoir essay 'Such, Such Were the Joys', in which he writes about his years as a student at a preparatory school:

> A child believes that the school exists to educate and that the school-master disciplines him either for his own good, or from a love of bullying. Flip and Sambo had chosen to befriend me, and their friendship included canings, reproaches and humiliations, which were good for me and saved me from an office stool. That was their version, and I believed in it. It was therefore clear that I owed them a vast debt

270

of gratitude. But I was not grateful, as I very well knew. On the contrary, I hated both of them. I could not control my subjective feelings, and I could not conceal them from myself. But it is wicked, is it not, to hate your benefactors? So I was taught, and so I believed. A child accepts the codes of behaviour that are presented to it, even when it breaks them. From the age of eight or even earlier, the consciousness of sin was never far away from me. If I contrived to seem callous and defiant, it was only a thin cover over a mass of shame and dismay. All through my boyhood I had a profound conviction that I was no good, that I was wasting my time, wrecking my talents, behaving with monstrous folly and wickedness and ingratitude — and all this, it seemed, was inescapable, because I lived among laws which were absolute, like the law of gravity, but which it was not possible for me to keep. (Orwell 1953)

Orwell was offered a place in the school because the owners of the school, nicknamed Flip and Sambo, believed he was deserving of a scholarship as his intellectual capabilities meant he would boost their reputation. In the quote above, Orwell describes a conflict between his deduction that Flip and Sambo are acting in his best interest, and his experience of strong emotions against them.

271

'Smart drugs' seems like the perfect solution for someone like young George, as the pills might have helped him to concentrate, do well within the rigid guidelines of the school, releasing him from emotions such as hate, shame and dismay. But this is precisely what is worrying about PCE. When taking PCE pills Orwell would not be escaping from the truth that he was still being used by the school owners for their selfish ends. His subsequent realization that this was what was going on when he was a child, emerged from the strong feelings he had at the time. His later, and more accurate, interpretation of the school owners' intentions might not have been achieved had he taken pills to numb his feelings.

The concern here is that if a student were being manipulated to do something while using PCE this might cloud their judgment about whether they were being forced and controlled. These pills are experienced as 'medication' or 'enhancement' rather than as coercion. The feelings of hate and contempt that Orwell felt towards his benefactors were not identified at the time but were described very strikingly many years later. This might not have been possible if he had been taking Ritalin, Adderall or any other PCE. Regardless of his childhood questions, later in life, he judged the emotions he felt to be accurate, and confirmed his supposition that the school owners were not his benefactors. Thus, if PCE does influence emotions and

272

behaviour, it might have distorted Orwell's understanding of reality, removing a significant inner conflict.

If the student does not feel that a certain educational technique is oppressive or shaming, then this might lead to claims that it is permissible. To an extent, I agree with this line of reasoning. However, the ethical issue here is intermingled with the epistemic challenge of understanding the situation. The use of PCE might hide some important aspects of the situation from the student; for example, the extent to which they are being used for someone else's purposes. Thus, by encouraging PCE use, we might damage students' freedom to evaluate their reality. This would then make it much harder for students to understand the situation they are in, both at the time and when reflecting on it many years later. In Orwell's case, PCE use would have prevented him from learning a lesson about manipulation that seems to be much more important than any Latin class.

Less extreme examples are also of interest here. If PCE enhances cognitive processes, for example concentration, this might blur the differences between teachers, and subjects. If a student finds it easy to concentrate on her work with two teachers using very different teaching styles, this might harm her innate understanding of which style is best suited to her. Similarly, if she finds it easy to

273

concentrate in both mathematics and English, this might prevent her from knowing she likes mathematics better.

Correspondingly, something similar might happen if we view the effects of PCE use from the perspective of the teacher. If to deal with the constant itch in my arm, I apply a local anaesthetic, I lose valuable information in the process. Sometimes it might be the right thing to do, but not always. The itch might be a signal that something more serious is happening that I need to have checked. If to deal with unruly and noisy children in the classroom they are given pills to make them behave the way I wish, I am also losing valuable information. There are countless reasons for a child to behave in the way they do, and it is partly the educator's role to actively interpret this situation. Educators are usually not entrusted just to teach students, but also to care for them. If a hungry or abused child remains calm because she is using PCE, then it is harder for the teacher to intuit an important part of that child's reality.

One might claim that this is asking too much of teachers, and that we should not require teachers to do anything more than 'just teach'. However, even if we assume that the teachers' role is solely to facilitate learning, then we must still accept that they need to engage with countless additional environmental factors, from the availability of clean air to breathe to the feelings and thoughts of their students. To

274

'just teach', teachers need to actively interpret the responses of their students, which might be distorted if they were to use PCE. Often, students become restless when they do not understand what is being taught. On PCE pills they might not exhibit restlessness, rendering it more difficult for the teacher to know whether they have 'got it' or not. Many pedagogical challenges that should be addressed are likely to be simplified to the question of whether a student has taken her pills today or not.

Policy makers tend to view education through the prism of exam results and other performance measures. These results are often seen as targets to be achieved, and not as a possible indication of other important aspects of students' lives. Liberal use of PCE would exaggerate the existing trend, by removing even more important information about the reality of students' lives inside and outside of school. This can happen in at least two ways. The first was discussed above, that other problems and issues might be concealed behind the use of PCE. For those rare policy makers that think of exam results as possible indicators of more serious problems, liberal PCE use will contribute a further layer of difficulty because it has the potential to disassociate results from important aspects of students' lives, such as socio-economic disparities, racism, etc.

The second concern expressed from the perspective of a policy maker is that PCE might be another method of gaming the system. Christopher Hood describes

275

the example that setting pupil-attainment test scores has lead teachers in some schools in the UK:

> …to concentrate on a narrow band of marginal students who were close to the target thresholds and to give proportionately less attention to those at the extreme ends of the ability range or to aspects of education beyond preparing students for those particular tests that figured in target regimes. (Hood 2006)

This is a specific case of the more general 'Campbell's Law', which states: 'The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor' (Campbell 1979, p. 34). It is often overlooked that when a certain measurement becomes a target, there is then pressure to aim for the target in a way that might corrupt what the original measurement was meant to measure.

As instrumental economic considerations gain further authority in educational contexts, the pressure to use PCE will grow, as it is likely to be a very cost-effective method of raising exam results. Compare the cost of the production and distribution of mass-produced pills with other measures for 'enhancing cognition', such as smaller class sizes, offering competitive teacher salaries, giving state funded meals

276

to students, updating IT systems, etc. PCE might therefore become another way of 'gaming the system', similar to educational methods such as teaching to the test. Certainly, PCE might make it easier for students to learn how to get a higher grade on an exam, but this might not be the most valuable and worthwhile part of education. For example, if we assume that PCE improves one's memory, students might be inclined to memorize answers, instead of researching and answering for themselves.

Hence, students might know less relevant information, but still succeed in achieving higher attainment on testing measures. Thus, PCE will distort policy makers' ability to achieve the goal of educating the next generation to serve as useful members of the workforce (assuming this is their aim). It is imperative that supporters of the liberal use of PCE not only prove that PCE improves performance in exams, but that it genuinely promotes the acquisition of knowledge to benefit students' future lives. In other words, PCE should be proven to do more than just help students navigate the stressful classroom environment that we have created for them (think again about the owners of the young Orwell's school, who cared only about results).

277

6.5 Criterion C) Trust

As the use of PCE does not necessarily require a relationship, the issue of trust in not immediately relevant for the assessment of such a technique. None the less, one issue that might be problematic in relation to trust is that the use of 'smart drugs' by students would be encouraged by various agents, such as for-profit companies, states, headteachers, teachers and parents, supposedly because it is in the interests of the students, while the reality will be that it is more in the interest of the other stakeholders. Such a promotion of the use of PCE will be deceptive and can be assessed by the framework, much like other deceptions discussed in the previous chapter.

As regards teacher-student relationships, if it is correct that teachers should maintain a certain level of their commitments to be truthful, caring and competent, then it is worth thinking whether presenting education in positivistic terms breaches these commitments. In disability studies, a distinction can be drawn between a medical model of disability and a social model. Jonathan Wolff summarizes the difference between the two models. In the medical model, he writes, "disability consists purely of a physical or mental impairment." In the social model, on the other hand: "disability is, at least to a high degree, 'socially constructed', in the sense that

278

the way we have fashioned the world, both physically and conceptually, makes it more suitable for some groups than others" (Wolff 2012, pp. 149-50). To use simple examples, a paradigmatic case explained by the medical model is short-sightedness. Some people have a condition in their eyes that causes them to see far away objects as blurred. The solution is handy - they can wear glasses. A paradigmatic case explained by the social model might be left handedness. The extent to which left-handedness is a disadvantage is dependent on social norms and human decisions. To take a minor example, the London Underground gates have their ticket verification on the right side of the gate, which makes it easier for right-handed people to use. Significantly more problematic was forcing young children in schools to write with their right hand, even if they were left-handed, which was a common practice a few decades ago. It seems that many of the recommendations to use smart pills assume something like the medical model of disability. A lack of high functioning in the school environment is seen as a disability, and smart pills solve the issue. Under this model, low grades are similar to short-sightedness. Opponents of smart pills might adopt the social model of disability. They can claim that it is true that some students are less attentive or get lower grades than others, but what makes this a disadvantage is how we constructed our social environments in the first place. Their proposed solution to the problem might be to change some of the environment factors that make these attitudes and behaviours disadvantageous.

279

In relation to trust, I believe the problem is that proponents of liberal PCE use do not take seriously enough the possibility that the value of academic performance is at least in part socially constructed, rather than valuable in and of itself. If it is at least in part socially constructed, then not communicating this to students might constitute a breach of trust on the part of educators. Suppose we had a pill that would make left-handed people permanently right-handed. If an educator is not whole heartedly convinced based on sound reasoning that being left-handed is a problem that should be fixed, then withholding that information may indeed constitute a breach of trust.

6.6 Conclusion

The analysis of liberal PCE use as benign manipulation highlights problems with such a practice that might not be visible otherwise. Mostly, I have argued that, although PCE drugs appear to be safe, there is currently little evidence that they are effective in improving cognitive performance for 'healthy' individuals even in laboratory settings. Furthermore, their safety and efficacy in the long term and in real life settings have not been researched yet. However, even if more effective drugs with almost no side effects were to be developed in the future, we should still be critical towards liberal use of PCE on the basis that such use distorts the accurate

understanding of what is occurring in the educational setting for students, teachers, parents and policy makers (Criterion B – Understanding). The epistemic argument concerns the degree to which we might resort to PCE as a solution to boost learning outcomes, as opposed to enforcing a categorical prohibition against it. It appears from the evidence presented herein that large scale use of effective PCE drugs represents a serious threat for society. Questions such as: do PCE users tend to be more submissive to authority figures? demand further research, probably by using a qualitative methodology, as many of the issues that interact with PCE use are not readily quantified and measured.

