



Universiteit Utrecht

Akrasia in a causal world

Understanding Akrasia within the Causal Theory of Action

Hanneke Mol (3975789)

Thesis BA Philosophy

Tutor: Gerhard Bos

Second Tutor: Annemarie Kalis

Date: 25-6-2016

Words: 7250

Contents

Contents	1
Abstract	2
Introduction	2
1. The problem: understanding Akrasia within a causal theory of action	
<i>1.1 The Causal Theory of Action</i>	3
<i>1.2 Akrasia</i>	4
<i>1.3 Difficulties in reconciliation</i>	5
2. Understanding Akratic action	
<i>2.1 The common CTA model of action</i>	6
<i>2.2 Two ways of wanting</i>	7
<i>2.3 A Four-Component Model of Action</i>	8
3. Understanding the Akratic actor	10
Conclusion	13
Literature	14

Abstract

The phenomenon of Akrasia – free and intentional action against one’s own better judgment – is one of the most difficult phenomena to explain for the Causal Theory of Action. Critics have argued that the Causal Theory of Action is unable to explain why Akratic actions exist. The aim of this paper is to develop an understanding of Akrasia that enables reconciling it with the Causal Theory of Action. First, the common model of action of the Causal Theory of Action is revised in order to enable understanding of Akratic action. Subsequently, the concept of ‘Desire for Continence’ is developed to explain Akrasia from the perspective of the actor. The revised model of action and ‘Desire for Continence’ are then combined to show how the phenomenon of Akrasia can be understood within the Causal Theory of Action.

Introduction

The causal theory of action, also known as the ‘standard theory of action’ is one of the most common ways to explain action in philosophy.¹ The theory holds that action consists in intentional bodily movement that is caused in the right kind of way by mental states of the actor, such as desires and beliefs. Although common and influential in philosophy of action, the Causal Theory of Action has been criticised as well in the past decades, for diverse reasons. Two of the main types of criticism that have been targeted at the Causal Theory of Action are that the theory is too inclusive (i.e. that it labels events as actions that should not be seen as actions), and that the theory is not inclusive enough (i.e. that it cannot explain the existence of certain actions).²

The topic of this paper is a specific point of criticism of the latter type which points to the existence of Akratic actions. Akratic actions are free and intentional actions against one’s own judgment of what would be the best thing to do. According to critics the Causal Theory of Action cannot explain why such actions exist. The purpose of these paper is to develop an understanding of the phenomenon of Akrasia that enables reconciling it with the Causal Theory of Action.

In order to do this, I begin with a more in depth explanation of the assumptions of the causal theory of action, the concept of Akrasia, and the seeming difficulty of reconciling those two. Subsequently, I develop an answer to the question how the phenomenon of Akrasia can be understood within the Causal Theory of Action by answering to two sub-questions: (1) how Akratic actions can be understood within the Causal Theory of Action, and (2) how the Akratic actor can be understood within the Causal Theory of action.

To answer the first question, I argue that the model of action that is usually defended in the Causal Theory of Action – the Three-Component Model of Action – is too limited to enable explanation of Akratic actions. Subsequently, I propose an adapted model of action – the Four-Component Model of Action – and show how this model can explain and interpret Akratic action from the perspective of a Causal Theory of Action.

The second question will be addressed in the third part of this paper, where I develop the concept of ‘Desire for Continence’ in order to enable agency within the Causal Theory of Action and the understanding of Akratic action from the perspective of the actor.

Lastly, by combining the answers that were formulated to the two questions, an answer will be given to the problem of reconciling the Causal Theory of Action with the phenomenon of Akrasia.

1. The problem: explaining Akrasia within a causal theory of action

1.1 *The Causal Theory of Action*

The term ‘Causal Theory of Action (CTA)’ refers to a group of diverse theories that share certain assumptions about what it means to act. The three main assumptions of the CTA will be discussed below.

¹ Nicholas Bunning, *The Blackwell Dictionary of Western Philosophy* (Malden: Blackwell, 2004).

² Jesús H. Aguilar & Andrei A. Buckareff (ed.), *Causing Human Actions* (Cambridge: MIT Press, 2010).

The first shared assumption of the theories that are grouped as the CTA, is that action consists in *bodily movement*. This means that, according to the CTA, action should be understood as an embodied, rather than as a mental phenomenon. Thus, mental states such as thoughts and desires, are not seen as action, whereas embodied states such as walking somewhere are (if the other assumptions of the CTA are met as well). It is important to stress here that 'movement' does not necessarily refer to motion.³ A person can act when they are waiting for the bus or when they are meditating, for example, even though their body is motionless in these situations. Rather, bodily movement should be understood as a positioning of one's body in a specific way. 'Positioning' here refers to the placement of one's body in space and in relation to the physical environment.

