De-Idealising the Educational Ideal of Critical Thinking

Abstract
It is widely recognised among educational theorists, educators, and policy makers alike, that critical thinking should claim a superordinate place in our system of educational objectives. In the philosophical literature on this topic, critical thinking is often conceptualised as the educational cognate of rationality, which in turn is analysed as being comprised of the relevant skills and abilities to assess reasons and evidence, together with the intellectual dispositions to actively use these proficiencies in practice. The resulting picture is in many respects normative and idealised, following the style of philosophical theorising commonplace in the tradition of analytic philosophy of education. In contrast, certain recent empirical findings related to the rational performance of actual human beings seem to cast doubts on the extent to which we can expect people to fulfil these idealised normative standards of rationality. After introducing the relevant philosophical theories and psychological results I ruminate on the implications these ideas have on our pedagogical views pertaining to critical thinking education.

Keywords
Analytic philosophy of education, debiasing, critical thinking, educational ideals, heuristics and biases, rationality

Introduction

It is widely recognised among educational theorists, educators and policy makers alike, that critical thinking—meaning goal-directed rational and evidence- based belief- formation, judgment and decision-making—should claim a superordinate place in our system of educational objectives (Dewey, 1933; Scheffler, 1989; Siegel, 1988; Robertson, 2009).[1] Although different philosophers of education have advanced their preferred more detailed explications of this notion, the mainstream theorising in this field seems to have settled on a general two-part model of critical thinking (Bailin and Siegel, 2003; Hitchcock, 2018). This bipartite characterisation includes on one hand, the set of logic-related reasoning skills needed to assess the validity of arguments and the epistemological acumen to assess the credibility of individual pieces of evidence, and on the other hand, the intellectual dispositions and virtues needed to put these skills to regular use in a way that is appropriate for an individual.

An eye-catching quality in the mainstream theorising on critical thinking is the methodological and topical focus on formal logic and argumentation theory (or informal logic), together with related theoretical issues drawn from the branch of philosophy known as analytic epistemology (Burbules and Berk, 1999). What has mostly been left out of these philosophical

analyses are the insights of well-established psychological studies on actual human reasoning and information- processing (Rysiew, 2008). In this article, I argue, that this non-empirical orientation has led to the creation of an abstract and idealised account of critical thought, which mischaracterises the prospects and limitations of critical thinking as an educational ideal, both in theory and in educational practice. The problem is that the philosophical model of critical thinking encourages us to approximate a paragon of an ideal reasoner, who is not encumbered by the inbuilt properties of our cognitive machinery, such as the workings of our information- processing and the systematic tendencies for sub-optimal functioning contained therein. On the pedagogical side, the non-empirical emphasis has encouraged an emphasis on teaching sound logical reasoning as the backbone of critical thinking courses and textbooks, which in itself does not protect against these psychological issues (but might in certain extreme cases even amplify them).

To remedy the situation, I suggest that mainstream theorising (and its pedagogical imperatives) is on the right tracks in broad strokes, but taken alone, is only a half-truth of the matter. When we incorporate the state-of-the-art empirical insights regarding human belief-formation and reasoning—namely, the key ideas of the dual- process theories of human cognition and the closely related research program called heuristics and biases—we stand to gain several steps forward regarding our understanding of critical thinking and its efficient pedagogy. Firstly, according to several authors, the definition and substance of critical thinking remains obscure and underdeveloped, especially on the side of critical thinking dispositions. In light of the psychological theories, at least some of the proposed elements of this critical spirit—such as the virtue of open-mindedness—can now be interpreted as folk psychological attempts to mitigate the genesis and effects of certain biases in our reasoning and judgment, such as the problematic phenomenon of confirmation bias (Lilienfeld et al., 2009). Secondly, educators have lamented that the most efficient ways to translate the philosophical model of critical thinking into pedagogical practice remain unclear. Here, too, the input of psychological literature can help us to better re-orient the focus of critical thinking didactics, as the experimental research from cognitive and social psychology puts us into a better position to understand the purpose and potential of these critical thinking dispositions, and also what can be done to foster them in our students. In short, debiasing strategies, such as the incentive to 'consider-the-opposite', can be used to counter at least some of the harmful effects of confirmation bias (Lord et al., 1984).

The general spirit of argumentation in this article parallels two lines of thinking, that are already familiar from broader contexts outside of the philosophy of education. The first of these is the criticism aimed at the idealised notion of economic rationality often assumed in the theories of neoclassical economics. This criticism was inaugurated by Herbert A. Simon (1955), and his notion of *bounded rationality*—the idea that we must approach the economical agents of these theories as being bound in terms of their cognitive capabilities and their surrounding situational factors—which ultimately led to the creation of the discipline of behavioural economics. The second parallel is more specific to philosophy, as it relates to the distinction between ideal and non-ideal theories, which has been brought up most notably in the sub-fields of ethics and political philosophy (Jaggar and Tobin, 2013). The basic premise is that philosophising, with its focus on universal analyses and abstract thought experiments,

creates idealisations regarding ethical norms and just societies, which in the process lose those particularised details about these phenomena that have crucial bearing on their nature. What I propose here, then can be considered an attempt to extend these familiar ideas into the realm of philosophy of education and to the topic of critical thinking more specifically.

The past century in critical thinking scholarship

In the literature on critical thinking, there exist numerous competing conceptualisations of this notion, as is to be expected. Nonetheless, it is possible to delineate a generic two-part blueprint of critical thinking, which is already present in embryonic form in Dewey's thinking (1933)[2], and which has since become largely endorsed as a starting point by many of the most prominent theorists in the field of philosophy of education (Paul, 1981; Siegel, 1988, 1997, 2017; Scheffler, 1973; Facione, 1990; Bailin et al., 1999; Hitchcock, 2018; Bailin and Siegel, 2006). Where the differences between various more specific formulations come in, is the way theorists articulate and flesh out the details of their particular versions of this general two-part layout. Even then, it can be argued that they are simply offering variant conceptualisations of the same basic concept, and the divergences are to a large degree verbal (Ennis, 2016).

