The Reconceptualisation of Akrasia as a Character Trait: An Exploration of Akrasia Through the Interplay Between Recurrence, Self-Deception, and Self-Control

Giannaris, Vasilis

The Reconceptualisation of Akrasia as a Character Trait: An Exploration of Akrasia Through the Interplay Between Recurrence, Self-Deception, and Self-Control

Vasilis G. Giannaris

Department of Philosophy, Leiden University

Philosophy (MA) (60 EC)

Philosophical Perspectives on Politics and the Economy

MA Thesis Philosophy 60 EC: 5014VTP6

Dr. Marijana Vujosevic

17/6/2024

*Note.*From *Prometheus and Digital Sculpture,* Eaton, 2007.

Abstract

This paper explores one of philosophy's longest and most enduring puzzles: the puzzle of akrasia. Akrasia, or weakness of will as it is often called, has challenged our understanding of human behaviour for over two millennia. Most literature treats akrasia as an isolated instance of a weak will. However, unlike most approaches, this thesis aims to reconceptualise akrasia as a character trait based on the tripartite relationship between its recurrence, self-deception, and self-control. The proposed view sees recurrent akratic behaviour as an essential factor in that process, which is often overlooked. Agents repeatedly fail to adhere to their better judgement, which supports the hypothesis of reconceptualising akrasia as a trait since it isn't a one-time event but part of a person's character. In this process of rethinking akrasia as a trait, self-deception is conscripted against the agent's better judgment. This creates a bidirectional relationship between akrasia and self-deception, allowing the agents to deceive themselves and further reinforce both. Over time, as the juncture between akrasia and self-deception becomes stronger, it undermines the agent's self-control. As the two manifest, the agent's ability to develop and maintain self-control is hindered, reinforcing akratic behaviour.

Keywords: akrasia, character trait, recurrence, self-deception, self-control

Οὐκέτ᾽ εἰμὶ προσβλέπειν οἵα τε πρὸς ὑμᾶς ἀλλὰ νικῶμαι κακοῖςκαὶ μανθάνω μὲν οἷα δρᾶν μέλλω κακά,θυμὸς δὲ κρείσσων τῶν ἐμῶν βουλευμάτων, ὅσπερ μεγίστων αἴτιος κακῶν βροτοῖς.

Και καταλαβαίνω βέβαια το κακό που πάω να κάνω, όμως πιο δυνατό από τη λογική μου είναι το μένος της ψυχής, αυτό που ευθύνεται για των ανθρώπων τα δεινά τα πιο μεγάλα.

And I understand the evil I am about to do, but stronger than my logic is the anger of the soul, that which is responsible for people's greatest sufferings.

—Euripides, *Medea*

Contents

I. Introduction

How often have we found ourselves in a situation where we knew we should be doing one thing, but for some reason, we did not? Our dedication to exercise frequently has faltered, the once highly impactful paper we were supposed to write has become a product of the imagination, and our resolve to eat healthier and maintain a healthy way of living has withered away. Haven't we all experienced similar situations in our daily lives? We often make judgments of what we should be doing, knowing why we should be doing it, but then we fail to keep up with these decisions. If someone can relate to these, then they experienced the phenomenon of akrasia in one way or another. Akrasia, or weakness of will, is one of the most enduring puzzles in philosophy. Ancient and contemporary philosophers have tried to solve it from various perspectives and approaches for over two millennia.

The topic of akrasia is not just a philosophical puzzle but a complex everyday phenomenon with a long reach and high (philosophical) relevance. Akrasia is a multifaceted phenomenon that intersects with other phenomena and internal (and external) mechanisms shaping how individuals act in different aspects of their lives, from mundane everyday phenomena to moral issues. It constitutes a fundamental part of our human condition, and its implications are evident in various fields; from political decisions such as voting behaviour to moral dilemmas and everyday topics like failing to adhere to a New Year resolution, akrasia is highly relevant and applicable. Its richness, influence, and nuances make it equally relevant for moral and political philosophers. Any case that deals with the complexity of human behaviour can be beneficial for advancing progress philosophically, scientifically, and societally, especially on a topic like akrasia, in which empiricism cannot always identify the cause. Thus, a philosophical investigation seems the most appropriate way to tackle this phenomenon without relying too much on empirical observations.

This thesis argues that akrasia should reconceptualised as a character trait due to the recurrent nature of akratic behaviour, indicating that akrasia is not an isolated instance but occurs recurrently supplemented by the agent's self-deception, which ultimately results in a person's deficiency of self-control further enhancing the akratic trait. The thesis aims to provide a deeper and more nuanced philosophical understanding of human behaviour in this tripartite relationship. Therefore, the goal of this thesis is twofold. First, I aim to answer the question: How does reconceptualising akrasia as a character trait enhance our philosophical understanding of recurrent akratic behaviour? By focusing on it, this thesis not only 'fills a

gap' in the literature but also presents a neglected alternative. The current suggestion in this thesis does not treat akrasia as a logical puzzle that many analytic philosophers, such as Davidson, do. The reconceptualisation of akrasia as a trait I offer is an alternative inspired by Aristotle, Alfred Mele and Paulien Snellen. To our knowledge, Aristotle was the first to look at akrasia and treat it as a trait but explain it as part of a deeper internal conflict between reason and passion. Mele, although he builds upon Aristotle, takes a different route. Although he recognises akrasia roughly as a trait, he explains it through various relevant and concepts throughout his work, such as irrationality, self-deception, self-control. Finally, Thomas Hill and, later, Snellen resurfaced the interest in akrasia as a trait. While Hill focused more on the moral assessment of a weak-willed person and the will and lack of willpower, Snellen focused more on the recurrence of akratic action while criticising previous approaches. While these approaches provide a foundational stepping stone, I wish to develop this line of thought further, introduce new angles, and see what implications it may have on some influential theories.

Second, I aim to explore the implications of considering akrasia as a trait of the influential theories of Davidson, Mele, and Holton. Some of the implications are: I challenge Davidson's assumption that akrasia is always irrational as akrasia as a trait can pave the way for it being potentially rational, therefore undermining Davidson's theory (that Mele also supports). Regarding Mele's account of self-control, it would need to be readjusted as self-control alone would not be able to counter the akratic trait if it is not complemented by strategies targeting an agent's character. Finally, although Holton's theory differs, it would be equally affected. Holton's assumptions would need to be reassessed to accommodate the possibility of both akrasia and weakness of will stemming from character.

Therefore, the structure, as already hinted, will be the following. First, I will introduce some influential accounts of the literature on akrasia. From the ancient to the modern, I will provide some contextual background, from the Socratic rejection until Holton's ingenious shift to intention violation; various arguments will be explored as the discourse around the topic has evolved and transformed throughout the aeons. Second, I will showcase how the reconceptualisation of akrasia as a trait is possible by relying on three arguments, namely the recurrent nature of akrasia, which is often supported by the agent's self-deception, which ultimately results in a person's lack of self-control. Finally, I will examine the possible implications of a character-based approach to akrasia in the discussed theories.

II. A Sketch of Two Ancient Views on Akrasia

The Socratic Rejection

Akrasia, often translated as 'weakness of will', 'incontinence' or 'lack of self-control', has perplexed philosophers (and, more recently, psychologists) for over two millennia[1]. It is one of the most enduring puzzles in studying human behaviour, with its first mention in Plato's Protagoras in the discussion between Socrates and Protagoras. The etymological origin of *Ακρᾱσῐ́ᾱ* latinised as akrasia, combines the prefix ἀ- (a-), which in both ancient and modern Greek denotes a lack or absence followed by the verb *κρᾰ́τέω* (kratéō) meaning to rule or command or having power and connotates the power over oneself in its most common sense[2]. The term akrasia is usually contrasted with its opposite *enkrateia (ἐγκράτεια),* which is composed of two parts: the prefix *ἐν- (en-),* followed by the same verb κρᾰ́τέω (kratéō). One of the first official *philosophical* mentions of the concept of akrasia is in Plato's Protagoras, the famous dialogue between Socrates and Protagoras. Among the themes discussed in Protagoras are knowledge and its status as a virtue. Socrates asks if virtue can be taught and what Hippocrates would gain if he studied under Protagoras. After the discussion went on, they examined what constitutes a pleasant and unpleasant life and what is good and bad, which he (Socrates) argues that most people argue:

> "Many people who know what's best to do are not willing to do it, though it is in their power, but do something else. And those whom I've asked about the cause of this say that people who act in that way do so because they are overcome by pleasure or pain or under the influence of one of the things I mentioned just now" (2002, 352e, p. 46)

Both Socrates and Protagoras seem sceptical of the views of the many. However, the passages highlight the instance of akrasia without naming it explicitly. Later, Socrates asks how a man does evil, knowing it is evil. To disprove the assumption of the many about akrasia, he employs his famous dialectical method to try to undermine the opinion of the many and expose the paradox of how someone can be acting against his better judgment knowingly. He

[1] For the sake of clarity and simplification, I will use akrasia, weakness of will and incontinence interchangeably throughout the text up until the point I discuss Holton. Holton differentiates akrasia and weakness of will.

[2] Other easy examples are apatheia/απάθεια (without pathos), the lack of passion, aporia/απορία a and poros, the expression of doubt or not knowing or aphasia, the lack of ability of one faculty of speech and so on.

argues that if they truly knew, they would act accordingly. He then introduces the hedonistic equation, arguing that good equals pleasant and evil equals painful and then discusses the nature of pleasure and pain in action and ethical behaviour. He uses his dialectical method to test Protagoras and his consistency in his views. However, there is a debate about whether Socrates endorses hedonism in Plato's Protagoras, as many passages supposedly show, but I won't delve into this. However, as Zeyl argues, Socrates wasn't a hedonist and used hedonism against the many to disprove akrasia (Zeyl,1980, p.260). Ultimately, Socrates rejects the assumption that akrasia exists and points out that akrasia is a state of ignorance since the people who act akratically don't know what is good.

Both Allen (1996) and Tenenbaum (1999) agree that Socrates' explanation relying on the hedonistic calculus is absurd (p.142). However, they imply that Socrates holds the position and does not just use it for the sake of argumentation. They argue that Socrates' argument in Protagoras relies on his identification of the good with pleasure. However, Socrates' identification of pleasure with the good means that the agent aims to fulfil his desires because he perceives them as good, which seems to infer two things, according to Tenenbaum. First, "people that are overcome by pleasure to pursue an evil is a ridiculous opinion," and second, if we are to accept the Socratic conclusion, "we would deny an extremely common phenomenon" (Tenenbaum, 1999, p.876). The hedonistic equation that relies on the premise that what is good is synonymous with pleasure and what is bad is synonymous with evil is what is often criticised. As the two authors have claimed previously, it is difficult to think that someone would knowingly choose pain because they are attracted to the good.

Moreover, I believe the difficulty lies in the fact that someone would willingly choose a course of action which they know to be painful (however, pain in itself is not the central issue here) but the inherent characteristic of evil that predetermines pain motivated by a pleasure that is not even worth pursuing to justify this choice. Most people find this explanation problematic because of the logical inconsistency of being overcome by what is good to do something evil. Ultimately, Socrates thinks that akrasia is a state of ignorance. In his view, knowing the good leads one to do good. Therefore, those who do not do good or act against their better judgment are ignorant. The problem for Socrates is that people lack a 'correct' understanding of what is genuinely a better course of action. This ignorance leads to choices that do not align with the true good.

Aristotle, Moral Character and Akrasia

The English word character derives from the ancient Greek verb *charàssō* [3]*(χαράσω)*, meaning to engrave or sharpen, which was later used to denote a mark, a distinctive quality to the word character. In later usage of the term, the concept came to include features or qualities that distinguish between individuals and is often associated with personality. In the philosophical context, Aristotle used the term in-depth, with a deeper and more complex meaning. Aristotle accepts the metaphysical assumption that Plato introduced, which is the distinction of the soul into rational and non-rational parts, and the virtues are associated with the two parts of the soul (1103a1-10). Building upon this notion, Aristotle's virtues are divided into intellectual and ethical. Ethical derives from êthos *(ήθος),* which is closely related and associated with (moral) character etymologically and conceptually. The intellectual virtues are affiliated "with the part of the soul that associates with reasoning and intellect (virtues of mind) while the other part was not devoid of reason but capable of following reason and were associated with character (virtues of character/ethical)" (Homiak, 2019, section 1).

