UNIVERSIDADE FEDERAL DE MINAS GERAIS
Faculdade de Filosofia e Ciências Humanas - FAFICH
Programa de Pós-Graduação em Filosofia

VERONICA DE SOUZA CAMPOS

AKRASIA EPISTÊMICA E INEFICÁCIA EPISTÊMICA:
UMA ABORDAGEM BASEADA EM VIRTUDES

Belo Horizonte
2022

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Faculdade de Filosofia e Ciências Humanas - FAFICH
Programa de Pós-Graduação em Filosofia

VERONICA DE SOUZA CAMPOS

EPISTEMIC AKRASIA AND EPISTEMIC INEFFICACY:

A VIRTUE-BASED APPROACH

A Dissertation presented to the Philosophy Department of the Faculty of Philosophy and Human Sciences, at the Federal University of Minas Gerais, in partial fulfilment of the requirements for the degree of Doctor of Philosophy.
Supervisor: Professor Ph.D. André Joffily Abath.

Belo Horizonte
2022

UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE FILOSOFIA E CIÊNCIAS HUMANAS
PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA

FOLHA DE APROVAÇÃO

Akrasia Epistêmica e Ineficácia Epistêmica: uma abordagem baseada em virtudes

VERONICA DE SOUZA CAMPOS

Tese submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em FILOSOFIA, como requisito para obtenção do grau de Doutor em FILOSOFIA, área de concentração FILOSOFIA, linha de pesquisa Lógica, Ciência, Mente e Linguagem.

Aprovada em 08 de setembro de 2022, pela banca constituída pelos membros:

Prof. André Joffily Abath - Orientador (UFMG)

Prof. Ernesto Perini Fizzera da Mota Santos (UFMG)

Prof. Alexandre Meyer Luz (UFSC)

Prof. Arthur Viana Lopes (UFPB)

Prof. Breno Ricardo Guimarães Santos (UFMT)

Belo Horizonte, 08 de setembro de 2022.

ACKNOWLEDGMENTS

I've been through some difficult times during the course of this PhD, god only knows. Alas, god and my hairdresser, for she's been the one helping me manage the grey.

Nevertheless, it wouldn't be wrong for me to say that these have been my best years so far, because I feel that I've learnt so much and grown so much, and that's the true beauty of being alive.

I would like to say thank you to a whole lot of people without whom this PhD wouldn't have been possible. I am the luckiest person of all for, in times of need, I never found myself alone. For starters, to my parents, that have been putting up with me being always busy with things they don't even understand for so long at this point, and who have done everything within their reach to support me, in every way. To all my extended family members, on mom and dad's side, for helping me every time I needed, especially my uncle Paulo. To all the carers I needed to hire, in the darkest times. To all the healthcare professionals I relied on. All of those people have helped me not give up.

To my supervisor, Andre, who has been supporting me for more than ten years at this point, since I was an undergrad, and who's always pushing me into growth. To my other supervisor, Quassim, who gifted me with his kindness, encouragement and so many insightful conversations. I'm most indebted to them for all the lessons I've learnt in both discussions and print. I couldn't have asked for better supervisors. I'm also obliged to professors Leonardo Ribeiro, Ernesto Perini and Alexandre Meyer, for their valuable advice; to all the members of staff at UFMG and at Warwick university, and to CAPES and the CAPES PrInt programme, that made this PhD study materially possible.

Last, to all my friends and colleagues at UFMG: Ludmila, Carol, Pedro, Eduardo, Maylson, Júlio; to all the undergrad students that came to my group of studies on argumentative writing; and specially to my friends Debarry, Carlos and Rochelle, with whom I shared so much of the joy of being at university. Thank you, all of you.

ABSTRACT

In this dissertation I make the case for a "new" epistemic vice, the vice of epistemic inefficacy. While baptizing and sketching profiles to vices is a bold philosophical enterprise, the task is made unpresuming by the fact that the vice I attempted to shed light on is actually correlated with a longstanding and much known problem in the history of philosophy: the problem of weakness of willpower, or akrasia. I make the case for this interrelatedness by showing that, while they're not, strictly speaking, the *same* problem, or mirror problems, weakness of willpower and epistemic inefficacy have a lot in common. In fact, epistemic inefficacy, I submit, is as close as you can get to weakness of willpower in intellectual, or epistemic, contexts, since in those contexts the notion of willpower (ability to control thoughts and actions) is not preponderant. The preponderant notion when it comes to epistemic activity is efficacy (ability to yield the intended results). To accomplish this, I show that there is a good deal of problems with the thing that is normally thought to be, or meant to be, weakness of willpower's mirror-concept, the so-called "epistemic akrasia"; and I devise a comprehensive presentation of vice epistemology and inquiry epistemology, and of why epistemic inefficacy squares of as an epistemic vice following the tenets of those disciplines.

Key words: epistemic akrasia, weakness of the will, epistemic inefficacy, vice epistemology, epistemology of inquiry.

RESUMO

Nesta tese defendo um "novo" vício epistêmico, o vício da ineficácia epistêmica. Embora batizar e traçar perfis para vícios seja um empreendimento filosófico ousado, a tarefa é tornada despretensiosa pelo fato de que o vício sobre o qual tentei esclarecer está, na verdade, correlacionado com um problema antigo e muito conhecido na história da filosofia: o problema da fraqueza de força de vontade, ou akrasia. Defendo essa inter-relação mostrando que, embora não sejam, estritamente falando, o mesmo problema, ou problemas-espelho, a fraqueza da força de vontade e a ineficácia epistêmica têm muito em comum. Na verdade, a ineficácia epistêmica, eu afirmo, é o mais próximo que você pode chegar da fraqueza da força de vontade em contextos intelectuais ou epistêmicos, uma vez que nesses contextos a noção de força de vontade (capacidade de controlar pensamentos e ações) não é preponderante. A noção preponderante quando se trata de atividade epistêmica é a eficácia (capacidade de produzir os resultados pretendidos). Para conseguir isso, mostro que há muitos problemas com a coisa que normalmente se pensa ser, ou deveria ser, o conceito-espelho da fraqueza de vontade, a chamada "acrasia epistêmica"; e eu concebo uma apresentação abrangente da epistemologia do vício e da epistemologia da investigação, bem como de por que a ineficácia epistêmica se enquadra como um vício epistêmico seguindo os princípios dessas duas disciplinas.

Palavras-chave: akrasia epistêmica, fraqueza da vontade, ineficácia epistêmica, epistemologia de vícios, epistemologia da investigação.

TABLE OF CONTENTS

INTRODUCTION

I went to university for the first time at the age of seventeen. Back then, at the School of Architecture, my feelings floated between a lukewarm disposition towards the career and a more or less stable conviction that there was where my natural talents landed, so that a brilliant and much rewarding future was just around the corner.

Soon after reaching half of the undergraduate course, I started to consider quitting. Upon a good deal of reflection, however, I concluded that since I was already halfway from the finish line, and since things were not particularly terrible at school, the most reasonable thing to do was to keep going until graduation. This way the investments my family and I had already put into that enterprise wouldn't have gone to waste. Also, with a degree, I would at least be able to get a decent job to earn a living, and from there I could figure out other things to do in life, including pursue a second degree. I was still in my teens. I had plenty of time. In sum, it was clear to me that the pros to keep going very much outweighed the cons.

It turned out, however, that shortly after I reached this resolution I didn't show up for class. For no reason, just because. This particular day wasn't meant to be a game changer in my life, or the beginning of a new era, by any means; except that I missed class the next day too. And then I never stepped foot in the School of Architecture again. In spite of feeling upset about it and having second thoughts, and guilt, after a while I simply reached a point from where there was no turning back. Upon hearing this story, one of my dearest colleagues (rightfully, I suppose) charged me with *akrasia*. "It's a shame you lack the willpower to carry on", he said, in a patronizing tone.

Though this anecdote of mine might sound unexceptional, the idea of akrasia – voluntarily doing something while thinking it would be best to do something different instead – is traditionally puzzling to philosophers[1]. Traditionally because it is one of the most ancient and longstanding philosophical perplexity-triggers, inasmuch as a

---

[1] I say "puzzling to philosophers", as opposed to just "puzzling", because non-philosophers do not seem to find incidents such as the anecdote I recounted particularly difficult to make sense of. They seem prone to quickly assuming that I didn't really want to get that degree, that's why my resolution was so short-lived; and apparently they feel alright with generalizing this explanation. Richard Holton (1999) acknowledges a similar point. He says that when ordinary people talk of cases such as the abovementioned, they are talking about abandoning a previously formed intention.

good account of what is going on is utmost elusive and we seem to not be able to fully understand it. Did I drop my resolution to get my degree, without actually realizing it? Or did I, having kept that resolution, fail to comply with it? Is it technically right to say that I could have done otherwise? In cases like this, "something has evidently gone quite wrong, but it is not immediately apparent what" (Hoffman 2008: ix).

Some contemporary philosophers took interest in the phenomenon of akrasia because they observed that it poses a challenge for their theories of action. Those are theories that in some way or another entail a form of motivational internalism, the idea that judgments about action correspond directly with motivation, or are sufficient to motivate. Akrasia challenges such theories because it presents the possibility of a person evaluating a certain course of action as being best overall, and yet not being motivated to take it, more or less in the same fashion as teenage me at the School of Architecture.

This is the case with Donald Davidson, for instance, insofar as the idea of one's acting against her better judgment challenges the theory that reasons both rationalize and cause actions, developed by him in his article "Actions, Reasons and Causes" (1963). And it is also the case with Richard Hare, whose theory of moral prescriptivism devised in *The Language of Morals* (1952) and *Freedom and Reason* (1963) is also defied by the idea that one can act in certain way while judging it is best to act differently. Another contemporary philosopher who devoted himself to discussing akrasia is Alfred Mele. Mele, however, took interest in the topic not because it challenged his theory of action, but rather because he wanted to develop a theory of action that took the possibility of akratic action seriously from the very beginning. This was what he set out to do in *Irrationality* (1987).

Though these scholars endorse very different theories of action, and actually very different world views, what is common amongst them is that they approach akrasia from an action theory standpoint: they take as their object of analysis akrasia in the form of a single and detached episode of acting against one's better judgment, and are set out to explain either *how* this episode is possible or *why* it is not.

In so doing, they work from what is possibly the most acutely paradoxical characterization available, one that stems from a set of inconsistent principles. The classic so-called "akrasia puzzle" (Davidson 1969: 95) comprises the following three principles:

P1. If an agent wants to do X more than he wants to do Y and he believes himself free to do either X or Y, then he will intentionally do X if he does either X or Y intentionally.

P2. If an agent judges that it would be better to do X than to do Y, then he wants to do X more than he wants to do Y.

P3. There are akratic actions.

There are roughly three basic ways to solve this puzzle: drop **P3**, denying the actual occurrence of akrasia; drop **P2**, denying that the connection between evaluation and motivation is as tight as it seems; and spell out the three principles again in alternative forms, showing that they can be consistent, after all. Hare took the first path, dropping **P3**; Mele took the second one, dropping **P2**, and Davidson took the third one, reinterpreting the principles in a way that they no longer contradict one another.

Due perhaps to lack of consensus (or perhaps to Hare and other akrasia skeptics not having been the most persuasive), the idea of there being akratic action as an article not yet fully understood remained. In the final years of the 20th Century, this idea made its official debut into the domain of epistemology, as some scholars raised a concern about there being forms of akratic believing, as well as of akratic acting. The pioneering article was Amélie Rorty's "Akratic Believers" (1983). Rorty and others ventured the suggestion that akrasia is a phenomenon that can also befall one's theoretical rationality and, therefore, challenge certain theories in the field. This form of akrasia – officially: "epistemic akrasia", or "akrasia of belief" – is commonly defined as one's holding that **p** against his own understanding that **p'** is what he should hold instead (were **p** and **p'** are inconsistent propositions, or rival hypothesis).

The suggestion brought together with it an epistemic version of the puzzle laid out above, featured in terms that are as much acutely paradoxical as the original one. In fact, this epistemic version of the puzzle basically mirrors the practical version, but with "action" being replaced by "belief". It finds a quite straightforward expression in John Heil (1984: 63-67):

R1. Where P and P' are epistemically incompatible for S, if S holds P to be more warranted than P', S will hold P if he holds either P or P'.

R2. Where S holds R and R' and takes these to be all that is relevant to the warrant of P and P' respectively, then if S holds R epistemically to outweigh R', he will hold P to be more warranted than P'.

R3. Akratic belief is possible.

In the face of this epistemic version of the puzzle, again, solutions envisaged by philosophers varied, verging from dropping **R2** (as, for instance, Mele 1986: 219), to taking the principles as being in need to be rewritten, because "the relation between a belief and its warrant is not so simple" (Heil 1984: 69; and also Hookway 2001); to a myriad of theories that in one way or another support dropping **R3** (explicitly, Owens 2002; and implicitly Hurley 1989; Pettit & Smith 1996; and Adler 2002a, 2002b).

Like with the puzzle about practical akrasia, here too there is a common element underlying the many approaches: though scholars endorse different theories of how theoretical rationality works, they take as their object of analysis epistemic akrasia in the form of a single and detached episode of acquiring a belief against one's better judgment, and are set out to explain either how this single and isolated event is possible or why it is not.

Now I am in position to clearly state the spirit in which this dissertation comes into play: it is deliberately intended to *not* attempt solving the epistemic version of the puzzle. Rather, it will give its contribution to the philosophical debate by introducing a character-based approach to the issue of akrasia in the intellectual domain. I propose a whole different way of looking into the phenomenon, one that does not take as its object of analysis single and isolated episodes of holding a belief against one's better epistemic judgment. That's ultimately because I don't think that this is what weakness of the will is. And even if such episodes do exist, they are not where the really interesting problems that akrasia raises are to be found. Therefore, my approach is one that circumvents the puzzle, in a sense, rather than one that tries to dismantle it.

I believe that venturing explanations of how single and isolated akratic episodes are possible, or of why they are not, though an interesting philosophical endeavour in its own right, is one that completely bypasses three of what I take as being the most significant issues pertaining the very idea of weakness of willpower, or akrasia.

First, concrete experience shows us that people can, and typically do, go against what they judge best, in practical as well as in intellectual contexts. This should be taken prima facie, that is, we should accept that people do it, and work from there. Aristotle was perhaps the first to make this explicit, in his response to Socrates' idealizing denial of practical akrasia: "this view [of Socrates' that akrasia is impossible] plainly contradicts the observed facts" (Aristotle, NE, VII 2 1145b27–28). Descriptive theories must be modelled after experience and be capable of shedding light on it, not the other way around. When it comes to akrasia, though, most of the post World-War II theories do one of two things. They either try to sort of artificially shape reality in order for it to fit their theoretical claims; or the try to build up on top claims that, in spite of making perfect sense relatively to one another, do not resonate very much with our concrete experience and, therefore, have a very limited potential to expand our understanding of ourselves and our lives.

This, the understanding of ourselves and our lives, is sometimes referred to in short by means of a flamboyant German word, the word *Verstehen*, as in, for instance, Cassam (2021). I believe Verstehen is the ultimate end of philosophy. Therefore, I believe an approach to the topic of intellectual forms of akrasia is only as interesting as its potential to serve as a tool for Verstehen.

Secondly, the biggest problem akrasia poses in everyday life (practical and intellectual) stems from it being a diachronic phenomenon, that is, something that extends itself over time, tainting one's whole way of dealing with some important matter (practical or intellectual). In general, it is not much of a problem if a person who deems it best to stick to a healthy diet indulges herself at a time, or from time to time. The real problem arises when it becomes a *lifestyle*. Likewise, most of the time it is not too bad if a person postpones drawing a conclusion, for instance, that she knows to be what evidence best supports. The real problem comes with iteration. Modulating akrasia episodically, however, misses this point entirely. It focuses solely in the synchronic aspect of akrasia (doing A while *at the same time* judging it would be best to do B), and leaves its diachronic aspect aside[2].

This diachronic aspect is, after all, an ethical component of the phenomenon. "Practical reasoning", Hoffman says, and I would extend his sayings to cover

---

[2] This aspect is highlighted and extensively discussed by Snellen (2018) in her much inspiring PhD dissertation concerning practical akrasia as a character trait.

theoretical reasoning as well, "is not a disembodied calculation. It does not have a sharply definable beginning and a clear-cut end, but rather extends over time" (Hoffman 2008: x). By virtue of its power to scatter over time, akrasia compromises one's life, practical as well as intellectual. That is, it turns one's life into something worse than it could otherwise be. Moreover, it turns *us* into worse people than we could otherwise be, not only as practical agents (citizens, friends, parents, and so forth) but also as epistemic agents (scientists, judges, police detectives, medical doctors, politicians, and all sorts of other agents that happen to be in charge of some epistemic duty within society). *This* is the biggest problem akrasia creates, not the conceptual riddle.

This ethical feature of akrasia is salient in the many accounts the notion has been given throughout history, but it has been somehow forgotten in modern times. Ancient and medieval philosophers viewed akrasia primarily as an ethical problem – they saw it as a character trait that gets in the way of an ethical ideal, such as living a virtuous life, or a life exempt from sin (Snellen 2018: 14-15). Renaissance, in turn, saw it as a sort of curse, something associated with disgrace and tragedy in one's life (Vasiliauskas 2016). Notwithstanding, the general approach to practical akrasia taken by the contemporary philosophers I described above (which is concerned either with the possibility or the impossibility of single and isolated events) disregards this aspect. It de-characterizes akrasia, in a sense, by wiping its ethical component off[3].

Third, a remarkable feature of akrasia that is almost completely overlooked by contemporary action-based literature is akrasia being an attitude or pattern of attitudes an agent takes when she faces dilemmas, of both small and big proportions. If we turn to the literature and look for the classical examples of akrasia that are the least stipulative and episodic (and, therefore, the ones that are more concrete, in essence), we see that what is going on is this: the agent finds herself cornered within a situation of which it is not easy for her to get out, even though she wants to. Like Shakespeare warns us in *The Merchant of Venice*: "If to do were as easy as to know what were good to do, chapels had been churches and poor men's cottages princes' palaces" (2.1.11–12.). The akratic person has a general understanding that tells her what the best course of action is, but her other states of mind are strong too, so the

---

[3] It's always nice to remember that there are exceptions. Rorty, for instance, acknowledges that "Akrasia is typically not episodic. Of course it can in principle occur as a single momentary event, a kind of motivational or epistemic sneeze, a single absent-minded lightfingered questionable bond sale or an isolated flare of rage. But it rarely does" (Rorty 1997: 649).

situation is not one that unfolds naturally, or in a way that is easygoing. It does not simply "result" in action that happens to be at odds with that understanding. What happens, rather, is that the agent undergoes conflict.

For instance, take Alcibiades, in Plato's *Symposium*. Alcibiades is probably the first character in western literature that is meant to represent a real life case of a person that suffers from weakness of willpower. A flamboyant man with great love of power, he is made to feel ashamed of himself and of the life he had been living, upon hearing Socrates speech on the value of living a virtuous life (216a-b). He then "vacillates between the 'good' that he had previously taken to be an accurate measure of his life, and the good of virtue that Socrates helps him to see" (Shanahan 2019: 137). He states: "I know perfectly well that I can't prove him wrong when he tells me what I should do; yet, the moment that I leave his side, I go back to my old ways: I cave in to my desire to please the crowd" (216b). According to Shanahan, this passage is quite significant, because it demonstrates that "Alcibiades has the ability to re-evaluate the quality of his life (…) but, there is still an internal conflict between this new good that Socrates has shown him and the old 'good' " (Shanahan 2019: 137). The conflict is so intense and it upsets Alcibiades so much that he declares, in the passages following that speech, that he couldn't help the feeling that he would rather die.

Or see, for instance, the conflict recounted by Saint Paul's in his Letter to the Romans: "What I do, I do not understand. For I do not do what I want, but I do what I hate" (Romans 7). This example is discussed by Saint Augustine as a classic case of akrasia related to sexual drive (Stowers 1994: 279). What Saint Paul is dealing with is a dilemma, in the sense that the choice that is required from him presents itself to him as being a difficult one to make. It is not easy for him to choose correctly, in accordance with what he believes to be correct. Even if in the end he chooses correctly, refraining from giving in to sexual temptation, internal conflict around the need to make this choice is very much vivid.

Or take the akratic protagonist in Dante's autobiographical *Vita Nuova*. Dante describes the so-called "accensio amoris", the onset of love, upon his encounter with Beatrice. Love comes through as an involuntary, sudden experience that leaves the protagonist perplexed, pulling him into two different directions: on the one hand, he is set out to grasp the "mystery of love", to understand whether the advent of Beatrice is a fortuitous accident or the revelation of a providential path for his life. He needs to

figure this out, in order to decide how to handle the situation. On the other hand, love engenders within himself an overwhelming flood of fantasies, which gets in the way of his intellectual quest. Every time he thinks of Beatrice or sees her, he is diverted by those fantasies and ends up sweetly indulging in them. He is determined to figure love out, but finds that he is not strong-willed enough to carry this inquiry on. Unsurprisingly, he suffers. The poetry of *Vita Nuova* arises precisely from this, the "double-edged, contradictory nature of his experience" (Mazzotta 2008: 140).

Or take Macbeth's killing King Duncan in Shakespeare's *Macbeth*. In spite of him having decided that it was best not to, he keeps being haunted by thoughts of regicide. Finally, when Lady Macbeth actually voices a proposal for regicide, Macbeth replies "We will speak further" (1.5.88). He didn't reply "No way", but he did not say "Let's do it" either. As Shugar notices, this evasive swerve is the indication that Macbeth's reasons for opting against killing King Duncan are "a source of discomfort, or embarrassment, to him" (Shugar 2006: 54). He has formed a resolution not to kill, but is not completely at peace with this resolution. He is hesitating.

The individuals in all those examples face, and not without struggle, a clash between opposing inclinations. That is the reason why, as Aristotle noticed, akrasia is commonly followed by self-awareness of one's wrongdoing as well as some form of regret (EN VII 7 1150a 20). I don't take it as a necessary feature of akrasia that the agent experiments dramatic suffering, nor that she experiments a *feeling* of regret, but I do consider it significant that there is a reflexive nature to the experience, without which neither suffering or regret would come by. A person who goes against her own better judgment without even noticing it exemplifies other forms of irrationality, not akrasia[4]. The akratic person reflects on the difficult choice she is facing. She experiments dilemma.

But it is actually a little more complicated than that. The akratic agent experiences a dilemma *that she shouldn't be experiencing.* Think like this: if you have a general guideline for living, or a better judgment regarding how to handle some

---

[4] Aristotle calls this other form of irrationality "vice" (EN VII 4 1148a-15; EN VII 7 1150a-20). It is very likely that he is referring to the vices of intemperance (lack of moderation), though it is possible that what he says there apply to other vices as well. His famous differentiation of the two is that the vicious agent does what she deems right, but errs, because her conception of what the right is flawed. The akratic agent, in turn, does what she deems wrong and therefore errs (her conception of what the right is correct). The vicious agent faces no internal struggle, because she believes she is doing the right thing.

particular matter, according to which the right thing for you to do is A, and yet you struggle when you find yourself in a situation in which you can do either A or B, then this is a dilemma you shouldn't be facing. If you have *really* committed yourself to a general guideline, it's admissible that you hesitate for one second, maybe two, but that's all. That's the most lasting your dilemma should be, were you to behave as an impeccably reasonable person who always abides by your principles. An impeccably reasonable person is one who already got over the point in life where option B offers real temptation.

Nevertheless, many times people have not got over it. They're still living in that place where they are tempted by option B, in spite of seeing themselves as committed to a principle according to which the right choice is A. The fact that one has not got passed this point is a red flag in terms of the strength of her will, even if in the end she ends up choosing in accordance with her general principle. The crucial intuition I'm trying to extract here is the following: the most emblematic form of akrasia has to do with being caught up within dilemmas *for longer than you should*.

I'm not trying, obviously, to stipulate a maximum duration for one's struggle in the face of a dilemma in seconds (who am I to?). What I'm saying, from a very pragmatic standpoint, is that there must be a *reasonable* limit for dwelling on a particular dilemma. There must be a point in your dealing with a dilemma such that, once you go past this point without having achieved success (resolution), you begin to be considered unreasonable. So unreasonability of the relevant kind is marked by exceeding this point. If you have sincerely committed yourself to the principles of a Christian monastic life but then you keep finding it difficult to abide by monastic rule, being tormented by temptations beyond that reasonability point (wherever and however the bar turns out to have been set), what happens is that after a while we begin to suspect there is something wrong: either your adhesion was not really collected and sincere, or you lack the will power to actualize it in action. If your adhesion was indeed collected and sincere, then you lack the power to actualize it in action. You are akratic.

Summing this third point up, the idea of akrasia is tied with dilemma of some sort, and with timing. Notwithstanding, most of the contemporary scholars of which I spoke before are silent of this matter. To them, it makes no difference whether the occurrence of akrasia unfolds with or without internal struggle, which means with or without dilemma; and consequently it is also of no particular significance to them for

how long the akratic agent dwells on her struggle. What matters to them is that the agent has fallen, somehow, in a situation of incontinence, of some sort. It is basically this event that matters.

So those are the three aspects that have been significantly overlooked by recent philosophical literature on akrasia, and which I intend my account to be able to contemplate: concreteness, diachronicity and longer-than-due dilemma facing. I focus exclusively on epistemic occurrences. So what I'm actually going to account for are concrete, real-life inspired, cases of epistemic agents (that is, agents that are engaged in some distinctively cognitive activity, such as knowledge seeking) undertaking a particular form of intellectual misconduct that i) involves going against their better judgment in some way; that ii) extends itself over time in some way, and that iii) encompasses one's dealing with the dilemma, or quandary, thus created for longer that it was due. This is very much a character-based approach, inasmuch as I shift the focus from beliefs, or doxastic states, to features of agents' epistemic performances and their life experiences.

I like to think that when I set myself out to introduce this character-based approach to the phenomenon of intellectual weakness of willpower I'm bringing something original to the table, in the sense that a particularly epistemic variety of the phenomenon is a recent topic in Anglo-American analytic-inspired philosophy. But in terms of my focus on the three aspects devised above, I'm definitely not inventing the wheel. I'm rather rehabilitating certain original features of the phenomenon of weakness of willpower that appear to have been forgotten, or overlooked, and exploring the ways they show themselves within intellectual contexts.

Nevertheless, there is a fourth aspect that my approach contemplates that is in fact innovative. That's because I work from a virtue-based epistemological background and adopt a virtue-based framework to deal with weakness of willpower; and virtue-based epistemology is a relatively new field. I underwrite contemporary responsibilist-consequentialist vice epistemology, of the sort developed by Cassam (2016, 2019), Kidd (2017) and Crerar (2017), which is sometimes referred to as *obstructivism*. The specific theory I submit is that intellectual weakness of willpower characterized in the way alluded (with the due preponderance of the aspects I've highlighted) squares off as an epistemic vice, in the straightforward obstructivist sense. I call this vice *epistemic inefficacy*, to preserve the distinction with the idea of

"epistemic akrasia" that has been the focus of most of the contemporary debate. This, I believe, is altogether original.

By now, a number of questions might have arisen. First, where does my discussion of intellectual weakness of willpower stand, relatively to the debate about epistemic akrasia that has been taken place for, roughly, the last thirty years? Because if there is no coupling, if the two discussions are completely insulated from one another, then how do we know that the question of epistemic akrasia *vs.* epistemic inefficacy is not, actually, a verbal dispute? A difference of opinion between two parties is said to be verbal when the two parties agree on the relevant facts about a domain of concern and just disagree about the language used to describe that domain (cf. Chalmers 2011: 515). So how do we know "epistemic akrasia" and "epistemic inefficacy" aren't actually two names for the same thing?

The answer to this question is: because if we withdraw labels, that is, if we eliminate the use of both expressions, to use Chalmers' terminology, we find that what is said to be examples of epistemic akrasia in the traditional debate (the debate that has been going on under the epigraph "epistemic akrasia" in recent years) and what I am describing are different things, although they are still related. They are different things because they possess different properties. The former is a matter of episodes in which a combination of conflicting attitudes appears to co-exist at a given instant in a person's mind, and it is a problem because it is contended that such co-existence is impossible. The latter is a personal flaw, that extends itself over time, in which a person undertakes conflicting attitudes for an extended period of time; and it is a problem because it generates bad epistemic consequences (it leads to the deterioration of epistemic goods, such as pieces of knowledge, for instance). But these two things are related to the extent that both seek to offer answers to the same question, the question "does it make sense to speak of epistemic weakness of willpower?"

I believe that the way in which the phenomena that are said to be examples of epistemic akrasia have been characterized in the traditional debate has a number of problems. Apart from the neglect of important aspects pertaining the very idea of weakness of willpower from which it derives its grounding assumptions, that I've already mentioned, the traditional debate also suffers from problems of vagueness and imprecision, conceptual misunderstandings, and a low "margin of profit", so to speak, when it comes to epistemology. Those issues are going to be clarified in a

great deal of detail throughout Chapters 1, 2 and 3. It will then become clear that my approach comes into play as a means to overcome those deficiencies. In particular, I believe that the way to overcome them is to try and get specific. Pick one variety of the phenomena being envisaged (which is clearly varied), and seek to get into the detail of what is going on. It then becomes clear that what is going on is not the same thing that the traditional debate portrays as being what is going on in the cases they are discussing (although the two things are related).

Second, what is obstructivism, exactly, and where does this branch of vice epistemology stand relatively to other accounts of epistemic vice within contemporary virtue-based epistemological theories? And why take a vice epistemology standpoint, in the first place? What are the main advantages it has over alternative accounts, and what are the main challenges it faces? This is the general theme of Chapter 4. There, I bring together three main topics that are the building blocks of my view and I put them into perspective relatively to one another: intellectual weakness of willpower, vice epistemology and epistemology of inquiry. This chapter lays the foundations that will allow me to address the pivotal question of what makes weakness of willpower an article of interest to epistemologists. Or, if we put it in sharper terms, the question of why weakness of willpower poses a problem that is distinctively epistemological, in addition to the ethical problem that has been an object of debate for so many centuries.

Then, in Chapter 5, I begin to develop my personal answer to these questions. Intellectual weakness of willpower poses an epistemological problem because unreasonability is an epistemic vice in its own right. It's the vice of epistemic inefficacy. In defending this view I go against the classic, Aristotelic-inspired view that being enkratic, i.e., having (strength of) willpower is not a virtue, and that lacking it is not a vice[5], by submitting that epistemic inefficacy meets the requirements for being a vice in the obstructivist sense presented in Chapter 4. I sketch a preliminary profile for the vice of epistemic inefficacy, highlighting its most remarkable traits, how they usually show up and how they're different from other epistemic vices and misconducts.

This vice is a defect of a person's thinking, therefore, it is a type of personal trait that turns one into a worse inquirer than she could otherwise be, by causing her

---

[5] Aristotle devises this view in the *Nichomaquean Ethics* (EN VII 1 1145a 20).

to perform worse than she could otherwise perform in some specific cognitive tasks. I propose that the consequences this demeanour brings to inquiry are the most important issue concerning epistemic inefficacy (more important than internal issues, such as the issue of motivation): this vice obstructs the acquisition, the retention and the enjoyment of epistemic goods (such as knowledge and understanding, but also others) in a variety of manners. I then delve into some of these manners, in Chapter 6, thus completing the profiling of the vice.

Now, another question that might certainly be popping up is: *what* vice is this, specifically? Vices are specific. Arrogance is the vice of underestimating one's own limitations and overestimating one's own talents and abilities. Gullibility is the vice of believing unlikely propositions that are not supported by evidence. Epistemic insouciance is the vice of being unconcerned about whether something is true or false, etc. Epistemic inefficacy is the vice of what, then? Because it's not clear that we possess "epistemic willpower", insofar as forming beliefs and drawing conclusions is not usually a matter of our will to do so.

In response to this I submit the following: indeed, in the biggest part of our epistemic lives, we're not really exercising the faculty of willpower (even though we might have it). Within epistemic endeavours, the key concept is efficacy, rather than the will. In the enterprise of solving a dilemma, exercising efficacy means being able to do away with the inconsistent terms, on due time, which can be done either by exonerating one of them, or the other, or by finding that they are not really inconsistent, after all. Epistemic inefficacy, on the other hand, is the failure to do this.

That is, epistemic inefficacy is the vice of tolerating inconsistency in inquiry for longer than one should. It can take many shapes, but, roughly, what the epistemically inefficacious agent does that he shouldn't be doing is: he fails at settling the matter of the inquiry he is engaged with, when settling is required. Rather than dismissing one of two inconsistent hypotheses at the right moment, he keeps putting up with the two, as if the matter was still open.

How is this idea of "putting up with inconsistency" to be understood, precisely? And what does this have to do with weakness of willpower? Chapters 5 and 6 also comprise an extensive clarification of this. Roughly, when an agent inquires into a matter, or engages with a question that needs to be answered, he envisages more than one hypothesis, where only one of them can be true, or the better one. Thus, he finds himself before a sort of dilemma, an epistemic dilemma, having to figure out

which hypothesis is the true one, or the better one. He begins at a point in which he doesn't yet have what it takes to settle the matter and moves, inquisitively, towards an (ideal) point in the future in which he already has all the information that is jointly sufficient for thus settling. The vice of epistemic inefficacy shows up when the agent continues inquiring *past* that ideal point. In so doing, the agent dwells on the quandary: he treats it as an open case, when in reality it is a case that he already has the means to close, and should have closed. The dilemma thus subsists for longer than it should. This action of iteratively postponing closure is the epistemic equivalent of the akratic person's act of doing again and again what she acknowledges as being inadvisable. It's not the exact same thing, or a perfectly square mirror attitude, but it's as close as you can get to akrasia, given that, as said before, the key notion when it comes to intellectual contexts is not the will.

One final observation before we move forward. One might be wondering whether my proposal amounts to a form of revisionism about the concept of akrasia, that is, whether or not I am suggesting a modification in the very scope of the concept of akrasia, or recommending that we begin using the word "akrasia" to refer to something different than what it has been used to refer to, or even that we replace the word "akrasia" for "efficacy" in epistemological discussions. To this I will say that my account is not revisionist, unless in the Carnapean (deflationary) sense in which abandoning a vague and imprecise concept and focusing on a more specific and precise one is a form of revision.

I am not suggesting that we change the traditional meaning of the expression "epistemic akrasia", nor that we perform a mere replacement of names. What I am suggesting is that the discussion about whether or not there is an epistemic equivalent of weakness of willpower in our epistemic lives is rendered more profitable if we change to a virtue-based approach, because this approach allows for more profitable results in terms of the goal of sense-making (Verstehen: making sense of the seemingly alien conducts, the candidates to being the relevant epistemic equivalents to weakness of willpower). Taking this shift involves addressing *specific* virtues and vices. So my proposal is a call for attention to the fact that there is a specific vice of the intellect that is a close relative of weakness of willpower, within our epistemic lives; and we make sense of this conduct if we understand this vice.

I believe that the choice between different perspectives, or conceptual frameworks, depends more on our purposes than on matters of fact. So what I am

doing is not changing the subject plain and simple, but rather examining it from a different perspective. This perspective demands that we get more specific, and getting more specific, in practice, ends up not being very different from picking a more specific sub-subject within the original subject. But that is as far as it goes, in terms of revisionism. My motivation for doing so, as stated before, is that the traditional perspective epistemologists have been taking while examining the subject hasn't taken us very far in terms of the specific purpose of sense-making (which I intend to show); and my main insight is that adopting a different strategy is a way of fixing this.

1. AKRASIA OF THE INTELLECT: A PANORAMIC VIEW

> *Xenophanes says: the Ethiops say that the gods are flat-nosed and black, while the Thracians say that they have blue eyes and red hair.*
> Clement of Alexandria, **Miscellanies V**, xiv.109.1-3.

There isn't one single thing being referred to by scholars when they use the expression "intellectual akrasia"[6]. In this introductory chapter I am going to present a number of examples found in the literature tagged with this term. What I aim to accomplish here is to get clear about the extension of this concept, that is, the set of things it applies to. I'll sketch preliminary answers to the following questions. Is there a common core shared by all of the examples? If there is, what does it consist of? And how should it be spelled out? Can it be spelled out in terms of a set of necessary and sufficient conditions? If there is, then what are those conditions? Also, what's the relationship between epistemic akrasia and the phenomenon of practical akrasia? Are the alleged cases of epistemic akrasia supposed to mirror cases of practical akrasia in some meaningful way? If they are, then what meaningful way is this?

1.1. Preliminary Remarks

There is no shortage of alleged examples of intellectual akrasia within the writings of epistemologists from, roughly, the last three decades. If we have a close look among scholars who believe akrasia to be an article of epistemological interest, we are left to wonder whether what they are discussing is a range of phenomena, or a heterogeneous and plastic phenomenon that can take many different shapes. It is as though the concept of akrasia is being used to capture a number of considerably different things.

---

[6] Throughout this dissertation I'll follow Turri, Alfano & Greco (2021) in alternating freely between the locutions "intellectual", "epistemic", "doxastic", and "theoretical". Though I acknowledge that these expressions can be given completely distinct senses, choosing among them for the purpose of the discussion carried out here is more of a matter of stipulation than of technical precision.

In the next sections I'll present and discuss some examples that appear in Rorty (1983), Mele (1986), Heil (1984), Hookway (2001), Ribeiro (2011) and Christensen (2014), which scholars refer to by using the same name, *epistemic akrasia,* or equivalent terms such as "akrasia of belief", "doxastic incontinence" and, less frequently, but also importantly, "level-splitting". It's not completely clear whether those terms are synonymous, strictly speaking, or whether they capture slightly different things; nor where these things stand in relation to each other. I'm less worried about laying sharp distinctions between them than about grasping what, if anything, they have in common. What interests me the most is that it looks as though most of the things portrayed by the examples, if not all of them, display some degree of proximity to run-of-the-mill examples of the phenomenon of practical akrasia, or weakness of willpower.   Practical akrasia is a mismatch between judgment and practical conduct, that is, between what a person thinks that she should do and what she actually ends up doing, as we can see from the example bellow.

> Here is a classic example, one that, in different versions, can be found in any book on the subject. Imagine Jill, a woman who is very partial to cream, but whose diet absolutely forbids her to have any dairy products. One day Jill enters a restaurant, where on the menu is fruit salad topped with whipped cream. She thinks for a while, weighing the pros and cons of ordering her favourite dish, and finally she decides that, all things considered, she will definitely not have the cream topped salad. At that very moment the waiter arrives, and Jill orders the cream-topped salad. (Peijnenburg 2000: 286)

Many, if not all, of the examples of epistemic akrasia presented by epistemologists are re-makings of Jill's tale with the whipped cream to some extent (some more visibly, others not so much). That's because they appear to involve a mismatch between judgment and conduct, but the relevant conducts there are epistemic, rather than practical. They are conducts that impact the acquisition, management or profitability of knowledge. Rather than ordering food at a restaurant, say, agents in those examples are conducting scientific experiments, evaluating criminal evidence presented on court, giving speeches, deciding whether or not to double-check information, and so forth.

Because the selected examples are much varied, there are many ways of cutting this cake, so to speak. For the sake of easing exposition, I've separated them in groups, depending on which epistemic activity is being directly affected by the

alleged akratic conduct. Section 1.2. deals with the alleged cases of akrasia in which the conduct affects the way one engages in inquiry, while section 1.3. presents cases that affects one's way of responding to evidence. I've also included some borderline cases in the discussion, in section 1.4.. Those are the cases in which the agent's problematic conduct affects neither his inquiry nor his response to evidence, but rather something else in the broader sphere of his intellectual life. This is usually something on the edge, so to speak; because the borders between one's intellectual and practical life are sometimes fuzzy.

1.2. Cases of akratic inquiring

The first article that appeared explicitly addressing akrasia in the domain of theoretical rationality was Amélie Rorty's "Akratic Believers" (1983). There she distinguished between akrasia that affects different areas of our intellectual lives, such as perception, description, inquiry, etc. (1983: 177-179).

Akrasia of inquiry, according to her, affects inferential transitions. In it, one "comes to a conclusion following a pattern of inference that he regards as illicit" (1983: 179). One of Rorty's examples of this form of akrasia is that of the person who "conducts an experimental inquiry in ways that will predictably confirm his hypotheses" (1983: 179). While Rorty does not provide a much detailed depiction of this phenomenon herself, by means of a concrete example, it looks as though what she has in mind is a type of cherry-picking conduct. Cherry-picking is when a person selects, from a range of valid options, the one or the ones that she thinks will suit her better, in terms of her aims and projects; and deliberately ignores the bigger picture of what does not suit her.

In the context of an inquiry, to cherry-pick is to have a previously formed belief that, for some reason, needs to be confirmed or reassured (for oneself), or proved (to others). Instead of genuinely submitting the belief to a fair trial, examining which evidence exists that goes for that belief and which goes against it, the person adopts a different demeanor. She selectively extracts evidence from the world, or extracts points from an argument, *because* she thinks they support that belief; and at the same time she disregards data, or plausible interpretations of arguments, that she thinks have the power of weakening it.

As Fox & Hotch put it, "cherry pickers opportunistically take the best and leave the rest" (2005: 46). While this might be a legitimate way of conducting oneself in

some non-epistemic contexts (for instance, in grocery shopping the habit of cherry-picking leads customers to save money on the long run, cf. Fox & Hotch, 2005), when it comes to our epistemic lives it will almost invariably have bad consequences. What Rorty has in mind, I believe, when she calls this "akrasia of inquiry" is the potential this conduct has of leading to flawed inquiring. Cherry-picking evidence or points in arguments is not a genuine way of finding things out. It tailors inquiry to yield predictable results, which may be partially misleading or even completely contrary to the reality of things.

It is debatable whether the person who does the cherry-picking knows that she is doing it or not. While it might be contended that people engage in this sort of behaviour unconsciously, the case can also be made that that the person who does it does know that she is doing it, and that that's precisely why she does it. This person is interested in finding a predictable response, rather than the truth; and she cherry-picks because she knows that this conduct is effective for this particular purpose. If she didn't know or at least didn't believe that by cherry-picking she is more likely than not to attain the desired result, she wouldn't have a reason to spend valuable time and energy on the thorough sorting that cherry-picking requires, that is, her conduct would be irrational. In other words, it makes sense to think that in order for cherry-picking to be intelligible, the cherry-picker has to know, at least in some way, or to some extent, that her conduct is going to yield the predicted response, rather than the legitimate answer to a legitimate question; but she draws conclusions from this distorted reasoning and present those conclusions as legitimate nonetheless.

There is also a form of akrasia of inquiry in which one fails to inquire, rather than at inquiring. Instead of coming to a conclusion following steps that the person regards as illicit, like in Rorty's description, a person can take certain things for granted, that is, she can reason from premises that she knows she is not entitled to hold, or justified in holding. In the paper "Epistemic Akrasia and Epistemic Virtue" (2001), Christopher Hookway presents what he takes as being a case of akrasia more or less along these lines, the airplane passenger's case.

AIRPLANE PASSENGER
I can judge that the available evidence is insufficient to support some belief I hold, or believe that the methods used to acquire it were unreliable, yet still fail to form a resolution to examine the matter further. Perhaps I do not decide to phone the airline to check that my plane is not delayed despite my being aware that

> delays are common and there would be serious practical implications were I to miss a later connection. (Hookway 2001: 183)

It might not be immediately obvious how this case is to be distinguished from cases of practical akrasia (action, or it's lack, against one's better judgment). However, the distinctively epistemic component here appears to be that the airplane passenger knows that he should not reason, practically as well as intellectually, *from the assumption* that his flight is going to be on time. He knows that it is not wise to take this for granted, because he knows that many times flights are delayed by airline companies without previous warning, and that not bearing this in mind could ruin one's holiday – a harm that could be prevented with minimum effort (assuming that the effort taken to have this information checked is indeed small). So this person knows that, rather than assuming that his flight is going to be on time, the overall best thing to do is to check, to go after the information, that is, to inquiry. But, for whatever reason, he or she fails to.

Hookway's example makes the suggestion that it's not just inquiring poorly – *not inquiring* when one knows that he should be inquiring can be a form of epistemically akratic behaviour too[7]. Like in the cherry-pinking case, there is a conflict between the conduct a person judges adequate, or most adequate, and the conduct she actually undertakes.

For this to be interesting from an epistemological standpoint, it needs to be accepted prima facie that a component of what is going wrong with the airplane passenger is that he knows, at least to some extent, that he is proceeding in a suboptimal way. He might not be paying due attention to his own demeanour, he might be failing to fully appreciate the consequences of choosing not to check flight schedule, and checking flight schedule might come at some cost, which he may or may not be willing to take. However, if the airplane passenger was completely clueless, that is, if he didn't know that delays in flights are common, to start with, then Hookway would have no reason to present this as a case of akrasia, I suppose. The problem would not be a mismatch between judgment and conduct, but rather a case of mistaken judgment; or, simply, a case of ignorance.

Now have a look at another of Rorty's examples. She calls it perceptual akrasia. This variety involves one's perceptual habits, which "do not conform to the person's

---

[7] Take "behavior" here, and in every other time I use this word, in the non-technical (lay) sense.

views about what ought to be salient, or to his principles about what is important"
(1983: 177). In perceptual akrasia, the person knows she should be focusing at
certain aspects of the situation, but instead she focus at things entirely else.

> **MILITARY COMMANDER**
> For some people, at least some aspects of seeing involve *looking for*. When
> scanning the visual field is a way of answering questions, it is possible to
> reorganize one's perceptions by reorganizing one's questions. A painter who has
> become a military commander might akratically look at a landscape as a
> composition, endangering his troops by his aesthetic musing, forgetting to look
> for defensible strongholds because he is absorbed in the reflection of the sky and
> trees in the stream. (Rorty 1983: 178, italics original)

Here there is an emphasis on the perceptual aspect, which makes it look as
though this form of akrasia is a problem that affects perception. That is, it makes it
look as though the problem with the painter turned into military commander is that he
looks at a landscape with the wrong mindset and, therefore, sees the wrong things,
so to speak. But this is, nevertheless a problem regarding how inquiry is conducted.

If we expand our concept of inquiry in order for it to comprehend more than the
typical notion of investigations (which are usually thorough, systematic and explicit
attempts to find things out), it makes sense to think of cases of perception framed by
the wrong mindset as cases of poor inquiry. That's because implicit attempts to solve
practical problems are inquiries too. Inquiring is not just a means to validate scientific
hypotheses or discover perpetrators of crimes, but above all, as Jane Friedman
notes, it is a means of problem-solving. "Figure out where we put our keys, or which
country uses '86' as its country code, or whether the restaurant takes credits cards or
is cash only" (Friedman 2019: 296) – every simple cognitive task that involves raising
questions is a form of inquiring.

Like the airplane passenger from Hookway's example, part of the problem with
the painter turned into military commander is that he fails at raising some important
questions. Specifically, he fails at raising some important *subordinate questions*. The
idea of subordinate question is devised by Hookway in a previous paper (Hookway
1994). According to him, when we're engaged in a complex task, practical as well as
cognitive, we normally have to raise questions and answer them. For roughly every
minimally complex thing that we try to accomplish we execute a patterned sequence
of actions that enable us to carry out that task satisfactorily; and deciding which
actions to do, as well as how and when to do them, frequently involves raising a

number of questions, the "subordinate questions" (1994: 214). It is as though many of our non-habitual tasks carried a hidden inquiry within, whose central question is "how to accomplish this successfully?". This question ramifies into a number of simpler, non-systematic and context-specific subordinate questions whose answering will support problem-solving, in each case.

For one who is ahead of a military campaign, those might include, but not be restricted to, questions such as "where is the best spot to position those snipers?", and "are those darker and scattered shades at the top of the hill reinforcements being brought in by the enemy, or is it just how the bushes look like from afar?". Of course the agent need not be explicitly asking herself the central question, of how to accomplish victory, at each and every moment. But failing at raising the adequate subordinate questions, at the correct timing, and to answer them in a responsible manner may decisively contribute to failure at knowing how to accomplish victory, which more often that not will translate into defeat.

Rorty's idea appears to be that the military commander's failing at raising some questions of this sort while starring at the landscape bears some sort of relationship with his lingering in the contemplation of the landscape as a composition, though it is not clear what causes what. Maybe if he had been asking the right questions, his perception would have been organized differently, perhaps more adequately, in terms of his military goals. Or maybe if he had purposefully abbreviated his artistic appreciation of the landscape, he would then have been able to ask the right questions.

Be that as it may, the contemplation of the landscape is deemed *akratic* by Rorty (as opposed to just "a contemplation", or "a lingering contemplation") because we expect that the military commander should know that he should stop it at some point, and start viewing the landscape in terms of the practical issues needed for the campaign. We expect of him that he has an understanding of how he should be acting, what he should be focusing on and when (a better judgment). Nevertheless, he is not meeting this expectation. He is failing either at effectively having this sort of judgment, or at redirecting his attention accordingly. So again, as in the previous cases, we have a type of mismatch between judgment (or the lack of a judgment where there should be one) and conduct.

As in the previous cases, a person in this situation could simply be distracted, and this could be the reason why she fails at performing the relevant inquiring in the

appropriate way. It is not immediately clear how this type of case is to be distinguished from cases of pure and simple distraction.

1.3. Cases of akratic response to evidence

Now compare the cases discussed above to this example of epistemic akrasia that appears in Ribeiro (2011):

> **SKEPTIC ENTHUSIAST**
> Consider my belief that I have two hands. There are well-known skeptical arguments which purport to show that this belief of mine is not justified. These arguments purport to show that, instead of believing that I have two hands, I ought to suspend judgment about whether I have two hands, if I approve of the skeptical argument(s). Now suppose further that there is some skeptical argument that I find completely persuasive. In other words, suppose I believe some skeptical argument succeeds. It doesn't matter at all if that makes me a dullard by your lights. What matters, for present purposes, is only that I judge the skeptical argument to be completely persuasive. Finally, suppose that, despite the fact that I find some skeptical argument completely persuasive, I continue to believe that I have two hands (instead of suspending judgment about whether I have two hands). (Ribeiro 2011: 21)

Like the examples from the previous section, Ribeiro's example involves a mismatch between judgment and conduct too, except that the relevant conduct here is not inquiring, but rather evaluating evidence. What we have in this case is: a person has a certain belief, and he also has a conceptual understanding about the reality of things, i.e., a judgment. According to this person's judgment, a certain skeptical argument succeeds, that is, skepticism (the doctrine that we can never really *know* things from perceptual experience) is true.

Now, this person's belief that he has two hands is formed on the basis of evidence from perceptual experience. This belief is in stark opposition with his understanding of reality, his judgment. Because if the skeptical argument says that we don't know whether or not we have hands, and if he finds that the skeptical argument does succeed, then he shouldn't be having the belief that he has hands, in the first place. We would expect that evidence from perceptual experience indicating that he has hands be interpreted by him as misleading and, therefore, as not offering support to this belief, which should be, then, dismissed. Notwithstanding, the belief subsists, that is, it doesn't go away just because he thinks it should. It's a recalcitrant belief – a belief that persists even though it conflicts with the agent's overall views of reality.

This case is different from the cases involving inquiry, from the previous section, in some important respects. Here it is less a matter of contention than it was in the previous cases whether or not the agent knows about the mismatch. Because of the way Ribeiro spells things out, it is as though the agent is telling the story first-hand; so it is clear that this is not a case of unconscious belief, nor of distraction.

Also, in this case, what is in fact required from the agent in order for the conflict (between judgment and conduct) to be undone is different from what was required from the agents in the previous cases. While it is a matter of debate whether or not mismatches of this sort indeed need to be undone (that is, whether or not we are rationally required to try and undo them), it is clear that the adequate way to do away with the mismatch, if it indeed needs to be done away with, varies between this case and the ones from the previous section. In the examples from the previous section, the right way to resolve the conflict between judgment and conduct would be to amend conduct, that is, to make conduct conform to judgment, rather than the other way around. Here, however, it looks as though it is the person's current judgment that needs to be altered in order to accommodate the odd belief, because a person cannot choose to alter her perceptual beliefs. But a person can do away with the mismatch by revising her judgment, switching from "the skeptical argument is completely persuasive" to "the skeptical argument is not completely persuasive". This person's akrasia lies in the fact that his appreciation of the evidence doesn't suffice to motivate him to perform this adjustment.

Now, the next two cases do not involve perceptual evidence. They're Hookway's (2001) case of the credulous mother, and Mele's (1986) case of the jealous husband.

CREDULOUS MOTHER
(…) the mother has the goal of preserving the reputation of her family. She has sufficient reason to adopt this goal and has reasonable views about the relative priority of her different goals. Suppose she also believes: 1) If her son is innocent, the reputation of the family is best preserved by declaring her belief in his innocence and doing all she can to secure his acquittal. 2) If her son is guilty, the reputation of the family is best preserved by denouncing his immorality, disinheriting him and announcing that she no longer sees him as a son of hers. The occasion has now arisen when she must decide whether to stand by her son or denounce him. When she works out how to act on this occasion, what would we expect her to do? (…) Reflection can always short cut the expected effects of her akratic belief upon her behaviour. It begins to look as if we have conflicting subjective conditionals of the problematic kind: If she believes that her son is innocent, she will defend his reputation; If she believes it is rational to believe that her son is guilty, she will denounce him. (Hookway 2001: 185-187)

Hookway presents this as an example of epistemic akrasia because the mother "both believes in her son's innocence and believes that the weight of reasons supports his guilt" (Hookway 2001: 187). Like in Ribeiro's Skeptic Enthusiast, the problem here lies in the fact the agent's response to evidence does not suffice to motivate her to change her mind; hence the mismatch between her actual conduct (of believing that her son is innocent) and the conduct she thinks she should have (of acquiring the beliefs that evidence supports). Compare it to Jealous Husband.

JEALOUS HUSBAND
Consider (…) the insecure, jealous husband whose initial suspicion that his wife is having an affair develops into a genuine belief, even though he knows that his evidence that she is being unfaithful is quite weak, that he has much better evidence that she is not having an affair, and that he would be much better off not having the belief in question. (Mele 1986: 217)

Jealous Husband displays the mismatch between judgment and conduct too, to the extent that the husband knows that evidence does not support the belief that his wife is having an affair; so he knows he should not have this particular belief, but he has it anyway. In this sense, it is a case of recalcitrant belief too: belief (that his wife is having an affair) that persists even though it conflicts with the agent's overall appreciation of what the evidence supports.