To begin with, all theorists agree that critical thinking shares at least some *prima facie* overlap with the more general notion of epistemic rationality (known alternatively as theoretical rationality or evidential rationality; Stanovich, 2011: 6), although different theorists disagree on the extent of this conceptual overlap. Some philosophers, such as Siegel (1988; 1997; 2017), view critical thinking as the educational cognate of epistemic rationality, so that these two co-extensional terms are basically interchangeable and it is only in educational contexts where rationality is framed in terms of critical thinking— perhaps because it is then easier to conceptualise rationality as a teachable skill set. For Siegel, our account of critical thinking is thus firmly grounded in our general conception of epistemic rationality *simpliciter*. For others (e.g. Perkins et al., 1993: 3; Davies, 2015), the relationship between critical thinking and rationality is asymmetric, as the skills and propensities of critical thinking do not yet exhaust all facets of rational thought (which might additionally include, for example, aspects of creativity and problem-solving). In any case, for present purposes nothing of importance hinges on the outcome of this disagreement, as all theorists can agree that to be a critical thinker is to be ultimately a rational person—even if for some philosophers there might be more to the notion of rationality beyond mere critical thinking.[3]

As stated, the mainstream model of critical thinking shares the basic framework in which this multifarious phenomenon consists structurally of two equally important and mutually supportive parts. The first, and perhaps a more obvious part of this model, is a cluster of intellectual abilities and reasoning skills needed to carry out evidence-based belief-formation, sound decision-making, and reflective judgment. This cluster thus contains a host of standard higher-order thinking-skills, such as the capacity to analyse and evaluate statements and pieces of evidence, together with the logical skills needed to draw out the implicit conclusions contained therein. The more dialectical elements in this part of critical thinking relate to our aptitudes to participate in discussions and debates, where we are required to construct our own

reasoned arguments and, conversely, to evaluate the arguments given by others. Failures in this element of critical thinking are usually identified as fallacies, whether they are against the norms of formal (e.g. circular reasoning, affirming the antecedent, etc.) or informal (e.g. *ad hominem* counter-argument) logic. This skills-part of critical thinking thus requires at least rudimentary acquittance with various substantial matters related to logic and arguments, such as knowing the differences between induction, abduction and deduction, being able to identify the constituents of arguments, mastering the basics of statistical thinking and so on (indeed, such matters often form the foundation of critical thinking courses, textbooks, and tests).

In the pedagogical literature on this topic, certain authors have urged that comprehensive critical thinking instruction must incorporate numerous epistemological themes in addition to the typical focus on learning to form logically sound arguments (Bailin and Battersby, 2015: 124). Such an epistemologically oriented conception of teaching for critical thinking includes several philosophical issues related to truth and justification, like the question of relativism (Siegel, 1997: 14, 23).

As stated, this skills-part is perhaps somewhat more immediately ~~recognizable~~ recognisable as a component of what we consider good and rational thinking. Theorists have noted, however, that taken alone, the cluster of these reasoning skills and judgment abilities are insufficient to capture the full intuitive meaning of critical thought, since it leaves open the possibility that an individual knows very well what the relevant skills of critical thinking are and how they could be put to use, but for some reason does not follow through with the trouble of actually employing those skills in everyday life (Paul, 1981; Siegel, 1988). What is needed, then, is a group of propensities, character traits, epistemic virtues, and habits of mind, which together constitute the second part of critical thinking often styled as the 'critical spirit', following the lead of Passmore (1967). For Siegel (1988), a central ingredient in critical spirit is a master disposition, which he dubs 'critical spiritedness', and it is meant to animate a critical thinker to use her skills actively and willingly. Other frequently mentioned constituents in the critical spirit part of critical thinking are meant to prevent sophistry, that is, situations where a critical thinker knows how to think critically and will also do so actively, but employs her skills selectively in order to defend a self-serving agenda or other claims, which she knows to be untrue. We thus need qualities of character like love for truth (Scheffler, 1989), open-mindedness (Hare, 1979; Hamby, 2015), fair-mindedness (Bailin et al., 1999), self-confidence (Facione, 1990), and intellectual courage (Paul and Elder, 2006) to describe what kind of a *person* a critical thinker is.

Several theorists have lamented that the critical spirit part has remained underdeveloped compared to the skills-part, as the intricate rules of logical thinking—and the host of fallacies their violations produce—are often emphasised disproportionally in both theory and pedagogical practice (Burbules, 1995; Holma, 2015). The instruction of logic and argumentation has a long pedagogical tradition to draw from, which can make this emphasis understandable. Moreover, understood straightforwardly in terms of particular skills and their mastery, critical thinking prowess is also more testable, for example, in various measurements of higher education study skills (Hyytinen, 2015). However, this emphasis on critical thinking skills is also unfortunate, as the critical thinking dispositions can be described as 'at least half the battle of good thinking, and arguably more' (Perkins et al., 1992: 9). As I will argue later, I

believe that at least part of the explanation for this underdevelopment is the fact that philosophers have often used aphoristic language and folk-psychological explanations of sorts to describe the purpose and mechanisms of this element. As can perhaps be expected, it can then be hard to fathom how we should approach fostering these poetic dispositions, such as open-mindedness or love for truth, in our students. When these same ideas are rephrased in terms of psychological research, we can put this project on a firmer footing. In this case, the promotion of the virtue of open-mindedness can be seen as an intuitive antidote for certain malignant forms of confirmation bias and belief perseverance, where our mind tends to instinctively gravitate towards information that corroborates our prior beliefs and ignores conflicting pieces of evidence. These effects can be countered, at least to a degree, by so-called debiasing strategies, such as in this case following the advice of the consider-the-opposite strategy, where we actively challenge our entrenched beliefs by imagining alternative viewpoints in our minds (Lord et al., 1984). However, I will put these ideas on hold for a moment and turn next to exploring the academic background of critical thinking theorising.

