**A Dual Systems Theory of Incontinent Action**

**Aliya R. Dewey**

Forthcoming in *Philosophical Psychology*.

**Abstract**

In philosophy of action, we typically aim to explain action by appealing to conative attitudes whose contents are either logically consistent propositions or can be rendered as such. Call this “the logical criterion”. This is especially difficult to do with clear-minded, intentional incontinence since we have to explain how two judgments can have non-contradicting contents yet still aim at contradictory outcomes. Davidson devises an innovative way of doing this but compromises his ability to explain how our better judgments can cause our continent behaviors. In this essay, I preserve Davidson’s approach to the logical criterion but deviate from his broader theory of action by developing a default-interventionist dual systems theory of action. To do this, I focus on the dynamical relationship between System 1 and System 2: (1) the logical construction of value judgments in System 2 from System 1 and (2) the imaginative construction of non-propositional conative attitudes in System 1 from System 2. I draw on Street’s Humean constructivism and Peacocke’s theory of imagination for logical and imaginative construction, respectively. Within this framework, I provide a new definition of continence and incontinence that satisfies the logical criterion and explains how our better judgments can cause our continent behaviors.

**Introduction**

In *How is Weakness of the Will Possible?* (1970), Davidson argues that all of our propositional attitudes at any point in time must be logically consistent. This constraint on our psyche makes it very difficult to understand how incontinent action is possible: how can we simultaneously judge with clarity-of-mind that something ought to be done and yet ought not to be done without making contradictory judgments? Davidson dedicated his paper to solving this problem by advocating a nuanced understanding of the judgments that subtly defused the contradiction. He argued that the all-things-considered judgment (hereafter, “ATCJ”) has a conditional form ‘*a* is better than *b*, given a set of reasons *r*’ and is not practical in issue while the all-out judgment (hereafter, “AOJ”) has an unconditional form ‘*b* is better than *a*’ and is practical in issue. Although this solves our original problem, it does raise a new one: if the ATCJ is not practical in issue, how is it that we can act continently on the basis of the ATCJ, rather than the accidental alignment of the ATCJ with the AOJ? In this essay, I propose a causal mechanism that empowers the ATCJ so that the agent can *intentionally* align the all-out judgment with the ATCJ.

This is the agenda for this essay: in §1, I discuss Davidson’s (1970) approach to satisfying what I call “the logical criterion”. I argue that his approach is worth taking but that it leads to the new problem of explaining how ATCJs often cause continent behavior. In §2, with the aim of solving this new problem, I diverge from Davidson and articulate a default-interventionist, dual systems framework for explaining action. In §3, I draw on evidence and experience to develop three of the six components of the framework. In §4, I develop the fourth component of the framework, which is a series of proximal causes that starts with the ATCJ, continues with the AOJ, and terminates in behavior. In §5, I argue that this process offers an explanation for how ATCJs often cause continent behavior. In so doing, I offer a new definition of continence and incontinence. In §6, I defend my dual systems theory of incontinent action against worries that it generates an infinite regress. In §7, I conclude by surveying further challenges and opportunities that come with this theory.

**§1. The Logical Criterion**

There are several criteria that any theory of incontinent action ought to satisfy: e.g. it must give an explanation that relates mental states via causation, it must be empirically plausible, it must be phenomenologically plausible, etc. However, there is one criterion that is especially difficult for theories of incontinent action to satisfy. I call it “the logical criterion”: that a theory of incontinent action contradicts every sentence that states that a mind *M* at time *t* contains two attitudes *A1* and *A2* of the same type *T* whose contents *S1* & *S2* are contradictory (i.e. *S1* ∧ *S2* ⊢ ⊥).

The logical criterion has intuitive support: it seems psychologically impossible for a person to possess two attitudes of the same type whose contents are contradictory at the same time. Imagine that I am looking at my favorite dessert: a vegan lemon meringue pie. I form two belief-desire pairs: one that compels me to have a slice and one that compels me to put it back in the fridge. Despite the fact that the belief-desire pairs aim at satisfying contradictory sentences (i.e. ‘I am eating the pie’; ‘I am not eating the pie’), we insist that they do not contradict. Instead, we analyze them with a finer grain. We say that the first is the belief that what I would be eating is lemon meringue pie coupled with the desire that I experience the gustatory pleasure of consuming lemon meringue pie. The second is the belief that what I would be eating would violate my dietary commitments coupled with the desire to honor those commitments. We expect that the apparent contradiction between attitudes will dissolve under finer-grained analysis. Accordingly, we never expect to find anyone with the belief that what I would be eating is lemon meringue pie coupled with *both* the desire that I experience the gustatory pleasure of consuming lemon meringue and the desire that I do not experience that same pleasure. Given this disposition of ours, we have reason to require that every theory of incontinent action assume that simultaneously contradictory attitudes of the same type are impossible.

While the logical criterion is appealing, it is difficult to maintain when we are talking about continence or incontinence. After all, in the case of intentional yet incontinent action, it seems as though we simultaneously make the contradictory judgments that *X* is worth doing and that *¬X* is worth doing. Although we found that my belief-desire pairs of the previous example do not really contradict, we cannot diffuse this contradiction in the same way. Let us say that I judge that not eating a slice of the pie is worth doing once I have deliberated about both my belief-desire pairs. However, I find myself reaching forward, taking, and eating a slice of the pie (often quickly, before I change my mind). I do not doubt for a moment that my eating a slice of the pie is intentional so it would seem that, in some sense, I made the contradicting judgment that eating a slice of the pie is worth doing. To uphold the logical criterion, we need to find some way of distinguishing the propositional contents of these two judgments.