Now, positioning of a body is not always seen as action within the CTA. For example, if someone trips over a stone and falls on the ground, their body is positioned in a certain way. Nonetheless, this would not be regarded as action within the CTA. What, then, is it about bodily movement that makes it action? This question is answered by the second assumption of the CTA, which is that action is *intentional*. This means that, for bodily movement to qualify as action, it should be performed by the actor with a certain aim in mind. Thus, if someone blinks, this would normally not be regarded as action, since it is an automatic activity, with no intention of the actor behind it. However, if the actor winks at someone to make that person smile, this is intentional bodily movement. Therefore, the latter would normally be seen as action within the CTA.⁴

According to the CTA, for action to be intentional, it is necessary that it is *caused by mental states of the actor*. This is the third main assumption of the CTA. What this means, is that that which is seen as the reason for an action (for example, having the desire to make another person smile, and the belief that winking will make them smile) is thought to be related to actually performing the act, in a causal way. It is important to note here that the CTA states that bodily movement should be caused by mental states *in the right kind of way* in order to qualify as action. The reason for this addition is to avoid the CTA classifying situations as action that intuitively do not seem to be action. For example, cases of deviant causation, where mental states indirectly cause bodily movement (e.g. when a desire the actor has to drop something leads to nervousness, and the nervousness leads to the dropping the object) should not be seen as action according to the CTA. Thus, the mental states that cause the action should also give the reason for it.

The three main assumptions can be summarised in the following definition of the CTA:

CTA. The causal theory of action is the view that action is intentional bodily movement that is caused in the right kind of way by mental states of the actor.

In the next section the concept of 'Akrasia' will be explained in more depth. This will help to subsequently clarify why Akrasia seems problematic for the CTA.

1.2 Akrasia

The term 'Akrasia' originates from ancient Greek, and literally means 'lack of control'. In ancient Greek philosophy the idea of 'Akrasia' was closely related to the idea of knowledge.⁵ It was thought that a person who acts Akratically, acts against what they *know* to be the right thing to do, because of a lack of self-control. In the modern philosophical debate, 'Akrasia' is usually understood from a slightly different perspective. Here the emphasis is on motivation, and understanding the rationality

³ Michael Smith, "The Standard Story of Action: An Exchange (1)," in *Causing Human Actions*, ed. Jesús H. Aguilar & Andrei A. Buckareff (Cambridge: MIT Press, 2010), 45-55.

⁴ Jesús H. Aguilar & Andrei A. Buckareff (ed.), *Causing Human Actions* (Cambridge: MIT Press, 2010), 2.

⁵ Plato, "Protagoras," in *Protagoras and Meno*, ed. Robert C. Bartlett, (New York: Cornell University Press, 2004), 1-66.; Aristotle, *Ethica Nicomachea*, translated by Christine Pannier & Jean Verhaeghe, 7th edition (Groningen: Historische Uitgeverij, 2009).; Cristhopher Bobonich & Pierre Destrée, *Akrasia in Greek Philosophy* (Leiden: Brill, 2007).

of actions from the point of view of the actor. 'Akrasia' is then understood as free and intentional action against *one's own judgment* of what is best to do.⁶

This understanding has been largely influenced by Donald Davidson, who sparked the modern debate about Akrasia with his essay "How is Weakness of the Will Possible?".⁷ Davidson argued that the idea of 'Akrasia' as being related to knowledge, emphasises the cognitive too much. What this means is that he did not want to restrict his discussion of Akrasia to factual knowledge. Instead he wanted to include all aspects of acting and evaluating actions, including prescriptions and emotional aspects.

To adopt a broader perspective, he focussed on the attitudes and beliefs of the actor instead, describing Akratic action (or 'incontinent action' as he calls it) as acting in disaccord with what one considers to be the better thing to do. Based on this idea, he defined Akrasia as follows:

D. In doing x an agent acts incontinently if and only if: (a) the agent does x intentionally; (b) the agent believes there is an alternative action y open to him; and (c) the agent judges that, all things considered, it would be better to do y than to do x.⁸

To gain a more concrete understanding of what such action could look like, imagine Michael who is a student in secondary school. Michael has to do his finals, and has an important chemistry exam coming up. He wants to do well on the exam, as he needs a good grade to be admitted to the university major that he wants to do. It is only a few days until the exam, and Michael believes that he should really start revising. However, instead of revising he chooses to play videogames all day.

In this situation, Michael acts in a way that goes against his judgment of what would be best for him to do. He believes that revising would be the best thing to do, yet he chooses to play videogames instead. These kind of actions seem to exist, but are rather puzzling at the same time. They are especially difficult to explain within a Causal Theory of Action. Why reconciliation of Akratic actions and the CTA is difficult, will be elaborated on in the next section.

1.3 Difficulties in reconciliation

As discussed above, the CTA assumes that actions are caused by mental states of the actor, such as desires and beliefs. Now, what makes Akrasia problematic for the CTA, is that actors seem to act against their own mental states when they act Akratically. That is, when acting Akratically their actions are not in alignment with their own ideas of what is best to do. If our actions are caused by our mental states, how is a situation possible in which our mental states point towards doing x, yet we do y instead?

Davidson has captured this problem of reconciliation in three principles. The first two principles are assumptions of the CTA, whereas the third principle states that Akratic actions exist.

- 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.⁹

The combination of P1 and P2 suggests that an actor will always do what they consider to be the best thing to do. However, when an actor acts Akratically, they do y, while they consider it better

⁶ Annemarie Kalis, *Failures of Agency. Irrational behavior and self-understanding*, (Lanham: Lexington Books, 2011), 25-26.