Analytic philosophy of education and the idealisation of rationality

Although it is possible to outline a broadly neutral bare bones account of critical thinking such as the one suggested above, it is also an undeniable fact that there have been deep internal disagreements within the scholarship on this topic.[4] Despite such quarrels, however, what has been distinctive in these debates—and to the theory-building on critical thinking more generally—is that the quality of this research program has been decidedly non-empirical in its approach, being conducted almost exclusively through philosophical considerations and arguments without any significant input from the neighbouring fields of cognitive and social psychology, which investigate human cognition and reasoning experimentally. There have been, it must be said, a limited number of publications on the subject of critical thinking that incorporate psychological matters—often related to metacognitive skills and their development—into their viewpoints.[5] These texts have remained in the minority, and they have typically not stimulated follow-up comments or cross references in the more philosophically oriented literature.

This non-empirical leaning in the mainstream critical thinking literature is explained by the broader meta-theoretic background of this research project. Namely, this work has generally been carried out by authors who can be most fittingly described as being aligned with the school of analytic philosophy of education, which has been, especially in English speaking countries, the preeminent approach in the philosophy of education since the 1960s (Burbules and Berk, 1999). In brief, the analytic philosophy of education values conceptual clarity, sharp distinctions, and unambiguous arguments as the end-product of philosophising (Curren et al., 2003). In this particular case of critical thinking, the philosophical analysis has—with the aid of thought-experimenting on the question of what it means to be a critical thinker—delineated the necessary and sufficient conditions for the normatively adequate application of this concept. What has been left outside of the analyses, in contrast, are the empirical studies on

features of human cognition, as these philosophical studies have placed methodological and topical focus on matters of formal logic, argumentation theory, and analytic epistemology.

It must be stated that this exclusion of empirical sources is not a symptom of any scientific ignorance or naive anti-scientism, as these philosophers are aware of the relevant scientific research (e.g. Siegel, 2015). Rather, this approach is backed by meta-theoretical considerations and a long philosophical tradition, which maintain that the normative and descriptive levels of inquiry should be kept separate.[9] When we are outlining the rules of proper critical thought, we are clearly tracing the contours of a normative notion. These normative standards of rationality cannot be derived from the observations on how actual people behave—'no ought from is', as the old Humean adage states. Therefore, mixing logical laws and psychological laws is perceived as a major blunder in philosophical inquiry, often derided under the name of psychologism (Kusch, 2020). The proper guiding role of empirical information comes at the stage of implementing these philosophical models in pedagogical practice.

Nevertheless, I argue that this 'philosophy first' doctrine can be criticised on the basis that following this methodological approach, our account of critical thinking becomes overtly idealised. To be more precise, if we base our standards of adequate critical thinking on abstract rules of logic and theories of analytic epistemology without taking into account the prospects and limitations of humans as cognitive agents, our view mischaracterises both the phenomenon of critical thinking and the pedagogical techniques needed to best attain this ideal. The 'epistemic agent' who is actually trying to fulfil the normative standards of critical thinking is not an immaterial Cartesian thinker or a disembodied 'brain-in-a-vat', but an actual human being of flesh and blood with limited computational capacities, restricted ~~amount~~ amounts of psychophysical energy resources, subconscious processes, finite memory, inbuilt biases, and so on. Moreover, this thinker does not exercise her critical thought in any 'philosophical bubble', but in real-life situations with time-constraints and other situational factors coming into play. All these issues place reservations on what we can hopefully try to achieve, and what the best methods are trying to reach this potential. Hume's 'no ought from is' is a valid doctrine for normative theory building, but at the same time the Kantian doctrine of 'ought implies can' is equally true regarding the outcomes of this theorising—we must apply normative standards and educational ideals which are realistically attainable for creatures like us (Stich, 1990: 27; Bishop and Trout, 2004).

Of the notable critical thinking theorists, Siegel (2017: 104–5n1), for example, is aware of this potential tension between the abstracted requirements of critical thinking and the potential of actual persons. To remedy the situation, he proposes a view which might be described as *gradualism*: Thinking critically is not a clear-cut affair, but rather, it must be seen as an ideal which we, as actual human beings with our fallible cognitive capacities, can attain only more or less. In other words, at some point there is an indefinable threshold, above which our thinking becomes good enough to qualify as truly critical—there exists no straightforward checklist of requirements against which the rationality of our performance can simply be measured and judged.[6]

What is important to note here, however, is that Siegel's thinking is motivated primarily by epistemic fallibilism, that is, the idea that we should allow people to err in their judgment

from time to time without judging them as complete failures as a result. To use the Chomskyan distinction between competence and performance, we can say that episodes of individual failures in performance do not necessarily invalidate the person's competence (Chomsky, 1965). However, it seems that for Siegel, we still exist somewhere on the same spectrum as the ideal reasoner, so that we just fall short of these normative expectations—as long as we do not commit any thinking mistakes, it is in principle possible to attain this ideal state. We are situated somewhere along a gradient from completely failed forms of uncritical thinking at one end to ideal critical thinking at the other.

I believe that the mainstream model of critical thinking and Siegel's views regarding our gradual attainment of this ideal are good starting points in our understanding of this phenomenon, but they are only half of the truth. The other half comes from empirical studies on human cognition. What I am suggesting here, however, is *not* that we simply take any piece of science at face value and try to derive our norms of critical thinking from this research. The division of labour between philosophy and the empirical sciences and the is/ought distinction still remain, so that the conceptual, methodological and normative matters related to this topic are of a philosophical nature, even if they must be informed by the empirical research. What the empirical research now does, is give us pointers on what the realistic endpoint of the spectrum is when we take into account the particular kind of mental hardware we have and the quirks it involves. The scientific research in this area is not finalised, and the follow-up discussions about the interpretation of individual experiments and even their larger meta-theoretic foundations continue actively within the scientific community. Nevertheless, the broad lines converge on certain views, which bear directly on our model of critical thinking. I now turn to this research and its implications for the theory and practice.