The main difference between Credulous Mother and Jealous Husband appears to be that, in the former, the odd belief was formed previous to her assessment of the evidence, that is, the mom's belief in the son's innocence was already there when she was presented with evidence indicating otherwise, but assessing the evidence wasn't sufficient to motivate her to revise this belief. Jealous Husband's belief that his wife is guilty of adultery, in turn, appears to have been formed after he assessed the evidence, in spite of the evidence not supporting it. Apart from this, because those two cases are very similar in structure, I'll focus my discussion on one of them, Jealous Husband, and the remarks I make about it I take to apply to Credulous Mother as well.

As pointed out before, one relevant difference between jealous Husband and Skeptical Enthusiast is that in the former the relevant evidence, as well as the relevant belief, are not of the perceptual kind, which means they might be subjected to different norms of evaluation. The Skeptical Enthusiast might be said to be justified in holding the belief that he has two hands (since he indeed has two hands) from an

externalist viewpoint, in spite of this belief conflicting with his own views. As for the Jealous Husband, in turn, it's difficult to see how he could be justified in holding the belief that his wife was having an affair either from an internalist or an externalist viewpoint (since the wife was not having an affair, and since he knows that evidence does not support the belief that she was). So the mismatch between judgment and conduct makes Jealous Husband look irrational, more so than Skeptical Enthusiast. Another way of putting this is: in both cases we have the mismatch between judgment and belief, but in Jealous Husband we have a correct judgment and an incorrect belief, while in Skeptical Enthusiast we have an incorrect judgment and a correct belief. It's easier to understand the mismatch in Skeptical Enthusiast, because one could say that he is basically mistaken: his reasoning is being informed by a mistaken judgment.

Now, another part of what makes the mismatch look irrational in Jealous Husband is the assumption that he knows he should not be concluding, on the basis of this evidence, that his wife is having an affair; but this is exactly what he is concluding, nevertheless. Is there a way of reading this case in which the jealous husband doesn't look irrational? Well, an alternative way of viewing it is: the agent is not concluding that his wife is having an affair *on the basis* of evidence that he knows to be insufficient to support this conclusion, but rather *in spite of* this evidence. That is, he believes his wife is having an affair not because he has seen evidence that indicates this (he hasn't), but for some other reason that has been left out of the description (some external reason, possibly).

It could be, for instance, that his wife has been arriving home impregnated with an after smell of male's perfume that is not his. When she arrives, he cannot point his finger at what it is that hints to him that she has been unfaithful; but he can sense it, so *just knows.* Robert Brandom (1998) presents a similar case when he discusses chicken sexers. Expert chicken sexers are professionals that reliably determine the sex of day-old chicks with close to 100 per cent accuracy. If you ask them, however, they will tell you that in many cases they cannot point their fingers at what it is that gave them the basis for decision. They just look at the rear end of a chick and see that it is male, or female. In this case, the husband would be justified in believing that his wife is having an affair against his appreciation of the evidence as much as a chicken sexer is justified in his deliberations.

One obvious problem with this alternative account is that it depends on the wife being indeed having an affair or, at least, arriving home with the funny smell. However, many of the real life cases of jealous husbands that resemble this case (that is, cases of men thinking that their wives are having affairs) do *not* have this component. In them, the wives are not having any affairs and there is no element hinting otherwise that suffices to make the husband's belief that they are having an affair any reasonable. It looks as though those are the cases Mele was primarily interested in approaching, then.

Another downside of this alternative interpretation is that, even in the eventuality of the husband being right, the point remains that his response to the available evidence is unsatisfactory. Because in this case, it appears, he should *not* be accepting "that he has much better evidence that she is not having an affair". If the after smell equips the husband with reason that suffices to render the belief that his wife is having an affair justified, or rational, then we would expect from him that he dismissed the evidence that she is not having an affair, at this point. We wouldn't expect from him that he remained in a state of dissonance. We expect coherence from rational people when they're dealing with evidence. So it is as if what Mele is calling "akrasia" in this case is one's failing to fulfil this expectation, or to comply with this norm, by failing at attaching the right weights to the available evidence. The person fails at dismissing evidence or sticking to them, at the right timing.

When people fail at attaching the right weight to evidence, they err in a way that is different from the way they err when they fail at conducting a profitable inquiry, although the two ways of erring can of course overlap. In fact, many, if not most of the times, these two ways of erring go hand in hand. But not necessarily. Without having incurred in any critical evaluation mistakes, candidly assessing the worth of all the evidence you happen to have access to, you can still fail at the level of inquiry. One example of this is given by unsound inductive generalizations. It could happen when, say, your evaluation of the *available* evidence was pristine, but what you really needed was to go after more evidence, before generalizing. Other examples concern cases of asking the wrong questions and following the wrong leads. On the other hand, you cannot inquiry in an effective and responsible manner while at the same time misevaluating critical evidence. If you misevaluate critical evidence, your inquiry is going to be flawed, even if you end up with true beliefs. That's because correctly

evaluating evidence is part of what a profitable inquiry is, and flaws in the former are going to indirectly affect the latter.

Now, there is yet another important difference between forms of akrasia that affect one's evaluation of evidence (you respond to evidence in a certain way while knowing that this is not the way you should be responding) and forms of akrasia that directly affect inquiry, that is, akratic ways of inquiring (you reason in ways you know are not ideal, or even not acceptable). Inquiring has a sort of diachronic dimension that the simple evaluation of evidence lacks. That is, inquiries typically are something that extends over time. They take some time, because they are constituted by a series of coordinate sub-actions and deliberations. The act of attaching weights to evidence, on the contrary, can be a prompt, in loco response. So it makes sense to think that akrasia that affects one's evaluation of evidence is a flaw in the way one responds, whilst akrasia that directly affects the way one inquires is a problem in the decisions one makes, or fails to make. A consequence of this is that the epistemic (ir)responsibility involved in each type of case is going to be of a different type, since the type of control we have over our decisions is different than the type of control we have over our response to data.

Last in this group, there are the cases involving the so-called *misleading high-order evidence.* In them, a person has some evidence *e* indicating that *p*, so she forms a belief that *p*; and then she comes across some further evidence indicating that maybe *e* does not really support *p*. In other words, the high-order evidence sometimes gets in the way of a person's evaluation of evidence, because it casts doubt on the correctness of her original assessment. Some of those cases have been referred to as examples of epistemic akrasia by scholars, depending on the way the person deals with the high-order evidence.

One of the most famous examples of high-order evidence getting in the way of a person's evaluation of evidence is Christensen's case of hypoxia (2010b: 126-127)[8]. Hypoxia is a condition in which the body is deprived of adequate oxygen supply. It commonly affects mountain climbers and pilots who are exposed to the low partial oxygen pressures encountered at altitudes higher than 10,000 feet. It causes impaired judgment, but it is very rarely recognized by the sufferer at its initial onset.

---

[8] Similar cases, involving seemingly misleading high-order evidence, have been presented by Horrowitz (2014) and Daoust (2018). Because those cases are not significantly different from Pilot in structure, I've opted for presenting just one of them.

Thus, the realization that one is at altitudes higher than 10,000 feet is evidence that one might be under deep but undetectable cognitive malfunction, by virtue of which her reasoning could be flawed and, therefore, untrustworthy.

Here is how Christensen builds the case:

PILOT

You're alone, flying a small plane to a wilderness airstrip. You're considering whether you have enough fuel to make it safely to an airstrip 50 miles further away than your original destination. Checking the relevant factors, you become extremely confident that you do have sufficient fuel — you figure that you've got a fair bit more than the safety-mandated minimum. So you begin your turn toward the more distant strip. But then you notice that your altimeter reads 10,825 feet. You feel completely clear-headed and normal; however you're fully aware of the insidious effects hypoxia can have. Should you trust your recently formed confident judgment about having sufficient fuel, and continue on your path toward the more distant airstrip? (Christensen 2010b: 126)

The pilot's case can be given a more technical description as follows. A pilot wants to know whether "fuel is sufficient to get to the more distant airport" (p). So she does the math (calculations as to distances and fuel efficiency, taking all the relevant parameters and data into account); and the result of these calculations, m, indicates that p. That is, m is evidence that p. But then, she looks at the altimeter and is reminded of the phenomenon of hypoxia. The information displayed by the altimeter, m_2, suggests that she might be suffering from hypoxia and, therefore, might have reasoned poorly (for instance, by miscalculating, or by failing to take into account all the data that happens to be relevant for the calculations). So m_2 is high-order evidence: it doesn't concern p directly, but rather the support relation between m and p; or the judgment that m supports p. If this judgment was correct in the first place, then m_2 is misleading high-order evidence – it's evidence wrongfully indicating that m doesn't support p[9].

In this situation, the pilot could respond in different ways. She could

1. keep p as well as her judgment that m supports p.
2. drop p as well as her judgment that m supports p.
3. keep p, and drop her judgment that m supports p.
4. drop p, and keep her judgment that m supports p.

---

[9] It is worth noting that m_2 suggests that p is not supported by m even if the pilot doesn't really have hypoxia. That is, m_2 exerts pressure against the support relation held between m and p independently of m_2 being true. If it were true, it would be what is called, within the literature on evidence, an "undercutting defeater"; but it doesn't have to be true neither to be known to be true by the pilot for its effects (of setting up the state of self-doubt) to come by.

As Sherrilyn Roush (2017, §3) observed, there is something conceptually problematic about scenario 4. It's not a plausible way of responding to the situation, for if one is definitely sticking to the judgment that m supports p, then she has no reason to drop p (dropping p would be irrational by all accounts, even if p is false, in the end). We set scenario 4 apart, then, and we're left with scenarios 1-3.

Scenario 1 is what epistemologists often refer to as "the steadfast response" to evidence (Kelly 2005, Schoenfield 2014). The agent in this scenario "sticks to her guns", so to speak. She thinks that the second-order evidence, m_2, does not impact the support relation between m and p at all. Scenario 2, in turn, is termed "the conciliatory response" (Feldman 2005, Elga 2007 and Christensen 2014). In it, the pilot sees m_2 as a *defeater* to the support relation between m and p[10]. As a consequence of seeing things in this way, she is compelled to change her mind. She abandons her initial judgment that m supports p and then rejects p, because now p is no longer supported by evidence.

Scenario 3, in turn, depicts what scholars sometimes refer to as the "level-splitting" response to evidence (Wedgwood 2011, Williamson 2011, 2014; Coates 2012, and Lasonen-Aarnio 2014); and it is considered an example of epistemic akrasia (Roush 2017). The pilot's akratic response to evidence in this scenario is due to the mismatch between his judgment and his conduct. He sees m_2 as a defeater to the support relation between m and p, but he doesn't reject p. On the contrary, he sticks to p, even though he now thinks that p is no longer supported by the evidence. As Roush (2017) remarks, "the state in which you believe 'F and my evidence does not support F' is a case of 'level-splitting', also called *epistemic akrasia*, because you believe you ought not to have a particular belief state but you have it anyway".

Like Jealous Husband, the Pilot is failing at meeting our expectation of coherence, or, in another way of putting it, he is violating a coherence constraint. It is a matter of dispute whether or not such a constraint exists as a real requirement of rationality (some scholars contend that it does not, for instance, Wedgwood 2011, Williamson 2014, Weatherson 2008 and Coates 2012). This is a reflection of the lack of consensus over what the correct way of responding to misleading high-order

---

[10] In John Pollock's (1989) terminology, some defeaters of justification for a conclusion are rebutters, that is, are simply evidence against the conclusion, while other defeaters are undercutters; they undermine the relation between the evidence from which the conclusion was drawn and the conclusion. In PILOT, if m_2 is not misleading, it is an undercutter. It undercuts the relation between the result of the pilot's math and the belief that fuel is sufficient ($m \therefore p$).

evidence is. Apparently, there is more than one plausible way of responding to high-order evidence, and the adequate way might vary according to the situation's specifics.

This renders Pilot a somewhat exotic case of akrasia, because when we think of practical akrasia, there isn't more than one plausible way of responding to whatever it is that is causing the disruption of coherence in action. Think of classical cases of practical akrasia, such as the cases of people giving in to temptation, like Austin's (1979) example of the person who couldn't help picking one more piece of chocolate cake, in spite of knowing that she was not supposed to. The remarkable difference between this and Pilot is that it looks as though Pilot could be purposefully compromising, whereas the person in the chocolate cake example is simply losing control. She can't help it. The desire for the extra piece of cake overtakes her.

1.4. Borderline cases

Some of the cases that appear in the literature as examples of epistemic akrasia have to do neither with one's evaluating evidence nor with erring while conducting an inquiry. They are borderline cases, to the extent that it is not completely clear that they are not to the other side of the "line" between our conscious cognitive activity and our non-conscious, or unconscious, cognitive processing.

One of those cases is what Rorty calls descriptive akrasia. It is mostly about habits of speech. In this form of akrasia, "the phrases that a person uses to describe his situation carry categorical and classificatory implications and presuppositions that affect action; they are often strongly but latently evaluative" (Rorty 1983: 178). The person incurring in this form of akrasia falls into "pre-fabricated patterns of speech, self-deceptively thinking that such lapses don't matter" (1983: 178). Rorty's example is that of a man engaged in conversation who falls into a way of speaking that he himself disapproves of.

CLOSETED SEXIST
Someone committed to non-sexist attitudes usually can avoid following his cohort in describing women in demeaning ways, knowing that if he talks of women as broads or chicks he is less likely to listen to what they say, less likely to interpret their remarks in the same charitable way that he would interpret the same words as spoken by a man. What he calls imaginative initiative in a man, he calls conniving manipulation in a woman; akratically he comes to think the man

requires and deserves respect and cooperation while the woman is to be belittled, avoided (Rorty 1983: 178).

Here we have a person who sees himself as committed to a general non-sexist principle. He has a resolution, or a general understanding of things (say, a judgment), according to which women are not inferior to men and should not be treated as inferior. Apparently, this commitment is strong enough for him to avoid calling women "broads" or "chicks", but not strong enough for him to avoid describing a woman's imaginative conduct as "manipulative". So the mismatch here is between the meaning of the things he says and the meaning of his resolution, or his overall understanding (as opposed to a mismatch between beliefs and judgment, like in the previous cases). There is a clear epistemic dimension to this mismatch, which lies in the fact that it brings epistemic injustice about, hence it being deemed a case of *epistemic* akrasia (as opposed to one of practical akrasia).

From all alleged cases of epistemic akrasia presented so far, this is the one I find the most difficult to make sense of, even though the attitude it depicts is remarkably unexceptional. I find it difficult to make sense of this as a case that belongs to the same general category as the previous ones, or that is more similar to them than dissimilar.

For one thing, differently from the other cases, it doesn't look as though the person here *knows* that he is doing something he shouldn't be doing. He might as well not be aware of the abovementioned mismatch at all. This sometimes happens to us – we do things whose consequences we didn't fully understand in the moment of action. A person who is currently on a diet, trying to lose weight, might chose to have a big bowl of caesar salad for lunch instead of a sandwich, not knowing that the former has three times more calories than the latter. This person is not aware of the inadequacy of her current action to her resolution. It's not clear that the closeted sexist from Rorty's example is not one of those. He might simply not know that by describing a particular women's creative conduct as manipulative he is wounding his own non-sexist principle.

What's more, there seems to be plausible alternative explanations for what is going on in the closeted sexist case, that need not involve the concept of akrasia at all. One of them is that it is a simple and plain case of ignorance, like I've just suggested by means of the caesar salad analogy. Another one is: maybe the case

could be accounted for in terms of implicit cognitive bias. People under implicit cognitive bias haphazardly say things they didn't mean; words simply escape their mouth in a non-reflexive way. And they might not know that their words mean what they in fact mean, that is, they might not be aware that they are coming out as prejudiced, for instance. If so, then they could hardly be said to know that they're doing something they shouldn't be doing. But in all the other cases presented so far, it is easier, or at least less strange, to accept that the person knows that something is odd in the way she is reasoning, and this is what their akrasia is all about: the person knows about the mismatch, but knowing doesn't suffice to elicit change.

Also, it is not strictly speaking correct to say that people under implicit cognitive bias do what they do, or say what they say, deliberately. It is not as if they deliberately chose to act in a way that goes against their general commitments, or their better judgment, like typical akratic agents (practical akrasia) and agents in the other examples do. A typical akratic agent does **a** in circumstances under which it was completely possible for him to have done **b** instead. But it's not clear that the same could be said of the agent in Closeted Sexist.

Another alternative explanation for what is going wrong with Closeted Sexist that doesn't involve the idea of akrasia could be that this is a case of self-deception. According to Mele (2001) when we are self-deceived, we believe something because we want it to be true, even though it is false. So maybe the person in Closeted Sexist wants to preserve his self-image as a non-sexist person. That is, he wants it to be true that he is not a sexist, and that's why he believes that he is committed to gender-equality principles, whereas in fact he is not. While I acknowledge that there might be substantial overlap between self-deception and akrasia[11], it looks as though we can make sense of Closeted Sexist as a simple and plain instance of self-deception *instead* of akrasia, which leaves us to wonder where the line should be drawn between these, as well as other, neighbouring concepts.

Be that as it may, one of the reasons Rorty might have had to term Closeted Sexist an example of epistemic akrasia is the fact that in this case there is a mismatch between the way the person thinks that he should think, that is, his non-

---

[11] Mele (1986: 221-222) attempts at qualifying this difference by saying that, for self-deception, one's belief that **p** must be false; and non-epistemic considerations in favour of holding or not holding **p** must be irrelevant to the charge of self-deception. In contrast, one may akratically endorse a proposition **p** that is true, and non-epistemic considerations (such as preferences and habits) might be quite relevant to a charge of akrasia.

sexist resolution; and the way he (apparently, based on what he says) actually thinks; even though this might be a momentary mismatch, only. Maybe this person is in the middle of the process of becoming non-sexist. If so, then it is likely that he already knows what to do, or what would be right for him to do, but still isn't capable of doing it thoroughly. That's because the process of becoming a non-sexist person takes time. From the moment he formed the resolution to honour the principles of gender equality to the moment he is a person that conducts himself in full accordance with those principles there is a time gap. It's the gap of education, or self-improvement, so to speak. Meanwhile, he is prone to fall into old patterns, sometimes. At those times, the mismatch comes to surface. It's a mismatch between his better judgment and his actual conduct.

Another alleged case of epistemic akrasia that could be termed borderline is Heil's psychoanalytic patient, which is similar to Closeted Sexist to some extent.

PSYCHOANALYTIC PATIENT
The incontinent believer is typified by the psychoanalytic patient who has acquired what might be termed an intellectual grasp of his plight, but whose outlook evidently remains unaffected. Such a person has failed somehow to integrate his appreciation of certain facts into his overall psychological state. He continues to harbour beliefs, desires and fears that he recognizes to be at odds with his better epistemic judgment (Heil 1984: 69).

Here we don't know exactly what plight the agent is involved with (but provided that he is a psychoanalytic patient, I'm guessing it could be something of the sorts of "do I actually resent my mom?", "am I actually bisexual?" or "what happenings from my childhood shaped me into who I am?"). Be that as it may, we have the following situation: some information has just been disclosed for the patient that is of relevance for the prospects of him making sense of his plight. However, the process by means of which this new information is incorporated into his "chart" of everything that is relevant for that specific end does not self-complete immediately. Rather, like in Closeted Sexist, it takes time. This is the process by means of which the newly acquired information would result in some "amendments" within that chart (for instance, the dissolution of certain fears, the intensification of certain desires, the dismantling of certain beliefs, etc). Whilst this process is not completed, some things in the patient's life remain as if that information hasn't ever been assessed, hence the

mismatch between his current understanding of things and the newly acquired information.

This example I find difficult to make sense of too, because, like in Closeted Sexist, I'm under the impression that, in real life cases that resemble this tale, we might be talking about information that is not held consciously; and also that the patient does not have control over the process by means of which the newly acquired information would be incorporated. So it is not as though the patient is deliberately refraining to incorporate, or deliberately harbouring her old beliefs, desires, fears, etc. In other words: he might not be aware of the mismatch and, therefore, not be able to undo it.

Setting this issue aside, it is interesting to take notice of the fact that timing plays an important role in the case having been described by Heil as a case of akrasia. If the process of incorporation (of the new information to the person's inner "chart" of everything that is relevant for her quest) had begun and ended quickly, or immediately, there would be no "space", so to speak, for the information to be held against the agent's overall judgment. The same holds true of Closeted Sexist: if self-improvement was a process that began and ended quickly, or instantly, it wouldn't be possible for her to fall into the lapses of speech discussed.

1.5. Some pressing issues

The previous three sections made it explicit that things being called "akrasia" in the domain of our theoretical rationality, or our epistemic lives, are quite varied. The types of situations depicted by the examples are diverse, and what is being pointed at, in each case, as being the *akratic conduct* also differs considerably from case to case. This makes us wonder: is there a common core shared by all of those conducts? If there is, what does it consist of?

Though the "architecture" of the many examples varies, they appear to have a common denominator: in all of them, there is visible *tension* between how the person proceeds intellectually in the situation, that is, the way she reasons, on the one hand; and her own general conception about what would be the right way for a person to reason, on the other hand. How she thinks does not fully match the way she thinks that she should think. So when we ask whether there is a common core shared by the many examples, a good candidate is: a mismatch between a person's judgment and her epistemic conduct.

Now this, in turn, raises a number of questions. First, can we say that the person in each and all of those examples has a fully-formed judgment about conducts, that is, that she knows that a certain course of action (concerning inquiry, evidence evaluation or other aspects of her intellectual life) is correct, or the most correct, overall? In some of the examples, this is very clear (for instance, in skeptical enthusiast) whereas in others it is not (for instance, pilot).

If knowing that a certain way of acting is correct, or the most correct, overall is necessary for akrasia, then it would have to have been *stipulated* by us, or by each and every scholar that came up with the examples, that those are indeed cases in which the person knows what the right call is, for them to be examples of akrasia, to start with. The possibility will always remain, however, that in the real life cases that resemble the many examples given here to some extent the person does *not* know that a certain course of action is correct, or the most correct, overall. The concept of akrasia wouldn't be much helpful for us to deal with those cases, then.

A consequence of this is that it's always going to be difficult to tell, as to each real life case that resembles the examples given here, whether or not it is a case of akrasia. That's because it's always going to be difficult to determine whether or not the person did know that the correct, or the most correct, thing to do was different than the one she actually did. If we can determine that she knew that the overall most correct thing to do was *a* but did *b* nonetheless, then this will probably be a case of akrasia. But can we always determine this? It's really hard to tell.

Second, is the concept of akrasia *equivalent* to "mismatch between judgment and conduct", or is rather meant to *explain* (some) cases of mismatch between judgment and conduct? Many scholars thought that akrasia was something that needed to be either explained or explained away, because the concept appears to be self-contradictory (for instance, Davidson 1969; and Hare 1963). Others, in turn, thought that it was rather something that possessed explanatory power over other things, over phenomena that we experience in our ordinary lives (for instance, Mele 1987). Even though both seem to be plausible unfoldments of the discussion, I follow Cassam (2021) in holding the general view that what we really should aim to achieve when dealing with seemingly alien outlooks and practices, like the ones depicted by the examples, is something called *Verstehen.*

Verstehen is an activity in sense-making. To try and expand our understanding of other people's, conducts, including their epistemic conducts. To acquire Verstehen

of other human beings is to put oneself in their shoes, so to speak; to be able to see things from their perspective, in terms of the reasons they have to respond to a situation in the way they respond. What philosophy in general, and epistemology in particular, should be trying to do, the way I see it, is to attain this epistemic goal, Verstehen. So when philosophers set themselves to examine cases of apparent mismatch between judgment and conduct, and when in so doing they deploy a concept such as "akrasia", if they're doing epistemology the most dignifying way, so to speak, what they're hopping is that that concept is capable of illuminating those cases, of expanding our understanding of what is going on. It might as well be true that the very concept of akrasia still stands in need for clarification (need to be either explained or explained away, like Davidson and Hare say), but this is only because we expect this concept to play an explanatory role on our understanding of other things.

Can the concept of epistemic akrasia live up to this expectation? That is, can it be a tool for Verstehen? If this concept is meant to help us understand cases of seemingly alien conduct, then it needs to involve more than the mere statement that there is a mismatch going on between a person's judgment and her conduct. That's because there are many possible ways in which the idea a person has about what should be done can fail to match her actions and beliefs, that is, fail to coincide with what she actually ends up doing or concluding.

I can end up doing something different than what I thought should be done because when the time came for me to act I felt lazy and did it another way. Or because I got distracted and wasn't paying attention. Or because I suddenly found myself deprived of the proper means to act, or compelled by external forces to do something different instead. Or because I was in the process of changing my mind. Or because I haven't made up my mind in a resolute way to start with, that is, I had a vague idea of how things should unfold, but couldn't come up with a steady and specific plan of action. Or, it could have happened because of some sort of weakness on my side, of the sorts that we talk about when we say that we couldn't resist temptation, or that we've *failed* because we've allowed ourselves to be overtaken by something stronger than our will (for instance, passions). Those are fairly different types of mismatch between judgment and conduct, with different root causes and different implications. Usually, the notion of akrasia is deployed to account only for the last type mentioned. But for the notion of epistemic akrasia to be

a tool for *Verstehen*, the mismatch to which the concept is being applied has to be *better explained by* this notion *than* by other explanatory resources.

Now, how are we to differentiate between the situations in which a mismatch between judgment and conduct is better explained by the concept of akrasia and the situations in which it is better accounted for by alternative explanations? Sometimes a mismatch between judgment and conduct is better explained by structural factors. Employees subjected to poor working conditions might know that their situation is unfair and might rationally acknowledge that their best chance of improving it involves an all-out strike. Notwithstanding, they keep showing up to work day after day. This is a clear example of mismatch between judgment and conduct. Even though someone could claim that this case involves agents allowing themselves to be overtaken by something stronger than their will (habit, routine or even fear), the case is much better accounted for by means of an explanation that appeals to structural factors: organizing a strike takes a considerable amount of social coordination and collective effort, and this might get in the way of the workers actually doing what they think will deliver the desired outcome.

Likewise, sometimes a mismatch between judgment and conduct is better explained using a sub-personal terminology and conceptual repertoire. Take as an example blindsight patients. Those patients are cortically blind due to lesions in their striate cortex, also known as the primary visual cortex; but they're capable of responding to visual stimuli that they do not consciously see, nonetheless. Those explanations are satisfactory enough without involving the idea of "akrasia", or related concepts, at all.

But how are we to tell those cases apart? That is, how to distinguish the situations in which the mismatch between judgment and conduct is better explained by the concept of akrasia, on the one hand, and the situations in which it is better accounted for by alternative explanations, such as structural and sub-personal explanations, on the other hand? Well, coming up with a rule of thumb, or a pre-set list of criteria, would certainly prove itself a tricky move that I wouldn't attempt at, here. I do believe, though, that a good start might be the following: judging by the myriad of examples, akrasia is clearly meant to be a personal concept. And personal concepts are typically useful where non-personal explanations don't seem to make sense, or don't feel satisfactory. We resort to personal concepts when appealing to

context, structure and chemical or biological factors either doesn't make sense or isn't particularly illuminating on its own.

Let me explain it in a more detailed manner. Akrasia is meant as a personal concept because it places the quibble, the flaw, at the level of the individual's space of reasons. The borderline cases might make this a little obscure, because in them other sorts of explanation for the alien conduct are available *as well*; but the thing is that, in each and any of the cases presented, the authors were making the case that people in those examples either did something that they shouldn't have or didn't do something that they should have, for what appears to be (but probably isn't) no reason whatever. That is to say, the authors were making the case that the *person* failed in making her conduct conform to her judgment. In addition to that, they're implying that the person knew that she was thus failing. We resort to personal concepts, concepts that are meant to capture a *person's* flaws, when other types of concepts don't to the job. In other words: we blame people when blaming nature, or the circumstances, only take us so far.

So this is a good clue about when the very idea of akrasia is meant to be epistemologically valuable: when some reference to the person's power, or to her autonomy, appears to be missing from other otherwise reasonable explanations of her epistemic conduct. This takes us directly to the next matter of concern, which is: speaking of power, or of its lack thereof, suggests that what is being talked about is a form of personal *weakness*. Now, what's the relationship between what's being called epistemic akrasia here and the phenomenon of lack of willpower, or practical akrasia, that we find in the philosophical literature since Plato? Is it being implied by the alleged examples of epistemic akrasia that that which is going on in those cases mirrors the phenomenon of practical akrasia in some meaningful way? If it is, then what meaningful way is this?

Practical akrasia is the name given to a person's conduct when she is placed before a bifurcation, of some description. There's the possibility of action A, and there's the possibility of action B. The person judges, say, A to be best overall. So she thinks that A is what she should do. Nevertheless, when the time comes, she ends up doing B, in complete disregard of her own resolution. Those cases are classically equated with lack of willpower, or weakness of the will, and the paradigmatic examples always involve the person being overridden by appetites of all sorts (sexual drive, the pleasures of food and drink, and so forth). So it is

suggested that a person's appetites can be stronger and more powerful than her reasonability, or her rational capacities, which, in turn, are rendered weak. It's because there is this mighty force, the appetite, that the person's will to do what is right becomes weak and ineffective. Now, how is this paralleled by the many examples of puzzling epistemic conducts presented throughout the chapter? Is there a real parallel?

For there to be a parallel, the epistemic examples must involve a bifurcation, of some sort, between options A and B; and they must contain an ingredient which would be the epistemic equivalent to that "mighty force", the appetite. The bifurcation between options A and B could be a bifurcation between inquisitive measures, like we've seen from section 1.2. (to inquire towards some matter or not to; to perform a certain set of experiments or not to, etc.). It could also be a bifurcation between the entailments of plausible interpretations to be attached to evidence, like we've seen from section 1.3. (does evidence **e** means this, or does it mean that? And so forth). And it could also be a bifurcation between other aspects of one's intellectual performance, like the borderline cases from section 1.4. suggest (if they are indeed at this "side" of the border, rather than to the other side). For instance, to describe this thing using these words, or those words? And so forth. Therefore, that the agent in all those cases is placed before a bifurcation is pretty visible.

Now, that there is that other ingredient, a "mighty force", that renders the person's resolution weak, is more difficult to see in the epistemic examples. What mighty force could this be? The idea of there being "epistemic appetites" sounds a bit quaint. Could it be that some non-epistemic reasons are at play, some very strong external reasons, that are capable of overtaking one's will to act in accordance with his or her own appreciation of what the right way of reasoning is? Or could it be the person's passionate commitment to a theory, a faith, a previously formed belief, or even a habit of doing things in a certain way? It's really hard to tell. And, if there is such a thing, a "mighty force" that gets in the way of a person thinking straight, then it's probably something that varies from case to case. This hints to us that the parallel between practical akrasia and alleged cases of epistemic akrasia, if it is fair to say that there is one, is probably only of a mild type; and that the latter form of akrasia is not meant to perfectly mirror the former (a topic that will be examined in greater depth later on).

Be that as it may, even in the absence of something identifiable as this "mighty force", we've seen from the examples laid out throughout the chapter that akrasia of the intellect is a matter of an insufficiency, and in this it bears a relevant resemblance with cases of weakness of the will (practical akrasia), because those are a matter of an insufficiency too. In practical akrasia, one's practical judgment does not suffice to motivate them to do the right thing, whereas in epistemic akrasia it is one's epistemic judgment that does not so suffice.

Epistemic akrasia of the sort that is illustrated by the many examples given in this chapter appears to be a matter of one's previously formed judgment not being enough to motivate them either to undertake the adequate inquisitive demeanour (cherry-picking, airplane passenger, military commander), or to revise an inadequate judgment (skeptic enthusiast, jealous husband, credulous mother, pilot), or to adjust other inadequate aspects of their intellectual life, such as speech (closeted sexist) or self-knowledge (psychoanalytic patient). The common denominator is that in virtually all of those cases of akrasia there is *insufficiency* of some sort. Something that a person knows or thinks that she knows is not sufficient to elicit the adequate state of mind and/or action. Another way of putting this is: in both epistemic and practical akrasia the person fails at being adequately moved by a certain type of knowledge – knowledge of what the right thing to do or think is.

1.6. Final remarks

With this we've managed to sketch some preliminary answers to some of the most pressing questions about the very idea of intellectual akrasia.

● **What's the extension of this concept, that is, what does it apply to?** It applies to a broad range of cases of seemingly alien intellectual conducts that appear to involve a mismatch between judgment and conduct, affecting either a person's inquiring measures, her evaluation of evidence or other aspects of her intellectual life, such as speech and self-knowledge.

● **Is there a common core shared by all of the examples which the concept has been used to refer to?** Judging only by the examples, there appears to be a common denominator, but one that is very difficult to be spelled out in terms there are anything more than vague. In virtually all of the examples examined, the

mismatch between judgment and conduct takes the form of an insufficiency, of some description. It looks as though the person's judgment does not suffice to motivate her either to adopt the adequate inquiring practices, or to adequately appreciate evidence, or to make the adequate adjustments to other aspects of her intellectual life, such as speech and self-knowledge.

● **What's the relationship between epistemic akrasia and practical akrasia?** Most of the situations portrayed by the examples examined have been construed in a way that displays some degree of resemblance to run-of-the-mill examples of the phenomenon known as lack of willpower, or weakness of the will (practical akrasia). The fact that the two phenomena resemble one another at least superficially suggests that they bear some meaningful relationship. In both of them the person is placed before two possibilities (of intellectual conduct, or of practical action); she judges that one of them, say, A is the best overall, but ends up choosing B nonetheless. It looks as though the person's judgment that A was best overall was too weak – it was not sufficient to motivate her to act accordingly. It is not clear, however, why it is that this judgment is not strong enough to motivate in the epistemic cases. That's because in those cases, contrary to the practical ones, it doesn't seem very illuminating to say that the relevant judgment was rendered weak by virtue of some more powerful element, a "mighty force" that overtakes it (like the appetites). So the parallel between epistemic akrasia of the sort illustrated by the examples and practical akrasia is only clear to a certain extent.

If we reflect on the examples devised throughout the chapter, we see that many of them depict situations everyone is likely to have seen or heard of at least once, and that appear puzzling at first sight. It is difficult to attain Verstehen of those situations, that is, they are difficult to understand. We know that there is something wrong with the scientist that cherry-picks data, as well as with the credulous mom who refuses to acknowledge her son's guilt; but it is not quite clear *what*. The underlying issue is slippery and difficult to grasp, so that the conducts of people in those many cases appear to resist a simple, straightforward explanation.

That's because either you explain them by invoking hidden and overpowering external reasons (such as the after smell, that I suggested when discussing Jealous Husband), or by invoking causes that are not reasons for what they cause (like

prejudice, in Closeted Sexist); or you explain them by suggesting that the person's judgment, the one that conflicts with her actual conduct, is not really as stated in the description of the case, but rather slightly different (maybe the person just changed her mind). But those explanations only get you so far, they're not completely satisfactory. In all of them, there seems to be something missing. What's missing appears to be some reference to the agent's "parcel" in this. That is, some reference to the agent as an autonomous entity that can be held accountable, once the underlying intuition is that he could have done otherwise but (for some reason) didn't.

If the concept of epistemic akrasia is meant to be a tool for Verstehen, that is, if it is meant to help us understand cases of seemingly alien conduct, we need to get past the preliminary characterization given here. That is, we need to get more precise in terms of what the concept actually involves, as just stating that it refers to mismatches between judgment and conduct, or to weakness of some sort, won't suffice. One way of getting more precise is trying to spell the specifics in terms of a set of necessary and sufficient conditions. Some scholars did that, and the conclusion they've reached is that the concept is self-contradictory, and, therefore, useless. That's what the next two chapters deal with.

2. SKEPTICISM OVER EPISTEMIC AKRASIA: PART I

> *Heraclitus says: doctors cut and burn and torture their patients in every way, and then complain that they do not receive the reward they deserve, for doing to one's body the same thing diseases do.*
>
> Hippolytus, **Refutations X**, 10.

Epistemic akrasia is named after the Greek concept of (practical) lack of willpower (*a-krasia*, with *a-* meaning "lack" and *kratos* meaning "power", or "strength"). The discussion about whether or not it is real reaches as far back as the *Protagoras* dialogue (351a-358d), where Plato has his characters find themselves marvelled at the following query: how come certain people oftentimes seem to be deliberately picking what they themselves deem to be the worst of two alternative courses of action? Are people really choosing what they deem to be the worst, or is this rather something else? Socrates' famous response in this dialogue is that it must be something else. "No one," he says, "who either knows or believes that there is another possible course of action, better than the one he is following, will ever continue on his present course" (cf. *Protagoras* 358b–c).

Following more or less the same trend, many contemporary epistemologists deny the soundness of the very idea of epistemic akrasia, and state that in cases that appear to instantiate it, such as the ones presented in Chapter 1, what is going on must be something else. This chapter is about them. Here we'll get more technical and specific, as opposed to the panoramic outlook that has been adopted so far. We'll go deeper into the details of the discussion on specifically epistemic occurrences of the phenomenon – whether they are real or merely apparent, and what the concept of epistemic akrasia needs in order to be of some usefulness to contemporary epistemology.

2.1. Preliminary remarks

The whole point of speaking of *epistemic* occurrences of akrasia is that we find a query similar to the one discussed by Plato arising from epistemic situations as well. After all, it is not altogether uncommon to see people who appear to be holding on to certain convictions that they themselves acknowledge as being the less likely of a set of alternative conclusions to be drawn from the available data. Take, for instance, the anecdotal case of Jeran Campanella, one of the hosts of a popular Flat Earth YouTube channel and one of the most prominent advocates for the flat Earth theory in the U.S. Campanella disproved himself with his own experiment, but then went on with his belief that the Earth is flat nonetheless[12]. Somebody could say, about Campanella, that it is as though he was not strong-willed enough to face the epistemic consequences of the experiment, or to accept the conclusion that imposes itself on the basis of the available evidence. If someone said that, even though it would sound a bit weird, we would understand what motivated he or she to think that. It would make some sense.

The idea that there is an intellectual variant of the akrasia phenomenon became an object of philosophical debate in its own right from Amélie Rorty's article "Akratic Believers" (Rorty 1983) onwards. In this article, Rorty recruited the idea of "akratic believing" with the aim of accounting for what is at stake within those somewhat queer incidents which, like Campanella's, are marked by the appearance that there is a split between "a person's principles and commitments, on the one hand, and his interpretations of the situation in which he finds himself, on the other" (Rorty 1983: 177).

There is a myriad of cases that appear to fall under this description, so the examples brought about as alleged instances of akratic believing are quite varied. Campanella's case involves belief against evidence. Many of Rorty's original examples involve neither belief nor evidence, though, but rather other components of

---

[12] In 2018, Jeran Campanella came up with an experiment designed to be the ultimate proof that the Earth is flat. It involved a high-powered flashlight and three poles of the same height, with holes cut on them also at the same height. The idea was to set up the three poles over a 4 mile width apart, all three 17 feet of the ground, and then activate the flashlight at the first pole, by shining it through the pole's hole. The light was expected to show up on a camera behind the hole of the second and the third poles, indicating that light travelled straight across the surface of the flat Earth. If, however, either the second or the third poles needed to be raised to a different height in order for the light to pass through their holes, it would indicate a curvature in the surface of the Earth, rebutting Campanella's theory. That was exactly how the experiment turned out, but Campanella's belief remained unshaken nevertheless. For more on this, see Daniel Clark's (2018) documentary *Behind the Curve*.

our intellectual lives, such as patterns of inference, perception and habits of speech, which suggest that epistemic akrasia is meant to be an umbrella concept. It might not refer to a single thing, but rather to a cluster of phenomena.

Though the "architecture" of the many examples varies considerably, as we've seen from Chapter 1, they appear to have a common denominator: in all of them, there is visible *tension* between how the agent proceeds intellectually in the situation, that is, the way she reasons, on the one hand; and her own general conception about what would be the right way for a person to reason, on the other hand. How she thinks does not fully match the way she thinks that she should think. This is a type of intellectual shortcoming, a flaw in one's intellectual performance. To have one's intellectual performance spoiled by such flaw is what Rorty called "to be an akratic believer".

So this is ultimately what the notion of akrasia was intended to stand for, when it was brought to the domain of epistemology, or to the studies of theoretical rationality in the early 1980's: to figure in the description of this flaw. In other words, it was meant to help us understand the nature of the problem we witness when we see those people whose intellectual conduct displays the sort of tension mentioned[13].

Nevertheless, the discussions that have been carried out under the epigraph "epistemic akrasia" since Rorty's pioneering article ended up diverting themselves from this initial aim, in a sense. They've distanced themselves from concrete, real-life inspired cases that puzzle us, and became more metaphysical. That's because rather than trying to account for cases of epistemic misconduct, some scholars became more interested in discussing the very concept of akrasia. Those discussions never took the form of a straightforward debate, though, with papers directly tackling and responding to each other. What we have, rather, is a number of scholars independently trying to refine the concept. That is, trying to account for what epistemic akrasia is, or must be, if this concept is to have some epistemological usefulness (and eventually reaching the conclusion that it has none, at the end of the day).

Now, one obvious way to try and refine a concept is to try and determine a set of necessary and sufficient conditions: what the thing must have in order to fall under that concept, and what a thing must not have in order to be excluded from the

---

[13] Besides Rorty (1983, 1997), some of the scholars who engaged in this enterprise include Heil (1984), Mele (1986), Hookway (2001) and Ribeiro (2011).

concept's extension. This is exactly what those scholars tried to do, a task that led them to the observation that epistemic akrasia's conditions are self-contradictory. By virtue of this they stood out for claiming that the whole idea of there being an epistemic or in other ways intellectual form of weakness of the will is a misconception; that the very concept of epistemic akrasia is empty and, therefore, epistemologically useless. Those scholars include Hurley (1989), Pettit & Smith (1996), Owens (2002) and Adler (2002a, 2002b)[14]. These will be my main interlocutors in the present chapter, as well as in the next one.

In what follows I aim at showing that you are only drawn to the conclusion that epistemic akrasia is an impossible thing (and therefore useless) if you model it after the concept of practical akrasia in a narrow way, that is, if you take epistemic akrasia to be the *exact* parallel to practical akrasia. It's not very clear, though, that this is the most productive or the philosophically most interesting way of modelling the concept. Also, I'll argue that even if you work from this narrow, parallelism-based conception, demonstrating that epistemic akrasia is impossible is not an easy task. Some authors thought that they were showing it, whereas in reality they were not.

Before we proceed, a quick word of clarification, about methodology. Someone could object the following: if the phenomenon I intend to talk about (epistemic inefficacy) is not what is usually meant by epistemic akrasia, as I've stated in the introduction, then why is it important to argue against those who have argued that epistemic akrasia, in its usual sense, doesn't exist? That's because of the following. Epistemic inefficacy, like epistemic akrasia, involves the agent undertaking contradictory epistemic attitudes. Due to the fact that the description that is usually given to epistemic akrasia is very general and vague (it basically restricts itself to this: an agent simultaneously undertaking contradictory epistemic attitudes), it wouldn't be difficult for someone to contend that the phenomenon I'm seeking to capture is a variety of epistemic akrasia, at the end of the day. But if epistemic akrasia is altogether impossible, then epistemic inefficacy would be impossible too.

---

[14] Of those, only David Owens is explicitly in the business of denying epistemic akrasia. Susan Hurley is concerned with the parallel between reasons for belief and reasons for action; Phillip Pettit & Michael Smith are concerned with the parallel between freedom of will and freedom of thought; and Jonathan Adler is committed to defending a version of evidentialism according to which one cannot believe against evidence. All five of them, however, are committed to fundamental claims that in one way or another entail epistemic akrasia being conceptually impossible.

So showing that epistemic akrasia is not conceptually impossible (what I'm doing in this Chapter and in the next one) is important precisely because, without this, every discussion about related topics, such as epistemic inefficacy, would always be under the shadow of the skeptical claim that what's being discussed doesn't really exist, or is a misconception. That's why it is part of my project to *show* (rather than just state) that the main arguments presented against epistemic akrasia, arguments purporting that it doesn't exist, do not apply to neighbouring notions, such as epistemic inefficacy, insofar as they don't even apply to all types of epistemic akrasia, which is what they intended. This is, at the end of the day, part of the biggest enterprise of showing that epistemic inefficacy is not the same thing as epistemic akrasia, although the two notions are related by virtue of the fact that both seek to offer an answer to the question "does it make sense to speak of epistemic weakness of willpower?"

This being said, here is the plan for the chapter. In section 2.2., I lay down this parallelism-based conception of epistemic akrasia, highlighting its assumptions and its entailments. In sections 2.3. and 2.4., I discuss two of the main arguments put forward by authors who believe that epistemic akrasia spelled out according to this parallelism-based conception is impossible. It will then become clear that they thought they were showing the impossibility of epistemic akrasia, whereas in reality their arguments do not succeed in the way they expected.

2.2. A parallelism-based (ISO-based) characterization

Talk of "epistemic akrasia" among epistemic akrasia skeptics is informed by a grounding assumption, what I like to call the *presumption of isomorphism* or, simply, ISO. Basically, it is the preconceived idea that epistemic akrasia, if there is such a thing, is a mirror concept to practical akrasia. That it must be something that impairs one's belief system in the *exact* same way that practical akrasia impairs one's practical economy. Or, in other words, that it is the exact same thing as practical akrasia, but affecting theoretical rationality, instead of actions.

a. on being a mirror concept

The presumption of isomorphism underlies a big part of the literature on epistemic akrasia, and it is explicitly stated by authors such as Owens (2002) and Mele (1986). They begin by providing a certain conceptual definition of practical

akrasia as weakness of one's will to act, or action against one's better judgment; and then they proceed to defining epistemic akrasia in the exact same way, except for the terms "action" and "to act", which are then replaced for "belief" and "to believe"[15].

Isomorphism (ISO)
Epistemic akrasia episodes are isomorphic to practical akrasia episodes: they are, to one's beliefs, what practical weakness of the will is to one's actions.[16]

This presumption of isomorphism goes side by side with the idea that what is there to be discussed, or what must be discussed, is whether or not it makes sense to talk of weakness of one's will to form beliefs; or of lack of willpower in theoretical rationality, in general. Weakness of one's will to form beliefs would be the exact parallel to the practical misconduct that we call akrasia, but does it make sense? Many of the initiatives to approach the topic of epistemic akrasia, thus, revolve around this pivotal question, what I sometimes refer to as the *logical question* (borrowing the word "logic" here in its lay sense): the question of whether or not "there is any phenomenon *analogous* to akrasia in the epistemic realm" (Owens, 2002: 381, italics added). "Analogous" is meant as isomorphic, in the sense laid out above: the exact parallel to.

Now, an interesting next step for a philosopher to take from there would be to pick the ISO-based characterization and go look into the world (into our epistemic lives, so to speak), to see whether one finds phenomena that might match that characterization in some relevant way. Finding good candidates for that match would possibly yield a positive answer to the logical question, whereas not finding any would yield a negative answer.

---

[15] Owens, for instance, defines practical akrasia as being "an agent's ability to do, freely and deliberately, something that he judges he ought not to do"; and then proceeds to defining epistemic akrasia as "a phenomenon analogous to akrasia in the epistemic realm" (…), which requires that "(a) a person's (first-order) beliefs can diverge from his higher-order judgements about what it would be reasonable for him to believe and (b) these divergent (first-order) beliefs are freely and deliberately formed" (Owens 2002: 381). Mele, in turn, defines practical akrasia as incontinent action. To act incontinently is to "lack (…) the power to act as one judges best, or to resist the temptation to act otherwise"; and then states that what he is after in his paper is "an account of a comparable, full-blown, variety of incontinent believing" (Mele 1986: 212).

[16] As Snellen (2018: §2.1) rightly emphasizes, talk of practical akrasia within contemporary philosophy is mostly made from an action theory standpoint, that is, scholars approach practical akrasia in the form of a single and isolated episode of acting against one's better judgment. That's why I've made ISO about *episodes*. In the next chapter I'll call this aspect into question, but for now let's stick to the episodic characterization.

However, instead of taking this pragmatic approach, some scholars would rather try to answer the logical question by a different means: by showing that there *could not* be any such matches. In other words, they concentrate on showing that the *concept* of epistemic akrasia itself has problems of the sort that ultimately turn it into an empty concept, or a philosophical fiction. It is as though the conceptual characterization of epistemic akrasia from which they're working is like that of a "square ball": it describes a metaphysically impossible thing, so that looking for concrete correspondences is of no use. Their work then comes down to showing *that* the concept of epistemic akrasia is like that of a square ball.

I believe those scholars got some of it right, and some of it wrong. They are right in that there is no phenomena, in the real world, that are the epistemic *isomorphic* to practical akrasia, in the sense sketched. There isn't, and there in fact could not be, something out there that is capable of impairing one's belief system in the *exact* same ways that practical akrasia impairs one's practical economy. So the exact parallel to practical akrasia within the domain of theoretical rationality doesn't exist. In this, the abovementioned scholars are right, I suppose.

But they are wrong in terms of what the real problem is. The whole point of the logical question, of discussing things like epistemic akrasia, is to account for the (or some of the possible) ways in which people fail epistemically. Disclaiming epistemic akrasia in the way these scholars do is not of much aid, when it comes to this philosophical task. That's because by simply verifying that it does not exist, we are still left without an account of what does exist, or of what is actually going on when we see a person being taken over by what appears to be some intellectual form of weakness of willpower.

To break this down, lets analyse the ISO-based characterization of epistemic akrasia, making its entailments explicit. It goes as follows. Practical akrasia is the conduct displayed by an individual while taking or being inclined to take a course of action that is admittedly at odds with her better judgment, in contexts in which it would have been possible for her to do it differently. Therefore, epistemic akrasia has to be, or is supposed to be, the conduct displayed by an individual in having or maintaining epistemic attitudes (for instance, forming beliefs) that are admittedly at odds with her better intellectual judgment (i.e., her judgment as to what she should believe), in contexts where it would have been possible for her to do it differently.

Strictly speaking, for one's practical conduct to be correctly captured by the concept of practical akrasia, the person must acknowledge, implicitly or explicitly, that the action she performs or is inclined to perform is at odds with her better practical judgment, i.e., her judgment as to what would be the best thing for her to do at the time. That is, it does not suffice that there be dissonance between what one does or is inclined to do and her better judgment, neither it suffices that there be dissonance between the action in question and her other inclinations, or even other actions – to talk of akrasia is to talk of cases in which one acknowledges the dissonance as such. This is how the notion is defined in contemporary Anglo-American philosophy, by authors such as Davidson (1980: 22-23); and this is also how it has been defined classically, by Aristotle, for example[17]. Call this the "acknowledgment condition" for practical akrasia.

Acknowledgement (ACK)
The action involved in a practical akrasia episode is acknowledged by the individual herself as dissenting from her better practical judgment.

Concomitantly, only cases in which things are such that it would have been possible for the action in question to not be or not have been performed can be instances of practical akrasia. If one feels she is about to pass out in the presence of a rabid dog whilst her better practical judgment tells her that the best thing to do is to run, her behaviour is not akratic, insofar as passing out, though blatantly (and admittedly) at odds with her better judgment, is not such that she could have done otherwise. Call this the "control condition" for practical akrasia.

Control (CTRL)
Things are such that it was possible for the action involved in the practical akrasia episode to not have been performed.

Making these conditions explicit matters for the sake of delimitation. The first one, the acknowledgment condition, renders akrasia distinguishable from ordinary mistaking (i.e., the performing of an inappropriate action which is not so acknowledged by the agent whose action it is, or which is only so acknowledged at a later time). It also prevents cases of hypocrisy from being described as akratic (i.e., cases in which one acts contrary to what she appears to deem best, but doesn't).

---

[17] Although Aristotle was ultimately an akrasia skeptic himself: he denied that it is indeed possible for the subject we typically refer to as akratic to act intentionally, voluntarily and in full knowledge of what she is doing.

The second one, the control condition, in turn, prevents the behaviour exhibited by thermometers, smoke detectors and non-rational entities from being described as akratic.

In short, if we are entitled to state, about a particular episode α, that α meets these two conditions, then α is an instance of practical akrasia. At the same time, if we claim, regarding a particular episode α, that α is a case of practical akrasia, we are committing ourselves to the claim that α meets both of the conditions, insofar as this is implicit in the very use of the concept. This is just another way of saying that the acknowledgment condition and the control condition are traditionally viewed as the two necessary conditions that an episode has to meet in order to be an instance of practical akrasia; and together they yield the sufficient condition for practical akrasia.

As I've been saying, authors such as Pettit & Smith (1996), Adler (2002a), Hurley (1989) and Owens (2002) endorse certain epistemological theories according to which an epistemic analogue of practical akrasia is impossible. Their argumentative path stems from the following considerations: if it indeed exists, within the epistemic domain, a phenomenon that is isomorphic to practical akrasia, for a person's epistemic conduct to be correctly described as one of its instances, first, she must implicitly or explicitly acknowledge that a certain belief of hers is at odds with what she thinks would be the right belief to be acquired or held. Secondly, concomitantly to this, things must be such that it would have been possible for her to not acquire or hold such belief. In sum, there must be *epistemic equivalents* of the acknowledgment condition as well as of the control condition; and an episode of epistemic akrasia must meet both.

In addition, due to the presupposition of isomorphism between the practical and the epistemic varieties of the phenomenon, cases involving unconscious beliefs cannot, prima facie, be meaningfully described as cases of epistemic akrasia, even if the content of those beliefs conflicts with what one thinks she should believe. That's because the fact that they are unconscious keeps the individual from acknowledging the dissonance. Likewise, variations of cases in which one has a belief inoculated into her doxastic mesh by an evil genius also cannot be meaningfully described as instantiating the phenomenon, as these are emblematic cases of which we are willing to say that one was completely oblivious to the acquisition of the beliefs in question and, therefore, was not exercising any form of control.

In other words, if there is an epistemic phenomenon that is isomorphic to practical akrasia, for one's behaviour to be correctly described as instantiating this phenomenon it has to be true that such an individual "appreciates her perversity" (Heil 1984: 67). That is, the person has to acknowledge that she is acquiring or holding a belief that is at odds with her better epistemic judgment; and she has to have been capable of not acquiring or holding such a belief. Thus, if we are willing to state, as to a particular episode β, that β is an episode of epistemic akrasia, we commit ourselves to β meeting the two abovementioned isomorphic conditions: the epistemic acknowledgment condition and the epistemic control condition. The heart of the matter, then, turns out to be whether there could be any β episode that simultaneously meets the two isomorphic conditions (logical question).

By rewriting the considerations presented so far in a more technical fashion, we have

Acknowledgment (ACK)
The action involved in a practical akrasia episode is acknowledged by the individual herself as dissenting from her better practical judgment.

and

Control (CTRL)
Things are such that it was possible for the action involved in the practical akrasia episode to not have been performed.