To satisfy the logical criterion, Davidson (1970) argues that the difference between the two kinds of judgments is that my ATCJ is conditional but my AOJ is unconditional.[[1]](#footnote-1) He claims that conditional judgments have two constitutive characteristics: they are never proximal causes of behavior (i.e. “practical only in its subject, not in its issue,” p. 39) and their contents satisfy the form “*a* is better than *b*, given a set of reasons *r*”. When an agent makes a conditional judgment that is relative to *all* reasons, Davidson calls it an “all-things-considered judgment” (“ATCJ”). He also claims that unconditional judgments have two constitutive characteristics: they are proximal causes of behavior and their contents satisfy the form “*a* is better than *b*”. Since they necessarily cause behavior, Davidson calls all unconditional judgments “all-out judgments” (“AOJ”).

According to Davidson (1970), incontinence results when the content of the AOJ is ‘(*b* > *a*)’ (thereby causing the agent to do *b*) and the content of the ATCJ is ‘(*a* > *b*, *r*)’, where the first relatum of the ATCJ is contradictory with the AOJ. Davidson satisfies the logical criterion because even though the ATCJ and AOJ aim at satisfying different sentences, their *whole* contents do not contradict: (*a* > *b*, *r*) ⊬ (*a* > *b*) and (*b* > *a*, *r*) ⊬ (*b* > *a*). Imagine that I give in and have a slice of pie. The content of my ATCJ is something like “putting the pie back in the fridge is better than having a slice of pie, given the desire that I experience the gustatory pleasure of consuming lemon meringue pie and the desire that I honor my dietary commitments”. By comparison, the content of my AOJ is something like “having a slice of pie is better than putting the pie back in the fridge”. Literally speaking, my AOJ and my ATCJ do not contradict because the ATCJ includes a clause (i.e. “given a set of reasons *r*”) that the AOJ does not.

This distinction is phenomenologically plausible. As I take a piece of pie, I think to myself *that my reasons tell me not to*. That is, when I attend to my ATCJ, it seems that I also end up attending to its relationship with my reasons. The first way of accounting for this is to build the relationship with reasons into the content of the ATCJ, as Davidson (1970) does. Then we can say that attending to the content of ATCJ involves attending to the relationship with reasons. The second way of accounting for this is to build the relationship with reasons into the functional role of the ATCJ. Then we could say that I end up attending to the context of my ATCJ whenever I attend to my ATCJ. The problem with this approach is that the contents of the ATCJ and AOJ remain contradictory if we account for the ATCJ’s relationship with reasons at the level of the function of the ATCJ rather than at the level of the content of the ATCJ. Therefore, we must account for the ATCJ’s relationship with reasons at the content-level in order to satisfy the logical criterion.

By comparison, the AOJ does not seem to have this same contextual dependence: we can attend to our AOJ without attending to its causes. For example, I tend to experience multiple reasons backing my ATCJ against a single AOJ. Although I could diagnose the causal history of my AOJ, I often don’t. On the contrary, the experience tends to be much more visceral: I typically experience the AOJ as a motivational push towards some action.

There is a challenge here, though. If the ATCJ is not “practical in issue”, as Davidson (1970) says, or not a proximal cause of behavior, then how can action be continent? We do not want to say that continence is achieved whenever the AOJ is equivalent to the ATCJ, lest we render continence a mere accident. Of course, some cases of continent action might be this kind of accident but it surely is not always, or even usually, an accident. We want to say that continence is sometimes, or usually, achieved in virtue of the ATCJ functioning as a distal cause of continent behavior.

As far as I can tell, there are three responses that we can make. The first response is to say that the ATCJ is often a distal cause of the AOJ (when the AOJ is consistent with the ATCJ) and the continent behavior that follows. To be convincing, among other things, this response needs to restate the causal relationship between the ATCJ and AOJ in terms of a series of proximal causes. The second response is that the ATCJ does not cause behavior but perhaps it is the reasons themselves (that the ATCJ concerns) that issue action by issuing an AOJ that it is consistent with the ATCJ. The third response is that we ought to keep Davidson’s distinction between the contents of conditional and unconditional judgments and discard his distinction between their causal roles. I cannot hope to do justice to comparing the three responses in this essay, so I have decided to take the first response simply in virtue of my knowing (with the benefit of hindsight)that the response is promising. (Of course, I realize that does not preclude the promise of the other two responses.)

**§2. A Dual Systems Theory of Action**

An obvious way to flesh out the first response to Davidson’s challenge would be to immediately list the series of proximal causes that constitute the distal causal relationship between an ATCJ and continent behavior. I doubt this could be done convincingly, though. Instead, I dedicate this section to introducing a dual systems theory of action and later sections to showing how the causal relationship between an ATCJ and continent behavior is embedded in this model. This is where I may deviate from Davidson (1970): although I think that one could make an argument that my theory of action is a refinement of his theory of action, I have my doubts and so I will not make an unnecessary commitment. What I am committed to, though, is that my theory of action vindicates his distinction between conditional (or ATCJ) and unconditional (or AOJ) judgments. In other words, I am only committed to satisfying the logical criterion in the way that Davidson does.