The science of human rationality

The dominant current in the modern experimental study of human reasoning is the research program known informally as *heuristics and biases* (Gilovich et al., 2002), which has been applauded as 'one of the crowning achievements of modern psychological science' (Lilienfeld et al., 2009: 390). Going back to the pioneering work of Peter Wason in the 1960s and followed by several influential studies by Amos Tversky and Daniel Kahneman (Wason, 1960; Kahneman et al., 1982), this paradigm has produced countless studies on various aspects of human reasoning over the subsequent decades, resulting in three Nobel laureates and some of the most- cited scientific publications of the previous century. By now, there is a strong indication of internal consistency and cross-task correlation between the results of these studies. This corroborates the general conclusion of this research, which is that human belief-formation, reasoning, and judgment quite often deviates systematically from the expected normative guidelines of rationality (Evans and Over, 1996; Stanovich and West, 2000).

The body of research accumulated on heuristics and biases converge on a view of human cognition known as *dual process theory*.[7] It is helpful to summarise the main points of this model, as it can give us a better understanding of critical thinking, its challenges, and what we can do to respond to these issues. As its name suggests, dual process theory maintains that our

cognitive system contains two distinct varieties of mental processes, which have been given various pairs of names in the literature, such as analytic versus intuitive modes of thinking (Kahneman 2003). The most atheoretical and descriptive way of drawing this distinction, however, seems to be to speak of Type 1 and Type 2 processes. The core idea is that belief can arise in our mind in two qualitatively distinct ways. Type 1 processes can be described as unconscious, concrete, context-specific, implicit, rapid, low effort, parallel, and automatic. Type 2 processes, in direct contrast, are conscious, explicit, high effort, slow, controlled, abstract, sequential, and deliberative. This description of Type 2 processing matches the traditional philosophical accounts of critical thinking skills and abilities, so the relevant normative prescription issued by our account of critical thinking can be understood as an instruction to engage in the logical reasoning and reflective belief-formation distinctive to Type 2 processing. However, as several authors have stressed (Stanovich, 2011; Bonnefon, 2015), in a more fine-grained analysis the relationship between the philosophical notion of critical thinking and the psychological phenomenon of Type 2 processing is not as straightforward as it might appear at first glance. The reason for this is that the mere shift into Type 2 processing is not in itself enough to improve the quality of our rational performance, as intelligent people can use their mental capacities to rationalise biased and emotionally motivated beliefs. What is actually needed for rational thought, is a more specific sub-type of Type 2 processing (I return to this matter in the next section).

Because engaging these Type 2 processes is comparatively more time-consuming and ~~it also~~ drains our psychophysical energy by taxing our central processing and nervous system, our mind usually operates by default ~~operates on the level of~~ on Type 1 thinking and switches to Type 2 processing only when required. The mainstream account of critical thinking and the notion of an ideal reasoner do not take this into consideration, and assumes that we can switch to Type 2 at any given moment without any costs. Here I follow Bishop and Trout (2004, 79), who opine that '[a]ny epistemological theory that aspires to guide reason must recognize that we are limited creatures, and as such we have to make choices about how to spend our cognitive resources.' These matters should be taken into account~~,~~ when we ~~state~~ determine our expectations for critical thinkers in the pedagogy of this topic.

To make the most of these limited resources, our mind utilises heuristics in belief-formation and judgment at the level of Type 1 processes. These heuristics can be defined as strategies to reach a satisfying belief or decision quickly and cost-effectively. The way they achieve this is to simplify the vast amount of information and stimuli available to us at any given time by focusing on the relevant features of this information flow. What they might lose in accuracy, they save in time and energy. These heuristics additionally protect the integrity of our internal belief- system by avoiding the introduction of conflicting beliefs into the mix, which might cause stressful and paralysing situations of cognitive dissonance. The problem is that although these heuristics get things right most of the time—or, at least close enough—and are integral for our everyday lives, they also lead to harmful biases (Tversky and Kahneman, 1974: 1124). These biases can be described as systematic and persisting deviations from the normative standards of critical thinking, producing distorted beliefs and fallacious reasoning. As such, they are sometimes also described as cognitive illusions (akin to the optic illusions inherent in our visual system).

A paradigmatic illustration of the dual process theory in action is the phenomenon of belief bias. This effect has been produced in studies, where participants are asked to assess the validity of the two-statement syllogisms presented to them. In these syllogisms, the conclusion is highly believable, although the premises do not validly support it, such as in the case of the syllogism 'All flowers have petals, roses have petals, therefore roses are flowers' (Stanovich and Stanovich, 2010: 197). When the subjects are asked to assess these syllogisms, they typically jump to the conclusion on the basis of its believability, and thus cut past the laborious task of using Type 2 processes to examine the actual validity of the argument. Our mind does not follow the linear logical script of going through the premises one by one and then assessing them on the whole, but rather, reaches the conclusion 'from the hip' based on the believability of the entire statement (Thagard, 2011).

Belief bias is an example of a cognitive bias (or 'cold' bias), but there are also related instances of affective biases (or 'hot' biases), which result from motivated reasoning when our fears and wishes sway our judgment (Correia, 2011). A well-known example here is the case of confirmation bias, also known as myside bias. We cannot take the time and use the energy to assess every piece of evidence, so we gravitate towards those stimuli that corroborate our prior beliefs and interpret ambiguous evidence favourably. Conversely, a related phenomenon here is belief perseverance, which means that we tend to hold on to our existing beliefs even in the face of contradicting evidence by practicing 'slothful induction'.