⁷ Donald Davidson, "How is Weakness of the Will Possible?" (1970), in *Essays on Actions & Events*, (New York: Oxford University Press, 1980), 21-42.

⁸ *Ibid*, 22.

⁹ Davidson, *How is Weakness of the Will Possible?*, 23.

to do x than to do y. If an actor is caused to act by their mental states, and their mental states seem to support action x, how is it possible that the agent nonetheless chooses to do y?

In the rest of this paper I will address the seeming difficulty of reconciling the CTA and the phenomenon of Akrasia by (1) answering the question how Akratic action can be understood by revising the common model of action of the CTA, and (2) answering the question how the Akratic actor can be understood by exploring the meaning of agency within the CTA in interaction with the ideas of Velleman.

2. Understanding Akratic action

2.1 The common CTA model of action

As explained in 1.1, the CTA assumes that action is caused by mental states of the actor. In order to see how Akratic action may arise within this model, it is important to understand what mental states exactly cause action, and how they do so. Now, within the CTA, it is usually thought that actions are caused by a combination of two different types of mental states: affective mental states and cognitive mental states.¹⁰ Affective states are the emotions of the actor that set the goal of acting. They are directly related to what the actor wants or desires. Cognitive states are rational evaluations of what kind of action will lead to achieving the desired goal.

This distinction between affective and cognitive states is further clarified by Davidson, who argued that whenever a person acts for a reason, they can be characterised as having a pro attitude towards actions of a certain kind and believing that the performed action is of that kind.¹¹ ‘Pro attitude’ should be understood broadly here, as any kind of desiring, valuing or being positive towards a certain action. Believing should be understood broadly as well, as encompassing remembering, knowing, perceiving and noticing something.

The way these affective mental states (or pro attitudes) and cognitive mental states (or beliefs) are connected to action according to the CTA, is depicted in Figure 1.

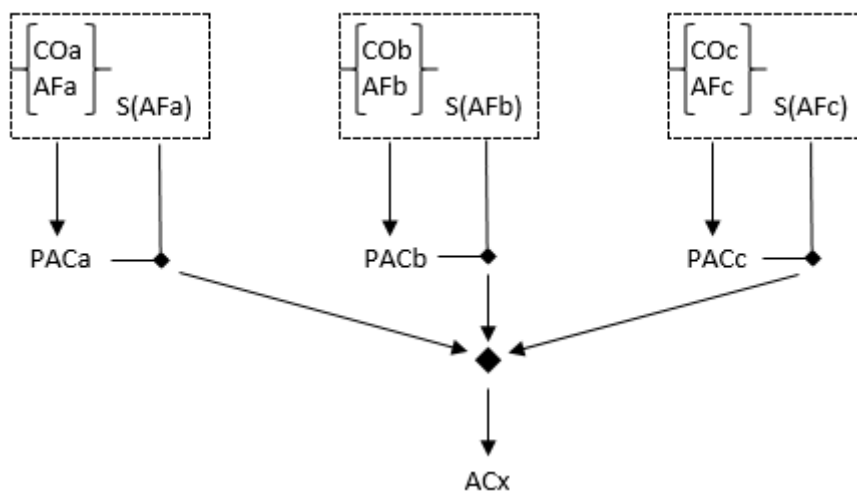


Figure 1. The Three-Component Model of Action.

The idea is that affective mental states (AFx) set the goal for acting, whereas cognitive mental states (COx) set the instruments for achieving that goal.¹² Each AFx-COx pair gives rise to a potential action (PACx) that the actor could perform. That these actions are *potential* means that actors do not necessarily act each time they desire something and believe that a certain action will lead to fulfilling

¹⁰ Berent Enç, “How Causes Can Rationalize: Belief-Desire Explanations of Action,” *Turkish Studies in the History and Philosophy of Science* (2005), 231-247.

¹¹ Donald Davidson, “Actions, Reasons, and Causes” (1963), in *Essays on Actions & Events*, (New York: Oxford University Press, 1980), 3-19.

¹² Michael Smith, *The Standard Story of Action*.

that desire. Instead, they have many different desires and beliefs about how to fulfil these desires, that cannot all be performed and may sometimes be conflicting. Therefore, all the potential actions (PACx's) are weighed first before an action is performed.

The weighing is done (according to the common CTA-model of action) based on the strength of the affective states ($S(AFx)$). Thus, how much the actor wants the different options open to them, or how strong their pro-attitude towards the different possible actions is, guides the weighing of actions. This weighing process does *not* necessarily have to be an explicit process. It can consist of the actor simply choosing one out of possible actions, without consciously thinking about every potential action and the strength of each of their desires. However, what *is* assumed, is that if an actor wants A and believes that *x*-ing will lead to A, but at the same time also wants B and believes that *y*-ing will lead to B, and can only perform either *x* or *y*, they will perform the action that leads to that which they want the most. Thus, if the actor wants A more than B, they will *x* rather than *y* and vice versa. Note that this is similar to the assumption of the CTA that is described in principle P1 of Davidson (see 1.3).