Relatedly, evaluating educational contexts in purely positivistic terms leads to a very limited understanding of them. Indeed, some aspects of students' lives are law-like and mechanistic, in that they should be managed and 'engineered' to ensure the availability of nutritious food, clean air, or prescription glasses. However, other aspects require the ability to interpret students' emotions and behaviour to access the significant data present in a situation. Part of the educational process involves discovering what the problems and issues are, not simply solving difficulties conceived of in advance. I have claimed that a decision to choose whether to use PCE or not, should be made by considering that the value of academic performance is at least in part socially constructed. Therefore, a parent or a teacher who presents

281

and recommends PCE use in purely positivistic terms, might be misleading those under their care, in a way that breaches the commitments of truthfulness, goodwill and competence.

If we see the main problem of using PCE as an epistemic one, this might explain some of the arguments against liberal use of PCE that were discussed earlier. For example, the claim that PCE creates an inauthentic learning environment arises from the concern that PCE use will change students' behaviour in ways that make it harder for others to interpret their needs in context. However, sometimes it would be justified to use PCE; i.e. in cases where a specific student might be seriously harmed by not using PCE, as there is less epistemic concern in such cases regarding the obfuscation of important communicative behaviour. If evidence proves this argument to be valid, then the recommendation would be that PCE use should be well regulated, like other pharmaceutical drugs. Therefore, at times, for specific individuals, using PCE is likely to be overall beneficial, and not harm their understanding of important aspects of their reality, as well as the understanding of those who educate them and care for them, and in those cases PCE use is appropriate. In a sentence, the argument presented in this chapter makes a negative verdict against liberal, widespread use of PCE, but supports a well-regulated minimal use of such substances.

Chapter 7   Gamification

7.1   What is digital gamification?

Games have been hailed as the ideal learning environment (Kapp 2012, p. xxii). The underlying assumption is that games offer engaging, motivating and safe learning opportunities. For this reason, there has been an ongoing interest in using certain aspects of games in order to promote learning and other educational aims. The use of game aspects to influence behaviour in the world outside of the game, has been termed 'gamification' (Deterding *et al.* 2011). Kapp, an avid supporter of gamification, provides the following definition for it:

> Gamification is using game-based mechanics, aesthetics and game thinking to engage people, motivate action, promote learning, and solve problems. (Kapp 2012, p. 10)

He further explains the different elements of the definition. A 'gamified' process is 'game-based', so that it 'creates a system in which learners… engage in an abstract challenge, defined by rules, interactivity, and feedback that results in a quantifiable outcome ideally eliciting an emotional reaction.' (Kapp 2012, p. 11) The mechanics

283

involved in playing a game include levels, points, scores and time constraints, and according to Kapp these are insufficient but crucial in the making of a game-like engaging experience. Aesthetics are also of importance, so that the design aspect of the game-like experience needs to promote engagement. Engagement in itself is of paramount importance for gamification, as an explicit goal of the process is 'to gain a person's attention and to involve him or her in the process'. Furthermore, claims Kapp, the most important element of gamification is so called 'game-thinking'. This is 'the idea of thinking about an everyday experience… and converting it into an activity that has elements of competition, cooperation, exploration and storytelling.'

Kapp, and other proponents of gamification, see it as a novel and effective technique to influence behaviour, including to promote learning:

> Gamification can be used to promote learning because many of the elements of gamification are based on educational psychology and are techniques that designers of instruction, teachers, and professors have been using for years. Items such as assigning points to activities, presenting corrective feedback, and encouraging collaboration on projects have been the staples of many educational practitioners. The difference is that gamification provides another layer of interest and a new way of weaving together those elements

284

into an engaging game space that both motivates and educates learners. (Kapp 2012, p. 12)

Although Kapp provides some cautionary remarks, those are mostly related to the challenges of creating a game that is 'both fun to play and instructional' (Kapp 2012, p. 14).

Many have bought into the idea of gamifying education. One example can be seen in 'Quest to Learn', a public middle school and high school in New York set up in 2010 that is based on an 'innovative educational philosophy developed by top educators and game theorists'. The school's pedagogical methods and curriculum are centred around games and game design:

Educational games are at the core of Quest's curriculum. Sixth graders use Dr Smallz, where they play the role of designers, scientists, doctors and detectives as they explore cellular biology and the human body. And ninth graders use Storyweavers, a collaborative storytelling role-playing game. These games not only engage students in the learning process, but also allow teachers to assess students in real time and provide feedback on learning experiences immediately. (Quest to Learn)

285

Proponents of gamification tend to focus on the novelty and effectiveness of gamifying education. The basic idea is to use the apparent allure of games and harness it to influence students towards achieving the aims of the educator. In a different context, gamification is being hailed by marketers as a way to influence the behaviour of potential customers: 'gamification… allows marketers to focus on what they know best — convincing consumers to take loyalty and purchasing actions — using a powerful toolkit of engagement gleaned from games. Armed with a new understanding of what makes people tick, and how to wind them up, marketers can build experiences that are enduring and engaging.' (Zicherman 2011) A similar claim seems to be made by proponents of gamifying education, only instead of promoting sales of various products, gamifying education is meant to promote learning.

The use of digital gamification in education has been receiving a great deal of attention in the past two decades. The vast majority of this attention follows two strands, one focused on the nature of gamifying education, and the other on the effectivity of gamifying education. However, although there has been some critical discussion of normative issues related to gamification, this question has received far less attention. As digital gamification is a method which influences students by not overtly coercive techniques nor does it involve rationally persuading the students, I suggest it is a good candidate to be thought of as manipulation. As that is the case, I

will use the manipulation assessment framework that was established in Chapter Four to assess the use of digital gamification in education.

Although not every gamification in education requires the use of an electronic device, and the two can be assessed as separate methods, there is a general tendency to bind these together, as I will also assume in this chapter. My reason to assume that is that digital games are a very popular form of gamification, as they offer many of the supposed benefits of gamifying education and are efficient in terms of cost and effort on the part of educators. Therefore, I will henceforth use the term 'digital gamification' to refer to the methods I am analysing here.

Specifically, I will focus on Kahoot! as a case study. Kahoot! is an online platform that allows users to create multi-participant quizzes. The platform is incredibly popular, and reports over 1 billion players a year (Kahoot.com). Though there is some empirical research about the platform, to the best of my knowledge the normative theoretical analysis presented here has not yet been performed. I believe that focusing on one specific digital game allows for a clearer view of the suitability of the manipulation assessment framework for assessing digital gamification and highlights the issues involved in this technique that are often hidden from view.

To begin our critical analysis of Kahoot! and gamification in general, it is worth providing some preliminary remarks. Unlike what appears to be a commonly held

287

view, even by experts of the field (Kapp *et al.* 2014, p. 40), there is little new about the use of games and gamification in education and learning. The use of games for training and education is at least 5000 years old, starting with the Chinese game that is known today as Go, in which players attempt to encircle each other's pieces as a simplified form of war manoeuvres on a battlefield (Keys and Wolfe 1990, p. 309). The Hindu game Chaturanga is another historical example of a war game simulation, and both of these games are somewhat similar to the 17th century war game simulation - Chess (Keys and Wolfe 1990, p. 309). Therefore, it is safe to assume that games have been used for purposes of learning, training and education for many years.

Likewise, the use of gamification techniques, i.e. the use of so-called game-based mechanics and game-thinking to change behaviour in real life is similarly an old practice. If we consider the techniques of measurement and reward, then even the weighing of good and bad deeds which appears in a variety of world religions, might count as gamification aimed at changing behaviour. One example is the Christian fable about Santa Claus as an accountant and gift giver, which is sometimes used to train children towards behaving well.

Similarly, the common school, with its established practices of accounting and measuring presence, academic performance, behaviour, and other aspects of the

288

students' lives, together with a system that rewards and punishes students for these, already incorporates many aspects of what may be thought of as gamification. As Neil Selwyn maintains:

> Of course, a logic of reward, measurement and competition is not unique to the use of games in education but has been a long-standing element of classroom organization and control. Nevertheless, digital games certainly extend and intensify its presence in education, as well as expressing learning in overtly value-based and market-driven terms. (Selwyn 2014, pp. 100-01)

Furthermore, gamification techniques meant to produce a predictable change in behaviour are not confined to the school. Collection of data and the ability to reward and punish users is being implemented by social media web-sites, credit score companies in modern liberal democracies, and state institutions in more authoritarian regimes, such as the suggested 'Social Credit Score' system in China (Chorzempa 2018).

The recent popular interest in gamifying education is mostly oblivious to these connections of gamification techniques to larger social and political issues, as well as to the history of games. Even in the narrower field of learning and education, the effects of game-based learning and teaching has been researched and critiqued for

at least six decades (Harviainen 2014). Nonetheless, there is renewed interest in the use of games and gamification in education as if these are novel ideas. It is possible that the recent hype in favour of game-based learning and gamification is riding on the current popularity of commercial digital games.

Therefore, one problem of gamification and the use of digital games in education is that the field is submerged in what Neil Selwyn calls 'Bullshit Ed-Talk'. Following Harry Frankfurt's famous essay 'On-Bullshit', Selwyn maintains that much of the discussion of digital technology in education is based on a limited language that "routinely normalizes matters of oppression, inequality and injustice. There is little – if any – acknowledgement of differences of class, race, gender, disability or other social ascription." (Selwyn 2016, p. 441) To a significant degree, this applies to gamification and digital games, in ways I will elaborate on below. A multi-faceted assessment of gamification is often lacking when various stakeholders in education are encouraging gamified educational processes. Game developers, alongside educational policy makers and practitioners on all levels, use optimistic and convoluted language to discuss the merits of gamification and digital games, and routinely downplay the inefficient and possibly harmful effects of gamifying education.

290

If we disregard the hype, the truth appears to be that we are still largely unaware of the possible merits and flaws of digital games in education (Harviainen 2014). Although little is known about when and why gamification promotes educational aims, at least some digital games and gamification of learning seem to genuinely engage with educational challenges in an effective, creative and novel way. To say the least, some of these have become incredibly popular with both educators and students worldwide. Through focusing on one case study, the analysis presented here aims to allow for a better understanding of what is involved in using digital gamification in education, by analysing it as an instance of manipulation.