As Homiak argues, when we speak about moral excellence, "the emphasis is not on mere distinctiveness or individuality, but on the combination of qualities that make an individual the sort of ethically admirable person" (Homiak, 2019, section 1). Therefore, moral character and its derivates should be understood as something beyond mere simplifications of individuality that offer distinctive attributes to one's person. It should be understood as a way that influences people's thoughts, feelings, and actions, which people can develop through training and (moral) education. How the individual becomes this ontological being represents a moral entity of many traits, virtues, and ethical considerations.

Then Aristotle further divides intellectual virtues into two kinds: those that relate to theoretical reasoning and those that relate to practical reasoning (1139a3-8). The second kind of virtue is a matter of character (and vices are, too). Virtues and vices are products of habituation or hexeis (ἕξεις). Aristotle argues, "By habits, I mean our good or bad relationship towards feelings or passions"[4] *(*Aristotle, 350/1926, 1105b, p.87*)*. However, hexeis can also be translated as a disposition. Whether one acts between the extremes (excess or defect) depends on his virtuous character or lack thereof. Aristotle's solution was to

---

[3] Therefore, the derivatives of χαρακτήρας (character), χαρακτηριστικό (characteristic) and χαρακτηρίζω (characterise).

[4] *Εξεις δὲ καθ' ἃς πρὸς τὰ πάθη ἔχομεν εὖ ἢ κακῶς, οἷον πρὸς τὸ ὀργισθῆναι*: Some English translations translate ἕξεις (habits) as states.

introduce the concept of the golden mean, which would be the agent's moral 'compass' and help him decide and act accordingly to avoid potential defects or excesses. Of course, for one to be virtuous (according to Aristotle), one had to be habituated into acting virtuously, deliberate rationally and had to have a certain level of moral education (which can be counted as a limitation if one is to consider who could have a proper education in ancient Greece) and the acquisition of phronesis or practical wisdom as a consequence.

Moreover, for Aristotle, acting virtuously was not just taming your feelings and passions according to reason's command. The feelings need to work in accordance with reason in a symbiotic and harmonious relationship for the individual to act virtuously. The disposition to act virtuously was a product of a moral character pertained to certain qualities and habits. The question and discussion about moral character was prevalent in ancient philosophy and has gained attraction not just in the context of moral philosophy. From Marx to Rawls and Anscombe, the concept of moral character has been used in various contexts and fields. However, there are also sceptics of moral character not just as an opposition to virtue ethics, which are directly contrasted through other ethical theories such as Kantian deontology, utilitarianism, or other ethical theories, but through a sentiment of expressing scepticism of the ontological status of moral character or in other words if the character exists at all. A scepticism over the overall existence of a moral character, one might say.

In book seven of Nicomachean Ethics, Aristotle begins his quest to explain the paradox of akrasia[5]. He developed the most systematic analysis of akrasia among the ancient philosophers. The puzzle of akrasia is investigated in the context of his virtue ethics, which relies on the assumption that moral goodness can be cultivated based on habituation and rational deliberation. In the first lines of book seven, Aristotle mentions the three conditions of character that should be avoided: vice, akrasia and bestiality. When it comes to Aristotle and his views on akrasia, there are many readings and interpretations, but this essay aims to present only some relevant views. Unlike the Socratic rejection, Aristotle believed in the existence of akrasia. However, he did not consider it as blameworthy as a vice; he placed it somewhere between a vice and a virtue. Robinson argued that Aristotle rejects the Socratic thesis on the ground that "the man whose conviction is weak is pardoned, whereas the acratic is blameworthy" (Robinson, 2010, p.150). In addition, Aristotle divided akrasia into two sub-levels: impetuosity (προπέτεια) and weakness (ασθένεια). While the latter deliberates and fails to keep up with their deliberation because some passion undertakes them, the former

[5] My view does not exhaust all the potential readings of Aristotle on akrasia.

doesn't even deliberate, gets overtaken by some passion, and acts akratically. It is essential to mention that both agents are considered akratic; however, the second, the weak (ασθενής), goes through a process of deliberation, which is most often associated with akratic actions, and it is this version that is of interest to this essay rather than the former the impetuous (προπέτης) agent which might be acting akratically unconsciously which some modern theories attribute to certain strands of self-deception. As Bratman pointed out, the cases of impetuosity are to be disregarded since these cases are devoid of deliberation (Bratman, 1979, p.153), and to be considered akratic, the agent must act consciously, as Davidson's requirements will shortly show.

The distinction between the two indicates that the impetuous agent experiences no internal conflict like the weak, which rationally deliberates and chooses the worst action. The impetuous only retrospectively realises his mischief and, therefore, regrets it. Aristotle's distinction between the impetuous and the deliberative indicates another essential aspect—the various levels of moral responsibility. The weak could be held morally accountable since there was an instance of rational deliberation and eventual failure, while the impetuous acted impulsively and gave in to their passions. This distinction is important because for the akratēs to be considered akratic, his actions must be free and autonomous to be held accountable and (morally) responsible. The autonomy of the action is essential for keeping the agent accountable and responsible through his agency and the assessment of his akratic actions. Without this distinction, the argument of akrasia as a trait would lack a foundation in accountability.

III. Several Contemporary Approaches to Akrasia

Donald Davidson's Logical Puzzle

Contemporary approaches to conceptualising akrasia (or weakness of will, as many modern philosophers call it) follow a different path than the ancient view. Western analytic philosophers have scrutinised the paradox of akrasia and introduced different perspectives that have influenced the nature of the debate. Davidson, while a pioneer who contributed to expanding the discussion, raised a lot of criticisms. I focus mainly on one of Davidson's essay, which is the most famous and influential regarding akrasia. There, he sets the stage to explain how incontinent actions are possible by offering a solution to what he deems as a logical puzzle. It is considered a puzzle because it seems impossible that all of Davidson's

principles are correct while not contradicting each other. Critics of Davidson have expressed their concerns about P1 and P2 and the connection to a person's motivation to action. In his essay, *How is Weakness of the Will Possible?* Davidson discusses for the possibility of a weak will. His argument goes as follows:

"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 incontinent actions." (Davidson, 1969, p.23)

Already in P1, Davidson mentions many important concepts such as wanting, believing, freedom and intentionality. Out of the four, intentionality is crucial here. Intentionality is one of the prerequisites for an action to be considered akratic. Intentionality is paramount because it indicates akratic action, which separates it from compulsive action or addiction. Intentionality is a prerequisite in Davidson's terms for someone to be considered akratic and claim (moral) responsibility. Davidson argues (and it is generally accepted) that an agent acts akratically if three conditions are fulfilled. If the agent does x intentionally, the agent believes there is an alternative open to him, and the agent judges, all things considered, it is better to do y over x (Davidson, 1969, p.22). Having said that, I don't consider an agent acting compulsively as akratic as he lacks Davidson's criteria, which I mentioned earlier. Intention is a complex and multifaceted concept to which many thinkers have devoted many pages. Anscombe famously explains intention in three senses: intention for the future, intentional action, and intention in acting (Anscombe, 2000, p.1). The intention is not just important because it connects with future long-term goals and whether we act towards them but also because intention in acting shows the agent's awareness during the performance of the action that Davidson pointed out. Ergo, intentionality presupposes the agent's freedom of choice (and therefore the possibility of akrasia) since an unfree, coercive, or any other form of compulsive action does not qualify as akratic. Believing is also crucial as it sets the ground for the agent to cement his reason for acting on something. The subjective belief is enough of a motivation from an internalist perspective for the agent to act, as shown in Davidson's second premise.

In P2, the element of subjective good or imagined good, as Davidson calls it when he quotes Aquinas (Davidson, 2001, p.22), is contrasted against some universal standard of good

or virtue, for instance. Moreover, as Davidson admits, P2 "states a mild form of internalism" (Davidson, 1979, p.26). Davidson hints at internalism because it shows a link between an agent's judgment and their motivational states. I believe that Davidson calls it a mild form of internalism because it presupposes some coherence between an individual's evaluative judgement and their motivations. Proposition two hints at the idea that the source of motivation lies inside the individual, and its motivation to action is because of the agent's rational assessment or desires without the need for external factors to bridge the gap between judgment and action. Davidson's argument emphasises the internal conflict between what an agent judges best and acting against his better judgment. Davidson doesn't seem to adopt the strict Aristotelian view of reason vs passion, which is also a form of internalism. He takes a different route that highlights the internal conflict and focuses on the process. However, he mentions a competing image found in Plato involving three actors: reason, desire, and the one who lets desire get the upper hand. The will or conscience ultimately decides who wins the battle (Davidson, 1969, p.35).

Davidson, being unhappy with the ancient explanations revolving around akrasia, argued that the problem is deeper than what Aristotle thought and resembles more of a logical puzzle. Given his solution to the problem, critics of Davidson argued that P1 and P2 contradict P3. Throughout the text, Davidson argues against this and claims he wasn't happy with giving up any of the propositions, so he had to find another way of solving the issue (Davidson, 1969, p.23). To tackle the problem and showcase how akratic action is possible, Davidson distinguishes between two different types of judgments: conditional (p3) and (p2) unconditional[6]. As the name implies, conditional judgements are conditioned based on the premises, whereas unconditional judgements are independent of any premises. The conditional judgement is expressed as in *pf* or *prima facie* judgement.

Suppose it is lunchtime, and I judge that buying a burger is the best option as it's faster than cooking, and I am starving. My judgment that "if I order a burger, I will satisfy my hunger because it is faster than cooking" is a *prima facie* conditional judgment. It holds as long as my conditions (or, in Davidson's terms, are expressed as r for reasons, which can be counted as causes) for quick eating and satisfying my hunger. However, imagine I came home from a long day at university, commuting for over an hour and spending a lot of money on ticket trains, coffee and cigarettes. I am now reconsidering my choice of buying a burger (since my unconditional judgment is that 'I should not eat junk food). My prior *prima facie*

[6] Davidson seems to consider an unconditional judgment as an intention. (Davidson,2001, p.99)

judgment was initially valid but conditional and subject to change against my unconditional judgment. Therefore, Davidson argued that judgements must consider the relationship between other judgements, meaning how they compare different aspects of other judgements.

The second kind, the unconditional judgement, is independent of any premises and expresses some universality without any qualification, just like logical truths. For example, the statement (often used in epistemology), 'all bachelors are unmarried men, which expresses a tautology', can be considered an unconditional judgement based on a logical truth. So, Davidson's solution to the problem of akrasia was the distinction between two different judgements and, ultimately, the agent acting against his conditional judgment. The agent judges that it is conditionally better to do x over y but judges unconditionally to do y over x. Ultimately, Davidson argues that akrasia or weak will is possible because the agent acts against his conditional or all-things-considered judgment (x is better than y) based on his unconditional judgment (y is better than x). So, in the prior example, I act based on my unconditional judgment, 'I should not eat junk food ' against my conditional judgment, 'if I buy a burger, it will satisfy my hunger'. Davidson's solution doesn't immediately solve all the problems, such as the role of practical reason, which is usually understood as "the capacity for resolving, through reflection, the question of what to do" (Wallace, 2020, Section 1) in akrasia, or as Snellen argued, why isn't it possible to act against the unconditional judgment? (2018, p.43) Mele has raised same concerns. These questions remain relevant today.

Alfred Mele's Scepticism Towards Davidson

In this section, I will explore some of Mele's arguments regarding akrasia. I will focus on his most relevant arguments and provide explanations when necessary. To begin with, Mele wrote extensively on akrasia, self-deception, and self-control, amongst other topics. In his paper, *Akrasia, Reasons and Causes,* Mele argues for the possibility of akratic action against the Socratic rejection. To tackle the paradox of akrasia, Mele argued that a causal theory of action was an attractive solution, and it could bridge the gap between akratic actions and what he called between here-and-now intentions, "an intention of the agent to do an A here-and-now" (Mele, 1983, p.345).

Furthermore, Mele was unhappy with Davidson's solution to akrasia. Mele rejected Davidson's P2 because he argues that the relationship between unconditional judgements and an agent's motivation is much more complex than Davidson thinks (Mele, 1983, p.366).

Motivation is a different force that does not always align with an agent's evaluation. When an agent is evaluating a course of action or even a particular choice, it is not certain whether or not he will abide by his evaluation. In instances of akrasia, the agent is usually aware of his evaluative judgement that *a* may be a better course of action than *b* but still fails to adhere to its realisation. Supporting Mele's remarks, Watson argued that Davidson's P2 "may be true if understood in the language of evaluation but false if understood in the language of motivation" (Watson,1977, p.321).

