

I knew I Shouldn't Do It; But I Did It: Davidson on Causal Strength and Weakness of Will

Rafael Martins¹

¹ University of Kansas

E-mail: ramartns@gmail.com

ABSTRACT: *Reasons for action is a widely employed methodology in practical philosophy, and especially in moral philosophy. Reasons are facts that explain and justify actions. But, conceptually, if reasons were causes, incontinent actions would be impossible. When an agent ranks an evaluation about what to do as his best judgement, it entails that he has a reason for acting as that judgement prescribes. But when an agent acts incontinently, he acts in accordance to an intention that is not aligned with his best evaluative judgement. Yet, if the agent's best evaluative judgement provides him a reason for action, this reason should also be his strongest reason, and therefore, the strongest cause. How then can it be possible that an agent incontinently acts according to a reason of inferior causal strength? In this paper, I analyze how Davidson's argument for the possibility of incontinent actions interacts with his causal theory of actions. I argue that Davidson's proposal does not fully respect the two principles of intentional rationality, that he himself claims to be compelling. Lastly, I sketch some initial steps that might be helpful to drawing more precise conceptual distinctions in terms of the rationality of incontinent actions.*

Keywords: *Mental Causation - Rationalism - Akrasia*

RESUMO: Razões para ação é uma metodologia amplamente empregada na filosofia prática e, especialmente, na filosofia moral. Razões são fatos que explicam e justificam ações. Mas, conceitualmente, se as razões fossem causas, ações incontinentes seriam impossíveis. Quando um agente classifica uma avaliação sobre o que fazer como seu melhor julgamento, isso implica que ele tem uma razão para agir conforme o julgamento prescreve. Mas quando um agente age incontinentemente, ele age de acordo com uma intenção que não está alinhada com seu melhor julgamento avaliativo. No entanto, se o melhor julgamento avaliado do agente lhe fornecer uma razão para ação, essa razão também deve ser sua razão mais forte e, portanto, a causa mais forte. Como então pode ser possível que um agente incontinentemente aja de acordo com uma razão de força causal inferior? Neste artigo, analiso como o argumento de Davidson para a possibilidade de ações incontinentes comporta-se conceitualmente com sua teoria causal das ações. Argumento que a proposta de Davidson não respeita totalmente os dois princípios da racionalidade intencional, que ele mesmo afirma serem convincentes. Por último, esboço alguns passos iniciais que podem ser úteis para desenhar distinções conceituais mais precisas em termos da racionalidade das ações incontinentes.

Palavras-chave: Causação Mental – Racionalismo - Akrasia

Introduction

Donald Davidson proposes that the distinction between all-things-considered judgments and unconditional judgements prevents the incontinent agent from being accused of logical contradiction. Second, he argues that practical reasoning resembles probabilistic reasoning, and therefore, all-things-considered judgements are best understood as *prima facie*. So, a *prima facie* best evaluative judgment would

compete with an opposite unconditional judgment that may explain the incontinent action. Davidson claims that his solution respects two principles about intentional actions and practical reason. P1) That what one wants to do, one does it intentionally and; P2) One always wants more that what one evaluates as best. But the possibility of incontinent action implies that one wants to do something that is not what one evaluates as best.

Davidson raised the problem I address here in the introduction of *Essays on Actions and Events*: “causal theories of action are challenged by intentional actions that are contrary to the actor’s best judgement. For if reasons are causes, it is natural to suppose that the strongest reasons are the strongest causes.” (2001, p. 08). In this paper, I analyze how Davidson’s argument for the possibility of incontinent actions comports with his causal theory of actions. First, I introduce the problem of weakness of will as proposed by Davidson with some clarifications about the virtues of this approach. Then, I sketch the causal theory of action and develop a detailed consideration of how Davidson proposes to maintain that the possibility of incontinent actions respects P1 and P2. For this purpose, I describe the distinction between conditional and unconditional judgements, and the analogy between practical reasoning and probabilistic reasoning. At this point of my considerations, I tentatively claim that Davidson’s proposal does not fully respect either P1 or P2. Lastly, I sketch some initial steps that might be helpful to drawing more precise conceptual distinctions in terms of the rationality of incontinent actions.

The problem of incontinent action (not a moral problem)

People often do not behave in the way they judge to be the best. Although agents seem to *know* what is best to do, it is possible that they are *not motivated* to act in accordance with their judgements. Moreover, sometimes agents act *contrary* to their best evaluative judgments. The first case exemplifies an inconsistency between evaluative judgments, on the one hand, and motivation on the other. The second displays an inconsistency between actions and judgments. These motivational or behavioral asymmetries in relation to evaluative judgments have been studied under various methods, and in philosophy they are known as either *akrasia* or weakness of will. Scholars from different fields like philosophy, moral psychology, and neuroscience debate about what is the correct characterization of an akratic or weak-willed action, and whether the contemporary concept of weakness of will equates to *akrasia* as understood in ancient philosophy.¹ In spite of those discussions, this paper focuses on weakness of will in the works of Donald Davidson, and his main argument for its possibility, given his causal theory of action.

One of the virtues of Davidson’s philosophical approach is that he analyses the phenomenon in a nonmoral style. This enables a more precise identification of weakness of will in the world and prevents us from confusing it with *moral* weakness of will. Incontinent actions, as Davidson labels acts that display weakness of will, are simply actions contrary to one’s best judgement, regardless its practical content. Best judgements may correspond to norms from different dimensions of practical reason like morality, prudence, etiquette, law, aesthetic, economic, and others. For instance, a best judgment may be the evaluation that it is better to act according to a legal norm, or to a prudential principle, and so forth. But what is intriguing about weakness of will is that the agent recognizes that there is a course of action open to him,

¹ As a sample of that discussion in the case of philosophy of action, see debate over the actual nature or, say, correct view of what is *akrasia* and/or weakness of will in Holton, R. Intention and weakness of will. *The Journal of Philosophy*, 96, (1999): 241–262; Mele, A. Weakness of will and *akrasia*. *Philosophical Studies*, 150(3) (2010): 391–404, and May, J. & Holton, R. What in the world is weakness of will. *Philosophical Studies*. Forthcoming (2011).

which he judges to be the best, and yet he ends up acting in accordance with some other lower-graded judgment. Thus, Davidson proposes to dissociate the problem of incontinent actions from its moral dimension in order to get to a clear framing of the phenomenon in question:

In approaching the problem of incontinence, it is a good idea to dwell on the cases where morality simply doesn't enter the picture as one of the contestants for our favour. (...) Then we shall not succumb to the temptation to reduce incontinence to such special cases as being overcome by the beast in us, or of failing to heed the call of duty, or as succumbing to temptation. (2006b, p. 79)

The problem of incontinence is not restricted to morality, but pertains to philosophy of action in general.² Therefore, the relevant instance of incontinent actions is simply an action contrary to one's best judgement, be it moral, prudential, aesthetic, or of any other substantial practical nature.