Assuming that

Isomorphism (ISO)
Epistemic akrasia episodes are isomorphic to practical akrasia episodes,

we thus have

Epistemic Acknowledgment Condition ("C1"), from ACK and ISO
The belief involved in an epistemic akrasia episode is acknowledged by the individual herself as dissenting from her better epistemic judgment.

and

Epistemic Control Condition ("C2"), from CTRL and ISO
Things are such that it was possible for the belief involved in the epistemic akrasia episode to not have been acquired or held.

The fundamental concern of the debate around the possibility of epistemic akrasia is whether or not there is a way for C1 and C2 to be simultaneously met. As I've pinpointed, a number of authors advanced arguments purporting that there isn't.

To be sure, two outcomes are possible here. Either the arguments intending to show that C1 and C2 cannot be met fail, or these arguments succeed. If they fail, it follows from this is that epistemic akrasia is not a conceptual impossibility; but no particular consequences, or consequences for particular cases that were thought to be of epistemic akrasia, follow. The question then becomes: is this concept adequate to discuss those cases? Is it epistemologically valuable? If, on the other hand, those arguments succeed, if they prove that either C1 or C2 or both cannot be met, then it follows that epistemic akrasia doesn't really exist, because it's conceptually impossible. Or, at least, it doesn't exist in the way it has been characterized, the ISO-based way. (A question that is raised, in this case, refers back to ISO: is ISO an adequate assumption to work from? Alternatively, we might ask: if there is no epistemic akrasia, why, then, do we have the impression that people behave in ways that are intellectually akratic?)

In this chapter I'll focus on arguments tackling C1. Before we proceed, let's get clearer on the main entailment of the ISO-based characterization that gave rise to those arguments.

b. theoretic evaluative judgments

As the notion of epistemic akrasia has been spelled out, for C1 to be met, the agent must have a better epistemic judgment. That is, ISO must be interpreted as entailing that agents have theoretical evaluative judgments, as much as they have practical evaluative judgments. Specifically, agents must have theoretical evaluative judgments in general, and a specific variety of such judgments, in particular: the better epistemic judgments, aka judgments *all-things-considered* as to what should be believed, or what conclusion should be drawn.

Now, there is more than one way of understanding the nature of evaluative judgments, of which "better judgments" are specimens. Traditionally, evaluative judgments (such as the judgment that certain course of action is better than another)

are thought to enjoy a special status relatively to other types of judgment, such as perceptual judgments (for example, that A is yellow, or that A is more yellow than B) as well as descriptive judgments (for example, that A is expensive, or that A is more expensive than B). What evaluative judgments have that other kinds of judgment don't is the property of being action-guiding: of purporting an answer to the practical question "what should I do?". They provide one with a motivation to act.

As the idea of epistemic akrasia is spelled out by the ISO-based characterization laid out above, C1 requires evaluative judgments that are epistemic isomorphic to practical evaluative judgments. That is, it requires evaluative judgments that have, as their object, not practical attitudes, such as actions, but rather epistemic attitudes, such as beliefs. Specifically, there must be theoretical evaluative judgments that are *belief-guiding* (that serve as a guide or motivation to belief acquisition); and among those, the special kind of evaluative judgment, the so called "better judgment". In other words, there must be a special kind of theoretical evaluative judgment that specifies something of the sort "under such and such circumstances, *m* is what one is supposed to believe, all things considered".

In more technical terms, the interpretation that must be given to C1 within an ISO-based characterization of epistemic akasia is one in which it entails theoretical evaluative judgments, TEJ:

> **Theoretical evaluative judgments (TEJ)**
> There are theoretical, or doxastic, evaluative judgments that play, in theoretical reason, the same role played by practical evaluative judgments in practical reason.

That said, the most straightforward way of arriving at the conclusion that epistemic akrasia is impossible is to reject the very idea of theoretical evaluative judgments as it stands, TEJ. It's to say: for there to be such a thing as epistemic akrasia C1 has to be met. C1 entails TEJ, but TEJ is false. So, if TEJ is false, then there is no epistemic akrasia.

The reason why TEJ looks false is that if practical evaluative judgments are action-guiding by definition, strictly speaking, the idea that there are *theoretical* evaluative judgments, TEJ, sounds quaint. This is because we do not ask ourselves questions such as "what am I supposed to believe?", at least not ordinarily, nor in the

same spirit as when we ask the practical question "what am I supposed to do?"[18]. This is because beliefs are not the kind of thing that is acquired from, or by, the following up of recommendations. Beliefs seem to be, rather, the kind of thing that is "naturally" given out of our accessing of evidence. Say q is evidence for p once q, being true, renders p true or increases the probability that p be true – we acquire the belief that p (or a degree x of belief that p) directly from our appreciation that q is true, no "recommendations" being involved in the process.

When we ask ourselves questions of the form "what am I supposed...?" we're actually seeking recommendations. What evaluative judgments do is to purport recommendations, and this is precisely what makes them the kind of judgment that they are. Purporting as well as following recommendations presupposes that we take the very thing that is being recommended as something we could either follow or refrain from following – there are no recommendations on what we take as following by necessity, or invariably, independently of our will. There are no recommendations on, for instance, whether or not we're supposed to see infrared, or agree that every bachelor is unmarried, for these things hold completely independent of our endorsements.

If practical evaluative judgments provide us with recommendations, and if there are only recommendations on what we can endorse, then the idea that there are theoretical evaluative judgments playing in theoretical reasoning a role similar to the one played by practical evaluative judgments in practical reasoning (TEJ) sounds difficult to understand – there might be a vague similarity between these two roles, but not a real analogy, one that suffices for isomorphism. And if this is so, then the one thing one's belief would have to be at odds with for C1 to be met simply doesn't exist.

Arguments claiming that C1 cannot be met on account that TEJ is false, more or less within the contours just outlined, can take two forms. One can deny TEJ simpliciter, i.e., by stating that there are no theoretical or doxastic evaluative judgments altogether (call it strategy I). Or one can deny TEJ with qualification, by claiming that there are theoretical or doxastic evaluative judgments, but that they do not play in theoretical reason the *same* role played by practical evaluative judgments

---

[18] An exception would be the situations in which one is dealing with conflicting testimony. Under circumstances of this sort, in which we are puzzled about whose narrative to believe, we ask ourselves "what am I supposed to believe?" more or less with the same spirit and the same expectations as when we are confused about what to do.

in practical reason (strategy II). In the next two sections I'll present, discuss and offer objections to two of those arguments.

2.3. Pettit & Smith's argument

A specimen of argument claiming that C1 cannot be met on account that TEJ is false has been put forward by Pettit & Smith (Pettit & Smith 1996: 448):

> Imagine that your beliefs run counter to what evidence and fact require. In such a case, your beliefs will not allow those requirements to remain visible because the offending beliefs themselves give you your sense of what is and your sense of what appears to be. You are therefore denied an experience whose content is that you are believing such and such in defiance of the requirements of fact and evidence. This is why, as G. E. Moore observed, you cannot simultaneously think that while you believe that p, yet it is not the case that p.

It is not quite clear from Pettit & Smith's excerpt which of the two strategies, I or II, was intended. The passage is interpretable as stating that the reason C1 cannot be met is that beliefs are acquired not by means of a "better epistemic judgment", but rather from our direct apprehension of what facts and evidence require, which is a variation of strategy I. Or it can be interpreted as stating that the reason C1 cannot be met is that beliefs are acquired on the basis of better epistemic judgments, but that these, in turn, *correspond* to our direct apprehension of what facts and evidence require, and once this kind of judgment is not belief-guiding, TEJ does not hold, which, in turn, is a variation of strategy II. Or the excerpt can yet be interpreted as stating that TEJ is false because better practical judgments can be countered, while theoretical ones cannot (countering a better theoretical judgment requires incurring in a paradoxical state of mind, one that can only be expressed by a Moorean sentence; and is, therefore, an impossible state of mind), which is another variation of strategy II.

From this later available interpretation the argument is not a very promising one and it has already faced objections from David Owens (2002: 382-383). According to Owens, there is no reason to think that countering a better theoretical judgment couldn't come out as a much less "disturbing" and non-paradoxical

sentence, rather than a Moorean one. Pettit & Smith's case against C1 sounds more promising when spelled out from the other two interpretations.

According to the other variation of strategy II, these authors are to be taken as stating something along these lines: I'm supposed to have the beliefs that facts and evidence support. And the beliefs I have are given by what I take as being what I'm supposed to believe, that is, by my better epistemic judgment. And my better epistemic judgment, in turn, is given by my apprehension of facts and evidence. So if I acquire a belief *m* that goes against what facts and evidence actually support, this will have happened because I've misapprehended facts and the evidence (after all, had I apprehended them correctly, my belief *m* and them would not be dissonant, but rather consonant). But if I've misapprehended facts and evidence, and if it is from this misapprehension that I come to acquire *m*, then I simply cannot realize that *m* is at odds with what the facts and evidence actually support. From my own viewpoint, *m* is perfectly consonant to what the facts and the evidence support, for I think that what they support simply is what I took them as supporting. Thus, while it is possible for an individual to believe something that is at odds with what the facts and evidence actually support, it is not possible for her to believe something that is at odds with what she *herself* takes as being what the facts and evidence support. Since beliefs are formed by virtue of our own apprehension of facts and evidence, and of what we take them as supporting (better epistemic judgment), it is impossible for anyone to have a belief admittedly at odds with what she herself takes to be what she is supposed to believe.

Within this variant, the argument states that beliefs are acquired on the basis of a better epistemic judgment, but that this, in turn, does not play the role of guiding belief acquisition. That is, better epistemic judgments are not belief-guiding. Their role is best described as *belief-constraining*: they constrain (i.e., they determine, rather than recommend, or motivate) belief acquisition. And they do this by restricting belief content to the realm of what facts and evidence were taken as requiring. Therefore, these judgments are not isomorphic, in the epistemic field, to better practical judgments (TEJ is false).

There are two main problems with this version of the argument. First, it looks like in order for this argument to work, belief acquisition must be a sort of causal, or "mechanic" process, that is, a process by means of which one's better epistemic judgment literally *produces* belief (rather than one by means of which this sort of

judgment rationally motivates belief acquisition). The contrast here is between conceiving the belief-forming process as a process based on a sort of "given", on the one hand, and conceiving it as a process based on a rational coupling between epistemic (rational) faculties and the world, on the other hand. The former process results in beliefs for which the agent ultimately cannot be held accountable, whereas the later allows for the idea of epistemic accountability, or epistemic agency.

It looks as though this version of Pettit & Smith's argument will only work if the process in question is of the former type, the mechanic process, in which once the agent has a better epistemic judgment according to which **m** she invariably ends up with the belief that **m**. But this is a somewhat unwelcome entailment, because we ordinarily talk about beliefs using a normative vocabulary: we blame people for holding beliefs that seem to be undue, for not being able to explain (i.e., to present the reasons for) why they hold such and such beliefs when they seem to be undue, and we also blame them for not holding such and such beliefs when they seem to be entirely due.

The second problem with this version of the argument is that it depends on this belief-constraining process (the process by means of which one's better epistemic judgments produce beliefs) being infallible[19]. If an agent's better epistemic judgment *always* comes out as the acquisition, by the agent, of the correspondent belief, then it in fact looks as though there's no way the beliefs acquired by an agent could ever be at odds with her better epistemic judgment, for the content of her beliefs will always be tacitly given in accordance with what such a judgment specifies.

If, however, one's better epistemic judgment can fail in its task of generating the correspondent belief, an eventual flaw in the process may result in a state of affairs in which the acquisition of a belief whose content is specified by one's better judgment simply does not occur. I'm talking about cases in which one already holds a belief **not-m** and, at some point, her better epistemic judgment indicates **m**.

─────────────────────

[19] By saying that this process must be *infallible* in order for Pettit & Smith's argument to work I'm not evoking the classical notion of epistemic fallibility, which is basically the fact that we can believe justifiably (i.e., rationally) and yet falsely. What I mean here is simply that for Pettit & Smith to succeed in their intent of showing that theoretical reason is not prone to configure the state of affairs characteristic of C1 the process by means of which one's better theoretical or epistemic judgment regarding a certain matter results in the incorporation of a belief on that particular matter into the agent's doxastic mesh needed to be immune to flaws, for otherwise (if this process, for whatever reason, does not translate into the acquisition of the relevant belief) the agent will end up precisely in the state of affairs characteristic of C1 in case she has a precedent belief on that very matter.

In this situation, the person must acquire the belief that *m* and disbelieve **not-m**, in order to preclude the state of dissonance that would otherwise configure C1. However, it looks as though the constraining force exerted by this judgment failing (so that *m* ends up not being incorporated to one's doxastic mesh, thus leading to **not-m** not being replaced) is a live possibility. In such a case, one ends up in a situation where she has a belief **not-m** at the same time as her better epistemic judgment tells her to believe *m*, which is exactly the state of dissonance described by C1.

If such cases are possible, we should concede to Pettit & Smith that the isomorphism between epistemic and practical better judgments may not be complete (if the latter are action-guiding whereas the former are belief-constraining, as opposed to belief-guiding). But, yet, these two types of judgment are isomorphic in the relevant sense that they may or may not succeed in their respective tasks of informing actions and beliefs, which restores TEJ. The question with which Pettit & Smith's argument needs to be confronted, then, is the question of whether cases like the abovementioned one are possible.

At first, one might be inclined to say they are not possible. It simply is not possible for one's better epistemic judgment to fail in this way, that is, in a way that the belief it calls for, *m*, ends up not being incorporated to one's doxastic mesh, and the previously held belief **not-m** ends up not being replaced.

The reason we are inclined to saying that this is not possible is that we are too used to taking basic perceptual beliefs – such as the belief that the wall on my left is white, for example – as the default specimen of the belief category. In fact, hypothetical cases like the one we've been discussing (i.e., cases in which one's better epistemic judgment fails to translate into the actual acquisition of belief) are hard to conceive if we think of them as involving perceptual beliefs. But not if we think of them as involving a myriad of other types of belief or belief-like attitudes.

This is because perceptual beliefs are straightforwardly given out of what we see, feel and hear. For one thing, if I see that the wall on my left looks white and if I judge all things considered that it looking white under adequate light conditions is a sufficient reason for me to believe that it is white (*m*), then even if I have been believing until then that the wall was beige (**not-m**), it is as though I naturally acquire the belief that *m* and exonerate **not-m**; and the idea that things could turn out differently sounds pretty counter-intuitive. So it is reasonable to say that, in the case

of perceptual beliefs, forming a better judgment naturally comes out as the acquisition of the corresponding belief.

But now take a case involving non-perceptual beliefs, such as this one, presented by Mele (1986: 218-219):

WILMA
Suppose, e.g., that Wilma judges not only that the evidence that her twelve year old son, Basil, has been experimenting with narcotics is very strong and much stronger than the contrary evidence, but also that her believing that he has been doing so is much better supported by nonepistemic considerations than is her not believing this. She thinks that even if the epistemic evidence were significantly weaker it would be best, practically speaking, for her to believe that Basil has been taking drugs, for she fears that in the absence of this belief it would be very easy for her to fail to give him guidance that he may well need. To be sure, Wilma wants it to be false that her beloved Basil has been using drugs and she recognizes that believing that he has been doing so would be very painful for her, but she judges that the pain is outweighed by other practical considerations. Undoubtedly, many more details can be added to the story of Wilma's finally coming, on the basis of epistemic and non-epistemic considerations, to hold the judgment that there is good and sufficient reason for her not believing that Basil is innocent of drug use. However, let us suppose now, for the sake of brevity, that she does come to hold this judgment but that she nevertheless believes that he has not been using drugs.

The Wilma case was entirely settled by Mele on the basis of assumptions ("suppose Wilma judges that..." and "...suppose that she nevertheless believes…"), so it might look stipulative, at first sight. Notwithstanding, the case resonates quite a lot with our mundane experience (who never met someone like Wilma, at least with respect to some particular matter of concern?). And I believe we all share the intuition, from which the example derives its strength, that cases alike are not only possible but also a commonplace. Such cases happen in the real world.

Wilma ends up with her better judgment telling her *m*, that Basil is experimenting with drugs, while preserving (that is, while still holding onto) the belief **not-*m*** that he is not so experimenting. One good and simple explanation for what is at stake seems to be that her better judgment is failing at effectively translating into (either by constraining, motivating, or by whatever means) the acquisition of *m*. And, because it is thus failing, Wilma's final situation is precisely one in which the state of dissonance characteristic of C1 prevails.

Someone could object that if Wilma did not actually acquire the belief *m*, it is because she did not effectively judge *all things considered* that she should acquire *m* at that time, for, had she so judged, she *would have* acquired *m*. This objection

sounds a little exotic, once the case, as it is spelled out by Mele, already specifies that Wilma's judgment that she'd better believe that Basil is experimenting with drugs *is* her judgment all things considered – it was formed on the basis of epistemic and non-epistemic considerations, that is, it took into account "all things"[20].

Such an objection strikes out the hypothetical character of the example by suggesting that what Mele stipulated as being the case is actually not possible, though the intuition we drive out of our ordinary experience, it seems, is on Mele's side. This objector could claim that better judgments are of such nature that one's acquired beliefs *express* them (and not that they play a role in belief acquisition, either by constraining or by guiding it). So that if Wilma's judgment that she would be better believe that *m* doesn't translate into the actual acquisition of the *m* belief, then this simply wasn't her better judgment on the matter at that time, but rather an ordinary (all-out) one[21].

Now, this has the potential to actually render it impossible for one's beliefs to be at odds with one's better epistemic judgment: it establishes a sort of impediment on principle, in that it fixes acquired beliefs as being an expression of one's better judgments and, therefore, as inexorably consonant to them. If this is right, then the Wilma case would have been spuriously spelled out by Mele – it would have been spelled out in a biased, partial manner.

But in this case scenario the objector is disagreeing about the very nature of evaluative judgments – he is taking evaluative judgments, as states of mind that play

---

[20] Jonathan Adler raised the point that cases such as Wilma's merely appear to be, but are not, cases of epistemic akrasia because the consideration that led the agent into the acquisition of the belief that is at odds with evidence is not included in the evaluative basis of judgment, that is, it is not an epistemic reason. "Any normal [believer] agent," says Adler, "needs to recognize, and be competent to apply, a distinction between epistemic and non-epistemic reasons, since non-epistemic reasons are constantly present for many of our beliefs. Its pull must therefore be regularly resisted (…)" (Adler 2002a: 12). Thus, Adler believes that someone like Wilma is merely behaving as a non-normal epistemic agent, but not incurring in epistemic akrasia. I find this interpretation non-illuminating, insofar as the problem posed by epistemic akrasia is precisely that the epistemically akratic agent doesn't behave in the way expected from a rational epistemic agent. So it is as if Adler is saying that cases like Wilma are not cases of akrasia because they're cases of akrasia, instead. Moreover – apparently Adler did not consider this – it is not one of the requirements for epistemic akrasia that the belief contrary to one's better epistemic judgment be acquired *only* from epistemic reasons; just as it is not one of the requirements for practical akrasia that the action contrary to one's better practical judgment be undertaken only from reasons that are non-epistemic. Prima facie, any belief that one has or holds that dissents from what she takes to be what she should believe is eligible to frame the situation as an epistemic akrasia episode, as well as any action that one takes which dissents from what she takes as what she should do can frame the situation as a practical akrasia episode.

[21] This would be a "detached judgment", or an "all-out judgment", in Davidson's terminology (1986, 1999). A detached judgment is a judgment "detached from its all-things-considered evaluative grounding", cf. Adler 2002a: 3.

a role (of whatever kind) in belief acquisition to be non-existent. This is precisely the first interpretation available to Pettit & Smith's argument against C1 charging TEJ to which I've alluded, the one that denies that evaluative judgments understood as judgments that play a role in belief acquisition (TEJ) exist.

A similar position with respect to practical evaluative judgments was held by moral philosopher R. M. Hare (1952, 1963). Hare advanced a general theory on the nature of evaluative judgments according to which one's better (practical) judgment is revealed by one's actions. To Hare, from the action actually taken by a person it follows that she judged that action to be the most appropriate, all things considered, at the time (and not the other way around). Thus, for him, "it is analytical to say that one always does what she thinks best all things considered" (Hare 1952: 169)[22].

One could hold, as to our epistemic lives, a position analogous to the one Hare holds regarding our practical-moral lives; and this may have been what Pettit & Smith intended, if they are to be interpreted according to strategy I. This would be equivalent to saying the following: one's better epistemic judgment is revealed by one's beliefs. So, from what a person actually believes at a time, it follows that she judged that belief to be the most appropriate, or the one better supported by evidence, at that time. For one who holds such a view, theoretical evaluative judgments have nothing to do with belief acquisition; they are but the formal mirror to one's beliefs.

The problem with the idea that evaluative judgments play no role within belief acquisition is that it doesn't resonate with the way we ordinarily think and talk about our own states of mind. In conducting ourselves within society, we often take it for granted that the beliefs people express at a given time are at least in part a function of what they deemed reasonable to believe at that time, not the other way around. That is, we reason from the assumption that what a person believes to be the case is (at least in part) a consequence of his or her judging that that was the most reasonable conclusion to be drawn, or the best explanation, or the most meaningful account, and so forth.

For one thing, we live our epistemic lives with the background assumption that it is on us, for the most part, to exercise our critical skills when dealing with

---

[22] Arguably, anyone who defends such a view is committed to there not being any practical akrasia of the sorts described by the set of conditions we've discussed, due to the impossibility of meeting the acknowledgment condition. Hare, for instance, argues that what happens in episodes that appear to be of practical akrasia is that the individual gives in to strong emotions.

information and deliberate as to what each situation requires, epistemically. And many of the beliefs we come to hold *derive from* (rather than simply mirror) such analyses and deliberations, at least in part, and at least in some of the cases. We, for instance, decide which inquiries are worth pursuing and which are not; we deliberate as to whether we should close a present inquiry or go after more data; we weight evidence, by comparing what goes for certain hypotheses and what goes against them *before* we close our minds upon a particular conclusion. All of these measures make a difference in terms of which beliefs one ends up holding. What is more, we even rebuke others when we sense that they are refraining from taking these measures.

That is to say, we treat one another, as much as we treat ourselves, as belief-forming *agents*, rather than as opinion-issuing organisms. We might want to let go of this fundamental assumption, that at least in some cases we exercise judgment before we reach conclusions and form beliefs, but only if we're willing to give up on seeing ourselves as sensible entities. Not being willing to bear this cost indicates that adopting a position analogous to Hare's within the epistemic domain will not reflect all that there is to the nature of epistemic evaluative judgments.

2.4. Adler's argument

Another argument claiming that C1 cannot be met on account that TEJ is false was presented by Jonathan Adler (2002a, 2002b). Take the following two excerpts:

> Theoretical reasoning is likened to practical reasoning because the former seeks to determine which hypothesis is best supported by the available evidence. The analogy [the alleged analogy between practical and theoretical reasoning] implies that, just as the objective of practical reasoning is to discover which option is best all things considered, the objective of theoretical reasoning is to discover which hypothesis is best supported on the total available evidence. It is this implication that is misleading. For the objective of theoretical or empirical reasoning is determining whether a hypothesis is true, not whether it is best supported. (Adler 2002b: 69)

> If evidence is adequate to accept a hypothesis, no (previously) conflicting evidence can retain its (epistemic) force against that hypothesis. So no evidence remains that can motivate belief against the hypothesis justified by one's evidence, as the desires disfavored in an akratic judgment [...] still retain a pull on the agent, and so can motivate him to act against it. So weakness of belief – as believing in opposition to one's evidence – is motivationally unintelligible. (Adler 2002a: 8)

Adler's argument is a variant of strategy II. He acknowledges that there are theoretical evaluative judgments and that they do play a role within belief acquisition, but he denies the isomorphism between them and the practical variety. That is not exactly because the kind of role played by each of them in their respective domains is different, but rather because the very structures of theoretical and practical reasoning, according to Adler, differ significantly: they operate according to different norms, or rules, and have different purposes, so that there is no parallel between them.

Adler's core idea is that the norms of theoretical reason are fundamentally different from the norms of practical reason, since the alternatives entertained in a theoretical reasoning are mutually exclusive in a way that is different from the way in which alternatives are mutually exclusive when entertained in a practical reasoning[23]. When two concurrent possibilities of action, *A* and *A'*, are placed before a subject so that she feels inclined to both but can only undertake one, accessing motivational elements that reinforce her inclination to *A* does not weaken her inclination to *A'*, and vice-versa.

Thus, if one comes to the conclusion that it would be better all things considered to pursue a career in engineering and settle in a large urban center (*A*), her inclination to join the hippie movement and go live a nomadic or rural lifestyle (*A'*) does not just vanish. Even if she does undertake *A*, her inclination to *A'* remains, being experienced in the form of "a kind of regret" (Adler 2002a: 6). This is because although *A* was deemed the best life choice all-things-considered, *A'* was judged better all-out, that is, with respect to some (partial) parameters – better from the specific viewpoint of being in touch with nature, for example – and it continues to be so deemed even in the face of the all-things-considered conclusion that states otherwise. Thus, in a case which one undertakes *A'* contrary to her better judgment according to which she is supposed to do *A* (an archetypal case of practical akrasia)

---

[23] Virtually the same idea is presented by Hurley (1989), who argues that whenever there's epistemic conflict between reasons to believe – for instance, when we access evidence *e* favouring **not-p** and evidence *f* favouring *p*, we find out whether *p* is more likely than **not-p** all-things-considered by figuring out what is more likely relatively to the conjunction *e* & *f*. If, by means of this demeanour, we find out that *p* is more likely than **not-p**, there is no remaining reason to believe **not-p**, even if **not-p** is more likely than *p* relatively to *e* (contrary to what goes on within practical conflict). This argument is endorsed by Raz (2007, p.6) to the effect of denying the very possibility of epistemic akrasia. Other authors that see the matter similarly include Dretske (1971), Harman (1980) and Owens (2000).

the undertaking of **A'** can only occur because the motivational elements that inclined her to **A'** haven't disappeared after the better judgment has been established.

The same does not hold for theoretical reasoning, according to Adler (2002b: 263). When two concurrent claims **P** and **P'** are placed before someone, the set of reasons **f** that she deems good and sufficient for **P** automatically defeat[24] the reasons **f'** in favor of **P'**, since **P** and **P'** cannot be both true at the same time. Thus,

> The fundamental disanalogy is that the goal of theoretical reasoning is all-out or full belief, and so the (threshold) acceptance of a proposition. When theoretical reasoning reaches that goal, contrary or undermining evidence is nullified. So there is no evidence to play the role of conflicting desires in drawing the agent away from his better judgment. (Adler 2002a: 18).

Essentially, therefore, when good and sufficient reasons favour certain belief, "the disfavoured belief evaporates, since it has been determined to be false" (Adler 2002a: 6). This entails that if reasons **f** are deemed good and sufficient to warrant (for example) the truth of **P**, then one's all-out judgment (**f** ∴ **P**) is the ultimate goal of theoretical reasoning. There is not – as there is in practical reasoning – an "all-things-considered" judgment from which the "all-out" judgment could dissent. Then, when a person judges all-out that **f** ∴ **P**, according to Adler, she automatically *i*) acquires the belief that **P**; and *ii*) deems any rival hypothesis **P'** false (insofar as **f** defeats **f'** which would otherwise serve as reason for the acquisition of **P'**) and immediately abandons it (Adler 2002a: 6-7).

Apart from the question of whether the process (of whatever sort it may be) involved in (*i*) can fail, which was already addressed when we discussed the Wilma case, if Adler is right that (*ii*) is a consequence of the all-out judgment that certain reasons are good and sufficient to support a given belief, then a person can't ever acquire a belief that is admittedly at odds with what she herself takes as being what she is supposed to believe (her better epistemic judgment). That is to say, the situation described by C1 cannot occur, simply because there are no better judgments at stake in theoretical reasoning – forming the all-out judgment that reasons **f** are good and sufficient for **P** (the all-out judgment **f** ∴**P**) is all that it takes for one to acquire the belief that **P** as well as for the abandonment of any rival

---

[24] Adler himself does not develop the idea of "defeater", limiting himself to highlighting that it amounts to "an intuitive and unavoidable notion" (Adler 2002b: 328). He mentions, as a reference, the work of Klein (1981).

hypotheses **P'**. In Adler's own words, "to judge all-out is to satisfy the conditions for believing that" (Adler 2002b: 61).

Two problems with respect to this argument become salient. The first one concerns Adler's idea that in practical reasoning one's inclination to the deferred option "survives" being experienced in the form of "a kind of regret", whereas in theoretical reasoning this simply does not happen. This doesn't look correct, because oftentimes, in theoretical reasoning, something akin to this "survival" (of the dismissed option) does happen, although it is not experimented as a *feeling* of regret.

We sometimes experience the survival of our inclination to accept **P'** even after the endorsement of **P** (where **P** and **P'** are rival hypothesis) in the form of a kind of recalcitrant and subtle doubt, or a lack of certainty, that is felt as a call for verification. It is quite likely that we all have been through situations in which a residual uncertainty remained underpinning an inference and manifested itself, for instance, in an urge for double-checking the inference's steps, juxtaposing each one with the impression that alternative hypothesis put away moments ago had at some point appeared to be good. I believe we've all been through this while, for instance, taking a difficult multiple-choice math test.

Another way we might experience temporary "survival" of our inclination to accept a rival hypothesis is through a feeling of wonder, or astonishment, in the presence of information that we know is most likely true than not true, but which breaks expectations. As a simplified low level example of this, consider the amazement of a lay person upon first hearing from an expert, for instance, that Cleopatra was not Egyptian (**P**). She accepts what the scholar says, but feels perplexed for a moment, ruminating "How come this? What about what I've always been taught before, on Cleopatra being that Egyptian queen (**P'**), and all that? Should it be all thrown away?".

Secondly, non-isomorphism as conceived by Adler is at least problematic. According to him, within theoretical reasoning one seeks the alternative which is true, not the one that is better than the others in the face of the total set of available data. That is, for him, within theoretical reasoning, the only good enough alternative is the one corresponding to the truth, so that for an alternative to be accepted or believed on certain grounds *has to* equal its being acknowledged as the only true one – unlike in practical reasoning, where we seek to discern between several seemingly good alternatives which one is the best. Within practical reasoning, what it is for an

alternative to be taken as the one we're supposed to select is not equivalent to it being the only good one but simply one that is overall superior to the others. There seems to be something right about these considerations, but there is something strange and seemingly incorrect too.

It seems right to assume that the goal of all theoretical reasoning has to have a strong connection with truth – after all, if we reason from facts and evidence and we're looking for the perpetrator of a crime, for example, we want to get to the right person, the one who actually did it, and not merely one who is the most suspicious, or the most likely to have done it, over other suspects (for obviously the fact that one is the most suspicious among others does not suffice to justify an attribution of authorship – the real author may simply not be in the pool of suspects, to begin with). The same holds for cases in which we want to know the cause of a disease, or solve a mathematical problem – the demand is for the truth, not simply for something that fits better than the alternatives. So in this Adler is right.

On the other hand, it is not correct to assume that every theoretical reasoning works this way. Abductive inferences are inferences (that is, they are specimens of theoretical reasonings) and they work rather differently. In comparing two scientific hypotheses which one offers the best explanation for a set of phenomena, we seek to specify which one is superior; and the criterion may be the reach of explanatory power or parsimony in contexts in which it is either not appropriate or possible to determine the truth.

The same holds for a myriad of situations, such as when we want to determine which product has the best cost-effective ratio within a certain market branch – the search may even end up picking up two rival candidates, instead of just one. Adler seems to equate theoretical reasoning with inquiries of a very specific kind: those that aim at reaching the amount of evidence that would suffice to establish the *truth* of one single hypothesis. But not only is this far from amounting to all sorts of theoretical reasonings that there are – it does not even amount to all sorts of inquiries, not to mention the fact that both theoretical reasoning and inquiries are largely subject to operational flaws.

Be that as it may, when Adler says that theoretical reasoning "aims at what is right", he is establishing the "rules of the game" of this variety of reasoning, that is, he is establishing that theoretical reasoning is ruled by certain specific norms that involve not selecting one of the alternatives unless it's *the* right one. The illustrative

analogy of the difference between the norms of practical and those of theoretical reason, then, is the difference between the rules of poker and those of rummy: "in poker the best hand wins, while in rummy only the right sequence of cards wins" (Adler 2002a: 4).

The nature of those rules, however, can be interpreted in more than one way. The rules can be taken as stating that within theoretical reasoning, one only comes to acquire a belief that *P* whenever she judges all-out that there are good and sufficient reasons for the truth of *P* (call it "Rules of the Game – I"); or it can be taken as stating that within theoretical reasoning one *should* only acquire a belief that *P* whenever she judges all-out that there are good and sufficient reasons for the truth of *P* ("Rules of the Game – II")[25].

RG-II is prescriptive: it specifies how an epistemic agent is supposed to proceed in order to reason well. RG-I, in turn, is descriptive: it specifies what a mental procedure must have in order to be a theoretical reasoning. If what Adler means is that RG-I is the norm of theoretical reason, he must be willing to accept that any procedure that dissents from what RG-I specifies is not a proper theoretical reasoning. In contrast, if what he means is that RG-II is the norm of theoretical reason, then he must be willing to accept that such is the procedure by which theoretical reasonings must be operated to be good, or sound, so that failing to be in accordance with the norm renders the performed reasoning flawed.

Therefore, if Adler's argument is to have the (strong) effect of blocking in principle the occurrence of a state of mind that is the epistemic analogous to practical *akrasia*, it must be read as meaning RG-I, since it is RG-I, but not RG-II, which makes theoretical reasoning immune to the relevant flaw (RG-II specifies the way an agent should proceed, but leaves it open the possibility that she will not so do, namely, failing).

The problem with RG-I is that it sounds "overly optimistic about human rationality", as Mele (1992: 114) observed. It is not inadequate to capture the way we,

---

[25] I say it can be interpreted in these two ways because the overall theory Adler is committed to, especially in 2002b, is a sort of evidentialism that purposefully conflates (or in other ways equates) *is* and *ought to be*. According to him, belief has its own normativity, or is itself its own normativity, because the ethics of belief is rooted in the very concept of belief. Therefore, according to him, it is not accurate to state that one ought not believe something of which there is no evidence, such as, for example, that the number of stars is even – because actually one *cannot* believe it, according to him, due precisely to the absence of evidence. The "cannot" represents a conceptual barrier, not just an inability; so what beliefs one has and what beliefs she should have turn out to coincide.

ordinary human beings in general, think and proceed. Our experience as ordinary human beings is subjected to bad inferences of all kinds. Anywhere you look, you see imperfect reasoning. For instance, cases in which the subject forms a judgment with respect to P, for example, $f \therefore P$; and acquires the belief that P, but without f being good and sufficient evidence for the *truth* of P, or under circumstances where it is at least very controversial that it is.

This seems to be especially common regarding hasty generalizations, which are bad inductive inferences. Take as an example the Timothy Evans case:

TIMOTHY EVANS
In November 1948 the British police arrested and accused Timothy Evans of the murder of his wife, Beryl. Beryl was pregnant and the couple had agreed to have an illegal abortion, allegedly as a result of which Beryl had suffered complications that led to her death. Evans was stunned and, feeling guilty for agreeing to the abortion, eventually confessed having killed his wife. As a result of his confession, the police never investigated the possibility that someone else was responsible for the crime. Four years after Evans was sentenced to death and executed, the real killer was discovered: John Christie, neighbor of the Evans couple, a serial killer who had already killed and concealed the bodies of seven women, confessed having strangled Beryl on the day of the alleged abortion and then told Evans she died due to complications in the procedure. At the time of Beryl's murder, John Christie already had criminal records, had made two other victims and concealed their bodies in his front yard. That is, if only the police had conducted an investigation at his adjoining house and found these two bodies, Evans's conviction would probably have been avoided.[26]

In the procedure whereby the British authorities concluded that Timothy Evans murdered Beryl (P), they reasoned from Evans' confession (f) that Evans murdered Beryl ($f \therefore P$). Provided RG-I, we would have to say that the British police judged all-out that Evans' confession (f) offered them reason that was good and sufficient for P.

Nevertheless, it does not seem that the British authorities had good and sufficient reason to take P as *true* – namely because P was false, as well as because the only reason they had for P at that time (i.e., f) does not seem to be anywhere near good and sufficient. So if RG-I were the rule of thumb for this variety of theoretical reasoning, then either $f \therefore P$ was not really the all-out judgment by which the British authorities came to acquire P at the time, or the demeanour through which they came to acquire P cannot really count as a specimen of theoretical reasoning. That the demeanour in question cannot count as a theoretical reasoning is highly

---

[26] Adapted from Woffinden (1988).

disputable (if not a theoretical reasoning, what else would it be? A case of evil genius, perhaps?). However, if $f \therefore P$ was not the all-out judgment whereby the British authorities came to acquire P, then the acquisition of P prompts a state of dissonance which is precisely what Adler was trying to show was impossible.

On the other hand, if we suppose RG-II instead of RG-I, the prognosis is not much better to the effect of blocking C1. In this case, it does not seem that the British authorities had good and sufficient reason to take P as true either – not just because P was false, nor because it was entirely within their power to carry on a broader investigation that would have brought about that P was false, but above all because to carry on a broader investigation seems to correspond precisely to what is *expected* from law enforcement authorities in circumstances such as this. This is what is expected from them *to be* law enforcement authorities, to begin with.

So it is reasonable to assume that these authorities had a better general intellectual judgment, an even broader general principle regulating their epistemic conduct, that they were violating. According to this broader principle, conclusions such as P should only be drawn after broad investigations have been properly conducted, since only this sort of investigation is capable of equipping one with good and sufficient reason for the truth of P. Nevertheless, the conclusion that P was drawn (and indeed the belief that P was acquired) in the absence of such an investigation, and in the full awareness that it was being drawn on the basis of (only) f, precisely setting up, once more, the state of dissonance that configures of C1[27].

2.5. Final remarks

The bottom line of the analysis carried out in this chapter is the following. Some authors, namely Pettit & Smith and Adler, put forward arguments claiming that the acknowledgment condition for epistemic akrasia, C1, cannot be met on account that C1 entails TEJ and TEJ is false. Those arguments do not succeed. The most Pettit & Smith and Adler's arguments can successfully show is that the so-called "evidential akrasia" is impossible, i.e., that it is impossible for a subject to *acquire* a belief *P' in the face of* evidence that she deems good and sufficient for P, where P and P' are rival hypothesis. But they do not show that it is impossible for a subject to

---

[27] Of course, the Tim Evans case can be interpreted as a case of hypocrisy or malice on part of the law enforcement authorities. There is no doubt that this does happen and that, in it, the final situation does not configure C1. However, unless one can give reasons why the case should be interpreted as a case of pure and simple evil, there is no reason to think that it cannot be read as a bona fide case of flaw.

fail to acquire **P** in the face of evidence that she deems good and sufficient for **P**, as in the Wilma case; neither that it is impossible for a subject to acquire **P** *in the absence* of evidence that she deems good and sufficient for **P**, as in the Tim Evans case. And since cases like Wilma's and Evans' meet the acknowledgment condition for epistemic akrasia, C1, in order to claim that C1 cannot be met, one needs to show that such cases are impossible, which Pettit & Smith and Adler's arguments do not manage to do. The enterprise of disclaiming epistemic akrasia depends, thus, on the successfulness of the arguments purporting the impossibility of meeting the control condition, C2. This is what we're turning to now, in Chapter 3.

3. SKEPTICISM OVER EPISTEMIC AKRASIA: PART II

Anaxagoras says that what is referred to as "the milk of birds" is actually egg-white.

Athenaeus, **Deipnosophistae** II 57 B.

In the previous chapter it has been discussed how that talk of "epistemic akrasia" among scholars is normally informed by a grounding assumption, which I called ISO. ISO is the idea that epistemic akrasia is the *exact* same thing as practical akrasia, but affecting one's theoretical reasoning, instead of his actions. As practical akrasia is weakness of one's will to act, or action against one's better judgment, epistemic akrasia is defined in the exact same way, but with the terms "action" and "to act" being replaced for "belief" and "to believe".

In recent decades, some scholars have called this ISO-based concept of akrasia into question, among which Pettit & Smith (1996), Adler (2002a), Hurley (1989) and Owens (2002), on account of their skepticism about its very possibility. They claim this is a phenomenon whose conditions of possibility are impossible to meet, in principle. This chapter carries on the discussion of this skepticism, right from where last chapter left us.

3.1. Preliminary remarks

The view advanced by the epistemic akrasia skeptics has the consequence that epistemic akrasia amounts to nothing but an empty name; and things that have been described as instances of this pseudophenomenon in the past (like the myriad of examples laid down on Chapter 1) have certainly been twisted and distorted, or even made up, to fit its description. What is really going on in those cases, if those scholars are right, has to be something else. Accordingly, the very notion of epistemic akrasia is epistemologically useless.

If epistemic akrasia is spelled out in the exact same terms as practical akrasia, but only with "action" and "to act" being replaced for "belief" and "to believe", its

possibility comes down to the satisfaction of two jointly sufficient conditions: the agent must acknowledge the relevant belief as dissenting from her better epistemic judgment; and things have to be such that it was possible for the agent to not have the relevant belief. In Chapter 2 I've made the case that arguments purporting that epistemic akrasia is impossible on account that the first condition cannot be met do not succeed. That's because they're only capable of demonstrating that one form of epistemic akrasia is impossible (evidential akrasia: to acquire a belief *p* in the face of evidence that one deems good and sufficient for the truth of **not-*p***). But epistemic akrasia could take many forms that fit the description, as much as practical akrasia can, and not all of those forms have been ruled out.

Practical weakness of the will can happen when one takes course of action B in spite of judging that course of action A was best all-things-considered. But is can also happen when one *fails to* take A in spite of judging A to be the best all-things-considered. And it can also happen when one takes B *in the absence* of a proper all-things-considered judgment (when one's all-things-considered judgment is very poorly formed, or vague).

Think like this: I fall into the akratic trap if I choose to have a chocolate cake for dinner in spite of judging that a bowl of vegetables would be best for me all things considered. I also fall into the trap if I simply fail to have the bowl of vegetables, even if I'm able to resist the cake. That is, if I skip dinner and I starve myself, this is a way of not doing what I deem best too. And I also fall into the trap if I choose to have a chocolate cake for dinner in the absence of a specific judgment all-things-considered that recommends the vegetables. If all I have is a vague and unspecific judgment, like "eat healthy", this judgment says nothing about a bowl of vegetables, or about today's dinner; but since chocolate cakes do not fit the description of "healthy" by any accounts, then if I choose to have the cake for dinner, I'm having an issue with my willpower too. I was not strong enough to do the right thing, so to speak.

Likewise with epistemic akrasia. The typical description is the case in which a person acquires the belief that *p* in the face of evidence deemed good and sufficient for the truth of ***not-p***. This is impossible, like Chapter 2 demonstrated. But it can also happen that one *fails to* acquire the belief that *p* in the face of evidence deemed good and sufficient for the truth of *p*. And it can also happen that one acquires the belief that *p in the absence* of evidence deemed good and sufficient. All those cases

square off as cases of epistemic akrasia according to the definition given, that is, they are ways of being epistemically akratic too.

Therefore, in order to make the case that epistemic akrasia is impossible, one needs to show either that such cases are impossible in principle, or that what's keeping epistemic akrasia from being possible is rather something else. Arguments devised in Chapter 2, Pettit & Smith's and Adler's, do not manage to prove that the abovementioned cases are impossible in principle. So for epistemic akrasia to be impossible, there has to be something else getting in its way. In this chapter, then, I'm going to dig into this "something else". A couple of arguments can be put forward purporting that epistemic akrasia is impossible on account that when an agent acquires a belief, things are such that it was *not* possible for the agent to not have acquired it. That is, they purport that the epistemic control condition for epistemic akrasia is impossible to meet. Do those arguments succeed?

3.2. Doxastic control

As last chapter made explicit, the ISO-based characterization according to which epistemic akrasia is usually modelled can be spelled out like so:

Isomorphism (ISO)
Epistemic akrasia episodes are isomorphic to practical akrasia episodes: they are, to one's beliefs, what practical weakness of the will is to one's actions.

It follows from this ISO-based characterization of epistemic akrasia that for a person's epistemic conduct to be correctly described as an instance of the phenomenon, this conduct must meet a couple of conditions that are *epistemic equivalents* to the conditions for practical akrasia.

Practical akrasia, like we've seen, happens when the individual takes or is inclined to take a course of action that is acknowledged as being at odds with his better judgment (acknowledgment condition for practical akrasia), in contexts in which it would have been possible for him to act differently (control condition for practical akrasia). Therefore, epistemic akrasia has to be, or is supposed to be, the conduct displayed by an individual in having or maintaining epistemic attitudes (i.e., forming beliefs) that are acknowledged as being at odds with her better intellectual judgment (acknowledgment condition for epistemic akrasia), in contexts where it

would have been possible for her to have or maintain a different attitude (control condition for epistemic akrasia).

More technically speaking, we have

Epistemic Acknowledgment Condition ("C1")
The belief involved in an epistemic akrasia episode is acknowledged by the individual herself as dissenting from her better epistemic judgment.

and

Epistemic Control Condition ("C2")
Things are such that it was possible for the belief involved in the epistemic akrasia episode to not have been acquired or held.

As I hope to have been able to demonstrate throughout Chapter 2, the main arguments claiming that C1 cannot be met fail. Now I would like to examine some of the arguments tackling C2. C2 is a clause about the possibility of things turning out differently, in terms of what belief the agent ended up acquiring, or holding, in a given epistemic situation. Thus, it's about the possibility an agent has of interfering in the process by means of which her beliefs are acquired and lost, that is, dismissed. Or, simply, it's a clause about the exercise of *doxastic control*.

Doxastic control can be conceived in more than one way. We routinely influence our doxastic states indirectly, by selectively exposing ourselves to evidence, by deciding to carry on or deciding to close investigations, by double-checking inferences, etc. However, given ISO, doxastic control would have to be, literally, a form of control of the same type as the one we have over our actions. That is, given ISO, C2 is to be interpreted as entailing CB:

Control over beliefs (CB)
Beliefs are something over which we have as much control as we have over actions.

We need to have at least as much control over beliefs as we have over actions, for epistemic akrasia to be possible. The problem is: it looks as though we have much more control over what we do than we have over what we believe or accept. If you offer one thousand dollars for me to stop doing something that I currently am doing, say, for instance, dancing, or writing, I can accept it; whereas if

you offer one thousand dollars for me to stop believing something that I currently believe, such as that the Earth is round, or that 2 + 2 = 4, I simply cannot.

In this way, one draws the conclusion that epistemic akrasia is impossible by rejecting CB. That is, by reasoning: for there to be an epistemic akrasia C2 has to be met. C2 entails CB, but CB is false – we do not have as much control over beliefs as we have over actions. If CB is false, then there is no such a thing as an epistemic akrasia. So, let's try and analyse CB to see whether it is indeed false.

There is more than one way of breaking CB down. The default way of breaking it down is as follows: if we have over our beliefs as much control as we have over our actions (CB), and if we have direct and indirect voluntary control over our actions, then we must have direct and indirect voluntary control over our beliefs. This idea, that beliefs fall under our voluntary control, is known as *doxastic voluntarism*. Doxastic voluntarism is a *subjection-to-the-will-based model of control* and it is contemporaneously advocated by authors such as James Montmarquet (1986).

An alternative way of looking at CB is: if we have over our beliefs as much control as we have over our actions (CB), and if we exercise control over our actions as far as we are capable of forming a view on the merits of each of the possible courses of action placed before us, then we must exercise control over our beliefs by forming a view on the merits of each of the plausible *hypotheses* placed before us. This is a *judgment-based model of control*, within which it is practical judgment, not the will, the means by which we exercise control over agency, but practical and epistemic. Such a model has been defended by David Owens (2000, 2002).

An argument claiming that C2 cannot be met on account that CB is false, then, can run two basic strategies. It can reject CB by denying that the subjection-to-the-will-based model of control applies to beliefs (strategy I); or reject CB by denying that the judgment-based model of control applies to beliefs (strategy II). In the next two sections I'll present, discuss and offer objections to two of those arguments.

3.3. Williams' classic argument

For one who subscribes to a subjection-to-the-will-based model of control, it might seem somewhat obvious that we do not control our beliefs voluntarily in the same way we control our actions. That's because only actions, but not beliefs, are subjected to our will both direct *and* indirectly.

I can choose to raise my arm now and my arm will be raised as a consequence of my willpower alone, in which case I'll be exercising direct voluntary control over this action. And I can choose to make this room dark right now, by choosing to bring about the circumstances under which the lights will be turned off. The lights will be turned off as a consequence of me performing the series of actions that will produce it as a result, for instance, getting up from my chair, walking to the switch and pushing my finger against it. In this case, I'll be exercising indirect voluntary control over the relevant action.

Beliefs, on the other hand, are subjected to our will only indirectly. I can choose to expose myself to a certain amount of evidence that speak in favour of a certain theory, in which case I might end up embracing the theory as a result, that is, believing that it is true. But I cannot bring myself to believe a theory by means of my willpower alone, that is, just by wishing. This is a somewhat self-evident, or obvious fact about beliefs. We lack direct voluntary control over them.

So in this interpretation what we have is this. The ISO-based idea of epistemic akrasia entails Control over Beliefs, CB ("Beliefs are something over which we have as much control as we have over actions"). For CB to be true, direct doxastic voluntarism has to be true. But direct doxastic voluntarism is obviously false. So to someone who subscribes to a subjection-to-the-will-based model of control, CB is obviously false, and strategy I actually does not even have to be run – that it does succeed is somewhat self-evident. It is as though there was no need to build this idea on top of arguments.

Nevertheless, Bernard Williams (1973) famously advanced an argument against direct doxastic voluntarism, the so-called "classic argument"[28]. According to the classic argument,

> It is not a contingent fact that I cannot bring it about, just like that, that I believe something (…). Why is this? One reason is connected with the characteristic of beliefs that they aim at truth. If I could acquire a belief at will, I could acquire it whether it was true or not; moreover I would know that I could acquire it whether it was true or not. If in full consciousness I could will to acquire a 'belief' irrespective of its truth, it is unclear that before the event I could seriously think of it as a belief, i.e. as something purporting to represent reality. At the very least, there must be a restriction on what is the case after the event; since I could not then, in full consciousness, regard this as a belief of

---

[28] This argument was also presented in Williams (1970).

> mine, i.e. something I take to be true, and also know that I acquired it at will. (Williams 1973: 148)

Williams' reasoning is as follows. If people could believe propositions just by wishing, they would be able to judge a proposition as true regardless of whether they think that that proposition is true or not. Concomitantly, people would know that they have this capacity. That is, they would know that it is within their power to make truth-value judgments about propositions regardless of whether they think that these propositions are true or not.

So, if direct doxastic voluntarism was true, Mary could believe, for example, that John is by the door regardless of her judging that the proposition "John is by the door" is true. In addition, Mary would know that she could acquire such a belief in this way, i.e., regardless of whether she thinks that it corresponds to reality or not. This, however, contradicts the very nature of doxastic acquisition and obliterates the very notion of belief – believing that *p* just is taking *p* as true[29]. You cannot have a belief that *p* unless you judge that *p* corresponds to reality, because believing something just *is* judging that it does correspond to reality.

For one thing, if a person believes that *p*, she will be surprised, or experience some form of awkwardness, or uneasiness, whenever she finds out that *p* is false. If direct doxastic voluntarism was true, however, someone who believed that *p* would not be surprised at all to find out at some point that *p* is false, because the truth of *p* was never taken into account for the very acquisition of the belief that *p*, to start with. But this is not what we see happening. People do get emotional responses, like surprise and others, when they realize that they were wrong about something. What's more, if direct doxastic voluntarism was true, people would not be able to face the beliefs that they have acquired from will as proper beliefs, that is, as mental states

---

[29] In fact this idea – "to believe that *p* is to believe that *p* is true" – is not the only possible interpretation of Williams' passage where he says that it is necessarily (and not as a matter of contingency) that "beliefs aim at the truth". Aiming at the truth is something that can be understood in more than one way, and there is a whole separate discussion devoted to understanding what these ways are and which of them would be the proper one. See, for example, Velleman (2000), Owens (2003), Shah (2003), Shah and Velleman (2005), Vahid (2006), Zalabardo (2010) and Whiting (2012). For present purposes, however, I'll refrain from getting into this discussion by assuming, without committing myself to any particular view, that what Williams means (and what the central intuition of authors who, like him, deem direct doxastic voluntarism false) is that there is a nexus between believing and taking as true; a nexus that, of whatever kind, makes it impossible for a subject to believe that *p* without judging that *p* is true.

intended to correctly represent the actual world. Accordingly, if Williams is right, direct doxastic voluntarism is false; and that is why Control over Beliefs, CB, is false.

Now, the main problem with this argument is that there is a difference between belief acquisition and belief fixation. Forming a belief "from scratch" is not the same thing as reassuring a belief that has already been formed, and that was being kept provisionally. When we acquire a belief, say, that *p*, we move from not having an attitude towards *p* to a state of mind in which we believe *p*, or in which we have certain degree of confidence that *p* that surpasses the relevant threshold, whatever the threshold turns out to be. When we fix a belief, on the other hand, we establish the truth or correctness of something that was previously accepted, or suspected, on provisional grounds. That is, we move from having a doxastic attitude towards *p* that falls short of a belief (for instance, suspicion, acceptance or hope) to a state of mind in which we believe *p*, or have confidence that *p* to a degree that surpasses the relevant threshold.

Some authors who have called attention to this difference include Johnston (1995: 438), Winters (1979: 253), and Scott-Kakures (1994). According to them, it is possible for a person to form a belief that *p* for reasons that are entirely practical (i.e., regardless of their stance on the truthfulness or the falsehood of the proposition *p*); in a way that they can become aware further down the line, or acknowledge later on, that there is evidence in favor of *p*. At this later moment, the person pins *p* into her doxastic mesh, tying it to other beliefs in the proper way. That is, she affixes it.

Consider, for example, the Helen case:

HELEN
Helen is at the hospital's hallway waiting for news about the emergency surgery Jonas is undergoing. She is then approached by a fortune teller who tells her that the surgery is going well and that nothing bad is going to happen to Jonas. Helen is skeptical of clairvoyant skills in general and knows that the words of that fortune teller constitute no evidence in favor of the claim she is making. Nevertheless, after listening to the fortune teller, Helen sets her mind at ease, thinking that the surgery is going well and that nothing bad is going to happen to Jonas. Moments later, one of the surgeons enters the room and announces that the surgery is going well and that nothing bad is going to happen to Jonas. Helen finally breaths a sigh of relief, for now she knows that the surgery is going well and that nothing bad is going to happen to Jonas.