To conclude, three different components, or aspects, can be distinguished in the common CTA model of action: (1) a cognitive mental state (e.g. an instrumental belief), (2) an affective mental state (e.g. a desire), and (3) the strength of that affective mental state. Therefore, I call this model the 'Three-Component Model of Action'.

The question that needs to be answered now, is how the Three-Component Model of Action relates to the phenomenon of Akrasia and to what extent it can explain this phenomenon. In order to answer this question, the next section will investigate the mental states that are involved in Akritic action in more depth. It will do this by addressing a distinction between two ways of wanting made by some philosophers.

2.2 Two ways of wanting

In order to explain the phenomenon of Akrasia, Watson has made a distinction between two ways of wanting: evaluational wanting and motivational wanting.¹³ Wanting to do *x* more than wanting to do *y* in an evaluation sense refers to preferring *x*-ing to *y*-ing or ranking *x*-ing higher than *y*-ing on some scale of values. It concerns attaching more importance to a certain action or valuing it more. Wanting to do *x* more than wanting to do *y* in an motivational sense, on the other hand, refers to being more motivated to actually *x* than to actually *y*. It can be understood as having a stronger pro attitude (e.g. a stronger desire) towards this option than towards the other options. Thus, in the standard CTA model of action it consists in $S(AFx)$.

Mele has used this idea to develop an answer to the question how Akritic action can exist if the CTA is true.¹⁴ He has argued that it is possible to defend both the CTA and the existence of Akrasia, but only if principle P2 of Davidson – which says that if an agent judges that it would be better to do *x* than to do *y*, he wants to do *x* more than he wants to do *y* – is rejected. According to Mele when a person acts akritically, this is because their motivational wanting is in disaccord with their evaluational wanting. Thus, they then judge *x*-ing to be better than *y*-ing, but are more motivated to do *y* than to do *x*, and therefore *y*.

Distinguishing between two ways of wanting is a useful way of understanding Akrasia for the CTA, as it suggests that Akritic action arises because of *mental states* being in disaccord, rather than because of *mental states and action* being in disaccord. More specifically, it suggests that motivational wanting and evaluational wanting are in disaccord during Akritic action, but that Akritic action is still caused by mental states (i.e. the mental state of motivational want). By doing this it allows keeping the assumption of the CTA that action is caused by mental states.

However, a difficulty with the understanding that Watson and Mele have of the two types of

¹³ Gary Watson, "Skepticism about Weakness of Will," *The Philosophical Review* 86 (3) (1977), 316-339.

¹⁴ Alfred R. Mele, "Akritic Action: Causes, Reasons, Intentions", in *Irrationality*, (Oxford: Oxford University Press, 1987), 31-49.

wanting is the way in which they conceive of evaluational wanting. As explained above Watson and Mele describe wanting something more in an evaluational sense as valuing it more. In section 2.1 it was described that pro-attitudes are conceived of broadly in the CTA, as any kind of desiring, valuing or being positive towards a certain action. Thus, on the standard CTA-conception, evaluational wanting as understood by Watson and Mele can be classified as a pro-attitude. Consider the example of Michael again. He may look at his different desires, and decide that he values working on his future by revising more than playing videogames. He therefore wants to revise more than he ways to play videogames in an evaluational sense. This kind of 'valuing' is similar to having pro-attitudes as they are commonly conceived of within the CTA. Valuing working on one's future can be understood as having a stronger pro attitude towards activities that are beneficial for one's future than towards activities that provide short-term entertainment. 'Wanting something more in an evaluational sense' can therefore be seen as the S of an (AFx) state compared to the S's of other (AFx) states. Thus, evaluational wanting, as Watson and Mele describe it, is related to the emotions of the actor and what goals they want to achieve, and has an affective nature.

This understanding of evaluational wanting poses a difficulty for understanding discrepancies between evaluational wanting and motivational wanting *within a CTA perspective*. The reason for this is that within the CTA it is thought that having an affective nature is exactly the defining feature that makes some states motivational, and that distinguishes them from states that are not motivational. From this perspective evaluational states should motivate if their nature is affective. However, if evaluational states motivate, it is no longer clear how the discrepancy between evaluational wanting and motivational wanting, that arises during Akrasia, could exist.

In conclusion, although it is useful to distinguish between two ways of wanting, on the understanding of that Watson and Mele have of evaluational wanting either (1) it is unclear in what way the mental states that evaluational wants are based upon differ from the mental states that motivational wants are based upon, or, if they are based upon the same mental states (2) it is unclear where the difference between evaluational wanting and motivational wanting comes from. What is needed, then, is an understanding of the nature of evaluational wanting that enables it deviating from motivational wanting within the CTA perspective. In the next section such an understanding will be developed.

2.3 A Four-Component Model of Action

Above it has been explained that the CTA distinguishes between two different types of nature of mental states: affective and cognitive. I have argued that understanding the nature of evaluational wanting as affective is problematic. Therefore, to enable explanation of the phenomenon of Akrasia within the CTA, the nature of evaluational wanting should be perceived of as cognitive instead. More specifically, evaluational wanting should be seen as a *cognitive assessment* of what would be the best thing to do.