Tentative ideas about the future of critical thinking pedagogy

The experimental findings and theories outlined above have provoked two noteworthy follow-up discussions regarding the proper interpretation and further implications of these results. I mention these topics briefly here, because they seem to bear directly on the future relevancy of critical thinking education. The first of these controversies has been called 'the great rationality debate' (Tetlock and Mellers, 2002), 'rationality wars' (Samuels et al., 2002), and 'rationality disputes' (Rysiew, 2008) by various authors. At the centre of this disagreement has been the question of what kind of conclusions these results point towards regarding the true quality of human cognition. On one side, there have been those theorists who have opined that these disconcerting results now prove—at least to some degree—that humans are actually more irrational than what has been generally assumed (this broad camp includes most of the major researchers in the heuristics and biases tradition, such as Kahneman). On the other side of this argument are those theorists who want to use these results to reinterpret our conception of rationality so that these heuristics and biases now demonstrate how our minds have adapted as part of human evolution to perform optimally in natural settings (Gigerenzer, Todd and The ABC Research Group, 1999). What might thus appear irrational about the theoretical standards of logic and epistemology can still in some sense be viewed as fit behaviour relative to past environments where our ancestors have tried to thrive. These interpretative differences aside, all parties can broadly agree that the aforementioned studies are significant and demonstrate that our performance is often problematic, especially in the context of our contemporary and 'unnatural' situations, where our behaviour often negatively impacts our

wealth and health (Stanovich, 2011). It is therefore clearly in our best interest to try to minimise the effects of these malfunctioning heuristics and biases in our lives.

Another pedagogically relevant question has focused on the issue of whether people still have the full epistemic obligation to try to think straight, as it has been demonstrated that at least some of our erroneous thinking might be beyond our conscious control and caused by deep inbuilt tendencies in our cognitive machinery. On this issue, some philosophers have argued that an individual can still be held responsible for their epistemic actions and that they should try to think rationally to the extent that it is possible (Correia, 2014). Our situation is not entirely incurable, as there are various ways to overcome these documented difficulties. It therefore seems that critical thinking education still has a genuine goal and a purpose to serve.

I now sketch in a few words some tentative wisdom regarding the contents, methods, and limits of critical thinking pedagogy, based on the literature in the field. As was noted earlier, the scope of critical thinking education has undergone several expansions in the past five or so decades. The first expansion extended the toolbox of critical thinking from the principles of formal logic to the additional issues belonging to the everyday 'marketplace' argumentation studied under the banner of informal logic. More recently, philosophers have argued that proper critical thinking instruction must also cover at least some epistemological ground (for example, the topics of truth, justification, and relativism). What is needed now, it seems, is a similar expansion into the domain of cognitive and social psychology. Of course, this move does not depose formal or informal logic from the equation, as the best antidote to illogicality is still learning the basics of logical thinking (Correia, 2014). It has been argued that internalising the logical way of looking at things can help in bringing down the psychophysical costs attached to the thinking process when we do not have to rely on the slow Type 2 processing as much (Hogarth, 2001: Stanovich et al. 2016: 133). However, this can only get us so far, as the psychological literature specifically points out that even the best logical thinking can fall prey to biased information processing and motivated reasoning. We therefore need a more versatile set of counterstrategies.

The obvious first step is to simply include information about biases in education, so students know what to look for in their own performance and in that of others – as is already done by some textbooks in the field, such as those of Warburton (2007) and Halpern (2014). This can help to facilitate what is called conflict detection (Stanovich et al., 2016): we recognise situations where we must be on our toes and not trust our first instincts. A bigger step forward in this area is to introduce so-called debiasing strategies in our textbooks and courses. Debiasing refers to those strategies (or mix of strategies) that are designed to suppress or mitigate either the genesis of biases or their harmful effects—typically so that certain proposed debiasing strategies are meant to mitigate a specific type of bias (Lilienfield et al., 2009: 391). Unfortunately, there are generally much fewer studies on debiasing than there are on heuristics and biases, and the studies that do exist have displayed mixed results. However, even if the research is still open and subject to change, we can still draw some moderate optimism for critical thinking pedagogy here (Ritola, 2015; Correia, 2018). One such corrective debiasing strategy is the advice of 'consider the opposite'. This recommendation means that we should actively try to imagine in our mind what it would mean (and how it could be argued for) if our certain belief was in fact false and the opposite was true. This causes us to reconsider our

entrenched beliefs, such as those which can form the basis for confirmation bias (Lord et al., 1984). In Stanovich's (2011) explanation such a move helps us to detach and decouple us from the misleading details in the information available to us. Mere logical thinking—named the algorithmic mind by Stanovich—is insufficient, as it might simply rationalise our biased beliefs and judgments with Type 2 processing. What is needed, is the input of our reflective mind, which is responsible for simulating alternative scenarios with our Type 2 processes. In the example here, we engage in such simulation to debias our tendencies for the tunnel vision and belief perseverance associated with confirmation bias.

Regarding the state of critical thinking education, putting an emphasis on these debiasing strategies could perhaps help to put critical spirit dispositions in our account of critical thought on a firmer footing. Indeed, several authors have noted, that the empirical results of the heuristics and biases tradition seem to provide empirical support for the idea that critical thinking includes dispositional elements (Ritola, 2015). In my assessment, much of the philosophical literature on critical thinking dispositions—and the problematic types of thought they are presumed to counter—is already folk psychological in its approach, in the sense that philosophers have used poetic language to identify and correct biased thinking as it happens observably in everyday situations. Thus, for example, the disposition to be open-minded can help to suppress egocentric beliefs and self-serving illusions, which arise as a result of confirmation bias and belief perseverance (Paul, 1990). Similar dispositions are often given a prominent role also in the psychological theories on rationality (Stanovich et al., 2016; Evans, 2008: 262–3).[8] What has potentially held back the pedagogical progress on this front has been the avoidance of the influence of empirical theorising in our philosophical models. Being more welcoming to the experimental results of these psychological studies can therefore give us a clearer picture of what we are actually trying to suppress and how it should be best pursued in a practice-oriented manner.