The basic idea of a dual systems theory of action is that there are two different systems, which are distinguished by their different constitutive types of processing. Evans & Stanovich (2013) argue that System 1 is constituted by autonomous processing, which makes minimal demands on working memory and does not vary between individuals, and System 2 is constituted by hypothetical reasoning, which makes significant demands on working memory and does vary significantly between individuals. In addition to these constitutive features, there are also many typically correlated features. *Given that there are many pairs of dual systems*, Evans & Stanovich note that it is only the case that System 1 *typically* consists of cognition that is fast, high-capacity, nonconscious, associative, and contextualized and that System 2 *typically* consists of cognition that is slow, low-capacity, conscious, serial, and abstract. I happen to advocate a pair of dual systems that seem typical in most of these regards.

Following Evans & Stanovich (2013), I advocate a default-interventionist dual systems theory of action. On this picture, System 1 has the default responsibility for causally determining action. Most of the time, it discharges this responsibility at very low cognitive cost and in a way that we would endorse if we were to reflect on the resulting actions. Davidson (1963) might not want to call this “intentional action” but I do, following more recent, empirically-motivated theories of action (e.g. Katsafanas 2013).[[2]](#footnote-2) Some of the time, though, System 1 discharges this responsibility in a way that we would not endorse (i.e. that System 2 would contradict) if we were to reflect on the resulting actions. When the risk of this happening is sufficiently high, a number of mechanisms are activated: e.g. we experience the meta-cognitive feeling of wrongness (Thompson 2009; Thompson, Turner, & Pennycock 2011). The activation of these mechanisms prompts System 2 to intervene and assume (at least part of) the responsibility for causally determining actions. It discharges the responsibility at a very high cognitive cost but in a way that we constitutively *do* endorse as we reflect on the resulting actions.

Even though System 2 is responsible for intervening in System 1 when the latter deviates from what System 2 would endorse, System 2 is also capable of error.[[3]](#footnote-3) The first kind of error occurs in the internal processes of System 2: e.g. in practical reasoning, in reasoning about endorsement. The second kind of error occurs in the intervening processes of System 2: System 2 fails to successfully intervene in System 1 when the latter deviates from what System 2 would endorse.

There are six parts of this model: (1) the attitudes that constitute System 1, (2) the attitudes that constitute System 2, (3) the causal functions that constitute System 1 (which count as some sort of “default”), (4) the causal functions that constitute System 2, (5) the causal influences of System 1 on System 2, and (6) the causal influences of System 2 on System 1 (which count as some sort of “intervention”). A complete and specific theory of action answers each of the following questions about each of these six parts:

1. What kind(s) of attitudes does System 1 consist of?
2. What kind(s) of attitudes does System 2 consist of?
3. How are attitudes in System 1 processed?
4. How are attitudes in System 2 processed?
5. How do attitudes in System 1 causally influence attitudes in System 2?
6. How do attitudes in System 2 causally influence attitudes in System 1?

The first two questions (1–2) are especially philosophical insofar as they evoke concern about the non-causal descriptions of the attitudes. Since most of the discussion about dual systems theory has occurred within psychology, there has been relatively little discussion about these questions.[[4]](#footnote-4) The last four questions (3–6) are especially psychological insofar as they only evoke concern about the causal descriptions of these attitudes. As we might expect, then, there has been a lot more discussion about these questions. Still, though, since less is known than is unknown, psychologists have abstained from giving complete answers. Instead, they tend to focus on Questions 3–4.[[5]](#footnote-5) This leaves the theoretically-inclined among us plenty of room to propose new, comprehensive models.

In particular, I will focus on Questions 5–6 and especially on how they relate to the answers we give for Questions 1–2. There has been some recent discussion on how System 2 contributes to System 1, such as through perspective-taking, private reflection, and cognitive reappraisal to influence intuitions (Haidt 2001, 2003; Pizarro & Bloom 2003). There has been an even older discussion amongst even the earliest empiricists about how System 1 contributes to System 2 (although not always as such). However, neither of these discussions will suffice to solve Davidson’s problem and satisfy the logical criterion unless we understand how these causal processes are related to the contents of attitudes. In §3, I build a model to describe the causal mechanisms from System 1 to System 2—I call this process “logical construction”. In §4, I build a model to describe the causal mechanisms from System 2 to System 1—I call this process “imaginative construction”.

**§3. Logical Construction**

In their influential discussion of willpower, Metcalfe & Mischel (1999) describe System 1 as “hot” and System 2 as “cool”. Of course, this temperature metaphor does not do any explanatory work but it does describe something important: we experience attitudes in System 1 differently than attitudes in System 2. In particular, we experience attitudes in System 1 as “hot”—as active—and attitudes in System 2 as “cool”—as passive. If I were to speculate about the origin of the temperature metaphor, I would say that we associate heat with activity and coolness with passivity because our body temperature rises and lowers in these respective cases. Thus, our use of the metaphor is evidence of the similarities between our experiences of attitudes in System 1 as fast and energetic and our experiences of attitudes in System 2 as slow and careful.