Identifying the correct instances of incontinent action is not only a matter of understanding that the agent's best judgement may be about different parts of a pluralist conception of practical reason. It should also include that incontinent actions do not be reduced to cases where agents act motivated by desires contrary to moral judgements. As well as it may be the case that a desire for a particular pleasure is coincidentally aligned with the agent's best judgement in favor of acting upon a principle, which simply happens to be moral in content. To set distractions aside, Davidson avoids the moralized view³ of *akrasia*, where every agent's best judgement is always necessarily a moral duty, and the agent, supposedly, falls victim of a desire that pushes him to an opposite direction. As he puts it: "there are numerous occasions when immediate pleasure yields to principle, politeness, or sense of duty and yet we judge (or know) that, all-things-considered, we should opt for pleasure" (2006b, p. 79).⁴ So, Davidson proposes to "divorce that problem entirely from moralist's concern that our sense of the conventionally right may be lulled, dulled, or duped by a [desire for] lively pleasure" (2006b, p. 79). The problem of incontinent action is only an investigation of how an agent acts or is motivated to act contrary to his best judgment, independently of whether that judgment recommends that a desire should be pursued. It suffices to say that the agent evaluates it as his best, or at least as better than all the others.

To illustrate the sort of theoretical divorce proposed by Davidson, consider a professional diver seeking victims of a tragedy, whose best judgement is a prudential conclusion in favor of not diving in what seems to him dangerous waters. Still, as a matter of professional commitment, he acts contrary to his best judgement and dives in. It must be added that the diver *has* a desire to dive in the water, because, as the Humean⁵ picture suggests, he could not act based solely upon his belief about his professional commitment. But the diver's practical judgement about his professional obligations does not equate to his best judgement, rather it is actually contrary to it, as well as it is the cause of his incontinent action. Therefore, the problem of weakness of will hinges only on the agent's not being motivated, or not acting in ac-

² Broadly construed, action theory concerns physical and mental processes that explain or cause actions and the rational calculations that justify intentional behavior.

³ Often seen under the label of moral rationalism.

⁴ For instance, a best judgement may result from rational evaluation of moral principles in conjunction with a desire to act morally. In this case, the best judgement is a conclusion that it is better to satisfy a given desire, which happens to be in accordance with a moral principle. This view aligns to, among others, the hedonist tradition in which attending to some (to-be-specified) pleasure is sometimes symmetrical to the commands of morality. But that doesn't have to be the case. And, probably, it isn't.

⁵ More about Davidson's endorsement of Humean legacy is exposed in the next section.

cordance with her best judgement, it is not a problem of moral weakness, neither a simple surrender to a desire.

The Causal Theory of Action

When an agent ranks an evaluation about what to do as his best judgement, it entails that he has a reason for acting as that judgement prescribes. But not just that. It also implies that this reason is the strongest cause of the action, among all other possible accruing causal forces. Davidson advances a Humean theory of reasons, where someone having a reason implies that the agent has a desire and a cognitive mental state, namely, a belief that the action in question has the property of bringing about the satisfaction of that desire:

C1: *R* is a primary reason why an agent performed the action *A* under the description *d* only if *R* consists of a pro-attitude of the agent towards actions with a certain property, and a belief of the agent that *A*, under the description *d*, has that property. (2006a, p. 25)

At a first glance, an explanation of one's actions by pointing out one's reasons (i.e. pointing out the reason that was causally efficient for one's action) may not overtly express the agent's pair of a pro-attitude and a correspondent belief. For example:

“Why did Jane go to the restaurant?”
- “Because she was hungry.”

Saying Jane is hungry does not immediately represent what are her mental states involved in the causal explanation of her action. But, in any case of rationalization of actions, the explanation provides enough information so that an interpreter can reconstrue⁶ Jane's reason in the following way:

Pro-attitude: Jane desires to satisfy her hunger.

Belief: Jane believes that by going to the restaurant she would satisfy her hunger.

Next, Davidson defends that there is a causal connection between reasons and actions:

C2: A primary reason for an action is its cause. (2006a, p.30)

Together, Jane's desire to satisfy her hunger plus her belief that by going to the restaurant she would satisfy her desire, constitute the cause that explains Jane's physical behavior of going to the restaurant. Therefore, to mention the reason that justifies Jane's action is also a form of causal explanation.⁷

⁶ In Davidson's words: “In order to understand how a reason of any kind rationalizes an action it is necessary and sufficient that we see, at least in essential outline, how to construct a primary reason.” (2006a, p. 24).

⁷ For Davidson, a justifying reason is also an explanatory reason, but the opposite does not necessarily hold. When a belief-desire pair justifies an action, it causally explains the action. However, a causal explanation of an action does not always justify that action. This asymmetry is due to a distinction between teleological and nonteleological conceptions of causation. The first explains an event by citing goals or purposes, whereas the latter does not appeal to those. As an instance of a nonteleological explanation, imagine that someone asks, “Why is Josh hurt? and obtains the answer “Because he has been in a car accident.” The relevant facts of such an explanation do not involve the fulfillment of any goal or end like in Jane's example, so in this sense they do not justify the event that they explain. As a result, some philosophers have objected that a justifying reason pertaining to the domain of teleological explanations cannot be reduced to a causal reason from the realm of nonteleological explanations.