In light of Davidson's P2, Mele argues against it by saying that Davidson's P2 doesn't always settle what the agent should do because judging isn't always translated into motivation, as someone may often judge x over y but fail to act on x. He adds that his unconditional judgment is not an intention (Mele, 1983, p.356). The two key takeaways from Mele's causal theory of action paper are that incontinent action is compatible with reasons being causes of actions (Mele, 1983, p.364) and the concept of self-control being an essential factor between one's intention and motivation.

Changing the topic from his causal theory of action, Mele, in his book *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control,* attempts to explain the problem of self-deception that manifests itself into akrasia. Again, he claims that akrasia "strictly speaking is not a type of action but a trait of character " (Mele, 1987, p.3). Although Mele thinks that akrasia is a character trait in its strict sense, he is more concerned about the relationship between irrationality and akrasia rather than akrasia as a character trait and doesn't investigate it. In the book's sixth chapter, he explains how strict akratic action occurs as a paradox of irrationality. Mele again expressed his concerns regarding Davidson's P2 and the expressed internalism it presupposes. Having set the ground in his previous papers, Mele introduces his argument regarding the subjective irrationality in an agent's actions. In short, Mele claimed that if the agent intentionally and freely performs b while holding the all-things-considered judgement, it is best not to do b, then the agent is acting subjectively irrational (Mele, 1987, p.5). Mele seems to agree with Davidson on the point of irrationality as the agent is acting freely and intentionally against his all-things-considered judgment since they both presuppose that not to do b would be the rational order of things; thus, acting against it is irrational. In the last section of the paper, I will contest (Davidson's) point in the implications and discuss whether akratic action is *always* irrational.

Exploring Richard Holton's Theory of Weakness of Will

While Mele criticised Davidson's conceptualisation of akrasia, Richard Holton takes a distinct approach to the debate. Holton's theory doesn't consider the possibility of akrasia being a trait. Mele also noticed it and said that Holton "seems to be treating weakness of will as shorthand for something like 'displays of weakness of will'; he gives no indication that what is displayed must be a character trait" (Mele, 2009, p.395). Nevertheless, Holton's theory presents a new alternative to the debate.

Instead of focusing on criticising particular arguments of a philosopher, Holton chose to criticise the entire discourse regarding akrasia, and for good reasons. His contribution not only revealed some misalignments between the common conceptions of understanding akrasia against the philosophical one, but he also expanded the nature of the debate by offering a new view on the topic. In the first lines of his *Intention and Weakness of Will* paper, Holton claims that non-philosophers never mention better judgements or better or worse course of action (Holton, 1999, p.241). Instead, he discovered that people understood the weakness of will as being too irresolute and not persisting in their intentions. His argument resolves around the idea that people fail to act on their intentions or are ready to reconsider their existing intentions too soon. Therefore, Holton introduces a new approach to the debate the intention violation. For him, it doesn't matter whether there is a better course of action at a specific time but how quickly the agents abandon an already formed intention.

Unlike the previous authors, Holton distinguishes between the weakness of will and akrasia; for Holton, akrasia refers to an agent acting against his better judgment, while the weakness of will would be the failure to adhere to one's original intentions. Holton's distinction relies on the assumption that the common opinion differs from the one the 'experts' have. However, some authors like Mele expressed their scepticism regarding Holton's distinction. Mele argues that in the studies he conducted, "no indication that our ordinary notion is captured better in terms of intention or decision than in terms of knowledge or belief " (Mele, 2012, p.20) and that "people occasionally make theoretical errors that taint their judgements" (Mele, 2001b, p.41). Mele points out that ordinary people's understanding of complex theoretical concepts is often mistaken (not that experts can't also be wrong). But, from what I understand, the researchers have undergone at least some reflective research and deliberation procedures compared to intuition. But even Holton himself was unsure of his original findings on the common understanding of akrasia or weakness of will as he, in his joint paper with May when May argues that "Holton should be much less confident in the

existence of an ordinary notion of weakness of will that only involves resolution-violations" (Holton, 2012, p.357). Holton's idea that the common understanding of the weakness will *only* involve resolution violation seems to treat the topic of akrasia superficially, and the empirical data he relies upon appears to have questionable methodology. Therefore, we should not make hasty generalisations based on similar data; we should treat and scrutinise it cautiously.

IV. The Reconceptualisation of Akrasia as a Character Trait

As was shown in the previous section, although there was scepticism regarding the existence of akrasia from the likes of Socrates, there is a general (contemporary) consensus that akrasia or weakness of will exists in one way or another. Although there is no consensus on what causes akrasia from Davidson's logical puzzles and Mele's no connection between judgement, motivation, and action to Holton's intention violations constitute some current influential conceptualisations of akrasia from different contemporary standpoints. Sometimes, the current strands complicate akrasia rather than simplify it. However, there is room for improvement, and the reconceptualisation of akrasia as a trait can fill the literature gap and further expand the debate.

Most modern explanations share a common similarity about akrasia: they treat it as an isolated instance. Ancient and medieval literature treats akrasia as a moral problem because it obstructs the agent from living a good (ethical) life or a life without sin. Snellen (following Aristotle) treats akrasia as a character trait. Her account of akrasia refers "not to a single and isolated episode but to a tendency of judgement violation, that allows all sorts of failures to abide by one's better judgement" (Snellen, 2018, p.19). She proposes the treatment of akrasia as a character trait by emphasizing the repetitive nature of akratic behaviour, which I believe is often overlooked by modern approaches. Inspired by Snellen, I build my argument on reconceptualising akrasia as a character trait by exploring how the recurrent nature of akratic action indicates a connection with the character. However, I diverge in my analysis and explanation as my analysis, although it relies on the same starting point, shows how this trait manifests through the conscription of self-deception, which ultimately decreases the agent's self-control. My claim is a mixture of different arguments that I synthesise under the umbrella of a character trait.

Kraut argues that in book seven, "Aristotle investigates *character traits* [emphasis added] – continence and incontinence" (Kraut, 2022, section 7) and rightly so since he

(Aristotle) argues that "three states that must be avoided related to character"[7] (Aristotle, 350/1926, p.375) at the beginning of book seven. As already mentioned, the word ethos (ἦθος – ήθη in plural) is often translated as character. This view is reinforced when Mele argues, "What Aristotle called akrasia is, very roughly, a trait of character exhibited in uncompelled intentional behaviour" (Mele, 2012, p.3). Aside from Aristotle, Mele and Snellen arguing for akrasia being a trait, Hill, in his essay, *Weakness of Will and Character*, focuses on treating weakness of will as a character trait rather than an isolated act (Hill, 1986, p.94). These works, starting with Aristotle, show that a character-based approach to akrasia exists but is uncommon. Aside from Aristotle, only Snellen entertains this argument thoroughly. She is the only one who offers a more complete development of the argument of akrasia being a character trait, although she mainly relies on the argument of recurrence.

First, when I talk about character traits, I do not mean the global character trait Doris had in mind, nor do I adopt a strict virtue ethicist approach. For Doris, global character traits have two main features: consistency and stability (Doris, 2002, p. 22). I agree with Rorty's claim that "akrasia is characteristically regional" (Rorty, 1980a, p.205). Although akrasia can be rethought as a trait, it doesn't mean it is pervasive across all decisions and actions. It is usually expressed in specific areas or aspects of one's life.

Furthermore, I define a character trait as an amalgam of relatively stable causal dispositions that influence an individual's thoughts, feelings, and actions, which are prone to rational deliberation and have a moral underlying tone[8]. I choose the word disposition because the term is broad enough to encapsulate mood, inclinations, or a person's tendency to act in a specific way. The idea behind this definition isn't to show that a character trait is an absolute and robust construct that makes an agent behave in a specific way under any circumstance. With this definition, I want to show that traits, although relatively stable, can be influenced by different scenarios and contexts. The manifestation of a trait might vary depending on the situation and social or cultural context. Human behaviour is the product of a multitude of interrelations and isn't only determined by internal dispositions; however, this doesn't negate the existence of character traits that often define one's (moral) character, behaviour, and action.

---

[7] "(ήθη φευκτῶν τρία) 1145a"

[8] The definition I used was based upon Snellen's (2018, p.19), Kristjhansson's (2010, p.27), Adams's (2006, p.130), Mumford's (1998, p.119), and Brandt's (1970, p.25) accounts of character traits, moral self, and dispositions. However, I added an extra layer to fit my conceptualisation as Snellen treats them as patterns instead of dispositions. She also doesn't include rational deliberation and their causal nature.

Moreover, it recognises traits' power as not just descriptive elements but as dispositions that have a causal effect and drive how a person thinks, feels, and acts, which is inspired by Mele's causal theory of action as causality can often explain many psychological events of an underlying behaviour. Also, character traits differ from personality traits as the latter are not grounded in a moral dimension like the former. Kristjánsson's account of character traits illuminates the picture of distinguishing them from personality traits when he claims that "character traits are usually reason-responsive and are related to one's moral worth" (Kristjánsson, 2010, p.27). My conceptualisation of akrasia as a trait argues that causal mechanisms such as desires or beliefs or, in this case, a trait can sometimes override long-term goals or practical deliberation. People consciously and recurrently act akratically against their better judgements. The cause, in that case, can be attributed to their akratic character. The akratic behaviour an agent displays is usually caused by something. Even if conditions and context may change, a person can still act akratically and not just once.

When thinking about agents' past and present behaviour and, in this context, constant misalignment with their better judgement should trigger many reactions as a critique to contemporary understandings of akrasia and how this phenomenon is possible. How is it possible for people recurrently to fail to abide by their better judgment, whether that is an everyday action or a moral one and still fail to recognise the recurrent nature of this phenomenon? To attribute this to a character is not a stretch of the imagination. Building on the definition I provided above can accommodate the multifaceted nature of a trait, which can include mechanisms such as self-deception, which I will mention later. It is essential to think of a trait beyond mere consistent behaviour; although consistency is a factor, it is not the only one. My previous definition highlighted the multifaceted nature of a character trait's ability to affect a person's thoughts, feelings and actions inside a causal theory of action. Having this definition as a starting point, it makes sense to argue for the multifaceted nature of a trait, its long reach and its interwovenness with other concepts, as it doesn't operate in a vacuum.

Finally, an akratic character trait developed through habituation does not render it a form of a predetermined event, as the agent plays an active role in forming it. The agent can reflect and alter his behaviour. Moreover, it hints at the agent's moral responsibility before acting and deliberating. Even if akrasia is ingrained, the agent consciously makes these decisions. The capacity for reflection and deliberation indicates the agent's autonomy. Modern research demonstrates how conscious interventions such as mindfulness and moral education are effective in helping people battle negative traits or develop positive ones. Thus,

traits are not fixed determinants of behaviour, nor do they render the agent incapable of thought, action, and responsibility.

Recurrent Akrasia

In this section, I will argue that akrasia can be reconceptualised as a trait due to its recurrent nature, which indicates a formed pattern rather than an isolated event. Imagine the following example: New Year's resolutions. People often commit every new year to a healthier lifestyle, a new creative hobby, or commit to a particular judgement that they deem better or advantageous for themselves, "This year, I will do x". People know that these resolutions constitute their better judgement, and by following these, not only will they be sticking to their better judgement, but they will also improve their lives one way or another. However, most people fail to meet New Year resolutions by recurrently acting against their better judgements. It is essential to realise that they don't fail once but repeatedly. If a once-committed person who wanted to lose weight, for instance, only skipped the gym once, it wouldn't be a big deal since we all once in a while are not in the mood, tired from work or simply lazy; we are humans after all. However, most people will recurrently act against their better judgment to work out and instead stay home and do something else. Recurrence is a common feature of akratic behaviour, which is why I am proposing treating akrasia as a character trait. A common denominator in the scenario above is the agent's character, which remains relatively stable. Akrasia does not just happen in a vacuum, nor does it happen once. Suppose akratic behaviour occurs recurrently rather than in isolated instances. In that case, a relatively stable identifiable trait may manifest as a pattern at times of judgment violation rather than just an occasional incident, as the literature implies. One may ask how this relatively consistent recurrence can be attributed to a character trait. In other words, does it make sense to talk about a weak-willed character?

To answer this question, I want to focus on the consistency underlying this recurrent akratic behaviour. Although a trait may not be translated in every manifestation of akratic behaviour, the recurrence underlies a correlation (although one can argue that correlation does not infer causation) between a person's behaviour and character. An agent's (consistent) inconsistency expresses itself in a way that goes against one's better judgement, which suggests that when the agent experiences internal turmoil on the question of '*what to do'*, it is his character that intervenes and provides an easy way out since the recurrent akratic behaviour constitutes a familiar solution which has been formed and chosen several times

already thus making it easy for the agent to act akratically again. The agent has been conditioned through the recurrent nature of akrasia to act this way repeatedly, which demonstrates why the agent is susceptible to this kind of recurrent behaviour.