Cases like Helen's seem possible and even unexceptional. One way of accounting for what is going on in this case is supposing that Helen acquires the

belief that **p** ("the surgery is going well and nothing bad is going to happen to Jonas") after listening to the fortune teller not because she thinks such words constitute evidence for the truth of **p**, but rather because she somehow wants **p** to be the case and the fortune teller's talk dovetails with her wish. We can even treat this in terms of the fortune teller's talk having induced, or triggered, in Helen the belief that **p** insofar as it dovetails with her wish. This is often true about very charismatic individuals: they are able to persuade other people, mostly by manipulating affective states, rather than by presenting reasons.

Supposing that this could be the case, then at this point Helen forms the belief that **p** on purely non-epistemic grounds, that is, without having given proper consideration to the issue of whether **p** is actually justified or not. Then **p** remains tentatively installed (i.e., not properly affixed) within her doxastic mesh, until she comes across what she takes to be good and sufficient evidence in its favour (the surgeon's testimony). At this later moment, Helen affixes **p**, by acknowledging that **p** is now warranted; and she may even dismiss her earlier attitude as unreasonable, gullible, or as constituting a flaw in evaluating the available evidence (or its lack thereof).

At the same token, it is conceivable that, had the surgeon came and announced that the surgery was not going well and that terrible sequels were to befall Jonas, Helen would have been surprised, rather than indifferent. In fact, it is actually hard to picture this scenario without Helen being surprised, if Helen is to be an ordinary human being, rather than a merely abstract or fictional character[30].

If cases like Helen's are possible, then it seems it is possible for a person, at one point, to acquire a belief that **p** regardless of her consideration of the truth of the **p**-proposition (which is exactly what Williams had said was conceptually impossible). And it is possible that this person, at a latter moment, accesses evidence that enable

---

[30] It has been pointed to me that what is hard to picture, if Helen is to be an ordinary human being (as opposed to an abstract or fictional character) is not a scenario in which she is not surprised, but rather one in which she is not disappointed; and that surprise and disappointment are different states. Disappointment can be defined as involving a sort of "regret that things are not the way it was hopped they could be" (Couper-Kuhlen 2016: 100), whereas surprise is usually defined in terms of break in expectation. Meyer et. al. (1991: 295-296), for instance, speaks of it in terms of as a suit of reactions elicited by events that deviate from schema (an agent schema, roughly, is his theory about how things are). So it might be contented here that what happens to Helen is not that she is surprised by the announcement that Jonas' surgery is not going well, but rather that she is disappointed by it. It is not clear, however, that disappointment is not, at the end of the day, just negative surprise. Break in expectation, where expectation was preferred. For an example where disappointment is defined precisely in this way, see Smith& Wrinkler (2006).

her either to properly affix **p**, thus rendering it a proper, full-blown belief, or dismiss it completely. And it is also possible that at this latter moment she acknowledges that **p** had been acquired through will (and thus unreasonably) moments ago[31]. So the question is: are cases of this sort really possible? Is the description given to what is going on adequate? If they are possible, and the description is adequate, then William's argument against direct doxastic voluntarism ceases to appear so convincing.

Some cases akin to Helen's were presented by Ginet (2001: 64-65), in an attempt to make the case for direct doxastic voluntarism. In one of them, the individual was road tripping on holiday miles away from home when asked by his wife about whether or not he had locked the front door. He reckons he did, but he's not sure. Given that the only way to access evidence that would enable him to be sure is going back home and checking it out, and given the great inconvenience of so doing, he decides to believe he did lock the front door, for reasons entirely practical: just to avoid the inconvenience of going back as well as to prevent his worries from disrupting his mood during the trip.

What is common among these cases (Ginet's and Helen's) is that in them people find themselves in situations in which they are, at least relatively, *concerned* about something being the case or not, about not having conclusive evidence for or against something, or about having ambiguous evidence. In other words: the individuals in those cases are experimenting a form of epistemic discomfort. They are bothered by not knowing something. This discomfort is alleviated when they choose to act as if the thing was true, or to accept it provisionally, even in the absence of conclusive evidence. It seems, at least prima facie, then, that in such situations we have direct voluntary control over the beliefs in question, insofar as we only hold those beliefs due to (at least in part) our will, the will to mitigate the discomfort.

---

[31] I do not wish to commit myself here to any particular taxonomy of the doxastic family, for I take it to be sufficient, for present purposes, to commit myself to the idea that there is a variety of forms of assent, or epistemic pro-attitudes, among which belief is only one species. This idea reaches back at least the stoics and has been contemporarily defended by authors such as Cohen (1989, 1992), Bratman (1992), Engel (1998, 1999, 2018) and Schwitzgebel (2001). For a nice review on this, see Engel (2012). The idea that there is a variety of mental states that share with paradigmatic beliefs some features, but lack others calls for a distinction between full belief, or full-blown belief, and other belief-like attitudes, that deserve the title of "belief" only by courtesy. I also do not want to advance any technical definition of full-blown belief here. For present purposes, it suffices that I state, in very broad terms, that a full belief or full-blown belief differs from other belief-like attitudes in that the former is evidence-responsive in a way that the later are not. I'll say more about this in Chapter 4.

One might immediately object that in such cases what people are doing is not exercising control over *belief* acquisition, but rather over something else. That cases such as Ginet's and Helen's are cases not of the subject *believing* a proposition, but rather *accepting* it, or *wishful thinking* it. Accordingly, the agents in these cases decide to provisionally take something as being the case for the sake of fulfilling some practical need. But, contrary to what happens when proper, full-blown, beliefs are at stake, in so doing they do not undertake a proper commitment to the relevant proposition being true (see, for instance, Buckareff 2004, Bratman 1999 and Cohen 1989, 1992). Or, alternatively, somebody could claim that the cases in question are cases in which the person *acts as if* something was the case (see, for instance, Alston 1989: 122-127 and Steup 2000), and that to act as-if-is in itself an attitude one takes toward the relevant proposition. The person goes through the motions, so to speak, behaving as if something was the case, say, **p**. She does it for the sake of some practical purpose, regardless of whether or not she effectively believes, affirms, denies, or suspends judgment as to the proposition **p**.

If this objection is sound, it restores the power of Williams' classic argument, because it establishes that cases where the acquisition of a belief is distinct from its fixation (cases such as Helen's and Ginet's) are not really cases involving belief, but rather something else, that is, some other type of attitude. In this case, the objector is committed to a version of *doxastic pluralism* (something we'll go back to in section 3.4.). Be that as it may, if the objection is sound, Helen's and Ginet's cases are not real counterexamples to William's argument. Direct doxastic voluntarism continues to be false when it comes to proper, full-blown beliefs. And if this is so, then Control over Beliefs, CB, is false; and the epistemic control condition, C2 ("it was possible for the belief involved in the epistemic akrasia episode to not have been acquired or held"), cannot be met on account that CB is false.

Interestingly, however, it does not seem to make much difference for the sake of the epistemic control condition, C2, whether the attitude thus involved is proper, full-blown belief, or some other type of epistemic attitude. This is because C2, and by extension CB, are meant to be clauses about doxastic attitudes understood in a broad sense, rather than clauses about (only) full-blown belief. Le me explain this.

The ISO-based characterization of epistemic akrasia departs from the assumption that epistemic akrasia is the exact same thing as practical akrasia, but affecting one's theoretical reasoning instead of one's actions. Theoretical reasonings

are not constituted by full-blown beliefs alone. That is, we do not operate *only* by forming and voicing full-blown beliefs. When we reason, we manipulate propositions, by undertaking attitudes towards them and by drawing inferences from those attitudes. But there is a myriad of types of attitudes one can have towards propositions, apart from the attitude of believing them.

Inquiries provide such a good example of this. When we inquire towards, say, **p**, we do not have a full-blown belief either that **p** or that **not-p** (Friedman 2019: 309). We inquire precisely because we do not have a full-blown belief, but want to arrive at one. In those occasions, what we do have towards **p** is some other attitude, such as supposition, suspicion or, perhaps, provisional acceptance (when we intentionally work from the assumption that **p** is the case, that is, when we take **p** as an underlying hypothesis). Those are attitudes that bear some relationship with truth, which means truth matters to them, but not in the same way it matters to proper, full-blown beliefs. To believe that **p** is equivalent to judging that **p** is true, while to provisionally accept **p**, for instance, is not. Without at least some of these attitudes the set of theoretical reasonings that inform one's inquiry would simply be impossible to make sense of.

If our theoretical reasonings are made of a variety of epistemic, or doxastic, attitudes, and not only of full-blown beliefs, then episodes of akrasia can, at least in principle, involve any of those attitudes. They do not need to involve a person acquiring a full-blown belief that goes against her better epistemic judgment. They might as well involve a person undertaking any of those other attitudes, in a context where that particular attitude conflicts with her better epistemic judgment and where things were such that she could have done otherwise.

The bottom line of this section is: Williams' classic argument against direct doxastic voluntarism supports the case that a *subjection-to-the-will-based model of control* does not apply to full-blown beliefs. It convincingly makes the case that beliefs are not subjected to our will. In a subjection-to-the-will interpretation of CB, then, CB is false. Therefore, the control condition for epistemic akrasia, C2, cannot be met on account that it entails CB, and CB is false. As a consequence, this argument proves that episodes of akrasia involving full-blown belief (full-blown belief akrasia) are indeed impossible. But since full-blown belief is not the only doxastic attitude that there is, the possibility is left open that other types of epistemic akrasia might exist, involving a range of other doxastic (belief-like) attitudes, attitudes that

could be subjected to some form of control. Those types of epistemic akrasia have not been ruled out.

3.4. Owens' alternative argument

In the previous section I showed that William's argument purporting that a subjection-to-the-will-based model of control do not apply to beliefs is not enough to make the case that the control condition for epistemic akrasia, C2, cannot be met on account that CB is false, because it does not suffice to make CB false. There are interpretations available to CB that William's argument does not reach, namely, interpretations according to which CB is not about full-blown belief, but rather about any belief-like doxastic attitude.

Now, there is an alternative way of looking at CB, following a *judgment-based model of control* (Owens 2000, 2002). According to this model of control, it is practical judgment, not the will, the means by which we exercise control over epistemic agency. CB states that we have over our beliefs as much control as we have over our actions. The judgment-based model of control states that we have control over our actions as long as we form views on the merits of each of the possible courses of action placed before us. So for CB to be true, we must have control over our beliefs by forming views on the merits of each of the two or more concurrent hypotheses envisaged. The question then is: do we have this form of control?

As we've seen when approaching Adler's argument in the previous chapter, when two or more possibilities of action concur, the agent forms a view as to which of the two is best overall, that is, she forms an all-things-considered judgment. Nevertheless, this judgment still allows for the dismissed alternatives to retain some of their value, or some of their appeal. This is how things are when it comes to practical reason. I might think, for instance, that the overall best thing for me to do at this time is to not eat another piece of a chocolate cake (all-things-considered judgment). Nevertheless, the idea of eating one more piece now still retains some of its appeal, as it would be quite nice to eat more chocolate cake now from the viewpoint of my immediate pleasure (all-out judgment). Practical akrasia is possible precisely because I can act, or be inclined to act, motivated by this latter judgment, rather than the former.

David Owens thinks that this mechanism of judgment-formation described by Adler is, in itself, a type of practical rational control. That is, it is a means for the agent to exercise control over her actions: a person can do what she judges best, or she can do something else. Effectively doing what one judges best represents the full exercise of the agent's control – she proceeds in a way that her action is in full agreement with what she deems best all-things-considered. Doing something else, in turn, represents a loss of control: the agent fails to proceed in agreement with her all-things-considered judgment. It is as though she has failed to keep herself within the bounds of her better judgment, so to speak.

As Owens points out, this judgment-based model of control applies to actions due to the fact that actions are intentional, that is, goal-oriented; and because they serve a variety of purposes. Actions are goal-oriented: we do certain things to attain certain results. And they serve a variety of purposes: different actions have different aims. Eating one more piece of chocolate cake now, for instance, is oriented towards immediate pleasure, whereas refraining from so doing is oriented towards ensuring an indigestion-free near future, or some other future goal. In the end, choosing between these two possibilities of action is a matter of choosing between their two respective aims, or the outcomes they're meant to deliver. Therefore, for this model of control to be extended to beliefs, beliefs must possess the two attributes alluded: they must be goal-oriented; and they must serve a variety of aims.

That beliefs are goal-oriented is something easy to accept – beliefs are oriented towards the goal of getting things right, that is, of correctly representing the world; or, simply, of being true (Owens 2002: 390-392). While it is actually debatable whether or not beliefs aim at the truth, this idea seems acceptable, once it explains, for instance, why we simply cannot believe something for which we have no evidence just because someone offered us a large sum of money in exchange.

Now, that beliefs serve a variety of purposes, according to Owens, is more difficult to accept. It seems as though beliefs serve the sole purpose of corresponding to reality, of representing the world in a way that allows us to navigate it. And if this is so, then beliefs do not meet the minimum requirements that would allow for the judgment-based model of control to apply to them. As a result, in a judgment-based interpretation of CB, CB would be false.

One might confront Owens' considerations with a case such as the Zoologist, below:

ZOOLOGIST
A zoologist goes into a research endeavor aimed at learning about the behavior of certain species of wolf that is believed to inhabit certain territory. Resources are distributed and schedule is planned, as for the goal of the research, which is to cover the largest lot of territory as possible and to attain as many observations as possible, within the available (limited) amount of time. Knowing that wolves are nocturnal and that the territory inhabited is pretty large, the zoologist demarcates numerous spots within the terrain map, and arranges facilities so as to spend his nights awaken, in order to maximize the probability of success in observation. While spending the night in each spot, he must lurk and seek to behold wolves. If no wolf shows up, he must decide whether the best move is to spend one more night within the same spot or to move into the next one. Staying for too many nights in the same spot may result in the zoologist running out of time and delaying the research schedule, thus preventing him from sweeping a large enough area. Moving to the next spot too quickly, however, may also undermine the research efforts, for he might end up having covered all the demarcated spots but without accomplishing a fair amount of observations.

After spending one night in one spot without having beholden a wolf, the decision as to stay or move is driven out of a process within which the zoologist reasons as follows: if he believes that the chances of beholding a wolf next night in his current spot are high (say he has heard howls during the night, for example), stay. If, on the contrary, he believes that his current spot is not very promising for next night, move.

At a certain time, he might believe that his current spot is promising, so that he should indeed spend one more night there, while, at the same time, judging that, given the amount of time left and the number of observations accomplished, it is moving, not staying, that would maximize the success of the investigation (by increasing his chance of being able to sweep the larger area). Conversely, it may happen that the zoologist judges that, given the amount of time and resources left, and the relatively smaller than expected number of observations accomplished, he should spend one more night at his current spot, albeit believing that the current spot is not very promising (if, for instance, no howls have been heard there).

In both cases, following Owens's, the beliefs at stake are goal-directed (they aim at being acquired only if true). However, contrary to what Owens has suggested, these beliefs serve a double purpose – namely, getting it right about the current spot (or each of the spots) being promising or not, in itself; and getting it right about the most intelligent way of managing time, that is, the way that maximizes his chances of succeeding in the research endeavour.

Many of our epistemic activities have double goals, more or less in this fashion; and, again, inquiries offer such a good example. Very frequently, if not most of the times, the beliefs we form in the course of an inquiry serve the double purpose of being acquired only if true *and* only if relevant to the central question that the inquiry is trying to answer[32]. Investigative and judiciary cases, such as the Tim Evans case, discussed in the previous chapter, are concrete examples of activities with double epistemic goals, in this sense. In them, the beliefs about the suspect perpetrator acquired throughout the investigation serve, or at least are meant to serve, the double purpose of being acquired only if true *and* of preventing wrongful convictions, inasmuch as the inquiry itself is aimed at bringing to justice the right person while avoiding prosecuting the wrong ones. Your job as a detective constable is to find the person to whom all the leads point towards while following (only) the genuine leads. Each of the beliefs you form throughout the process have this double aim.

Scientific explanations offer another example. The statements we adduce in the course of formulating a scientific explanation serve the double purpose of being adduced only if true *and* only if relevant to the thing we're trying to explain. If we believe that a given explanation for a given phenomenon is true, we believe that the statements adduced are true *and* that they are relevant for that explanatory task. Say I'm trying to explain, for instance, why a certain sample **S** of table salt dissolved in water. I'll explain this by means of an inductive generalization, from the observation that all samples of table salt that have been blessed by the local church's preacher dissolved in water. The explanation, thus, is that **S** dissolved in water because it is composed of sodium chloride and it has been blessed by the local church's preacher. This explanation is a bad one because one of the statements adduced is irrelevant: all samples of table salt, regardless of whether or not they have been blessed, dissolve in water, so the fact that this particular sample has been blessed is irrelevant (even if it is true). Therefore, it might be true that **S** has been blessed by the local church's preacher, but my belief that this is so fails to fulfill one of its goals within the explanation: the goal of being adduced only if relevant.

---

[32] This might be thought about in terms of belief having a constitutive goals (being acquired only if true) and non-constitutive goals (being acquired only if relevant, for instance). Non-constitutive goals might subject beliefs to some judgment-based variety of control: I might set aside the issue of whether *p* is true or not (and therefore not acquire the belief that *p*) because I make a negative judgment on the merits of *p*: that *p* is not relevant for the central question the inquiry is trying to answer.

Now, if this is correct, if the beliefs a person forms throughout the course of activities like the abovementioned (the Zoologist's quest, investigative and judiciary cases in general, and scientific explanations) serve a double epistemic purpose, then why would those beliefs not be eligible to be subject to the judgment-based model of control described by Owens? Apparently, they are eligible.

Owens could object to this by claiming that the type of attitude we take towards the statements being adduced for an explanation is not belief, but rather acceptance. In a similar fashion, he could claim that the attitude undertaken by the Zoologist as a means of prompting a decision as to stay or move is not proper believing, but rather *guessing*. Like believing, guessing aims at the truth; but the two attitudes are not purposive "in the same way" (Owens 2002: 392). In making an individual guess such as, for instance, a guess as to whether the current spot is promising or not, the Zoologist is trying to get that particular matter right. However, the correctness of this particular guess is not the only goal he is pursuing. He is also after a practical goal. His judgments about when to move, about how much to wait before forming a final resolution as to stay or move, etc., will be informed by his need to maximize the productivity of the research endeavour (that is, by the need to attain as many observations as he possibly could while, at the same time, covering as many spots as possible throughout the terrain, within the available time window). The latter is not an epistemic goal, but rather a practical one.

This, according to Owens, is exactly what guessing is; as opposed to believing. In guessing, certain practical considerations matter that do not matter to believing, such as productivity. In believing, we're supposed to take only epistemic considerations into account, for truth does not care about productivity, or about any of our practical needs. Each of the individual attitudes the Zoologist undertakes at the end of each night, therefore, is more like an individual guess, an informed guess, rather than a belief[33].

So guessing-situations, according to Owens, are rich in akratic possibility precisely because this attitude, of guessing, unlike believing, falls within our sphere of direct control. We control our guesses, that is, we control whether or not to make a particular guess, by forming views on the merits of that guess. Guessing, thus, is subject to the judgment-based model of control, since it satisfies both the

---

[33] Owens presented a somehow akin case: the case of an individual who is going through a pop quiz by guessing answers. See Owens 2002: 392.

requirements for this model of control, whilst believing does not. It is as though guessing is more of a practical attitude than a doxastic, or an epistemic one. Therefore, akrasia of guessing is possible, following Owens; but not the real epistemic akrasia (full-blown belief akrasia).

Now, two things must be noticed. First, that Owen's argument, much like the objection against Helen's and Ginet's cases presented in the previous section, is committed to a version of doxastic pluralism, the doctrine that belief-like attitudes are varied, and has to rely on this doctrine being true.

Second, Owens' distinction between believing and guessing looks considerably stipulative (provided that pretty much all of the epistemic attitudes we undertake in real life are embedded within a practical framework, how can I tell, in each case, whether the attitude a person is undertaking is believing or guessing?). So it is not clear why attitudes such as the Zoologist's, whatever it turns out to be, once falling under the judgment-based model of control, cannot figure in the very description of an epistemic akrasia episode. After all, the Zoologist's endeavour is an intellectual, or epistemic enterprise; and the attitudes undertaken by him as part of this enterprise (even if they indeed fall short of proper full-blown belief) are epistemic, or, at least, have an evident epistemic dimension: the concern with arriving at knowledge.

Like in the Helen and the Ginet cases, there is an epistemic dimension underpinning the attitude undertaken by the Zoologist at the end of each night. This attitude is not constitutive of his intellectual conduct any less than full-blown beliefs, on the contrary – the success of the scientific enterprise depends on the theoretical reasoning that underlies it including this attitude, as much as it depends on the final acquisition of a full-blown belief. That is, the Zoologist's attitude, like Helen and the character in Ginet's example, *are* epistemic. Acceptance, act-as-if and guessing have an epistemic dimension, they're part of our theoretical reasonings as much as full-blown beliefs are. Like full-blown beliefs, those are attitudes we take regarding propositions and they involve taking stances as to something being the case or not, albeit provisionally or hypothetically.

In sum, Owens' alternative argument against direct doxastic voluntarism supports the case that a *judgment-based model of control* does not apply to full-blown beliefs. It convincingly makes the case that we do not select which beliefs to acquire and which to dismiss by forming views on the "merits", or the "vantages", of each possible belief, because full-blown beliefs do not have double-aims. There is

one single relevant criteria on the basis of which to judge beliefs, and this criteria is truth. In a judgment-based interpretation of Control over Beliefs, CB, then, Owen's argument makes CB false.

However, this argument only makes CB false if CB is spelled out in a narrow way it has been spelled out following the ISO-based characterization of epistemic akrasia: in a way that specifies, or stipulates, that the attitude at stake in the akrasia episode *is* proper, full-blown belief. But the ISO-based characterization itself leaves room for the possibility that epistemic akrasia episodes could involve other sorts of attitudes, since ISO states that epistemic akrasia is akrasia affecting our theoretical reasonings, and theoretical reasonings are not constituted by full-blown beliefs alone.

The takeaway from this section is that when it comes to the specific task of making the case that C2 cannot be met on account that CB is false, Owen's alternative argument backfires for the same reasons why William's classic argument backfires: those arguments only suffice for proving that episodes of akrasia involving full-blown belief (full-blown belief akrasia) are impossible. But since those very arguments entail that full-blown belief is not the only epistemic or doxastic attitude that there is (inasmuch as they entail doxastic pluralism), the possibility is left open that, if doxastic pluralism is true, epistemic akrasia might exist involving a range of other doxastic (belief-like) attitudes, attitudes that could be subjected to one form of doxastic control or another. The next section is devoted to a detailed discussion of this issue, the issue of doxastic pluralism.

3.5. Doxastic pluralism

There are at least two ways of looking at cases such as Helen, Ginet's and Zoologist. One can think that in such cases, when the agents undertake an attitude towards the matter at hand ("Jonas's surgery is going well"; "The front door is locked"; and "the current spot is a must vantage point to behold wolves", respectively) what they are doing is not exactly forming certain beliefs, but rather something else. That those are *not* cases of people believing the relevant propositions, but rather accepting them, acting-as-if them, wishful thinking them, or simply guessing, or some combination of those. The general idea that attitudes such as acceptance and others are qualitatively, functionally and normatively distinct from beliefs has been held by people such as Cohen (1992), Bratman (1993), Audi (1994),

Engel (1998, 2000, 2018) and Dinges (2022). This quote from Engel summarizes the general point:

> The doxastic zoo contains many animals: belief, acceptance, belief in, belief that, certainty, conjecture, guess, conviction, denial, disbelief in, disbelief that, judgment, commitment, etc. It also contains belief's "strange bedfellows": credences, partial beliefs, tacit beliefs (…). The account of belief I propose (…) takes belief to have a distinctive nature, which allows us to set this attitude apart from other doxastic attitudes and from the bedfellows. (Engel 2018: 297)

Another way of seeing the abovementioned cases is to think that the attitudes involved (acceptance, and so forth) are varieties of belief. That is, that they are different forms a belief could take, or different aspects of believing. Rorty (1983) puts forward an idea along these lines:

> The phenomena standardly classified together as *believing* are in fact quite various and diverse. The propositional content that we detach as *the belief* occurs only as an isolatable aspect of a complex series of actions and activities, many of them habitual: attending, focusing, seeing as..., classifying, describing as...". (Rorty 1983: 181)

According to Rorty, we use the same word ("believing") to refer to a rather broad and heterogeneous variety of attitudes one could have toward propositions, that is, a variety of doxastic attitudes. In this way of viewing things, maybe the agents in the abovementioned cases do have beliefs towards the relevant propositions; but the point is that those beliefs are only intelligible against the background of a person's undertaking certain attitudes that are not beliefs, but that have an epistemic relevance, nonetheless. For instance, maybe the Zoologist does have a belief that his current spot is a must vantage point to behold wolves, but just by virtue of having guessed, through weighting epistemic and practical considerations.

Be that as it may, what is remarkable about these cases is that the attitudes involved are not quite the same thing as full-blown beliefs. They have a different etiology and a different normativities, in spite of having something in common with full-blown beliefs. The agents in those cases decide to provisionally take something as being the case for the sake of fulfilling some practical purpose, and in this they exercise some form of doxastic control. But, contrary to what happens when proper, full-blown beliefs are at stake, they do not undertake proper commitment to the

proposition being *true*, so that, if pushed into offering reasons (epistemic reasons) to justify those attitudes, their possibilities of so doing would be limited. Notwithstanding, it would still be possible to offer some, because there are at least some facts that speak in favour of endorsing the propositions in question.

The common ground underlying both ways of viewing the cases (Rorty's and Engel and others') is that there is a variety of forms of endorsement, or of epistemic pro-attitudes, with different etiologies, normativities and functions, playing different roles within our intellectual lives. This is the core of *doxastic pluralism*, the doctrine that human cognition comprises and depends on both full-blown beliefs and other sorts of belief-like attitudes towards propositions.

Some of the attitudes one can have towards propositions can be said to be "belief's cousins" (Engel 2018), in the sense that they bear some similarity to beliefs, or are based on beliefs, although distinct, in some respect. They may originate from different epistemic sources and bear different relations to knowledge (Hughes and Sims, 1997; Langdon, 2013). They can be held with different levels of conviction, ranging from very high, as in the case of things taken to be self-evident, such as basic physical laws; to relatively uncertain, as in the case of unfamiliar topics (Peters et al., 2004). They can be more or less steadfast, suffer more or less influence from one's affective states and epistemic considerations, and be more or less long-lasting (Young et al., 2003; Bisiach et al., 1991; Connors and Coltheart, 2011).

Here are some of those attitudes that have already been object of debate among philosophers and psychologists. There are "tacit beliefs" (Lycan 1985), whose content is a proposition about which the agent has never really thought about, but is disposed to believe when told. There are "subdoxastic states" (Stich 1978), which are "informations, and perhaps representations, processed in our perceptual or memory systems when we perform various cognitive tasks (…) but which are not properly believed" (Engel 2018: 299). There are "pathological beliefs", such as the Cotard's delusion, in which a person "believes" she is dead (Davies & Stone 1992; Bayne & Pacherie 2005). There are "feelings of knowing", which include feelings of familiarity, or "déjà-vus" (Koriat 2005). There are the so-called "aliefs" (Gendler 2008), which include various emotion-induced feelings and representations that influence our inferences.

Apart from those, there are also attitudes such as expecting, noticing, having intentions, regretting (Engel 2018), which are propositional attitudes; as well as

inquiring, investigating, wondering, curiosity and suspension of judgment, or agnosticism, which are "interrogative attitudes" (Friedman 2018). They resemble propositional attitudes, except that they are states and processes whose contents are questions, rather than that-clauses. Those are important because they play a regulative role in reasoning.

Now, what reasons do we have to think that doxastic pluralism is true? The main argument in favour of this doctrine stems from appeal to experience, common sense intuitions and to our ordinary ways of describing things. For instance, the distinction between acceptance and full-blown belief derives much of its strength from the observation that we basically only refer to a situation using the former when the situation involves a decision to not push inquiry further. There are cases in which one accepts a proposition without believing it, as well as cases in which one believes a proposition without accepting it. Van Fraassen (1980) has argued that acceptance is common in science: the scientist often does not think that some particular theory on which her work depends *is* the literal truth. Therefore, strictly speaking, she does not believe it. But she accepts it as the best explanation currently available nonetheless and, therefore, as an adequate ground for research.

Other cases may involve belief without acceptance. For instance, if one is about to use a ladder to climb to a height, one may check the stability of the ladder in various ways. At some point, one accepts that the ladder is stable and starts climbing it. One may genuinely believe, even before checking it, that the ladder is stable, but because so much depends on it and because caution is good general policy, one nonetheless does not accept that the ladder is stable until one has checked it more carefully (Bratman 1999).

Another argument in support of doxastic pluralism is that appealing to doxastic attitudes that allow for degrees (such as degrees of belief, or degrees of confidence, etc.) has proven useful in decision theory, game theory, and economics; and that those disciplines mirror reality because they produce reliable predictions. If that's the case, then at least some of our doxastic attitudes need to involve gradation, which full-blown belief doesn't; therefore, at least some of our doxastic attitudes have to be different from full-blown beliefs in a significant way[34].

---

[34] Standard philosophical treatments of this topic include Jeffrey (1983) and Skyrms (2000).

If the basic idea of doxastic pluralism is false or groundless, then any account of the impossibility of epistemic akrasia that relies on either Williams' or Owens' arguments will have failed, because both these arguments entail some distinction between belief and other doxastic attitudes, and, therefore, they both entail some form of doxastic pluralism. If, on the other hand, doxastic pluralism is true, than virtually any of those other doxastic attitudes could, in principle, figure in an epistemic akrasia episode, as long as they figure in our theoretical reasonings. And this, in turn, means that appealing to either Williams' or Owens' arguments will only allow one to prove that full-blown belief akrasia is impossible, but not the many other forms of epistemic akrasia that can, in principle, exist.

For other doxastic attitudes to figure in an epistemic akrasia episode, these other doxastic attitudes would have to be similar to full-blown beliefs to some extent, namely, to the extent that they involve taking certain facts as being the case and reasoning from them. But they would also need to be different from full-blown beliefs to another extent, namely, to the extent that they would have to be something over which we have at least some direct control. Can at least some of those other attitudes be similar and dissimilar to full-blown beliefs, in these respective senses?

One might object that even if we do have some sort of direct doxastic control over some of the doxastic attitudes mentioned, we do not have it over all of them. Many of the belief-like attitudes mentioned before are of such nature that it simply is not possible for them to *not be* undertaken, whenever they happen to be undertaken. We cannot control subdoxastic states, nor pathological beliefs, nor feelings of knowing, for instance. Accordingly, if one of those attitudes is involved in an alleged episode of epistemic akrasia, the episode will not be one of akrasia, after all, because C2 will not have been met.

While it is true that we lack control over many of those belief-like attitudes, attitudes such as acceptance and act-as-if, that offer alternative explanations to what is going on in Helen, Ginet and Zoologist's cases, appear to be of such nature that it is possible for them to not be undertaken. Until the surgeon showed up, Helen could have kept her judgment suspended about whether or not Jonas's surgery was going well. That is, she could have responded to the fortune teller's talk with doubt, or skepticism, instead of by forming the correspondent belief-like state. In fact, this is what we do many times, and a person in Helen's position probably agrees that it

would be the most sensible way of responding, since she is a skeptical about clairvoyance skills.

Likewise, the character in Ginet's example could as well have decided to go back and check the front door in spite of the inconvenience (what is stopping him?). When we act in a way that is similar to the character in Ginet's example, we normally know that, from a specifically epistemic viewpoint, our conduct is suboptimal. Though we might believe that it is the overall best thing to do from a practical standpoint, we know that, if asked specifically about matters of knowledge, or if the stakes were really high, we should go back and check (instead of assuming that the front door has been locked, which we don't know). And so forth.

The fact that we don't know about each and every of the belief-cousins whether we have direct doxastic control over them or not only means that not all of them will be suitable to figure in an akrasia episode. Those over which we have no control will not be. But as long as we have direct control over at least some of them, then those attitudes, whatever they turn out to be, can be undertaken at a time when they conflict with the agent's better epistemic judgment. Thus, they can figure in epistemic akrasia episodes; and in cases where this happens, the conditions for epistemic forms of akrasia are being fully met, in spite of those attitudes not being proper, full-blown belief.

Think like this: episodes of practical akrasia can occur involving practical attitudes that fall short of "proper actions", or "full-blown actions", such as dispositions to act, choices or intentions, as long as these attitudes satisfy the Control Condition, that is, as long as they are of such nature that they could as well not have been undertaken. For instance, say I have a strong disposition to have one more piece of a chocolate cake while at the same time I believe that the overall best thing to do is to refrain from so doing. In fact, I end up not having the extra piece. But the only reason why I end up not having it is because I have a broken ankle that keeps me from walking to the kitchen's table where the cake is. I did not have the extra piece, but I am guilty of weakness of the will, or akrasia, as much as I would be had I been able to consummate the action. Why is it? That's because, if only I could, I would have had it. The choice has been made to do something that conflicts with my judgment all things considered, even though I could have chosen otherwise; and the only reason why my choice wasn't actualized in action is because of external circumstances that got in my way. Internally, I've been weak willed.

This shows us that practical akrasia doesn't require "proper action", or the consummation of a "full-blown action". Why, then, would things have to be any different when it comes to the epistemic forms of this conduct? Why would they have to require full-blown belief, and only full-blown belief? They do not. And if we have the impression that they do, it's only because this have been stipulated by the ISO-based characterization of epistemic akrasia.

The bottom line here is that Williams' and Owens' arguments work for the sake of showing that the control condition for epistemic akrasia, C2, cannot be met, on account that Control over Beliefs, CB is false. But only if C2 is spelled out in the narrow way: in a way that stipulates that the epistemic attitude involved in akrasia episodes *is* proper, full-blown belief. But there is no reason to think that it has to be. That's because the control condition for practical akrasia, from which C2 is derived, is not, itself, that narrow. It's not entailed by the use of the relevant concept of an *action* that it needs to be proper, full-blown, consummated action. Or, if it is, it's just by a matter of stipulation.

If we accept that the person in the chocolate cake example above is not less akratic than she would be if she had had the means to consummate the intended action, then we have to accept that the concept of practical akrasia itself is broader, or at least less narrow, than the ISO-based characterization has made it look. Consequently, the concept of epistemic akrasia too has to be a broader one. Being broader, it is more difficult to be proven impossible – it might take more to prove its impossibility, than it would take if it was narrow, full-blown-belief-only-based. That's because now it can take many forms, involving many different doxastic attitudes; and to prove that it is impossible on account that C2 cannot be met one has to prove that C2 cannot be met in each and any of its many different forms. No arguments have been present so far that attain this.

3.6. Final Remarks

In the course of this chapter we've seen how Williams' and Owens' arguments succeed in purporting that full-blown belief akrasia is not a live possibility, by showing that the Control Condition (C2) cannot be met, on account that Control over Belief (CB) is false. They give good reason why one cannot acquire a full-blown belief that goes against evidence in any way that could be fairly called "free" or "deliberate".

These arguments, however, cannot be generalized to cover all the forms epistemic akrasia might take. That's because they cannot be extended to the full range of doxastic attitudes that make up our epistemic agency. The epistemic attitudes at stake in the akrasia episode don't have to be restricted to full-blown beliefs, as the discussion carried out at the end of last section clarify. They could be other forms of endorsement, or belief-like attitudes. If even practical akrasia of action could involve practical conducts and inclinations that fall short of a proper, full-blown, consummated action, there is no reason to think that an epistemic variant could *only* involve full-blown beliefs, and not these range of other doxastic attitudes and intellectual conducts. Accordingly, not all conceivable forms of epistemic akrasia have been ruled out as conceptually impossible.

Having shown that not all conceivable forms of epistemic akrasia have been proved conceptually impossible was of crucial importance because, without this, every discussion about epistemic akrasia or about related topics would always be under the shadow of the skeptical claim that what's being discussed doesn't really exist, or is a misconception. Therefore, proving that not all conceivable forms of epistemic akrasia are conceptually impossible rehabilitates this concept, in a sense, and frees me from the easily raisable objection that epistemic inefficacy is impossible because epistemic akrasia is impossible (insofar as the two notions are related – they are related to the extent that both purport answers to the question: "does it make sense to speak of epistemic varieties of weakness of willpower?").

However, not being conceptually impossible is one thing, and being epistemologically valuable is quite another. Just because it hasn't been proved that epistemic akrasia is a contradictory concept doesn't mean that this concept is capable of figuring in the best explanation of seemingly alien conducts that appear to bear some parallelism with weakness of the will, like the ones presented in Chapter 1. If those cases were now accounted for in terms of doxastic belief-cousins, instead of beliefs, they would be immune to the skeptics' arguments raised in this chapter and in the previous one. But would we understand them better now? It doesn't look like this would add much. Practically speaking we would be just playing with words. So even though epistemic akrasia is not impossible, approaching people's conducts in terms of epistemic akrasia doesn't take us very far, in terms of understanding (making sense of) these conducts.

This is not to say that all those cases are actually nonsensical, or that nothing in real life exists that parallels them, on the contrary – many of them seem remarkably unexceptional. They are the type of thing we've seen happening, if not to ourselves, then to those around us; and if not many times, at least a couple of. Yet, they are puzzling. But there are a few things keeping us from making progress in our goal of making sense of them.

One of those things is the fact that, in each and all of those cases, as well as in the discussion carried on throughout Chapters 2 and 3, to a certain extent, we are pressed into working from what Mele terms a "cold attribution" standpoint (1986: 221), which is, roughly, ascribing propositional attitudes to agents and zooming in. The cases are presented in a way that zooms in a person's epistemic performance and clips it at the exact place or moment when a tension between propositional attitudes (like judgment and belief) is identifiable, and then this zoomed-in clipping is tagged "akrasia". I believe this misses the point. Weakness of willpower is not that particular frame, or zoomed in bit. In other words: it's not a matter of detachable episodes.

This is easier to understand if you think of a timeline. Let's go back to the chocolate cake example, the one in which a person knows that the overall best thing to do is to not eat several pieces of cake in a row, but then proceeds to having another piece anyway. The problem with the person in this example does not confine itself to the *moment* in which she picked, say, the fourth piece of cake (assuming that more than three pieces in a row is "several" by all accounts). That small instant when this action is performed is a *part* of the problem, but this person's problem, what's wrong with her, so to speak, does not exhaust itself at that moment.

Therefore, the real problem is not simply the combination of that action with her (contrary) better judgment, co-existing at a given instant. This person was already being akratic before that moment, though it is difficult to identify precisely since when. It may have been since it first came into her mind that she shouldn't be picking the third piece, but it might as well have been since she first formed a resolution to stay away from all cakes. Likewise, the problem does not disappear from the moment she finishes to chew and swallow, because the silent disposition to iteration is always there. It's a problem within the *person*. But characterizing akrasia (and epistemic akrasia) episodically misses this point.

In other words: there is a diachronic aspect to akrasia that the typical examples (the ones from Chapter 1), as well as the discussion carried out by scholars (Chapters 2 and 3), completely overlook, because they are all too focused on the synchronic aspect, i.e., the coexistence of conflicting propositional attitudes. This deprives the concept of "epistemic akrasia" from most of the epistemological value it could otherwise have. I believe the concept of akrasia has potential epistemological value, because some of the "essence" of the problem of akrasia appears to have echoes everywhere in our lives, including our epistemic lives. Bur I also think that its value has much more to do with the diachronic dimension of the phenomenon than with its synchronic aspect.

The second thing keeping us from progressing in our goal of sense-making is the fact that, in the alleged cases of epistemic akrasia discussed, as well as in the biggest part of the discussion over the (im)possibility of epistemic akrasia, the conflict between judgment and conduct has been taken as *the* problem, the "something wrong", as if this was a bad thing in itself. That is, the mere fact that there is a conflict between one's judgment and her conduct has taken precedence over the epistemological and ethical consequences of there being such a conflict. But this is not, I suppose, the main issue with akrasia, the interesting issue that we should be focusing on.

The interesting aspect of the very notion of akrasia, the one that has the potential to be of epistemological value, is the fact that this conduct gets in the way of some important epistemic good. Ancient and medieval philosophers, as well as some of the contemporary scholars engaged in the discussion of akrasia, like Amélie Rorty (1983), acknowledged this fact: practical akrasia is above all an ethical problem, because it gets in the way of something that is important from an ethical standpoint. It has been historically regarded as a personal trait that stands in the way of a moral ideal, such as moderation, or living a virtuous life, or leading a life without sin (Snellen 2018: 13-14); and this is why it became a topic of discussion, in the first place. In the same fashion, what makes this notion worth discussing in epistemology is its potential to stand in the way of an epistemic ideal, such as knowledge, responsible inquiry, or the virtuous epistemic life.

For us to fully appreciate the ways in which akrasia is capable of generating the consequences it generates to our epistemic lives, we need to understand it as encompassing a much more comprehensive phenomenon than just particular,

episodic undertaking of a conduct against one's better judgment (even because one single episode of acting against what we judge best is unlikely to have far reaching consequences, unless, of course, what's at stake is something really big). And we need to get specific about this phenomenon, which means we need to examine the specific ways in which is presents itself, its more relevant traits, what type of consequence it generates, and so forth.

In sum, I believe forms of personal weaknesses that resemble (either superficially or not) the phenomenon of weakness of the will do show up in our intellectual lives, and they do so unexceptionally, but they are not intrinsically valuable to the epistemological discussion just by virtue of mirroring the conceptual definition of akrasia. Actually, most of the cases presented in the literature as paradigmatic examples of epistemic akrasia (the examples laid out in Chapter 1) are not, because we're left in the dark both about their diachronic aspect as about the damage consequences they provoke. We don't even know if they have a diachronic aspect, or noxious consequences. Nevertheless, they serve the important purpose of paving the way for the most substantial discussion, by allowing me to clear misunderstandings.

Stepping into what I take to be the most substantial discussion will require that a number of moves be made. First, that akrasia in the intellectual domain be seen as an epistemic *vice*. The fact that intellectual forms of akrasia amount to personal flaws that generate negative epistemic consequences, on the one hand, and the fact that the most pressing issue pertaining this very notion is precisely that those consequences get in the way of some epistemic ideal, on the other hand, both indicate that the discussion needs to be redirected towards a virtue-based approach, if it is to be profitable. This idea is not new, John Heil (1984: 70) has already suggested that intellectual forms of akrasia be viewed as epistemic vices, although he hasn't gone very far in this project himself. Chapter 4, then, will have the aim of getting clear on what an epistemic vice is, as well as on what the discipline of vice epistemology is all about.

Second, since not all cases of epistemic akrasia are rich in epistemological value, only the ones that negatively affect our epistemic ideals, another move that needs to be made is a shift of focus. I propose that we abandon the idea of "epistemic akrasia" broadly understood as practical akrasia's mirror concept and start dealing with more specific types of things in our intellectual lives that echo, or at least

appear to echo akrasia, to some extent. We grab a hold of the epistemologically more interesting cases by picking, from the many forms of epistemic weaknesses that bear resemblances to the phenomenon of weakness of the will, the ones that generate the worst epistemic consequences by virtue of being what they are, and by abandoning the ones that either don't generate any noxious consequences or that are such that we do not know whether or not they do generate some.

In other words, I'm suggesting that there is an epistemic vice that do bear similarities with the phenomenon of weakness of the will; and that concrete manifestations of this vice is what we should focus on from now on. Examples of this vice won't always resemble the cases presented in Chapter 1, because they won't always share the necessary and sufficient conditions discussed in Chapters 2 and 3. And they might even not be acknowledgeable as members of the same general class as those cases, specially because the latter have been spelled out from the "cold attribution standpoint" mentioned, which doesn't work very well for the purposes of describing vices.

To reflect this clipping, and to avoid conceptual misunderstandings, I propose using the name *epistemic inefficacy* to refer to this vice, and abandoning the umbrella-concept of epistemic akrasia[35]. Both have their roots on the idea of weakness of the will, but, contrary to "epistemic akrasia", which is a vague and episodic concept with a very diverse range of alleged examples and without a serious concern with consequences, "epistemic inefficacy" will be specific, less heterogeneous and with a bigger emphasis on detrimental consequences generated. In other words, epistemic inefficacy refers to a vice, and brings with it all the implications that the very notion of a vice carries, including the implication that the primary purpose this notion is intended to serve is the purpose of sense-making (Verstehen).

I call it "inefficacy" to be truthful to the intuition that in the biggest part of our epistemic lives we neither do nor fail to do what we *want* to (an insight Williams' and Owens' got right). Willpower is the ability to control our actions and thoughts. We

---

[35] This move, abandoning a concept in favour of one with a more specific scope, might be thought of in terms of a Carnapean process of explication. Carnap thought that "ordinary language and traditional concepts were to be overcome; they were to be replaced by better and more scientific ones" (Leitgeb & Carus 2022: §1.1). In more or less the same fashion, I submit that the notion of epistemic akrasia be overcome, that is, replaced by the notions corresponding to the (specific) vices affecting people's intellectual performance, when that performance involves the undertaking of contradictory attitudes. Epistemic inefficacy is one of those vices.

don't really exercise this ability a lot when coping with epistemic situations, even if we possess it. So in our epistemic lives, or in epistemic contexts, the notions of will and willpower are not the key concepts. Efficacy, on the other hand, is the ability to generate an adequate outcome, to produce an adequate result. Therefore, this is a more suitable concept to be used in epistemological discussions focused in evaluating people's performance as knowers, or as inquirers. Chapters 5 and 6, then, will be devoted to a detailed discussion of the vice of epistemic inefficacy.

4. EPISTEMIC VICES AND VICE EPISTEMOLOGY

> *When I examine myself, some vices appear on the surface, I almost can lay my hands upon them. Others are less distinct and harder to reach. And some are not always present, but recur in intervals. These I should call the most troublesome.*
>
> Seneca, **On the Tranquillity of Mind**, I.

Since someone first came up with the idea of examining what appears to be instances of weakness of willpower in people's intellectual behaviour, scholars have seen the topic from a myriad of perspectives. Some take them to be plain and simply cases of moral akrasia within epistemic contexts (knowing that something is the right thing to do, yet failing to do it). Others take them to be philosophical fictions: cases that have been twisted and distorted to fit a contradictory concept. Some have held that the idea of there being epistemic forms of weakness of willpower is a good thing.

That it can sometimes be the best conduct to be undertaken, under some special epistemological circumstances. Others take it that it cannot. This relatively new topic has a lot of controversy around it, and little agreement in the arena of Epistemology.

The question that should lie beneath every scholar's approach is this: what makes the idea of weakness of willpower interesting to epistemologists? Or, to be more precise, why is it that this idea poses a problem that is distinctively epistemological, as opposed to just an ethical problem? My short answer to this question is: because it can be an *epistemic vice.* Now, what is this? How is this idea to be understood? That's what this chapter is meant to tackle.

4.1. Preliminary remarks

I submit that there is an epistemic vice that affects our epistemic lives in more or less the same way in which weakness of willpower affects our ethical lives: the vice of epistemic inefficacy. Epistemic inefficacy is not the same thing as epistemic akrasia,

because it doesn't carry the same conceptual implications of this latter notion (doesn't have the same necessary and sufficient conditions, for instance). Nevertheless, its as close as you can get, if you think that weakness of willpower can affect every sphere of a person's life, including his intellectual life, but that the will is not the most important faculty being exercised in the latter.

Epistemic inefficacy is the vice of tolerating inconsistency in inquiry for longer than one should. To be sure, we, as epistemic agents, often need to put up with conflicting propositions for *some* time. For instance, say I'm inquiring into the perpetrator of a crime. If one eyewitness tells me she is sure it was person A while another eyewitness is positive that it was person B, I need to live with the conflict until I find out which of the two eyewitnesses is not telling me the truth. By the time I have this figured out, I'll be in a position to exonerate the "evidence" against the innocent suspect, thus resolving the conflict. However, I might as well fail to live up to this, by holding on to the conflict even in the presence of evidence that is jointly sufficient to resolve the plight, even by my own standards. I submit that this would make a run-of-the-mil case of epistemic inefficacy – a distinctively intellectual way of being weak.

Tolerating inconsistencies for longer than I should is something that can, of course, take many forms. I can refrain from dismissing suspect A in the presence of evidence that is jointly sufficient to establish that the perpetrator is B. Or I can dismiss both suspects and start to look for other possibilities, in the presence of evidence that actually suffices to establish A or B's involvement in the deed. Or I can dismiss A and keep B under suspicion in the *absence* of evidence that suffices to link B to the crime, even after a long and diligent search. The list goes on. It doesn't matter what precise shape the conflict takes – as long as I put up with it for longer than I should, I'm acting out of a distinctive form of weakness. Insofar as this prevents me from being a responsible inquirer as to the matter at hand, epistemic inefficacy is an epistemic vice. By saying that it is a vice I mean that it is a cognitive defect of mine, a personal trait that makes me a worse inquirer than I could otherwise be, or that causes me to perform worse than I could otherwise perform in this specific cognitive task.

Now, to account for intellectual forms of weakness of willpower in this way is to move in a different direction relatively to the main accounts the notion has been given by epistemologists in the last few decades. I argue, nonetheless, that the philosophically most interesting way to approach it is to see it this way, as an intellectual vice in the terms just stated. I take it that intellectual weakness is a

problem, and that it is a problem to epistemology, specifically, because it is something that prevents effective and responsible inquiry.

So there are three topics being brought together here, that we need to put in perspective relatively to one another: epistemic inefficacy, virtue-based epistemology (of which the notion of intellectual vice is a chief component) and epistemology of inquiry. While I'll leave the detailed profiling of the vice of epistemic inefficacy to Chapters 5-6, in the following sections I'll delve into the idea of *vice* as an article of epistemological interest; the notion of *responsible inquiry* that figures at the overall account of virtue-based epistemology I underwrite, and the challenges such an account is mainly subject to.

4.2. Virtue-based frameworks and inquiry epistemology

To clarify the precise sense in which tolerating inconsistencies for longer than one should is an intellectual vice, it is important that I locate myself against the background from which this particular conception of vice stems. We are talking about contemporary virtue-based epistemology.

Contemporary virtue-based epistemology is a collection of approaches to epistemology that have in common with one another two basic claims: first, that epistemology is a normative discipline and, second, that the primary focus of epistemic evaluation are agents, rather than doxastic states of agents. A virtue-theoretic approach to epistemology was first brought about by Ernest Sosa (1980). Since then, virtue-based epistemology has branched into different directions, depending mainly on whether authors see virtues as skills (virtue reliabilism) or as character traits (virtue responsibilism)[36].

I pose no objection against treating some of our skills and faculties, such as good memory and accurate vision, as virtues, like virtue reliabilists do. I believe, however, that not all of the virtues and vices that are of epistemological relevance are of this kind. Some are, or at least are better accounted for in terms of, traits of an agent's character. Open-mindedness, curiosity, dogmatism and gullibility are instances of this variety. Epistemic inefficacy, which is the one that interests me the most, I believe to be among the latter. So my account of virtue is more of a character-based, or responsibilist, type. (I don't deny other varieties of vices, only I am focusing

---

[36] This nomenclature is originally by Lorraine Code (1984).

on the ones that happen to be character-based, as well as in the questions they raise to epistemology.)

Within responsibilist accounts, talk of intellectual virtue and other aretaic notions (such as intellectual vice, excellence and character)[37] is something that can serve different purposes and be done in a variety of ways. In the taxonomic terminology of Jason Baehr (2012: 12), virtue-based epistemological approaches split into two broad categories depending on how they relate to traditional epistemology (*conservative* and *autonomous*), and can be of two basic types as to the reach of explanatory power attributed to aretaic notions (*strong* and *weak*).

Conservative approaches appeal to the concept of intellectual virtue as a way of addressing traditional epistemological problems[38]. According to them, virtue and other aretaic notions provide tools to solve traditional problems, which means this family of approaches dwells on traditional epistemology's main concerns, so to speak[39]. *Autonomous* approaches, in turn, focus on matters of intellectual virtue in ways that are predominantly independent from the most longstanding problems in epistemology, but that are still broadly epistemological in nature[40].

Now, both conservative and autonomous approaches can be *weak* or *strong*, depending on whether one believes that a virtue-based framework is supposed to add to or replace the classic epistemological framework. Thus, *weak conservative* and *weak autonomous* approaches propose that aretaic notions complement (but not replace) classic epistemological notions in tackling the problems each of them are meant to tackle. *Strong conservative* and *strong autonomous* approaches, in turn, take it that aretaic notions must replace classic epistemological notions entirely.

------

[37] As Kraemer (2015) notes, virtue epistemologists have by and large been more interested in intellectual virtues than in intellectual vices. It is useful to speak, more broadly, of *aretaic notions* rather than of virtue or vice alone, because virtue epistemology and vice epistemology within the responsibilist branch share this common ground view according to which virtues have their opposite counterparts, the vices. Usually virtue is seen a means between two extremes, where this two extremes are the vices. So what we have, actually, is an aretaic framework, with an aretaic vocabulary, where virtue and vice correlate.

[38] By which I mean problems such as skepticism, the nature of perception, the problem of induction, the Gettier problem, the dispute among internalists and externalists about epistemic justification, among others. For more, see Sosa (2003, Chapter 9).

[39] For examples of conservative approaches to virtue epistemology, see Zagzebski (1996), Fairweather (2001), Axtell (2007, 2008, 2010), Napier (2009) and Baehr (2012).

[40] For examples of autonomous approaches to virtue epistemology, see Kvanvig (1992), Code (1987), Hookway (2000, 2003) and Roberts and Wood (2007).

I believe that the problem posed by epistemic inefficacy and other intellectual character vices, as I understand them, is by and large independent from, and runs parallel to, virtually all the central issues within traditional epistemology. That's why, in tackling this problem, my grip on a virtue-based framework is of the *autonomous* variety. Also, I don't believe much would be gained from repelling the many attempts that have been made to address this problem through a classic epistemological framework *because* of this framework's deficiencies, but I don't think much is to be gained from insisting in them either. I believe that the most interesting way to see the problem, the one that is philosophically most compelling, is to see it as a problem of agents and their conducts, not of doxastic (momentary, or episodical) states. So I do not propose replacement, but rather a shift of focus. As such, I am in a *weak autonomous* virtue-based epistemological enterprise.

Now, the question that yearns to be answered is: why is it that intellectual character vices such as epistemic inefficacy and others pose a problem that is distinctively epistemological (rather than ethical, or a problem to psychology)? Contemporary analytic epistemology is overall interested in analysing key epistemic concepts, such as knowledge and justification, with the aim of answering questions such as "how is knowledge different from true belief?" and "what is justified belief?". In the face of this, it may not seem evident why epistemologists should be interested in intellectual character traits, in general, and in vices, in particular.