Metcalfe & Mischel (1999) are not the only ones to pick up on this difference. Korsgaard (1992) describes this difference in characteristically dramatic terminology: “I perceive, and I find myself with a powerful impulse… But I back up and bring that impulse into view and then I have a certain distance. Now the impulse doesn’t dominate me.” Although I do not agree with the causal elements of her description, I do think she is right to say that our attitudes in System 1 “dominate” us while we “dominate” our attitudes in System 2. In other words, prior to the intervention of System 2, we are in the “grip” of our attitudes, but once System 2 intervenes, our attitudes are in our “grip”. This describes my experience very well as I decide from different perspectives whether to have a piece of lemon meringue pie.

As both of these metaphors indicate, there is a difference in our experiences of our attitudes in Systems 1 and 2. I think it would be a mistake to say these are the same kinds of attitudes that we have different experiences of. After all, that would require us to say that our attitudes are *contents* of our experience, rather than *constituents* of our experience. Hence, this would add a layer of phenomenological complexity that we probably would not want to deal with. Instead, it is better to insist that our attitudes constitute our experiences. Then we can explain the differences in our experiences by saying that these are different kinds of attitudes altogether.

This kind of argument has been made before, in the context of representations of the future. Trope & Liberman (2003) argue that “temporal distance changes people’s responses to future events by changing the way people mentally represent those events,” (p. 403). Following Trope & Liberman, McClure & Bickel (2014) add that “construals” (as Trope & Liberman call them) in System 1 “facilitate rich emotional responses” whereas construals in System 2 lack “the elaboration that seems essential to make them viscerally compelling,” (p. 72). The point is that sufficiently different kinds of experiences of the same objects is evidence of different kinds of “construals” or *different kinds of mental representations* of the same thing. This is the same kind of argument as the one that I am making.

For these kinds of reasons, when moral psychology became more popular in philosophical circles in the 1980s, many philosophers of action accepted that there is a difference between propositional and non-propositional attitudes of approval. For instance, Street (2012) argues that desires are not constrained by instrumental rationality and the like but that value judgments are. She also suggests that desires contain a single object (to which they are aimed) whereas value judgments contain many objects: they “involve experiencing very specific features of the world as “calling for” or “demanding” or “counting in favor of” other very specific things,” (p. 44). This is consistent with Davidson’s (1970) move to say that the ATCJ (a value judgment) has “given a set of reasons r” built into its content.

This is the picture that I am suggesting: System 1 consists of non-propositional conative attitudes, such as desires, impulses, urges, etc. Although the propositional descriptions of these non-propositional attitudes are logically consistent (as per the logical criterion), they are not rationally constrained: as Korsgaard (1997) says, it is possible to desire to live and yet not desire the amputation that is necessary to live. Another feature of non-propositional conative attitudes is that they are highly motivating, in such a way that, when they are inhibited, they tend to demand our attention. By comparison, System 2 consists of propositional conative attitudes: value judgments. Value judgments are both logically consistent and rationally constrained (except in cases of dumbfounding: e.g. Haidt 2001). Setting these cases aside: Street (2012) agrees that it is impossible to value to live and yet not value the amputation that is necessary to live. Another feature of value judgments is that they are less motivating. This is often accompanied by our experiencing them with more structural and phenomenological nuance. When we reason with value judgments, we feel we do so at a distance since they do not “grasp us” and “pull us in”. These are my answers to Questions 1–2.

Now we are prepared to answer Question 5: how do desires in System 1 causally influence value judgments in System 2?[[6]](#footnote-6) That there is causal influence is presumed by the default interventionist model: in order for System 2 to intervene in System 1 in the way that we observe it doing, System 2 has to receive information from System 1. Street (2008, 2012) goes further to say that we *construct* value judgments from our desires.[[7]](#footnote-7) Refining Humean and broader empiricist intuitions, Street suggests that we construct value judgments from our desires by selecting some subset of our desires that are rationally constrained and then generating additional value judgments, as rationally required. For example, she would say that when we reflect on our desire to live and our desire not to be amputated, we form either the value judgment that living is good or that amputation is bad but not both, since our value judgments are constrained by instrumental rationality. Furthermore, if we construct the value judgment that living is good, we end up constructing the further value judgment that amputation is good, since it is a necessary means to living. Hence, I call this process “logical construction”.

As I have argued, it is not just rationality (or the lack thereof) that distinguishes desires from value judgments: it is also the motivational power. The constructive process diminishes motivation (i.e. “cools” or “creates reflective distance”) when it converts desires into value judgments.[[8]](#footnote-8) This suggests an important question, though: what is the advantage of a process that diminishes the motivational power of its inputs? Since logical construction is an adaptive process, we would expect it to offer some adaptive advantage that redeems its motivation-diminishing character.