The akratic trait is not just a fleeting occasional preference the agent tends to choose or an instance of a momentary impulse. It is consistent and manifests differently across various contexts, thus justifying recurrent akratic behaviour. Through his recurrent mental conditioning, the agent repeatedly fails to adhere to his better judgement and succumbs to the influence of his akratic trait. Recurrent mental conditioning is a crucial factor here because the influence of akrasia as a trait makes the agent repeatedly fail to adhere to his better judgment. The process of mental conditioning begins with the formation of habitual patterns of behaviour where these repeated behaviours reinforce the akratic pattern of thinking and acting. If the agent consistently chooses to overindulge in drinking or procrastination over his better judgement(s), then his choices would become habitual and deeply ingrained into his character. When the agent consistently succumbs to his desire to check his social media, the behaviour might be reinforced by his brain reward system or his dopamine excretion, making it more likely to occur again in the future. Despite the agent's better judgement about his course of action, traits with already established conditioning can override these judgments. What is important to mention here is that the more the agent fails to abide by his better judgments, the more the akratic trait is reinforced, ensuing in akratic behaviour. As a result, when it comes to critical moments in which the agent must choose what to do, his actions are influenced more by his trait rather than his rational deliberation. That shows how akrasia as a trait creates a pattern of undermining rational deliberation, further enhancing akrasia.

Think of lying as an example. An agent tends to repeatedly lie about his personal life or work and deceive others in some ways for his gain while recognising that lying is wrong and he shouldn't be lying. Although lying may be expressed differently in each agent (withholding information, straightforward lying, or ignoring facts), there is a common understanding of what lying is and the moral connotation such behaviour holds. While there can be exceptions, such as the so-called '*white lies*,' lying usually has a negative connotation from a moral standpoint and societal standpoint, so the better judgement of a person would be not to lie as it is immoral and may have negative (ethical, personal, societal) consequences on the agent. While the agent knows all this, he consciously lies instead of telling the truth. When this behaviour occurs repeatedly, it suggests that the agent has a particular underlying disposition to act in such a way and doesn't happen out of the blue. When lying occurs recurrently, an individual with a disposition towards lying justifies that his actions are, to a

certain extent, conditioned by his akratic character trait (absence of moral education would be another factor interrelated to character). This agent repeatedly chose to lie despite knowing the ethical implications of his actions, which didn't happen once after an isolated incident of poor judgment.

The behaviour highlighted above can be argued that it is instigated by his akratic trait, which has a causal power to make the agent act in a specific way regardless of the context (whether it is in his personal relationship, work or idle chat) since the agent had already acted like this and is familiar with it. His trait can initiate an 'appropriate' akratic behaviour when deemed beneficial for him to do so. After all, akrasia provides a solution in cases of uncertainty. The agent in the above example (should not be mistaken with a pathological liar, which is considered a compulsive disorder and not an instance of akrasia) possesses an akratic trait manifesting in lying, which makes him act in a certain way against his better judgement. Of course, the example of lying has more profound implications because it deals with moral weakness, moral education, moral relativism, ethics, etc. However, the fact that an individual who is fully capable of ethical reasoning and doesn't lack any cognitive capacity but consciously and recurrently engages in akratic behaviour while knowing that lying is 'wrong' suggests a deeper underlying issue that the character approach seeks to address.

Self-Deception and Akratic Character

Self-deception, whether pursued independently or in relation to akrasia, is a rich topic with many nuances[9]. In the philosophical literature, the likes of Davidson, Mele, and Rorty, who have also written on the subject of akrasia, have contributed to the topic and the intersection between the two. In the previous section, I argued how the recurrent nature of akratic actions justifies the assumption of the reconceptualisation of akrasia as a trait. This section highlights the bidirectional relationship between akrasia as a trait and self-deception. The account of self-deception I rely upon is the deflationary one that Mele proposed, which I will explain shortly. The general implications of this relationship are important in understanding how akrasia can be rethought as a trait through this bidirectional relationship.

[9] My research is limited, so I can only partially judge whether self-deception is an aspect of akrasia. Self-deception is an entirely different puzzle on its own. While the two can work together and are related, they can exist without the other. That doesn't undermine my argument, as it shows the complexity of human psychology, and acknowledging their independent existence enriches my approach by showing how multifaceted both are.

The systematic use of self-deception as a means to justify akratic action by the agent reinforces not just my hypothesis but also the interconnection of the two.

By highlighting the relationship between the two, I don't just focus on the weak lapse in willpower that Davidson hinted at or on the conflict between reason and passion that the ancient Aristotelian view promoted. Instead, I aim to show how a mechanism such as self-deception plays a crucial role by arguing that the akratic trait conscripts self-deception to achieve its goal, which is the fulfilment of the akratic tendencies. This two-way relationship reinforces both akrasia and self-deception. Self-deception as a mechanism conscripted by the akratic trait explains how individuals recurrently fail to act by their better judgment. Self-deception is crucial for understanding akrasia as a trait because it shows how a psychological process links the behaviour as a biased (often rationalised) way of processing information or misinterpreting data in favour of akratic behaviour. The akratic trait uses self-deception to sustain itself, as the maintenance of recurrent akratic behaviours needs solid ground to justify them, which self-deception provides. My hypothesis challenges how we view the debate and hints at potential solutions. The habitual reliance on self-deception after the initial conscription by the trait to justify akratic actions leads to a repeated pattern of behaviour (since it helps the agent to act akratically) making these actions more stable and identifiable as possible parts of a character trait. Due to the difficulty and the vastness of the topic of self-deception, I will mention some background information and the essential elements around the debate of self-deception before explaining the connection between akrasia and self-deception.

Demos defined self-deception as "when a person lies to himself, that is to say, persuades himself to believe what he knows is not so. In short, self-deception entails that B believes both *p* and *not-p* at the same time" (Demos, 1960, p.588). This definition of self-deception showcases that instances of self-deception must meet strict requirements. But then the following questions arise. How is it possible for someone to hold two contradictory beliefs? After Demos, many became interested in the concept. Davidson characterised it similarly to incontinent actions as clouded judgments (Davidson, 1969, p.29). Although they share many similarities, self-deception can be seen as another weakness, albeit a different one than akrasia, manifesting in a different form. Still, self-deception is also a process the agent uses to achieve a goal when faced with a complex scenario, and it can exist independently outside of the akratic paradigm. Self-deception doesn't always have to be towards oneself. It can be directed towards others, but as Rorty argued, "self-deception involves deception of the self, by the self, for or about the self" (Rorty, 1972, p.393).

There are three main approaches to self-deception. The *lexical* one, in which proponents use a standard lexical definition of deception (and usually reject that self-deception exists), the *exampled-based* when one looks at representative examples and standard features and a *theory-guided* which is guided by common sense principles and aetiology (Mele, 2001a, p.5). Hybrids also exist. They all come with different assumptions and limitations. The lexical approach has two assumptions:

1. "Person *A* deceives person *B* (where *B* may or may not be the same person as *A*) into believing that *p* only if *A* knows, or at least believes truly, that -*p* and causes *B* to believe that *p*.

2. By definition, deceiving is an intentional activity: non-intentional deceiving is conceptually impossible" (Mele,2001a, p.6)

However, criticism has been raised towards both definitions. Regarding the first one, knowing *p* and believing ~*p* seems problematic for many critics. How is it possible for one person to consciously hold two contradictory beliefs? According to Mele, this is the *static* paradox, which "requires being in an impossible state of mind" (Mele, 2001a, p.7). The second assumption is about a *dynamic* paradox, in which Mele asks what prevents the guiding intention from sabotaging its functioning (Mele, 2001a, p.8). These two paradoxes caused the further development of the concept of self-deception and its different approaches. According to their stance towards these paradoxes, self-deception is further categorised into the intentionalist and revisionist approaches, each having further sub-divisions. The former looks at the concept of self-deception as a way to accept the responsibility of self-deceived agents and distinguish it from wishful thinking. In the traditional intentionalist model, the deceiver believes *p* while also truly believing ~*p.* So, the deceivers must hold a contradictory belief (the so-called dual belief requirement) and intentionally get to have a belief that they know is false. To accommodate their views, they introduce various divisions and partitioning of the self and the mind to achieve their goal. Mental exotica, as Mele calls it (Mele, 2001a, p.4).

On the other hand, revisionists are sceptical regarding these divisions, and they revise the intention or belief requirements. They try to simplify their approach by saying that by revising the belief requirement or the intention requirement, they can offer a more straightforward solution that avoids the paradoxes of self-deception. So, by revising the

intention requirement, the revisionists treat self-deception as a motivationally biased belief, thus avoiding the problem of the agent deceiving himself with an intention.

I choose Mele's deflationary model (a revisionist approach) for self-deception as it is the most appropriate to complement akrasia as a trait. It takes an interesting route away from the unconscious belief and the paradoxes. For instance, Mele argues that *desire* may be a strong motivation in cases of self-deception supplied by ignorance or uncertainty (Mele, 2001a, p.25). Other causes of false belief include emotions, anxiety, or different attitudes, according to Archer, Egan, and Gendler[10] (Deweese-Boyd, 2006, Revisionist Approaches). After consideration and a reconfiguration some years after the original version, Mele stated that certain conditions must be met for the agent to enter the state of deflationary self-deception. Namely:

1. "The belief that *p*, which *S* acquires, is false
2. *S* treats data relevant, or at least seemingly relevant, to the truth value of *p* in a motivationally biased way.
3. This biased treatment is a nondeviant cause of *S*'s acquiring the belief that *p*
4. The body of data possessed by *S* at the time provides greater warrant for ~*p* than for *p*
5. *S* consciously believes at the time that there is a significant chance that ~*p*
6. *S*'s acquiring the belief that *p* is a product of "reflective, critical reasoning," and *S* is wrong in regarding that reasoning as properly directed." (Mele, 2012, p.51)

Mele added the last two conditions in his 2012 paper *When Are We Self-Deceived?* Given the criticism of his first four conditions when they appeared in 2001 and to differentiate it from wishful thinking.

Self-deception is used in many cases as the self often wants to avoid unpleasant scenarios or justify specific actions contrary to one's better judgment. By self-deceiving themselves, agents can justify instances of procrastination by toning down its negative consequences, overestimating the agent's capabilities or rationalising specific actions. Self-deception is often a panacea for the agent to cope with difficulties and avoid mental breakdowns. However, Mele argues that agents "rarely act with the intention of deceiving themselves" (Mele, 1987, p.123). For Mele, the agent has control over the information

[10]A character approach to akrasia interacts with all these states and emotions and doesn't need to be bound to one cause or theory. Its flexibility can also incorporate other deflationary models.

process, so, at some level, he also accepts the agent's responsibility (albeit on a different level than the intentionalists), to which I agree. From my point of view, self-deception through the conscription by the akratic trait is not an unconscious process but a rational, conscious choice in which the benefits of believing it outweigh his better judgment.

It is not like the agent is saying, 'Today, I will deceive myself'. The akratic trait in which recurrent actions have previously established conscript self-deception as a method to achieve its goals, how agents handle information, or to help with coping through rationalisation or other means. Self-deception helps the agent to evaluate his actions in support of reinterpreting facts, often based on rationalisations and biased explanations, while possessing knowledge that points towards the other direction. When the agent acts recurrently akratically, he does this consciously because he knows he acts against his better judgment. To cope with this, he employs self-deception to provide some relief. Self-deception is often employed as a coping mechanism so the agent can withstand the weight of his akratic actions. This relationship reinforces both akrasia and self-deception, as they are frequently interconnected.