I uphold, however, that one should broaden one's view of epistemology. From a more comprehensive standpoint, we might talk of epistemology as being in the business of understanding cognitive success (which comprises, correspondingly, cognitive failure)[41]. Shifting to this more comprehensive perspective is a move worth making because notions such as knowledge and justification are embedded within a broader picture, and that's the picture of human's quest for finding things out. Here enters the second key element we're bringing together, the notion of *inquiry*. Because

---

[41] To be sure, broadening one's view of epistemology is something one can do while having different aims in mind. The move I'm suggesting has the clear aim of turning to our intellectual flaws and other defects in order to determine what not do and how not to be, while in a quest for finding things out. Other authors whom, like me, work with this aim in mind are, for instance, Alfano (2015), Battaly (2014) and Cassam (2016). But this is one aim among others. Some authors make this broadening move more with the aim of reframing epistemology as a discipline whose purpose is to promote intellectual well being. For instance, McDowell (1994) and Pritchard (2016) are more concerned with helping us overcome "anxieties" due to defective presuppositions about knowledge. These many aims need not be inconsistent.

cognitive successes and failures, more often than not, are a matter of how we engage in the activity of inquiring.

There is a (relatively small, but expressive) number of contemporary epistemologists who view the activity of inquiring as being the epicenter of our epistemic lives. Those include Hookway (1994), Friedman (2019) and Kelp (2021). Inquiry, they say, is "the attempt to find things out, to extend our knowledge by carrying out investigations directed at answering questions, and to refine our knowledge by considering questions about things we currently hold true" (Hookway 1994: 211). Their approach deals predominantly not with the evaluation or justification of information already acquired by an agent, but rather with the problem of how information is acquired in the first place. Within the conception of epistemology underwritten by them, it is an epistemological priority "to understand, guide, and improve human inquiry, with the aim of enhancing the effectiveness and responsibility of our investigations" (Cassam 2016: 161)[42].

From the perspective of epistemology of inquiry, it becomes much clearer why epistemologists should be interested in intellectual vices, in general, and in vices such as epistemic inefficacy, in particular: understanding the activity of inquiry is partly a matter of understanding how and why our quests for knowledge and understanding backfire, when they do backfire. And this, in turn, is partly a matter of grasping the influence of the various flaws we are constantly prone to in our attempts to find things out. Many vices are intellectual flaws, therefore, are a matter of epistemological concern.

Character vices, specifically, are flaws because they compromise us in our attempts to find things out. They make us into worse inquirers, by getting in the way of what Hookway calls "effective and responsible inquiry" (Hookway 2003: 198). While effectively inquiring is a matter of attaining the right results, responsibly inquiring is somewhat like responsible driving, to use Cassam's analogy: "it takes a combination of knowledge, skill and attitude" (Cassam 2016: 166). An effective and responsible inquirer is one who not only guides herself by the evidence, but also

---

[42] Though inquiry epistemology is a relatively small and recent trend, it too has strong and weak variants, if we apply to it a table similar to Baehr's. Hintikka's approach to epistemology of inquiry in Hintikka (2007), for instance, is of the strong variety: he holds the view that the entire notion of knowledge, so central to traditional epistemology, should be replaced by the concept of information. Hookway's approach in Hookway (2008), in turn, is of the weak type: he does not propose replacement, but rather a shift of perspective. According to him, we best understand concepts such as knowledge and justification by examining the role they play in the regulation of inquiries.

honours "the obligations that come with being an inquirer. These include the obligation not to be negligent and exercise due care and attention in the investigation of the matter at hand" (Cassam 2016: 166).

Under the influence of character vices, however, we are kept from honouring these obligations while carrying out our epistemic endeavours, which, in turn, have their effectiveness diminished. For instance, the vice of epistemic gullibility makes an agent prone to believe unlikely propositions that are not supported by evidence (Cassam 2016). Epistemic insouciance makes the agent unconcerned about whether something is true or false, thus leading her to not investigate matters that need to be investigated, or to investigate recklessly (Cassam 2018). Epistemic injustice wrongs another person's epistemic credentials, or her credibility as an inquirer, making it harder for knowledge to be shared (Battaly 2017). And so forth.

Epistemic inefficacy, specifically, impairs effective and responsible inquiring through making the agent hold on to conflicting propositions at times when she should seek the means to resolve the conflict, if she doesn't yet have them. The agent under this vice knows that deadlocks pose a threat to the effectiveness of inquiries but, when faced with a deadlock herself, she fails at seeking the means to resolve it, nonetheless. Like others of its kind, this vice poses a problem that is epistemological in essence: by hindering virtuous inquiring, it undermines knowledge acquisition, as well as knowledge retention and knowledge use, thus jeopardizing cognitive success.

To sum up, "when things go wrong, when our inquiries go badly, we want as inquiry epistemologists to understand why we go astray" (Cassam 2016: 174). Which means part of what we want to understand, as inquiry epistemologists, is what flaws we possess. That's ultimately the reason why vice epistemology is "a component of inquiry epistemology" (Cassam 2016: 161), and that's also the reason why epistemologists should be interested in studying vices: their types, their nature, the way they operate, and so forth.

In the next session I'll discuss in some more detail the very idea of an intellectual character vice.

4.3. Intellectual character vices

What are aretaic (virtue-based) intellectual character traits? And how is the notion of an intellectual character vice, specifically, to be understood? I submit that

intellectual character vices are traits of a person's cut-out that play a significant role in the explanation of why people intellectually (mis)behave as they do. I'll go further into this by means of a concrete example.

HURRICANE CARTER
On June 17, 1966, when two men and a woman were shot to death in the Lafayette Bar and Grill in Paterson, New Jersey, Carter was a twenty-nine-year-old professional middleweight boxer. (…) Though Carter and his friend John Artis were not identified by the surviving victims, and though both passed lie detector tests, they were eventually charged, tried and convicted of the crime by an all-white jury and sentenced to life imprisonment. (Roberts 2000: 1167-1168)

The prosecution's case against Hurricane Carter was built upon the claim that the boxer committed the slayings to revenge the murder of a black bartender by a white man that happened earlier that same evening, in the same neighbourhood. In support of this claim, they presented the contradictory testimony of two local burglars who claimed to have seen the shooters leave the crime scene; and an unspent shotgun shell together with a .32 calibre bullet, that were allegedly found by the police inside of Carter's car. It later turned out that neither the shell nor the bullet was of the same kind used in the crime, and that both the burglars who testified had actually made deals with the prosecuting attorney, and then recanted.

The case gained enormous publicity, and so did the fact that the rest of the state's evidence was incredibly weak: no weapons were ever found, no bloodstains on Carter's clothes, no fingerprints, no paraffin tests of the defendant's hands (for traces of gunpowder) were ever conducted. Also, the only surviving victim who actually saw the shooters positively identified Carter as *not* being one of the gunmen. In fact, it was disclosed to the jury as well as to the media that this witness had on more than one occasion stated that the two shooters were tall light-skinned men with middle-eastern traits. Neither Carter nor his friend fit that description. Carter in particular was short and black, with very dark complexion.

In the end, however, the jury was not persuaded by any of the evidence (or its lack thereof). They apparently just went with the prosecution's narrative. It was not until 1985 that Carter's wrongful conviction was overturned by a federal court, after several appeals. Yet, some people in and out of the criminal justice system believed then, and some believe even to this day, that he was in fact guilty.

Now the question one must ask is: why would someone genuinely believe something like this, if the evidence linking Hurricane Carter to the crime was so weak? In a way, to ask why someone believes *p* is to ask for her reasons to believe *p*. People who believe that Carter was guilty could say that a prime reason was Carter's being close to the scene in a car that happens to match the getaway car's alleged colour. After all, his car was stopped by the police in the same street of the Lafayette Bar ten minutes past the time of the crime. "He being there is simply too suspicious. Why would he be there, if he had nothing to do with it..?" They might also point fingers at Carter's choleric personality, Carter's past criminal records, or even Carter's political views. "He was an admirer of Malcolm X who preached the use of violence. He was very much the kind of guy who would shoot white people in a bar", they may contend.

This explanation, however, only gets us halfway. Part of the problem is that these reasons are bad reasons. All of the abovementioned facts might well be veridical, yet none of them actually links Hurricane Carter to *that* particular incident. The case built by the prosecution can be conclusively refuted and the exculpatory evidence is overwhelming. So the problem with these people's account of what happened back in 1966 is that it is baseless and false, and in this sense it does not provide us with a complete explanation of why is it that they believe as they do. If we want to understand this, then merely knowing what these people call their "reasons" is not enough.

The other pieces that appear to be missing from a more inclusive explanation of their intellectual conduct are traits of the structure (say, racism and police corruption[43]), sub-personal factors (say, confirmation bias[44]), and personal traits (for instance, tunnel vision: some people have a tendency to cling to the first hypothesis envisaged and undervalue alternatives, even in the presence of counter evidence).

As Cassam (2019: Chapter 2) points out, it is extremely difficult to decide whether a particular outcome is better explained in personal, sub-personal or in

---

[43] It's understood that the United States in the mid 1960's is a society marked by the prevalence of racial feuds, that all-white juries are twice as likely to convict innocent black people, etc. For more on structural racism in the U.S., see Lawrence & Keleher (2004).

[44] Errors in information processing and decision making are a well established scientific fact, the so-called cognitive biases. Confirmation bias, specifically, is the propensity to give special treatment to evidence that confirms one's pre-existing beliefs. Such biases are predictable, universal, mostly unconscious, and not person-specific, meaning we all have them. For more on cognitive biases, see Nickerson (1998).

structural terms, or which of these factors plays the most prominent role in each particular case. What is not that difficult to see is that all three types of factors tell something that helps us make sense of why some people happen to mistakenly believe that Hurricane Carter is guilty of the Lafayette Bar shootings. For, arguably, while people are under the same superstructure and while sub-personal factors affect us all, yet it is just some people, not all of them, that make up terrible, as well as great, inquirers. And this is what creates an opening for the notion of *intellectual vices* and *virtues* (personal traits par excellence) to come along: people who are convinced of Hurricane Carter's guilt are not the most exemplar intellectual agents. They do not excel in the art of inquiry and they do not act responsibly either in evaluating evidence or in tying it to conclusions.

Responsible inquirers have the ability to recognize whether something is an answer to a given question or to a slightly different question (Hookway 2008). So they know that being close to a crime scene and being the one who pulled the trigger are two different things. They also have a good sense of when they are in danger of jumping to conclusions. They understand the importance of physical evidence, as well as of the significance of extensively searching for this kind of evidence and finding none. They know that when someone is promised money to testify, this testimony might not be as trustworthy as it would have been had the person not been granted any advantage. And so forth.

People who are convinced of Hurricane Carter's guilt, however, seem to either not understand or resist complying with these. In forming their view, they rely on a sloppy narrative and often times on gossiping, prejudiced and even conspiratory acquaintances. They fail at reading evidence (or its lack). They slip to conclusions. All of these facts tell us something about these people, about the kind of people that they are, and this, in turn, helps us understand why they believe as they do. In short, these facts tell us about some of these people's intellectual traits, which are detrimental traits, therefore, vices. (Actually, it is not just people who believe in Carter's guilt that possess traits like these. Many of those who protested on the streets proclaiming his innocence also do: they closed their mind upon Carter's having been falsely accused and falsely tried from the very beginning without ever wanting to know in detail what the prosecution had. Rubin Carter's story and its aftermath make up a nice case to illustrate this discussion because they're nothing short of a festival of intellectual vices, on both sides.)

So if we ask again "why is it that someone genuinely believes that *p* when the case for *p* is so weak..?", besides the explanation that points to the reasons these people claim to have, and besides structural and sub-personal explanations, we might as well envisage an explanation that appeals to their personal traits. This intellectual character-based explanation might be, for example, that people believe in Hurricane Carter's guilt *because* they are gullible, in a sense (they attach too much credibility to the prosecution's narrative when such degree of credibility is unmerited). Also *because* they are closed-minded (have tunnel vision). Also *because* they are insouciant (have too little regard for the truth, or for the idea that there is a need for taking the appropriate steps in order to seek the truth). Also *because* they are prejudiced (tend to disregard people perceived as dissimilar, as well as their narratives). And also, maybe, because they are not strong-willed enough to do the right thing (they put up with contradictory information that they should resolve).

To sum up, according to the conception that is being underwritten here, intellectual virtue-based character traits are personal traits of agents that are invoked to partially account for the way in which people think and reason, and for how they conduct themselves regarding matters they are trying, or should be trying, to find out. "They are habits or styles of thought or inquiry (…). They are distinctive ways of seeking out and evaluating evidence, and assessing the plausibility of explanatory hypotheses" (Cassam 2016: 164).

Rationalizing explanations, structural and sub-personal explanations, and intellectual character-based explanations are all limited (usually, none of them can account for every aspect of a story on its own); and they need not be inconsistent. They might add up to one another, thus deepening and widening our understanding of why people act the way they do, why they believe certain things as well as of why and in what sense their beliefs are "unjustified". Also, as Cassam (2016: 163) remarks, they are related. The reasons a person gives for her belief that *p* only strike her *as reasons* because she is gullible, cynical, prejudiced, etc., which, in turn, only comes to be because the person is within the bounds of a certain structure. If the people we are talking about were not prejudiced, for instance, they wouldn't take Carter's admiration for Malcolm X as a reason for believing he is guilty.

Now, the account I've presented faces mainly four types of challenges. Two of them are internal to virtue-based epistemology, while the other two are external. In

the next two sections I'll present and discuss these four objections, as well as my prospects for dealing with them.

4.4. Conceptual challenges

The conception I've been advocating for is, broadly speaking, an expression of the view commonly referred to as *consequentialism* within virtue-based epistemology. Roughly, virtue consequentialism is the view that what turns virtues and vices into what they are, essentially, are the way these traits impact knowledge acquisition, or cognitive success, broadly understood. This view is to be contrasted with virtue *motivationalism*, according to which what makes virtues virtuous and vices vicious is not a matter of their consequences, but rather a matter of the agents' internal motivations and dispositions. Insofar as my view is based on the idea that intellectual vices are vicious due to their negatively impacting effective and responsible inquiry (which is an attempt to gain knowledge), it is a form of vice consequentialism.

Notwithstanding, it differs from standard vice consequentialism in some important respects. As such, it might face challenges from standard vice consequentialists. But it might receive criticism from motivational approaches too, inasmuch as motivationalists starkly disagree with me about what makes intellectual vices vicious. These are the two main internal sources of criticism to my framework, which I'll tackle first.

The external sources of criticism, in turn, both come from scientifically-driven skepticism about the idea of character (situationism and anti-globalism). Skeptics about character state, roughly, that the very idea of character and character trait is a misconception, which I take to be a more serious attack on theories such as mine. Therefore, the last and biggest section is devoted to dealing with them.

a. Motivationalism

Virtue-based epistemologists eventually disagree with one another about what, exactly, makes a vice vicious. While authors such as Swank (2000), Cassam (2016), Crerar (2017) and myself submit that what makes an intellectual vice vicious is mainly a matter of its consequences, others, such as Zagzebksi (1996), Baehr (2010) and Tanesini (2018) think that it is mainly a matter of the agent's psychology. These views are known as *vice consequentalism* and *vice motivationalism*, respectively.

The quarrel between consequentialists and motivationalists is broader than their grip on intellectual vice – it originally stems from the way they see virtue, in general. Consequentialists see virtue as being ultimately a matter of attaining good effects, or producing good results. Accordingly, qualities are as valuable as they are a good means to deliver relevant results. Success rate need not be absolute, but it must be reliable. So from a virtue consequentialist's standpoint, results are necessary for virtue, and also sufficient: any quality that reliably delivers good results counts as a virtue, whether it is a hardwired capacity, an acquired skill or a character trait. Good motives may of course accompany any quality, but they are not required in order for this quality to be a virtue.

It follows from this view, to use Battaly's examples, that a venture capitalist who consistently succeeds in helping others via charitable donations has the virtue of benevolence, even if he is solely motivated by the selfish concern of paying less taxes in the future. Likewise, a student who reliably arrives at true beliefs by exercising logical skills has epistemic virtues, even if her only motivation is to get good grades (Battaly 2014: 52-53).

Motivationalists, on the other hand, contend that simply attaining good results does not make a person excellent. It also matters *why* one attains, or tries to attain, those results. In short, having the proper motives matter. A reliable success rate in attaining good results does not suffice, and may not even be required. Therefore, the venture capitalist from the example above does not have the virtue of benevolence, neither the student who only cares about getting good grades (but not about truth) has epistemic virtues. Motivationalists disagree with one another as to whether having the proper motives is sufficient for virtue, but they agree as to its being necessary.

When it comes to intellectual vice, however, things get a little more complicated. Consequentialists stick to their appeal to results: they believe that what turns qualities into intellectual vices, essentially, is their resulting in a state of affairs that precludes or erodes some variety of epistemic good (be it true belief, knowledge, understanding, wisdom, or whatever other sort of distinctively epistemic good). Motivationalists, in turn, claim that it is not a quality's delivering bad results in terms of one of these goods' acquisition or retention that turns it into a vice, but rather its involving a *bad motivational component*.

Now, what exactly is it, for a vice, to have a "bad motivational component"? Motivationalists usually think that a good motivational component, of the kind that is crucial for intellectual virtue, is some kind of desire for knowledge (Zagzebksi 1996). Insofar as they understand vice as an inversion, or mirror image of virtue, they believe that intellectual vices are distinctively marked either by the absence of a desire for knowledge (Baehr 2010: 209), or by the presence of a motivation "to oppose, antagonize or avoid things that are epistemically good in themselves" (Tanesini 2018; Battaly 2016: 106).

If motivationalists' way is the right way to see things, then me explaining why, for instance, some people mistakenly believe in Hurricane Carter's guilt by appealing to what I've been calling their intellectual vices cease to look much enticing, because at no point I made reference to these people's *wanting* to oppose or antagonize knowledge, or to their lacking the desire to acquire knowledge. So unless I'm willing to add to my story that these people are ill-motivated in some way, a motivationalist will be in her right to turn to me and say that my explanation explains nothing.

The problem is: this would be quite bold an addition to make, and I'm not sure whether it would be better or worse on balance. I could say, for instance, that the problem with people who believe in Hurricane Carter's guilt is that they want to get a conviction, or what they call "justice", a lot more than they want to get knowledge of the truth. So, in their inquiring, they are motivated by the wrong desire. While this could of course be true, I think this explanation leaves aside some important issues regarding those people's intellectual behaviour. For even if they are mainly driven by a desire for justice, this does not come out as their positively wanting to oppose, antagonize, null or avoid the distinctively epistemic good, namely, knowledge of the truth. After all, many people want justice, and yet, in their quest for justice, they play it by the book of responsible inquiry – which the Hurricane Carter's guilt believers do not. I think that these people want justice and want the truth. So the problem is not that they oppose knowledge or despise it, but rather that the particular way they proceed in trying to get knowledge is bad. And it is bad because of the effects it has on responsible and effectively proceeding.

To be sure, I don't believe motivationalists are altogether wrong in highlighting the role played by defects of motivation in vicious conduct. Many vices indeed involve some sort of defect of motivation as one of their essential components. For instance, epistemic malevolence and epistemic self-indulgence are examples of vices in which

the agent is distinctively ill-motivated. Epistemic malevolence is defined by opposition to knowledge as such, or opposition to other people's shares in it (Baehr 2010: 192). The epistemically malevolent individual either believes knowledge is in fact a bad thing, or resents other people having access to it. As such, this individual is driven by the wrong motives.

Intellectual self-indulgence, likewise, is marked by the absence of a proper desire for knowledge. The epistemically self-indulgent person is one who consistently pursues, consumes or enjoys inappropriate epistemic objects. For, arguably, not all epistemic goods are equally valuable. "Trivial truths about sports, or the whereabouts of celebrities, are less valuable than truths about science or the world economy" (Battaly 2014: 68). A person who spends most of her resources in the pursuit of these low-value goods is an epistemically self-indulgent individual. It is clear that the problem with such an individual is her having the wrong motivations.

The point, however, where I believe motivationalists got it wrong is in their taking bad motivational components as being a necessary feature of intellectual vice per se, or a feature of *all* vices. That's because many vices either do not have such component, or only have it in a contingent way. It is weird to think of, for instance, gullible and close-minded people as either lacking a desire for knowledge or possessing a desire to oppose it. These people aren't ill-motivated or non-motivated towards knowledge. On the contrary, more often than not they are curious, they want to find things out. Think of the person who believes a conspiracy theory and spends considerable time searching for information that she believes will add up to solving the "puzzle". She wants to deepen her understanding of the "truth". The problem with these people is not that they are not curious, but rather that they give their curiosity a bad use, meaning the particular way they go about trying to fulfil it is bad.

Many other traits that vice epistemologists include in their inventories of intellectual vices have more or less the same outfit. Take, for instance, epistemic prejudice and most forms of epistemic injustice. People who possess these traits want to have knowledge, but they fail at seeing certain individuals as being potential sources of it. Take, to make the examples even more abundant, arrogance. Arrogant people aren't either indifferent to knowledge acquisition or set up to go in the opposite way – they want to have knowledge, only they fail at recognizing that they themselves make mistakes (at times in which they are indeed mistaken) in their pursuit of it.

So being ill-motivated or not-at-all motivated towards knowledge isn't a necessary condition of vice. Part of the reason why motivationalists believe otherwise is that they erroneously see vice as a perfect inversion, or mirror image, of virtue. This is what Charlie Crerar (2017) calls *the inversion thesis*. Roughly, it's the thesis that, in a range of theoretically significant ways, virtue and vice are straightforward opposites. I agree with Crerar in that approaches to vice that are informed by this thesis, such as the motivational approach, cannot accommodate the full range of intuitive and important vice cases.

b. Standard vice consequentialism

As we've just seen, the view I've been defending is a form of vice consequentialism. But it is not standard vice consequentialism. Standard vice consequentialism states that there is a precise sense in which what makes a virtue virtuous is a matter of its consequences – it says that virtues are truth-conductive. So there is a particular sort of consequence a trait must generate in order for it to be a virtue: it must positively impact the agent's ratio of true to false beliefs, and it must do so systematically. That is to say, it must reliably produce true beliefs. Likewise, from a standard vice consequentialism standpoint, what makes vices vicious is their being truth-obstructive. A trait must reliably produce false beliefs in order for it to be a vice.

On my account, however, intellectual virtues and vices are not required to generate this specific consequence in order for them to be delineated as virtues and vices. The consequences that matter to me are the consequences a trait generates for effective and responsible inquiry, rather than the way it impacts the simple ratio of true to false beliefs[45]. Virtues are traits that abet effective and responsible inquiry, they enhance one's ability or disposition to meet her obligations as an inquirer. Vices, in turn, are traits that undermine, rather than enhance, one's such ability, or disposition.

So here comes the quarrel: the standard vice consequentialist might turn to me and say: traits whose consequences enhance the degree to which an inquiry is conducted in a responsible way are not real virtues, because "being a responsible inquirer" is no guarantee of effectiveness, that is, of truth. A scientist might spend

---

[45] Of course part of what makes up an effective inquiry is that you acquire a considerable number of true beliefs. But you can acquire a fair amount of true beliefs following irresponsible inquiring, such as, for instance, when you let wishful thinking shape your conclusions and for a matter of luck you get things right.

decades working some scientific question out only to end up with the wrong answers, because some of her underlying assumptions were false – in spite of her conducting herself within decent scientific parameters[46].

Likewise, one can conduct oneself in ways that are far from ideal, in terms of meeting one's "obligations as an inquirer", and yet she can systematically attain truth. Take, for instance, the chicken sexers (Brandom 1998), alluded to in Chapter 1. They don't know on what basis they make their decisions. This could hardly be deemed a responsible way to try to find things out, so according to accounts like mine, chicken sexers should be charged with possessing some form of epistemic vice – which doesn't make sense at all, insofar as they are indeed reliably attaining truth. What they have, the standard vice consequentialist might contend, is a variety of epistemic virtue, not vice.

So the standard vice consequentialist would be basically claiming that my account of intellectual virtues and vices is flawed because it allows for mistaking virtues for vices, and vice-versa. Some of the traits that in my account should be classified as vices can actually sometimes enhance truth acquisition, rather than block it; and some of the traits that I regard as virtues can actually sometimes obstruct truth, rather than lead to it.

I don't think, however, that this is a correct way of putting things. Yes, I believe that the expert chicken sexer reliably attains truth, and I also believe that she is intellectually reckless, in a sense. But I believe she attains truth *despite*, rather than *because of*, her intellectual recklessness. If she knew and could explain the basis on which she makes her discriminations, this would by no means make her into a worse epistemic agent, on the contrary, it would only add up to her qualities as a knower. Lucky enough for her, not being able to provide such an explanation does not do her a great harm (it does not lead her, for instance, to make incorrect discriminations of chicks' sexes)[47].

---

[46] For a nice historical tale of this sort, take one of the great scientific minds behind the so called "chemical revolution" in the 18th Century, Joseph Priestley (1733–1804). Priestley is credited as the man who first reported the discovery of oxygen (O_2). Unfortunately, however, he never fully understood the nature of his discovery and ended up drawing a good number of false conclusions out of it. He did this because he was wedded to a set of erroneous underlying assumptions: the so-called "phlogiston theory" (which was the dominant theory about the nature of combustion at the time), in which oxygen as an (what is today called) oxidizing gas had no place to be. For more on Priestley, see West (2014).

[47] Though it causes no harm to *her*, I'm still inclined to say it causes some harm if we think in terms of an epistemic community – it makes it harder for her to share knowledge with other people who want to become professional chicken sexers.

But there are countless other scenarios in which not being able to explain on what basis one makes her endorsements is a bad thing: it makes it harder for her to teach what she knows, for instance, so it obstructs the transmission of knowledge. Also, it functions as a red flag, that is, an indication (a pretty reliable one, I suppose) that the person is at risk of believing unlikely propositions that are not supported by evidence, that is, of being gullible, prejudiced, or mistaken in some other form. A medical doctor who is not capable of explaining why he thinks that his patient has such and such disease, for instance, hints to us that it is likely that he doesn't know what he is talking about.

For similar reasons, the unfortunate scientist from the example above fails to attain truth despite, rather than because of, her intellectual carefulness and thoroughness. If she was not careful and thorough in her epistemic dealings, this would by no means turn her into a better epistemic agent, neither in my terms or the standard vice consequentialist's terms (i.e., it would not increase her likelihood of attaining truth). Unfortunately for her, possessing intellectual virtues and exercising them throughout an epistemic endeavour sometimes isn't sufficient for us to attain the correct result, and this is what happened to this scientist; but it was not these traits that caused her to be unsuccessful either.

Needless to say, in a myriad of other scenarios, possessing and exercising intellectual virtues, as they were described by me, contributes to attaining the right result instead of the wrong one, and it is in this sense that virtues are said to help explain cognitive success. But never did I claim that all cognitive successes need to be explainable by reference to the agent's virtues, as I understand them. Picture a detective, for instance, sitting still at a party doing nothing, when all of a sudden someone bumps into him accidentally – someone that happens to possess the exact information that is missing from the detective's files. The two of them start chatting and the stranger eventually discloses to the detective the crucial piece of information, which decisively helps him solve an important criminal case that would otherwise go cold[48]. In this scenario, success is much more due to the detective's luck than to his personal merits. Likewise, not all cases of cognitive unsuccess are explainable by the agent's vices, as I conceive of vices. Many of those cases are much more a result of

---

[48] To see many veridical stories of criminal cases being solved after detectives come across crucial information completely by a matter of chance, watch *Cold Case Files*, the 1999 TV show/docuseries produced by A&E Network cable channel, rebooted on Netflix in 2017.

the agent's bad luck than of her personal flaws (the case of the unfortunate scientist being an example of this). My account precludes no such cases.

So the chicken sexer example and the unfortunate scientist example do not show that I'm mistaking virtues for vices, and vice-versa, in my overall account. What they show is, rather, that abiding by responsibility standards do not guarantee attaining truth, as well as that violating such standards do not necessarily block access to truth. The question that is really worth asking, thus, is why follow them. Why following these standards would be of some use, if they are no guarantee of truth?

My answer to this question jots down at the problem of what is to be considered "the right result" of a cognitive enterprise, after all, or "the distinctive epistemic good" at stake. Though responsibility standards do not guarantee attaining truth, they are important for knowledge. I believe that we, as epistemic agents, should aim not at truth, but rather at more comprehensive epistemic goods, like knowledge, understanding and solutions to problems. The ideally effective and responsible inquiry, in my view, is one that produces one of those goods, rather than mere true beliefs. That's because, arguably, those goods are more valuable than mere true belief. And, in order to attain them, one must be able to justify her beliefs, as well as the decisions she constantly needs to be making through the course of her investigation. What responsibility standards offer is a means for the agent to be able to justify such beliefs and decisions and, consequently, a means for her to acquire and make the most appropriate beliefs and decisions. Not having this sort of justification renders one a less effective inquirer than she would otherwise be, even if her beliefs turn out to be true.

Accordingly, I take it that standard vice consequentialists and vice obstructivists, like myself, are possibly speaking of different varieties of virtues and vices. The standard vice consequentialists is concerned with traits that impact the acquisition of true beliefs, while I'm concerned with the ones that impact the acquisition of other epistemic goods, like knowledge, understanding and solutions to problems.

4.5. Skepticism about character

Some authors hold the view that character traits as ordinarily conceived are either non-existent (Doris 1998, 2002; and Harman 1999, 2000) or explanatorily empty (Alfano 2012; Carter & Pritchard 2015). If these authors are right, there is no

reason to appeal to character traits such as virtues and vices in seeking a more inclusive account of why people intellectually behave the way they do. Intellectual vices would be a form of philosophical fiction. How should vice epistemologists like me deal with skepticism about character?

a. Situationism

The primary motivation for skepticism about character stems from the view known as *situationism*. In short, it is the view that how people behave is better explained by reference to traits of the situation than by reference to personal traits, as we ordinarily think of them. Situationists usually do not deny that character and character traits exist – what they reject is the idea that character traits have as much explanatory power as virtue ethicists and virtue epistemologists believe. Though it is most usually brought to the table to be used against virtue ethics, recently it has also been used to criticize virtue epistemology as well.

Writers in this tradition draw inspiration from findings in neuroscience and social psychology. They contend that empirical sciences provide us with data that suffices to establish that "how a subject responds to a situation turns out to be in fact highly sensitive to specific features of the situation, including ones of which the subject may be consciously unaware" (Carter & Pritchard 2015: 168). Intellectual virtues and vices, accordingly, lack explanatory power over epistemic conduct, because behaviour is, as Doris writes, "extraordinarily sensitive to variation in circumstance" (Doris 2002: 2). Such variations in circumstance include mood swings, ambient sounds and smells, social distancing or proximity, the weather, and presence or absence of rewards, such as candy and the right to watch to comedy films. In making this case, these authors are endorsing a version of *epistemic situationism*: the view that factors such as the abovementioned influence one's epistemic performance a lot more than whatever individual traits the person might possess. In fact, these factors are so influential that personal traits become devoid of any explanatory power whatsoever.

One garden-variety empirical test that is usually presented in support of epistemic situationism is the Duncker's test (Duncker 1945). Roughly, the experiment involves assigning the participant the task of fixing a lighted candle to a vertical cork board in such a way that no wax drips on the floor. To this, the participant is given access to four items: the candle, a cardboard box full of thumbtacks and a book of

matches. The key to successfully completing the task is to empty the thumbtacks box and place the candle inside, and then fix the box with the candle in it on the cork board using the thumbtacks, and finally light the candle using a match. This way all the wax pouring from the candle will be deposited in the box and none of it will drip on the floor.

The problem is that a very expressive number of subjects tested in this experiment have extreme difficulty in performing the task, simply because they see the box as a mere container for the thumbtacks, so that the idea that they could use it as an item in itself doesn't even come to mind. Scientists know this because in the control group, where the same items are delivered to the participant, but with the box already emptied (thumbtacks delivered outside of it), the success rate is higher – participants quickly realize that the box is to be used, and in what way.

What Ducker-like experiments are taken to establish is that merely situational factors, such as differences in the way items are presented, play a satisfactory explanatory role in terms of our successful or unsuccessful cognitive performances, in spite of factors of this sort being epistemically irrelevant, so that no appeal to personal traits is needed.

In calling attention to scientific findings of this sort, epistemic situationists go on to criticize virtue-based epistemologists' tendency to exaggerate the extent to which character traits are to be recruited for explanations of people's cognitive performance and, especially, for predictions of how they will perform in future situations. According to them, character factors are being overrated, while situational factors are being underrated in terms of such explanations and predictions, which is what they call the "fundamental attribution error" (Harman 1999).

So the question is: are virtue-based epistemologists like me making a version of this error? Could it be that the epistemic performance of people who believe in Hurricane Carter's guilt, for example, is better explained by situational factors than by their character traits? Because if it could, then there would be no point in ascribing such traits to those people, once they would serve neither to explain nor predict their epistemic conduct.

Besides the obvious fact that not all the cases of bad epistemic performance are embedded within ducker-like problem-solving scenarios (where the solution is known in advance by an external judge), the first thing that need to be made clear is that many of the personal traits that are said to enhance one's epistemic performance

are ideals to which "we might aspire but few of us attain other than in highly attenuated forms" (Cassam 2016: 170). That is to say, the traits to which virtue epistemologists attribute some explanatory power over cognitive success are *rare* traits. We can explain the successful epistemic conduct of excellent agents by pointing to such traits (their intellectual virtues). But, since there are very few excellent agents out there, this strategy does not allow us to account for the epistemic conduct of most of the people, simply because most of the people do not possess the relevant traits. So the situationists are right in that virtues might not have as much explanatory power over cognitive success as it has been attributed – only not for the reason they present.

The same does not apply, however, to intellectual vices, and here we have an interesting asymmetry. Many of the intellectual personal traits that worsen people's epistemic performance are distressingly common. Traits like gullibility and closed-mindedness, for instance, are all too prevalent. I would not be exaggerating if I supposed that every single one of us know at least one person (if not many) who conduct herself, as to some matter, in a way that is very similar to the way the people we discussed conduct themselves regarding the matter of Hurricane Carter's guilt. This of course does not equal saying that other factors aren't influential, including situational factors, but gullible and closed-minded people do not cease to behave the way they do depending on ambient sounds, the presence of sugar, etc. Part of the problem is that they consistently behave in this way, at least regarding some subject matters.

So cognitive failure, unlike cognitive success, cannot be properly accounted for without reference to personal traits. That's why even if situational factors are as relevant as situationists claim, the role and status of virtue-based epistemological enterprises like mine prevail, inasmuch as this sort of enterprise is defined more as a (practical) matter of understanding how epistemic vices work and how it would be possible to mitigate them than as a (theoretical) matter of explaining excellent conduct by appealing to personal excellences[49].

---

[49] Some authors claim that research on heuristics and related cognitive biases shows that humans are better understood as agents that manifest natural epistemic flaws (i.e., that type failure that is made explicit in the Duncker test is actually one of those); so that the task of becoming a better epistemic agent is, at the end of the day, a matter of cultivating means to circumvent these flaws, or to overcome them. See, for instance, Roberts and West (2015) and Samuelson and Church (2015).

Now, another thing epistemic situationists might contend is that I'm talking in very general terms about "people who possess said and said vices", or "the people who believe in Hurricane Carter's guilt", which is by and large a matter of stipulation. That is, they might contend that it is *me* who stipulated that these people believe what they believe because of such and such personal traits. But who am I to say that these people have such and such personal traits, to begin with, if I haven't, by any means, tested them?

Alas, I cannot demonstrate by means of empirical evidence that *the* people who believe in Hurricane Carter's guilt have the personal traits I ascribed to them, because no study has been conducted with a sample of this particular group of people with the aim of determining which traits they in fact possess. Though it could be very useful, no such study is forcefully necessary. That's because I can defer to a more general question: the question of whether there is empirical evidence from social psychology in support of the claim that people have individual traits that can affect their epistemic performance in cases of questionable belief. If there is this sort of empirical evidence, this would mean vice epistemology's background assumption is correct. Then my bet that people who believe in Hurricane Carter's guilt believe what they do due (at least in part) to their personal traits would no longer be like a far cry from actual (non-stipulative) cases of questionable belief. So, is there such evidence?

If we look into the literature in social psychology from recent decades, we see that there is in fact a good number of studies focusing on individual differences and reaching interesting conclusions. For instance, gullibility. A recent study conducted by social psychologists from Macquarie University (George et. al. 2020) found that something called "the gullibility scale" (Teunisse et al. 2020; see also Mericer 2017) can in fact predict behaviour. According to them, some people – namely, those who score high in this scale – are more susceptible to fall prey to scams than others. The results of these studies suggest that gullibility is measurable, plays a role in explaining individual behaviour (the behaviour of being deceived) and can make for accurate predictions.

Psychologists have also designed mechanisms to measure other individual difference traits of the sort that is relevant for my present purposes, for instance, dogmatism (Altemeyer & Hunsberger 2004) and close-mindedness (Roets & Van Hiel 2011a). Close-mindedness is especially interesting, because research suggests it

might be at the root of other cognitive defects, thus contributing to the formation of what we might call a "dogmatic mentality". The trait is measured by means of what they call a "need for closure scale". A person's need for closure is the strength of her desire for infinite knowledge on some issue. Studies found that the higher one's need for closure, the higher her likelihood to "seize and freeze" on information (de Dreu, et. al. 1999; see also Kruglanski & Webster 1996). That is, people who score high in the need for closure scale tend to seize information that reduce ambiguity and uncertainty, and subsequently freeze on it, refraining from processing further data that might jeopardize their attained certainty – which is a very close description to what we referred to earlier as "tunnel vision". So the presence of this trait (high need for closure) explains or at least helps to explain people's tunnel vision-behaviour.

Another finding concerning close-mindedness is that the more an individual possesses the trait, the higher her disposition to perceive the social world as offering threat and danger is (Perry & Sibley 2013, see also Roets & Van Hiel 2011b). That is, people tend to form social biases as a function of their need for closure. This, again, sounds pretty akin to what we've referred to as epistemic prejudice and injustice: the tendency to cast individuals perceived as different in a worse light.

As the last paragraphs left explicit, social psychologists do not talk about "intellectual character", or "intellectual character traits", like virtue-based epistemologists do – they refer to the relevant traits as "individual differences", or "individual difference traits". They have nonetheless found that individual difference traits such as the abovementioned are normally distributed (few people score super low, most score the average range, and few score super high); are relatively stable across time, and are more observable when averaged across a wide sample of occasions (McCrae 2004). This sum of characteristics suggests they are enduring individual traits that are significantly correlated with (meaning: capable of explaining) some forms of bad cognitive performance. As such, these individual difference traits just *are* character vices as virtue-based epistemology understands them.

Like I said earlier, these scientific findings do not rule out epistemic situationism, that is, they do not point towards situational factors such as sugar and ambient sounds having no bear on the acquisition of questionable beliefs or others forms of bad cognitive performance. What they do rule out is the "fundamental attribution error", the idea that personal traits are overwhelmingly less significant than virtue-based epistemologists hold – because these findings show that personal traits

are pretty significant. So in light of these scientific findings, virtue-based epistemologist's core idea that intellectual vices offer a substantial nonsituationist explanation of some of our poor epistemic performances stands.

b. Anti-globalism

Still, skeptics about character traits may come up with another complaint. They may contend that in order for the idea of a character trait to make sense, character traits need to be overall *consistent*. So for instance, if a person who believes in Hurricane Carter's guilt acquired this belief in part because of her gullibility, then she should be no less gullible while accessing other matters and acquiring other beliefs as well. Or, if someone came to the conclusion that Hurricane Carter is guilty as a result of this person having tunnel vision, then we should expect this trait to be pervasive in her epistemic conduct: she should display a tunnel-vision mentality in other domains of inquiry too. This idea – that virtues and vices must be behaviourally expressed in a wide range of trait-relevant situations, thus forming a pattern – is the main component of what is called *globalism*.

According to authors such as John Doris (2002), the general idea of character, and the one that underlies virtue-based ethics as well as virtue-based epistemology, is a globalist conception. This conception is mainly defined by the requirement for consistency just alluded. In Doris' words, "[character-]trait attribution is associated with a conditional: If a person possesses a trait, that person will exhibit trait-relevant behaviour in trait-relevant eliciting conditions" (Doris 2002: 15-16).

Notwithstanding, it is conceivable that a person acts genuinely gullibly in inquiring into the matter of Hurricane Carter's guilt, while acting not gullibly at all when inquiring into the matter of, say, which companies' shares have the best benefit-cost ratio within the stock market at a given time. (To be sure, both are trait-relevant eliciting situations: situations in which a person has the opportunity to act gullibly, so to speak). People don't systematically manifest all of their personal intellectual traits in all of the domains of inquiry of their lives. That is to say, globalism doesn't look true. In fact, Doris and other anti-globalists[50], following studies in experimental psychology conducted by Walter Mischel (1961, 1978), claim that this is

---

[50] Flanagan (1991), Harman (1999), Ross and Nisbett (1991).

empirically established: people are too inconsistent from one situation to another for global traits to be usefully invoked in the explanation of behaviour[51].

So the anti-globalists' point, to sum up is: for a certain behaviour of yours to be an instance of a genuine character trait, it needs to come out as a pattern across varied situations. As long as science shows that no such patterns arise, no genuine character traits exist. If Doris and these other anti-globalists are right, then it follows, it seems, that gullibility, as well as any of the other vices I've been discussing, cannot be a genuine thing by reference to which any of a person's epistemic conduct is explained. Anti-globalism puts the very notion of a vice *qua* personal trait on the check.

In the face of this challenge, I believe virtue-based epistemologists like me should bite the bullet, in a sense. Yes, character traits require *some* consistency in order to make sense as what they essentially are – for, if not, then behaviours that we traditionally see as expressing these traits would be no different from random or isolated occurrences, not worthy of being called a "trait". Notwithstanding, who said that all of the so-called character traits need to express themselves consistently *to the degree* the anti-globalists expect, and *in the way* alluded (that is, by forming a pattern), in order for them to be explanatorily valuable? They need not.

First, say you are a very outspoken individual. You are pretty well known by your fellows as someone who has a consolidated habit of stating your opinion out loud in a range of situations, even in the presence of people who are set up to disagree. You have a taste for polemics, and you have a high tolerance of dissent, meaning you cope well with being challenged. Yet, you do not speak your mind *every* opportunity you have. Maybe sometimes you just sit quiet and enjoy yourself watching people disagreeing with one another, even though your motivations for standing up in favor of one of them are exactly the same as your motivations for taking part in discussions in every other occasion. Also, it might as well happen, for instance, that you are a very outspoken person regarding matters of politics, economics and international relations, broadly understood, while at the same time you do not dare speak so freely about faith and religion, art, sports, sociology, etc.

---

[51] Skepticism about consistency begun in the first decades of the XX Century, with studies of honesty conducted by Hartshorne & May (1928). Taking off from these studies, Mischel's work became the standard reference. It has been presented as evidence that personality, or character, is fragmented; that people simply do not have global, consistent characters – either good ones or bad ones.

In a case like this, provided that you do not display the trait invariably, but only routinely, does it follow that you are not an outspoken individual, or that your frequent outspoken behaviour can't be explained by reference to you possessing a correspondent trait..? I do not think it does. If being outspoken a lot of the times, or more often than not, within a considerably large domain of inquiry, is well enough for you to be acknowledged by your peers as possessing the trait (meaning you don't need to express it *every* time in order to be granted the epithet), I don't see any reason for us to not say you possess the genuine trait.

The anti-globalist might object to me that while the example above sounds plausible, the idea does not generalize to other traits. Some traits require adamant consistency to make sense as character traits, which in turn makes my view look ad hock. If a husband showers his wife with gifts pretty much all the time and beats her just every once in a while, would it still be fair to say that he is a loving and caring husband? It certainly would not. Or take, for instance, honesty. If you are honest most of the time and refrain from being honest just seldom, then isn't it the most accurate description that you simply were never genuinely honest, to begin with?

I agree with the anti-globalists in that there must be limits on how much the requirement for consistency put on a trait can be flexibilized before it ceases to be ascribable, or recognizable as a personal trait. And I agree with them in that different global traits require different degrees of consistency too. But I also believe that this very point unveils another important asymmetry between virtues and vices, an asymmetry that is ultimately the reason why the anti-globalist's objection is misdirected.

The asymmetry lies in that vices are not like virtues in terms of how much consistency they require, as well as of how much inconsistency they tolerate before they cease to be ascribable. No vice, I submit, require adamant consistency; though I accept that some virtues may. Like the examples above make explicit, it takes a minimum amount of gift-giving for you to be considered a loving husband, but it only takes one single time of beating your wife for you to be considered violent and dangerous. Same with truthfulness: you must tell the truth a fair minimum of times to gain the trust of your fellows. While the issue of how many times you must tell the truth before you begin to be considered a truthful person is debatable, you likely don't need more than one slip-up to be distrusted and deemed a liar. And in order for this to happen, this slip-up can be of whatever sort, in whatever situation: whether you

are caught being dishonest at your business or at a bowling game or at the church makes no difference. The slip-up doesn't require specific conditions to deliver its noxious effects.

To understand this asymmetry, think of book II of the *Nichomaquean Ethics*, where Aristotle states that "man are good in but one way, but bad in many" (NE II, §7). Virtues, as he has them, are means between extremes, the vices. Hitting the mean (that is, attaining a particular virtue) is difficult, because there are lots of ways to get it wrong and only one way to get it right. You can get it wrong by hitting either of the extremes, or by hitting any spot in between each extreme and the mean. But you can only get it right if you hit *the* mean. That is, if you do the right thing, at the right time, for the right reasons, in the right way. So the spectrum of virtue and vice in best conceived of as like a dartboard – there is only one small bullseye, and a wide area around it. To be virtuous, you need to hit the bullseye. But since there are not *specific* spots on the board corresponding to the vices, hitting anywhere in the wide area is to miss the bullseye already, and, therefore, is to be vicious, to some degree. You don't need to follow a pattern, of always hitting at some (particular) spots.

It follows from this picture that virtue is a lot more demanding, while vice comes quite easily. Thereafter, unless you are consistently virtuous, you are vicious to some degree. Provided that being consistently virtuous is excruciatingly difficult, you likely will miss the bullseye sometimes, perhaps many times; which means you will be vicious, overall, to a significant degree. That does not mean particular vices will show up consistently in your behaviour across varied situations – but they need not.

Think of it through a concrete example. If you are presented with a collection of evidence that is jointly sufficient to establish that the likelihood of Hurricane Carter being the searched-for criminal is overwhelmingly low, there is only one way for your epistemic performance to express virtue: you must conclude that Hurricane Carter is not the searched-for criminal. Any other conclusion that you reach renders your epistemic performance at this inquiry flawed and, accordingly, vicious. Now, while there is only one right response, there is a myriad of other possible responses. You might, for instance, conclude that Hurricane Carter is the criminal, and the evidence – the "real" evidence – does show it; but the evidence that was presented to you was tainted and partial. The "real" evidence remains yet to be found. Or you might conclude that Hurricane Carter is the criminal, because if he was not the criminal, this

would mean that the Paterson Police Department is corrupt, which is impossible, hence, he has to be the criminal. Or you might conclude that Hurricane Carter is the criminal, because, you know, who knows..? Maybe he is the criminal. What difference does it make, though? And so on.

In these case scenarios, your conduct would be close-minded, gullible, and insouciant, respectively. Now here is the thing: you don't need to adopt any *specific* of these conducts, nor to repeat alike conducts in a range of other occasions, such as when you reason about other matters (thus displaying a pattern across varied type-situations) for these conducts to be genuinely pernicious. Adopting any of them, even if only at the present situation, will make you vicious to some degree – because within my account of epistemic vice, such traits are vices because they make you less effective at trying to understand the events you are trying to understand. They don't need to disturb *every* inquiry that you happen to engage in.

I personally doubt that people who adopt the abovementioned conducts do so in an isolated way – I believe this kind of person displays, or is inclined to display, the same vices across *similar* inquiry-type situations. For instance, situations in which there is another black person sitting on the dock, or another person being mistried, and one is to inquiry into their guilt. This, however, is a different component of globalism: it's what Doris calls the requirement for stability. Unlike the requirement for consistency, the requirement for stability states only that the overall idea of character demands that "character and personality traits be reliably manifested in trait-relevant behaviours over iterated trials of *similar* trait-relevant eliciting conditions" (Doris 2002: 22, italics added). In other words, vices must show up in similar situations. And the requirement for stability, Doris himself accepts, is supported by empirical evidence. It poses a significantly smaller threat to the idea of character I underwrite. It can arguably be circumvented by the same sort of empirical evidence I adduced to address situationism.

To sum up: anti-globalism doesn't represent a real matter of concern to me because my view does not commit me to an intellectual vice needing to affect all of a person's investigative practices. People might do fine when considering some matters, and display none of the vices they do display in their consideration of, say, the matter of Hurricane Carter's guilt. We can be effective and responsible thinkers about some matters and not others, "just as someone can be a careful driver but a careless cook" (Cassam 2016: 174).

4.6. Final remarks

In the course of this chapter, I hope to have been able to locate myself against the background of contemporary virtue-based epistemologies, by clarifying that the general account I underwrite is obstructivism, a weak autonomous vice epistemology of the responsibilist-consequentialist variety. I also presented and dealt with the main sources of objections this account is primarily subject to.

I also showed that my account is to be understood as a component of a broader epistemological enterprise, which is inquiry epistemology. As one of its main areas of interest, inquiry epistemology wants to understand why our inquiries go awry when they do go awry. It is for this reason that the general nature of intellectual character vices, as well as identifying and studying specific vices, are articles of great epistemological interest. My personal contribution in this arena is through describing and discussing what I take to be a much overlooked, yet quite controversial, intellectual vice, epistemic inefficacy; a task to which I will now turn.

Before we move forward, a quick word of clarification. I has been pointed to me that vice epistemology of the sort I underwrite is not really a variety of virtue-based epistemology, specially after it has been mixed up with inquiry epistemology. That's because this blended approach to epistemology doesn't have the notion of virtue as an *end*. Rather, the ends that I talk about are the responsibilities of a person as an inquirer. Therefore, the objection goes, this framework is much closer to a model of epistemic deontolgy, with virtues being mere auxiliars in the attainment of the obligations that come with being an inquirer.

I acknowledge that there is reason to this objection, insofar as I've indeed spoken about epistemic obligations, and as virtues being a means to fulfil them (rather than an end in themselves); as well as vices as being something that keeps the person from fulfilling those obligations, and so forth. Nevertheless, the defining feature of virtue-based epistemology is that it concentrates on the intellectual virtues and vices (as opposed to the evaluation of belief); and not the claim that virtues are the end of the epistemological enterprise, or the end of our lives, as epistemic agents. This might be a relevant difference between virtue-based epistemology and virtue-based ethics, from where the former derives inspiration. They are not mirror disciplines. This quote from Alvin Goldman sums this point up.

> Some proponents of 'high church' virtue epistemology might find elements of consequentialism or deontology anathema to their hopes for a distinctive, virtue-based epistemology. By 'high church' virtue epistemology, I mean a form of virtue epistemology that models itself closely after virtue ethics, which many theorists view as a rival to ethical consequentialism and deontologism. (Goldman 2001: 31)

Virtue-based ethics is presented, many times, as an alternative both to deontology and to consequentialism. As such, it proposes abandoning the ideas of an "obligation" and of a "calculus of consequences" as sources of normativity, and replaces them by the notion of virtue. As such, it doesn't absorb much from the deontological and consequentialist conceptual and normative repertoires. Most of the time, the same is not true of virtue and vice epistemologists[52].

Virtue-based epistemology is characterized by giving aretaic notions a fundamental role. This is not the same as holding that such role is teleological, or that it is axiological. As Heather Battaly remarks (1998), virtue-based epistemology *focuses* on the intellectual virtues and vices, and it does so in many ways. According to her, there are four main ways: a conceptual focus, an ontological focus, a focus on the virtues as indicators of the evaluative status of beliefs; and, finally (the default position, which is also my position) a focus on the virtues as protagonists of one's cognitive experience.

Having a conceptual focus on virtues and vices means that the *concepts* of knowledge and justified belief are analysed in terms of virtue-based notions. Having an ontological focus on virtues and vices, in turn, means that the *nature* of justification and/or knowledge is explained in terms of virtues, without the explanation of the nature of the virtues making reference to the ideas of justification or knowledge. Here, justified beliefs and knowledge themselves, rather than our concepts of them, are explained; and the virtues themselves do the explaining.

Alternatively, one can focus on virtues as indicators of justified belief and of knowledge, while denying that they are what *make* beliefs justified. On this view, virtues are broke down in terms of knowledge and justification, contrary to what happens in the previous approach. Last, virtues might be protagonists in one's philosophical works even though one is not interested in constructing a virtue theory, per se. On this view, knowledge and justified belief cannot be reduced to virtues, neither can virtues be reduced to knowledge or justified belief. Instead, they compose

---

[52] For excellent discussion on this, see Lockie (2008).

an independent part of the task of expanding our grasp of the human cognitive experience, as well as a tool for promoting our efforts to be "epistemically better" people.

A focus on the vices as protagonists of unsuccessful cognitive experiences, which is my focus at the moment, need not involve the claim that vices are explanatorily basic nor that virtue is the end of our epistemic lives, and it need not be incompatible with the idea that fulfilling one's responsibilities as an inquirer is one of those ends. In the next chapters, I'll be making the much more modest claim that vices enjoy some explanatory power over conduct. That is, that they are suitable to figure in the best explanation of certain unsuccessful cognitive experiences.

5. THE VICE OF EPISTEMIC INEFFICACY: AN INTRODUCTION

A man encountered a tiger in the forest. Unable to flee or subdue the animal by force, he chose a third option, and leapt on the tiger's back. The man knew that if he was careful and patient he could ride the tiger until it was old and weak. Then he could clutch his neck and begin to squeeze.

East Asian Parable

"Akrasia" is a greek word literally translatable as "lack of self control", or "weakness of the will". A parochial example of ordinary akrasia found in contemporary Anglo-american philosophy features a person who decides that she wants to contribute to fighting famine in Africa, but then finds that the check she intended to write to Oxfam never gets written, either because she is diverted by other concerns, or because every time she is about to write it, other expenses upsurge that make her decide to postpone her donation for yet another month (Hookway 2001: 179-180)[53]. Another famous example stares an individual who couldn't help picking one more piece of chocolate cake, in spite of knowing that she was not supposed to (Austin 1979: 198).