I argue that logical construction offers at least three advantages. The first advantage, which I have already discussed, is a minimal form of rationality. The second advantage is that the construction of value judgments enables individual episodes of diachronic reasoning. Logical construction grants conative attitudes access to the superior working memory that is constitutive of System 2 (Evans & Stanovich 2013). Since working memory is short-term memory, this improves individual episodes of diachronic reasoning, in part by ensuring that they are rationally constrained. The third advantage is that the construction of value judgments enables the storage of value judgments as well as the retrieval of them when they are required for practical reasoning. The evidence for this claim is that System 2 processes (i.e. practical reasoning) often require value judgments that do not correspond to any experienced desires, whereas System 1 processes do not require desires that are not being experienced or that do not correspond to any value judgments. Thus, System 2 seems to have a greater capacity for storing value judgments than System 1 has for storing desires.[[9]](#footnote-9)

We might wonder why these advantages must come at the cost of motivational power. I can only speculate why this is so but the phenomenological evidence does suggest an answer. It is fits with our experience to say that value judgments acquire the “stability” that is necessary for System 2 to operate with and realize its adaptive advantages *in virtue of* having diminished motivation whereas desires retain the motivational power that is necessary for System 1 to operate with and realize its adaptive advantages *in virtue of* having diminished “stability”.[[10]](#footnote-10) To invoke Korsgaard (1992) again, it seems right to say that it is only possible to control our conative attitudes once we have escaped their grasp and, in turn, grasped them. Thus, my suggestion is that there is some underlying psychological necessity that creates a trade-off between motivational power and the three advantages of rationality, enhanced working memory, and stored value judgments. Logical construction makes it possible for the two systems to deal with the same material on both horns of the trade-off so that the agent, as a whole, does not have to face the dilemma.[[11]](#footnote-11)

**§4. Imaginative Construction**

Understanding logical construction helps us make better sense of Davidson’s (1970) challenge, which I discussed in §2. The ATCJ is a value judgment that is formed on the basis of other value judgments. These value judgments are logically constructed from desires in the past and present. As such, all of these value judgments have weak motivational power. It is not clear how the ATCJ can compete against the desires, which have strong motivational power. Thus, it is not clear how action can be continent, except by mistake. If logical construction is all there is of the relationship between System 1 and System 2, then the adaptive advantages of System 2 are not real advantages: rationality, enhanced working memory, and value judgment storage pose no advantage because they cannot influence action. Of course, though, logical construction is not all there is of the relationship between System 1 and System 2: there is also imaginative construction, which is the causal relationship from System 2 to System 1. Imaginative construction makes sense of the adaptive advantages of System 2 by revealing a causal relationship between the ATCJ and behavior.

There are two mechanisms by which System 2 causes action, of which imaginative construction is one and the most important. For example, let us consider Nagel’s (1970) example of the agent who treats the belief that, in the future, they will desire *x* as grounds for acting to bring *x* about now. I take Nagel’s agent to be judging that *x* is a value without desiring *x*. There are two possible cases, which involve the two different mechanisms by which System 2 causes action: (1) the value judgment is stronger (in the motivational sense) than all of his agent’s other desires or (2) the value judgment is weaker (in the motivational sense) than some of his agent’s other desires.

The first case is straightforward yet rarer. We need to elaborate on Nagel’s (1970) example a bit first, though. Assume the agent has little else to do, so that she has only few and weak desires. For example, she might weakly desire resting more than doing something. However, she thinks to herself that she has to present at a conference in a few days and she believes that she will desire to present successfully. In the present, this belief *somehow* triggers the construction of a value judgment in them that presenting successfully is good. This value judgment has as much motivating strength as a value judgment usually can and so, given that their desires are especially weak (in the motivational sense), the value judgment achieves causal efficacy and causes them to prepare for the presentation. Clearly, though, we are not usually so inactive and so most cases of continence are not going to be like this.[[12]](#footnote-12)

The second case is less straightforward yet much more common. In this case, we deliberate before or during action and we find that our desires are more motivating than our value judgments. Consider Davidson’s (1970) example of Dante’s adulterous lovers, Francesca and Paolo, who indulge in a secret affair until Francesca’s husband discovers their adultery and murders them both. Knowing her husband well, Francesca knew that adultery posed a significant risk to both her and her lover. However, Francesca was motivated more by her desire for the torrid romance than she was by her value judgment that committing adultery is worse since it might lead to the terrible outcome of being killed by her husband. As Davidson asks of us, we can even imagine Francesca incontinently indulging in her affair with full clearness of mind.

To pump our intuitions, let us consider the problem of continence phenomenologically by asking ourselves how we might turn things around to behave continently if we were in Francesca’s situation. We might have an inner monologue much like the following: “I already know that there is a significant risk that my husband will kill me if he discovers my affair with Paolo. Given this, I judge that the best thing to do is to stop my affair with Paolo. However, I know that I do not have the motivation to do this. That said, I believe that if I were to imagine my murder as vividly as I am experiencing my affair with Paolo, I would gain the motivation that I need to stop my affair with Paolo and ensure my safety by remaining faithful to my husband.”

Williams (1980) makes a similar suggestion: “[An agent] may think he has reason to promote some development because he has not exercised his imagination enough about what it would be like if it came about. In his unaided deliberative reason, or encouraged by the persuasions of others, he may come to have some more concrete sense of what would be involved, and lose his desire for it, just as, positively, the imagination can create new possibilities and new desires,” (p. 104–105). I intend to cash this anecdotal mechanism of imagination out in such a way that it involves a contribution of value judgments by System 2 to System 1.

Francesca’s hypothetical (and admittedly stilted) inner monologue is an example of what Haidt (2001) calls “private reflection”. As I mentioned in §3, Haidt follows Selman (1971) by suggesting that role-taking is “the most widely discussed method of triggering new intuitions,” (p. 819). Role-taking consists in imaginatively constructing *new* desires from *already existing* value judgments (i.e. *in* System 1, *from* System 2). In this way, *imaginative construction* is the reverse process of the *logical construction* of value judgments from desires, which I discussed in §3.