What is central to this 'cognitive assessment' is the ability to consciously look at oneself and one's own actions and reflect upon them. I call this *reflective ability*. Reflective ability enables actors to take distance from their current and immediate desires and beliefs about how to fulfil them, and to consider things in the larger context of oneself as a person that exists over a longer period of time. In Figure 2 an adapted CTA model of action is depicted in which reflective ability (R) is added as a fourth component. As displayed in the model, evaluational wanting (or the Will, W) is formed based on the weighing of mental states that occurs through reflective ability (R).

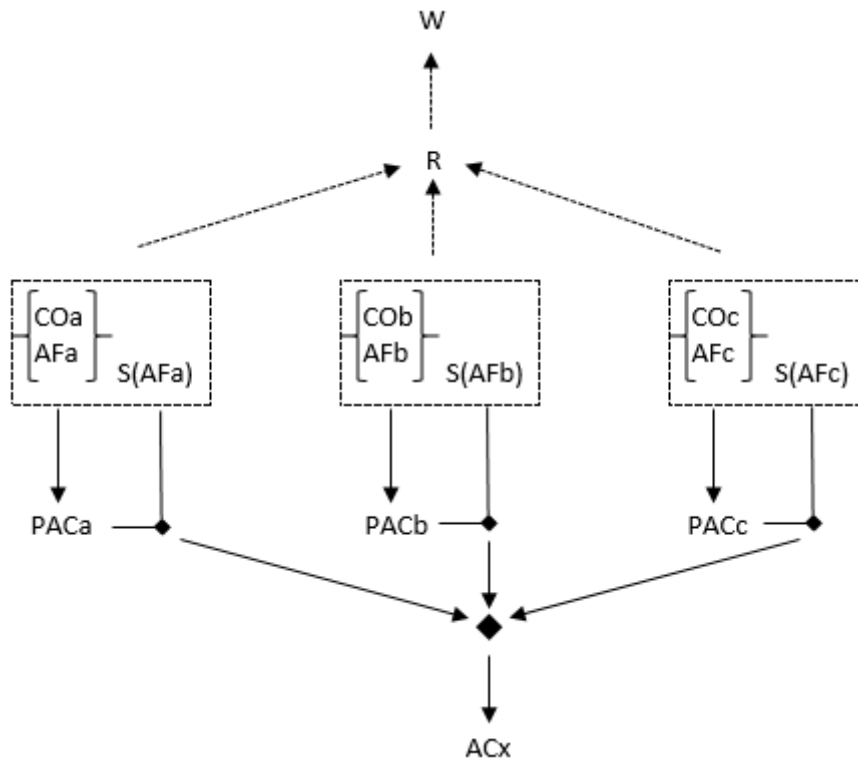


Figure 2. The Four-Component Model of Action

There are many different forms reflection could take, but three main ones that I believe to be important in the context of Akrasia will be discussed below. These are: reflection on oneself as a person, reflection on current mental states, and reflection on mental states through time.

First, reflection on oneself as a person could affect the way in which mental states and actions are weighed. How a person perceives themselves (their self-perception) can influence their interpretation of what mental states and actions should mean to them and how they should weigh them. For example, imagine a person who both desires to eat a cake and to be healthy, and wonders whether or not they should eat the cake. If this person reflects on themselves as a person, and concludes that they are a person that is very considerate about their health, this would support giving a stronger weight to not eating the cake than to eating the cake. Similarly, someone who perceives themselves as a moral and caring person may reflect on their identity and decide that they should give stronger weight to moral actions than to other actions.

The way in which reflection on the self and identity informs the weighing of actions is closely related to the principles that a person holds. More specifically, in order for reflection on the self to affect the Will, the person should not only perceive their self in a certain way but also believe that this is the way their self should be. For example, a person that is very concerned about their health but believes that they should not be, will probably not form the evaluational want (Will) to always perform health-concerned actions. If, on the other hand, they hold a principle that says that they should be concerned about their health (e.g. 'always perform the action that is most beneficial for your health') they are likely to give a stronger weight to health-beneficial actions than to actions that are bad for their health. Thus, reflection on oneself as a person with a certain identity that is formed according to specific principles, can affect the weighing of mental states and actions in favour of those that are in accord with the principles the person holds.

A second form that reflection could take is reflection on current mental states. A person may reflect upon the desires they have and decide that the thing which they have the strongest desire for is not the most rational thing to desire most. For example, imagine that Michael reflects on his

mental states. He has the desire to be entertained and the belief that playing videogames will entertain him, as well as the desire to have a good future and the belief that revising is important for this. He perceives that his desire to play videogames is stronger than his desire to revise. However, he also believes that it is not rational to want to play videogames more. The reason for this comes from his ideas about how much each action leads to a desired outcome. He can see that playing videogames will give him some entertainment, but only for a little while. Thus, he believes that playing videogames only gives him a small amount of his desired outcome of entertainment. Revising, on the other hand is very important for a good future, as his performance on his exam decides whether or not he can do the university major that he wants to do. Thus, he believes that revising has a great impact on his desired outcome of having a good future. Based on these considerations, Michael can weigh his options and conclude that he wants to revise more than he wants to play videogames in an evaluational sense. Thus, a person can reflect on their desired ends and on *how much* of these desired ends different actions (as given by the COx's) will bring about. These reflections can then help them to weigh actions and to decide what would be the most rational thing to do.