7.2 Case Study - Kahoot!

Kahoot! is the brand name of an online platform providing users with the opportunity to set up and participate in a trivia game. As a platform, it allows educators and students to create quizzes, discussion and surveys. Creating a quiz with Kahoot! is easy and free, and it is a common tool used by many educators in different educational institutions, with over 1 billion yearly users worldwide (Kahoot.com). To further explain the popular quiz feature, which is the one I shall focus on, it is useful to think of it as an online pub-quiz or a family game of trivia. As

291

the quiz maker, the user can write her own questions and answers. Players can use a variety of electronic devices connected to the internet (phones, computers, tablets, etc.), to log into the specific quiz by using a unique pin number and are then asked to choose a nickname to be known by for the quiz. Once the quiz maker starts the quiz, questions and answers appear on a main screen (computer, TV, projector, etc.) and the participants are able to select what they believe is the correct answer on their own devices. After all the participants have selected an answer, or the time allocated to the question by the quiz maker expires, the correct answer appears on the main screen as well as a graph showing how many participants chose which answer. Simultaneously, participants who answered correctly will see a green screen with a V on it on their device while participants who answered incorrectly will see a red screen with an X on it. Participants receive a numerical score based on whether they answered correctly and the speed at which they have done so, and the platform aggregates and keeps track of their overall score for the quiz. Next, the main screen shows who scored the most for the particular question and then shows the current overall leaders in the quiz. The same process repeats itself for all the questions, and the final score and winners are revealed after the last question has been answered by the participants.

292

Kahoot! has been hailed by researchers and educators as 'a great assessment tool to add to any classroom to enhance student learning through motivational, real-time assessment!' (Johns 2015), and as 'a promising formative assessment tool that is feasible, practical and makes learning fun and enjoyable' (Ismail and Mohammad 2017). Research of students' perception of Kahoot! finds that students believed Kahoot! promoted learning, influencing classroom dynamics, engagement, motivation and the learning experience (Licorish *et al.* 2018). In my own personal experience, middle school students are extremely enthusiastic about using Kahoot! and repeatedly ask to play it. In this sense, Kahoot! is a good example of the extent to which gamification of education can be engaging and entertaining. However, little to no attention has been given to possible harmful consequences, and other reasons that might undermine the increasing use of the platform. To do so, I shall use the manipulation assessment framework.

As a digital game, Kahoot! is a technique that might be used by educators to influence the behaviour of their students. It is not coercive, nor is it an attempt to influence the students by reasoning with them. Kahoot! was designed to engage students in a very entertaining way, by using game elements such as competition, points, leader board, time constraints, attractive aesthetics and catchy music. All of this assist in attracting the attention of players, much like other online digital games.

293

However, unlike many other digital games, Kahoot! is specifically marketed as 'making learning awesome' (Kahoot.com 2020). As I argued earlier, methods of influence which do not use coercion or reasoning are good candidates to be considered manipulation. Therefore, by using the manipulation assessment framework to assess the use of Kahoot! as manipulation I suggest we might gain new insights about Kahoot! and other digital games, which can easily be missed otherwise.

7.3 Criterion A) Consequences

In his metatheoretical analysis of the use of digital games and gamification in education, Harviainen raises the question of 'whether we currently know enough about gamifying education to safely and beneficially utilize that gamification' (Harviainen 2014). He criticizes the tendency to automatically present gamification as superior to traditional methods of teaching and instruction. In regard to the effectiveness of gamifying education, some of the problems he recognizes in the field are as follows: proponents of gamification rely on empirically unsupported claims made by other proponents of gamification. There is little empirical research to show that gamification is effective, what are the contexts in which it is effective and what are the processes involved. Furthermore, he claims that games and

294

gamification must be tailored to the specific context to be both beneficial and safe. However, because games and gamification need to be tailored to the specific context, there are difficulties in generalising empirical findings (Harviainen 2014).

The worries Harviainen raises can be usefully arranged within the categories of the manipulation assessment framework established in the previous chapters. I will use the case study of Kahoot! to demonstrate how the framework might be used to analyse a specific digital game as an instance of benign manipulation.

7.3.1 Intended consequences

Educators might use Kahoot! for various reasons. The main ones that are mentioned in the literature are that it 'adds fun' to the classroom, promotes students' engagement and motivation, enhances learning and comprehension and promotes interaction between educators and students. There is a tendency by researchers to group these various reasons together, and to assume there are causal links between them. For example, Plum and LaRosa claim that: "Such eLearning tools add positive energy, support concept exploration, and add fun to the classroom, which seems to translate into increased comprehension and motivation" (Plump and LaRosa 2017, p. 157). For the purpose of assessing Kahoot!'s effectiveness, i.e. the extent to which it

is successful in achieving the educator's desired result, I shall briefly discuss each of these possible positive effects.

There appears to be good evidence that Kahoot! does 'add fun' to the classroom. Firstly, the fact that it is played by hundreds of millions worldwide is a good reason to believe that many educators and students do find it entertaining, as they actively choose to use the platform and are not forced to do so. My own personal experience is that students repeatedly ask to use it and enjoy doing so. Furthermore, it appears that these effects of Kahoot! do not wear out even if it is used frequently. A research from 2014 found that single use of the platform during one lecture is similar in its effect to a repeated use in every lecture in a class for five months (Wang 2015). Therefore, it appears reasonable to infer that Kahoot! does 'add fun' to the classroom. In this sense, students are engaged to play the game, and are motivated to win while they are playing it.

On the other hand, it should be emphasized that at least some students do not enjoy playing the game. For example, Plump and LaRosa recorded the following comments by participants: "The same students won each time, which wasn't fun for the rest of us," "It was stressful because I had to read the question and answer it so quickly I didn't have time to think," and "It seemed a little gimmicky." (Plump and LaRosa 2017, p. 156). This contradicts claims made by the same researchers in

their conclusion which reads '…the "gamification" of learning increases student engagement by appealing to all students, even the most introverted, combining both a cooperative fast-paced learning environment and friendly competition' (Plump and LaRosa 2017, p. 157). Indeed, participation in a public quiz show in which you do not know most of the answers can be rather daunting. That said, the overwhelming majority of students do report enjoying the game.

A different question is whether playing the game in class causes students to increase their engagement and motivation to study the content that they are being quizzed on. I could not find empirical research looking into this question, and I imagine the answer can go many ways. Assuming the students are quizzed on something they were supposed to study earlier, some might use the friendly environment of the platform, its anonymity and 'built in permission to fail', to play by guessing the correct answer without being motivated to study the content. For them, the game will become 'how well can I guess the correct answer without learning the content', which also appears to be fun. In other words, Kahoot! might increase students' engagement, but this does not necessarily mean it increases engagement with the content to be learned. Whether it does or does not motivate students to engage with the content is an empirical question which has not been investigated yet. Selwyn echoes this sentiment as well, mentioning that players tend to find

297

limited strategies in order to do well in the game, without necessarily engaging in higher-level learning:

> Thus despite their association with intense excitement and thrill seeking, games tend to be played along rational and repetitive lines, with players adhering to strict rules and developing bounded strategies in order to succeed. (Selwyn 2014, p. 97)

Furthermore, there is little evidence to support the view that Kahoot! enhances learning and comprehension. Many empirical studies assess students' perceptions of using the game, but this is not the same as assessing the actual learning that is going on. The extent to which it is the learner's perception of her learning that should inform us about the actual learning process is probably small. The pervasive 'myth' that 'learners know best' is discussed and criticised by Kirschner and van Merriënboer (Kirschner and van Merriënboer 2013), who conclude that there are few reasons to think that students' perceptions of their learning are reliable. Unfortunately, most of the research into the use of Kahoot! is based on student surveys, and therefore should not be considered very reliable as an indication of learning and comprehension.

There is some evidence that Kahoot! is somewhat effective in vocabulary acquisition (Abrams and Walsh 2014; Ismail and Mohammad 2017; Licorish *et al.*

298

2018; Pede 2017; Poláková and Klímová 2019), but also that students think it is 'not the best tool to simplify complex subjects as perceived by medical students' (Ismail and Mohammad 2017). It appears that the empirical findings so far allow for a tentative conclusion that Kahoot! is mostly useful for assisting in rote learning, as might be required in vocabulary acquisition.

To conclude, whether Kahoot! is effective in achieving the manipulator's desired result depends to a large degree on the expectations educators have from the platform. If it is seen mostly as an entertaining addition to classroom instruction, Kahoot! does appear to 'add fun' to lessons. However, as far as encouraging students to engage with the content and promote their overall learning is concerned, it is doubtful whether Kahoot! has any significant positive effect. This conclusion can be generalised to other uses of gamification in education. While we should not downplay the importance of entertaining students every once in a while, even if that entertainment does not promote further learning, we should not automatically equate engagement with learning, as many proponents of gamification appear to do.

7.3.2 Unintended consequences

Digital games which use a competitive framing, such as Kahoot!, risk reinforcing these tendencies in competitive students. Some players will be over-

299

motivated to win which will hinder their learning as they are likely to concentrate their efforts on winning the game without applying the learning task at hand. Highly competitive players will also contribute to an environment in which other players might feel like failures. Furthermore, competitive framing also encourages cheating. This is because of the supposedly safe environment of play where consequences do not carry much weight, and the desire to win. In turn, cheating reduces the value of achievement for everyone. That said, while in a gamified environment some consequences are neutralized, not all are so – some of the social consequences are real and might affect the level at which players play the game. For instance, an employee might lower their level of play in order not to beat their boss. Moreover, some students might prefer not to play, i.e. games might not fit everyone's learning preferences, and perhaps especially capable students will find these tedious.

As was briefly mentioned above, the competitive design of Kahoot!, alongside the game feature that allows for anonymity while within the 'play zone', might encourage students to cheat in order to win. A few recurring examples I have noticed myself were students who used more than one device and so competed under various nicknames, students who picked an answer only after seeing what other students had done, students who researched the answers online, and more. In cheating, students are learning how to win the game, but the value of the

300

achievement, for them as well as for others, diminishes. A player who won by cheating, might still enjoy the beneficial social consequences that come with winning the game, and in this case have internalized the value that winning is more important than anything else, including honesty. Again, this might be in contradiction to the intended consequences of using Kahoot! that the educator had in mind.