There is an immense bulk of work on problems raised against Davidson's Causal Theory of Action (CTA), but my aim here is not to challenge CTA. Rather, I focus on how Davidson's defense of the possibility of incontinent action comports itself with his CTA. The problem starts with recognizing that when an agent acts incontinently, she acts in accordance to an intention that is not aligned with her best evaluative judgement about what to do. Yet, if the agent's best evaluative judgement provides her a reason for action, this reason should also be her strongest reason, and by CTA, the strongest cause. How then can it be possible that an agent incontinently acts according to a reason of inferior causal strength? In the introduction of *Essays on Actions and Events*, Davidson makes this point explicit: "causal theories of action are challenged by intentional actions that are contrary to the actor's best judgement. For if reasons are causes, it is natural to suppose that the strongest reasons are the strongest causes."⁸ (2001, p. 08) Mele equally observes this difficulty in asking whether "occurrences of akratic actions" are "compatible with the truth of CTA" (1987, p.31). In *How is Weakness of Will possible?* (2006b), Davidson proposes a way out of this problem. I introduce it in the next section.

A Distinction Between Conditional and Unconditional Judgements

Davidson defends incontinent actions as actions that an agent, knowingly and intentionally, performs counter to her, *all things considered*, best judgement. According to Davidson's definition, 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 her.
- (c) The agent judges that, *all things considered*, it would be better to do y than to do x .

(2006b, p. 72)

That an agent does x intentionally means that x is desirable.⁹ If it is an intentional action, there is a belief and a pro-attitude that sustain it, but this fact does not imply that the pair of mental states is a reason. In (b) the agent believes she has simultaneous belief-desire pairs before her, and they compete along the agent's scale of evaluation. And to understand (c), consider Davidson's illustration of someone who has just relaxed in bed after a hard day when it occurs to her that she forgot to brush her teeth. She then evaluates the alternatives according to some facts like: "my teeth are strong and at my age decay is slow, [and] if I get up, it will spoil my calm and may result in a bad night of sleep." So, "sensual indulgence" suggests her to skip that for once. But the deliberation is done "in light of all her reasons", and some of the facts along this list, say, concerning her own health, induce the agent to judge that it is better to get up and brush. After pondering the force of the alternatives, she judges, everything considered, it is better to

⁸ I shall clarify that I am only discussing *internal* causal conflicts as opposed to conflicts between internal and external causal forces. In this manner, external obstacles, material or otherwise, such as a fence or a person holding another will not be considered.

⁹ This distinction is essential for Davidson's argument that an agent acts incontinently with an intention, but not rationally. The incontinent agent acts intentionally but based on a desire that she has no reason for acting in accordance. Davidson is very careful: "To know a primary reason why someone acted as he did is to know an intention with which the action was done. If I turn left at the fork because I want to get to Katmandu, my intention in turning left is to get to Katmandu. But to know the intention is not necessarily to know the primary reason in full detail. If James goes to the church with the intention of pleasing his mother, then he must have some pro-attitude toward pleasing his mother, but it needs more information to tell whether his reason is that he enjoys pleasing his mother, or thinks it right, his duty, or an obligation." (2006a, p. 26)

stay in bed. Even so, the agent's "feeling that she ought to brush her teeth"¹⁰ takes her out of bed and she brushes. Davidson classifies such action as incontinent because it is against the agent's judgement that it is better to stay in bed. Although her all-things-considered judgement was in favor of self-indulgence, her intention to do something else was stronger. Thus, the formation of an all-things-considered judgement coexists with the agent's competing intentions. (c) is clear now: in the case of incontinent action, even though the agent deliberates that *y* is the best course of action, she acted in accordance with her intention to *x*. (2006b, p. 79)

There is here an asymmetry between the agent's intention that moves her and her all-things-considered judgement. The agent's judgments about her reasons are not necessarily ranked in correspondence with the actual motivational force of those reasons. So, the judgement that lies behind the agent's incontinent action has a different force than her all-things-considered judgement. In order to explain this asymmetry, Davidson introduces the distinction between conditional and unconditional evaluative judgements. The agent who incontinently does *x* rather than *y*, does indeed judge that *y* is better than *x*, but this is a conditional judgement that, all things considered, "Doing *y* is best". An all-things-considered judgement results from the agent's deliberation about competing reasons for action given what she knows about the world (and its functioning) and about herself (the moral theories she holds, her prudential principles, her professional obligations, ties of love, her capacities and abilities or lack of them, etc.). This is one way to exemplify it: "Given reasons *r*₁, *r*₂, *r*₃,...: Doing *y* is best." As such, deliberation is a process in which the agent marshals evidence and principles, and judges about what is best to do in light of all available relevant reasons.¹¹ It is far from clear what are all the factors that must count in this epistemic process. However, it is plausible to assume that the proviso "all things considered" means that the agent engages in a process of pondering competing courses of action.¹²

On the other hand, an unconditional judgement is a *sans phrase* or unqualified judgement of the form "Doing *x* is best". In a rational (and thereby continent) agent, unconditional judgements always follow from conditional judgements. Davidson seems to imply that because unconditional judgments are not tied to a thoughtful treatment of competing reasons, they are more readily reflected in the agent's actions.¹³ So in the case of incontinence, the agent judges that, given all the available reasons, doing *y* is best; but given some other reason or subset of available reasons, she judges that doing *x* is best. She then forms the unconditional all-out or partially considered judgment that doing *x* is best on the basis of her second judgement, and ultimately acts on that. In this manner, she violates the principle of continence.

Thus, conditional and unconditional judgements are operational at different epistemological levels,

¹⁰ It is curious that here Davidson does not mention facts that justify her reason to brush her teeth that enter into her deliberation, except for a moral feeling that she ought to brush her teeth. This helps him to make the case for unconditional judgements.