Third, a person could reflect on their mental states through time. They could transcend their wants and beliefs about how to fulfil them at the current moment, and consider what the best thing to do is given their mental states throughout time. For example, take a person who is addicted to smoking, but also wants to improve their health. At each separate moment in time that they desire to take one cigarette, acting on this desire will only have a small impact on their health. However, through reflection they may connect all these separate moments and realise that in sum their smoking habit does have a significant impact on their health. This could bring them to decide that not smoking at all these specific moments is the best thing to do, even though these moments on their own do not have a big impact on their health.

These three forms of reflection do not provide an exhaustive list of reflective ability, but they give an impression of what reflective ability can look like. The question that needs to be answered now is how Akrasia can be understood from the CTA perspective using the new model in which reflective ability has been added. I have explained that affective states (AFx's) are the main focus of motivational wanting. This means that what the agent will be most motivated to do, and what they will therefore do, is that which they have the strongest affective state about. Cognitive states only affect motivation by telling which action should be performed to achieve the outcome that is desired most. In the model coming to action is depicted by the little diamond at the bottom. Here the potential action for which $S(AFx)$ is the largest is turned into actual action. For evaluational wanting, on the other hand, cognitive states are the main focus. People can evaluate mental states and actions through their reflective ability and thereby form a cognitive belief about what the best thing to do is. As the nature of evaluational wanting is cognitive, and not affective, evaluational wanting does not motivate the actor, and therefore discrepancies between evaluational wanting (cognitive) and motivational wanting (affective) can arise.

This chapter has given an answer to the question how Akratic action can be understood within the CTA. The next chapter will focus on the question how Akratic action can be understood from the perspective from the actor. It will also give an answer to the question why people sometimes act akratically, whereas at other times people do act in accordance with their Will (evaluational wanting).

3. Understanding the Akratic actor

The Four-Component Model of Action enables the possibility of action that diverges from evaluation. However, a point of criticism could be that the actor, or the person that acts, seems to be completely disconnected from their actions in this model. That is, there is a person that has a Will and ideas about what would be best to do, but this does not seem to be connected to the way in which their

body behaves.

A point that is similar to this has been targeted at the CTA by Velleman.¹⁵ He has argued that in the CTA as it is usually advocated, the agent disappears. This reason for this is that the common CTA model only takes mental states and the actions they cause into consideration. Because of this, the actor only has a passive role in which something happens to them, but no role in influencing the course of events. As he words it: "(...) my objection is that the occurrences it [the CTA] mentions in the agent are no more than occurrences in him, because their involvement in an action does not add up to the agent's being involved."¹⁶ However, contrary to this place that the CTA gives to the agent, Velleman argues, "(...) in our conception of ourselves, at least, if not in reality – is our perceived capacity to interpose ourselves into the course of event in such a way that the behavioural outcome is traceable directly to us."¹⁷ How then, could such a conception of agency, be reconciled with a causal world?

The idea that Velleman develops to solve this problem, is that the possibility for agency to exist, arises when a person has a desire to act in accordance with reasons. Such a desire enables them to act in the way which they believe to have most reason to, and can at the same time fully be reconciled with a causal world view and the CTA, as the action of the person is *caused* by their desire to act in a rational way. This seems to me a useful perspective for understanding the meaning of agency in the CTA. In what follows I will address the question of how Akrasia can be understood from the perspective of the actor, by creating a version of this argument that is adapted to the phenomenon of Akrasia.

As observed by Velleman, people have a *conception of agency*, or the idea that they are somehow involved in their actions. In order to understand the meaning of this agency in more depth, it is useful to analyse the concept and to define what it consists in. The way I define agency is as consisting in (1) the ability to *have* a Will, and (2) the ability to *exercise* one's Will in action. The ability to *have* a Will refers to the ability to reflect on one's own mental states and to form an evaluational want by weighing them cognitively. This is depicted by the R and W in Figure 2.

The ability to *exercise* one's Will refers to the ability of a person to connect their Will to their bodily movement by acting according to their Will. The Four-Component Model of Action as depicted in Figure 2 does not enable these kind of actions; the Will (W) is separated from motivation to act and eventual action (ACx). Therefore, in order to enable agency, the model needs to be expanded. What needs to be added is an overriding desire that enables one to act according to the Will rather than according to individual desires. I call this overriding desire 'Desire for Continence'. The adapted model is depicted in Figure 3.

¹⁵ J. David Velleman, "What Happens When Someone Acts?," *Mind* 101(403) (1992), 461-481.

¹⁶ Velleman, *What Happens When Someone Acts*, 4.

¹⁷ *Ibid.*, 7.