Therefore, one unintended consequence of the game might be that players will learn from it different things than the ones the educator wished them to learn. As Selwyn remarks:

> Most learning that takes place within a digital game might be more accurately described as supporting repetitive adjustments to the rules, rhythms and expectations of the game design, rather than genuinely educational engagement. (Selwyn 2014, p. 104)

While it does not appear reasonable that students will be addicted to Kahoot! in the same way that some players are addicted to other video games, the engaging nature of Kahoot! might make it a constant disruption to classroom learning. Students might repeatedly ask to use the game which will interrupt more traditional, but perhaps no less effective, learning methods. Also, at times of low motivation to learn, such as the end of the school day, end of the week, or end of term, Kahoot! might become a solution that is 'too readily available' for teachers – akin to a microwave meal - not

301

too bad if eaten once in a while, but with significant downside effects if used too often. This way, Kahoot! becomes a constant distraction which will result in wasted time in the classroom that could have been used more efficiently. This is related to the students' and the teacher's understanding of the situation, because unlike other distractions, the company that promotes Kahoot! repeatedly claims that the game is 'making learning awesome', a claim that is believed by many students and teachers, but with very little empirical evidence to support it. I will discuss this issue further in the next section.

To conclude the discussion of Criterion A) Consequences: Kahoot! offers some advantages over more traditional teaching methods. For one, Kahoot! is very engaging and in most cases 'adds fun' to the educational process for most students. That said, it is questionable if this fun translates into effective learning. Moreover, increased engagement might hinder learning because of the effects of over-competitiveness and cheating, and de facto offering an easy 'fast-food' option for classroom interaction that might be distracting from more efficient methods. The important conclusion is that a digital game is not separate from the context in which it is being used. Some digital games in some contexts are probably benign, fun and even effective, but in other contexts they might hinder one's learning and have other harmful effects.

302

7.4 Criterion B) Understanding

Proponents and researchers of gamification tend to focus on the short-term positive consequences of using gamification techniques, and disregard or downplay the possible harmful unintended consequences. However, even in cases where the overall consequences of the specific gamification are positive, a wider perspective reveals important reasons to adopt a critical attitude towards the use of gamification in education. Some of these criticisms are based on the understanding that gamification is not detached from the context in which it is being used:

The deployment of a game in an educational context reflects more than just the needs of the curriculum. While gameplay does allow players some leeway to act freely within the constraints of the game, the subject matter is not treated in a truly neutral fashion. It always contains values embedded in the game by its designer, as well as the value sets of the teacher who has chosen to use that particular game for that particular topic (Graham & Gray, 1969). This is no different from a course book and the teacher's choice to use it. Because of the educational framework, and the player desire to do well (often: win), players have limited chances and/or motivation to explore alternative options (Carlson & Misshauk, 1972, Harviainen 2014).

303

In other words, as was discussed earlier, because means in education are never value neutral, they should be assessed as part of a wider social structure within which the gamified process promotes certain values. Along these lines, Selwyn discusses digital games and gamification as a technology we should distrust, at least to a certain extent:

> Most significantly – and returning to the core theme of this book – these arguments obscure a set of ideological criticisms that should be levelled at digital games. These all relate to the contention that it is a mistake to see digital games as somehow 'apart' or 'separate' from the social, political, cultural, and economic contexts of education. Rather than drawing the user into some sort of decontextualized 'play world' or 'magic circle', the playing of games remains embedded firmly in specific contexts… Four broad sets of objections can be made against the increased use of digital games in education: firstly regarding their rationalization of education; their entrenchment of dominant values and structures; their commodification of learning and education; and their imbrication of education with the interests of capital. (Selwyn 2014, p. 96)

In what follows, I will use Selwyn's ideas as they apply to Kahoot! in order to show that the platform demonstrates the applicability of these objections. Notably, in most cases both the students and the teachers are unaware of the underlying values embedded in the game which in turn causes them to have an inaccurate understanding of their situation.

There are at least four aspects of the hidden values in Kahoot! that may hinder the teacher's and students' understanding of their situation. First, the alluring nature of the game might hide the fact that the game does not hold pedagogical value beyond entertainment. Second, the game makes questionable assumptions about what is knowledge and what it means to be knowledgeable and encourages users to internalize those. Third, the game reinforces dominant social values such as individuality and competitiveness. And fourth, the game promotes the interests of for-profit companies, without acknowledging this fact. All of the reasons above are interlinked, but I shall discuss them as separate in order to provide clarity.

The reasons listed above are relevant for both the teacher and students who use Kahoot! and fall within the scope of Criterion B) Understanding together. As a reminder, the two aspects of this criterion are: 1) To what extent does the manipulation prevent the manipulee from having an accurate understanding of the

situation? and 2) To what extent does the manipulation prevent the manipulator from having an accurate understanding of the situation?

7.4.1 Internalization of false or limited beliefs

Selwyn argues that the 'rationalization of action inherent in digital games is at odds with freer, individually determined forms of learning' (Selwyn 2014, p. 97). Put simply, although games do allow some free action, this freedom is determined by the rules of the game. Therefore, in order to succeed, players must adjust their actions to the demands of the game. In other words, players are strongly encouraged to internalize the values embedded in the designed rules of the game, which they do not have a say over and cannot change. This issue is especially significant with digital games, that are produced and designed far away from where they are being used. Unlike a game designed by the teacher in the classroom, or by the students themselves, it is much harder to tweak the rules of a digital game, as the rules are set up in advance.[34]

The values embedded in Kahoot! stem from the designed elements of the game. In other words, the player must accept certain assumptions in order to play

---

[34] To draw further insights about the difference between digital games and games set up by a teacher, I will discuss below an example of a game set up by a teacher.

306

the game and succeed. Firstly, Kahoot! is a competition between players, and so

players must adhere to a competitive individualistic attitude over a collaborative one.

Secondly, points are allocated based on answering correctly, and more points are

given if the player answers quickly. This assumes that questions have one correct

answer, and that being a fast thinker is superior to being patient and taking the time

to self-assess one's knowledge. Both of these assumptions are questionable, if not

utterly wrong in many contexts, but one cannot play without accepting them. The

extent to which players internalize these assumptions comes into view when some

players, in order to get on the leader board of the top players, will sometimes make a

quick guess, which might be a better strategy than taking the time to be surer of the

right answer. Also, making a guess is better than not answering at all – as there is no

option of 'I do not know'. Moreover, the questions are only multiple-choice questions,

or True/False ones, which again assumes a limited understanding of knowledge that

the player must accept in order to play. Even if it is correct that in some areas of

human knowledge there is a clear and true state of affairs, this is not the case in

many other areas of human knowledge, which will not be able to 'come into play' in a

quiz. These various 'rationalised assumptions' force players to internalize beliefs that

might not be desired by the educator, such as: answering quickly is more beneficial

than taking the time to consider the options, making a guess is better than

acknowledging not knowing the answer, every question has a right and wrong

307

answer, etc. In order to play the game, and win it, players must internalize these rules of the game. This is mostly done implicitly, and without full awareness of what is really going on.

Moreover, by measuring and rewarding students for answering quickly, Kahoot! assumes that knowledge can be easily quantified, assessed and compared among different knowers. In some subjects and contexts this might be the case, but not in others. For example, it is much easier to measure whether a student recognizes the correct spelling of words, than it would be to measure the quality of a literary analysis or a creative creation. This diversity of the quantifiability and measurability of knowledge is absent in Kahoot! whose design risks promoting the message that 'if it is not measurable it does not exist'. For this reason, Kahoot! appears more suitable to be used in very limited areas of human knowledge, but it risks promoting the false message that those are the only or the most important areas of human knowledge. While perhaps not all students and teachers will acquire this way of viewing learning and knowledge, some might. At least in the official marketing slogans of the company, these false assumptions about learning and knowledge are promoted very forcefully, such as in 'Kahoot! is a free game-based learning platform that makes it fun to learn – any subject, in any language, on any device, for all ages!' (Kahoot.com 2020).

308

Some criticisms of the dominant values of digital games focus on their 'banal' portrayal of gender, race, age and sexuality and their 'battle ground' mentality that has players rely on tactical and strategic thinking (Selwyn 2014, p. 99). This criticism does not directly apply to Kahoot!. However, in a more general way, Kahoot! does ask players to adhere to other conventions of dominant values, such as individualism and competition (Selwyn 2014, p. 99). A related issue is that students might equate 'winning the game' with being knowledgeable or smart. As was explained above, a player wins the game if they are good at the specific skills the game rewards, and if they are willing to internalize the values of the game. For example, students might internalize the assumption that fast thinking is an inherent attribute of intelligence. But, even while playing the game, would we consider as less intelligent a player who takes her time to consider the various answers more carefully? The equation the game makes between speed and intelligence is indeed questionable. This might pose a danger to students who tend not to do well under time constraints or students who struggle with reading. These students will feel that they are constantly losing, and hence are not as intelligent as other students. That is, the game promotes a narrow understanding of what it means to be 'smart' which is inaccurate in many other contexts. Student who internalize this understanding of intelligence are in danger of holding false views about their situation.

The discussion above does not mean that players necessarily internalize the assumptions of the game through the long term nor that these beliefs carry on to other contexts 'outside' of the game. Indeed, people play many games, with different rules, and it is not obvious that adhering to the rules of the game in order to play it, must encourage one to hold and apply the assumptions embedded in the rules in other contexts. As an example, one might participate in a game of trivia without transferring the rules and assumptions embedded in the game to other spheres of one's life. However, three factors might make Kahoot! more problematic than a weekly pub-quiz. First, it is played by many people, on many occasions, in many different disciplines. The sheer quantity of players playing the game, and the diversity of the educational processes in which it is being used, might make it a far more powerful transmitter of values and beliefs than if it was limited in scope. Second, Kahoot!, as claimed by the owning company (Kahoot.com 2020), is being played by school children, whose beliefs might be more malleable than those of adults. And, as it continues to be used in universities and on the job training, the game will accompany the players from a young age to adulthood and beyond. The repetition of the messages inherent to the game might become very powerful as a result. Third, the assumptions and beliefs the player must accept in order to play the game are similar to some common beliefs in non-game contexts (equating fast thinking with overall intelligence, that knowledge is quantifiable, all questions have a

single correct answer, etc.) which makes the game an amplifier of these non-game beliefs, rather than requiring the players to transfer these beliefs from the game context to the non-game context. Therefore, it is reasonable to think that the beliefs and assumptions embedded in Kahoot! will be transferred to situations that are beyond the game itself, to a greater degree than those of other games.