¹¹ Davidson calls this the "principle of continence" which is to judge what is best to do on the basis of all available relevant reasons. (2006c, p. 88)

¹² Davidson talks about this point in different passages, here is one example: "Genuine conditional intentions are appropriate when we explicitly consider what to do in various contingencies." (2006c, p.131)

¹³ Here is how Davidson makes the point: "*Prima facie* judgements cannot be directly associated with actions, for it is not reasonable to perform an action merely because it has a desirable characteristic." (2006c, p.134) Besides, "We cannot suppose that whenever an agent acts intentionally, he goes through a process of deliberation or reasoning, marshals evidence and principles, and draws conclusions." (2006c, p.124) Therefore, "The judgement that corresponds to, or perhaps is identical with, the action cannot be a *prima facie* judgement; it must be an all-out or unconditional judgement which, if we were to express it in words, would have a form like 'This action is desirable'." (2006c, p.134)

where the former type is not *necessarily* accompanied by motivation or action.¹⁴ Thereby they do not stand in logical contradiction. In this manner, Davidson's incontinent agent may simultaneously and consistently hold: i) a conditional judgement that, all-things-considered, *y* is best to do with ii) the unconditional judgement that *x* is best to do. Indeed, if the judgement (i) "*y* is best to do" were not somehow different in its epistemic nature from the judgement (ii) "*x* is best to do", it would be impossible to accept that the agent simultaneously holds both without any contradiction.

Accommodating Principles and *Prima Facie* Judgements

However, it is not yet explained why the agent performs *x* after deliberating that, all things considered, *y* is a better course of action. The difficulty here is that for most of the views, the action that one evaluates as best to do is also the action that one most *wants* to do. For instance, Socratic intellectualism claims that under rational conditions, agents necessarily act accordingly to their best judgement. So, if an agent believes *y* is a better course of action and, in spite of that, she performed *x*, she actually did not truly consider *y* better, but *x* instead. This view is committed to a strong relation between the agent's judgements, motivations, and actions. As the view suggests, if an agent rationally deliberates about what to do, her actions will be natural consequences of her best judgements. In other words, if an agent rationally deliberated about what to do, it is part of being rational that her actions are in accordance with her best judgements. The one who fails to act so is irrational. Likewise, it is part of this trend that motivations are symmetrical to evaluations, i.e. that what the agent ranks as best to do is also what she is most motivated to do. Consequently, the one who fails to present such symmetry is irrational.¹⁵

Davidson credits the apparent impossibility of incontinent actions to the inconsistency between two principles of intentional rationality. But he believes them to be quite compelling. So, without dropping either one, he wants to hold that incontinent actions are possible in a causal framework of reasons. These are the principles:

P1: If an agent wants to do *x* more than he wants to do *y* and he believes himself free to do either *x* or *y*, then he will intentionally do *x* if he does either *x* or *y* intentionally.

P2: If an agent judges that it would be better to do *x* than to do *y*, then he wants to do *x* more than he wants to do *y*.

P3: There are incontinent actions.
(2009b, p. 73)

P1 states that what a free agent wants more is what she does intentionally. It tells that choosing among different courses of action connects ranking an option (a belief-desire pair) and acting on its basis. P2 means that an agent wants more that what she believes to be better or, in other words, it says that an agent is more motivated for the action that she ranks higher. These two principles together block P3 as it is not possible that one wants something that is not believed to be better (contradiction of P2), and even

¹⁴ I call these views normative judgment externalism, *viz.* The thesis that all-things-considered reasons are not necessarily motivational.

¹⁵ I propose some adjustments in the next sections in an attempt to explain how the agent chooses to act for other reasons instead of his best and may not be accused of irrationality.

less possible that one intentionally acts in disagreement with what one wants most (contradiction of P1). However, Davidson wants to defend the existence of incontinent actions without rejecting fundamental ties between judgements and actions. (2006b, p. 74) For that purpose, he intends to show that P1, P2, and P3 do not contradict one another.

Davidson proposes to solve the problem by using an analogy between the logical forms of practical reasoning and probabilistic reasoning to explain how, as he puts it, “the will judges one course of action better and yet chooses the other?” (2006b, p. 83) Davidson’s analogy goes as following (I am just using parts of it):

- (M4) If the barometer falls, it almost certainly will rain. *pr* (Rx, Fx)
 (m4) The barometer is falling. *Fa*
 (C4) It almost certainly **will** rain *pr* (Ra, pr (Rx, Fx) and *Fa*)

And, says Davidson, “we may at the same time be equally justified in arguing that”:

- (M5) Red skies at night, it almost certainly won’t rain. *pr* ($\sim Rx, Sx$)
 (m5) The sky is red tonight. *Sa*
 (C5) It almost certainly **won’t** rain. *pr* ($\sim Ra, pr$ ($\sim Rx, Sx$) and *Sa*) (1969, p. 85)

The analogy shows that conditional evaluative judgements are *prima facie* like the conclusions of probabilistic reasoning, which may be either supported or overturned by new available evidence. So, a *prima facie* practical conclusion can be supported or overturned by new available reasons. But the introduction of “*prima facie*” should not be done exclusively at the level of the conclusion (for instance: 1. *prima facie* it will certainly rain, 2. lying is *prima facie* wrong, etc.) for that would produce the logical impossibility that an action is both *prima facie* right and *prima facie* wrong. Rather, “*prima facie*” should apply as connecting the evidence with some likely conclusion, for example: that the skies are red tonight makes it *prima facie* probable that it will not rain. Likewise, in the practical reason case, that an act is a lie, *prima facie* makes it wrong.

So, for Davidson’s view a moral principle, like “Lying is wrong”, should be taken as “that an act is a lie, *prima facie* makes it wrong”. In this case, the fact that the connection is *prima facie* means that it can be overturned by further reasons or by competing evaluative judgments based on defective sets of reasons.¹⁶ An unconditional judgement contrary to “lying is wrong” can be formed (for instance: in the circumstances of getting one’s mother out of jail, one judges that it is *prima facie* right to lie, and then forms the unconditional judgement that “it is right to lie”) and thereby overturn her best *prima facie* judgement that lying is wrong. Therefore, given a particular reason (that the agent’s mother is in jail), the agent’s all-things-considered judgement is overturned by an unconditional judgement that it is right to lie. This allows her to form an intention to lie and so act “incontinently”.¹⁷

Thus, we have the following scenario:

¹⁶ Sets that do not fit the Principle of Continence.