I suggest that we call “imaginative construction” what Haidt (2001) calls “role-taking”, because it explicitly mentions the mode of thought that is constitutively involved in the construction process: imagination. Without delving into the broad literature on imagination (so as to stay on course), let us simply pick out and draw from Peacocke’s (1985a) conception of imagination. According to Peacocke, imagining starts with a diachronic string of synchronic images, which we should generalize to include all sensory experiences. Next, the imaginer makes a set of suppositions about these sensory experiences. Peacocke illustrates this step by comparing two imaginations with the same images of a box that are differentiated by the imaginer supposing that there is a cat that is hidden behind the box in one imagination and not behind the other. The product of these two steps is what Peacocke calls “S-imagining”, where ‘S’ represents ‘supposition’.

Imaginative construction also consists of three steps.[[13]](#footnote-13) First, like Francesca, the agent begins with an ATCJ in favor of a particular belief-value pair, recognizes its motivational deficiency, and therefore knows that the following action will be caused by a different AOJ unless either the motivation of the ATCJ can be sufficiently increased or the motivation of the AOJ can be sufficiently decreased. Thus, the agent decides to resolve the motivational discrepancy (a resolution that follows from second-level deliberation, which I will discuss in §5).

Second, the agent does this by S-imagining two things. The agent S-imagines the first thing by (1) imagining a string of sensory experiences that corresponds to the future consequences of the action that the AOJ motivates and (2) supposing that the events that are represented in that string of sensory experiences *will* or *may* occur in the agent’s future experience. The agent S-imagines the second thing by (1) imagining a string of sensory experiences that corresponds to the future consequences of the action that the ATCJ motivates and (2) supposing that the events that are represented in that string of sensory experiences *will* or *may* occur in the agent’s future experience. The first step in both cases is actually grounded in the second: the string of sensory experiences is created to “fill in” the agent’s beliefs about what will or may occur in the agent’s future experience.

Third, these two instances of S-imagining activate the subconscious machinery that constructs desires in response to the imagined sensory experience.[[14]](#footnote-14) That is, S-imagining gives rise not only to a string of sensory experiences but also, upon the activation of perceptual processing, to a string of perceptual experiences, which constitutively include desires (Haidt 2001). Thus, beliefs about what will or may occur in the agent’s future experience cause desires to occur in the agent’s *current* experience by engaging perceptual processing through S-imagining.

Following Haidt (2001), I suggest that this process of imaginative construction converts the belief-value pair in the ATCJ into a belief-desire pair of the same kind (i.e. a belief paired with an imagination-induced desire) in the AOJ, thereby eliminating motivational discrepancies and enabling the ATCJ to “overpower” the original AOJ and cause continent action. To return to our example of Francesca: if Francesca were to S-imagine both the possible consequences of her affair and the consequences of abandoning that affair, then she would find herself with new desires, which would cause her to reject Paolo and remain faithful with her husband.

It seems quite plausible that one could agree that this S-imagining mechanism could construct desires from value judgments without agreeing that this mechanism is sufficient to *guarantee* continence. However, for reasons that will become apparent, I do not want to back down from making this claim.

Instead, consider the incontinent agent who has imaginatively constructed some weak desires from their value judgments, yet which are insufficiently motivating to cause continent action. For example, Francesca S-imagines her husband killing her and Paolo out of jealousy. This creates an aversion in her but this aversion is only weakly motivating and, thus, insufficient to stop her from pursuing her affair. Peacocke (1985b) would suggest that the problem here is an “imaginative asymmetry” between her future situation and her current situation.

Unfortunately, Peacocke (1985b) does not say what imaginative *quantity* there is an asymmetry (or difference) in. However, I think the difference he had in mind is the vividness of the imagery, where vividness is defined as the breadth (i.e. quantity) and depth (i.e. detail) of features in the mental imagery (e.g. Marks 1973). I conjecture that the vividness of sensory imagery is positively correlated with the motivational strength of the desires (or aversions). Trope & Liberman (2003) agree, saying that imagined events are less motivating when they are “represented in terms of a few abstract features that convey the perceived essence of the events (high-level construals) rather than in terms of more concrete and incidental details of the events (low-level construals),” (p. 403). This would make sense since we would expect the perceptual processing machinery to be much more engaged when processing vivid sensory imagery.

Thus, when Francesca remains incontinent even after imaginative construction, we can push back by saying that her S-imagining is insufficiently vivid to fully engage System 1. We can continue to say this until her imagination has as many relevant (i.e. motivating) details as her actual experience would/will. At this point, we might say that her constructed desires are in “fair competition” with her pre-given desires. More formally, we can say that System 2 manipulates System 1 in such a way that the AOJ is in favor of the same action as the ATCJ, even though they are not identical.