7.4.2 Hidden interests of for-profit companies

Although the stated goal of the company behind Kahoot! is to 'improve education all over the world and help everyone – of any age, aptitude or circumstance – unleash the magic of learning.', it should be acknowledged that Kahoot! is a private company which has raised hundreds of millions of dollars in investment, is worth approximately 1.4 billion dollars and is expecting to make a profit in the future (Lunden and Mascarenhas 2020). Currently, although it is offering a basic free version of the game to teachers and home users, it also offers a professional paid version and is actively attempting to monetize the game. One can imagine how students who enjoyed playing the free version in class would later purchase the professional version as corporate employees. Thus, educators who choose to use Kahoot! in class are also providing a service to a private company. In this way, not only might they promote a private company without necessarily

311

intending to do so, they are also participating in what Neil Selwyn calls the 'commodification of learning and education' (Selwyn 2014, p. 96). Unlike a teacher that produces or designs a game for her students, a teacher that uses Kahoot! as a ready-made product, is providing a service to a for-profit company without being compensated or acknowledged for this service. The teachers who promote the game are not being paid to do so, although the company does profit from their actions. The almost obvious fact that the use of Kahoot! is embedded within an economic system is hidden from most users of the platform.

Moreover, and in light of the above, it would make sense to design the game exactly to achieve the purpose of the company, i.e. to make money – in the case of Kahoot! that would be to create a very engaging game, so that more people would be likely to use it as often as possible. This bears a strong resemblance to other non-educational for-profit companies which use various psychological mechanisms to maximize customer engagement, such as slot-machines in casinos, social media companies, media websites and more. Furthermore, in marketing the game as a 'learning' experience, the company undermines possible criticisms that are directed at other digital games – such as that they are a waste of time, even if they are very engaging. By doing this, the marketing of Kahoot! as 'making learning awesome' is

312

very clever, but also misleading and ironic – as it leaves users understanding less about their reality.

In sum, the game-like experience that Kahoot! offers hides from the students who play it important aspects of their situation. By drawing their attention to the fun and exciting elements of the game, to the competition and joy of winning, it hides the fact that they are being used by a for-profit company for its own purposes. This simple truth has not only escaped countless of students who routinely play the game, it has also escaped a huge number of educators as well as educational researchers. While the criticism presented above could be arrived at without the notion of manipulation, when we assess Kahoot! as an instance of manipulation, it sheds light on these aspects of the use of the game in education.

Similar to other instances of manipulation, because gamification hides important aspects of reality from the educators and students, one way to counter these effects of the game, and indeed to perhaps make the experience of the game truly educational is suggested by Harviainen:

> One final thing that needs to be mentioned is that when using game-based or gamified methods, educators should take time during briefing and debriefing to discuss with the students the ways in which the game is different from the real world. This is partially for the

313

purpose of improving the students' information literacies, but it has a
deeper purpose: the fields in which educational games have been
used the longest - military and business sciences - are also those that
are especially associated with treating real world phenomena (human
lives and entire national economies included) as if they were games,
as well as getting away with doing so. This is an extremely risky,
unwelcome type of gamification, the support of which through an
educational system is not recommended. (Harviainen 2014)

I would add that educators should discuss with students the various ways by which
Kahoot! and other digital games used in education incorporate assumptions, values
and aims that are not clearly apparent to players. For example, that Kahoot! is a for-
profit company that aims to engage as many users as possible and is designed to
achieve this in the ultimate aim of making money, that the designed mechanism of
the game makes assumptions about what it means to be intelligent and successful
that are not necessarily true in other contexts, that it assumes that knowledge can be
easily quantified, that engagement equates to learning and more.

314

7.4.3 Too much fun

On one side, the appeal of Kahoot! is exactly why students are happy to play it and teachers are keen to use it in class. However, the strong appeal of the game might also hinder the students' and teachers' ability to have an accurate understanding of their situation. The aspects of the game mentioned above, that it rationalizes education alongside certain questionable assumptions about learning and intelligence, that it is not a proven effective learning method, that it encourages the internalization of dominant values, and that it serves the purposes of a for-profit company, are all 'hidden' from plain view of both teachers and students. This might be likened to a casino highlighting the fun and exciting aspects of gambling for the purpose of maximizing their intake of a larger share of players' money. In fact, similar cognitive mechanisms are being used by so called 'gaming' companies who specialize in maximizing the engagement of gamblers to the ones that are being used in Kahoot! and other digital games. As the valuation of the company is at least partially dependent on the number of active players of the game, the game's ability to draw students' engagement is a way for the company to profit.

All of these issues are hidden from plain view, as the company and proponents of the game repeatedly promote it as if engagement equates to learning, and as if the game is 'value-neutral'. In the words of one of its co-founders: '"We as a

315

company are not interested in what they are teaching", he said. "We are interested in how they are playing.' (Singer 2016) This remark not only makes the questionable assumption that Kahoot! is somehow effective for learning, it also hides the profit-making aspect of the endeavour.

7.4.4 The manipulator

It appears to me that the most significant way in which Kahoot! might hinder the understanding of the educator is that the educator, much like the students, might unquestionably accept the messages that are promoted by the company. That is, that "Kahoot! is a free game-based learning platform that makes it fun to learn – any subject, in any language, on any device, for all ages!" (Kahoot.com 2020) In many ways, a teacher who uses the platform is very much in the same boat as the students who play it. First, the assumptions the game makes about what it means to be smart (fast thinker), what is knowledge (one correct answer, quantifiable and measurable), what it is to be successful (individualistic and competitive, winning above all) are taken up by the teacher who chooses to use the game, at least while using it. For some teachers, these false and limited assumptions might be internalized and obscure the more complex reality of these issues. Second, the fact that Kahoot! is a for-profit company with a very clever marketing message as a

316

'learning' game may also be hidden from the view of the teacher. The teacher might promote the use of the game without realizing that they are providing a service to the company.

Another reason to think that using the game might hide important aspects of reality from the teacher is that often students are more technologically savvy than the teacher. In instances when that is indeed the case, students might cheat without being caught, they might hack the game in ways hidden from the teacher, or 'mess with' the teacher in other ways. One example I've seen documented online was students posting the pin number of the game online and asking random people to join. This way a huge number of people might join the game, without even being in the classroom.

7.5   Criterion C) Trust

I've discussed above some of the reasons we should consider when assessing whether a certain gamified learning process is appropriate, first focusing on the intended and unintended consequences of the method, and secondly on the possible effects the gamification might have on the understanding of both the students and the teacher. Lastly, I wish to consider whether the gamification of

317

learning is appropriate within the norms of behaviour which constitute the relationship between the manipulator and manipulee. Specifically, is the use of gamification a breach of trust that is important and meaningful in the teacher-student relationship?

As discussed earlier, two of the main attitudes that constitute trust in the teacher-student relationship are goodwill and truthfulness on the part of the teacher. In this sense of trust, is the use of gamification a breach of trust? At least some game theorists think that it is. Ian Bogost suggests that digital gamification should be renamed 'exploitationware':

> In particular, gamification proposes to replace real incentives with fictional ones. Real incentives come at a cost but provide value for both parties based on a relationship of trust. By contrast, pretend incentives reduce or eliminate costs, but in so doing they strip away both value and trust. (Bogost 2011)

I believe this overarching generalisation is inaccurate. As was discussed earlier, to assume that the only moral way to influence the behaviour of others is by reasoning with them is unreasonable. Just as I would like to move and be moved by appeals to emotions, I might not object to being moved by game mechanics. Specifically, in cases where there are no serious harms, and where the gamification does not

318

prevent me from having an accurate understanding of reality, there appears to be little reason for me to object to such influence. Bogost appears to see as a problem the fact that gamification might hinder one's understanding of their situation, which I agree with, but this does not necessarily apply to every gamified process. For example, some gamified processes might assist me in achieving long term goals, such as physical health, where the 'pretend incentives' such as badges or points align with my real incentive of being healthy.

If we examine Kahoot! it appears that many students, if not all, would not object to lessons being engaging and entertaining, by various means the teacher might employ, gamification being one of those. This is especially true in cases where the gamified experience has overall good consequences and does not hinder the students' understanding of their situation. This might be the case when a teacher chooses to prepare a game by herself to be used in the classroom. For instance, there are many examples of creative writing exercises which may be gamified, such as a pass around story-writing. In this gamified writing exercise, students are asked to start a story and after a minute or two pass the paper to another student, who will continue writing the story for another minute and then pass it on again. In such a game, students are being incentivised by the game-mechanics involved, such as the time constraint, but this does not appear to be a breach of their trust. Such a writing

319

game would align well with the aim of the learning process (creative writing). Moreover, the teacher can easily tweak the rules to make the game suitable to the specific context, for example by allowing more or less time for each round, or by forbidding the use of swear words or students' names . On top of that, such a game would not hide the interests of a for-profit company, or encourage the students to internalise false beliefs. That is to say, we can imagine a game that would incorporate many of the benefits of gamification, without the negative aspects of it. Indeed, many educators routinely make use of such playful and 'gamified' learning experiences.

As mentioned above in Kahoot! there are more serious aspects that are hidden from view, and we might need to be more careful in our use of the platform. For one, if allowed time to reflect upon their situation, teachers and students might not wish to be taken advantage of by a for profit company, or at least would like to be made aware of the ways the company profits from their use of the platform. In that sense, a teacher that chooses to use Kahoot! might be bringing into the classroom the interests of economic factors, even if these are for free. This might be akin to teachers handing out a candy to the student because they were given it for free by the company that manufactures these candies. In such a case, this might constitute a breach of trust. The more general lesson here is that though at times gamification

320

does not appear to be problematic at all, it might be wise to be critical towards more overarching promotions by self-interested parties to use it across different contexts.

7.6 Conclusion

In his assessment of educational games and gamification, Harviainen concludes that:

> As a result, successful deployment of games and game elements as learning tools is first and foremost an issue of checks and balances, between educational content and playability, between too little and too much challenge, between insufficient and excessive enjoyment - and between a very wide range of personal and interpersonal variables affecting those. Yet even if one balances all those correctly, the goal may not be fully obtained, because well-learned does not necessarily mean easily assessed, and the educational establishments require proper assessment. (Harviainen 2014)

In this chapter I have used the manipulation assessment framework as it applies to the case study of Kahoot!. This has resulted in similar findings as the more general

ones mentioned by Harviainen in the quotation above. Kahoot!, and other digital games, are not incredibly beneficial as the company which produced it would want educators to think, nor are they necessarily exploiting their users. By using the manipulation assessment framework, I was able to consider the intended and unintended consequences of using the game, the extent to which the game might hinder pupils' and teachers' understanding of their situation, and whether the use of Kahoot! might constitute a breach of trust.