¹⁷ Because her conditional judgement “given that my mother is in jail, it is *prima facie* right to lie” is supported by a defective set of relevant reasons available to the agent. It is defective due to its lack of important and accessible reasons to the agent, as for instance that it is illegal to lie under oath. Consequently, her unconditional judgement that it is right to lie does not follow a conditional judgement that respects the principle of continence. She is blamed for overlooking reasons that are relevant and accessible to her.

- i) That an act is a lie, *prima facie* makes it wrong. (all-things-considered judgement)
- ii) It is right to lie. (unconditional judgement)

In theory, an agent would have reasons in support of both (i) and (ii). For example, the reasons in support of (i) may include the fact that she is testifying under oath, plus all the reasons that are relevant and available to her. Thus, the agent judges that in such circumstance not lying would be, *prima facie*, the best action. Let us consider the set of reasons in favor of (i) as s1. However, there is a less complete set of reasons for which the agent may judge that lying would be the *prima facie* best action. For example, the defendant is her mother and the agent knows that lying would keep her mother out of jail, and this set of reasons lacks some relevant one or more considerations. In other words, the agent, knowingly or not, neglects one or more reasons. Let us refer to this latter set of reasons in favor of (ii) as s2. We have then (i) and (ii) as follows:

- iii) *Prima facie*, telling the truth is the best to do, given s1.
- iv) *Prima facie*, not telling the truth is the best to do, given s2.

Metaphysical and epistemic virtues of s1 and s2 aside, the considerations (iii) and (iv) are tied in competing for the agent's decision at this point. Now, consider that the agent ranks (iii) as her best judgement after deliberating about all available reasons and their likely consequences, for instance: lying and also going to jail if she gets caught, lying and not getting caught, thereby releasing her mother, not lying and not having her mother close to her, hiring other lawyers, etc. However, she *really* would like to have her mother out of jail, no matter the risks it takes. So, she departs from (iii), an all-things-considered option and acts upon (iv), an unconditional judgement that "not telling the truth is the best to do". In the same manner that in probabilistic reasoning "almost certain" allows that statements like "given that the barometer has fallen, it almost certainly will rain tonight" be overturned by "skies are red tonight", in practical reasoning, "*prima facie*" allows that statements like "given s1, *prima facie* telling the truth is the best to do" be overturned by s2. The incontinence results from the appearance that s2 violates the principle of continence.

The proposed solution is therefore given in two phases. First, Davidson distinguishes between conditional and unconditional judgements in order to answer the problem that an individual cannot consistently hold two diametrically opposite judgements. The second move is the analogy between probabilistic reasoning and practical reasoning. The analogy reveals that evaluative judgements are *prima facie* judgements and thus revisable in the face of further available reasons or of competing unconditional judgments. So, the agent's best judgement does not clash with her incontinent judgement because the latter may have been brought about by new evidence against the former.

By making the two moves, Davidson hopes to have reconciled P1, P2 and P3. He closes "Intending" with the conclusion that "to intend to do something is to want to do something in some very broad sense." (2006c, p. 136) So, I take this to provide support for the conditional that if the agent wants more to act based on a detached judgment (iv), which means to act in disagreement with her continent best judgement (iii), then she does it intentionally, which matches P1. Second, she wants more that what she evaluates best, respecting P2. The problem that I see here is that she has an intention that matches an unconditional judgement detached by a conditional judgement that does not respect the principle of continence. If I am right, this seems to imply that P2 is not fully honored. Originally, according to P2 the agent wants

more that what she evaluates best all-things-considered. But, what Davidson has shown is that the agent wants to do something that is not in accordance with what she judges best all-things-considered. He has not in fact explained how this want supersedes the agent's motivation related to her best all-things-considered judgement. The *prima facie* condition only shows how an unconditional judgement may supersede a conditional judgement, but it does not explain *why* the agent wants to do some other action more than that what she evaluates as best.

The Rationality of Incontinent Actions

Davidson's solution implies that in acting incontinently the agent is more motivated to act upon an unconditional judgement that does not follow from her all-things-considered judgement. So, the strategy dropped P2 because the agent does not want more that what she evaluates best. Also, it is not clear how does an all-things-considered judgement ever motivate the agent, if unconditional judgements are more readily reflected in the agent's behavior.¹⁸

All Davidson's account has made so far was to show how an epistemic asymmetry between judgements about what reasons the agent ranks as best and the ones that set her in motion could explain seemingly incontinent actions. But we do not have yet an explanation for the question "why anyone ever performs an action when he thought that, everything considered, another action would be better?" (2006b, p. 88) His best answer is that "the akratic does not hold logically contradictory beliefs, nor is his failure a moral failure. What is wrong is that the incontinent man acts, and judges, irrationally, for this is surely what we must say of a man who goes against his own best judgement." (2006b, p. 88) Davidson denies that those actions are part of a rational pattern and, based on CTA, he concludes that "if the question is what is the agent's reason for doing x when he believes it would be better, all things considered, to do another thing, then the answer must be: for this, the agent has no reason." (2006b, p. 88) The agent has an intention for acting incontinently, but no genuine reason for doing it.

Davidson's conclusion is not satisfactory, however. Does it not seem perfectly rational to lie in order to get one's mother out of jail? I shall think so. At this point, a better answer can be construed according to how we interpret the rationality of actions. In virtually all cases, philosophers have considered incontinent action a byproduct of irrationality, since this kind of behavior contradicts a fundamental thesis of practical reason that an agent always wants more that what she evaluates as best to do.¹⁹

The question whether incontinent actions are rational necessarily depends on the conception of rationality one holds. Hence, I suggest that incontinent actions can be explained in terms of a pluralist account of rationality and value. For instance, an agent may judge that the course of action x is best under a conception of the good μ and end up acting in accordance with the concept of good π , thus accused of irrationality (and thereby incontinence) on the grounds of μ . So incontinent actions are irrational insofar as

¹⁸ After Davidson establishes the distinctions that prevents competing evaluative judgements (as conclusions of practical reasonings) from logically clashing with each other, he recognizes that "there has been a loss of relevance, for the conditionalization that keeps (C6) from clashing with (C7), and (C8) from clashing with either, also insulates all three from action." (2006b, p. 86) As Davidson has previously argued, intentional action is propelled by unconditional judgements.