Discussions around cases alike are longstanding within normative ethics and the studies of practical rationality, because even though these tales seem all too familiar, many of the most popular theories in the field simply cannot escape the conclusion that akrasia is a misconception. In a similar way, in recent decades, many scholars became concerned with whether a parallel situation arises in the study of theoretical rationality, and the dominant position within those discussions was that the idea of an intellectual, or epistemic, form of weakness of willpower is a misconception.

In Chapters 2-3 I showed that whereas arguments presented so far by its skeptics fail to rule out all the conceivable forms of the phenomenon, on the one

---

[53] This example is originally by Velleman (1989: 138).

hand, the very concept of "epistemic akrasia", on the other hand, is not a very useful resource to epistemologists, when the end envisaged is sense-making (making sense of the seemingly alien conducts in which an agent appears to be undertaking conflicting epistemic attitudes). My approach to the question about intellectual forms of weaknesses of willpower, therefore, will be an unorthodox one. Stemming from a background in vice epistemology, my aim here is to purport that there is an epistemic vice that is a form of weakness, parallel to weakness of willpower in terms of its consequences. I call it epistemic inefficacy.

Epistemic inefficacy is the vice of tolerating inconsistency in inquiry for longer than one should. While not the equivalent to "epistemic akrasia", as the latter is typically characterized (by a set of sufficient and necessary conditions that mirror the conditions for episodical practical akrasia), epistemic inefficacy is as close as it gets to practical weakness of willpower, if we think two things. First, that weakness is a pervasive trait that can affect many aspects of a person's life, including her intellectual, or epistemic life. And, second, that the most important notion in epistemology is not our will, or our willpower (the ability to control thoughts and actions), but rather efficacy (the ability to bring about the right results).

5.1. Preliminary remarks

Conflicting propositions often need to be put up with for *some* time by an inquirer – the problem arises when the agent reaches the point in which she has enough evidence to resolve the conflict, and yet fails to. For instance, say I'm inquiring into whether it was Julius Caesar or someone else who burnt the library of Alexandria. I need to live with the dilemma until I eventually reach the point where I have gathered enough evidence from ancient Egyptian textbooks and other historical records and the evidence is jointly sufficient to conclusively support one of the two hypotheses. By the time I reach this point, I'll be in a position to exonerate the alternative hypothesis, thus resolving the conflict. But I might as well fail at this particular step, for some reason or another, and treat the quandary as prevailing when it actually should have been settled. I submit that this distinctive way of failing is an epistemic form of weakness, and it corresponds to an epistemic vice, in the obstructivist sense sketched in Chapter 4. I call it the vice of epistemic inefficacy.

This way of failing could take different forms. I could, for instance, refrain from dismissing hypothesis A and keep it as a live possibility in the presence of evidence

that suffices to rule it out. Or I can keep hypothesis A as a live possibility in the presence of evidence that suffices to establish the truth of hypothesis B, where A and B are rival hypothesis. It doesn't matter what precise shape it takes – as long as I treat the quandary as prevailing for longer than I should, I'm being weak in a distinctively epistemic sense.

My primary motivation for thinking that this is an intellectual form of weakness that is parallel to weakness of willpower is my belief that what I've just described has some remarkable similarities with the anecdotal cases of practical akrasia alluded, rather than the assumption that its conceptual definition should be the exact parallel to practical weakness of the will.

Many of the epistemic virtues and vices that borrow their names from terms pertaining to the classical virtue ethics' glossary are not their exact epistemic counterparts. For instance, courage. Classically understood, courage is the virtue of facing threat to one's life in exchange for a greater good, for example, in warfare (*Nichomaquean Ethics* 1115a-1128b). Epistemic courage, on the other hand, is the virtue of standing up for one's ideas in the presence of people who will disagree (cf., for instance, Zagzebski 1996: 17-18). An intellectual form of weakness of the relevant intellectual capacity, as I take it, is only parallel to practical forms of weakness of willpower as much as facing opposition by disagreeing peers is a parallel to facing death in the battlefield. It captures some of the essence of the original notion, by repeating some of its crucial, distinctive features; but it is not the original notion simply with "action" replaced by "belief".

This doesn't mean, of course, that I'm associating the two things (practical weakness of willpower and epistemic inefficacy) just because. As I hope to be able to make clear, my grip on the topic encompasses the understanding that the philosophically most interesting way of approaching the issue of weakness in intellectual contexts is approaching it from a virtue-based perspective. That is to say, rather than asking the question of whether one can consciously accept a proposition while also accepting that it is epistemically wrong to do so (Hookway 2001: 178), I propose that we focus on the most captivating question of how to account for what is epistemically unsatisfactory about people who conduct themselves in inquiry in ways that are similar to the agents' in the Oxfam-check example, or the chocolate-cake example, to a certain extent – ways that assimilate some of the essence of the original problem.

This being said, here is the plan for the chapter. In section 5.2. I'll get clear on the idea of dilemma facing and on why this is important for weakness of willpower. Then, in section 5.3. I'll present my overall picture of epistemic inefficacy by means of a concrete and detailed example, and my reasons for thinking that this is weakness of willpower's closest intellectual equivalent. Then I'll proceed to discuss the relevant ways in which epistemic inefficacy is different from other vices and other epistemic misconducts with which it can be all too easily conflated. Section 5.4. delves into the differences between epistemic inefficacy and the vice of close-mindedness, while section 5.5. presents the differences between epistemic inefficacy and sub-personal states that might resemble it superficially. I'll conclude, in section 5.6., by presenting a preliminary profiling, with the vice's most remarkable prototypical traits seen so far. The precise ways in which epistemic inefficacy squares of as a vice, following the obstructivist approach to vice epistemology presented in Chapter 4, will be dealt with in a detailed way in Chapter 6.

5.2. Weakness of willpower and facing dilemmas

A remarkable feature of akrasia that is almost completely overlooked by Anglo-american contemporary philosophy is that akrasia is an attitude or pattern of attitudes an agent takes when she faces dilemmas, of both small and big proportions. If we pick from the literature the least stipulative examples of the phenomenon (that is, the ones that are more concrete, in essence; the ones that seem more akin to what we see in real people, in the real world[54]), we find that what is going on is: the agent finds herself cornered within a situation that she struggles to get out of, even though she does want to get out. The person has a general idea of what the right thing to do is, but she also has an inclination towards doing something entirely else, so the situation is not one that unfolds naturally, or in a way that is easygoing. It does not simply "result" in the person choosing an action that conflicts with her general idea of what is right. Instead, the process involves struggle, because the person experiments the clash between two opposing inclinations. In other words, the agent undergoes a dilemma.

---

[54] Like the ones presented in the Introduction of this dissertation: Dante's Vita Nouva protagonist, who is kept from thinking straight due to the temptation of contemplating images of Beatrice; Saint Paul's dealing with the issue of resisting sexual temptation; Sheakspeare's McBeth dealing with the question of killing or not killing King Duncan; and Alcibiades vacillating between his love of power and his love of virtue.

It's not a mere coincidence that even the most archaic descriptions of the phenomenon mention that the experience is commonly followed by self-awareness of one's failing to abide by his or her own general principles, as well as by some form of regret, like Aristotle remarks in the *Nichomaquean Ethics* (EN VII 7 1150a 20). I don't interpret this as entailing that people that are guilty of weakness of willpower necessarily experiment dramatic suffering, nor that they necessarily have the *feeling* of regret, but I do consider it worth noticing that there is a reflexive nature to the experience, and that the person who undergoes it is in good faith. For, without these, neither suffering or regret would be possible to come by.

If follows from these observations that it doesn't make sense to apply the label "weakness of willpower" to people who are acting in seemingly alien ways but who are not, literally, undergoing conflict while dealing with the conflicting possibilities of action placed before them. That is, weakness of willpower does not apply to people who are not struggling in the face of a dilemma. To face a dilemma is to envisage opposite possibilities of action and to hesitate before them[55].

Many of the dilemmas we face in our lives are epistemic. They require deciding between two rival epistemic possibilities: to draw conclusion A instead of conclusion B; to underwrite theory A to the detriment of theory B; to abandon line of investigation A in order to pursue line of inquiry B, and so forth. Sometimes those dilemmas go away quite smoothly, and quickly. Upon reassessing the evidence, people decide that conclusion A is what follows, instead of conclusion B. Upon performing more tests, scientists decide that theory A is more accurate than the rival theory B. And so forth. Those are the cases that do *not* interest me here, precisely because they involve no struggle on the agent's part. Cases of dilemmas that resolve themselves do not test a person's willpower, or her efficacy as a problem-solver.

Not all epistemic dilemmas, however, unfold like this. That's because sometimes a person finds herself being unable to resolve a particular conflict between two rival epistemic possibilities. She accesses more evidence, but she finds that the evidence doesn't settle the matter. She conducts more tests, but then finds that the tests too don't conclusively rule out either of the two possibilities envisaged. And so forth. What happens in situations of this sort is that the person is forced to put

---

[55]It is worth noticing that it follows from these observations that most, if not all, of the cases discussed in Chapter 1 are not really equivalents of weakness of willpower in the epistemic realm, in spite of their having been presented by scholars as alleged examples of "epistemic akrasia".

up with inconsistency (the inconsistency that exists between the two rival possibilities, since they cannot both be accepted, but none of the two have been ruled out yet). This is the type of situation that often tests a person's intellectual willpower. That is, this is the type of situation in which intellectual weakness of the relevant kind can become salient.

But it actually gets a little more complicated than that. Weakness of willpower is not just a matter of the person experiencing dilemma, but rather a matter of he or she experiencing a dilemma *that she shouldn't be experiencing.* Think like this: if you have a general guideline for living, or a principle regarding how to handle matters of some particular type, according to which the right thing for you to do is A, and yet you struggle when you find yourself in a situation in which you can do either A or B, then this is a dilemma you shouldn't be facing. If you have *really* committed yourself to that general guideline, it's admissible that you hesitate for a little while, but that's all. That's the most longstanding your dilemma should be, if you were an impeccably reasonable person who always abides by your principles. An impeccably reasonable person is one who already got over the point in life where option B offers real temptation.

Nevertheless, many times people have not got over it. They're still living in that place in life where they are tempted by option B, in spite of seeing themselves as committed to a principle according to which the right choice is A. When a person is that place in life, so to speak, she will be tempted by option B even though she has much better reason to choose A, and this raises a red flag in terms of her strength of willpower, even if in the end she ends up choosing in accordance with her general principle. The crucial intuition this observation gives us is: weakness of willpower has to do not only with facing dilemma, but with being caught up within the dilemma for *longer* than you should.

This doesn't mean, of course, that we can stipulate a maximum duration for one's struggle in the face of a dilemma, so that if the person exceeds that duration she automatically qualifies as lacking willpower. What I do mean, though, is that some dilemmas are of such nature that they have a "turning point". There is a point in the unfoldment of person's dealing with some types of dilemmas, in a way that, once you go past this point without having achieved success (resolution), you begin to be considered unreasonable. Usually, that's the point where epistemic reasons in favour of option A has irreversibly or nearly irreversibly outweighed epistemic reasons in

favour of option B. So weakness of the relevant kind happens when the person keeps treating the dilemma as an open case *past* this point. Where exactly this point is, or where the line is to be drawn, is usually circumstantial and will vary in each case.

For instance, if you have sincerely committed yourself to the principles of a Christian monastic life but then you keep finding it difficult to abide by monastic rule, being tormented by temptations beyond that reasonability point (wherever and however the bar turns out to have been set), what happens is that after a while we begin to suspect that there is something wrong: either your adhesion was not really collected and sincere, or you lack the willpower to actualize it in action. If your adhesion was indeed collected and sincere (meaning: if you are in good faith), then you lack the power to actualize it in action. You are akratic.

Likewise with epistemic situations. If you are an expert in a field and you have a general guideline, or a rule of thumb, about how to handle issues of some particular type, according to which the right thing for you to do when you have to deal with these issues is A, if you are an excellent problem-solver in your field, that is, a person with maximal epistemic efficacy, you wouldn't feel tempted to deal with the relevant problem in way B, or in any other way that you consider sub-optimal. If you struggle when you find yourself in a situation in which either A or B look possible, then this is either because you are not *really* committed to that general guideline that prescribes A, or because you don't have maximal epistemic efficacy as a problem-solver in the field (and this is regardless of your general guideline being right or wrong).

To be sure, nobody has complete epistemic efficacy in any field, or, at least, nobody should see themselves as enjoying any perfect skills. Which means, everybody could find themselves facing dilemma in any field, that is, everybody, even experts, could find themselves considering option B and weighting it against option A at some time. What is crucial in this case is how long this is going to last for. If it extends itself past the point in which reasons pro-A have outweighed reasons pro-B in a way that is very unlikely that the weigh balance can reverse itself, there you have it.

Here is a more illustrative example. Say you are a novice physician and you've learnt that there is no serious scientific evidence linking antiperspirants to breast cancer, like a popular theory disseminated in the 1990's used to hold. But then you keep finding it difficult to *not* think that your patients' breast cancer is a reflex of

antiperspirants use, when you see them arriving at your clinic wearing a ton of it. It's alright if it happens once, or for a short while. But after it has lasted long enough (too much inquiring towards a patient's hygiene habits to see if her disease can be traced back to antiperspirants), or after iteration (too many patients having their antiperspirants use inquired about), we'll begin to suspect that there is something wrong with you. Either you haven't really understood or accepted the data, or it is as though you lack the intellectual "guts", so to speak, to actualize it in your conduct. If you've understood the data and you've fully accepted it, i.e., you have no caveats about the data's authenticity, for example then it is as though you lack the guts to actualize that knowledge in your own conduct.

Now, practically speaking, what we do observe when we see cases of seemingly stubborn physicians in real life is that nobody that adopts this conduct, that is, no physician that spends valuable time and resources tracking specific factors like this does so from a neutral standpoint. That is, nobody does this without believing, at least to a certain extent, that the factor-hypothesis has *not* been completely ruled out. Those people accept the data, but they don't think that the data is strong enough to *completely* undermine the old theory either. So they think that there is still a chance that that old theory is valuable or, at least, not completely mistaken. That there is some truth to it, that just hasn't been backed up by the data yet.

This is another way of saying that those people have some degree of commitment to the theory, *in addition* to their commitment to the general principle that a medical practitioner should base his practice on theories that are backed up by the best evidence in the field. Some degree of commitment to the idea that represents the weaker alternative in a dilemma (in this case, the antiperspirant hypothesis) is what causes the person to linger herself in facing that particular dilemma past the reasonable point. We'll get back to this in section 5.3.

In the next section we'll examine epistemic inefficacy it in a more detailed way, from the analysis of an extended concrete example.

5.3. An overview of epistemic inefficacy

What do epistemically inefficacious people look like, when their vice is made the most visible throughout their extended intellectual conduct? Too get a glimpse of it, meet Bob, the palaeontologist.

PALAEONTOLOGIST
Bob, a respected palaeontologist, is called to examine the bones of a tyrannosaurid skeleton of controversial taxonomic status. The aim of the enterprise is to determine whether it is of a *Tyrannosaurus rex*. The skeleton's main features strongly suggest a T-rex, except for a few morphological differences, of which stature is the most remarkable: it is way too short compared to virtually every other T-rex skeleton that they know about. After conducting a series of measurements and tests, Bob concludes that the skeleton should be described as belonging to another taxa, thus coining the term *Nanotyrannus* to refer to what he takes to be a new scientific finding: another sympatric tyrannosaurid species of markedly smaller adult body size – a pygmy relative of the mighty *Tyrannosaurus*.[56]

A while after *Nanotyrannus* having officially entered the records, a series of studies popped up, presenting substantial evidence in favour of the so-called ontogenetic niche partitioning theory in dinosaurs. This theory holds that dinosaurs have varied considerably in size and shape throughout their life span, mainly according to resource abundance.

Roughly, according to this theory, dinosaurs would have been somewhat like modern water snakes. The water snake is an interesting reptile that eats predominantly fish from its birth to early adulthood. When it reaches about 50 cm body length, however, its diet gradually switches to frogs and other amphibians, and remains like so to the end of the animal's life. The shift in diet is accompanied by discrete, but noticeable, morphological changes, insofar the technique required to hunt for frogs is different from the one required to catch fish. So a fish-eater small water snake that looks different from a frog-eater bigger water snake may both belong to the same species, *Nerodia erythrogaster*[57]. Dinosaurs, these new studies suggest, vary likewise. So it is possible (though it is not certain) that the fossil examined by Bob was actually that of a small juvenile T-rex at some point in its development from egg to early adulthood.

Given that the ontogenetic niche partitioning theory raises an alternative way of looking at the fossil examined by Bob a while ago, it's understood that in light of the new studies alluded he should now reconsider – a duty he does not run from. So, after carefully accessing the data, Bob finally accepts that there is a possibility that that particular fossil belonged indeed to a juvenile T-rex, rather than to an adult "Nanotyrannus".

---

[56] This story was inspired and freely adapted from the veridical case recounted in BBC's article *Meet Nanotyrannus, the dinosaur that never really existed*. Available at https://bbc.in/3ezLPzz .

[57] For an interesting review of the ontogenetic niche partitioning theory see Werner & Gilliam (1984).

Notwithstanding, because the studies in question amount to evidence that pertains to other species and that is silent on the specific issue of tyrannosaurids, Bob takes it that those studies do not *conclusively* establish that he was wrong; therefore, he does not rush to withdraw his claim to having found and baptized a new species of dinosaur. He remains, rather, waiting for further studies to either confirm his discovery or debunk it. "In time, the theory that the *Nanotyrannus* actually existed as a separate species will be either proved or disproved", he says, undisturbed.

In a situation of this sort, there is clearly a conflict going on between

p: The skeleton in question likely is of a new, uncatalogued species.

and

p': Adult Nanotyrannuses likely are juvenile Tyranossauruses.

Though technically speaking there is no contradiction between **p'** and **p**, the conflict stems from the fact that one is not supposed to be evenly confident in both, at least not in the long run. One must give up one of the claims, at some point[58], inasmuch as they are epistemically incompatible. By saying that two claims are epistemically incompatible I mean simply that accepting one of them would constitute a reason for not accepting the other (Heil 1984: 63).

What hangs in the bar here is the validity of *Nanotyrannus* as a separate biological species. Nanotyrannus can be an adult relative of the T-rex *or* it can be literally a synonym to "juvenile T-rex", in which case "Nanotyrannus" is not a species on its own right. However the case may be, in order for this debate to be fully resolved, palaeontologists have to discover the definitive teenage Tyrannosaurus, that is, a fossil that is acknowledgeable as the young T-rex by a set of independent criteria. If this fossil is morphologically very distinct from the alleged adult Nanotyrannus (distinct enough for the two of them to be unmistakable), then it

---

[58] This expectation (that one must give either **p.1.** or **p.2.** up, at some point) is in itself a source of controversy among contemporary scholars. Some take it that this is indeed a requirement of rationality (for instance, Broome 2013 and Gibbons 2013); while others take it that it is not (for instance, Sinnot-Armstrong 1996). The latter believe, instead, that rationality is bounded by dilemmas – situations in which one is not capable (and therefore not required) to take a rational stand. We'll have the opportunity to discuss this issue at length in Chapter 5. For now, I'll take it that there are at least some situations in which suspension of judgment in the face of the conflict, or dwelling on the conflict, is acceptable for some time, but not indefinitely; and that the palaeontologist's situation is one of those.

is likely that Bob was right and the fossil examined by him was of a different species. If not, then it is likely that he was wrong. Until this specimen is found, the question remains open of what exactly the skeleton examined by Bob is, because only then will scientists posses the means to tell an adult pigmy from a teenage giant, so to speak (if they are, in fact, different). In other words, only then will they have something to point to while saying "Nanotyrannus is not a juvenile T-rex, because this [pointing to the newly discovered skeleton] is what a juvenile T-rex looks like: see, they're different"; or else "Nanotyrannus is exactly a juvenile T-rex, because this [pointing…] is what a juvenile T-rex looks like: see, they're identical".

Whilst this fossil does not show up, Bob's suspension of judgment in the way described is a way of coping with the conflict. That is, by so doing he is putting up with the conflict, allowing the case to remain open. His demeanour is not an epistemic misconduct at this point, on the contrary. Provided that the conundrum depends on the upsurge of new evidence in order for it to be solved, a provisional suspension of judgment seems fine. But – and this is crucial – it is fine *if* it is indeed held provisionally. In Bob's case, this is not, however, how things actually unfold.

Two years go by. Bob keeps diligent track of the publications in the field and the efforts made to uncover new fossils, and he sees that no new fossils have been uncovered that could possibly be the much wanted juvenile T-rex acknowledgeable by independent criteria. In contrast, some of the fossils that were formerly registered as adult Nanotyrannuses have been reclassified as juvenile T-rexes here and there. The palaeontologists who authored such reclassifications claim that as the ontogenetic niche partitioning theory grew more sophisticated, the theory itself provided them with the basis for a change of mind (in spite of this theory being silent on the specific Nanotyrannus conundrum). So perhaps Bob should do the same, shouldn't he? After all, those are Bob's epistemic peers, and it is pretty clear that he is outnumbered by them at this time. But he is not yet ready to. He'd rather wait some more. He's confident that the definitive proof will show up sooner than later.

Then, a decade goes by. Again, no new fossils come up of the sort that could settle the matter for good, but more and more former adult Nanotyrannuses' fossils become juvenile T-rexes through reclassification. It looks like Nanotyrannuses are becoming rarer each time. In fact, it is not just the Nanotyrannuses that are becoming juvenile T-rexes. A good deal of other taxas are "disappearing", under the understanding that some of the skeletons that were formerly regarded as belonging

to adults of a myriad of recently baptized exotic species are in fact juveniles of a much smaller number of "classic" species. So perhaps it's about time for Bob to withdraw his "discovery", isn't it?

Well, yes, but… he doesn't look at it this way. For one thing, Bob agrees that, normally, ten years is a long enough time frame for a controversial theory that hasn't received any further evidential support to be welcomed to the hall of scientific mistakes. After all, the continuous absence of evidence where evidence was expected is not a neutral fact. It means something, for the prospects of the theory – it means that the theory becomes each time less likely to be true, not more. So he knows that the most reasonable thing for a scientist to do when the theory she endorses reaches a critical point in this process of becoming less and less likely is to straightforwardly withdrawn support to it, rather than keep judgment suspended.

Notwithstanding, even though Bob accepts that this is a general truth about scientific theories, he remains uncertain that *this* particular matter, the Nanotyrannus conundrum, has reached the point from which it could be considered settled. He wants assurance, but assurance would only come when certain specific conditions are satisfied, that haven't yet been, and perhaps never will. While it's not possible to obtain assurance at this time, he believes that the quarrel could well stand for longer. So he'd rather wait some more before taking a definitive stand. And thus he stays: sitting on the fence, hopeful that after the next expedition, perhaps, he'll be able to either voice a denial or an affirmation with (what he takes to be) the fair degree of certainty. In the meantime, he won't concede that his *Nanotyrannus* hypothesis has been ruled out.

Now, how do we account for what is epistemically unsatisfactorily about Bob's conduct? What Bob does is: he stretches the dilemma[59]. He tolerates epistemic conflict more or less "in the same way we endure a hangover" (Rorty 1997: 644): by telling ourselves that the discomfort will soon pass and that things will be different next time. Arguably, some degree of tolerating epistemic conflict is necessary, if one is to be an effective and responsible inquirer, and this can be attained in several different ways. But if every scientist, for instance, coped with epistemic conflict with the same measures and to the same extent that Bob does, scientific enterprise would

---

[59] For practical purposes, I'm taking the expressions "epistemic conflict", "epistemic incompatibility", "dilemma", "quandary" and "inconsistency" to be synonyms, and I alternate freely between them. I understand that these expressions might be ascribed different meanings within different contexts, but for the purposes of this work this is more a matter of stipulation than of real concern.

stall. Likewise, if police detectives, judges, lawyers, medical doctors and prosecutors did the same, diseases would be diagnosed too late, lawsuits would take forever; offenders would simply never be caught and innocent suspects would never be able to clear their names.

I submit that Bob's attitude is a garden-variety case of an epistemic vice: the propensity to tolerate epistemic conflict in inquiry for longer than it would be reasonable. This, I argue, is the vice of being epistemically inefficacious or, simply, the vice of epistemic inefficacy.

Now, why think that this is somehow related to the problem of practical weakness of willpower (akrasia)? Just because in both we have an agent dealing with a dilemma doesn't seem a strong enough reason. If it's just this, the relatedness between those two things wouldn't be more than anecdotal. Nevertheless, it isn't.

It has been said, in Chapter 1, that one of the difficulties of conceiving of akrasia of the intellect is that, classically understood, akrasia is something that happens when a person is overridden by appetites of all sorts (sexual drive, the pleasures of food and drink, craving for power, and so forth). It's because there is this mighty force, the appetite, that the person's will to do what is right becomes weak and ineffective. The problem I remarked there was that, in epistemic contexts, we don't seem to be able to locate this ingredient, an ingredient that would be the epistemic equivalent to that "mighty force", the appetite.

In the account of epistemic inefficacy I'm offering here, however, we can see with clarity what ingredient this is: it's commitment to a previously formed idea, or a theory. Commitment to ideas can work like a passion. It's Bob's unwillingness to abandon his Nanotyrannus hypothesis that keeps him from eventually doing what every scientist knows, or at least should know, that must be done: to settle the matter after the amount of evidence gathered in inquiry has surpassed the reasonable threshold. He doesn't have the "guts" to do this, so to speak, but he doesn't want to go all-in in rejecting the data either.

Much like the person in the chocolate cake example, who knows she is supposed to close her mouth after having eaten the reasonable "portion", a person like Bob knows, or at least should know, that a scientist is supposed to consider the matter settled after certain reasonable conditions have been met. When an overwhelming amount of evidence has been gathered that supports a hypothesis to the detriment of its rivals, or when an overwhelming amount of evidence has been

gathered that simply do not support a hypothesis that they were expecting it would support, scientists have eaten the reasonable portion of their cake. It's time to stop. But Bob cannot stop, because is overridden by this "mighty force": his sense of commitment to that old hypothesis.

This commitment is crossdressed as caution: wanting to examine more, keeping inquiry open, demanding more data, and so forth. As much as it can look like skepticism, what it actually is is a personal difficulty to abandon an idea. Contrary to cases where caution is due, in a case like Bob's it is excessive and pointless (and, like we'll see in a detailed way in Chapter 6, it's also noxious).

At its essence, (strength of) willpower is the ability to resist short-term temptations in order to meet long-term goals (Duckworth 2011). In a similar way, the reason why Bob's epistemic efficacy is not strong is that he is unable to resist the instant gratification that can be obtained from thinking that his old hypothesis still retains some of its value, that it hasn't been conclusively ruled out yet. There is, indeed, a minuscule chance that his hypothesis hasn't been conclusively ruled out yet; he clings to this minuscule chance. Being unable to resist this instant gratification, he is kept from meeting the long-term goal of science, which is to hold only the theories that are true or, at least, the ones that offer the best explanation currently available for the phenomena they are meant to explain.

I take it that epistemic inefficacy, in the way it has been characterized above is, if anything is, the equivalent to weakness of willpower in distinctively intellectual contexts: a struggle in the face of a dilemma that exists for longer that it should because the agent is overridden by this "mighty force", his residual commitment to an idea (that happens to be the weaker side of the contest). It is clear that it doesn't mirror the set of necessary and sufficient conditions for practical akrasia. For instance, cases of epistemic inefficacy don't always meet what has been described in previous chapters as "the acknowledgment condition" for akrasia. It is not clear that people like Bob fully acknowledge that they're undertaking a conduct that goes against what general scientific mindset prescribes, or that they see themselves as being excessively cautious. Many times those people see themselves as being cautious to the right extent, that is, neither too little nor too much. They think that they're doing nothing wrong, that that dilemma is one that needs to be endured until the day more conclusive evidence shows up that allows for it to be adequately settled (a day which, of course, is bound to never come).

Nevertheless, this is as close as you can get to practical akrasia in the intellectual domain because full-blown belief akrasia is impossible, as we've seen from Chapters 2-3; and the other candidates, the examples of epistemic akrasia with a mirroring structure, laid out in Chapter 1, either were too vague or episodic, or they were dependent upon the cold attribution standpoint to look like akrasia.

Given the characterization of epistemic inefficacy presented, a number of clarifications are needed, which concern mainly the ways in which it differs from other epistemic vices and from other seemingly alien intellectual behaviours. Epistemic inefficacy might be all too easily confused with other epistemic vices, of which close-mindedness is perhaps the most remarkable. Also, it might not be immediately distinguishable from other epistemic misconducts that technically speaking are not vices, such as selective skepticism and indecisiveness. In the next two sections I'll address the question of how epistemic inefficacy is different from those other things. Answering these questions will allow me to discuss in more detail the specifics of why epistemic inefficacy is an epistemic vice.

5.4. Epistemic inefficacy *vs.* close-mindedness

Let me borrow an idea from Marta Nussbaum (1988), in order to break this question down. Nussbaum says that virtues and vices are ways (good and bad) of responding within certain domains, or spheres of human experience. According to her, the reference of virtue and vice terms is fixed by these spheres, by means of what she calls "grounding experiences". Those are experiences that figure in virtually every human life, and in which virtually every human being will eventually have to make some choices rather than others, and act in some way rather than some other (if not adequately, then inadequately).

For instance, generosity and greed are two ways of responding to a same grounding experience, which is the experience of possessing extra resources while knowing that someone else is in need. Every person who happens to find herself possessing some extra resource while knowing that someone else is in need will respond to the situation either virtuously, by being generous, or viciously, by being greedy. So the virtue of generosity and the vice of greed are the adequate and the inadequate responses, respectively, to the same grounding experience[60]. Courage

---

[60] A grounding experience need not fix the reference of only two corresponding attitudes. Usually, a grounding experience allows for one adequate attitude that can be undertaken in response to it and

and cowardice, likewise, are the virtue and the vice concerning the same grounding experience, which is the experience of facing a threat to one's life. And so forth.

I'm sympathetic to Nussbaum's view of how virtue-based terms have their reference fixed, and I take it that it can be extended to epistemic virtue-based terms. So, for instance, curiosity and insouciance might be said to have their reference fixed by the same grounding intellectual experience, that is the experience of realizing that there is something unexplained, or a piece of knowledge missing. They're the adequate and the inadequate way, respectively, of responding to this experience. The curious person tries and goes after that piece of knowledge, whereas the insouciant is not bothered. Epistemic courage and epistemic cowardice have their reference fixed by the experience of having to voice and defend an opinion against an audience who will disagree. The courageous person will try and respond to the disagreeing peers' objections, whereas the coward will not. And so forth. Knowing what the grounding experience of a vice is helps us get clear about what vice it is and how it works. That is, it gives us insight into the vice's specifics, which in turn allows us to distinguish it from neighbouring vices and other misdemeanours.

This being said, one is to wonder what the exact grounding experience to which people like Bob are responding is. One plausible answer is to say it is an experience related to what is often referred to by epistemologists and social psychologists as *closure*. In the face of questions that were raised but not yet answered, or not answered properly, an agent can respond in different ways, adequate and inadequate, depending on the extent to which she craves closure. Need for closure is an "individual's desire for a firm answer to a question and an aversion toward ambiguity" (Kruglanski & Webster 1996: 264).

This makes sense especially because Bob seems to exhibit an ultra-high need for closure: he desires a firm answer to the conundrum he's facing and is not willing to consider the matter settled until he has had one. It looks as though the inadequateness of his epistemic response is at least in part a function of this desire. Specifically, he is only willing to accept that his Nanotyrannus theory is mistaken when he finds himself in one of two firm-answer scenarios: either when the "definitive" proof is dig off the ground *or* when he is presented with a potentially

---

more than one inadequate ones. For instance, a person can respond to the experience of having extra resources while someone else is in need by donating the right amount, by overgiving, and by not giving enough. The latter two are inadequate ways of responding.

infinite amount of evidence supporting the rival theory, the ontogenetic niche partitioning theory.

The problem is that those are utterly unreasonable demands. For one thing, who in the world knows whether fossils that have been laying underground for 65 million years will ever be uncovered? Who knows whether these fossils even made it preserved to modern times? What's more, every time the ontogenetic niche partitioning theory receives some support, Bob deems it not enough, and demands that the theory receives even further support for him to acknowledge it as a defeater to his hypothesis. He demands that it be supported by infinite evidence, or else he won't give in.

Arguably, what would be reasonable for Bob to do instead is accept that all the evidence provided by the ontogenetic niche partitioning theory in its current state is jointly sufficient to back up the claim that Nanotyrannus is actually a juvenile T-rex rather than a different biological species; and be prepared to change his mind if (not exactly when, but if) new evidence pops up that indicates otherwise. In other words, he should have made up his mind at some point by reasoning from the available data and been open towards the possibility of changing his mind in the future. His troubled and apparently higher than reasonable craving for certainty kept him from doing his.

Arguably, this bears important similarities with the vice of close mindedness, which is certainly one of the most discussed of its kind. Close-mindedness is marked by "an unwillingness or inability to engage seriously with relevant intellectual options" (Battaly 2018: 261). The close-minded individual fails to engage seriously with relevant intellectual options because she is closed to these options being true, or legitimate, or better than the ones she herself endorses. Like Bob – he looks like he simply is closed to the possibility that the ontogenetic niche partitioning theory in its current state suffices to back up the claim that "Nanotyrannuses" just are juvenile T-rexes.

Now, if the grounding experience Bob is responding to is an experience related to need for closure, then isn't it natural to think that the vice displayed by his conduct simply is the vice of close-mindedness? Why coin and baptize an exotic vice, using a fancy word, if the conduct that is to be explained can already be explained as an instance of the much well-known vice of closed-mindedness? It is as if, in so doing, I'm guilty of the exact same sin as Bob himself. "Epistemic inefficacy" looks pretty much like a philosophical *Nanotyrannus*, someone could argue.

Let me clarify why this is not so. While it might well be true that Bob has issues with his unrealistic expectations relatively to closure, the aspect of his conduct that I want to call attention to is not the fact that he desires a firm answer, per se, but rather the fact that in his quest for a firm answer, he postpones closure. Specifically, he postpones closure for way longer than it would be due. He inadvertently delays taking a stand, when he already has enough ground on the basis of which to take a stand. So every time he is supposed to step forward and take a stand, he sort of finds an excuse not to do so at that particular time. Excuses are always pretty much the same: "this is not conclusive, we cannot jump to a conclusion yet"; "let me wait some more, to see whether new evidence pops up that gives me more certainty…".

But every time new evidence does pop up and certainty is increased, he again comes up with an excuse for further postponement. This is another sense in which his conduct resembles parochial cases of practical akrasia, such as the tale of the chocolate cake person, who knows she is supposed to be dieting, but is always finding herself in front of another piece of chocolate cake again, saying something to the effects of "just this one more", or "this is going to be my last…".

What does this mean, exactly? It means that while the grounding experience Bob is responding to is indeed one related to need for closure, and while his individual need for closure is indeed troubled and unrealistic, his conduct is not close-minded in a strict sense, as it shows signs that are quite awkward for a typical close-minded mentality. In fact, he might as well be seen as displaying an excessive, or in other ways inappropriate, disposition towards *open-mindedness*, due to his excessive caution. In keeping the case open for longer than it would be right, Bob might be seen as precisely *avoiding* closure (as opposed to craving it), which means accepting to live with uncertainty and/or ambiguity instead of firmness.

In fact, ultra-cautious people like Bob cope with uncertainty and/or ambiguity for way longer than typical close-minded individuals would be willing and able to. Need for closure is inversely correlated with tolerance of ambiguity: generally speaking, the higher one's need for closure, the more badly she wants firm answers, and the less she tolerates having ambiguity and/or uncertainty instead (DeRoma et. al. 2003; see also Berenbaum et. al. 2008). Bob tolerates ambiguity and uncertainty for years. This is utterly atypical of closed-minded individuals.

Notwithstanding, his conduct is not to be described as a genuine instance of open-mindedness either, at least not in the sense in which open-mindedness is

classically talked about, as being something praiseworthy, a virtue (the virtue of carefully considering alternatives and being aware of one's own fallibility as a believer, cf., for instance Riggs 2010, Baehr 2012a and Kwong 2016). Let's see this through.

From a practical perspective, open-mindedness is at the opposite side of closed-mindedness in the need for closure spectrum – open-minded individuals are those with a low need for closure and high tolerance of ambiguity. So we might think of open-mindedness and close-mindedness as a pair of alternative and opposing responses to the same grounding experience, just like courage and cowardice, generosity and greed, etc.

According to Kruglanski (2004, see also Kruglanski & Webster 1996), one becomes closed or open minded depending on her epistemic motivations, that is, her goals relatively to knowledge. Particularly, it depends on whether the individual wants to approach or avoid closure. Close-minded individuals are closure seekers: they seek closure by demanding certain information ASAP. Open-minded individuals, in turn, are closure avoiders: they avoid closure by refraining to adhere to information (what Kruglanski & Webster 1996 call "seizing and freezing"). Closure seekers are disturbed by the idea of being without a firm answer for too long, they're inclined towards jumping to conclusions. Closure avoiders are disturbed by the idea of taking a stand all too soon; they're inclined towards caution[61].

As we've seen, Bob's conduct has some of the traits of the closure seeking pattern, in that he too wishes for firm answers to certain questions and demands certain information to be disclosed. But it also has traits of the closure avoiding pattern, in that Bob, like open-minded individuals, suspends judgment due to concerns about prematurely taking a stand. Though open-mindedness is often talked

---

[61] Kruglanski (2004) actually distinguishes two orthogonal dimensions of motivation: closure seeking versus closure avoidance; and specificity versus nonspecificity. If I've just arrived at the airport and I want to know to which gate I might proceed, this is an example of a need for non-specific closure, once I do not have preference for any specific gate. If I have just taken a SAT test and I want to know my results, this in turn is an example of specific need for closure – I want to know that I succeeded, and finding out that I didn't is not going to make me as happy. Both examples are Kruglanski's (2004: 5–8). In terms of closure avoidance, his examples are less concrete, but he speaks of circumstances under which one is concerned about committing a costly error of judgment, and thus "may be prone to suspend judgment altogether or avoid premature closure", or "to avoid or postpone it until having considered ample further information" (Kruglanski 2004: 10). We might think of, for instance, the person who witnesses a feud and, worried about the costs of taking the wrong side, decides to remain neutral. This would be an example of avoidance of non-specific closure in case the person does not particularly esteem any of the contending sides, and an example of avoidance of specific closure in case she did.

about as being an epistemic virtue, like we said, it's clear that Bob's conduct is not praiseworthy, by any accounts. Context makes all the difference here: his excessive caution would perhaps have been adequate, had it been circumscribed to a certain period of time, that is, had it been held until a certain point in the quarrel's unfoldment: the point where significant amount of evidence has been gathered that overwhelmingly supports the rival hypothesis. While the exact point where the line is to be drawn is debatable, I believe Bob went past it, by all accounts. So his conduct is neither close-minded nor open-minded, but some rather strange hybrid of the two.

Yet, someone might still object to me the following: provided that this type of conduct is strongly (albeit ambiguously) tied to a troubled need for closure, isn't it to be better accounted for as a *variety* of close-mindedness? According to Battaly (2018: 261), there are many varieties of close-mindedness. Close-mindedness can come out as dogmatism, the unwillingness to engage seriously with particularly relevant alternatives to the beliefs one already holds. Or it can come out as the lack of curiosity over alternative hypothesis. Or it can come out as testimonial injustice (actively dismissing what others have to say as to the matter that is under consideration), to name a few. Why could it not come out as epistemic inefficacy, of the sort I've been describing?

Within this view, basically any way of not doing the right thing when it comes to seeking closure is a possible way of being close-minded, including being too open. That is to say, close-mindedness is turned into a quite broad concept: it's a vice that can show up by a variety of means, including by means of other vices. I would pose no objection to describing Bob's epistemic inefficacy as a variety of close-mindedness thus understood, provided that one caveat is added. This caveat is the following: apart from the desire for a firm answer itself, epistemic inefficacy does not relate to most of the general tendencies associated with high need for closure.

Need for closure, as Cassam points out, is associated with some general tendencies. Close-minded individuals, which are those with a high need for closure, typically display reluctance to consider novel information once a given conception has been adopted; they have a poor appreciation of perspectives different from their own; they have high levels of self-confidence and self-assuredness; they are intolerant of others whose opinions contradict their own, to name a few (Cassam 2019: 29).

Not every scientist who acts in the same way as Bob, however, possesses those traits. Many of them are in good faith; they see themselves as open to consider novel information and they engage with people and with sources that are not likeminded. Bob appears to be with them in this. After all, he doesn't resist accessing the novel studies on the ontogenetic partitioning theory. He keeps up with the research in the field, wanting to know about his intellectual rivals' findings as well as his intellectual peers' position, something the paradigmatic close-minded person wouldn't do. Bob at any point dismisses his peers, or downplays their qualifications.

And he doesn't even state his own position with a sense of certainty, or confidence, on the contrary – in acting like he does he is actually fostering uncertainty, keeping the flame of ambiguity burning, so to speak, because he keeps casting doubt on a matter that is actually becoming each time less doubtful, rather than more.

In other words, epistemically inefficacious individuals genuinely want to solve the conundrum. In spite of their having a residual commitment to an idea, they are not dogmatic – they don't let this commitment blindfold them completely. So they engage with relevant alternatives. Where they fail is at determining *when* to stop such an engagement. Therefore, to speak of their conduct as being a variety of close-mindedness might not do them justice after all, unless the concept of close-mindedness we're referring back to is indeed a very deflationary one.

This allows us to take notice of a crucial fact about epistemic inefficacy: this vice is a failure in timing. Epistemically inefficacious people fail in their assessment of when the correct time to make certain deliberations is. While conducting an inquiry, they form a resolution to deliberate, but fail to carry such resolution on. They set themselves up to approach the matter by reasoning more or less in the following way: "I'll deliberate thus and so whenever conditions C are met" (for instance "I'll exonerate hypothesis **p'** as soon as I have sufficient evidence in support of **p**"). This is fine, as far as it goes. The problem is that they fail at acknowledging that the stipulated conditions C have been met, when they have already been met. In this sense, it becomes clearer how the epistemically inefficacious agent is different both from the typical close-minded agent (who fails at forming the alluded resolution) and the typical open-minded one (who forms the resolution and carries it on within the adequate timing).

5.5. Epistemic inefficacy *vs.* other epistemic misconducts[62]

Epistemic inefficacy can also be easily conflated with things such as motivated skepticism, indecisiveness, cognitive dissonance and recalcitrant belief. Those are not epistemic vices, technically speaking, inasmuch as the literature in which such terms are found is more nested within the scientific (rather than the philosophical) bookshelves, where talk of "vices" is rather uncommon. In this domain, authors speak more often of "individual differences", "biases" and "tendencies". Notwithstanding, the scope of these terms is not always clearly distinguished from the scope of terms pertaining to other idioms, including vice epistemology's idiom, so one is left to wonder whether they are not, actually, coextensive.

In this section I'll make the case that epistemic inefficacy does not coincide with any of those phenomena, though they might be in other ways related. I'll focus primarily on motivated skepticism, but I believe the reasons I'll present for why epistemic inefficacy is not the same thing as motivated skepticism also justify why it is not the same thing as any of the phenomena mentioned.

In her book *On Death and Dying* (1969), psychiatrist E. Kubler-Ross describes the reaction of a patient upon first receiving the diagnosis of a terminal illness. First, the patient first considered the possibility that her X-rays had been accidentally switched for someone else's. After confirming that they had not, she left the hospital to go on a journey of seeing several different physicians, in hopes of finding a "better explanation" for her medical condition (Kubler-Ross 1969: 38).

This patient's conduct is an instance of what is commonly referred to by social psychologists as *motivated skepticism*, the "use of differential decision criteria for preferred and nonpreferred conclusions" (Ditto & Lopez 1992). That's the idea that

> when individuals encounter information with unfavorable implications (…), they are more likely to generate multiple hypotheses for testing, engage in a more extensive search for mitigating information, and devote greater processing capacity to evaluating relevant evidence than when confronted with information that is more palatable. (Ditto & Lopez 1992: 570).

The conduct of the Kluber-Ross' patient resembles Bob the palaeontologist's conduct in some important respects – after all, both might be said to have received

---

[62] I'm using the term "behaviour" here in the lay sense, rather than in any technical (i.e., behaviourist) one.

"bad news", in a sense ("information with unfavorable implications", also called "preference-inconsistent information"); both are unwilling to assert a conclusion after coming across these news, and would rather go after further information of a certain kind, despite the costs and the odds of lengthening such enterprise. But do these two conducts, epistemic inefficacy and motivated skepticism, amount to the *same* thing?

Stemming from research carried out in 1970's on the structure of human reasoning (e.g., Tversky & Kahneman 1973, 1974; Shiffrin & Schneider 1977), a variety of scientific studies have been presented as evidence in support of the claim that information consistent with a preferred conclusion tend to be examined less critically by a person, while information inconsistent with a preferred conclusion tend to be examined with more caution and criticism (see for instance Schwarz 1991). In other words: science shows that one's preferences have a bearing upon how skeptical her response to evidence turns out to be. Specifically, people tend to be more skeptical when accessing data that goes against what they, for whatever reason, would rather be the case.

So motivated skepticism is basically a variety of confirmation bias, the "seeking or interpreting of evidence in ways that are partial to existing beliefs, expectations, or a hypothesis in hand" (Nickerson 1998: 175). Here are two important things to take notice about this: first, that this response is understandable, if we think that when a newly accessed information is inconsistent with existing beliefs, expectations, or a hypothesis that a person already has, it will require from that person that she makes certain adjustments in her doxastic mesh, so to speak, and making those adjustments takes cognitive effort. It costs us energy. In order to make sure we're not spending energy in vain, double-checking the information is useful. So it's not always, or not necessarily, a bad thing. Second, confirmation biases are sub-personal mechanisms: they're natural, pervasive, not person-specific (meaning we all have them), not content-specific (meaning they can affect different areas of one's life); they're also predictable and mostly unconscious.

Now, if we think of vices as being a matter of habit (as opposed to a product of nature),as being personal (meaning some people have them and others don't), as being content-specific (meaning specifically tied to certain grounding experiences) and systematically bad (they systematically get in the way of us attaining some epistemic good, like knowledge, understanding, etc.), then vices and sub-personal

mechanisms are two very different things. If motivated skepticism is a variety of cognitive bias, then it's not the same thing as epistemic inefficacy.

However, as Cassam (2019 Chapter 2) emphasizes, it is often difficult to decide whether a particular phenomena is better accounted for in personal, sub-personal or other terms (for instance, supra-personal, or structural, terms); or which of these factors plays the most prominent role in each particular outcome, if they are indeed different factors. Those can actually be different levels of explanation, or levels of description (as, for instance, in Dennett 1969), as opposed to a whole/part distinction[63]. In other words, motivated skepticism could be the correspondent sub-personal explanation for the same phenomena I'm trying to explain in personal terms. I cannot rule this out.

Notwithstanding, I would like to highlight some aspects due to which I believe epistemic inefficacy is not, or at least not necessarily, the exact personal counterpart of sub-personal motivated skepticism, but rather a different thing.

In motivated skepticism, individuals faced with preference-inconsistent information are more motivated[64] to critically analyse the available data, compared to individuals who are faced with preference-consistent information (Ditto & Lopez 1992: 581). However, the crucial fact about the information at stake in the palaeontologist case (contrary to what we see in the Kubler-Ross' patient case, as well as in other paradigmatic cases of motivated skepticism) is not that it is preference-inconsistent, but rather that it is ambiguous, or creates ambiguity.

Let's concede that Bob has a preference for Nanotyrannus being a separate species rather than a juvenile T-rex (because this, if true, would mean he has, say, a scientific accomplishment: he would be the author of the mainstream theory in the field). Accordingly, let's suppose that the series of studies pertaining the ontogenetic niche partitioning theory indeed presents him with data that is preference-inconsistent. Even so, the critical fact here is that this theory actually creates genuine

---

[63] For an excellent discussion on whether the personal-subpersonal distinction is a level-level distinction or a whole-part distinction see Drayson (2014).

[64] It was later suggested by authors such as Friedman (2012) that the cases discussed by Ditto & Lopez (1992) are cases of selective skepticism not due to genuine motivational bias, but rather due to non-motivational influences, including cognitive bias related to break in expectation. According to Friedman, people tend to be more skeptical when dealing with preference-inconsistent information because it fails to confirm to their prior expectations, not her prior motivations. For the purposes of my discussion, it makes no difference whether the selective skepticism seen in those cases is due to motivational bias, expectational bias or a combination of the two (as proposed by Ditto et. al. 2003). What matters is that it is a matter of cognitive biasing.

ambiguity and, therefore, genuine dilemma. Before this theory came into play, the fossil examined by Bob and others of its kind were to be acknowledged as Nanotyrannuses right away. It was after this theory entered the stage that those fossils begun to be seen as Nanotyrannuses *or* juvenile T-rexes. So Bob is responding to something that is, at least initially, a real conundrum: the conundrum represented by the claims **p** and **p'**, "The examined skeleton is of a new species" and "Adult Nanotyrannuses are juvenile Tyranossauruses", respectively. This is the remarkable difference between Bob's case and parochial cases of motivated skepticism.

To be sure, there is a loser sense in which every time you face the question of whether **p**, you are actually facing a dilemma, in a loose sense: the dilemma of whether **p** or **not-p**. But it is not the same thing to be faced with the question of whether **p** and to be faced with an array of evidence that is ambiguous between **p** and **not-p**. The diagnosis received by the Kubler-Ross' patient, for instance, is not ambiguous between "terminal illness" and "non-terminal illness" – the R-rays did not present themselves (or in other words, were not presented to her) as possibly meaning "either terminal illness or something else". It is the patient who treated them as ambiguous without being entitled to, by setting herself off on an odyssey aimed at uncovering other possible meanings for them. In this sense, the X-rays yield genuine preference-inconsistent information, because the meaning of the X-rays unambiguously and directly counters the patient's preference for being healthy (**p**) rather than terminally ill (**not-p**), in a scenario in which she has no evidence whatsoever in favour of **p**.

In a similar fashion, in one of Ditto & Lopez experiments, half the subjects received evidence that they had a genetic trait indicative of a high likelihood of developing a pancreatic disorder (Ditto & Lopez 1992: 574-580; see also Ditto et. al. 2003). Here the individuals who received bad news, but not those who received good news, doublechecked the evidence to see whether the trait was indeed correlated with higher likelihood of the disorder. Like the Kubler-Ross' patient, they treated the bad news as ambiguous, even though it was not.

These subjects act differently from Bob the palaeontologist, who takes advantage of the fact that the information yielded by the ontogenetic niche partitioning theory creates real ambiguity. Bob keeps yearning, or, in a sense, waiting, for the disambiguation to be established, for him to take a stand. The problem is that

he is never willing to concede that the disambiguation has been established. As a result, he waits too long. He keeps waiting even after disambiguation has already been established, because he fails to acknowledge that it has already been established.

It could be objected here that from this moment on (from the moment disambiguation has already been established onwards), Bob falls into motivated skepticism, because there is no more ambiguity, therefore no more reason to keep the quandary open. All there is is the "bad news" for Bob, but yet he keeps refusing to take them at face value. But my point is precisely that Bob fails to acknowledge that this crucial point has come at the time it comes. He transitions from a moment in time in which he indeed cannot close the case (because there is genuine ambiguity, which justifies inquiry being kept open), to a moment in time in which he definitely can and should close it (because there is no more ambiguity). But the transition between these two points in time is *soft*. Bob does not keep up with the exact moment in which he is not anymore in the first "phase" and has already entered the second.

This is different from what we see in the motivated skeptic, to whom there has never been a first phase, to begin with. The Kluber-Ross patient sees ambiguity where there has never been any. Bob fails to see disambiguation where there previously has been ambiguity. So it makes more sense to think of motivated skepticism as accounting for a component of epistemic inefficacy conducts, rather than that these conducts, broadly understood, *are* cases of motivated skepticism.

There is also another significant difference. The Kubler-Ross patient is faced with preference-inconsistent information in a scenario where she is entitled to having a preference, contrary to Bob who, as a scientist, is expected to not have this sort of preference. In other words, while it is understandable that people prefer being healthy to being ill, it is not so understandable that scientists prefer certain theories to be true to the detriment of rival theories. Bob should not prefer that the Nanotyrannus was a separate species, because having this sort of preference pushes one into infringing conventional standards of the research community with the aim of arriving at a particular result, and this is the opposite of what science should be (Wilholt 2009).

As such (because this is epistemically harmful in itself), having preferences of this sort could be considered an epistemic vice in its own right. So part of the reason why Bob's epistemic conduct comes out as blameworthy (certainly more

blameworthy than the Kluber-Ross patient's) and almost anecdotal has to do with the fact that, if he does have said preference, than his epistemic inefficacy is mixed up with an array of other vices including vanity, wishful thinking and, perhaps, self-righteousness.

I don't think that it's always the case that epistemically inefficacious people have *preferences.* What they do have is residual commitment to some idea, but this is not the same as preferring that idea to be true rather than false. We can be committed to ideas for numerous reasons, including out of habit and out of fear, in which cases it's not strictly correct to say that we prefer them to be true. Nevertheless, I do think that people like Bob usually have residual commitment to certain ideas because they have personal preferences, and that this adds up to their vicious pattern. To many of them, the dilemma is personal – something that they hold dear hangs in the line, so the stakes are high, and this is reflected on the proportions of their caution. Perhaps, if Bob, for instance, did not have a preference, it would have been easier for him to be convinced of the ontogenetic niche partitioning theory at some point after being exposed to fair amount of evidence in its favour[65].

Last in the list of things that allow us to compare epistemic inefficacy to cases of sub-personal mechanisms, it has been pointed to me that Bob's case also resembles the gambler's case in the "gambler's fallacy" discussed by Tversky & Kahneman (1974). The more bets the gambler has lost, the more he feels a win is now due, even though each new turn is independent of the last. It resembles Bob's case because Bob, just like the gambler, appears to feel that (or to be reasoning from the assumption that) the more the fossils dig off the ground are of other types of dinosaur, the closer the much wanted fossil that can solve the conundrum for good is to be dug off. His expectation that this fossil is going to be dug doff in the next expedition apparently raises as time goes by, rather than diminish, and that is why he sits on the fence for so long.