In sum, there are good reasons to maintain a critical attitude towards Kahoot!. Although the game does 'add fun' to lessons, it is not clear that beyond a very limited scope it promotes actual content learning. Some of the unintended consequences might include over competitiveness and time wasted. Moreover, both the educator's and the students' understanding of their situation might be jeopardised, due to players being required to adhere to questionable assumptions the game makes forcefully, such as that being a fast thinker is superior to being patient and reflective, or that knowledge can easily be quantified and measured. A serious worry is that the game does not make it clear that the engaging nature of the game serves the interests of a for-profit company, a fact which is hidden from most educators and students. These worries are also relevant to the trust that is part of the teacher-student relationship, because if the educator's role is in part to be truthful and

promote a better understanding of reality, then using the game in an uncritical way might jeopardise this commitment. In that, the use of a digital game that is produced by a for-profit company is significantly different than other gamification techniques which the educator might choose to use.

However, the discussion in this chapter should not be read as an argument against the use of gamification more generally in educational settings. In many cases, educators who choose to use games and gamification as part of the educational process will not be doing anything objectionable. To be better aware of whether the use of games and gamification is problematic, educators may use the assessment framework for benign manipulation. If the game has an overall positive consequence, is not harming the students' and the teacher's understanding of the situation, and is not betraying the students' trust in the teacher – such use will be appropriate. I believe that in many, if not most, cases in which educators use games in the classroom, their actions are commendable, rather than blameworthy.

Chapter 8   Conclusion

My main goal was to provide a conceptual map with which to assess the appropriate use of benign manipulation in educational processes and relationships. In order to do that I provided an analysis of the concept of benign manipulation, articulated a framework for the assessment of the use of benign manipulation and demonstrated the application of the framework in three case studies. To summarize briefly: based on recent theoretical work on manipulation, I have maintained that there is currently no consensus regarding the nature of manipulation. However, for the practical purpose of assessing instances of manipulation, it would be reasonable to assume that an attempt to influence others which involves deception, trickery, subversion of rationality, pressure or a failure to track reasons can be considered an instance of manipulation. Assuming that manipulation, in general but at least in the case of benign manipulation, should be understood as a non-moralised concept, I established a framework for the assessment of benign manipulation in education.

The framework consists of three criteria: A) the consequences of the manipulation, B) the effect the manipulation has on the understanding of the situation for both the manipulator and the manipulee, and C) whether the manipulation

324

breaches the trust between the manipulator and the manipulee. As part of the conceptual analysis, I have also discussed why the view that manipulation harms autonomy and the view that there is a qualitative difference between manipulating children and manipulating adults are mistaken.

In the second part of the thesis I applied the framework for assessing benign manipulation to three case studies. The first was the use of deception in education, and specifically the attempt to raise academic performance by communicating false expectations to students, which is sometimes known as 'the Pygmalion Effect'. Using the framework, I concluded that although in certain contexts it might be appropriate to deceive students, the thought that deceiving students should be a default option is misguided. Next, I looked into the suggested liberal use of pharmacological cognitive enhancement. By using the framework, I demonstrated that even if the consequences of such liberal use are overall beneficial, we should still be reluctant to use pharmacological cognitive enhancement on a wide scale due to epistemic reasons, which are captured under Criterion B) Understanding of the framework. Thirdly, I assessed the suggested use of gamifying learning in order to promote learning and academic performance. I focused on the widely popular quiz game 'Kahoot' to show that although game aspects can encourage engagement and 'add fun' to the educational process, they might also mask serious issues, such as hiding

325

important aspects of reality from the students, and influence students in ways other than what they claim and intend to.

I believe the application of the conceptual framework to assess the three case studies, demonstrated that I was able to provide a workable conceptual map for the assessment of benign manipulation in education. It is my hope that I was also able to demonstrate the wide range of possible application of the framework. In terms of suggested interventions in education, the framework can also be used in assessing interventions such as 'cash for grades' in which students are offered financial incentives to raise their academic performance (Warnick 2017), changing the composition of school meals from processed food to healthier options (Belot and James 2011), the use of weighted vests (Orenstein 2018) and many other policies and practices that involve benign manipulation.

The motivation for the present work originated from personal experiences as a young teacher. When I started teaching, I found myself using different kinds of manipulation very often. Intuitively, I felt that many of my manipulative actions were appropriate, but I did worry that I am abusing students' trust or acting unethically for other reasons. When I started reading about the use of lies and deceptions in education, the recommendations stemming out of the empirical research I had read, specifically on the issue of the Pygmalion Effect, struck me as morally dubious.

326

Deceiving students as a policy meant to raise performance seemed problematic. Yet here I was, acting in a way that is very similar to the policy that I intuitively felt is wrong. This is the puzzle the present work aims to resolve. Why is it that some manipulative techniques used by teachers are appropriate, and yet recommendations to use these techniques as default policies are mistaken?

Very often, the mistake is that these policies are based on empirical research that only looks into the positive consequences of the techniques, without considering other relevant features of the action. On the other hand, it would be a mistake to do away with manipulation altogether, as manipulation can have many benefits, and is sometimes required. For this reason, following recent work in analytic philosophy, I have analysed manipulation as an un-moralised concept. The result of this analysis shows that if and when manipulation is wrong, it may not be because of the ethical nature of manipulation, but because there are other wrong making features of the action, such as negative consequences, epistemic reasons and issues related to the trust that is a constitutive part of teacher-student relationships. The framework for assessing benign manipulation in education is meant to serve as a map for educators who might be struggling with similar puzzles.

A more general result of the analysis is that there should be significant room for teachers' professional judgement, rather than a general requirement to follow

327

recipes or prescriptions for effective teaching. In this, my work serves as an addition to such works as Jane Green's 2011 *Education, professionalism and the quest for accountability: hitting the target but missing the point* and Joseph Dunne's 1997 *Back to the Rough Ground*. Green claims that the present system of accountability focuses on measured performance. In turn, this focus encourages a rigid bureaucracy and inflexibility and obligatory compliance by practitioners (Green 2011). As an addition to this line of criticism, the work presented here adds that the focus on measured performance may also lead practitioners to use manipulative techniques when they should not use them. Such use may have various unintended results, like bad consequences, diminished understanding of reality by both teachers and students, and broken trust. In order to provide teachers with a more nuanced approach, one that does not base decisions solely on measured performance, but also does not try to get rid of all manipulative actions, I have offered an alternative account of manipulation and its use in education.

Dunne's project is wider still, as he aims to criticise the dominant view that professional knowledge consists in applying scientific theory and technique to solve problems of professional practice (Dunne 1997). Dunne's approach is to trace the discussion of the issue in the writings of various past philosophers, such as Aristotle, Jurgen Habermas and Hannah Arendt. Though I share the worries and ultimate

328

motivation of Dunne's work, my approach was to understand and resolve specific cases practitioners might encounter in their daily routines. In that, it may serve as a humble addition to Dunne's theoretical work. Moreover, the analysis and recommendations presented in this work, as they stem from teachers' practices and were designed to inform their practice, may offer practical advice to teachers who engage with these dilemmas in their classrooms. I am confident the present work can be developed to encourage fruitful discussions and insights in the context of professional development workshops for teachers or teachers' training courses. In the few opportunities I have had to present my work to teachers and other educators, I realized that teachers often find themselves in similar situations, and yet do not have the time and space to think and discuss these with other professionals. I encountered curiosity and enthusiasm, as well as much confusion. I hope this work can serve to dissipate some of that confusion, as well as make clear that some confusion and uncertainty are inherent to the educational profession, and teachers should hone and trust their own professional judgement. As assistance, rather than acting on intuition alone, or according to recommendations made out of context, teachers can use the analysis for manipulation and the framework for assessing benign manipulation to make better decisions. To say the least, the work presented in this thesis can serve as a starting point for a long due conversation about the use of manipulation in education.

329

Chapter 9 Bibliography

Abrams, S.S. and S. Walsh. 2014. 'Gamified Vocabulary', *Journal of Adolescent & Adult Literacy*, 58: 49-58.

Affeldt, S.G. 1999. 'The Force of Freedom: Rousseau on Forcing to be Free', *Political Theory* 27(3): 299-333.

Aristotle, 1954. *The Nicomachean Ethics of Aristotle.* W.D. Ross. (trans.), London: Oxford University Press.

Arpaly, N. 2003. *Unprincipled Virtue:An Inquiry into Moral Agency*, Oxford ; New York: Oxford University Press.

Bagot, K.S. and Y. Kaminer. 2014. 'Efficacy of Stimulants for Cognitive Enhancement in Non-attention Deficit Hyperactivity Disorder Youth: A Systematic Review', *Addiction* 109(4): 547-57.

Baier, A. 1986. 'Trust and Antitrust', *Ethics* 96(2): 231-60.

Bakhurst, D. 2007. Pragmatism and Ethical Particularism, in *New Pragmatists*, Cheryl Misak (ed.), Oxford University Press, pp. 122-141.

Bakhurst, D. 2009. 'Minds, Brains and Education', *Journal of Philosophy of Education* 42(3-4): 415-32.

Bakhurst, D. 2011. *The Formation of Reason*, United Kingdom, Wiley-Blackwell.

Baron, M. 2003. 'Manipulativeness', *Proceedings and Addresses of the American Philosophical Association* 77(2): 37-54.

Baron, Marcia. 2014. 'The Mens Rea and Moral Status of Manipulation', in *Manipulation: Theory and Practice*, C. Coons, & Weber, M. (eds.), Oxford University Press, pp. 98-120.

Belot, M. and J. James. 2011. 'Healthy School Meals and Educational Outcomes', *Journal of Health Economics* 30(3): 489-504.

Bennett, J. 1974. 'The Conscience of Huckleberry Finn', *Philosophy* 49: 123--34.

Bergstorm, C. and J. West. 2021. 'Calling Bullshit', Callingbullshit.Org. https://callingbullshit.org.

Berliner, D.C. 2002. 'Comment: Educational Research: The Hardest Science of All', *Educational Researcher* 31(8): 18-20.

Biesta, G. 2007. 'Why "What Works" Won't Work: Evidence-Based Practice and the Democratic Deficit in Educational Research', *Educational Theory* 57(1): 1-22.