¹⁹ In Davidson's "How is Weakness of Will Possible?", I think it can be considered as contradicting either P1 or P2. However, discussion about the issue concentrate on P2, which roughly means that what the agents evaluates best is also that what he wants more. Or, in other words, a thesis about the symmetry between the agent's evaluations and motivations.

they are interpreted on the basis of a particular conception of rationality. Only when we choose a rational standpoint, grounded on a particular value, can we argue that a particular action is irrational or not.

Let me introduce three conceptions of formal rationality. First, an agent may be considered as rational from the standpoint of internal consistency, like for instance the view of Michael Smith (1994). For this view, an important sign of rationality is coherence between sets of beliefs and sets of desires. In this case the agent who does not desire to do that what she believes is best to do is irrational. Rationality as internal consistency is not about consistency between practical judgements and actions. Rather, it posits a conceptual connection between one's best normative judgement and the motivation to perform it.²⁰ In this manner, an agent is irrational when the course of action that she is motivated to perform does not match what she judges, all things considered, to be the best for her to do. For instance, if the agent achieved the conditional judgement that it is better to abide by the law and yet fails to do so on the basis of an unconditional judgement, she is irrational either due to her lack of the desire to abide by the law or because she has another desire in an opposite direction, not equally well grounded. Therefore, not desiring to act as one's best judgement prescribes, or acting against it, is irrational.

Second, for an external consistency approach exemplified by Arpaly (2003), a rational agent is the one who demonstrates that her beliefs coherently track evidence. That means the agent infers correct or highly probable conclusions given her available evidence. Arpaly's argument suggests that it is possible that an agent has a false belief about what she would do, if she were fully rational, because she lacks important evidence. So, the agent falsely believes she is rationally achieving an evaluative judgement, but because, unknowingly to her, her evaluation lacks important evidence, it does not produce a fully rational judgment about what to do, since not all things have been considered. The focus of this view is not the consistency between evaluative judgments and correspondent desires, rather, it is the consistency between evaluative judgments and the evidence that supports it that makes it rational.

Finally, rationality can be construed as behavioral consistency between the agent's judgments and her actions.²¹ For this view, an agent is rational insofar as her actions coherently reflect her evaluative judgments. Hare, for instance, defended rationality as logical consistency between judgement and action: "it is a tautology to say that we cannot sincerely assent to a command addressed to ourselves, and at the same time not perform it, if now is the occasion for performing it, and it is in our (physical and psychological) power to do so." (1963, p. 79) If that is a tautology, then once the agent reaches an evaluative judgment, it is a logical necessity that either she acts accordingly, or she is irrational. This view is sometimes seen as a global result of the previous two conceptions in case they are true. If one's best judgement is conceived of under any of such rationalist views, then acting against one's best judgement or even not being motivated to act accordingly, is irrational. The only form of saving the agent's rationality is to show that the judgement moving the agent towards another direction is also somehow grounded.

²⁰ For the theory proposed by Smith, it is conceptually necessary that if an agent judges that he is rationally required to perform action X in circumstance Y, then he will be equally motivated to perform X in circumstance Y, or he is irrational from the practical viewpoint. Smith characterizes practical irrationality as subjection to "confusing influences of weakness of will" (1994, p 61) due to "psychological compulsions, physical vices, emotional disorders, depression, mental tiredness and diseases." (ibid, p. 154)

²¹ Social interaction requires that we identify such relation. We often explain and predict people's behaviors by tracking their thoughts, namely, their beliefs and desires that apparently lie behind their actions. Behavior explanations and predictions are useful for our everyday lives. That is how we know, for instance, that given our professor's beliefs and desires, he will come to class tomorrow. However, the requirement imposed by social interaction does not imply that the relation between people's beliefs and their desires to act accordingly is necessary.

Before I move on to the next section, I must mention that I am not postulating multiple rational structures different in kind. If I am not wrong, my view only assumes that these three views are different tokens of one rational structure. In this more overarching structure, rationality equates to consistency. After all, it is this sort of relation that remains as demanded in common among the three conceptions. In fact, they present three different ways in which rationality as consistency is attributed to agents.

Degrees and Plurality of Values

If incontinent actions are not rational, Davidson may be losing his claim that they are intentional. Sure, if an action is rational, then it must be intentional as well. As he defended in “Actions, Reasons, and Causes,” when an action is performed for a reason, it carries the minimum condition that the action is intentional and “to know a reason why someone acted as he did is to know an intention with which the action was done.” (2006a, p. 25-6) But Davidson seems to be saying that incontinent actions are intentional and yet not rational. His text it is not clear whether possessing an intention does not necessarily imply having at least *some* reason. Moreover, it is clear that to identify the intention of one’s action does not equate to identifying one’s reason for action, i.e. intentions are not reasons. But if it were true that to have an intention is to have a reason and Davidson denies that an incontinent action was done for a reason, then P1 is blocked, namely, that such action is intentional. In other words, if he denies that the action was not grounded on a genuine reason, then it may be true that the action was not intentional either.

In order to maintain P1 and P2, I suggest that the incontinent action is grounded on a reason. There must be some reason for which the agent acts incontinently, and this reason should be justifiable under one of the three conceptions of formal rationality I mentioned in the previous section. In each of the cases I offered, professional diver, tooth-brushing, and mother in jail, there is something in the incontinent option that attracts the agent.²² When one sees an option as attractive, one necessarily finds something good in it. The good that the agent believes she will reach by acting incontinently, or the goodness of the action in itself, exerts some force in her considerations and draws the agent to neglect her once best evaluative judgements. The difference is that the course of action evaluated as best is, according to the value in question, *more* valuable than the incontinent option. Thus, two assumptions make the incontinent agent irrational: a) There is only one single kind of good, and b) One always performs the action one believes to be the most valuable or conducive to the most valuable state of affairs. Therefore, an action which does not lead to that good, or is not intrinsically good, is irrational.