The mechanism underlying the gambler's conduct is described by Tversky & Kahneman as a garden-variety example of heuristics: resort to shortcuts or rules of thumb that are prone to error, due to limitations either in cognitive ability or in terms of the resources required for reasoning according to probability theory, or simple

---

[65] It is interesting to note that preferences seem to play a role in many cases of practical akrasia. The chocolate cake person, for instance, might have a quite easier time quitting alcohol, cigarettes or other deserts, than she has quitting chocolate cake. This might be at least in part because the person is partial to chocolate cake, i.e., she has a strong preference for this specific treat.

logic. I pose no objection towards this sort of mechanism being a component of epistemic inefficacy or a description of some of what lies "underneath". But, again, vices and sub-personal mechanisms do not completely coincide, because the former are personal, tied to certain grounding experiences and they systematically produce noxious cognitive effects, whereas heuristics is neither.

5.6. Final remarks

What we have discussed so far allowed for a few important characteristics of the vice I've called epistemic inefficacy to become salient. These are meant not as a set of necessary conditions, but rather as prototypical marks.

1. Epistemic inefficacy is an attitude towards quandaries.

2. Epistemic inefficacy's grounding experience is the experience of lack of closure in the context of a quandary, that is, the experience of an unsolved dilemma.

3. Epistemic inefficacy is a flaw in timing. Inquiring involves making certain deliberations. Epistemically inefficacious people fail at deliberating within the correct timing, while in the process of inquiring towards resolution of a dilemma.

4. Epistemically inefficacious people are in good faith. They want to solve the dilemma, but are kept from doing this in an effective manner because of their residual commitment to one of the terms in the dilemma (the hypothesis that is "losing").

We shall now turn to the details of what exactly turns epistemic inefficacy into an epistemic vice in the technical sense. As Chapter 4 left clear, I follow Cassam (2016, 2019), Kidd (2017), Crerar (2017) and others in underwriting an obstructivist approach to vice epistemology. As such, I believe the consequences of an epistemic vice matter more than is motives: a pattern of conduct is a vice inasmuch as it gets in the way of epistemic goods somehow, such as knowledge, understanding, and so forth. So in the next section I'll discuss the exact ways in which epistemic inefficacy does this. In so discussing, other important characteristics of this vice will be made explicit, allowing for a more comprehensive profiling.

6. THE VICE OF EPISTEMIC INEFFICACY: ITS EFFECTS

Sharpen a blade too much and its edge will soon be lost.
Lao Tzu, **Tao Te Ching**, IX.

This chapter is aimed at expanding the characterization of epistemic inefficacy provided before by including a detailed discussion of why, essentially, we should think that the pattern of conduct described in Chapter 5 is vice (as opposed to a misconduct of other sorts); and of why it is an epistemic vice (as opposed to an ethical misconduct), following an obstructivist framework in vice epistemology. From an obstructivist viewpoint, it is an epistemic vice because it systematically obstructs, if not completely, then partially, some important epistemic good. The paradigmatic epistemic good in virtue-based epistemology discussions is knowledge, but other things, such as understanding, wisdom and solutions to problems are epistemic goods too.

The chapter also includes important considerations that help us answer, in a more detailed way, the question raised in the end of Chapter 1 and at the beginning of Chapter 5: the question of why think that epistemic inefficacy is an interesting concept to epistemologists. It's because this concept is suitable to figure in the best explanation of what is epistemically unsatisfactory about real world people that go about in inquiries that involve dilemma in a way that lingers the dilemma for longer than needed. I provide and discuss real-life examples with the aim of showing how this works.

6.1. Preliminary remarks

From an obstructivist standpoint, bad epistemic attitudes, thinking styles and character traits are vices because they cause some sort of epistemic damage, that is, they get in the way of our epistemic goods somehow: either by obstructing its acquisition, retention, transmission or use. As Cassam points out (2019: 7), "knowledge is something that we can gain, keep, and share"; and I would like to add

to this list *use*. Knowledge is also something that we enjoy, i.e., it is something that we benefit from in our practical as well as intellectual lives. We use knowledge to plan action, to go after more knowledge and to better ourselves, that is, to become knowledgeable individuals. Therefore, vices make it difficult either for one to learn something, or to keep what one learns, or to share it with others, or to profit from it, or some combination of these.

Now, what are the exact harms caused by epistemic inefficacy? Does just the mere fact than an inquiry is being kept open for longer than ideal suffice to produce *harm*? I submit that, in many cases, it does. Here I would like to point to and briefly discuss two main sorts of harm that this conduct generates: immediate and ulterior.

As an immediate consequence, epistemic inefficacy delays the drawing of conclusions, which, in turn, delays knowledge. I believe it affects both knowledge acquisition and its retention, because it impairs something that stands more or less in between acquisition and retention, which I'll call *consolidation*. As an ulterior consequence, epistemic inefficacy threatens to destroy pieces of knowledge that have already been acquired, which represents, mainly, damage to knowledge retention and to its use.

6.2. Immediate harms: delayed consolidation of knowledge

One way in which epistemic vices get in the way of knowledge is by leading to systematic errors in thinking (Cassam 2019: 66). Specifically, they disturb what is termed the confidence condition for knowledge (Cassam 2019: 67). For a person to know the truth of some proposition **p** she must be reasonably confident that **p** and have the right to be confident If it is wishful thinking, superstitious thinking, conspiracy thinking or any other thinking vice rather than evidence that gives that person the confidence that **p**, then she doesn't have the right to be confident, even if **p** turns out to be true after all[66].

---

[66] It is by no means a consensus among epistemologists whether confidence is necessary for knowledge, though many social epistemologists work from the assumption that it is. Assuming this has an advantage: it gives us something that can work as a provisional standard for the attribution of epistemic credentials. Nevertheless, cases in which a person has knowledge but doesn't have the right to be confident are tenable. Bonjour's (1980) perfect clairvoyant is one of those cases. A man has perfect clairvoyant skills. From a reliabilist viewpoint, he has knowledge of the facts that are revealed to him by means of his clairvoyant faculty. But he doesn't have the right to be confident that those facts are true, because he has no evidence that clairvoyance skills are real, or that he possesses such skills. I'll remain neutral on the issue of whether or not the confidence condition is a necessary condition for knowledge in the strict sense, but I'll assume it by virtue of its practical advantage, that is, to beacon knowledge ascription.

The confidence condition generally applies to knowledge attribution: we use it as criteria to evaluate whether someone indeed knows something, or is justified in holding a belief, or has the right to claim knowledge, etc. It encompasses a default and challenge conception of entitlement, more or less in Williams' terms (Williams 2001): you are entitled by default to hold certain position as long as there are no justified challenges to that position. Conversely, you can raise a challenge to a position held by default only if the challenge is justified.

It is entailed by this default and challenge model that one cannot cast doubt on a proposition that is supported by evidence just out of the blue. One must own the right to suspend judgment, and only evidence countering the proposition being evaluated (or lack of evidence in support of it where evidence was expected) can equip one with this right. Thus, if it is a thinking vice rather than evidence or its lack that leads the person to doubt some proposition, then this is wrong, even if that proposition turns out to be false in the end.

It follows from this that when a person keeps judgment suspended relatively to a proposition out of inefficacious thinking or any other thinking vice (rather than because evidence is indeed not sufficient to support a higher degree of confidence in that proposition), this is wrong. She didn't have the right to. So, epistemic inefficacy leads to a systematic error in thinking: it leads one to challenge without being entitled to.

Apart from this, we might as well think about the ways in which epistemic inefficacy is more concretely detrimental. That is, we might think of the concrete immediate consequences that challenging without being entitled to produces that negatively impact inquiry. Epistemic inefficacy has the potential to hamper inquiry, which compromises its effectiveness and responsibility; and this, in turn, has the potential to obstruct knowledge. Specifically, epistemic inefficacy delays the drawing of conclusions that are rather crucial for knowledge consolidation, and whose delaying has the potential to further put epistemic goods already attained into jeopardy. Let's see this through.

If we think specifically in terms of a conduct's consequences, people like Bob, the palaeontologist from Chapter 5, have a way of coping with the conflicting propositions, namely, by tolerating them for longer than it was reasonable. This has something in common with the much known story of the tobacco strategy, recounted by Oreskes & Conway in the book *Merchants of Doubt* (2010). This book describes

how big tobacco companies in the 1950's paid scientists to cast doubt on the link between smoking and cancer.

Though the intentions, contexts, demeanours and proportions of these two actions are obviously different, some of the effects delivered by Bob's conduct and the tobacco industrials' conduct in terms of how they jammed scientific knowledge are alike. In both cases, a scientific fact to which there was sufficient evidential support has been treated as if it hadn't received sufficient evidential support yet. As a result, the fact had its establishment threatened, or, at least, delayed (if we assume sooner than later people would find the truth out). In other words: through both Bob's and the tobacco industrials' conducts, what is actually a warranted fact, something that is *known* by certain people, is denied the social status of knowledge.

We might use the word "consolidation" here, to refer to the process by means of which a warranted scientific claim, one that has been drawn as a conclusion from an adequate set of data, is "officially" proclaimed as knowledge within a community. Consolidation is a social process and, as such, it requires that certain social conditions be met. For a piece of information to become official knowledge, it has to have been officially presented and agreed upon by those who possess the adequate entitlement, and it has to have been accepted by those who don't.

For instance: for it to become official that, say, there is a smallpox pandemic going on, the medical science community has to declare that there is a smallpox pandemic going on using an official communication channel; this has to enjoy a minimum consensus among them, and the rest of the population has to accept it. If the medical science community hasn't declared it, or has done it, but not through an official communication channel, then the fact that there is a smallpox pandemic going on doesn't enjoy a solidity status – it's still only hearsay, even if a lot of people believe it, and even if it is true. On the other hand, if the medical science community has declared it, through an official communication channel, but the rest of the population doesn't accept it (because, say, they have been convinced that this is a media hoax, or fake news), then this fact doesn't enjoy a solidity status either, even if it is true.

In both cases, the fact feels flimsy – it feels like the type of thing it would be completely ok to not believe. Solid facts, on the other hand, are feel sturdy and robust – the kind of thing it would be ok to believe even if you don't fully understand it, and even if many people don't. Examples of consolidated facts in today's world

are, for instance, the fact that silver cannot be transformed into gold, and that smallpox has been eradicated. An example of a non consolidated fact in today's world, in turn, is the fact that cannabis has medicinal properties.

A fact can have its consolidation process obstructed for many reasons, including because of epistemic inefficacy and other vices of the agents involved. The conclusion that Bob, the palaeontologist from Chapter 5, kept avoiding to positively draw, as well as the conclusion that the tobacco scientists discussed in *Merchants of Doubt* made an effort to positively undrawn, were important for the social consolidation of knowledge of a certain kind (even though the tobacco case is obviously not a case of epistemic inefficacy). In Bob's case, knowledge that Tyrannosauruses varied in size and shape throughout their life spam, following ontogenetic niche partitioning patterns (or simply knowledge that Nanotyrannuses are juvenile T-rexes). And, in the tobacco case, knowledge that smoking causes cancer. Not drawing as well as undrawing such conclusions clogs the process through which those pieces of knowledge would attain proper consolidation.

Now, because consolidation is crucial for a piece of knowledge to be made available *as such*, that is, as a robust fact, as opposed to something uncertain, clogging this process prevents the relevant piece of knowledge to become a profitable epistemic good, especially outside of the community of experts. An epistemic good is less valuable if, for whatever reason, it cannot be enjoyed by people to whom it would otherwise be useful, or if we cannot benefit from it, for whatever reason. By delaying the process by means of which the abovementioned information would rightfully acquire the status quo of knowledge and become enjoyable as such by the members of a community, the intrinsic epistemic value of these facts is (unrightfully) held low.

The palaeontologists' example from Chapter 5 was a fictional example, although it has been adapted from the veridical case of the palaeontologist who launched (and insisted in) the Nannotyrannus hypothesis. Now, here is a completely veridical example of epistemic inefficacy delivering the immediate harm just discussed. It's the tale of Gilbert Levin, one of the engineers ahead of the Mars Viking lander, a spacecraft sent to Mars in 1976 with the mission of searching for extraterrestrial life.

THE VIKING LANDERS
Hopes of confirming the presence of life on Mars were riding high when the twin Viking landers touched down on Mars in 1976. The scientific payload included the Labeled Release apparatus, designed by Levin and his colleagues, as well as three other life-detection experiments. The Labeled Release experiment, or LR, was set up to take a bit of Martian soil and add a drop of water containing nutrients tagged with radioactive markers. The air above the mix was then monitored to see if it gave off a radioactive gas such as carbon dioxide or methane. That could be read as an indication that organisms in the soil were metabolizing the nutrients. If the experiment came up with a positive response, a duplicate soil sample — the control — was heated to a temperature that should have been high enough to destroy microbes, but not to destroy any strong chemicals that might have produced a similar response sans life. The good news for Levin and the other life-hunters was that the LR experiment came out positive, and the control experiment came out negative. The bad news was that two of the other experiments came out negative, but they were based on different assumptions about potential Martian life. The *really* bad news was that the fourth experiment, conducted by Viking's Gas Chromatograph Mass Spectrometer device, or GCMS, didn't detect any organic molecules in the soil. The failure to find any organics led most scientists to assume that there was nothing living in the soil. Most scientists assumed that the LR findings were just a fluke. But not Levin.[67]

Levin thought that after the LR experiment the question about life on mars remained inconclusive, and that new experiments were needed to either prove him right in the future, or to present an alternative explanation to the LR results – some inorganic substance mimicking life, but not life. At that point, he was probably right. More research was needed.

In the four decades following the Viking mission, several proposals were raised for inorganic substances that could mimic the metabolism-like processes observed in the LR. One possible candidate is formate, which is a component of formic acid found naturally on Earth. A 2003 LR-type experiment found that formate in a soil sample from the Atacama Desert produced a positive result, in an environment with virtually no microorganisms (Navarro-González et. al. 2003). Nevertheless, this was on Earth. As much as the Atacama Desert's soil is as close as you can get to the Martian soil composition, it is still Earth. So the question remained open.

---

[67] Excerpt extracted in full from the NBC New's report "The Quest to Find Life on Mars: Been There, Done that?", published on March 27th 2012, available at https://www.nbcnews.com/science/ cosmic-log/quest-find-life-mars-been-there-done-flna561253. For an interesting discussion, see also Chapter 6 of Michael Brooks' 2008 book *Thirteen things that don't make sense: the most baffling scientific mysteries of our time* (Brooks 2008: 83-96).

Then came perchlorate, another potential candidate. In 2009, the Phoenix mission to Mars detected perchlorate in the soil. Perchlorate molecules could yield a positive result in LR-type experiments because they produce the relevant gas under some circumstances. But they do not break down when heated, which means they would continue to give positive results in the control experiment, contrary to what was observed in the original LR (Hecht et. al. 2009). So the issue remained.

Then, a 2013 study found that cosmic rays and solar radiation can cause perchlorate to break down into hypochlorite, which would also produce positive results. But, unlike perchlorate, hypochlorite is destroyed when heated, which means it would give negative results in the control experiment, just like happened in the original LR experiment (Quinn et. al. 2013). For these reasons, hypochlorite is arguably the best explanation to the LR results.

Now, this is a good candidate to a "turning point" in the history of this conundrum, that is, a good bid for settling the matter, right? Not for Levin, though.

In 2014, the Mars Science Laboratory Curiosity rover detected organic molecules on Mars for the first time (Webster et. al. 2014). Finding organic molecules would answer the main objection raised against the Viking landers' results. Levin interpreted this as meaning that the dilemma prevails – that microbes could not be ruled out as possibly responsible for the results obtained in 1976. However – and this is crucial – finding organic molecules is not the same as finding life. Organic compounds have been detected elsewhere, including in places known to be sterile, like the Moon (Sephton 2004; Klotz 2009); and they are not the same thing as living microbes.

The general consensus remains that no living microbes have ever been found on Mars. It doesn't mean, of course, that microbes don't exist on Mars – just that they are not what the LR experiment found. They're not the most likely responsible for the results observed in 1976. Nevertheless, Levin died in 2021 at the age of 97, without ever having been convinced that his experiment on the Viking Mars landers to detect microscopic life on Mars yielded a false positive. In a 2016 article, he was still claiming that microbes cannot be conclusively ruled out as an explanation for the LR experiment results (Levin & Straat 2016).

This is a real world example of epistemic inefficacy provoking an obstructive effect upon the consolidation of an epistemic good. In this case, the relevant epistemic good is a piece of understanding, namely, understanding that NASA hasn't

found evidence for life on Mars. Levin was a very respectable and competent scientist (which scientist who is less than this would be ahead of a NASA mission?). He hasn't been paid to cast doubt on the inorganic hypotheses. And, it is worth mentioning, his last paper, the 2016 paper, was peer-reviewed, in spite of advancing a very heterodox view. Nevertheless, his posture towards the dilemma around the interpretation of the LR results, a dilemma that indeed existed for long enough, made it more difficult for the information "NASA hasn't found evidence for life on Mars" to be official, in a sense. Or, at least, it made it more difficult for this understanding to enjoy wide acceptance as a robust fact.

For it looks as though from 2013 onwards it could already be considered a fact, which means, back in 2013 it already had what it takes to be treated as an official, solid conclusion. But Levin's treating the quarrel as if it was still an open case gave it the appearance that this was something science was still in the dark about, or still working out. That scientists didn't know – when they in fact did know, or were at least entitled to claim that they knew, that no microbes were there. (This apart from his conduct having had a few minor side effects, namely, that of generating a hype amongst ufologists, conspiracy theorists and other pseudoscience of extraterrestrial life enthusiasts, which is not of the essence in terms of what is being discussed here).

Levin persistent conduct of casting doubt on the conclusion that inorganic substances offered plausible alternative explanations for the LR experiment result, even though it was fine up to a certain point (because, like we've seen, some of those substances answered some questions but raised others), comes at a cost. Not to him, personally, or to him as a scientist, but to the entire scientific and especially to the extra-scientific communities, that would otherwise have access to an easier and polemic-free understanding of a fact. If every time you need to do a fact-checking you find yourself before big and intricate polemics that shouldn't really be there (aka: if you find yourself before a matter that is not consolidated, when it should have been), you're being charged this cost. In sum: that conduct misleads people into thinking that the issue is still being debated, or that it at least should be; that scientists are still uncertain, or at least that they should be, and so forth.

Delay in knowledge consolidation is the immediate negative effect inefficacy causes, it is something that comes directly as a consequence of this vicious conduct.

In the next section I'm going to delve into some of the ulterior costs that this pattern of conduct might generate. Those are the harms it causes indirectly.

6.3. Ulterior harms: the deterioration of epistemic goods

Apart from the detrimental effects just described, epistemic inefficacy has yet another, more severe, way of being noxious. The palaeontologist's case is one in which inquiry doesn't reach its natural termination point and doesn't have other termination constraints: his wait for the definitive proof can indeed go on and on, and lag in knowledge consolidation is arguably the biggest trouble it can cause. In a similar fashion, the Viking landers' case is a case in which inquiry eventually reached its termination point, in spite of Levin not acknowledging it; but it wouldn't be a problem if the termination point had taken longer to be reached. No ulterior harm apart from delay in consolidation would ensue, had that been the case.

But the same is not true of cases where an inquiry cannot go on indefinitely because it has some sort of "expiration date". In those cases, epistemic inefficacy can be more harmful: in addition to delay in consolidation, it can lead to the spoil, and even to the destruction, of epistemic goods.

Some types of inquiry have no natural point of termination, neither are they bounded by any external termination constraints. By external constraint I mean any factor, epistemic or not, that may force the agent to close inquiry[68]. So, for instance, if I want to know what my favourite red wine in the world is, I might go on tasting new samples until the last day of my life. This inquiry has no natural point of termination, nor external termination constraints, for even if I eventually taste all the red wines that exist and find a favourite one, new ones will keep popping up into the market, so my quest can literally go on forever. In an inquiry of this type, epistemic inefficacy causes the lesser harm.

Other varieties of inquiry have a natural point of termination, but no clear external termination constraints. The inquiry undertaken by Bob, the palaeontologist, might be said to belong to this category: in investigating whether or not

---

[68] For practical purposes I'm taking it that closure of inquiry is an intentional act. I borrow from Martin (2014) the definition of closure of inquiry as being "what one does when one decides to quit looking for evidence regarding the truth of some proposition. When one closes inquiry, one does not expect to reopen it (although one may)". Also, I assume that one need not be conscious of it *as* closing inquiry, and that it need not always lead to the acquisition of some belief by the agent whose inquiry it is.

Nanotyrannuses are juvenile T-rexes, the inquiry's natural point of termination is the discovery of the exact type of fossil that would be capable of solving the dilemma for good. If this fossil, however, for whatever reason, is not found, Bob's quest can keep going, because technically speaking nothing other than rationality itself will push him to stop gathering evidence and keeping it as an open case. Likewise with the Viking landers' inquiry: it has a natural point of termination, namely, it will be terminated when astrobiologists find the thing that appeared to be ingesting nutrients, metabolising them and belching out radioactive methane. But it has no external termination constraints, meaning that, if scientists never find out a suitable candidate, the quest for what that thing was could as well keep going. This type of inquiry is more vulnerable to epistemic inefficacy's detrimental effects than the previous type, which had neither external termination constraints nor a natural point of termination, because in this type of scenario consolidation is obstructed by epistemic inefficacy past the termination point, as we've seen.

Now there is a third variety of inquiry in which epistemic inefficacy can be considerably more pernicious. Those are the inquiries that have a natural point of termination and external termination constraints. For instance, if I want to know which companies' shares have the best benefit-cost ratio within the stock market today, this inquiry is supposed to be closed when I have evaluated a certain amount of reports on all the relevant companies' shares. But it is also constrained by the fact that the stock market closes at 5 p.m. So here, in addition to its natural point of termination, this inquiry has an "artificial" one, because if I don't gather and process all the relevant data until before 5 p.m., the efforts I've put into trying to answer my central question will have gone to waste, literally. I can pick and buy the best shares found *so far*, but in the long run (if the flaw persists) my underlying goal of maximizing profit will be compromised. The closing of the stock market, therefore, is like an "expiration date" for my present inquiry.

In this particular case, the external termination point is quite an explicit one: I know that it will come at 5 p.m. But it might as well not be. For instance, if I want to find out what is the best strategic approach to save a business that is on the verge of bankruptcy, this inquiry will end when I do reach a conclusion, but the business might as well collapse *before* that. The collapse of the business is an external constraint on the termination of this inquiry, and it's a much less explicit one, inasmuch as this is not the sort of event than can be predicted with great precision. It is in this type of

inquiry that vices that are ultimately a matter of timing, such as epistemic inefficacy, cause the harsher effects.

The crucial element in those last two cases, the stock market case and the collapsing business' case, that is even more severe in the latter, is that the "artificial" point of termination can arrive earlier than the natural one, forcing inquiry to end before cognitive success is attained. It is as if their "expiration date" in unmarked. Here is where epistemic inefficacy can cause the greatest damage, to the extent that this vice is about postponing the drawing of certain conclusions past the reasonable point, and this will only increase the likelihood of the "expiration date" coming earlier than cognitive success.

What's being described is a type of "ticking bomb" scenario: a scenario in which the agent has to run against time, because something important is bound to be destructed, and your job there is to figure out a way of avoiding that. You can fail in many ways: because you needed resources that you couldn't access, because you fall ill and this keeps you from doing your job as an inquirer, because no ideas come to you and you run out of time, and so forth. And you can also fail because of personal shortcomings, that is, vices.

Now, in order for epistemic inefficacy, as well as to any other vice, to cause a damage in those scenarios that is a distinctively *epistemic* damage, the thing that is bound to be destructed has to be an epistemic good, of some sort. How tenable is this idea? Are there scenarios of this sort, that is, "epistemic ticking bomb scenarios"? Are there cases where the damage caused by failure in those scenarios can be explained by the vice of epistemic inefficacy? And how is this idea of "destruction of knowledge" to be understood, precisely?

The idea of "ticking bomb scenario" is draw out of the so-called "ticking bomb thought experiment", which is an argument that appears sporadically in academic literature and in the media, within discussions on the topic of torture. The scenario is normally brought about with the purpose of making the case for the necessity of waiving the prohibition against torture under exceptional circumstances (usually, in the case of terrorism suspects or as a means of intelligence-gathering in warfare). As Wolfendale (2006: 271) points out, in the standard ticking bomb scenario, a suspect possesses information that must be obtained quickly in order to avert massive civilian casualties – he knows, say, the whereabouts of a ticking bomb that has been planted in the middle of a big city by a terrorist group, but would rather die than collaborate.

The proponent of the argument then proceeds to a defence of the claim that under these circumstances the use of torture as a means to obtain the crucial information would be legitimate. Terrorism is claimed to pose such an extreme threat that the prohibition against torture cannot be maintained.

Variations of the ticking bomb argument have been put forward by writers such as Alan Dershowitz (2004, 2016), John Parry (2004) and Mark Bowden (2002), mainly as utilitarian justifications for admitting exceptions to the prohibition against torture. Other times, the ticking bomb scenario is presented with the aim of discussing the problem of dirty hands decision-making in politics as, for instance, in Walzer (1973). The expression "dirty hands" is inherited from Jean Paul Sartre's play *Les Mains Sales* ("The Dirty Hands"). The main character in this political drama, Hugo Barine, is in a quest for executing a tyrant leader before he proceeds to putting into practice a set of policies that will drag the nation into ruin. As such, he deals with the moral dilemma that results from getting one's hands dirty, i.e., doing something wrong in the process of doing "the right thing", the thing that will deliver the greatest good. Some people, like Walzer, think that, in politics, dirty hands are inevitable and, at times, even required. According to him, the very nature of politics sometimes demands that the politician transgresses his own moral code, and ours, "for the sake of the people who might otherwise die in the explosions" (1973: 167).

Even though descriptions of the ticking bomb scenario vary significantly, the many different accounts of this hypothetical scenario normally have three elements in common: (1) something very valuable is under threat; (2) the threat is imminent, i.e., time is of the essence; and (3) it is possible to prevent the catastrophe by doing something that is considered suboptimal. As Bufacchi & Arrigo (2006: 358-359) remark, the ticking-bomb argument is so hyperbolical that it has more affinities with science fiction than with political science, which leaves us to wonder about the true extent of its usefulness for those interested in understanding the real world. Nevertheless, ticking bomb scenarios are capable of giving us interesting insight about an important issue: the issue of how the norms of conduct can be different in contexts where time is crucial.

If a bomb is not ticking, so to speak, it becomes much more difficult for someone to make the case for the use of torture as a means of gaining knowledge, because in this case the authorities would not be deprived of the means to conduct a thorough search themselves. They would have plenty of time to locate the bomb

through their own efforts and means, without the collaboration of the alleged suspect and, therefore, without the need to violate any moral code in the process.

Setting aside the issue of the permissibility of violating human rights that originally lies at the root of the ticking bomb thought experiment, the interest insight that can be extracted from the experiment itself is the following: true scarcity of time has the power of altering our perception of the inherent values of things, as well as the reasonability of our conducts. Something that wouldn't be deemed the best thing to do if we had plenty of time could in fact become the wisest move if time becomes an issue, depending on how high the stakes are. When the likelihood of something very valuable (a good) ending up lost or damaged increases as time runs short, we feel pushed, in a sense, into doing things that we wouldn't normally do, or wouldn't normally consider necessary, for the sake of safeguarding the relevant good. Conversely, some things that we would normally dear can be instantly devoid of their perceived value, if we realize that time wouldn't be enough for them to be enjoyed. This is why retailers, for instance, put groceries whose sell-by date is drawing near on offer with substantial discounts.

The crucial insight that the idea of ticking bomb scenario gives us, then is the following. Whilst real ticking bomb scenarios are arguably rare, and whilst it is debatable whether or not strictu sensu ticking bomb scenarios even exist, it is true that many times, in our ordinary lives, we alter our appreciation of things, including our conducts, as a consequence of our appreciation of the exhaustibility of time. In other words: the intrinsic value of certain things change, according to the amount of time we have "left".

Sometimes we submit what turns out to be an underdeveloped manuscript because we realize we won't have time to perfect it any further, and the consequences of not presenting anything are far worse than those of presenting a suboptimal version. Likewise, any person who was once desperate for a job would agree that, if for whatever reason you are short on time, it is better to show up for a job interview with your clothes dirty from some coffee you spilled than to not show up or to take the risk of arriving late; to continue with the mundane examples.

The point I'm trying to make is: you wouldn't normally consider presenting a substandard manuscript to a call for papers, nor showing up for a job interview wearing dirty clothes. However, due to the scarcity of time, you start to see those upshots as tenable options, because rejecting them means jeopardizing your

chances of attaining some desired good (a publication, in the first case, or a job, in the second). If this is true, then we are encouraged to think that virtually every case in which time plays a role in decision-making is a soft variation of a ticking bomb scenario. This, in turn, means that, far from rare or fictional, ticking bomb scenarios are rather prevalent in our lives. This includes, of course, our epistemic lives.

Now, does it make any sense to talk of *epistemic* ticking bomb scenarios? For this to make sense, the good under imminent threat has to be epistemic, more so than practical. This happens, sometimes. That is, sometimes it is a piece of knowledge that is at risk of being destructed. The most obvious and low-level examples of this involve the burnt of sources. In *Burning the Books: A History of the Deliberate Destruction of Knowledge* (2020), Richard Ovenden comes up with a myriad of examples. His focus is on knowledge of the past. Starting with ancient Mesopotamia and ending in Facebook and Twitter, Ovenden details historical episodes of threat to our knowledge of past events that comes with the attempts at burning, burying or deleting the texts through which the story of humanity has been documented. When a unique written record is destroyed, it is not just the physical object that ceases to exist. A piece of knowledge is also gone.

Other times, rather than the past, it is knowledge of future matters (that is, knowledge that doesn't exist yet) that happens to be under threat. Solutions to problems of fact, or answers to questions, that haven't been found out yet, or worked out yet. For instance, when laws are passed prohibiting stem-cell research, an important epistemic good, or at least an important avenue to an important epistemic good, is made unavailable: increased understanding of diseases and of their potential cures. One of the points of stem-cell research is that it leads scientists to expand their understanding of how certain medical conditions develop and progress, which may or may not lead them to be able to work out a solution for them. When this type of research is interposed, potential answers to the question of what might work and what might not work to cure diabetes type one and lateral sclerosis (currently incurable), for instance, are interdicted. What would otherwise become a piece of knowledge never actually comes into being – it is nipped in the bud.

Also, there are cases where it is not avenues to knowledge, but rather knowledge usefulness, the ability that knowledge has of being of service, that is nipped in the bud. Somewhat like in the case of burnt sources, it can happen that a piece of knowledge that already exists has its possibility of being enjoyed

deteriorated. Some of the most expressive examples of this occur in lying or omitting. In *Midnight in Chernobyl: the untold story of the world's greatest nuclear disaster* (2019), Adam Higginbotham describes a situation of the like.

Three days after the explosion at the nuclear power plant in Chernobyl, in 1986, not one word had appeared in the Soviet press or been reported on radio or television. An emergency meeting of the Politburo was assembled. Their aim was to decide what – or whether – to tell the Soviet people about the accident. In spite of the broad agreement among the Party elders on the need to make a public statement as soon as possible, Yegor Ligachev, Gorbachev's emerging conservative opponent and the second most powerful man in the Kremlin, voted against it. Following him, others at the table argued that they didn't have enough information yet to tell the public, and feared it could cause panic. (And, of course, they feared that the statement would feed anti-communist propaganda in the west.) "By the time they took a vote", Higginbotham remarks, "Ligachev had apparently prevailed: the Politburo resolved to take the traditional approach. The assembled Party elders drafted an unrevealing twenty-three-word statement to be issued by the state news agency" (2019: 175-176).

Bottom line: the summit released information about the disaster, but in a way that concealed its true dimensions, as well as relevant facts in connection to it. For instance, the fact that a cloud of radiation was hovering above East Europe and moving south-west. This cloud carried microscopic fragments and particles composed of pure radioactive isotopes, including iodine 131, a chemical known to cause thyroid cancer in children. Because the cloud was not mentioned in the statement, the soviet people, as well as and other soviet leaders and leaders of western nations were denied knowledge of the true extent of the Chernobyl disaster and of its potentially harmful sequels[69].

Now, because you cannot see, smell, or taste the deadly radiation, unless somebody that knows that the radiation is there warns you, it is very unlikely that you are going to figure it out that you have been exposed, or is about to be exposed to it, until it's too late. What is more, iodine 131 has a half-life of 8 days. This means that it would take 8 days for the radioactive material rolling in the toxic cloud to be reduced to half of its initial amount, and then another 8 days for that amount to be reduced to

---

[69] The presence of this moving cloud was known by Soviet authorities at the time, because it had already been reported in Lithuania (Higginbotham 2019: 173).

half, and so forth, until it eventually wears out. Those initial days are critical, because it is when you get the highest exposure and, therefore, it is when you should be getting most protection. But all that the population was told is that there had been some accident at the plant – the Politburo's statement refrained from mentioning when (Higginbotham 2019: 179), so people were denied knowledge that they were within the decisive 8-day window.

The Chernobyl case is a case of deterioration of knowledge. That's because the value of that piece of knowledge, or its usefulness, is intimately tied to its possibility of reaching those to whom it may concern within the decisive time window, which never happened. After the first 8-day window, the value of this knowledge is drastically reduced, and this is the equivalent to its deterioration.

Also, the Chernobyl case is an epistemic ticking bomb scenario, because in it we can see all three distinctive marks of this type of scenario: first, some distinctively epistemic good has to be under threat. Roughly, either a piece of knowledge that already exists has to be at risk of disappearing, or a piece of knowledge that hasn't been found yet has to be threatened of never being found. In the Chernobyl case, its the former: it's information about iodine 131 that science already knew. Second, the threat has to be imminent. That was the case. After the relevant 8-day window, that information would deteriorate. And third, the damage has to be such that it could have been averted if somebody had been willing to lower his or her standards of what is acceptable. That's the case too. The event leading to the deterioration of the relevant piece of knowledge was preventable. It could have been prevented, had the Politburo taken the measure that would, in that context, be considered suboptimal: releasing the data.

The Chernobyl case shows us that the idea of epistemic ticking bomb scenario is tenable. There are real world cases of epistemic goods that deteriorate because they couldn't be saved from expiring. Nevertheless, in the Chernobyl example, what kept the relevant epistemic good from being saved from deterioration was not vices. Arguably, in this case, even though names have been pointed to (i.e., Yegor Ligachev), explanation of what went wrong, or of why "the bomb" hasn't been defused, is more truthful if we resort to structural factors rather than to personal flaws.

Higginbotham describes the standard soviet environment as a "deeply ingrained" culture of "secrecy and paranoia" that shaped the soviet attitude towards

communication at all levels (2019: 176). The truth about incidents of any kind that could undermine soviet prestige or cause public panic used to be routinely suppressed. They denied Holodomor, the great famine that hit Soviet Ukraine in 1932-1933, for more than five decades. They denied the 1957 disaster in the nuclear plant of Kyshtym, and kept denying it even thirty years after the explosion. They denied the Korean Air Lines incident, in 1983, when a Soviet air force pilot mistakenly shot down the jet, killing 269 people. They even denied people their own medical records, as a young soviet journalist once declared: "in the USSR we do not tell a patient if he has cancer" (Vitkovskaya 2016). So it makes sense to think that the explanation of why the relevant epistemic good couldn't be saved from deterioration, in this case, has more to do with this culture, than with personal traits of the agents involved.

Notwithstanding, there are cases with similar contours, that is, cases of epistemic ticking bomb scenarios, where failure to defuse "the bomb" is better explained by vices, in general, and by epistemic inefficacy, in particular. It is in those scenarios, I submit, that epistemic inefficacy has the potential to be the most harmful. The case I present in the next section exemplifies this.

6.4. The race for vaccines case

To see an illustration of epistemic inefficacy in an epistemic ticking bomb scenario, consider another case, the race for vaccines case.

THE RACE FOR VACCINES
On October 19, 2020, Brazilian medical doctor Alessandro Loiola appeared live in an interview broadcasted by the Brazilian TV channel RedeTV to speak against the race for an anti-COVID-19 vaccine. According to him, due to the process having been extremely rushed, immunization may not be safe. "To be concerned with effectiveness prior than safety (of the anti-COVID vaccine) is absurd", he said, referring to the Brazilian Health Regulatory Agency, the ANVISA (the Brazilian equivalent of the American FDA, or the British MHRA). "ANVISA needs to be incisive in its inspection. If there is someone with a minimum of moral decency within the agency, they will not be authorizing any vaccines for public use with a period of clinical trials shorter than four years", he asserted.[70]

Loiola's point is that before seeking to know a vaccine's benefits (that is, whether it is effective), ANVISA should seek knowledge of it risks (whether it is safe).

---

[70]The relevant portion of the interview can be watched at RedeTV's official YouTube channel, through https://www.youtube.com/watch?v=uD5NQW6CcTc .

Knowing that a vaccine actually prevents COVID-19 is of no use if in addition to preventing COVID it also causes cancer, modifies human DNA[71] or turns the person into an alligator[72].

Now let's pretend for a while, just for the sake of the argument, that the person ahead of ANVISA thinks like Loiola, or that Loiola himself is ahead of ANVISA. This person is certainly aware that inquiring into a vaccine's safety is a considerably bigger enterprise than inquiring into its efficiency (that's why he estimates that the former would take so much longer – no less than four years, according to his estimate).

The reason why this is so is that efficiency is more like a yes or no question, while safety is a much more open-ended question, since you want to understand which side effects the vaccine causes, if any; at what point (meaning: when) they upsurge, how bad, or how dangerous they are; whom they affect the most, and why each of them affects some people but not others. In sum, while determining efficiency means focusing on one issue (the issue of whether the substance meets a minimum pre-established requirement in terms of its ability to prevent infection), determining safety requires considering a constellation of issues. As such, determining safety requires a formidably greater amount of empirical data and, therefore, takes more time.

Being ahead of ANVISA, a person is well aware that the COVID-19 crisis is the worst humanitarian challenge ever faced at least since World War II. And he also knows that our best hope to overcoming this crisis lies at finding effective and safe substances to fight the disease, including vaccines, if, of course, we are to overcome it while avoiding mass human losses.

Attaining such a goal, in turn, involves dealing with a delicate equilibrium. On the one hand, creating and approving a vaccine too fast may result in releasing a substance that is neither much efficient nor safe, in the worst case scenario. This would mean acting recklessly, wasting resources and undergoing avoidable risks, which translates into increased losses, both human and economic.

---

[71] This claim was made by Loiola's in another talk, three days before his interview to RedeTV, which was widely spread through social media.

[72] This claim was made by Brazilian President Jair Bolsonaro the day before Loiola's interview to RedeTV, while (apparently) trying to make the same point about benefits *vs.* risks. For details on this one, see the cover by the *New York Post* at https://bit.ly/3mRs1dj .

On the other hand, trying to be too careful and perfectionist by only approving a vaccine after it has been tracked for *every* collateral damage conceivable, so that you actually know how risky the substance is in terms of each and any one of them, will inevitably take too long. This would involve allowing the pandemics to go rampant for an extended period of time, which also translates into increased losses, both human and economic. But this demeanour has an additional consequence, which is what interests me the most here: it involves an increased chance of the virus suffering mutations and becoming, essentially, a different pathogen. The virus becoming a different pathogen acts like a possible "expiration date" for this inquiry, because it changes the whole picture of the epistemic endeavours at play, and can possibly make things so that the correlations eventually established cease to hold true.

In the former case (releasing a vaccine too fast), one fails by not knowing enough. In the later, one fails by wanting to know too much. To do its job right, ANVISA needs to be *reasonable*. It cannot be either too demanding or not demanding enough. It has to be somewhere in the middle.

So a couple of months go by. As it turns out, some plausible vaccine candidates start to appear, as over 100 research teams worldwide put massive effort into it. Pfizer-BioNTech, in Germany, the University of Oxford-AstraZeneca in the UK and Sinovac in China are among the first to release positive results. After phase III clinical trials involving randomized, placebo-controlled international studies with thousands of participants are successfully completed, developers can finally submit applications for emergency use authorization of their vaccines within regulatory agencies.

The emergency use authorization protocol varies from country to country, as regulatory agencies are national and have autonomy to proceed according to their own norms and criteria. Now, Loiola must have been quite confident that no regulatory agency "with descent people within" would be authorizing any vaccines before at least 4 years of clinical trials. Therefore, it might have come as an absolute surprise to him when, on December 2nd 2020 (with less than 1 year of clinical trials, thus) the UK's MHRA granted emergency use authorization for the Pfizer-BioNTech vaccine, becoming the first in the western world to authorize emergency use of an anti-COVID-19 vaccine. He might also have received with surprise the news, less than 2 weeks afterwards, that the US and the European Union followed the UK in this deliberation.

Here is a piece of the statement made by Peter Marks, director of the FDA's Center for Biologics Evaluation and Research on December 11th, 2020:

> While not an FDA approval, today's emergency use authorization of the Pfizer-BioNTech COVID-19 Vaccine holds the promise to alter the course of this pandemic in the United States. (...) With science guiding our decision-making, the available safety and effectiveness data support the authorization of the Pfizer-BioNTech COVID-19 Vaccine because the vaccine's known and potential benefits clearly outweigh its known and potential risks. The data provided by the sponsor have met the FDA's expectations as conveyed in our June and October guidance documents. Efforts to speed vaccine development have not sacrificed scientific standards or the integrity of our vaccine evaluation process.[73]

Doctor Loiola must have looked at Peter Marks' statement with remarkable skepticism. "Well, no doubt the FDA is a decent institution. But there's no way for anyone in the world to truly know, at this time, whether this vaccine's benefits outweigh its risks. Because there isn't enough data", he may have thought to himself.

As it turned out, in the following days more and more statements similar to Peter Marks' kept popping up, as hundreds of other regulatory agencies around the world authorized the Pfizer-BioNTech vaccine for emergency use. Besides the UK, the EU and the US, by December 31st this vaccine had been granted emergency use authorization by regulatory agencies in Argentina, Canada, Chile, Costa Rica, Ecuador, Iraq, Israel, Jordan, Kwait, Mexico, Oman, Panama, Qatar, Saudi Arabia, Singapore, United Arab Emirates and Switzerland. It had also received a recommendation for emergency use authorization by WHO, the World Health Organization[74].

It soon becomes clear that Loiola is stuck within a sort of deadlock, between the following two claims:

p: the vaccine's benefits outweigh its risks.

and

p': at this time, nobody knows whether the vaccine's benefits indeed outweigh its risks.

---

[73] This was published in the FDA website, at https://bit.ly/3tqf4tn .
[74] This can be checked at the WHO website, via https://bit.ly/3rGMgLj .

The fact that those more than twenty regulatory agencies who deliberated in favour of the emergency use authorization for the Pfizer-BioNTech vaccine are high-standard institutions and AVISA's epistemic peers (meaning they are competent, reliable and decent, by all accounts) speaks in favour of **p**, while the fact that the problem being dealt with is all too new (meaning data might not be enough to sustain an optimally responsible deliberation on the matter) weights in favour of **p'** and against **p**.

While not contradictory in their own right, **p** and **p'** are in conflict here, inasmuch as even if one can be evenly confident in both at a given time, he cannot dwell on this situation for long. It can be tolerated for a while, but in time has to be resolved, because there will come a time when the individual will have to take action, and positive action can only conform to one side of the conflict, not both. In other words: a critical moment will come when the person ahead of ANVISA will have to answer Pfizer's submission. He'll have to either sign a document authorizing the vaccine for emergency use, *or* sign a document refusing to concede such authorization.

He decides, then, to wait until some further data from Pfizer-BioNtech is made available before following the other countries in their decision. Meanwhile, more and more countries go on authorizing the vaccine (and not just Pfizer's but also others) for emergency use. After a while, the requested data is made available. Does ANVISA have enough data to sustain a responsible deliberation now? He hesitates. After all, the information just received is not conclusive. It answers some questions, but raises others.

He is well aware that time is running short, and that what he is facing is an emergency. But he is afraid it's simply too early to take a stand. He acknowledges that the fact that a plethora of countries have already granted emergency use authorization for that vaccine means something – it means the jab is more likely than not to be safe enough, but he wants certainty. If adamant certainty is not possible, then he wants to have at least less uncertainty than he has now. So he decides to wait until some more data from Pfizer-BioNtech is made available. And thus the tale goes.

Fortunately, this man was not ahead of ANVISA and this extremely distressing situation didn't go on for too long in the real world. The agency actually approved the Pfizer-BioNtech vaccine on February 23rd, 2021. This anecdote, however, gives us

good food for thought: we can now picture the calibre of the damage that could have ensued had it been Loiola, or a like-minded person, ahead of ANVISA[75].

A scenario in which the quandary described above, between **p** and **p'**, is dealt with in the way Loiola does, would be a scenario in which he, or whoever was ahead of ANVISA with his philosophy in mind, would be displaying an array of vices, some of which epistemic in essence. One of them would be epistemic inefficacy, which is visible through excessive caution with which he deals with the dilemma, even past the point when it should be considered settled. This caution is possibly an echo of his residual commitment to the conservative idea that no vaccine should be approved with less than 4 years of data gathering, that's being echoed in **p'**, which is the side of the dilemma that has been overtaken. This residual commitment leads him to cope with the epistemic conflict for longer than it would be right. Specifically, the vice is visible through his posture of lingering evidence gathering past the point when it should have been phased out.

To be sure, this is different from both close-mindedness and excessive open-mindedness, for the same reasons discussed in Chapter 5: Loiola is not closed to the possibility that the vaccine is safe, but he is not open to this possibility in the right way either. His demeanour is also different from motivated skepticism, of the sort seen in conspiracy thinking, for instance[76]. This sort of skepticism, through blunt distrust that an anti-COVID-19 vaccine will ever be safe, was actually undertaken by

---

[75] A quick remark, as of a side note: interestingly, ANVISA's approval of the Pfizer's vaccine was made under the agency's standard protocol, rather than the emergency use protocol. That is, on February 23rd, 2021 this vaccine got the definitive registration, instead of the provisional one. With this, Brazil's agency made the third fastest definitive registration for the Pfizer-BioNtech anti-COVID-19 vaccine in the world. So apparently people ahead of ANVISA do not play it by doctor Loiola's book, on the contrary: it took them much fewer data and much less time than Loiola's initial estimate, and the process was indeed much quicker than the average definitive approval, which normally takes years of data gathering, not months. Even more interestingly, as Pfizer-BioNtech itself declared to the press, the company did not submit a request for emergency use authorization of its vaccine in Brazil and preferred to go ahead with the submission for a definitive registration because the latter would be quicker [!] than the former. Now, how come *emergency* authorization turns out to be slower than standard approval? In the end, we are truly left to wonder whether the "Beaureau of Emergency Authorizations" inside of ANVISA isn't indeed run by a Loiola-like individual.

[76] A much discussed and also paradigmatic variety of motivated skepticism is directed towards science and related to conspiracy thinking. It amounts to the cases in which lay people are especially distrustful of conclusions that scientists agree upon when it comes to problems such as climate change and rising crime rates. The structure of those cases does not seem to vary much relatively to the cases involving people receiving disease diagnosis that were discussed in the previous section: the individual prefers a world in which she didn't have to live with the solutions required to deal with such problems and, therefore, tends to see with more skepticism the evidence pointing towards the problems being real. For an interesting description of these cases, see Campbell & Kay (2014).

authorities from countries such as Tanzania and Madagascar[77]. These authorities didn't want more time or more evidence – they were set up, from the very beginning, to refuse vaccines.

And, finally, this is also different from ordinary indecisiveness, in spite of some convergences. If we look at the literature on indecisiveness from social psychology, we see that typical indecisive individuals worry too much about making mistakes, and postpone or avoid decisions in order to minimize the risk of making a mistake (Frost & Shows 1993; see also Salzman 1980). They also tend to interpret ambiguous situations in a "worst case scenario fashion" (Rassin & Muris 2005). Because of this tendency, they experiment anxiety which, in turn, leads them to be ultracautious, resulting in a "better safe than sorry" decision making style (Rassin & Muris 2005: 1286).

Loiola might have been indecisive in a loose sense, because he too seems worried about making a mistake, and he too is ultracautious. But his indecisiveness is of a peculiar kind: he acknowledges that the "best case scenario" is more likely than the "worst case scenario", that is, he understands that **p** is *more likely* true than not true. The problem with him is that he sets the bar too high, in terms of how much more likely he demands it to be, or of how much evidence of this higher likelihood he wants to see before he makes a final call (the call to drop **p'**). So it's not that he doesn't know what to do, like typical indecisive people – he does know. What he doesn't know is when to do it.

Additionally, he keeps wanting more reassurance, so his indecisiveness has a somewhat cyclical aspect to it. Every time he receives more data, he wants more data. To refer back again to the analogy with practical akrasia: he lacks the strength required to break the cycle of wanting more data, just like the classically akratic person lacks the strength of willpower required to break the cycle of always picking one more piece of a chocolate cake.

In sum, Loiola does know that it's best to have the vaccine than to not have it, which requires that he deliberates in favour of the vaccine's authorization for emergency use at some point. What he doesn't know, or seems to not know, is what point this is. What *the right time* to make such a deliberation is, and what the

---

[77]See *Tanzania refuses COVID-19 vaccines*, published in The Lancet (Vol. 397 February 13th, 2021: 566), available at https://bit.ly/3dxCKVI. See also *Madagascar takes last stand on Covid-19 vaccine, refuses immunization* at https://bit.ly/2PMgrDH .

requirements for reasonable deliberation in each circumstance are. Under normal circumstances, that is, if this highly mutagenic virus was under control and hadn't outbroken into a pandemic, perhaps his original estimate (4 years of data gathering) would be a reasonable one, in terms of how long it would take for a substance to have its provisional authorization approved. But in the actual circumstances, having this estimate as a parameter simply is not reasonable.

Now, why exactly does Loiola's conduct in this case pose a threat that is distinctively epistemic, rather than just practical, or moral? No doubt his conduct poses threats that go beyond the epistemic domain. The most remarkable of them is moral in essence: every new day spent in the quest for extra data that won't really make a difference in terms of the amount of certainty he wants to attain is an extra day without a vaccine. And without a vaccine, people die. But how exactly does this translate into *epistemic* damage?

An answer to this question has to do with the variety of inquiry Loiola is involved with. The race for vaccines scenario is an epistemic ticking bomb scenario. There is a distinctively epistemic good under threat of being deteriorated, the threat is imminent, that is, time is of the essence; and the catastrophe can be prevented if the agent is willing to take measures that he considers suboptimal. What distinctively epistemic good is this? It's knowledge of the current viral strain, that cumulated for the formulation of the current vaccines. This knowledge was attained after immense cognitive effort undertaken by scientists all over the world, and it is at risk of deteriorating if Loiola's is not willing to compromise. Let me explain this in a more detailed way.

Loiola's inquiry into whether the anti-COVID-19 vaccine's benefits outweigh its risks is supposed to be closed at a certain point in time, namely, when a big enough sample of the population have received the jabs and had their side effects monitored, because then it will be made clear whether its benefits actually surpassed its risks. But in addition to this natural point of termination, it is constrained by the possibility that the virus would suffer mutations. No one can predict when and if this will occur, but it is a known fact that the longer a virus spends circulating among a non-immunized population, the greater the odds of it suffering mutations in its DNA that would result in it becoming, essentially, a different pathogen[78].

---

[78] Science shows that the Sars-Cov-2 virus is significantly more stable compared with others of its kind, meaning it mutates less (Afonso 2021). This, however, does not make it exempt from generating

The virus becoming a different pathogen is "the bomb", or the "expiration date" (one of the possible ones) of the inquiry leading to the development of the Pfizer-Biontech vaccine, as well as the other COVID-19 vaccines. It becoming a pathogen that escapes vaccine's coverage, especially, would be a fatal expiration constraint. That's because if it becomes a different pathogen, important epistemic goods scientists all over the globe have worked relentlessly hard to attain in the months following the outbreak would have their intrinsic value drastically decreased – if they don't become completely useless.

These goods include, naturally, information that cumulates for the knowledge of how efficient and safe the current vaccines are (which was Loiola's primary goal), because in the face of a modified pathogen a vaccine might have its efficiency, as well as its safety, changed. And they also include the vaccines *themselves* qua an epistemic good. Just like a book is not simply a heap of paper sheets and ink, a vaccine is not just an aqueous solution infused with biocomponents; that is, it's not just a material good. It's also an epistemic good, insofar as it is the answer to the question of how to solve a problem[79].

Inquiries are "activities with distinctively epistemic goals: they are directed at solving problems of fact, at finding things out" (Hookway 2001: 178). Those anti-COVID-19 vaccines are the attained goal, that is, the successful result, of a long and complex chain of inquiries – the inquiry aimed at solving the problem of how to hold the COVID-19 pandemic back. In this sense they can be said to be an epistemic good in their own right: their formula is the solution to a problem, the answer to a question; one that that has to have been worked out by means of tremendous cognitive effort.

Many different things can be epistemic goods. The most popular candidates are things like justification, warrant, coherence (Roberts & Wood 2003); true beliefs

---

strains of concern – which actually happened. See for instance Plante et. al. (2021), Volz et. al. (2020) and Korber et. al. (2020). The disagreeable fact that viruses in general mutate a lot, that has been known by scientist for decades and has even become a frightening piece of folk biology, was somewhat turned into common sense knowledge since it was widely emphasized by media of various types following the COVID-19 outbreak. See, for instance, the BBC's July 2020 article *Coronavirus: Are mutations making it more infectious?* (available at https://bbc.in/35Vfb63). See also Deutsche Welle's June 2021 article *Virus variants in Asia threaten the whole world* (available at https://bit.ly/3jqUsPg).

[79] This might be rephrased in a more accurate fashion if we say that a book's *content*, as well as a vaccine's *formula*, are epistemic goods (rather than the book itself, or the vaccine itself). It doesn't make a big difference for the sake of the argument being advanced, though.

(Sosa 2003), particular intellectual virtues (Zagzebski 1996); understanding of important truths (Riggs 2003); wisdom (Baehr 2012b); and pieces of knowledge (Zagzebski 2003), including knowledge that is instrumentally valuable in the realization of particular ends (de Bruin 2013). All of these things are said to be epistemic goods because they promote, or are in other ways constitutive of, our intellectual well being. Intellectual well being too can be understood in a variety of ways[80]. Arguably, one of these ways has to do with the notion of care (Zagzebski 2004). When we care about something, we care about having knowledge in the domain of that something (Zagzebski 2004: 356). Having such pieces of knowledge increases intellectual well being, not having them decreases it.