The deflationary account of self-deception can be particularly useful in explaining akrasia and helping us understand how it emanates, at least from a character-based approach. Self-deception can be considered a process that works alongside the akratic trait, reinforcing each other[11]. During akratic episodes, agents often use self-deception to rationalise decisions that contradict their better judgment, suggesting a recurrent (and relatively stable) disposition. The employment of self-deception (among other means) is used to justify akratic actions against the agent's better judgement. Their akratic trait predisposes agents to engage in the process of self-deception as a way to commit akratic actions, reinforcing akrasia (the more agents are self-deceived, the more akratic they become) and the more akratic they become, the more they are self-deceived, which indicates a cyclical two-way dynamic relationship in which akrasia as a trait and self-deception complement and sustain each other as shown in figure 1. Recognising this bidirectional relationship highlights how complex the phenomenon of akrasia is and the means it uses to maintain itself. The causal dynamics between akrasia as a trait and self-deception work in two ways. The akratic disposition of the agent can predispose the agent to employ self-deception more frequently. The akratic trait conscripts self-deception as the preferred strategy to justify (and often rationalise) akratic actions. But it

[11] I won't deal with the mechanisms around the creation and preservation of self-deception as a cognitive process as it is a vast topic of its own accord. However, "its manifestation through different psychological processes of both conscious and unconscious efforts is possible". (Hippel & Trivers, 2011, p.1)

can work the other way around. Repeated self-deception, for instance, rationalising or bias towards crucial information, can become habitual, reinforcing akratic tendencies. If an individual can self-deceive himself, it becomes easier to recurrently act against one's better judgment without facing severe internal resistance. In that sense, self-deception reinforces akrasia's trait. The individual faces little (or none at all) internal turmoil due to his action, which allows for easier use of self-deception. This showcases the causal relationship between the two and the dynamic aspect of the character approach.

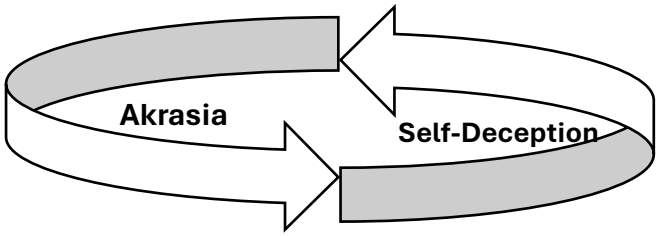

Figure 1. The dynamic bidirectional relationship between akrasia and self-deception.

Imagine the following scenario that highlights Mele's self-deception conditions. Hannah is a junior project manager working for a company and trying to get promoted. For that reason, she consistently and repeatedly takes on challenging projects all by herself despite knowing that she will fail without asking for help (which constitutes her better judgment: "Considering the difficulty of this project, I should ask for help') but proceeds without help, thus acting akratically. She convinces herself that she can finish them alone on time. Based on past experiences, she knows that's impossible (acquired belief is false). Still, she convinces herself that 'this time will be different', therefore treating the known data (past experiences) to her in a motivationally biased way, thus reinforcing her original belief p (which is false) while the evidence shows the opposite. Her desire to get promoted clouds her judgment and she consciously recognises the possibility of ~p (failing to finish the project). However, despite this awareness, she engages in what she thinks is critical and reflective reasoning, concluding that her belief for *p* is solid and, therefore, she can make it.

There are several important takeaways from this example. First, Hannah recurrently acts akratically and fails to adhere to her better judgment due to her akratic character, which isn't a one-time thing. Second, her self-deception is triggered by her desire (to get promoted), but her character precedes the desire. If her ambitions cloud her judgement, it's not because one day she woke up and chose to desire x. She already had the disposition of ambition, greed (usually attributed to virtues and vices under a virtue ethics framework) or anything

expressed through her desire. The previous example highlights the ambition or the greediness of the agent. It shows that her character has intervened in the process, especially if her behaviour was expressed in visible manifestations repeatedly. As argued earlier, habituation is paramount for the virtues and vices of an agent under an Aristotelian framework; therefore, they are acquired, not innate. So, an akratic trait, the more it is exercised and the more it is used through repeated action, the stronger it becomes. Moreover, it is crucial to recognise the effect on practical reasoning (which I will elaborate on in the second part of the thesis), which she initially thinks should be reflective and critical but misjudges what critical or reflective thinking entails and proceeds as usual. All in all, Hannah's example fulfils the required conditions of the deflationary account of self-deception and how it can lead to akrasia as she treats data in a biased, motivated way, a greater warrant for ~p against p, the conscious belief that she might fail and that her reasoning is 'critical'.

As a trait, akrasia is capable of using self-deception as a means to achieve its goals. This ability is part of the symbiotic relationship akrasia shares with self-deception since they operate in similar contexts. The bidirectional relationship between akrasia and self-deception showed that the hypothesis of akrasia as a trait is viable through Mele's deflationary account that strengthens and reinforces recurrent akratic behaviour. At the core, akrasia is a trait that gives orders as the primary legislative power and tells others how they should act. Self-deception is like the executive function, like a trustful agent of the legislative body that will adhere to the decisions of the higher-ups (Rorty,1980b, p.906). Akrasia gives orders to the function which seems more likely to achieve its goals, and then the appropriate faculty executes them. As part of the agent's character, akrasia has an extensive reach and a big arsenal.

At this point, one may argue that not all instances of akrasia are supplemented by self-deception. While in some cases this can be true, it only reinforces the assumption of how complex and multifaceted akrasia is and how traditional approaches may be limited. Even if cases of akrasia without self-deception exist, self-deception remains a significant process for recurrent akratic behaviour that accompanies akrasia as a trait. If the akratic cases become recurrent, seeing how the trait hypothesis accompanied by self-deception can be rejected is challenging.

Akrasia and Self-Control

Self-control is deemed the opposite of akrasia, both linguistically and conceptually. The two complement each other as diametrically opposing concepts. Aristotle contrasts *akrasia* to *enkrateia,* or self-control, highlighting self-regulation's importance in achieving ethical behaviour. Before Aristotle, Xenophon considered *enkrateia* not as a particular virtue but as "the foundation of all virtues" (Xenophon, 371/1923, p.67). This perspective shows self-control's fundamental role in cultivating moral character and living a virtuous life. Incorporating self-control is thus a logical continuation to accompany akrasia. The abilities of self-control can work as a leverage to mitigate akrasia. The long reach of self-control is not limited to akrasia, but it can work against other forms of weakness, such as self-deception. As Davidson (1969) argued, self-deception is a weakness similar to akrasia (p. 28), so it is reasonable to assume that self-control could also work against it.

In this section, I will rely on the two previous arguments, recurrence and self-deception, and complement them with self-control to create a synthesis that supports the central hypothesis of this thesis of reconceptualisation of akrasia as a trait. Employing these three arguments will make clear not just the interrelation between them but also how their interconnection makes the reconceptualisation of akrasia as a trait possible. As shown earlier, recurrent akratic behaviour indicates a recurrent phenomenon and not an isolated instance of weak will that hints at akrasia as a relatively stable and enduring pattern supporting the trait hypothesis. Moreover, self-deception is conscripted by this akratic trait as it facilitates these actions by allowing the agents to rationalise or produce biased, supposedly 'critical and reflective' decisions that contradict their better judgement, further enhancing akrasia into the agent's being. As a result, self-control is critically undermined in this process by these persistent, self-deceptive, and recurrent akratic actions, hindering the individual's ability to align his actions to his better judgement. The weakened state of self-control allows the akratic trait to manifest continuously with minimum resistance and shows how pervasive akrasia's influence is.

But what exactly is self-control? Mele's book *Irrationality* (which I was inspired by) deals with all the aforementioned concepts and treats both akrasia and self-control as character traits. He claims that to have a character trait of self-control is to be self-controlled and exhibit this trait under appropriate circumstances, such as when there is an opposing judgement. Therefore, he defines self-control as "the ability to master motivation that is

contrary to one's better judgment—the ability to prevent such motivation from resulting in behaviour that is contrary to one's decisive better judgment" (Mele, 1987, p.54). However, in Mele's view, this ability is just an approximation, not yet transformed into a trait, as it needs to be complemented by a motivational component. Although I agree with this notion expressed by Mele, I would add that transforming it into a long-lasting trait would require consistent practice and reinforcement of one's values, rational judgments, and understanding of better judgements. These are not inherent since we are all born in a clean slate, *tabula rasa*, but require a combination of methods that promote constant moral development and education starting from a young age, and they also need to align with an agent's values and principles. Internalising self-control as a fundamental principle through its embrace and recognition should come naturally by the agents, not just purely because it was taught to guide their affairs. A genuine and enlightened understanding of self-control would lead to a strong and long-lasting trait that can counter and mitigate the effects of akrasia.

Unsurprisingly, self-control has multiple positive effects on a person. Baumeister et al. argued that "people with high trait self-control were happier than other people. Their global life satisfaction was higher. They had more frequent positive emotions and less frequent negative emotions. They reported less stress and guilt" (Baumeister, 2020, p.27). Moreover, Mele believes that akrasia and self-control have a long reach and influence an agent in a dual sense. Their "influence extends beyond action to practical thinking and the formation and retention of beliefs" (Mele, 1987, p.50). This is important because Mele connects practical thinking and belief formation and sees the implications beyond the action level. It certainly feels possible for akrasia to affect practical reason as it deals with reasoning about 'what to do', and akrasia is primarily involved in disrupting this process. Conversely, self-control tries to position the agent's action to his better judgement, therefore standing opposite of akrasia. So, Mele's assumption about the implications of akrasia for practical reasoning seems plausible, which I will develop further in the second part of this thesis.

Self-control exists in agents to a certain degree, whether as an ability or a trait, although its development is necessary. Mele first pointed at this distinction. The ability to exercise self-control seems more situational or less coherent than self-control as a trait. For instance, I think the difference between a trait and an ability lies in the consistency of manifestation. Having the ability to exercise self-control over a specific instance may or may not be expressed consistently. In comparison, the trait should be relatively consistent throughout all instances.

A trait should be more stable and enduring, consistently influencing an agent's behaviour through various situations. When self-control transforms from an ability to a trait, it becomes a fundamental part of a person's character and can influence and guide their actions. Therefore, I think that a trait implies a deeper integration into an agent's character. To achieve this integration, I suspect that habituation and development can play a crucial role. By regularly exercising self-control, the trait becomes more apparent over time. Although it is difficult to pinpoint the exact place where this happens, it is not impossible that repetition of exercises such as resisting temptation or via other methods could gradually transform the ability into a trait. Still, as Mele mentioned, it needs to be complemented with a motivational component, but motivation is not the only factor. Although motivation is necessary, the cultivation of motivation to achieve pure self-control is insufficient if the agent does not align it with their values and principles, supported by moral education and self-reflection.

The multidimensional reach of self-control ranges from everyday scenarios like adhering to a simple judgement to complex ethical decisions. To exercise self-control, as Mele said, is to master the motivation to act against a better judgment. However, akrasia, which literally means the lack of self-control, intervenes in that process, preventing the agent from acting against his better judgment. In that sense, when an agent recurrently acts against his better judgement, it strengthens the assumption that akrasia may be more than just an isolated instance and highlights its negative relationship with self-control. But that doesn't mean that when an agent is self-controlled, has self-control over all the cases of his life, or it is expressed equally everywhere. Maybe this is what Mele meant when he differentiated between having the ability to be self-controlled and having self-control as a character trait. An agent may be able to restrain himself from overeating, but he may be struggling in other aspects, such as procrastination. If we consider Rorty's previous argument of regional traits, the agent's self-control may also be local.

The role of self-deception in the context of akrasia is also crucial. Agents often deceive themselves when faced with opposing judgements due to their akratic trait overshadowing self-control through self-deception. Self-deception is not only conscripted by akrasia to achieve its goals through rationalising or biased handling of information but also allows for the recurrence of akratic actions. The more frequently the agent is self-deceived, the more he is to act akratic, thus causing recurrence. As argued before, the relationship between akrasia and self-deception is bidirectional, a relationship in which one complements and reinforces the other, which allows for the recurrence of the action time and time again. It

is not a stretch to argue this when that notion of akrasia is expressed through various examples of biased handling of information of the agents. How many times did people know that eating this extra piece of cake has negative health consequences for raising blood sugar levels but choose to downplay the effects by rationalising it or by biased handling of information?

Could self-control have prevented this from happening? The answer is more complex than a yes or no since self-control has many branches and many ways to self-control. In *Backsliding*, Mele distinguishes between *orthodox* and *unorthodox* exercises of self-control (2012, p. 4) that, in short, show that there are instances of the agent exercising self-control in the service of a decision that conflicts with his better judgement. This is an instance of unorthodox self-control, which authors like Henden find impossible because, for him, self-control abides by "what one takes oneself to have most reason to do" (Henden, 2008, p.73). Kennett mentions other forms of self-control, like orthonomous self-control, which she argues is much like the commonsense notion of self-control (Kennet, 2001, p. 134). So, the discussion about self-control (even more in conjunction with self-deception) is not so straightforward. But to simplify things, although there are multiple versions of self-control with many layers and nuances, the common notion of self-control is generally understood as a positive characteristic/ability agents have, so coming back to the question, even if the agent tends to self-deceive themselves, self-control could have stood as a counter to it. It could have aligned the agent with his better judgment and prevented him from acting akratically. The reach of self-control is such that even if self-deception has been established as a habit, self-control can possibly break it. This view is reinforced by Baumeister et al. when they claim that "people scoring high on self-control... seem to operate by managing habits. They form good habits and break bad ones" (Baumeister, 2020, p.27). If self-control can break over established 'bad' habits even if the agents rely upon self-deception, then its reach can be paramount against self-deception and akrasia as a whole.