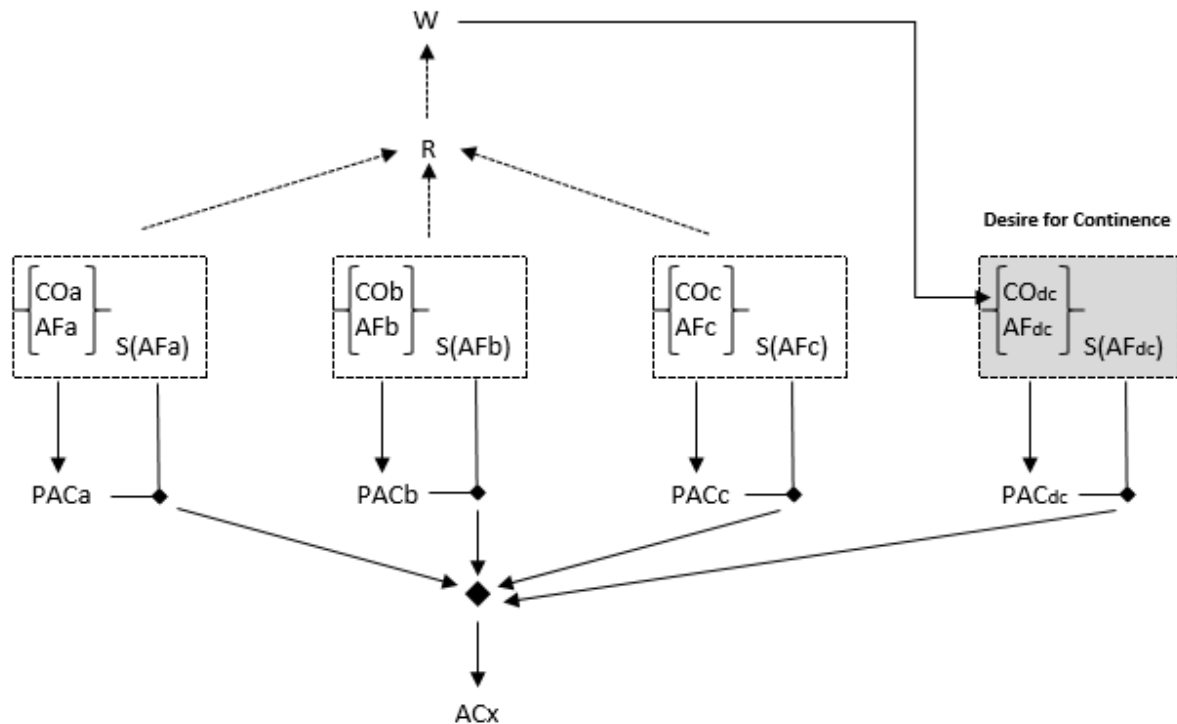


Figure 3. The Four-Component Model of Action with Desire for Continance

On the right an AFx-COx pair for Desire for Continance has been added that is connected to a potential action (PACdc). AFdc here refers to the Desire for Continance itself: it is the desire to act in accordance with the Will. COdc is an instrumental belief about what action will lead to fulfilling this desire. It is directly related to the Will, as fulfilling the Desire for Continance will be fulfilled by doing what is prescribed by the Will. S(AFdc) refers to the strength of the Desire for Continance. In order for a person to be able to act in accordance with their Will, their S(AFdc) needs to be larger than their other S(AFx)'s.

Based on this adapted model, Akratic action can now be explained from the perspective of the actor as a situation in which (1) in the definition of agency is developed in a person, that is, they have the ability to form a Will, and thus to have judgments about what is the best thing to do, but (2) is not developed enough to enable them to always act in accordance with that will, that is, their Desire for Continance is not strong enough to always override other desires (S(AFdc) is not always bigger than the other S(AFx)'s). Two questions still need to be addressed: (1) what does Desire for Continance entail specifically?, and (2) can Akratic action still be seen as action on this model? Both questions will be answered below.

First, what exactly *is* Desire for Continance; what does it entail specifically? It is important to understand that Desire for Continance does not necessarily have one specific content; it can have many different forms. What unites these forms, however, is that they are directed towards acting in accordance with the Will. They can differ with regard to (1) the specific *reason* for which they are directed at acting in accordance with the Will, and (2) the *aspects* of the Will they are directed at. The reason for wanting to act in Accordance with the Will, could for example be the type of person they desire to be, the image they want others to have of them, their valuing of continence and feeling that it is important in life, etc. Also, Desire for Continance can be directed at the entire Will, but also at specific aspects. Aspects of the Will that Desire for Continance could be directed at are numerous (e.g., family, work, study, morality, religion, sexuality, etc.).

Note that the nature of the Desire for Continance is *affective*. In this way it is opposed to more

cognitive versions of this argument, such as the Principle of Continence that Davidson proposes.¹⁸ Seeing the nature of the Desire for Continence as affective is needed in order to explain connection of the Will to motivation and to action. A Principle of Continence would be part of R in this model, and therefore would not solve the problem of connecting the Will to action.