A final, and a more defeatist, observation based on the literature here is that some biases might ultimately be beyond our pedagogical reach in the sense that we cannot correct them, even through the debiasing efforts of these critical thinking dispositions. What is needed, instead, is a qualitative change in the environments or 'choice architectures' where we conduct our thinking (Thaler & Sunstein, 2008). This move might be called contextual debiasing in contrast to the more individualistic forms of debiasing discussed above (Kenyon and Beaulac, 2014). The proposed forms of contextual debiasing include, for example, practices of blind review, where an individual is restricted from access to certain information that might influence her judgment and trigger forms of prejudiced thinking. In extreme cases, this might mean outsourcing our thinking to algorithmic tools (Bishop and Trout, 2004). As these matters are outside of critical thinking education, I leave them largely aside here. I highlight here only the idea that sometimes what is needed to debias us is engaging actively in discussion with others, as this can raise the quality of our thinking when we feel the positive pressure to be accountable to our interlocutors. Our audiences can obviously also provide valuable feedback and recognise our failures better than we do ourselves.[9] It seems that this suggestion gives some vindication to the proposal that critical thinking should be viewed more as a social and collaborative practice, rather than as a personal responsibility or achievement (Burbules and Berk, 1999). Overall, what this notion—and the approach of contextual debiasing more generally—seems

to demonstrate, is where the upper limits lie for us as the epistemic agents, who are expected to fulfil the ideals of critical thinking. Being aware of these limitations can therefore help us in some sense to de-idealise the educational ideal of critical thinking while retaining it as a valuable aim in our pedagogical practices.

Conclusion

This article has argued that the mainstream model of critical thinking in the philosophical literature is overtly idealised when we take into account certain recent empirical findings related to the rational performance of actual human beings. These findings seem to cast doubts on the extent to which we can expect people to fulfil these idealised normative standards of rationality. Although the philosophical view is a good basis for our theorising, it is only a half of the truth in this matter. When we incorporate the state-of-the-art empirical insights regarding human belief-formation and reasoning—namely, the key ideas of the dual-process theories of human cognition and the closely related research program called heuristics and biases—we stand to gain several steps forward in our understanding of critical thinking and its efficient pedagogy. Namely, we gain a picture of what the psychophysical costs of critical thinking in reality are, and how these costs might constrain our critical thinking activities. Additionally, the experimental results regarding cognitive and affective biases provide us with a clearer picture of what the challenges for critical thinking are and how they might be countered through debiasing strategies. This gives us a better understanding of critical thinking dispositions, in theory, and in pedagogical practice.

Notes

[1] The value of critical thinking as an educational ideal has been advocated on many grounds. Societally, and most importantly, critical thinking prowess has been linked to healthy and participatory democratic citizenship, as it is these intellectual skills and tendencies that enable citizens to follow public discussions and come to their own informed verdicts regarding the important matters discussed in public forums. Of the international educational organisations with interest in critical thinking education, this positive dimension of critical thinking has typically been highlighted by UNESCO (1995). The second major reason for the current promotion of critical thinking comes from the fact that these thinking skills are thought to improve our students' employability in the rapidly changing knowledge-based and digitalised post-industrial work markets of tomorrow, where—in place of narrow task-specific expertise—general-purpose transferable competences such as critical thinking become valuable and (supposedly) enable workers to adapt to future vocations that might not even exist yet. This vocational dimension of critical thinking is highlighted by OECD (2018). Critical thinking has also been linked to certain other currently popular educational objectives, such as information literacy, lifelong learning (Green, 2015), and creativity (Bailin, 1987).

² Dewey (1933: 133) described 'reflective thinking" as "active, persistent, and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it and the further conclusions to which it tends'.

³ Recognising this relationship between critical thinking and rationality is helpful, because it helps to bridge the gap between philosophical and psychological literatures, as these matters are more often discussed in terms of rationality in the latter form of research.

⁴ Perhaps the most notable centre of attention has been the question of whether there is just one valid kind of critical thinking fit for all contexts, or whether we should acknowledge that there can be some concurrent and domain-specific modes of reasoning and principles of assessing evidence (McPeck, 1981; Ennis, 1989). The outcome of this theoretical debate has tangible pedagogical repercussions as well, namely, whether critical thinking should be best taught in its own dedicated courses, or whether its themes should instead be blended into the syllabuses of individual academic disciplines in a way that strengthens their idiosyncratic forms of reasoning and argumentation. Besides this issue, another major internal disagreement regarding the nature of critical thinking has focused on the question of whether these ideals of critical thinking and reasoning simply privilege masculine or Western modes of thought while they are not truly universal and shared by everyone around the globe (Thayer-Bacon, 1992). The criticism on this point goes that the notion of objective and monological rationality is seen as a conservative and even dangerous idea, which can easily be misused (either deliberately or unintentionally) to suppress and marginalise alternative ways of thinking as 'irrational' or 'illogical'.

⁵ Some notable examples here are Halonen (1995), Halpern (1998), Kuhn (1999), Bishop and Trout (2004), Rysiew (2008), Kenyon and Beaulac (2014) and Lau (2015). As an anonymous reviewer helpfully pointed out, the normative character of critical thinking as an educational phenomenon is already intrinsically present in the very fact that it is valued as a desirable goal in educational practices. Speaking generally, our theoretical analysis of what critical thinking is becomes normativised when it is regarded as something we should try to attain, in this case as the endpoint of education. In some sense, this should perhaps be obvious, but it is not always sufficiently recognised in the empirical theorising on the matter, particularly in the field of higher education studies, where the developmental models of academic skills (such as critical thinking) sometimes explicitly claim to be purely descriptive and non-philosophical, yet their supposedly neutral accounts evaluate the improvement of students' maturation in a clearly normatively loaded fashion. For this criticism, see Holma and Hyytinen (2015). In comparison, the theorisation regarding human reasoning in cognitive psychology is often aware regarding of the interplay of normative and descriptive elements within their theories (Stanovich, 2011). However, even here, certain accounts of so-called ecological rationality seem to come close to attempting to derive our norms regarding rational behaviour from empirical observations (see Gigerenzer, Todd and The ABC Research Group, 1999).

⁶ A similar idea is espoused by Bailin et al. (1999: 287): 'Fulfilling relevant standards in thinking is, of course, not an all or nothing affair. This being the case, we sometimes talk about good and poor critical thinking to indicate the degree of fulfilment of relevant standards. When someone's thinking is very poor we may simply say that the person is not thinking critically, even though he or she may be striving to fulfil the relevant standards.'

[7] There is no single accepted account of dual processing, but rather a family of theories that share certain core tenets and disagree on others. For an overview, see Evans (2008). For a critical take on the subject, see Osman (2004).