While an agent is usually comprised of opposing traits, feelings or thoughts which generally exist side by side, one usually overtakes the other at the time of action. When an agent acts akratically, that means a deficiency in self-control. The two opposites can't coexist together at the same time. One can't be akratic and self-controlled at the time of action since there is tension between the two. When the agent acts akratically, his self-control is not working at that moment. So, if an agent acts against his better judgment at that particular moment, he is not exercising self-control. However, it is essential to mention here that I consider it plausible one can be self-controlled in one sense in controlling his thoughts, for

instance, while being akratic in another to fail to control his actions or vice versa. Maybe that can happen when self-control is only partially developed without the proper or imperfect methods. I think of self-control as being multi-faceted and multi-layered. For instance, an agent could be able to control his negative thoughts and stay positive, which could show that his self-control regarding cognitive functions may be more effective. In contrast, the same agent could struggle to manifest his self-control in controlling his temptations or procrastination, leading to a less developed behavioural self-control. So, self-control may not be as simple and straightforward as characterising everything under 'self-control', but it is more complex with many different layers.

In cases when the agent is confused and doesn't know how to settle the question of what to do in a given scenario, one is either akratic or self-controlled. Therefore, when there is an akratic action, self-control is deficient. As a result, recurrent akratic actions reinforce this deficiency. If self-control is a limited resource that tends to run out over frequent use, as Baumeister et al. argued in their 1998 paper *Ego Depletion: Is the active self a limited resource?* [12] A similar hypothesis about self-control depletion from an opposing trait doesn't seem far-fetched. Humans are made through numerous limited resources that get depleted and replenished repeatedly. Resources such as energy get depleted through various mental and physical activities and replenished by resting and eating; hydration goes through a process of depletion from sweating to replenishment through drinking or memory that depletes over time due to age, and through sleep and exercises, we can replenish it. So, the assumption that self-control is a limited resource is not a farfetched scenario.

Imagine the following example. An agent named Tommy is committed to losing weight, and he set a goal to follow a strict diet, which requires him to avoid high-calorie snacks like chips and cookies, which are his guilty pleasures. He knows that avoiding these constitutes his better judgment for losing weight. On a Friday night, he is invited to a birthday party in which all kinds of delicacies lie on the table, including his favourite snacks. Despite his better judgement and knowledge of his goals, he consciously decides to indulge this once as he rationalises by saying it's just once since it is a special occasion or by treating information in a biased way. This initial indulgence, although it may seem harmless, sets the stage. First, he rationalises and succumbs to the temptation, acting against his better judgement because it is a 'special occasion'. However, Tommy doesn't stop there. Over the

[12] Being tired (both physically and mentally) affects the depletion of self-control, known as ego depletion, and it is well documented in the psychology literature. Baumeister, Muraven and Vohs are notable names in this field, with many others confirming this hypothesis.

following weeks, similar occasions arise more frequently. Company meetings, reunion parties, baby showers. Each time, Tommy is faced with a choice. Having rationalised it once, he doesn't find it hard to succumb again to the temptation and justify his current and future lapses. With each occasion that arises, Tommy relies on self-deception to downplay the consequences of his actions or treat known data in a biased way. He convinces himself that he can compensate for snacking by exercising later or skipping meals (which he never does anyway).

How does all this relate to self-control? In every weak instance, Tommy reduces his self-control. The more frequently he succumbs to his temptations, the more depleted his self-control resource becomes. It's not only in the abstract sense that his self-control diminishes but also in deeper psychological fatigue that affects his decision-making capabilities (for a complete view, one should also consider the physiological dependency these snacks have on our being such as the long-term effects of sugar, but this is something for another discipline to explore). If self-control is a limited resource and works like a muscle, it gets tired after use when Tommy indulges in snack eating. Tommy still tries to rely on self-control, even unsuccessfully, in instances of indulgence. When the agent must constantly decide between two actions requiring mental effort, his supply diminishes even if he doesn't strictly exercise self-control. The tension that arises from this constant battle between akrasia and self-control (which ultimately favours akrasia) takes a toll on the psychological resources of the agent that self-control relies upon. The conflict between the two affects the agent, exhausting the agent whether he succumbs to the temptations or not. If self-control behaves like a muscle, then overextension is not necessary to tire it.

Moreover, I suspect the mental processes preceding a decision affect an agent's supply reserve. Even at the pre-conflict level, when the agent thinks about the future conflict, it affects his mental energy. This process may contribute to mental exhaustion if he constantly thinks about his reasons against his self-deceptive rationalisations. Consider, for example, how often people stress prior to a decision—thinking about potential consequences or any other imaginable scenarios in their head and the effect they have on their increased anxiety and general cognitive level. This is sometimes more stressful than the actual event. The pre-decision stage's mental toll on the agent is often more significant than the product. These kinds of internal deliberations affect the agent and his cognitive resources. Finally, one should also consider the cumulative effect these recurrent akratic actions have on a limited resource such as self-control. It's not just the conflict, decisional states, or unsuccessful tries to engage in self-control. The cumulative failures to abide by his better judgement could be a

significant factor in the depletion. Suppose Tommy, from the previous example, recurrently faces these decisions, experiences these internal conflicts, and keeps failing. In that case, this is a general decrease in his ability and self-confidence to manage and regulate his actions over time. Suppose the agent is demotivated, insecure or lacking self-confidence before acting. In that case, it will significantly decrease his chances of exercising self-control. Self-confidence is an essential variable in this equation as it increases the agent's chances to achieve his desired outcome through positive reinforcement or visualisations. If self-control behaves like a muscle and is overtrained or undertrained, it will fail progressively over time (maybe Aristotle was right after all, and a golden mean is the solution). Similarly, when the agent fails to engage in practices that enhance self-control cumulatively, the overall resource of this ability will be hindered.

All in all, the relationship between akrasia, self-deception and self-control is undoubtedly complex. If self-control is seen as a limited resource that can be depleted and undermined through various acts, not just through acts of resistance but through the relationship between recurrent akrasia and self-deception, then it expands the understanding and the hypothesis of akrasia being a trait. The recurrent akratic behaviour indicative of a character trait undermines self-control and slowly depletes it. Encountering temptations in the presence of the akratic trait accompanied by self-deception affects the agent's self-control and self-confidence capabilities. Self-controlled decisions don't occur in isolation but are part of a broader schema that shapes and is shaped by one's (akratic) character. The reconceptualisation of akrasia as a trait is justified through persistent self-control undermining. The explanation of the hindrance of an agent's self-control wouldn't have been possible if akrasia was an isolated instance of a weak will since the agent's self-control would have been able to 'recover'. But if akrasia is a trait characterised by constantly pressuring self-control, it makes sense for self-control to be undermined and the agent to be unable to exercise self-control. The agent's chronic inability to align himself with his better course of action, which, as showcased, happens due to the akratic trait interfering with the self-control trait via self-deception, reinforces my hypothesis. Therefore, understanding the akratic trait in conjunction with the aforementioned concepts not only helps us to imagine akrasia as a trait but also makes us think of strategies to enhance self-control and mitigate its effects of akrasia.

V. Implications

Implications for Davidson's Theory

Davidson's essay greatly influenced the debate on akrasia. Generally, it is considered one of the starting points for akrasia in modern approaches. Here, I would like to address a specific aspect of his theory that he seems to embrace throughout the text: the irrationality he assigns to the akratic action. Davidson starts his paper by identifying a weak will as an action against the agent's better judgment. As argued earlier, my reconceptualisation builds upon the hypothesis that akrasia can be a character trait, relying on several arguments, one of which is self-deception. He seems to have considered akrasia being a trait but ultimately rejected it. He also argued that self-deception is similar to incontinent action (Davidson, 1969, p.28). He concludes that the agent acts and judges irrationally since these go against an agent's better judgment (Davidson, 1969, p.41); apart from the observations above, Davidson seems to believe that in instances of akrasia, the agent is irrational at two levels: judging and acting. While he investigates the concepts of rationality and practical reasoning and establishes that the agent is irrational at two levels, Davidson's explanations are unclear since he argues that the agent acts against a *principle of continence,* judged best based on all available relevant reasons.

In contrast, his account of practical reasoning claims that unconditional judgements are detached from premises.[13] While I agree that akrasia can affect the agent's capacities for judging and acting, I disagree with this notion of irrationality that Davidson *always* applies to the agent. Isn't it possible that agents can act intentionally, akratically and rationally? It might seem counterintuitive, but agents can rationally justify akratic behaviour through self-deception and rationalisations against their better judgement. Throughout my essay, I pointed out how the akratic trait is characterised by persistent akratic behaviour supported by self-deception. I showed that agents can act intentionally, akratically, and rationally as part of the interplay between self-deception and the recurrence of the trait. The akratic trait, through the employment of self-deception, can provide rationalisation to the agent as he is about to act, meaning that there are occasions when the agent might act provisionally rationally against his reasons. Rather than an interaction between rationality and irrationality, there can be

[13] Davidson has been criticised for whether his unconditional judgments are intentions (e.g., Mele), whether one can act against them (e.g., Snellen), and their relationship to conditional judgments.

instances of competing rationalities. Thus, Davidson's account of akrasia and irrationality should be reconsidered.

Assuming the reconceptualisation of akrasia as a trait is feasible, what might the implications be to Davidson's theory? Could viewing akrasia as a trait involving self-deception challenge Davidson's assumption that akrasia is inherently and *always* irrational? I am not arguing that akrasia is *always* rational. There can be many cases where akrasia is irrational, but for a complete and nuanced account of akrasia, we should also accommodate for akrasia being potentially rational. If my hypothesis is plausible, agents often rely on recurrent rationalisations and biased, uncritical processing of information, creating a basis for akratic behaviour, which means that akrasia may not always be irrational or the means by which akrasia is achieved isn't irrational either. This indicates that akratic behaviour may not be simply a failure of rationality and the triumph of irrationality in a dualistic conflict but a more complex interplay between competing rationalities.

I identify three implications towards Davidson's theory. First, Davidson's framework that akrasia is always irrational would need to be re-evaluated to account for the complexity of reasoning and the possibility of akrasia being supported by internally rational processes instigated by self-deception. Davidson thought of akrasia as irrational due to the agent's failure to align his action with his evaluative judgement, as expressed in his P2, which states that "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". (Davidson, 1969, p.23). He relies on the assumption that agents always act on their rational/better judgment, and every other instance of non-alignment to this is considered irrational based on the ground that best judgment equals rational. However, the reconceptualisation of akrasia as a character trait characterised by self-deception can muddy the waters between Davidson's distinction and challenge his assumption. This nuanced understanding of akrasia complicates the relatively straightforward classification of akrasia as merely irrational that Davidson implies by showing that akratic actions can sometimes be potentially rational either through self-deception, which agents rely upon to rationalise their actions or through their biased handling of information.

Although Davidson characterises akrasia as irrational as being against a person's better judgment, we should not forget how cunning reason can be. Audi reinforces my assumption when he claims, "an intentional action can be explained as performed for a reason or that an intentional action is explainable by a reason" (Audi, 1990, p.274). Audi entertains the same thought and further investigates and debunks arguments that 'prove' the irrationality of akratic action. Moreover, McIntyre and Arpaly examined similar assumptions. McIntyre

(1990) argues that “in certain situations, incontinence may be more rational than continence and at the very least is no less rational than continence” (p.380). Arpaly (2000) argued that “acting against one’s best judgment can sometimes be rational” (p.491). Finally, Buss (1997) seems to imply that there might be instances of rational weakened will actions when she claims that “the weak-willed rationaliser contrives a reason for satisfying her desire to defy her reason” (p.36). All these examples show how the idea that someone can act on a reason against another reason is not entirely new or groundbreaking and has already been thought and analysed. Therefore, agents can go against their better judgment (for a reason) to challenge Davidson’s assumption about the inherent irrationality of akrasia. A sufficient reason to act akratically can still be rational. Reason, after all, can be cunning in its ways, goals and understanding (Hegel, 2001, p.47).

The reconceptualisation of akrasia as a trait involving self-deception muddy the waters of the Davidsonian distinction of what constitutes rational and irrational, to which I have argued a re-evaluation is necessary. There is a need to recognise the complexity of reasoning, where akratic actions and judgements are not just irrational failures but can often be supported by rational internal processes. My suggestion challenges this traditional dichotomy and obliges us to rethink the relationship between akrasia, practical reason and rationality, which, even if wrong, can bring a newfound appreciation for the topic.