In this sense, all products of scientific endeavours can be said to be epistemic valuable in themselves, insofar as they are the successful results of inquiries in domains that we care about. They are things worth preserving, items that must be protected against spoiling[81]. Thus, to the extent that we care about the COVID-19 pandemic, and to the extent that the anti-COVID-19 vaccines of which we are speaking comprise the answer to an important question in that domain, they are an epistemic good. Having this good adds to our intellectual well-being, not having it subtracts from it. Having once had it and then lost is, or being deprived from it for whatever reason, is an even greater subtraction: it means cognitive effort put into acquiring this good went to waste.

In conducting himself the way he does, Loiola's threatens to spoil this epistemic good – he risks the very vaccines he is trying to evaluate becoming less useful, or even useless, due to virus mutation. As the common flu illustrates quite well, because the influenza virus constantly mutates, a new version of its vaccine has to be made each year, since formulas that used to work become outdated[82]. As much as he wants to be cautious and conservative and to avoid damage, what Loiola's posture does is: it increases the risk of turning COVID-19 into the new common flu: a

---

[80] Intellectual well-being might be equated with the recognition of when we do or don't know something (Zagzebski 1996: 267). It can also be thought of in terms of the overcoming of anxiety due to incorrect presuppositions about knowledge, or about what one should know (McDowell 1994: xi; Pritchard 2016). It can also have to do with mitigating epistemic injustices (Fricker 2007), to name a few.

[81] Which does not mean, of course, that one must not try to update and transform those goods with the aim of improving them.

[82] An article published by Time Magazine in 2017 has an excellent review of this: *We Are Not Ready for the Next Pandemic* (Walsh, 2017).

disease which we cannot eradicate and to which we need, rather, to be constantly working vaccines out.

We might describe Loiola's conduct in an alternative manner by evoking Hookway's notion of a "subordinate question", alluded in Chapter 1 (Hookway 1994). When carrying out inquiry, we set ourselves a question and execute a patterned sequence of actions designed to enable us to arrive at a satisfactory answer to that question (Hookway 1994: 223). Those actions include raising a number of other questions, what Hookway calls "subordinate questions" (1994: 214). To ensure satisfactoriness of an inquiry, one must raise the correct subordinate questions and answer them in a responsible manner.

An epistemically inefficacious person such as Loiola errs either in not asking or in not answering, at the right timing, a rather crucial subordinate question: the question about whether the prospects of increasing certainty in the resolution of some particular quandary by lingering evidence gathering indeed overcomes the costs of so proceeding. Because Loiola reaches a point from which the responsible answer to this question, had it been raised, would be *no*: the costs of so proceeding are too high. By lingering the stage of evidence gathering in the way he does, he risks not only jeopardizing his own present inquiry, but also deteriorating important epistemic goods already attained by others.

This doesn't entail of course that Loiola's conduct has the power to turn knowledge into not-knowledge, or to make us forget things; is not in this sense that his conduct might lead to the deterioration of pre-existent epistemic goods. The fact, for instance, that said and said vaccines are effective against such and such viral strain does not cease to hold true in case new strains arise. This fact continues to hold true. What changes in case new strains arise is the broader picture in which this fact is embedded, and, thus, its significance, or its value. The upsurge of certain new strains makes things so that these particular formulas cease to offer a good answer to the question of how to hold the pandemic back. In other words, Loiola's conduct risks spoiling these vaccines qua an epistemic good because it risks removing from them the capacity to solve the problem they were meant to solve, and throwing scientists back to square one in terms of this very problem, which will have to be worked out all over again.

6.5. Final remarks

We're now in position to update the chart presented at the end of the Chapter 5, comprising the main prototypic traits of epistemic inefficacy as an epistemic vice.

1. Epistemic inefficacy is an attitude towards quandaries.

2. Epistemic inefficacy's grounding experience is the experience of lack of closure in the context of a quandary, that is, the experience of an unsolved dilemma.

3. Epistemic inefficacy is a flaw in timing. Inquiring involves making certain deliberations. Epistemically inefficacious people fail at deliberating within the correct timing, while in the process of inquiring towards resolution of a dilemma.

4. Epistemically inefficacious people are in good faith. They want to solve the dilemma, but are kept from doing this in an effective manner because of their residual commitment to one of the terms in the dilemma (the hypothesis that is "losing").

5. Epistemic inefficacy obstructs knowledge by retarding consolidation (the social process by means of which claims that enjoy sufficient evidential support are converted into profitable epistemic goods).

6. In epistemic ticking bomb scenarios, epistemic inefficacy puts epistemic goods that have already been attained at risk of deterioration.

In the course of this chapter, I concentrated on the discussion of topics 5 and 6, which expanded the characterization provided before by including a detailed description of why, essentially, we should think that epistemic inefficacy is a vice (as opposed to a misconduct of other sorts); and of why it is an epistemic vice (as opposed to an ethical misconduct). Following an obstructivist framework in vice epistemology, it's an epistemic vice because it systematically obstructs, if not completely, then partially, some important epistemic good.

The chapter also included important considerations that complemented the understanding of why epistemic inefficacy is an interesting concept to epistemologists, or of why it is involved with problems that are distinctively epistemological: it's because this concept is suitable to figure in the best explanation of what is epistemically unsatisfactory about certain people's way of going about in certain inquiries. Specifically, it is suitable to figure in the best explanation of what is unsatisfactory about the way people like Gilbert Levin, the man ahead of the Viking landers' mission, and Alessandro Loiola, the Brazilian medical doctor that said that no

decent regulatory agency would release anti-COVID-19 vaccines before 2024, reasoned.

Those people's conducts were heterodox conducts in their respective domains. This means that explanations that appealed only to structural or external factors, like the zeitgeist of the time, or corporative culture, would be insufficient (they would not account for why those individuals were alone in the midst of their peers). As to explanations that appealed only to sub-personal factors, like cognitive bias, those are not satisfactory either, because we want to preserve the intuition that it was within those people's reach to compromise on their unreasonable demands for certainty, or on their ultracaution, for the sake of the ulterior goal (scientific or sanitary). We want to think of them as examples not to be followed, which means we want to preserve the idea that there is a personal component involved.

A few other considerations might be highlighted here, concerning the characterization of epistemic inefficacy as an epistemic vice. First, that epistemic inefficacy, understood as a personal trait, is more a quality of a particular piece of thinking than a quality of a *type* of agent, even though we often refer to the agent in the cases discussed as "the inefficacious person".

Cassam (2019) distinguishes between three basic varieties of epistemic vice: epistemic vices can be character traits, ways of thinking or attitudes. Character traits are stable dispositions to think, act and feel in particular ways. Arrogance is an epistemic vice that is a character trait: the arrogant person has the stable disposition to think of herself as superior to others, to behave in ways that are aggressively assertive or presumptuous, and to feel superior (Cassam 2019: 12). Ways of thinking, in turn, are styles of reasoning, or inferential or inquisitive habits. They amount to what a person does, rather than what a person is like. Wishful thinking is an example of a thinking vice: it is a flawed style of reasoning, one in which the agent's desires have a greater bearing on the conclusions she draws than evidential considerations. And attitudes are orientations or postures toward something. Contempt is an example of an epistemic vice that is an attitude. To describe someone as contemptuous toward something or toward someone is to qualify her attitude, rather than her habits, or herself. One can be contemptuous toward a particular person without being a contemptuous person in general, or someone with a general disposition to be contemptuous.

Some vices can show up in more than one form. Arrogance, for instance, can show up both as an attitude and a character trait. This means it can show up as a posture toward something in a particular case without entailing the presence of the character trait. It is possible to be arrogant in certain respects or at certain times without being an arrogant person. When one is atypically arrogant, or arrogant in certain respects and humble in others, her arrogance is an attitude, rather than a character trait[83].

Within Cassam's table of distinctions, epistemic inefficacy, like arrogance, is a vice that can show up in more than one form. Specifically, it can present itself as an attitude and as a thinking vice; in both cases it is more a trait of one's conduct that of oneself. Compare driving dangerously with being a dangerous driver: driving dangerously is a (bad) way of driving; it is what a person does, whereas a dangerous driver is what a person is (Cassam 2019: 57-58). Just as one can drive dangerously without being a dangerous driver, a person can reason badly at some particular time or about some particular matter without being a bad epistemic agent herself. This encompasses the understanding that you need not be a special type of person to be subject to failing at your inquiries in an epistemically inefficacious way – it can happen to anyone.

As a thinking vice, epistemic inefficacy makes the agent conduct herself inadequately while dealing with dilemmas. In inefficacious thinking, the concern about certainty has a greater bearing on the drawing of conclusions (or its lack) than evidential considerations, so the person is ultracautious. This is usually a reflex of residual commitment to ideas or theories, that the person is only willing to give up on completely if faced with complete certainty. This leads her to an inadequate pattern of thinking. Bad thinking.

Thinking, Cassam says, is a process of "choosing among potential possibilities", while ways of thinking are "particular ways of including or excluding possibilities, of evaluating evidence, or drawing conclusions" (Cassam 2019: 61). A thinking vice is any particularly bad way of doing these tasks, or of failing at doing them. Epistemic inefficacy, specifically, might be said to be a bad way of excluding

---

[83] To be sure, the whole picture of one's character is a function of her character traits as much as of her ways of thinking and her attitudes, but ways of thinking and attitudes are not stable general dispositions to be certain type of individual. They're predicates of what a person does, rather than of the person herself. Because these three varieties of vice are different they must receive different treatments by a vice epistemologist.

possibilities: the person fails at excluding one of the sides of the dilemma she's involved with, within the correct timing.

As a posture, epistemic inefficacy is what is often termed a "local trait", as opposed to a "robust trait" (Doris 2012): for a person to have an epistemically inefficacious posture it suffices that she adopts, in some particular area of her epistemic life, or some particular domain of inquiry, a pattern of reasoning that involves coping with conflict for longer than she should. The vice need not show up consistently throughout the many areas of her intellectual life. Just like the person in the chocolate cake example, who need not display the same pattern of conduct regarding other flavours of cake or other varieties of treats in order to qualify as weak willed. She might, for instance, have a much easier time quitting drinking than she has refusing extra pieces of chocolate cake, perhaps, for instance, because she is very partial to chocolate cake and not to alcohol. Be that as it may, to display the attitude vice of epistemic inefficacy one need not pervasively display the relevant pattern of reasoning, but only locally.

CONCLUSION

In this dissertation I made the case for the existence of a "new" epistemic vice, the vice of epistemic inefficacy. While baptizing and sketching profiles to vices might sound like a bold philosophical enterprise, the task was made unpresuming by the fact that the vice I attempted to shed light on is actually correlated with a longstanding and much known problem in the history of philosophy: the problem of weakness of willpower. I did my best to make a compelling case for this interrelatedness, by showing that, while not, strictly speaking, the *same* problem, or mirror problems, weakness of willpower and epistemic inefficacy have a lot in common.

In fact, epistemic inefficacy, I submit, is as close as you can get to weakness of willpower in intellectual, or epistemic, contexts, since in those contexts the notion of willpower (ability to control thoughts and actions) is not preponderant. The preponderant notion when it comes to epistemic activity is efficacy (ability to yield the intended results). Efficacy, like willpower, can be weak. One of the ways in which those abilities are made weak is when the agent is overridden by a "mighty force". In the practical cases, the mighty force is the appetites, that is, craving for food, sex, drink, power. In the epistemic cases, it's the person's residual commitment to ideas, theories, hypothesis and hunches. In those cases, this weakness squares off as a vice. I devised a good deal of examples and mobilized a good deal of resources to explain why it squares of as an epistemic vice, in the obstructivist sense. One of those resources was, of course, an extended presentation of what obstructivism is in vice epistemology and inquiry epistemology.

The task of profiling epistemic inefficacy also required me to show that there is a good deal of problems with the thing that is normally thought to be, or meant to be, weakness of willpower's mirror-concept, the so-called "epistemic akrasia". I hope I've been able to clarify that dismissing this concept as self-contradictory is not that simple – prominent scholars thought they were successfully doing it, whereas in reality they were ruling out but one version, or one variety, of the thing. The reasons why this concept should be dismissed by epistemologists is not that it is self-contradictory (it is not), but rather that it doesn't capture important elements of the very idea of weakness that underlie cases of practical akrasia, such as the presence of a disposition towards iteration and the production of bad consequences. As a

result, it turns out to not be a very useful concept for the goal of Verstehen, that is, charitable sense making of other people's reasoning.

Now, many interrelated discussions have been left out of the scope of this dissertation that are, nevertheless, important for the biggest picture of how this vice operates. One of them concerns the problem of when it is reasonable and when it is not reasonable to abandon one's commitments. My discussion of epistemic inefficacy assumes that when a person is dealing with a dilemma, or inquiring towards the solution to a quandary, as the search progresses, she will naturally reach what appears to be a "no turning back point" – a point where evidence supporting, say, hypothesis B appears to overtake the evidence supporting hypothesis A in a way that this balance either cannot be reversed, or it's very unlikely that it can be reversed. When one reaches this point, I suggested, he should abandon his commitment to hypothesis A.

For instance, in the Viking landers' case, presented in Chapter 6, the "no turning back point" is the discovery that an inorganic substance, hypochlorite, is capable of delivering the exact same results in the LR experiment that they thought only living microbes could. From the moment this discovery is made, the dilemma of inorganic substances *vs.* living microbes reaches a "no turning back point": the inorganic substances hypothesis is "ahead" in the contest (the contest aimed at determining who is the best candidate to explain the original LR results), in a way that it is going to be very difficult for the living microbes hypothesis to overtake. That's because this inorganic substance, hypochlorite, fulfils all the requirements for a successful explanation *and* it has parsimony on its side. It has the advantage of offering a simpler account than the rival hypothesis. This is the point, I suggest, when the agent should abandon her residual commitment to hypothesis A; in this case, the living microbes hypothesis, and consider the matter settled.

So my account of what a responsible and effective inquiry looks like presupposes that there must be this "not turning back point", but it is silent on the specific criteria that will allow us to recognize, or to set, this mark. In the Viking landers' case it is difficult to dispute that the advent of hypochlorite is the mark, but it is not always clear where this "no turning back point" lies. In fact, in many cases this is the matter of dispute in itself. Cases involving peer disagreement and misleading high-order evidence offer the most paradigmatic examples that illustrate this difficulty.

This is a debate in its own right. Some scholars have argued that upon getting evidence suggesting that his or her evidence does not support the conclusion it appears to support (that is, upon assessing misleading high order evidence, or a disagreeing epistemic peer), the agent should stick to her previously formed commitment. Those include Kelly (2005) and Schoenfield (2014). Others have argued that in those situations the agent should abandon such a commitment. Those include Feldman (2005), Elga (2007) and Christensen (2014). Others have argued that it varies from case to case and that there is no rule of thumb; for instance, Kelly (2010). And there are also the ones that argue that the agent should decrease confidence in his previously formed commitment but not abandon it completely. That is, they propose that residual commitment be maintained, that dilemma be kept unresolved and that the agent remains in an incoherent state of mind, admitting of both rival hypotheses A and B. Very different lines of reasoning motivate this verdict, and authors who support it include Williamson (2011, 2014), Wedgwood (2011) and Coates (2012). Those scholars seemingly believe that by undertaking this conduct the agent is solving the problem, whilst I believe that by doing so he is putting up with it. Therefore, addressing this debate remains as a possible avenue for future inquiry.

Another issue that didn't receive a lot of attention in the account of epistemic inefficacy devised in this dissertation is the issue of preemption *vs.* caution. The problem posed by epistemic inefficacy, as we've been seeing, is a problem of misplaced excessive caution. We are always weighting benefits against risks and we are usually trying to take as few risks as possible. When we do not have a lot of time, however, caution *itself* can turn into a risky option, as the race for vaccines example suggests. Misplaced ultracaution puts a number of epistemic goods under threat of deterioration or even destruction. Now, this raises some important questions. Does the case for epistemic inefficacy as a vice supports the case that, when in doubt, it's better to take risks than to err on the side of caution? When time is of the essence, is it a good policy to lower our standards of caution? This is the problem of preemption *vs.* caution.

One might explain what preemptive action is by contrasting it with preventive action, say, for instance, in the context of warfare, where those terms are most commonly used.

> A preventive war is undertaken when a state sees its relative advantage in decline, sees the inevitability of war, and chooses to initiate the war now while it still has the advantage. History and international law frown upon preventive war, seeing it as a disguise for naked aggression. Preemptive war, on the other hand, involves the initiation of military action because an adversary's attack is believed to be imminent. A preemptive strike is directed against an adversary's capability before it can be used. It is not conducted for purposes of initiating war. (Worley 2003: vii)

Effective preventive action is action that has the power to avert future damage. Effective preemptive action is action that has the power to avert imminent damage, that is, damage that is just about to ensue, or at least that is believed to be just about to ensue. It is the type of action you undertake as an emergency measure – unorthodox, something that you wouldn't normally consider doing, but that needs doing because the consequences of not doing it would be felt immediately, and would possibly be worse than the consequences of having done it. An example of preemptive action in warfare is the United States entering Iraq in 2003, under the Bush administration. G. W. Bush declared, referring to the alleged threat that Iraq could possess and mobilize weapons of mass destruction against the US: "If we wait for those threats to fully materialize, we will have waited too long" (Bush 2002).

A ticking bomb scenario is, if anything is, a scenario in which preemptive action appears to be justified, to the detriment of cautious action. Now, can the same be said about epistemic ticking bomb scenarios? Because the account of epistemic inefficacy I advanced suggests that in epistemic ticking bomb scenarios caution is unadvisable. So does the case for epistemic inefficacy as an epistemic vice support the case for an epistemic version of the Bush doctrine? This is an avenue for future development too. A debate addressing a related issue, the issue of whether the criticisms frequently presented against the Bush doctrine can be generalized to cover all cases of preemptive action, including in epistemic contexts, has been started by Cassam (2020).

Last, the account devised throughout this dissertation is basically silent on the matter of epistemic vice-charging, the critical practice of charging other persons with epistemic vice. I didn't say much about when or under which circumstances charging people with the vice of epistemic inefficacy would be justified, as opposed to the cases when such a charge would be questionable.

In recent years, the desiderata for robust vice-charge have been discussed by scholars (that is, what a vice-charge must and must not comprehend), as well as the main obstacles to the general practice of epistemic vice-charging. As Ian Kidd (2016) discusses, the efficacy of a vice-charge is contingent upon a degree of consensus between critic and target that is difficult to attain precisely in the situations where vice-charging is most likely to be provoked. Those are the situations in which the target of the vice-charge disagrees with you about fundamental issues, such as the value of a certain epistemic good, for instance. So a robust critical practice of vice-charging is possible in principle, but very difficult in practice. Other scholars that have engaged in this discussion are Tanesini (2018) and Cassam (2021). In this way, a comprehensive discussion of the motivation and obstacles of charging people with the vice of epistemic inefficacy is another topic for future attention.

REFERENCES

Adler, J. (2002a). Akratic Believing? *In: Philosophical Studies*, 110: 1-27.

Adler, J. (2002b). *Belief's Own Ethics*. Cambridge: The MIT Press.

Afonso, A. (2021). The mutation risk. *Pesquisa Fapesp.* Issue 299, jan. 2021.

Alfano, M. (2012). Expanding the Situationist Challenge to Responsibilist Virtue Epistemology. *The Philosophical Quarterly*, 62: 223-49.

Alfano, M. (2015). Becoming less unreasonable: A reply to Sherman. *Social Epistemology Review and Reply Collective*, 4(7): 59-62.

Allen, R. (1993). *The Dialogues of Plato*, Volume 2: The Symposium. Yale: Yale University Press.

Altemeyer, B.; Hunsberger, B. (2004). A revised Religious Fundamentalism scale: The short and sweet of it. *International Journal for the Psychology of Religion*, 14(1): 47–54.

Alston, W. (1989). The Deontological Conception of Epistemic Justification. In: *Essays in the Theory of Knowledge*. Ithaca: Cornell UP: 115-52.

Aristotle. (2000). *Nicomachean Ethics*. Trans. Roger Crisp. Cambridge: Cambridge University Press.

Audi, R. (1994). Dispositional Beliefs and Dispositions to Believe. Nous 28: 419-434.

Austin, J. (1979). A Plea for Excuses. In: *Austin J. L.: Philosophical Papers*. J. Urmson and G. Warnock (eds.), Oxford: Oxford University Press: 175–204.

Axtell, G. (2007). Two for the Show: Anti-luck and Virtue Epistemologies in Consonance. *Synthese* 158: 363-83.

Axtell, G. (2008). Expanding Epistemology: A Responsibilist Approach. *Philosophical Papers* 37: 51-87.

Axtell, G. (2010). Epistemic Virtue. In: Jonathan Dancy and Ernest Sosa (eds.), *A Companion to Epistemology*. Malden: Blackwell: 116-17.

Baehr, J. (2010). Epistemic Malevolence. In: H. Battaly (ed.) *Virtue and Vice, Moral and Epistemic*. Chichester: Wiley-Blackwell: 189-213.

Baehr, J. (2012a). *The Inquiring Mind: On Intellectual Virtues and Virtue Epistemology*. Oxford, Oxford University Press.

Baehr, J. (2012b). Two Types of Wisdom. *Acta Analytica* 27: 81-97

Battaly, H. (1998). What is Virtue Epistemology? In: Twentieth World Congress of Philosophy. The Paideia Archive. Available at: https://www.bu.edu/wcp/Papers/Valu/ValuBatt.htm

Battaly, H. (2014). Varieties of Epistemic Vice. In: Jonathan Matheson and Rico Vitz (eds.), *The Ethics of Belief*. Oxford University Press: 51-76.

Battaly, H. (2016). Epistemic virtue and vice: Reliabilism, responsibilism, and personalism. In: C. Mi, M. Slote, & E. Sosa (Eds.), *Moral and intellectual virtues in Western and Chinese philosophy*. New York: Routledge: 99-120.

Battaly, H. (2017). Testimonial injustice, epistemic vice, and vice epistemology. In: I. J. Kidd, G. Polhaus, & J. Medina (Eds.), *The Routledge handbook of epistemic injustice*. New York: Routledge: 223-231.

Battaly, H. (2018). Close-mindedness and dogmatism. *Episteme*, 15 (3):261-282.

Battaly, H. (2019). A third kind of intellectual virtue: Personalism. In: H. Battaly (Ed.), *The Routledge handbook of virtue epistemology*. New York: Routledge: 115-126.

Bayne, T.; E. Pacherie. (2005). In Defense of the Doxastic Conception of Delusions. *Mind and Language* 20 (2): 163–188.

Berenbaum, H.; Bredemeier, K.; Thompson, R. (2008). Intolerance of uncertainty: Exploring its dimensionality and associations with need for cognitive closure, psychopathology, and personality. *Journal of Anxiety Disorders* 22(1), 117–125.

Bowden, M. (2002). Torture, if it saves lives, may be a necessary evil. In: Philadelphia Inquirer. 7 April 2002. Available at: https://freerepublic.com/focus/news/661373/posts.

Brandom, R. (1998). Insights and blindspots of reliabilism. *The Monist* 81: 371-392.

Bratman, M. (1992). Practical Reason and Acceptance in a Context. In: *Faces of Intention*. Cambridge: Cambridge University Press: 15-34.

Bratman, M. (1993). Practical Reasoning and Acceptance in a Context. Mind 102: 1-15.

Bratman, M. (1999). *Faces of intention.* Cambridge: Cambridge University Press.

Brooks, M. (2008). *Thirteen things that don't make sense: the most baffling scientific mysteries of our time.* New York: Doubleday.

Broome, J. (2013). Ration*ality Through Reasoning*. Oxford: John Wiley & Sons.

de Bruin, B. (2013). Epistemic Virtues in Business. *Journal of Business Ethics*, 113(4): 583–595.

Buckareff, A. (2004). Acceptance and Deciding to Believe. *Journal of Philosophical Research,* 29: 173-90.

Bufacchi, V.; Arrigo, J. (2006). *Torture, Terrorism and the State: a Refutation of the Ticking-Bomb Argument*. Journal of Applied Philosophy, 23(3): 355-373.

Bush, G. (2002). President Bush Delivers Graduation Speech at West Point. The White House. Available at https://georgewbush-whitehouse.archives.gov/news/releases/ 2002/06/print/20020601-3.html

Campbell, T.; Kay, A. (2014). Solution aversion: On the relation between ideology and motivated disbelief. *Journal of Personality and Social Psychology, 107*(5), 809–824.

Carter, J., Pritchard, D. (2015). Epistemic Situationism, Epistemic Dependence and the Epistemology of Education. In: M. Alfano & A. Fairweather (eds) *Epistemic Situationism.* Oxford: Oxford University Press: 168-191.

Cassam, Q. (2016). Vice epistemology. *The Monist*, 99: 159-180.

Cassam, Q. (2018). Epistemic insouciance. *Journal of Philosophical Research*, 43: 1-20.

Cassam, Q. (2019). *Vices of the mind*. Oxford: Oxford University Press.

Cassam, Q. (2020). Terrorism and Pandemics: Re-Assessing the Bush Doctrine. Available at https://www.quassimcassam.com/post/terrorism-and-pandemics-re-assessing-the-bush-doctrine

Cassam, Q. (2021). Misunderstanding vaccine hesitancy: a case study in epistemic injustice. Educational Philosophy and Theory: 1-16.

Chalmers, D. (2011). Verbal Disputes, 120(4), 515–566.

Christensen, D. (2010). Rational Reflection. *Philosophical Perspectives*, 24: 121–140.

Coates, A. (2012) Rational Epistemic Akrasia. *American Philosophical Quarterly*, 49 (2): 113– 124.

Code, L. (1984). Toward a Responsibilist Epistemology. *Philosophy and Phenomenological Research*, 45: 29–50.

Cohen, J. (1989). Belief and Acceptance. *Mind*, 98: 367-89.

Cohen, J. (1992). *An Essay on Belief and Acceptance.* Oxford: Oxford University Press.

Coates, A. (2012). Rational Epistemic Akrasia. *American Philosophical Quarterly*, 49: 113-124.

Collins, J. (2000). *Preemptive Prevention*. The Journal of Philosophy, 97(4), 223–234.

Couper-Kuhlen, E. (2016). A Sequential Approach to Affect: the case for "disappointment". In: Haakana M., Laakso M. & Lindström J. (eds) *Talk in Interaction: Comparative Dimensions.* Helsinki: Studia FennicaLinguistica, pp. 94-123.

Crerar, C. (2017). Motivational Approaches do Intellectual Vice. *Australian Journal of Philosophy*: 1-14.

Daoust, M. (2018). Epistemic Akrasia and Epistemic Reasons. Episteme 16 (3):282-302.

Davidson, D. (1963). Actions, Reasons and Causes. *The Journal of Philosophy* Vol. 60, No. 23, American Philosophical Association, Eastern Division, Sixtieth Annual Meeting (Nov. 7, 1963), pp. 685-700.

Davidson, D. (1969/1980). How is Weakness of the Will Possible? In: Joel Feinberg (ed.), *Moral Concepts*. Oxford University Press.

Davidson, D. (1970). Events and Particulars. *In: Noûs*, 4: 25-32.

Davidson, D. (1986). A Coherence Theory of Truth and Knowledge. In: E. LePore (ed.), *Truth and Interpretation.* Oxford: Blackwell's: 307–319.

Davidson, D. (1999). The Emergence of Thought. *Animal Mind*, 51: 7–17

Davies, M. (2000). Interaction without reduction: The relationship between personal and sub-personal levels of description. *Mind & Society* 1, 87–105.

Davies, M.; T. Stone (1992). *Pathologies of Belief*. Oxford: Blackwell.

Dennett, D. (1969) *Content and Consciousness*. London: Routledge and Kegan Paul.

DeRoma, V.; Martin, K.; Kessler, M. (2003). The Relationship Between Tolerance for Ambiguity and Need for Course Structure. *Journal of Instructional Psychology*, 30: 104-09.

Dershowitz, A. (2004). *Tortured Reasoning*. In: Sanford Levinson (ed.), "Torture: A Collection". Oxford University Press.

Dershowitz, A. (2016). *Is it Necessary to Apply Physical Pressure to Terrorists and to Lie about it?* Cambridge University Press.

Dinges, A. (2022). Degrees of Acceptance. The Philosophical Quarterly, (72) 3: 578–594.

Ditto, P.; Lopez, D. (1992). Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions. *Journal of Personality and Social Psychology,* 63(4), 568–584.

Ditto, P.; Munro, G., Apanovitch, A.; Scepansky, J.; Lockhart, L. (2003). Spontaneous Skepticism: The Interplay of Motivation and Expectation in Responses to Favorable and Unfavorable Medical Diagnoses. *Personality and Social Psychology Bulletin,* 29(9), 1120–1132.

Doris, J. (1998). Persons, Situations, and Virtue Ethics. *Nous* 32: 504-30.

Doris, J. (2002). *Lack of Character*. Cambridge: Cambridge University Press.

Drayson, Z. (2014). The Personal/Subpersonal Distinction. *Philosophy Compass,* 9(5), 338–346.

Dretske, F. (1971). Reasons, Knowledge and Probability. *Philosophy of Science*. 38: 216-20.

de Dreu, C.; Koole, S.; Oldersma, F. (1999). On the seizing and freezing of negotiator inferences: need for cognitive clousure moderates the use of heuristics in negotiation. *Personality and Social Psychology Bulletin*, 25(3): 348-362.

Dreyfus, H. (2013). The Myth of the Pervasiveness of the Mental. In: by J. Schear (ed.) *Mind, Reason and Being-in-the-World: The McDowell-Dreyfus Debate*. London: Routledge.

Duckworth, A. (2011). The significance of self-control. Proceedings of the National Academy of Sciences, 108: 2639–2640.

Duncker, K. (1945). On Problem Solving (trans. L. S. Lees). *Psychological Monographs* 58, Washington, DC: American Psychological Association.

Engel, P. (1998). Believing, accepting, and holding true. *Philosophical Explorations,* I(2): 140–151.

Engel, P. (1999). Dispositional belief, assent, and acceptance. *Dialectica* 53(3–4), 211–226.

Engel, P. (2000). Introduction: The Varieties of Belief and Acceptance. In: Engel, P. (eds) *Believing and Accepting*. Philosophical Studies Series, vol 83.

Engel, P. (2012). Trust and the Doxastic Family. *Philos Stud* (2012) 161:17–26.

Engel, P. (2018). The Doxastic Zoo. In: Annalisa Coliva, Paolo Leonardi & Sebastiano Moruzzi (eds.), *Eva Picardi on Language, Analysis and History*. Palgrave: 297-316.

Flanagan, O. (1991). *Varieties of Moral Personality: Ethics and Psychological Realism.* Cambridge, MA: MIT Press.

Forster, M. (2013). Wittgenstein on family resemblance concepts. In: Arif Ahmed (ed.) *Wittgenstein's Philosophical Investigations: A Critical Guide*. Cambridge University Press: 66-87.

Fox, E. J.; Hoch, S. J. (2005). Cherry-Picking. Journal of Marketing, 69(1), 46–62.

van Fraassen, B. (1980). *The scientific image.* Oxford: Oxford University Press.

Friedman, J. (2012). Motivated Skepticism or Inevitable Conviction? Dogmatism and the Study of Politics. *Critical Review: a Journal of Politics and Society*, 24(2), 131–155.

Friedman, J. (2013). Question-directed attitudes. *Philosophical Perspectives*, 27: 145–174.

Friedman, F. (2019). Inquiry and Belief. *Nous* 53: 296-315.

Fricker, M. (2007). *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford: Oxford University Press.

Frost, R.; Shows, D. (1993). The nature and measurement of compulsive indecisiveness. *Behaviour Research and Therapy*, 31(7), 683-692.

Gendler, T. (2008). Alief and Belief. *Journal of Philosophy.* 105 (10): 634–663.

George, M.; Teunisse, A.; Case, T. (2020). Gotcha! Behavioural validation of the Gullibility Scale. *Personality and Individual Differences*, n. 162.

Gibbons, J. (2013). *The Norm of Belief*. Oxford: Oxford University Press.

Ginet, C. (2001). Deciding to Believe. In: M. Steup (ed.), *Knowledge, Truth, and Duty*, pp. 63–76. Oxford: Oxford University Press.

Goldman, A. (2001). The unity of the epistemic virtues. In A. Fairweather, & L. Zagzebski (Eds.), *Virtue epistemology: Essays on epistemic virtue and responsibility*. Oxford: Oxford University Press.

Hare, R. (1952). *The Language of Morals*. Oxford Clarendon Press.

Hare, R. (1963). *Backsliding. In Freedom and reason*. Oxford: Oxford University Press

Harman, G. (1980). Reasoning and Evidence One Does Not Possess. *Midwest Studies in Philosophy*, 5: 165-182.

Hartshorne H.; May, M. (1928). Studies in the nature of character. Vol. 1. *Studies in deceit*. New York: Macmillan.

Harman, G. (1999). Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error. *Proceedings of the Aristotelian Society* 99: 315-31.

Harman, G. (2000). The Nonexistence of Character Traits. *Proceedings of the Aristotelian Society* 100: 223-6.

Heil, J. (1984). Doxastic Incontinence. *Mind*, 93: 56–70.

Hecht, M.; Kounaves, S.; Quinn, R.; West, S.; Young, M.; Ming, D.; Catling, D.; Clark, B.; Boynton, W.; Hoffman, J.; DeFlores, L.; Gospodinova, K.; Kapit, J.; Smith, P. (2009) Detection of Perchlorate and the Soluble Chemistry of Martian Soil at the Phoenix Lander Site. Science 325: 64-67.

Hieronymi, P. (2006). Controlling Attitudes. *Pacific Philosophical Quarterly,* 87: 45-74.

Higginbotham, A. (2019). *Midnight in Chernobyl: the untold story of the world's greatest nuclear disaster*. Simon & Schuster.

Hintikka, J. (2007). *Socratic Epistemology: Explorations of Knowledge-Seeking by Questioning*. Cambridge: Cambridge University Press.

Hoffmann, T. (ed.) (2008).*Weakness of Will from Plato to the Present*, Washington: Catholic University of America Press.

Holton, R. (1999). Intention and Weakness of Will. *Journal of Philosophy*, 96: 241–262.

Hookway, C. (1994). Cognitive Virtues and Epistemic Evaluations*. International Journal of Philosophical Studies* 2: 211-27.

Hookway, C. (2000) Skepticism and the Principle of Inferential Justification. *Nous*, 34(s1), 344-365.

Hookway, C. (2001). Epistemic Akrasia and Epistemic Virtue*.* In: A. Fairweather &L. Zagzebski (ed.), *Virtue Epistemology: Essays on Epistemic Virtue and Responsibility*. Oxford: Oxford University Press: 178-199.

Hookway, C. (2003). How to be a Virtue Epistemologist. In: M. DePaul and L. Zagzebski (eds.) *Intellectual Virtue: Perspectives from Ethics and Epistemology*. Oxford: Oxford University Press: 183-202.

Hookway, C. (2008). Questions, Epistemology, and Inquiries. *Grazer Philosophische Studien: internationale Zeitschrift fuer analytische Philosophie*, 77, 1-21.

Horowitz, S. (2014). Epistemic Akrasia. Noûs, 48(4), 718–744.

Hurley, S. (1989). *Natural Reasons: Personality and Polity*. New York: Oxford University Press.

Jackson, R. (2015). The epistemological crisis of counterterrorism. *Critical Studies on Terrorism*, 8: 33-54.

Jeffrey, R. (1983). *The logic of decision*, 2nd edition, Chicago: University of Chicago Press.

Jervis, R. (2003). *Understanding the Bush Doctrine.* Political Science Quarterly, 118(3), 365–388.

Johnston, M. (1995). Self-Deception and the Nature of Mind. In: C. MacDonald & G. MacDonald (ed.), *Philosophy of Psychology*. Cambridge, Mass.: Blackwell.

Kalis, A.; Mojzisch, A.; Schwizer, T.; Kaiser, S. (2008) Weakness of will, akrasia, and the neuropsychiatry of decision making: An interdisciplinary perspective. *Cognitive, Affective, & Behavioral Neuroscience*, 8 (4), pp. 402-417.

Kampa, S. (2019). Obsessive-Compulsive Akrasia. *Mind & Language*, 1:1–18.

Kelp, C. (2021). *Inquiry, Knowledge and Understanding*. Oxford: Oxford University Press.

Kenny, A. (1979). *Aristotle's Theory of the Will.* New Haven, 69.

Kidd, I. (2016). Charging Others With Epistemic Vice. *The Monist* 99 (3):181-197.

Kidd, I. (2017). Capital Epistemic Vices. *Social Epistemology Review and Reply Collective* 6, no. 8: 11-16.

Klein, P. (1981). *Certainty: A Refutation of Skepticism.* Minneapolis: University of Minnesota Press.

Klotz, I. (2009). Organic molecules found ... on the moon? NBC News, Dec. 16th 2009. Available at https://www.nbcnews.com/id/wbna34449441.

Kraemer, E. (2015). Virtue Epistemologies and Epistemic Vice. *Athens Journal of Humanities & Arts*, v.2, i.1: 19-30.

Korber, B. et. al. (2020). Tracking changes in Sars-CoV-2 Spike: evidence that D614G increases Infectivity of the Covid-19 virus. *Cell.* Vol. 182, pp. 794–95. Aug. 20, 2020.

Koriat, A. (2005). Knowing and the Feeling of Knowing. In Synergies, ed. W. Østreng, 89–92. Oslo: Centre for Advanced Study.

Kubler-Ross E. (1969). *On death and dying*. New York: Scribner.

Kruglanski, A.;Webster, D. (1996). Motivated closing of the mind: "Seizing" and "freezing". *Psychological Review,* 103(2): 263-283.

Kruglanski, A. (2004). *The Psychology of Closed Mindedness.* New York: Psychology press.

Kvanvig, J. (1992). *The Intellectual Virtues and the Life of the Mind.* Lanham, MD: Rowman and Littlefield.

Kwong, J. (2016). Open-Mindedness as Engagement. *The Southern Journal of Philosophy, 54(1), 70–86.*

Lasonen-Aarnio, M. (2014). Higher-order evidence and the limits of defeat. *Philosophy and Phenomenological Research*, 88.2: 314-345.

Lawrence, K.; Keleher, T. (2004). Chronic Disparity: Strong and Pervasive Evidence of Racial Inequalities. *Race and Public Policy Conference*. Available at < https://www.intergroupresources.com/rc/Definitions%20of%20Racism.pdf > Access in Feb. 18, 2021.

Leitgeb, H. & Carus, A. (2022). Rudolf Carnap. In: The Stanford Encyclopedia of Philosophy (Fall 2022 Edition), Edward N. Zalta (ed.), forthcoming URL = <https://plato.stanford.edu/archives/fall2022/entries/carnap/>.

Levin, G.; Straat, P. (2016). The Case for Extant Life on Mars and Its Possible Detection by the Viking Labeled Release Experiment. *Astrobiology,* 16(10): 1-13.

Liu, W. (2016). Aristotle on *prohairesis. Labirinth*, v. 18, n. 2, winter.

Lockie, R. (2006). Problems for virtue theories in epistemology. Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition, 138(2): 169-191.

Lycan, W. (1985). Tacit Belief. In: R. Bogdan (ed.) *Belief*. Oxford: Oxford University Press: 61–82.

Martin, K. (2014). *Justified Closure of Inquiry: a non-reductive account*. PhD Dissertation, Baylor University.

Mazzotta, G. (2008). Dante: Healing the Wounded Will. In: Hoffmann, T. (ed.), *Weakness of Will from Plato to the Present*, Washington: Catholic University of America Press: 138-158.

McCrae, R. (2004). Conscientiousness. *Encyclopedia of Applied Psychology,* v.1: 469-472.

McDowell, J. (1994). *Mind and World.* Cambridge, MA: Harvard University Press.

Mele, A. (1986). Incontinent Believing. *The Philosophical Quarterly*, 36: 212-222.

Mele, A. (1987). *Irrationality*. New York: Oxford University Press.

Mele, A. (1992). Acting for Reasons and Acting Intentionaly. *Pacific Philosophical Quarterly,* 73: 355-374.

Mercier, H. (2017). How gullible are we? A review of the evidence from psychology and social science. *Review of General Psychology*: 1-37.

Meyer, W.; Niepel, M.; Rudolph, U.; Schützwohl, A. (1991). An experimental analysis of surprise. Cognition & Emotion, 5(4), 295–311.

Mischel, W. (1968). *Personality and Assessment*. New York: Wiley.

Mischel, W. (1971). *Introduction to Personality*. New York: Holt, Rinehart & Winston.

Montmarquet, J. (1987). Epistemic Virtue. *Mind*, 96: 482–97.

Napier, S. (2009). *Virtue Epistemology: Motivation and Knowledge*. New York: Continuum.

Navarro-González, R.; Rainey, F.; Molina, P.; Bagaley, D.; Hollen, B.; De laRosa, J.; Small, A.; Quinn, R.; Grunthaner, F.; Cáceres, L.; Gomez-Silva, B. (2003). Mars-Like Soils in the Atacama Desert, Chile, and the Dry Limit of Microbial Life. Science, 302(5647), 1018–1021.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology,* 2(2): 175-220.

Nussbaum, M. (1988). Non Relative Virtues: an Aristotelian Approach. *Midwest Studies in Philosophy,* 13(1), 32-53.

Oreskes, N.; Conway, E. (2010). *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming*. Bloomsbury Press.

Ovenden, R. (2020). *Burning the Books: A History of the Deliberate Destruction of Knowledge*. Belknap Press.

Owens, D. (2000). *Reason without Freedom*. London: Routledge, 2000.

Owens, D. (2002). Epistemic Akrasia. *The Monist,* 85: 381-97.

Owens, D. (2003). Does Belief Have An Aim? *Philosophical Studies*, 115: 275-97.

Parry, J. (2004). *Escalation and Necessity: Defining Torture at Home and Abroad*. In: S. Levinson (ed.) Torture: A Collection. Oxford: Oxford University Press.

Plante, J.; Liu, Y.; Liu, J. et al. (2021). Spike mutation D614G alters SARS-CoV-2 fitness. *Nature* 592: 116–121.

Perry, R.; Sibley, C. (2013). Seize and freeze: openness to experience shapes judgments of societal threat. *Journal of Research in Personality*, 47(6): 677-686.

Pettit, P; Smith, M. (1996). Freedom in Belief and Desire. *Journal of Philosophy*, 93: 429-49.

Pritchard, D. (2016). *Epistemic Angst: Radical Skepticism and the Groundlessness of Our Believing.* Princeton: Princeton University Press.

Quinn, R.; Martucci, H.; Miller, S.; Bryson, C.; Grunthaner, F.; Grunthaner, P. (2013). Perchlorate Radiolysis on Mars and the Origin of Martian Soil Reactivity. *Astrobiology,* 13(6), 515–520*.*

Rassin, E.; Muris, P. (2005). Indecisiveness and the interpretation of ambiguous situations*. Personality and Individual Differences, 39(7), 1285–1291.*

Raz, J. (2007). Reasons: Explanatory and Normative. *Oxford Legal Studies*, 1: 1-25.

Ribeiro, B. (2011). Epistemic Akrasia*. International Journal for the Study of Skepticism,* 1: 18-25.

Riggs, W. (2003).Understanding 'Virtue' and the Virtue of Understanding. In: M. DePaul & L. Zagzebski *Intellectual Virtue: perspectives from Ethics and Epistemology*. Oxford: Clarendom Press: 203-226.

Riggs, W. (2010) Open-Mindedness. *Metaphilosophy* 41 (1-2):172-188.

Roberts, R. (2000). The Journal of American History. Oxford, Vol. 87, n. 3.

Roberts, C.; West, R. (2015). Natural epistemic defects and corrective virtues. *Synthese*, 192(8): 2557-76.

Roberts, R.; Wood, J. (2003). Humility and Epistemic Goods. In: M. DePaul & L. Zagzebski *Intellectual Virtue: perspectives from Ethics and Epistemology*. Oxford: Clarendom Press: 257-279.

Roberts, R.; Wood, J. (2007). Intellectual Virtues: An Essay in Regulative Epistemology, Oxford: Oxford University Press.

Roets, A.; Van Hiel, A. (2011a). Item selection and validation of a brief, 15-item version of the need for closure scale. *Personality and Individual Differences*, 50: 90-94.

Roets, A.; Van Hiel, A. (2011b). The role of need for closure in essentialist entitativity beliefs and prejudice: An epistemic needs approach to racial categorization. *British Journal of Social Psychology*, 50: 52-73.

Rorty, A. (1981). *Essays on Aristotle's Ethics.* University of California Press.

Rorty, A. (1983). Akratic Believers. *In: American Philosophical Quarterly*, 20: 175-83.

Rorty, A. (1985). Self-deception, akrasia and irrationality. In J. Elster (Ed.), *The multiple self* . Cambridge: Cambridge University Press: 115-131.

Rorty, A. (1988). Where does the akratic break take place? In *Mind in action: Essays in the philosophy of mind.* Boston: Beacon Press: 229-270.

Rorty, A. (1997). The Social and Political Sources of Akrasia. *Ethics*. 107(4), 644-657.

Roseano, P.; González, M.; Borràs-Comes, J.; Prieto P. (2015). Communicating Epistemic Stance: How Speech and Gesture Patterns Reflect Epistemicity and Evidentiality. *Discourse Processes*, 00:1–40.

Ross, L.; Nisbett, R. (1991). *The Person and the Situation.* Philadelphia: Temple University Press.

Runco, M. (2014). Personality and Motivation. *Creativity*: 265-302.

Sabini, J.; Silver, M. (2005). Lack of Character? Situationism Critiqued. *Ethics*, Vol. 115, No. 3: 535-562.

Salzman, L. (1980). *Treatment of the Obsessive Personality*. New York: Jason Aronson.

Samuelson, P.; Church, I. (2015). When cognition turns vicious: Heuristics and biases in light of virtue epistemology. *Philosophical Psychology*, 28(8): 1095-1113.

Schwitzgebel, E. (2001). In between believing. *Philosophical Quarterly*, 51, 76–82.

Schwarz, N. (1991). Feelings as information: Informational and motivational functions of affective states. In E. T. Higgins & R. M. Sorrentino (Eds.), *The handbook of motivation and cognition: Foundations of social behavior* (Vol. 2, pp. 527-561). New York: Guilford Press.

Sephton, M. *(2004).* Organic matter in ancient meteorites. Astronomy & Geophysics, 45(2): 8-14.

Shah, N. (2003). How Truth Governs Belief. *Philosophical Review*, 112: 447-482.

Shah, N.; Velleman, D. (2005). Doxastic deliberation. *Philosophical Review*, 114: 497-534.

Shakespeare, W. (1987). *The Merchant of Venice.* M.M. Mahood (ed). Cambridge: Cambridge University Press.

Shakespeare, W. (2013). *Macbeth*. Mowar, B.; and Werstine P. (eds). Folger Shakspeare Library: Simon & Schuster.

Shanahan, C. (2019). Alcibiades' Akrasia: Reason for Wrongdoing? *The International Journal of the Platonic Tradition*, 13(2), 131–152.

Shiffrin, R.; Schneider, W. (1977).Controlled and automatic human information processing: Perceptual learning, automatic attending, and a general theory. *Psychological Review*, 84: 127–190.

Shugar, S. (2006). *Knowing is not enough: Akrasia and self-deception in Shakespeare's Macbeth.* [Montreal]: McGill University. PhD dissertation.

Skyrms, B. (2000). *Choice and chance*, 4th edition, Belmont, CA: Wadsworth/Thompson.

Sinnott-Armstrong, W. (1996). Moral Skepticism and Justification. In: Sinnott-Armstrong, W. & Timmons, M. (eds.). *Moral Knowledge? New Readings in Moral Epistemology*. Oxford: Oxford University Press.

Smith, J. & Winkler, R. (2006). The Optimizer's Curse: Skepticism and Postdecision Surprise in Decision Analysis. Management Science, 52(3), 311–322.

Snellen, P. (2018). *Akrasia as a character trait*. [Groningen]: University of Groningen. PhD Dissertation.

Sosa, E. (1980). The Raft and the Pyramid: Coherence versus Foundations in the Theory of Knowledge. *Midwest Studies in Philosophy*, 5: 3-25.

Sosa, E. (2003). The place of truth in epistemology. In: DePaul, M., Zagzebski, L. (eds.). *Intellectual virtue: Perspectives from ethics and epistemology*. New York: Oxford University Press.

de Sousa, R. (1987). *The Rationality of Emotion*. MIT Press.

Steup, M. (2000). Doxastic Voluntarism and Epistemic Deontology. *Acta Analytica*, 15: 25-56.

Stich, S. (1978). Belief and Subdoxastic States. *Philosophy of Science* 45: 499–518.

Stowers, A. (1994). *Rereading of Romans: Justice, Jews, and Gentiles.* New Haven: Yale University Press.

Swank, C. (2000). Epistemic Vice. In: G. Axtell (ed.) *Knowledge, Belief, and Character: Readings in Virtue Epistemology. Lanham*: Rowman and Littlefield Publishers:195-204.

Tanesini, A. (2018). Epistemic Vice and Motivation. *Metaphilosophy*, v. 49, n. 3.

Teunisse, A.; Case, T.; Fitness, J.; Sweller, N. (2020). I should have known better: Development of a self-report measure of gullibility. *Personality and Social Psychology Bulletin*, 46(3): 408-423.

Tuomela, R. (2000) Belief versus acceptance. Philosophical Explorations: An International Journal for the Philosophy of Mind and Action, 3:2, 122-137.

Turri, J., Alfano, M., Greco, J. (2021) Virtue Epistemology. In: *The Stanford Encyclopedia of Philosophy.* (Winter 2021 Edition), Edward N. Zalta (ed.), available at <https://plato.stanford.edu/archives/win2021/entries/epistemology-virtue/>.

Tversky, A.; Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-113.

Vahid, H. (2006). Aiming at Truth: Doxastic vs. Epistemic Goals. *Philosophical Studies,* 131: 303-335.

Vasiliauskas, E. (2016). Mortal Knowledge: Akrasia in English Renaissance Tragedy. *Politics and Aesthetics in European Baroque and Classicist Tragedy*. Leiden: Brill: 221-38.

Velleman, D. (1989). *Practical Reflection.* Cambridge: Harvard University Press.

Velleman, J. (2000). On the Aim of Belief. *The Possibility of Practical Reason*. Oxford University Press.

Vitkovskaya, J. (2016). *How the Soviet Union stayed silent during the Chernobyl disaster*. The Washington Post.

Volz, E. et al. (2020). Evaluating the effects of Sars-CoV-2 spike mutation D614G on transmissibility and pathogenicity. *Cell*. Nov. 19, 2020.

Vranas, P. (2001). *Respect for Persons: An Epistemic and Pragmatic Investigation*. PhD diss., University of Michigan.

Walker, A. (1989). The Problem of Weakness of Will. *Noûs*, Vol. 23, No. 5, pp. 653-676.

Walsh, B. (2017). The World Is Not Ready for the Next Pandemic. *Time Magazine*, May 15, Vol 189, No 18.

Walzer, M. (1973). *Political Action: the Problem of Dirty Hands.* Philosophy & Public Affairs, 2(2): 160-180.

Watson, G. (1977). Skepticism about weakness of will. *Philosophical Review,* 86: 316-339

Weatherson, B. (unpublished manuscript). Do Judgments Screen Evidence? available online at [http://brian.weatherson.org/JSE.pdf].

Weatherson, B. (2008) Deontology and Descartes' Demon. *Journal of Philosophy,* 105 (9): 540–569.

Webster, C.; Mahaffy, P.; Atreya, S.; Flesch, G.; Mischna, M.; Meslin, P.-Y.; Farley, K.; Conrad, P.; Christensen, L.; Pavlov, A.; Martin-Torres, J.; Zorzano,

M.-P.; McConnochie, T.; Owen, T.; Eigenbrode, J.; Glavin, D.; Steele, A.; Malespin, C.; Archer, P.; Sutter, B.; Coll, P.; Freissinet, C.; McKay, C.; Moores, J.; Schwenzer, S.; Bridges, J.; Navarro-Gonzalez, R.; Gellert, R.; Lemmon, M. (2015). Mars methane detection and variability at Gale crater. Science, 347(6220), 415–417.

Wedgwood, R. (2012). Justified Inference. *Synthese*, 189: 273-295.

Werner, E.; Gilliam, J. (1984). The Ontogenetic Niche and Species Interactions in Size-Structured Populations. *Annual Review of Ecology and Systematics, 15(1), 393–425.*

West, J. (2014). Joseph Priestley, oxygen, and the Enlightenment. *American Journal of Physiology-Lung Cellular and Molecular Physiology*, 306(2): 111-119.

Whiting, D. (2012). Does Belief Aim (Only) at the Truth? *Pacific Philosophical Quarterly*, 93: 279-300.

Wiggins, D. (1987). *Needs, values, truth*. Oxford: Blackwell.

Wilholt, T. (2009). Bias and Values in Scientific Research. *Studies in History and Philosophy of Science Part A*, 40(1): 92–101.

Williams, B. (1970). Deciding to Believe. In: H. E. Kiefer & Milton K. Munitz (ed.), *Language, Belief, and Metaphysics*. Albany: Suny Press: 95-111.

Williams, B. (1973). Ethical Consistency. *In: Problems of the Self*. Cambridge: Cambridge University Press.

Williams, M. (2001). *Problems of Knowledge,* Oxford University Press, Oxford*.*

Williams, S. (1990). Belief, Desire and the Praxis of Reasoning. *Proceedings of the Aristotelian Society*. XC: 119-42.

Williamson, T. (2011). Improbable Knowing. *In:* Dougherty, T. (ed*.*), *Evidentialism and its Discontents*. Oxford University Press.

Williamson, T. (2014). Very Improbable Knowing. *Erkenn*, 79: 971-999.

Winters, B. (1979). Believing at Will. *Journal of Philosophy*, 76: 243-56.

Woffinden B. (1988). *Miscarriages of Justice.* London: Hodder & Stoughton.

Wolfendale, J. (2006). *Training Torturers: A Critique of the "Ticking Bomb" Argument*. Social Theory and Practice, 32(2): 269-287.

Worley, R. (2003). *Waging ancient war: limits on preemptive force*. Strategic Studies Institute, US Army War College.

Zagzebski, L. (1996). *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge: Cambridge University Press.

Zagzebski, L. (2003). The Search for the Source of Epistemic Good. *Metaphilosophy*, 34(1–2): 12–28.

Zagzebski, L. (2004). Epistemic Value and the Primacy of What We Care About*. Philosophical Papers* 33(3): 353-377.

Zalabardo, J. (2010). Why believe the truth? Shah and Velleman on the aim of belief. *Philosophical Explorations*, 13: 1-21.

Zheng, Y. (2001). Akrasia, Picoeconomics, and a Rational Reconstruction of Judgement Formation in Dynamic Choice. *Philosophical Studies,* 104: 227-251.