Birbalsingh, K. 2016. *Battle Hymn of the Tiger Teachers: The Michaela Way*, John Catt Educational.

Blakely, J. 2013. 'Returning to the Interpretive Turn: Charles Taylor and his Critics', *The Review of Politics* 75(3): 383-406.

Blumenthal-Barby, J. 2014. 'A Framework for Assessing the Moral Status of Manipulation', in *Manipulation: Theory and Practice*, Coons, C. & Weber, M. (eds.), Oxford University Press, pp. 98-120.

Bogost, I. 2021. 'Persuasive Games: Exploitationware', Gamasutra.Com. https://www.gamasutra.com/view/feature/134735/persuasive_games_exploitationware.php.

Bok, S. 1980. *Lying: Moral Choice in Public and Private Life*. London: Quartet Books.

Bortolotti, L. and E. Sullivan-Bissett. 2017. 'How Can False or Irrational Beliefs be Useful?', *Philosophical Explorations* 20(1): 1-3.

Bostrom, N. and A. Sandberg. 2009. 'Cognitive Enhancement: Methods, Ethics, Regulatory Challenges', *Science and Engineering Ethics* 15(3): 311-41.

Bovens, L. 2009. 'The Ethics of Nudge', in *Preference Change*, Springer, pp. 207-19.

Brighouse, H. and A. Swift. 2014. *Family Values : The Ethics of Parent-Child Relationships*. Princeton: Princeton University Press.

Buss, S. 2005. 'Valuing Autonomy and Respecting Persons: Manipulation, Seduction, and the Basis of Moral Constraints', *Ethics* 115(2): 195-235.

Callan, E. and Arena, D., 2009. 'Indoctrination', in *The Oxford Handbook of Philosophy of Education*, Siegel, H. (ed.), Oxford, pp. 104-121.

Campbell, D.T. 1979. 'Assessing the Impact of Planned Social Change', *Evaluation and Program Planning* 2(1): 67-90.

Carson, T.L. 2010. *Lying and Deception: Theory and Practice*. Oxford ; New York: Oxford University Press.

Chorzempa, M., P. Triolo and S. Sacks, 2018. 'China's Social Credit System: A Mark of Progress or a Threat to Privacy?', *Policy Briefs*, PB14-18, Peterson Institute for International Economics.

Chua, A. 2011. *Battle Hymn of the Tiger Mother*, New York: Penguin Press.

Cooper, D.E. 2008. 'Teaching and Truthfulness', *Studies in Philosophy and Education* 27(2-3): 79-87.

Dancy, J. 2017. 'Moral Particularism', in *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), Metaphysics Research Lab, Stanford University.

Depaulo, B.M. 1993. 'Nonverbal Communication of Expectancy Effects: Can We Communicate High Expectations if Only We Try?', in *Interpersonal Expectations: Theory, Research, and Applications,* Blanck, P. D., Manstead, A., & Oatley, K. (eds.) 261-75.

Deterding, S., D. Dixon, R. Khaled and L. Nacke. 2011. 'From Game Design Elements to Gamefulness: Defining Gamification', in *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, 9-15.

Dewey, J. 1929. *The Sources of a Science of Education: The Kappa Delta Pi Lecture Series no 1*. New York,: H. Liveright.

Dunne, J. 1997. *Back to the Rough Ground : Practical Judgment and the Lure of Technique*. Notre Dame, Ind.: University of Notre Dame Press.

Dweck, C. S. 2006. *Mindset: The New Psychology of Success*. Random House.

Dworkin, G. 2019. 'Paternalism', in *The Stanford Encyclopedia of Philosophy,* Edward N. Zalta (ed.), Metaphysics Research Lab, Stanford University.

Eden, D. 1992. 'Leadership and Expectations: Pygmalion Effects and other Self-Fulfilling Prophecies in Organizations', *The Leadership Quarterly* 3(4): 271-305.

Fendler L. 2006. 'Why Generalizability is not Generalizable', in Smeyers P., Depaepe M. (eds), *Educational Research: Why 'What Works' Doesn't Work*, Springer, Dordrecht, pp. 51-64.

Fischer, J.M. and M. Ravizza. 1999. *Responsibility and Control: A Theory of Moral Responsibility*: Cambridge University Press.

Fowler, R. 2011. 'Mossbourne Academy: A Tale Of High Expectations. . . And No Excuses', *The Telegraph*. https://www.telegraph.co.uk/education/8341428/Mossbourne-Academy-A-tale-of-high-expectations.-.-.-and-no-excuses.html.

Frankfurt, H.G. 2005. *On Bullshit*. Oxford: Princeton University Press.

Frierson, P.R. 2016. 'Making Room for Children's Autonomy: Maria Montessori's Case for Seeing Children's Incapacity for Autonomy as an External Failing', *Journal of Philosophy of Education* 50(3): 332-50.

Gheaus, A. 2015. 'Is There a Right to Parent?', *Law, Ethics and Philosophy*.

333

Glasser, W. 1992. *The Quality School : Managing Students Without Coercion*, New York: Harper Perennial.

Gorin, M. 2014a. 'Do manipulators always threaten rationality?', *American Philosophical Quarterly*, 51(1), pp. 51-61.

Gorin, M. 2014b. 'Towards a Theory of Interpersonal Manipulation', in *Manipulation: Theory and Practice*, C. Coons, & Weber, M. (eds.), Oxford University Press, pp. 73-97.

Greely, H., B. Sahakian, J. Harris, R.C. Kessler, M. Gazzaniga, P. Campbell and M.J. Farah. 2008. 'Towards Responsible Use of Cognitive-Enhancing Drugs by the Healthy', *Nature* 4569 (7223): 702-05.

Green, J. 2011. *Education, Professionalism and the Quest for Accountability: Hitting the Target but Missing the Point*. London: Routledge.

Greenspan, P. 2003. 'The problem with manipulation', *American Philosophical Quarterly* 40(2): 155-64.

Hand, M. 2003. 'A Philosophical Objection to Faith Schools', *Theory and Research in Education*, 1(1): 89–99.

Hand, M. 2006. 'Against Autonomy as an Educational Aim', *Oxford Review of Education* 32(4): 535-50.

Harviainen, J. 2014. 'Critical Challenges to Gamifying Education: A Review of Central Concepts, in *Proceedings of the Game On! Conference*, 22-23.

Hattie, J. 2009. *Visible learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. London ; New York: Routledge.

Hildt, E. 2013. *Cognitive Enhancement: An Interdisciplinary Perspective*. New York: Springer.

Hood, C. 2006. 'Gaming in Targetworld: The Targets Approach to Managing British Public Services', *Public Administration Review* 66(4): 515--21.

Horner, R.H. and K. Mcintosh. 2016. Reducing Coercion in Schools: The Impact of school-Wide Positive Behavioral Interventions and Supports, in *The Oxford Handbook of Coercive Relationship Dynamics,* Dishion, T.J. and Snyder, J.J. (eds.), pp. 330-40.

Hyslop-Margison, E.J. and M.A. Nasee. 2007. *Scientism and Education: Empirical Research as Neo-Liberal Ideology*. Dordrecht: Springer.

Ismail, M.A.A. and J.A.M. Mohammad. 2017. 'Kahoot: A Promising Tool for Formative Assessment in Medical Education', *Education in Medicine Journal* 9(2).

James, W. 1907. *Pragmatism: A new name for some old ways of thinking; popular lectures on philosophy*. [S.l.]: Longmans.

Jaworska, A. 2005. 'From Caring to Self-Governance: The Bare Bones of Autonomy and the Limits of Liberalism', *Newsletter on Philosophy and Medicine* 5(1): 19-23.

Johns, K. 2015. 'Engaging and Assessing Students with Technology: A Review of Kahoot!', *Delta Kappa Gamma Bulletin* 81(4): 89.

Jussim, L. and K.D. Harber. 2005. 'Teacher Expectations and Self-Fulfilling Prophecies: Knowns and Unknowns, Resolved and Unresolved Controversies. *Personality and Social Psychology Review* 9(2): 131-55.

Kahneman, D. 2011. *Thinking, fast and slow*. London: Allen Lane.

Kahoot.Com. 2021. https://kahoot.com/.

Kant, I. 1993. *Grounding for the Metaphysics of Morals ; with, On a Supposed Right to Lie Because of Philanthropic Concerns*. 3rd ed. Indianapolis: Hackett Pub.

Kant, I. and H.J. Paton. (trans.), 1967. *The Moral Law; Kant's Groundwork of the Metaphysic of Morals*. New York,: Barnes & Noble.

Kapp, K.M. 2012. *The Gamification of Learning and Instruction*, San Francisco, CA: Pfeiffer.

Kapp, K.M., L. Blair and R. Mesch. 2014. *The Gamification of Learning and Instruction Fieldbook : Ideas Into Practice.* San Francisco, CA: Wiley.

Kasten, V., 1980. Manipulation and Teaching. *Journal of the Philosophy of Education*, 14(1), pp.53-62.

Keys, B. and J. Wolfe. 1990. 'The Role of Management Games and Simulations in Education and Research', *Journal of Management* 16(2): 307-36.

Killen, M. and J.G. Smetana. 2015. 'Origins and Development of Morality', *Handbook of Child Psychology and Developmental Science*: 1-49.

Killen, M., & Smetana, J. G. 2015. 'Origins and Development of Morality', in, *Handbook of Child Psychology and Developmental Science: Socioemotional Processes*, M. E. Lamb & R. M. Lerner (eds.), Wiley, pp. 701-749.

Kirschner, P.A. and J.J.G. Van Merriënboer. 2013. 'Do Learners Really Know Best? Urban Legends in Education', *Educational psychologist* 48(3): 169-83.

Klein, P. 2008. 'Useful falsehoods', in *Epistemology: New Essays*, Q. Smith (ed.), Oxford University Press.

335

Kodelja, Z. 2021. 'Intellectual Doping and Pharmaceutical Cognitive Enhancement in Education: Some Ethical Questions', *Journal of Philosophy of Education* 55(1): 167-85.

Kozol, J. 2007. *Letters to a young teacher*, New York: Three Rivers Press.

Labaree, D.F. 2010. *Someone Has to Fail : The Zero-Sum Game of Public Schooling*. Cambridge, Mass. ; London: Harvard University Press.

Legg, C. and C. Hookway. 2019. 'Pragmatism', in *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), Metaphysics Research Lab, Stanford University.