This brings me to the second implication relating to Davidson’s P2. Davidson’s P2 is undermined in another way. As said earlier, his P2 states, "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” (Davidson, 1969, p.23). My hypothesis of reconceptualising akrasia as a trait complicates judgment in another way: I argue that it negatively affects the agent’s deliberation procedure, *eroding* the agent’s capability of judging clearly, thus weakening P2. When an agent judges, he goes through a process of deliberation through a process of (practical) reasoning to reach Davidson’s conclusion, as expressed in Davidson’s first part of p2. Davidson’s P2 relies on the assumption that the agent a priori judging x being better than y and, he argues that “practical reasoning does, however, often arrive at unconditional judgements that one action is better than another” (Davidson, 1969, p.39). Davidson’s explanations regarding unconditional judgments and how they motivate a person’s action remain vague and sometimes confusing. How does practical reasoning reach that conclusion? Does it come before or after the conclusion of a conditional judgment? I question Davidson’s explanation and usage of practical reasoning as my account of akrasia as a trait affects the agent’s capabilities of judging clearly.

Specifically, I see this erosion happening because akrasia is a trait characterised by its recurrent nature. Habitual patterns of self-deception and rationalisations that reinforce and justify recurrent akratic behaviour can erode the agent's capacity for effective and critical deliberation over time. As a trait, akrasia suggests that agents do not act akratically as isolated instances of weak will but as a continuous recurrent process. If it is a trait, then the recurrent existence of it should affect deliberation at some level. It is logical to assume there would be repercussions on the faculty of reason that helps us judge. Eventually, their ability to engage in critical and thorough reflective deliberation would weaken. It's not an exaggeration to talk about a diminished reflective capacity when the agent recurrently acts acratically. Erosion is a slow process that happens over time, and the more the agent acts akratically, the more his judging ability is compromised. Therefore, agents should address the underlying rationalisations and habitual patterns that drive these actions to counter the erosion. Methods such as enhancing self-awareness to critical reflection of self-deceptive narratives and self-control could be a step towards battling akrasia.

In that sense, Davidson's P2 is directly affected and undermined because the erosion of the deliberation procedure shows that the deliberative capacity which ends up in judging is compromised. How can the agent judge that it would be better to do x than y if the judging process is itself impaired? Which brings me to the third implication. I think the link that Davidson assumes in P2 between judgement, action and desire is hindered if my hypothesis is true because the agent wouldn't be able first to judge about x critically or even if she did, how would her evaluative judgement (that x is better than y) translate into a desire for action. That means that the erosion of the deliberation process affects the applicability and linkage between these actions of Davidson's P2. If true, there is no certain way that the agent can a) judge clearly, and b) even if she does, it isn't sure that this can directly translate a desire to action. Therefore, Davidson's P2 is compromised, weakening the link between judgment, desire and action. If akratic actions are often supported by self-deceptive and rationalised processes that affect the faculty of reason, then intervention strategies should go beyond the usual reinforcement of people's better judgments and holding on to that.

Implications for Mele's Theory

Mele's theory is more accommodating to akrasia as a trait than Davidson's. While Mele recognises akrasia as a trait, he does not focus too much on it. Mele wasn't happy with Davidson's solution either and criticised Davidson on the grounds that a better judgment

doesn't automatically translate to action. Mele rejects Davidson's P2 and explains akratic action because better judgment doesn't necessarily translate to motivation and, therefore, to action (Mele, 1992, p.49). He also thinks the relationship between intentions and motivations is more complex (Mele, 1983, p.366). He set out to explain how a causal theory of action is possible and dwells upon motivational balance and self-control as self-control can, to a certain extent, influence and counter our competing motivations.

While Mele's theory is more accommodating, it still has room for improvement. I think my account has two important implications for Mele's framework. First, self-control alone is insufficient to achieve motivational balance; deeper character development techniques should be addressed simultaneously, as the disconnection between judgement and action may stem from deeper ingrained patterns in a person's character. When it comes to internal conflicts between an agent and his better judgement, it might be inadequate to address deeper issues without accompanying them with more sophisticated methods, especially when they stem from character, which causes a persistent disconnection between judgment and action.

Self-control is undeniably a first layer of defence against short-term goals overriding long-term ones. Still, it remains a surface-level solution against most complex scenarios, such as akrasia as a trait. If akrasia reconceptualised as a trait manages to penetrate this first layer of self-controlled armour, then it can cause a recurrent disconnection between judgment and action. For example, a student might resolve to self-control techniques to finish his thesis on time; however, if he habitually rationalises procrastination through self-deception or other means, he may repeatedly fail to align his better judgement with action. That would show that the issue does not lie in the immediate failure of self-control or the inadequacy of self-control in itself but in a deeper character trait that can overcome self-control. As argued earlier, I understand the akratic trait not just as a trait lacking self-control but as a complex, multifaceted causal disposition capable of affecting a person's thoughts, emotions and actions. Of course, akrasia and self-control are related, but the akratic trait complex as it is goes beyond just a trait lacking self-control; that's why it can also accommodate self-deception and other processes. Its richness and multipolarity make it so vast and confusing in the first place. It wouldn't be a puzzle otherwise. Suppose we understand akrasia as a multipolar and multifaceted trait with a long reach. In that case, it justifies the hypothesis of this thesis, and it doesn't exhaust the possibilities and any new potential findings and understanding that might arise from future (philosophical) investigations.

That's why I think a character approach to akrasia with character and development of appropriate methods and empowering qualities such as self-awareness, enhancement of moral responsibility or some form of resilience building (both mental and physical) alongside self-control can contribute to the overall capacity of the agent following through his better judgements. Individuals become better equipped to recognise and respond to temptations and other factors by focusing on character development methods which support self-control. Perseverance in the face of desires or temptations can only be achieved through a combination of methods and not just through strict self-control. So, while Mele's contribution and reliance on self-control are essential, I think they are insufficient when dealing with akrasia as a trait. A more holistic approach is necessary to provide a more robust solution to the problem of akrasia.

The second implication concerns Mele's explanation of akrasia as a case in which the strength of motivation does not correspond to the strength of evaluation, resulting in why the akratic agent didn't exercise self-control while he could. Mele claims that an agent can have a 'slightly better reason' and choose not to exercise self-control without being a contradiction. Watson and Tenenbaum have criticised Mele's view that this "merely pushes the problem one step further" (Tenenbaum, 1999, p.886). Watson (1977) contends that Mele's appeal to self-control doesn't solve the problem because it is not evident why an agent having self-control failed to exercise it (p.324). My reconceptualisation of akrasia as a trait offers a new perspective in this discussion without oversimplifying the phenomenon by simply stating an 'easy way out' by merely blaming a trait. Instead, when akrasia is viewed as a trait, it systematically undermines the alignment between motivation and evaluation. The akratic trait perpetuates a conflict with the self-control trait, which will dominate at the moment of action if the agent is akratēs. So, when the agent cannot exercise self-control, it isn't simply because the agent no longer holds her judgment but because the akratic trait is so strong that self-control can't manifest freely.

This internal conflict suggests that the failure is not just a result of a lack of self-control or discrepancy between motivation and evaluation but due to the overpowering influence of the akratic trait over the trait of self-control. The akratic trait prevents the other traits from acting as they engage in a conflict to dominate each other. Of course, if the akratic trait has been acting recurrently, it would have more leverage than the self-controlled trait and vice versa. If an agent repeatedly practices self-controlled techniques from a young age, his trait of self-control will overpower the akratic one. This implication challenges Mele's theory by showcasing the internal conflict between the two competing traits. Mele emphasises the

role of self-control in overcoming akratic behaviour, but his proposed explanations that the agent either "misjudged the amount of self-control needed to prevent an akratic action or that she overlooked that it was possible" (Mele, 1987, as cited in Tenenbaum, 1999) misses the point. While Mele showcases the interaction between action, motivation and self-control and how it can help overcome akratic behaviour, my assumption challenges Mele's framework by nuancing the relationship between the three by highlighting the interplay between self-control and akrasia.

Moreover, this implication raises questions about the traditional self-control methods (that Watson and Tenenbaum argue, among others) and how they address akrasia. Not many approaches focus on the akratic tendencies as ingrained in a person's character, so mere appeals to self-control might be insufficient. All in all, it can be argued that the conflict between akrasia as a trait and the opposing trait of self-control challenges Mele's framework as it suggests that the inability to exercise self-control may not be solely due to an abandonment of judgment or due to misalignment between evaluation and motivation but rather from the interplay between the two traits rendering the self-controlled trait unable to manifest.

Lastly, I want to address a potential criticism that Mele faced, as expressed by Tenenbaum. He argued that "how an agent that fails to exercise self-control due to akrasia cannot help us understand how akratic action can be free or free from internal compulsion" (Tenenbaum, 1999, p.889). I need to address it because one may confuse my prior argument with some form of internal compulsion. I argue that the agent's actions are still free despite the internal struggle. Freedom of action doesn't necessitate the absence of internal conflict, but it requires that the agent retains her capacity for deliberation and free choice. In cases like akrasia as a trait, the agent can deliberate (even in a compromised state due to the erosion) and realise her better course of action, but her competing motivations influence her. The presence of akrasia doesn't eliminate the agent's ability to choose; it just complicates the understanding of the decision-making process.

The complexity of decision-making is such that many factors, including character traits, affect one's actions. The fact that an akratic trait influences the agent's behaviour (and deliberation process) doesn't mean the action is compelled or unfree. Instead, it shows the nuanced nature of human action and freedom, where free action results from deliberation[14]

[14] One can go one step further and argue how 'free' is the agent's deliberation process from external factors; however, this is a different discussion.

within the context of one's character. Therefore, by framing akrasia as a trait, one can recognise that an agent acts freely within the domain of his character. The agent still preserves his autonomy and moral responsibility. When the akratic trait overpowers the self-control trait, and the agent engages in akrasia, the agent remains capable of reflective self-evaluation and can work to align his better judgment with action.

Implications for Holton's Theory

As we have seen in *Intention and Weakness of Will,* Holton presents a compelling argument on the differences between akrasia and weakness of will and how the latter arises. His main argument revolves around the idea that weakness of will occurs when agents are too ready to reconsider their intentions (Holton, 1999, p.241). According to him, akrasia is a different kind of weakness stemming from a lack of self-control and concerned judgments—violations. In contrast, the weakness of will is concerned with intention violation. Holton's distinction shifts the debate in a different direction, which expands the discourse and challenges the traditionalist view on akrasia. Where do these weaknesses come from, assuming that Holton's distinction is correct and akrasia differs from the weakness of will? What if they derive from a common denominator? Namely, the weakness of character? What would this mean for his theory?

Reconceptualising akrasia as a character trait puts a character approach to akrasia first. It identifies it as, first and foremost, a particular weakness of character. Holton identified the expressions of weakness of will compared to akrasia, but he did not delve deeper into the origins of one or the other. What if both judgment and intentions violations are failures of character? Could then they be more intertwined than Holton's distinction shows? If they both occur recurrently and express a habitual behaviour, then the obvious implication would be that there is more overlap than Holton initially thought. The clear-cut separation that Holton advocates becomes blurry and more interrelated. Suppose both types of violations stem from the same source. In that case, emphasis should be placed on intervention strategies focusing on overall character strengthening alongside self-control rather than specific mechanisms, thereby simultaneously addressing both types of failures. That would make Holton's theory need refinement to account for broader character-based approaches and sufficient counterstrategies.

I begin by assuming that Holton's distinction is correct. Moreover, I think that, like akrasia, weakness of will is not an isolated instance, and people tend to revise their intentions

recurrently rather than just once. Chronic revisioning of intentions would indicate that weakness of will, although having a different effect, is ruled by the same principles. The intention violation agents undergo can be argued to be a similar weakness with deeper origins and implications. Not all weaknesses derive from character; there can be other causes, but there are too many similarities between the weakness of will and akrasia to disregard the assumption. To be clear, by similarities, I do not mean the content of both are the same, but the form and the pattern these two weaknesses take upon manifestation. They differ in the subject matter they are dealing with, namely an intention against a judgment, but they are still weaknesses as agents fail to abide by their means. In that sense, my interest lies in their origins. They both show similarities in their expression. To me, these weaknesses seem paradoxically similar, and I think a character approach can accommodate both.