Second, one could wonder whether Akratic action can still be seen as action on this model. In 1.1, action was defined as bodily movement that is *intentional*. It could be argued that if action is Akratic, rather than based on Desire for Continence, it is not intentional, and therefore does not qualify as action. In order to address this point, it is necessary to look at the meaning of 'intentional'. In 1.1 it was defined as *having an aim in mind* while acting. It is important to understand that this does not necessarily have to be an explicit aim. That is, the agent does not consciously have to think about this aim while performing their action. Nor does 'aim' refer to that which the agent decides to be the best thing for them to do after thorough consideration. Rather, 'aim' should be understood here as referring to the aim-directedness of affective mental states. It refers to the goal of acting that is set by desires or pro-attitudes of the agent. Understood thus, Akratic bodily movement can still be seen as intentional, as it is guided by an affective mental state (AFx). Therefore, it does qualify as action.

Conclusion

The purpose of this paper was to answer the question how the phenomenon of Akrasia can be understood within the Causal Theory of Action. In order to enable answering this question, two sub-questions have been addressed: (1) how Akratic actions can be understood within the Causal Theory of Action, and (2) how the Akratic actor can be understood within the Causal Theory of Action. By combining the answers that have been given to these two questions, the main question of this paper can now be answered.

Within the Causal Theory of Action the phenomenon of Akrasia can be understood as a situation in which the largest S(AFx) of the actor – which decides what they are most motivated to do, their motivational want – is not their S(AFdc), because of which their eventual action (ACx) differs from the potential action (PACdc) that is in accordance with their evaluational want, or Will (W). Thus, in these situations the actor will be caused to act in a way that goes against that which they believe to be the best thing to do, as the desire to act thus is stronger than their Desire for Continence.

In order to see how this solves the problem set out in the beginning of this paper, compare the following version of the principles of Davidson that were used to describe the problem of reconciling the CTA and the phenomenon of Akrasia, adapted according to the solution developed in this paper.

P1: If an actor wants to do x more than they want to do y and they believe themselves free to do either x or y, then they will intentionally do x if they do either x or y intentionally.

P2: If an actor judges that it would be better to do x than to do y, then they want to do x more than they want to do y, if they have a strong enough Desire for Continence that overrides their other desires.

P3: If an actor has a desire to do y, that is stronger than their other individual desires, and they do not have a Desire for Continence that is strong enough to override this desire, then they want to do y more than they want to do x.

Now, there is no problem anymore in accepting these principles, and accepting at the same time:

P4: There are Akratic actions.

For now Akratic actions can be explained by P3 in combination with P1. That is, in cases of Akrasia, the person acts on their strongest desire and the belief of what is a means to fulfilling that desire, because they do not have an overriding Desire for Continence that enables them to act on what they

¹⁸ Davidson, *How is Weakness of the Will Possible?*, 41.

consider to be the better thing to do.

Although a method of reconciling the CTA with the phenomenon of Akrasia has been developed in this paper, it has several limitations. In particular, the concepts of 'reflective ability and 'Desire for Continence' need to be further scrutinized and elaborated upon. The purpose of this paper has been to develop and formulate them such as to enable understanding of Akrasia within the CTA, but further research is needed to create a more in depth understanding of these concepts. Subsequently, future research could be done to investigate how these two concepts are related to other concepts, such as intention and free will. Thus they could be used to explore other questions, concepts, and philosophical problems, and to advance understanding in other areas.

Literature

- Aguilar, Jesús H., & Buckareff, Andrei A. (ed.), *Causing Human Actions*. Cambridge: MIT Press, 2010.
- Aristotle. *Ethica Nicomachea*. Translated by Christine Pannier & Jean Verhaeghe, 7th edition. Groningen: Historische Uitgeverij, 2009.
- Bobonich, Cristopher, & Destrée, Pierre. *Akrasia in Greek Philosophy*. Leiden: Brill, 2007.
- Bunning, Nicholas. *The Blackwell Dictionary of Western Philosophy*. Malden: Blackwell, 2004.
- Davidson, Donald, "Actions, Reasons, and Causes" (1963)." In *Essays on Actions & Events*, 3-19. New York: Oxford University Press, 1980.
- "How is Weakness of the Will Possible?" (1970)." In *Essays on Actions & Events*, 21-42. New York: Oxford University Press, 1980.
- Enç, Berent. "How Causes Can Rationalize: Belief-Desire Explanations of Action." *Turkish Studies in the History and Philosophy of Science* (2005), 231-247.
- Kalis, Annemarie. *Failures of Agency. Irrational behavior and self-understanding*. Lanham: Lexington Books, 2011.
- Mele, Alfred R., "Akratic Action: Causes, Reasons, Intentions." In *Irrationality*, 31-49. Oxford: Oxford University Press, 1987.
- Plato. "Protagoras." In *Protagoras and Meno*, edited by Robert C. Bartlett, 1-66. New York: Cornell University Press, 2004.
- Smith, Michael. "The Standard Story of Action: An Exchange (1)." In *Causing Human Actions*, edited by Jesús H. Aguilar & Andrei A. Buckareff, 45-55. Cambridge: MIT Press, 2010.
- Velleman, J. David. "What Happens When Someone Acts?" *Mind* 101(403) (1992), 461-481.
- Watson, Gary. "Skepticism about Weakness of Will." *The Philosophical Review* 86 (3) (1977), 316-339.