[8] Psychological discussions on critical thinking dispositions have centered on open-mindedness, as it is regarded as the most central disposition which also predicts an individual's success in critical thinking tests (e.g. Stanovich et al., 2016).

[9] Sadly, even this strategy is not a perfect silver bullet to uncritical thinking, as despite all our best intentions it can backfire and actually lead to the social form of biased thinking known as groupthink (Ritola, 2015).

References

Bailin, S (1987) Critical and creative thinking. *Informal Logic*, 9(1): 23–30.

Bailin S, Case R, Coombs, JR and Daniels LB (1999) Conceptualizing critical thinking. *Journal of Curriculum Studies*, 31(3): 285–302.

Bailin, S and Siegel, H (2006) Critical thinking. In: Blake N, Smeyers P, Smith RD and Standish P (eds.) *The Blackwell guide to the philosophy of education*, pp. 181–193. London: Blackwell

Bailin S and Battersby M (2015) Teaching critical thinking as inquiry. In: Davies M and Barnett R (eds.) *The Palgrave handbook of critical thinking in higher education*, pp. 123–138. New York: Palgrave Macmillan.

Bishop MA and Trout JD (2004) *Epistemology and the psychology of human judgment*. Oxford: Oxford University Press.

Bonnefon, J-F (2018) The pros and cons of identifying critical thinking with system 2 processing. *Topoi*, 37(1): 113–119.

Burbules NC (1995) Reasonable doubt: Toward a postmodern defense of reason as an educational aim. In: Kohli W (ed.) *Critical Conversations in Philosophy of Education*. pp. 82–102. New York: Routledge.

Burbules NC and Berk R (1999) Critical thinking and critical pedagogy: Relations, differences, and limits. In: Popkewitz TS and Fendler L (eds.) *Critical theories in education: Changing terrains of knowledge and politics*, pp. 45–65. New York: Routledge.

Chomsky N (1965) *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Correia V (2011) Biases and fallacies: The role of motivated irrationality in fallacious reasoning. *Cogency*, 3(1): 107–126.

Correia V (2018) Contextual debiasing and critical thinking: Reasons for optimism. *Topoi*, 37(1): 103–111.

Curren R, Robertson E and Hager P (2003) The analytical movement. In: Curren R (ed.) *A companion to the philosophy of education*, pp. 176–191. Oxford: Blackwell.

Davies, M (2015) A model of critical thinking in higher education. In: Paulsen, MB (ed.) *Higher education: Handbook of theory and research*, 30, pp. 41–92. Switzerland: Springer International Publishing.

Dewey J (1933) *How we think: A restatement of the relation of reflective thinking to the educative process* (revised edition). Boston: D.C. Heath and Company.

Ennis, RH (1989) Critical thinking and subject specificity: Clarification and needed research. *Educational Researcher*, 18(3): 4–10.

Ennis, RH (2016) Definition: A three-dimensional analysis with bearing on key concepts. In: Bondy P and Benacquista L (eds.) *Argumentation, objectivity, and bias: Proceedings of the 11th international conference of the Ontario Society for the Study of Argumentation (OSSA), 18–21 May 2016*, pp. 1–19. Windsor, ON: OSSA.

Evans JstBT (2008) Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59: 255–278.

Evans JstBT and Over DE (1996) *Rationality and reasoning*. Padstow: Psychology Press.

Evans JstBT and Stanovich KE (2013) Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, 8(3): 223–241.

Facione PA (1990) *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*. Millbrae, CA: The California Academic Press.

Gigerenzer G, Todd PM and The ABC Research Group (eds.) (1999) *Ecological rationality: Intelligence in the world*. New York: Oxford University Press.

Gilovich T, Griffin D and Kahneman D (eds.) (2002) *Heuristics and biases: The psychology of intuitive judgment*. Cambridge: Cambridge University Press.

Green P (2015) Teaching Critical Thinking for Lifelong Learning. In: Davies M and Barnett R (eds.) *The Palgrave handbook of critical thinking in higher education*, pp. 107–121. New York: Palgrave Macmillan.

Hamby B (2015) Willingness to inquire: The cardinal critical thinking virtue. In: Davies M and Barnett R (eds.) *The Palgrave handbook of critical thinking in higher education*, pp. 77–87. New York: Palgrave Macmillan.

Halonen J (1995) Demystifying critical thinking. *Teaching of Psychology*, 22(1): 75–81.

Halpern DF (1998) Teaching critical thinking for transfer across domains: Disposition, skills, structure training, and metacognitive monitoring. *American Psychologist*, 53(4): 449–455.

Halpern DF (2014) *Thought and knowledge: An introduction to critical thinking* (5th edition). New York: Psychology Press.

Hare W (1979) *Open-mindedness and education*. Montreal: McGill-Queen's University Press.

Hitchcock D (2018) Critical thinking. In: Zalta E (ed.) *The Stanford Encyclopedia of Philosophy*. Available at: https://plato.stanford.edu/entries/critical-thinking/ (accessed 11 June 2020).

Hogarth RM (2001) *Educating intuition*. Chicago: University of Chicago Press.

Holma K (2015) The critical spirit: Emotional and moral dimensions of critical thinking. *Studier i Pædagogisk Filosofi*, 4(1): 17–28.

Holma K and Hyytinen H (2015) The philosophy of personal epistemology. *Theory and Research in Education*, 13(3): 334–350.

Hyytinen H (2015) *Looking byond the obvious: Theoretical, empirical and methodological insights into critical thinking*. University of Helsinki, Institute of Behavioural Sciences, Studies in Educational Sciences 260. Helsinki: Unigrafia.

Jaggar AM and Tobin TW (2013) Situating moral justification: Rethinking the mission of moral epistemology. *Metaphilosophy* 44(4): 383–408.

Kahneman D (2003) A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, 58(9), 697–720.

Kahneman D, Slovic P and Tversky A (1982) *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.

Kenyon T and Guillaume B (2014) Critical thinking education and debiasing. *Informal Logic*, 34(4): 341–363.

Kuhn D (1999) A developmental model of critical thinking, *Educational Researcher*. 28(2): 6–25+46.