In short, Holton's argument revolves around the argument that the outcome of the weakness of will would be "over readily revising an intention when it is in the agent's power to desist from this revision" (Holton, 1999, p.262). While Holton seems to explain in depth his arguments around the distinction between the weakness of will and akrasia, he does not explain the origins of either. Perhaps that was not his aim, one might argue. However, a character approach to akrasia can even accommodate weakness of will. When agents over-readily revise their intentions while having adequate power, their weakness is far from accidental, especially if that happens recurrently. Agents who have the power to resist but, for some reason, do not show that the cause runs deeper. Intention and judgment violations show that the agent is weak, albeit in a different sense. However, these weaknesses are more likely to appear when the weakness runs deeper. I believe these weaknesses arise because of the agent's character, and the more conditioned the agent is towards relapsing rather than improving, the more the weaknesses perpetuate the problem. The desire to revise your intention or go against your better judgment shows a more integrated (in the sense of their origin) understanding of these phenomena.

Weakness of will, defined as Holton's, is the readiness to revise our intentions, which can also be viewed as a trait of character characterised by the weakness to maintain one's intentions. Holton's emphasis on the over-readiness of revising an intention can be understood as a manifestation of deeper character traits such as lack of commitment or due to the akratic trait. Being weak-willed, expressed through character, would explain why people fail to do either. An agent who frequently abandons their intentions is more likely to have some form of ingrained habits that reflect his character. Through this view, both akrasia and weakness of will are not merely situational lapses (although they can be affected by them) but

are more deeply rooted in an individual's character. That shows that both weaknesses derive from a common source that Holton may have missed and allows us to develop interventions and counterstrategies that address each underlying character trait more effectively.

All in all, the primary implication of Holton's theory is the need to reconsider the strict distinction between akrasia and weakness of will in terms of their origin and the potential claim that they stem from character, thus recognising them as more interrelated than previously thought. While there may be a distinction between intention violation and judgment violation, if one goes one step further, maybe both weaknesses derive from a deeper character flaw. My reconceptualisation of akrasia as a character trait calls for a shift in focus and understanding from isolated instances of the weak will, looking beyond these and focusing on more enduring character traits that drive these behaviours. So, although I welcome Holton's distinction and contribution to the debate, I cannot but notice the potential common factor in these weaknesses.

VI. Conclusion

This essay argued for reconceptualising akrasia as a character trait. The paradox of akrasia began with Socrates, who rejected it as ultimately a problem of knowledge. Following Socrates, Aristotle wrote extensively on the topic. He argued for the existence of akrasia contrary to the Socratic rejection and believed that akrasia occurs because individuals are overpowered by their irrational desires. Aristotle's works are the foundations upon which most contemporary work has been based, albeit taking a different turn afterwards.

One of the most influential accounts of akrasia started with Davidson's approach as a logical puzzle. While Davidson's towering influence on the topic is undisputable, different arguments exist. His argument relied on the ingenious distinction between two different kinds of judgments (conditional and unconditional) in which the agent ultimately acts against his conditional judgment. Following Davidson, Mele was unhappy with Davidson's solution, and his equally clever solution was that he denied the link between better motivation (what one wants) and evaluation (what one judges to be good). Finally, Holton turned the debate on its head by arguing that it is not about judgment violation but intention violation.

My hypothesis relied on the idea that akrasia can be rethought as a trait. Many have hinted at the possibility of akrasia being a trait, starting from Aristotle and then with modern scholars such as Hill, Mele and Snellen. Building upon these, the reconceptualisation of akrasia became a viable alternative, as argued in this essay. I argued that akrasia could be

reconceptualised as a character trait based on three main arguments. First, I argued how the recurrence of akratic behaviour indicates that akrasia is not an isolated instance of weak will that many traditionalist views have proposed but occurs recurrently. Second, self-deception often facilitates this recurrent akratic behaviour. The two form a bidirectional relationship that reinforces both. Then, the agent uses self-deception to rationalise his justifications to act against his better judgment through biased information handling or uncritical reflections. Finally, the two preceding arguments combined undermine self-control, hindering the ability of the agent to align their actions to their better judgments, thus rendering the ability to self-control vulnerable to the akratic trait. The view I proposed in this paper aimed at a more nuanced view of akrasia, which aimed to expand our philosophical understanding and deepen our view of this fundamental all too human condition. Not only did I try to argue for an often-neglected alternative, but I also tried to answer a topic relevant to philosophers and non-philosophers alike.

The second part of my paper examined the implications of my thesis for current existing theories, namely those of Davidson, Mele, and Holton. All three theories build upon each other but take a different approach to the problem of akrasia or weakness of will. I argued that if my hypothesis is correct, it would challenge Davidson's characterisation of akrasia as purely irrational since there can be instances that akrasia could be potentially rational. I did not claim that akrasia is *always* rational, but cases of rationality should be accommodated. The second implication argued that Davidson's P2 is undermined by the akratic trait, which erodes the agent's deliberation process, thus weakening their ability to judge, leading to the third implication, which complicates the linkage between judging, desiring and acting.

My hypothesis directly challenged Mele's assumption that self-control alone can manage akrasia. I argued that self-control alone could not counter the akratic trait if not complemented by strategies targeting the character. Moreover, I contested Mele's explanation of why agents fail to exercise self-control. My theory suggested the akratic trait conflicts with the trait of self-control, resulting in a struggle in which the latter is prevented from manifesting due to the former's chronic conditioning, resulting in the suppression of self-control.

Finally, although Holton's view presents a fresh view on the debate and challenges the traditionalist view on akrasia, his assumptions do not remain unchallenged. I questioned his distinction regarding the origins of akrasia and the weakness of will. I argued that both akrasia and weakness of will stem from a person's character. If they both derive or are

influenced by a weak character, they might be more interconnected than Holton initially believed.

The paradox of akrasia will remain perplexing even after this hypothesis; however, this essay tried to fill the gap between akrasia and character and expand our philosophical understanding by highlighting the importance of recognising akrasia as a trait, mainly from a philosophical perspective but borrowing elements from moral psychology and philosophy of action. The intersection of these disciplines provides a fruitful synthetic ground for further exploration of the topic. It remains to be seen how empirical evidence could support my thesis, expand it, and accommodate for potential limitations.

References

Adams, R. M. (2006). *A theory of virtue excellence in being for the good*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199207510.001.0001

Allen, R. E. (1996). *Plato: Ion, Hippias Minor, Laches, Protagoras*. Yale University Press.

Anscombe, G. E. M. (2000). *Intention*. Harvard University Press.

Aristotle. (1926). *Nicomachean Ethics* (H. Rackham, Trans.; Ser. Loeb Classical Library). Harvard University Press.

Arpaly, N. (2000). On Acting Rationally Against One's Best Judgment. *Ethics*, *110*(3), 488–513. https://doi.org/10.1086/233321

Audi, R. (1990). Weakness of Will and Rational Action. *Australasian Journal of Philosophy*, *68*(3), 270–281. https://doi.org/10.1080/00048409012344301

Brandt, R. B. (1970). Traits of Character: A Conceptual Analysis. *American Philosophical Quarterly*, *7*(1), 23–37. http://www.jstor.org/stable/20009331

Bratman, M. (1979). Practical Reasoning and Weakness of the Will. *Noûs*, *13*(2), 153–171. https://doi.org/10.2307/2214395

Buss, S. (1997). Weakness of Will. *Pacific Philosophical Quarterly* 78 (1):13–44.

Davidson, D. (2001). *Essays on Actions and Events: Philosophical Essays Volume 1*. Oxford, GB: Clarendon Press.

Davidson, D. (1969). How Is Weakness of the Will Possible? In Joel Feinberg (ed.), *Moral concepts*. London: Oxford University Press.

Demos, R. (1960). Lying to oneself. *Journal of Philosophy* 57 (18):588-595.

Deweese, B. I. (2023). *Self-Deception*, The Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/archives/fall2023/entries/self-deception/

Doris, J. M. (2002). *Lack of Character: Personality and Moral Behavior*. New York: Cambridge University Press.

Eaton, S. (2007). *Prometheus and Digital Sculpture*. Scott Eaton. Retrieved June 28, 2024, from https://www.scott-eaton.com/2007/prometheus-and-digital-sculpture

Hegel, G. W. F. (2001). *The Philosophy of History* (J. Sibree, Trans.). Batoche Books Kitchener.

Henden, E. (2008). What is Self-Control? *Philosophical Psychology*, *21*(1), 69–90. https://doi.org/10.1080/09515080701874092

Hill, T. E. (1986). Weakness of Will and Character. *Philosophical Topics*, *14*(2), 93–115. http://www.jstor.org/stable/43153982

Holton, R. (1999). Intention and Weakness of Will. *Journal of Philosophy* 96 (5):241.

Homiak, M. (2019). *Moral Character*, The Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/archives/sum2019/entries/moral-character/

Kennett, J. (2003). A Taxonomy of Agent-control, *Agency and Responsibility: A Common-Sense Moral Psychology* https://doi.org/10.1093/acprof:oso/9780199266302.003.0005

Kraut, R. (2022). *Aristotle's Ethics*, The Stanford Encyclopaedia of Philosophy. https://plato.stanford.edu/archives/fall2022/entries/aristotle-ethics/

Kristjánsson, K. (2010). *The Self and its Emotions*. Cambridge: Cambridge University Press.

May, J., & Holton, R. (2010). What in the world is weakness of will? *Philosophical Studies, 157*(3), 341–360. https://doi.org/10.1007/s11098-010-9651-8

McIntyre, A. (1990). Is Akratic Action Always Irrational? In O. Flanagan & A. O. Rorty (Eds.), *Identity, Character, and Morality: Essays in Moral Psychology (*pp. 379-400). MIT Press.

Mele, A. R. (1983). Akrasia, Reasons, and Causes. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *44*(3), 345–368. http://www.jstor.org/stable/4319642

Mele, A. R. (1987). *Irrationality: An Essay on Akrasia, Self-deception, and Self-control*. Oxford: Oxford University Press.

Mele, A. R. (2001a). *Self-deception Unmasked*. Princeton University Press.

Mele, A. (2009). Weakness of will and Akrasia. *Philosophical Studies*, *150*(3), 391–404. https://doi.org/10.1007/s11098-009-9418-2

Mele, A. R. (2014). *Backsliding: Understanding Weakness of Will*. Oxford University Press.

Mele, A. R. (2020). (Ed.). *Surrounding Self-Control*. Oxford University Press.

Mele, A.R. (2001b). Acting Intentionally: Probing Folk Notions, In Bertram F. Malle, Louis J. Moses, Dare A. Baldwin (Eds.), *Intentions and Intentionality: Foundations of Social Cognition.* The MIT Press.

Mumford, S. (1998). *Dispositions*. Oxford, GB: Clarendon Press.

Robinson, R. (2010). Aristotle On Akrasia (Book VII.1–11). In O. Höffe (Ed.), *Aristotle's "Nicomachean Ethics"* (pp 149–166). Brill. https://doi.org/10.1163/ej.9789004177628.i-259.62

Rorty, A. O. (1972). Belief and Self-deception. *Inquiry*, *15*(1–4), 387–410. https://doi.org/10.1080/00201747208601665

Rorty, A. O. (1980a). VI. Akrasia and Conflict. *Inquiry*, *23*(2), 193–212. https://doi.org/10.1080/00201748008601902

Rorty, A. O. (1980b). Self-deception. Akrasia and irrationality. *Social Science Information*, *19*(6), 905-922. https://doi.org/10.1177/053901848001900601

Snellen, P. J. (2018). *Akrasia as a character trait*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen.

Tenenbaum, S. (1999). The Judgment of a Weak Will. *Philosophy and Phenomenological Research*, *59*(4), 875–911. https://doi.org/10.2307/2653561

Von Hippel, W., & Trivers, R. (2011). The evolution and psychology of self-deception. *Behavioral and Brain Sciences*, *34*(1), 1–16. https://doi.org/10.1017/s0140525x10001354

Wallace, R. J. (2020). *Practical Reason*, The Stanford Encyclopaedia of Philosophy. https://plato.stanford.edu/archives/spr2020/entries/practical-reason/

Watson, G. (1977). Skepticism about Weakness of Will. *The Philosophical Review*, *86*(3), 316–339. https://doi.org/10.2307/2183785

Xenophon. (1923). *Memorabilia Oeconomicus Symposium Apology* (E. C. Marchant & O. J. Todd, Trans.). Harvard University Press.

Zeyl, D. J. (1980). Socrates and Hedonism: “Protagoras” 351b-358d. *Phronesis*, *25*(3), 250–269. http://www.jstor.org/stable/4182098