It is possible to defend that an agent has a reason to act incontinently because she sees at least *some* good in that course of action. There are some different ways to conceive of that difference. Keeping external accruing circumstances aside, we may remain monist and say that the difference is not of value, but of degree.²³ Or we can say that the good the agent sees in the incontinent action is not of the same type of the good she sees in the best evaluated action. In this case, the most relevant difference between the values of the actions, or of the value they bring about, is not in the degree of each, but that they are distinct kinds. The incontinent action promotes another kind of value than the value promoted by the agent’s best evaluative judgement. For the pluralist, there are values irreducible to each other, such as, liberty, equality, happiness, wealth, love, rationality, etc. that agents simultaneously hold as good. My view is that, like P2

²² This point is developed by David Wiggins in Wiggins, D., 1980

²³ For instance, Stuart Mill’s distinction between lower and higher pleasures.

suggests, intentions are often aligned with rational evaluations, but there are competing requirements of consistency accruing to the agent's decision. The agent is not merely charmed by the incontinent option, but also different requirements of consistency may generate conflicts between the values in consideration, or between options carrying or producing different degrees of the same value.

It could be said that the agent is mistaken about the value of the alternatives. So, if an agent opts for an action of lesser value, she does not really know that there is an option of superior value. She has false beliefs about the values of the alternatives. However, Davidson's incontinent agent knows the differences in value between the options. She knows that her incontinent action is in some way less valuable than her best evaluative judgement. Yet, given CTA, if the best reason is also the strongest cause, how can the agent intentionally act according to a causally weaker reason?²⁴

The answer to this question has to be made in at least two major stages. First, it is necessary to decide between monism and pluralism about the structure of value, so that we identify the very value grounding the incontinent action. Then, we need to establish a theory of rational choice providing some kind of principled method of decision between available options. This theory will have to deal with concerns about the comparability of bearers of value, and the commensurability of values. In this paper I chose to remain ecumenical about value pluralism and monism, because incontinence can be explained according to either one. In the case of monism, the difference between the actions is of degree, while in pluralism, the actions are grounded on different values. Hence, the monist has to explain the rationality of a choice carrying or promoting a lower degree of value²⁵, while the pluralist²⁶ needs say something about the commensurability²⁷, or lack thereof, between different values. It seems to me that both need to address issues of comparability.²⁸ Defending either monism or pluralism and explaining incontinent action in detail according to a rational choice theory are tasks beyond the scope of this paper. My goal here is only to suggest that the incontinent action is minimally rational. If I am correct, then both monists and pluralists have the task to accommodate the rationality of incontinent action.

Acting Against One's Best Judgement Is Sometimes Rational

There are three instances in which it is possible to interpret an agent as rational: (i) consistency between evaluative judgments and actions, (ii) coherence between best judgements and desires, and (iii) consistency between the agent's evaluative beliefs and available evidence. These are demands of consistency according to which an agent can evaluate reasons and also, as I said, according to which her actions are interpreted. To blame one of acting against one's best judgement must be made from one of these three standpoints. But if the incontinent action can be justified, or even required, by one of the standpoints, then the action is rational.

I would like now to dissipate some accounts of incontinent action. It is natural to think that conscious beliefs are rational. When people arrive at a normative belief and thereby form a correspondent de-

²⁴ Davidson notes that "in the case of incontinence, the attempt to read reason into behavior is necessarily subject to a degree of frustration." (2006b, p. 89).

²⁵ Stocker (1990)

²⁶ Wiggins (1980) explains incontinent action in terms of incommensurable values pluralism.

²⁷ Incommensurability is about the relation between the values. A structure of incommensurable values would lack a common master value or measure according to which we can rank values in conflict.

²⁸ Incomparability describes a relation between two or more bearers of value. Chang (1997)

sire, they suppose that this belief results from an process in which they have pondered reasons in favor and counter to that belief, and because they ran this process in their minds, they think the belief is rational and conscious. However, the agent's thinking of her belief as rational does not make it rational for her to have this belief, because unknown to her it may be that her deliberative process was defective at some point. Say it either lacks sufficient evidence or relies on biased evidence. So, an agent may be considering an evaluative judgment as her best one, when it is actually not. No one disputes that human rationality is limited and imperfect as a matter of natural fact, but this is not a strong account of incontinence, because the agent's best evaluative judgment is not fully rational. In this case, it is not the incontinent action that is irrational, but the best evaluative judgment. However, at least conceptually, an evaluative judgment obtained by a defective epistemic process cannot be a *best* evaluative judgment. Second, this move dissolves the concept of incontinent action by transforming it into the next rational option, once it annuls the rationality of the former best judgment. In other words, the action is not incontinent anymore, rather it *is* the best option itself. This view seems to backslide into Socratic intellectualism, a view that is too strong to have any explanatory power in cases of highly intuitive incontinent action. For those cases, it is more promising to use the value-based account of incontinent action, where both the best evaluative judgment and the incontinent one are minimally rational, since they are both grounded on either two different degrees of the same value or on two different values.

Another account, such as Smith's (1994), relies on the agent's awareness of her own rationality, i.e., her belief A that her normative belief B is a byproduct of a fully rational deliberation. An agent would be considered rational insofar as her desires coherently track beliefs that she believes to have rationally achieved. But for Smith it is part of the meaning of a rational evaluative judgement that it elicits a desire for a correspondent action, which again sounds like Socratic intellectualism, since the best normative judgment necessarily prompts action. Smith evades problems for Socratic intellectualism by appealing to empirical obstacles to continent action. Due to the neglect of relevant evidence an agent can falsely believe it is rational for her to desire x. In this case it may not actually be true that she is irrational in not having such desire because there is a fundamental difference between the agent believing she is rationally reaching a belief about what she should do and really being rational. Under such circumstance, it would not be irrational not to act or desire accordingly to a belief that you just believe to have rationally achieved, but you are actually mistaken. If the agent falsely believes that she has rationally reached belief B (if it is false that belief B is rational), then even if the agent believes it is rational, belief B is actually not rational and therefore not desiring to act in accordance with it will not necessarily render the agent irrational. Since the only hindrance for the agent acting continently is empirical, this view also dissolves the notion of incontinent action at its conceptual level.