Licorish, S.A., H.E. Owen, B. Daniel and J.L. George. 2018. 'Students' Perception of Kahoot!'s Influence on Teaching and Learning', *Research and Practice in Technology Enhanced Learning* 13(1): 9.

Loewen, J.W. 1995. *Lies My Teacher Told Me : Everything Your American History Textbook Got Wrong*. New York: New Press.

Lunden, I. and N. Mascarenhas. 'Kahoot raises $28m for its user-generated educational gaming platform, now valued at $1.4b', *Techcrunch*. https://social.techcrunch.com/2020/06/11/kahoot-raises-28m-for-its-user-generated-educational-gaming-platform-now-valued-at-1-4b/.

Mahon, J.E. 2009. 'The Truth about Kant on Lies', in *The Philosophy of Deception*, Martin C. W. (ed.), Oxford University Press, pp. 201-24.

Marples, R. 1999. *The Aims of Education*, London: Routledge.

Martin, C.W. 2009. *The philosophy of deception*, Oxford ; New York: Oxford University Press.

Maslen, H., N. Faulmüller and J. Savulescu. 2014. 'Pharmacological Cognitive Enhancement—How Neuroscientific Research Could Advance Ethical Debate?', *Frontiers in Systems Neuroscience* 8: 107.

Mcleod, C. 2015. 'Trust', in *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), Metaphysics Research Lab, Stanford University.

Mcnaughton, D. and P. Rawling. 2000. 'Unprincipled Ethics', in *Moral particularism*, Hooker, B and Little, Mo (eds.), Oxford University Press, pp. 256-75.

Moreau, D., B.N. Macnamara and D.Z. Hambrick. 2018. 'Overstating the Role of Environmental Factors in Success: A Cautionary Note', *Current Directions in Psychological Science* 28(1): 28-33.

Mullin, A. 2007. 'Children, Autonomy, and Care', *Journal of Social Philosophy* 38(4): 536-53.

336

Mullin, A. 2014. 'Children, Paternalism and the Development of Autonomy', *Ethical Theory and Moral Practice* 17(3): 413-26.

Noddings, N. 1992. *The Challenge to Care in Schools: An Alternative Approach to Education*, New York ; London: Teachers College Press.

Noddings, N. 2001. 'Care and Coercion in School Reform', *Journal of Educational Change* 2(1): 35-43.

Noggle, R. 2018. 'The Ethics of Manipulation', in *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), Metaphysics Research Lab, Stanford University.

Orenstein, N. 'Use of Weighted Vest at Berkeley School Raises Questions About Parental Consent. https://www.berkeleyside.org/2018/01/30/use-weighted-vest-berkeley-school-raises-questions-parental-consent.

Orwell, G. 1953. *Such, Such Were the Joys*. New York,: Harcourt.

Paul, A.M. 'How to Use the "Pygmalion" Effect', Time.Com. https://ideas.time.com/2013/04/01/how-to-use-the-pygmalion-effect/.

Pede, J. 2017. 'The Effects of the Online Game Kahoot on Science Vocabulary Acquisition', Rowan University.

Peters, R.S. 1966. *Ethics and Education*, Allen & Unwin.

Peters, R.S., 2015. *Reason and Compassion: The Lindsay Memorial Lectures Delivered at the University of Keele, February-March 1971 and the Swarthmore Lecture Delivered to the Society of Friends 1972*, Routledge.

Pinola, M. 'How to Manipulate Your Kids Into Doing What You Want?', https://lifehacker.com/how-to-manipulate-kids-into-doing-what-you-want-1455891399.

Plato, A. Bloom (trans.). 2016. *The Republic of Plato*, New York: Basic Books.

Plump, C.M. and J. Larosa. 2017. 'Using Kahoot! In the Classroom to Create Engagement and Active Learning: A Game-Based Technology Solution for Elearning Novices', *Management Teaching Review* 2(2): 151-58.

Poláková, P. and B. Klímová. 2019. 'Mobile Technology and Generation Z in the English Language Classroom—a Preliminary Study', *Education Sciences* 9 (3): 203.

Porsdam Mann, S., P. De Lora Deltoro, T. Cochrane and C. Mitchell. 2018. 'Is the Use of Modafinil, a Pharmacological Cognitive Enhancer, Cheating?', *Ethics and Education*: 1-17.

Rancourt, B.T. 2015. 'Better Understanding Through Falsehood', *Pacific Philosophical Quarterly*, 98(3): 382-405

Rawls, J. 1972. *A Theory of Justice*. Oxford: Clarendon Press.

Ray, K.S. 2016. 'Not Just "Study Drugs" for the Rich: Stimulants as Moral Tools for Creating Opportunities for Socially Disadvantaged Students', *The American Journal of Bioethics* 16(6): 29-38.

Rosenthal, R. and L. Jacobson. 1966. 'Teachers' Expectancies: Determinants of Pupils' IQ Gains', *Psychological Reports* 19(1): 115-18.

Rosenthal, R. and L. Jacobson. 1968. 'Pygmalion in the Classroom', *The Urban Review* 3(1): 16-20.

Rousseau, J.J. and A. Bloom (trans.). 1979. *Emile: Or, on Education*. New York: Basic Books.

Schapiro, T. 1999. 'What is a Child?', *Ethics,* 109(4): 715-38.

Selwyn, N. 2014. *Distrusting Educational Technology: Critical Questions for Changing Times*, New York ; London: Routledge, Taylor & Francis Group.

Selwyn, N. 2016. 'Minding Our Language: Why Education and Technology is Full of Bullshit … And What Might Be Done About It', *Learning, Media and Technology* 41(3): 437-43.

Siegel, H., 2004. 'Faith, Knowledge and indoctrination: A friendly response to Hand.' *Theory and Research in Education*, 2(1): 75-83.

Singer, N. 'Kahoot App Brings Urgency of a Quiz Show to the Classroom', *Ney York Times,* https://www.nytimes.com/2016/04/17/technology/kahoot-app-brings-urgency-of-a-quiz-show-to-the-classroom.html.

Singh, I. 2005. 'Will the "Real Boy" Please Behave: Dosing Dilemmas for Parents of Boys with ADHD', *The American Journal of Bioethics* 5(3): 34-47.

Singh, I. 2014. 'Authenticity, Values, and Context in Mental Disorder: The Case of Children with ADHD', *Philosophy, Psychiatry, and Psychology* 21(3).

Singh, I., I. Bard and J. Jackson. 2014. 'Robust Resilience and Substantial Interest: A Survey of Pharmacological Cognitive Enhancement Among University Students in the UK and Ireland, *PLOS ONE* 9(10).

Sisk, V. F., Burgoyne, A. P., Sun, J., Butler, J. L., & Macnamara, B. N. 2018. 'To What Extent and Under Which Circumstances Are Growth Mind-Sets Important to Academic Achievement? Two Meta-Analyses', *Psychological Science*, *29*(4), 549-571.

Smeyers, P. 2006. 'The Relevance of Irrelevant Research; The Irrelevance of Relevant Research', in Smeyers P., Depaepe M. (eds), *Educational Research: Why 'What Works' Doesn't Work*, Springer, Dordrecht, pp. 95-108.

Smith, Q. 2008. *Epistemology: New Essays*, Oxford ; New York, Oxford University Press.

Sunstein, Cass R. 2014. *Why Nudge?: The Politics of Libertarian Paternalism*. Yale Univeristy Press.

Sunstein, Cass R., and Richard H. Thaler. 2003. 'Libertarian Paternalism Is Not an Oxymoron', *The University of Chicago Law Review* 70 (4): 1159-1202.

Sunstein. Cass R., and Richard H. Thaler. 2008. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. Yale University Press.

Tauber, Robert T. 2009. 'What Teachers Expect From Students, They Generally Get', *Education Matters* 4: 1-5.

Taylor, Charles. 1985. *Philosophical Papers: Human agency and language (Vol. 1)* Cambridge ; New York, Cambridge University Press.

Taylor, R.M., 2017. 'Indoctrination and Social Context: A System-based Approach to Identifying the Threat of Indoctrination and the Responsibilities of Educators', *Journal of Philosophy of Education*, 51(1): 38-58.

Tillson, John. 2021. 'Wrongful Influence in Educational Contexts', in *Oxford Encyclopedia of Philosophy of Education*, Hytten, K. (ed.), Oxford University Press.

Tschannen-Moran, Megan, and Wayne K. Hoy. 2000. 'A Multidisciplinary Analysis of the Nature, Meaning, and Measurement of Trust', *Review of Educational Research* 70 (4): 547-593.

Vrecko, Scott. 2013. 'Just How Cognitive Is "Cognitive Enhancement"? On the Significance of Emotions in University Students' Experiences with Study Drugs', *AJOB Neuroscience* 4 (1): 4-12.

Wang, A. Inge. 2015. 'The Wear Out Effect of a Game-Based Student Response System', *Computers & Education* 82: 217-227.

Ware, Alan. 1981. 'The Concept of Manipulation: Its Relation to Democracy and Power', *British Journal of Political Science* 11 (2): 163-181.

Wareham, R.J., 2019. 'Indoctrination, delusion and the possibility of epistemic innocence.' *Theory and Research in Education*, 17 (1): 40-61.

Warnick, Bryan R. 2017. 'Paying Students to Learn: An Ethical Analysis of Cash for Grades Programmes.' *Theory and Research in Education* 15 (1): 71-87.

White, John. 1990. *Education and the Good Life: Beyond the National Curriculum*. London: Kogan Page.

Williams, Bernard. 2002. *Truth and Truthfulness: an Essay in Genealogy*. Princeton, N.J. ; Oxford, Princeton University Press.

Wilson, J., 2010. 'Indoctrination and rationality', in *Concepts of indoctrination*, Snook, I (ed.), Routledge, pp. 14-19.

Winch, Christopher. 2006. *Education, Autonomy and Critical Thinking*. Routledge.

Wittgenstein, Ludwig. 1967. *Philosophical Investigations*. 2nd ed. Oxford,: Blackwell.

Wolff, Jonathan. 2012. *Ethics and Public Policy: A Philosophical Inquiry*. Routledge.

Wood, Alan. 2014. 'Coercion, Manipulation, Exploitation', in *Manipulation: Theory and Practice*, Coons, C. & Weber, M. (eds.), Oxford University Press, pp. 17-50.

Zicherman, Gabe. 2011. 'The Purpose of Gamification', http://radar.oreilly.com/2011/04/gamification-purpose-marketing.html.