What about the case in which the AOJ remains different than the desire that corresponds to the ATCJ even when the situation that elicits that desire is imagined with maximum vividness? I do not think this is possible. After all, if Francesca imagines her murder as vividly as she experiences her affair with Paolo and yet still chooses to continue her affair with Paolo, it does not seem to make sense to say that Francesca makes an ATCJ in favor of terminating her affair. On the contrary, even though she might admit that others would condemn her for doing so, I would say that she makes an ATCJ in favor of *continuing* her affair. In other words, I think we must attribute the difference between the AOJ and the desire that corresponds to the ATCJ to the *content* of the sensory imagery (e.g. whether the represented situation is such that it affords strong desires or not) rather than to the *form* of the sensory imagery (i.e. whether the represented situation is sufficiently vivid). If this is true, as it seems to be, imaginative reconstruction can *guarantee* continence so long as the agent is able to commit and does commit to *full* imaginative construction. In other words, imaginative construction is the in-principle solution to incontinence. This makes the failure of incontinence intelligible: the agent failed at imaginative construction to a sufficient degree.

**§5. Continence & Incontinence**

Now we have the structures that we need to define continent and incontinent actions. It is already apparent that incontinent action typically occurs when there is a misalignment of System 1 and System 2, due to insufficient imaginative construction. Similarly, it is already apparent that continent action typically occurs when there is an alignment of Systems 1 and 2. In this section, I define these more precisely:

**(Def. 1) ATCJ:** an ATCJ is a conditional judgment *in System 2* that there is decisive reason for some action *A*. An ATCJ is weakly motivating.

I have already mentioned that I tentatively disagree with Davidson (1970) about this, since I suspect the ATCJ is practical in issue and he says that it is not. My reason for disagreement is that there is some psychological evidence that indicates that the ATCJ in System 2 can overpower desires in System 1 in short bursts and when the desires are sufficiently weak (e.g. Metcalfe & Mischel 1999, Holton 2003). My first interpretation of Nagel’s (1970) example illustrated the practical issue of the ATCJ.

**(Def. 2) AOJ:** an AOJ is an unconditional judgment *in System 1* that there is decisive reason for some action *A*. An AOJ is strongly motivating.

Davidson (1970) and I agree on this. However, it is worth noting that neither he or nor I have elaborated much on how the AOJ is related to desires and other elements of System 1. The only thing we need to say is that an AOJ is an intermediary variable between desires and behavior. I cannot speak on Davidson’s behalf but my reasons for doing this are so that I can avoid committing myself to a particular answer of Question 3 (just as I have avoided committing myself to a particular answer of Question 4).

**(Def. 3) “All-In” Judgment**: an AIJ is an unconditional judgment *in System 2* that there is decisive reason for some action *A*, when there is also an AOJ that is in favor of *A*.

This is my primary contribution to the discussion on incontinent action. Rather than talking loosely of an “alignment” between Systems 1 and 2, we can say that they are aligned when the ATCJ is all-in. This makes it easy to define ‘continent action’ and “incontinent action”:

**(Def. 4a) Continent Action:** an action that consists of behavior *B* is continent if and only if the ATCJ either proximally causes *B* or is all-in with the AOJ that causes *B*.

**(Def. 5a) Incontinent Action:** an action that consists of behavior *B* is incontinent if and only if the ATCJ neither proximally causes *B* nor is all-in with the AOJ that causes *B*.

There is a problem with these definitions, though, since they admit an important exception: actions that are caused by desires that System 2 would not endorse but that cause action anyways because the mechanisms that surveil System 1 failed to detect the discrepancies. For example, imagine that I am frantically pacing my apartment, thinking about a distressing philosophical problem. My attention is so engrossed in thought that I do not notice that I have done something I generally disapprove of by having a piece of pie. Energized by the uptake in blood sugar, I simply return to pacing and pondering. We want to say that I have acted incontinently but we cannot according to Definition 5a. Similarly, if I had almost had a piece of pie but thoughtlessly returned the pie to where it was without having a bite, we would want to say that I had acted continently, but we cannot according to Definition 4a. Thus, we need to modify these definitions as follows:

**(Def. 4b) Continent Action:** an action that consists of behavior *B* is continent if and only if one of the following conditions is satisfied:

1. If System 2 produces an ATCJ, then the ATCJ either proximally causes *B* or is all-in with the AOJ that causes *B*.
2. If System 2 does not produce an ATCJ, then the ATCJ that would have been produced by System 2 (had it been active) would have either proximally caused *B* or been all-in with the AOJ that would still have caused *B*.[[15]](#footnote-15)

**(Def. 5b) Incontinent Action:** an action that consists of behavior *B* is incontinent if and only if one of the following conditions is satisfied:

1. If System 2 produces an ATCJ, then the ATCJ neither proximally causes *B* nor is all-in with the AOJ that causes *B*.
2. If System 2 does not produce an ATCJ, then the ATCJ that would have been produced by System 2 (had it been active) would have neither proximally caused *B* nor been all-in with the AOJ that *may* still have caused *B*.

As a result of the pair of second conditions, both definitions can say that I have acted incontinently and continently in the modified examples, respectively. There is one thing worth noting in Definition 5b, though: although the AOJ may still have caused *B* if System 2 had produced an ATCJ, it is possible that the AOJ would not cause *B* since it is possible that System 2 may successfully alter the AOJ via imaginative construction such that it would no longer cause *B*.