According to another account, fostered by Arpaly (2003), which I call external consistency, takes evidence to be an important factor in interpreting an agent as rational, besides her internal consistency or the coherence between her evaluative utterances and her actions. The upshot of this view is that it is not irrational to act contrary to one's best judgements if such judgements are mistaken due to their lack of correspondence to given evidence. As I mentioned in the second section, parts of this evidence is essential for the agent's ranking of what is worth pursuing in life and concern the world (and its functioning) and the agent herself (the moral theories she holds, her prudential principles, her professional obligations, ties of love, her capacities and abilities or lack of them, etc.). So, the agent incorrectly believes that she rationally reached her best judgement when actually she has not, because she has overlooked some piece of evidence. Examples are not hard to imagine since everyone may conceive of something that one would

rank as best to do, even though one does not have the ability required to perform such action.

The externalist conception of rationality seems to have sufficient characteristics to accommodate some of Davidson's claims about incontinent actions and the causal theory of action. He considered that incontinent actions were products of unconditional evaluative judgements, so we cannot suppose that incontinent actions are products of full-fledged rational deliberations. The externalist conception can accommodate that claim since a best judgement may be superseded by a later belief acquired by the agent becoming aware of new evidence, instead of rationally choosing to act incontinently. Still, we have to make justice to Davidson's causal theory of actions and for that we need to show how other reasons open to the agent, which are not his best may be have causal strength.

Davidson argued that one's best judgement is a *prima facie* conclusion precisely because they may be overturned by an unconditional judgment. It seems to me that there is no theoretical commitment that impedes this judgement to correctly track evidence while the agent's best judgement does not. This would give the unconditional judgement sufficient causal power due to its empirical consistency. Second, relating the unconditional judgment to evidence blocks the consideration that incontinent action is always irrational since such action pass muster with the externalist view of rationality. So, one's best judgement may well be wrong given particular evidence, and in this case, acting according to an all-out unconditional judgement that better tracks evidence is not irrational. In this manner, an agent cannot be blamed to be irrational if she does not desire to act in accordance with a belief that even though she believes it to be her best rational belief, it is actually an incontinent belief. If this thesis is correct, then it may be even possible to argue, like Arpaly, that desiring to do an action in accordance with an incontinent belief that an agent mistakenly believes to have rationally achieved is more irrational than not desiring to act accordingly to that belief (sort of incontinent action that many consider irrational). The value-based account is fine with this view insofar as the initially best evaluative judgment was minimally grounded.

Conclusion

My only overall conclusion in this paper is that not all incontinent actions are irrational, at least conceptually, since some of them are minimally grounded on a value. CTA-wise, unless the motivational force of an all-things-considered conditional judgment is always necessarily stronger, which seems false to me, it is possible that unconditional judgments sometimes acquire sufficient motivational force to supersede the former. That said, the Davidsonian account is unsatisfactory because it explains the incontinent action according to an epistemic difference between the incontinent and the best evaluative judgment. This difference renders the agent less than rational, either about the incontinent judgment or the best judgment, and hence without sufficient control over her options, thereby dissolving the problem. This view may mask the actual difference, or equality, of value between the options. Thus, I suggest that an account of incontinent actions needs to say something about the values themselves and their relation, controlling for all other respects such as material and epistemic.

Lastly, I conclude proposing that the problem about the rationality of incontinent action does not have to wait, conceptually, on a decision about the number of values. Rather, if the phenomenon of *rational* incontinent action really exists, its account only requires a view about value commensurability or comparability.

Bibliography

- Arpaly, Nomy. *Unprincipled Virtue: An Inquiry into Moral Agency*. Oxford University Press, 2003.
- Chang, Ruth, (ed.), *Incommensurability, Incomparability, and Practical Reason*. Harvard University Press. 1997
- Davidson, Donald. "Actions, Reasons, and Causes". *Essential Davidson*. Ed. Lepore, E. and Ludwig, K. Oxford Clarendon Press, 2006a. pp. 23-37.
- Davidson, Donald. "How is Weakness of the Will Possible?". *Essential Davidson*. Ed. Lepore, E. and Ludwig, K. Oxford Clarendon Press, 2006b. pp. 72-90.
- Davidson, Donald. "Intending". *Essential Davidson*. Ed. Lepore, E. and Ludwig, K. Oxford Clarendon Press, 2006c. pp. 122-137.
- Davidson, Donald. *Essays on Actions and Events*. Oxford University Press, 2001.
- Hare, Richard M. *Freedom and Reason*. Oxford University Press, 1963.
- Holton, Richard. "Intention and weakness of will." *The Journal of Philosophy* 96 (1999): 241–262
- May, J. & Holton, R. "What in the world is weakness of will." *Philosophical Studies* (2012) *Philosophical Studies* 157 (3):341–360
- Mele, Alfred. *Irrationality: an essay on akrasia, self-deception and self-control*. Oxford University Press, 1987.
- Mele, Alfred. "Weakness of will and akrasia." *Philosophical Studies* 150(3) (2010): 391–404
- Smith, Michael. *The Moral Problem*. Oxford: Blackwell, 1994.
- Stocker, M., *Plural and Conflicting Values*, Oxford: Clarendon Press, 1990.
- Wiggins, D., "Weakness of Will, Commensurability, and the Objects of Deliberation and Desire" in *Essays on Aristotle's Ethics*, A.O. Rorty (ed.), Berkeley: University of California Press, 1980

Artigo recebido em: 25 de dezembro de 2019

Artigo aceito em: 26 de dezembro de 2019