Kusch M (2020) Psychologism. In: Zalta E (ed.) *Stanford Encyclopedia of Philosophy*. Available at: https://plato.stanford.edu/entries/psychologism/ (accessed 11 June 2020).

Lau JYF (2015) Metacognitive education: Going beyond critical thinking. In: Davies M and Barnett R (eds) *The Palgrave Handbook of Critical Thinking in Higher Education*. New York: Palgrave Macmillan.

Lilienfeld S, Ammirati R, Landfield K (2009) "Giving debiasing away". *Perspectives on Psychological Science*, 4(4): 390–398.

Lord CG, Lepper MR and Preston E (1984) Considering the opposite: A corrective strategy for social judgment. *Journal of Personality and Social Psychology*, 47(6): 1231–1243.

McPeck JE (1981) *Critical thinking and education*. New York: St. Martin's Press.

OECD Centre for Educational Research and Innovation (2018) *Fostering and assessing students' creative and critical thinking skills in higher education*. Paris: OECD. Available at: http://www.oecd.org/education/ceri/Fostering-and-assessing-students-creative-and-critical-thinking-skills-in-higher-education.pdf (accessed 11 June 2020).

Osman M (2004) An evaluation of dual-process theories of reasoning. *Psychonomic Bulletin & Review*, 11: 988–1010.

Paul RW (1981) Teaching critical thinking in the 'strong' sense: A focus on self-deception worldviews, and a dialectical mode of analysis. *Informal Logic*, 4(2), 2–7.

Paul RW (1990) *Critical thinking: What every person needs to survive in a rapidly changing world*. Rohnert Park, CA: Center for Critical Thinking and Moral Critique.

Paul RW and Elder L (2006) *The miniature guide to critical thinking: Concepts and tools, 4th edition*. Dillon Beach, CA: Foundation for Critical Thinking.

Passmore J (1967) On Teaching to be Critical. In: Peters RS (ed.) *The Concept of Education*, pp. 192–211. London: Routledge & Kegan Paul.

Perkins DN, Jay E, & Tishman S. (1992) *Assessing thinking: A framework for measuring critical thinking and problem solving skills at the college level*. Washington, DC: The National Center for Educational Statistics Workshop on the Assessment of Higher Order Thinking and Communication Skills of College Graduates: Preliminary Listing of Skills and Levels of Proficiency.

Perkins DN, Jay E, & Tishman S. (1993) Beyond abilities: A dispositional theory of thinking. *Merrill-Palmer Quarterly*, 39(1), 1–21.

Ritola J (2015) Deliberative democracy, critical thinking, and the deliberating individual: empirical challenges to the reasonability of the citizen. *Studier I Pædagogisk Filosofi*, 4(1), 29–54.

Robertson E (2009) The epistemic aims of education. In: Siegel H (ed.) *The Oxford Handbook of Philosophy of Education*, pp. 11–34. Oxford: Oxford University Press.

Rysiew P (2008) Rationality disputes: Psychology and epistemology. *Philosophy Compass*, 3(6): 1153–1176.

Samuels S, Stich S and Bishop M (2002) Ending the rationality wars: How to make disputes about human rationality disappear. In Elio R (ed.). *Common sense, reasoning and rationality*, pp. 236–268. New York: Oxford University Press.

Scheffler I (1973) *Reason and teaching*. New York: Bobbs-Merrill.

Siegel H (1988) *Educating reason: Rationality, critical thinking, and education*. New York, N.Y.: Routledge.

Siegel H (1997) *Rationality redeemed: Further dialogues on an educational ideal*. London: Routledge.

~~Siegel H (2010) Critical thinking. In: Peterson P, Baker E and McGaw B (eds.) *International encyclopedia of education*, Vol 6: 141–145. Oxford: Elsevier.~~

Siegel H (2015) New work on critical thinking: Comments on Frímannsson, Holma and Ritola. *Studier i Pædagogisk Filosofi*, 4(1): 55–62.

Siegel H (2017) *Education's epistemology: Rationality, diversity, and critical thinking*. Oxford: Oxford University Press.

Simon H (1955) A Behavioral model of rational choice. *Quarterly Journal of Economics*, 69 (1): 99–118.

Stich S (1990) *The Fragmentation of reason: Preface to a pragmatic theory of cognitive evaluation*. Cambridge, MA: MIT Press.

Stanovich KE (2011) *Rationality and the reflective mind*. Oxford: Oxford University Press.

Stanovich KE and West RF (2000) Individual differences in reasoning: Implications for the rationality debate. *Behavioral and Brain Sciences*, 23(5): 645–665.

Stanovich KE and Stanovich PJ (2010) A framework for critical thinking, rational thinking, and intelligence. In: Preiss DD and Sternberg RJ (eds.) *Innovations in educational psychology: Perspectives on learning, teaching, and human development*, pp. 195–237. New York: Springer Publishing Company.

Stanovich KE, West RF and Toplak ME (2016) *The rationality quotient: Toward a test of rational thinking*. Cambridge MA: MIT Press.

Tetlock PE and Mellers BM (2002) The great rationality debate. *Psyuchological Science*, 13(1): 94–99.

Thagard, P (2011) Critical thinking and informal logic: Neuropsychologic perspectives. *Informal Logic*, 31(3): 152–170.

Thaler RH and Sunstein CR (2008) *Nudge: Improving decisions about health, wealth, and happiness*. New Haven: Yale University Press.

Thayer-Bacon BJ (1992) Is modern critical thinking theory sexist? *Inquiry: Critical Thinking Across the Disciplines*, 10(1): 3–7.

Tversky A and Kahneman D (1974) Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157): 1124–1131.

UNESCO (1995) *The declaration of principles on tolerance*. UNESCO Publishing. Available at: http://unesdoc.unesco.org/images/0015/001518/151830eo.pdf (accessed 11 June 2020).

Warburton, N (2007) *Thinking from A to Z (3rd ed.)*. Devon: Routledge.

Wason PC (1960) On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology*, 12(3): 129–140.