**§6. Second-Level Deliberation about Imaginative Construction**

As I was explaining the process of imaginative construction in the case of Francesca, I claimed that the process began with her decision to resolve the motivational discrepancy by proceeding to the second step of imaginative construction. This gives rise to two problematic questions. First, assuming that imaginative construction counts as a so-called “mental action”, it is possible that an agent might judge that imaginative construction is the best action. Does this create the possibility that an agent could act incontinently by preventing imaginative construction? Second, if Francesca does not have enough motivation to override her desire to continue pursuing an affair with Paolo, how does she have enough motivation to initiate a process that is a means to the end of overriding her desire?[[16]](#footnote-16)

If we try to answer the first question by positing second-level dual systems and second-level imaginative construction (like we did to answer our original questions about incontinent action), then we are well on our way to generating an infinite regress. After all, if second-level deliberation can be continent or incontinent, then we could invoke a second pair of dual systems that compete over the mental action of imaginative construction and that, in turn, involve their own system of imaginative construction. This second system of imaginative construction would require its own explanation, thereby generating an infinite regress. To block this infinite regress, I claim that the second-level decision to proceed with (without) imaginative construction is necessary for continent (incontinent) action but it is not the sort of thing that can be continent (incontinent).

In principle, there is an easy way to block the infinite regress: claim that the second-level deliberation involves a *single* judgment (either an AOJ or ATCJ) within the unified operating context of a single system (either System 1 or System 2, respectively), eliminating the possibility of incontinence. The challenge, now, is to argue that this claim is true. Although doing this convincingly would require a greater effort than I have the space for, I do want to make a plausibility argument for it.

Imagine that, at the second level of deliberation, the agent tries to decide between two reasons. The first reason consists of an intrinsic pro-attitude to continence and a belief that imaginative construction guarantees continence. The second reason consists of an intrinsic pro-attitude to the incontinent course of action (its ends) and a belief that refusing imaginative construction guarantees that course of action. I cannot see how the agent will make this decision apart from merely weighing the strengths of these pro-attitudes. Accordingly, the action that ensues may be continent or incontinent but, surely, the second-level deliberation itself leads up to its own decision—to what the agent wants most, in a broad sense. Insofar as continence and incontinence are matters of taking some reasons seriously (i.e. giving them motivational power) and not others (i.e. withholding motivational power from them), I cannot see how the ultimate decision of whether to side with System 1 or System 2 cannot help but take both reasons seriously.[[17]](#footnote-17) Therefore, we have *prima facie* reason to believe that second-level deliberation is apt to neither incontinence nor continence.

If we take this model of second-level deliberation seriously, then the second question amounts to asking: how is it that the second reason I just mentioned does not always overpower the first reason? After all, if desires in first-level deliberation are so strongly motivating, then it seems strange to imagine that an intrinsic pro-attitude to continence could outcompete an instrumental con-attitude to continence on the basis of motivation alone. However, we must be careful to remember that motivation is not constrained by instrumental rationality. If we are motivated to achieve some end, we might still find ourselves lacking motivation to take the means that are necessary for achieving that end. After all, just consider how my motivation to optimize my body composition is frequently outweighed by my contravening motivation to eat vegan lemon meringue pie! Thus, at the rarified level of second-level deliberation, it is a toss-up how much motivation I will have to take the means of inhibiting imaginative construction to my incontinent end. Thus, I find that I often do have more motivation to engage in imaginative construction than I have motivation not to.

**Conclusion**

To satisfy the logical criterion, Davidson (1970) argues that we must distinguish between the logical contents of ATCJs and AOJs. Having distinguished two kinds of practical judgments, though, he does not explicate the competitive relationship between the two. In this essay, I have proposed a new theory of action that is built on the default-interventionist interpretation of dual systems models of action. According to this theory, System 1 consists of non-propositional conative attitudes, including AOJs, and which are highly motivating at the cost of irrationality, low working memory, and inferior storage of non-propositional conative attitudes. By comparison, System 2 consists of value judgments, including ATCJs, which are rationally constrained, have high working memory, and superior storage of value judgments at the cost of motivation. The relationship between the two systems is not only competitive, as one might expect, but mutually contributive: the contents of System 2 originate in System 1 via logical construction and the contents of System 1 are regulated by System 2 via imaginative construction. As a result of this dynamic relationship, the two systems can ensure that the trade-offs they individually face are not trade-offs for the agent as a whole. This model answers the original question by saying the ATCJ *typically* causes continent behavior by influencing System 1 via the relationship of imaginative construction (but allows for alternative mechanisms as well).

The solution this theory provides for Davidson’s problem raises more questions than it answers. With these new questions come new opportunities and challenges. Among these, one opportunity strikes me as especially worth mentioning. The example of my dessert cravings and Francesca’s affair are both cases in which imaginative construction is aimed at constructing desires that are aligned with prudential reasons. Can imaginative construction aim at constructing desires that are aligned with altruistic reasons? I suspect that it can. The only difference is that instead of constructing desires by imagining one’s future or possible experiences, an agent constructs other-oriented desires by imagining the future or possible experiences of other people. Thus, this theory provides promising groundwork for dual systems theories of *moral* action.

**Bibliography**

Baumeister, R. & Masicampo, E. (2010). Conscious thought is for facilitating social and cultural interactions: how mental simulations serve the animal-culture interface. *Psychological Review, 117(3), 945–971.*

Carruthers, P. (2009). An architecture for dual reasoning. In *In Two Minds: Dual Processes and Beyond*, (eds.) J. Evans & K. Frankish. Oxford: Oxford University Press.

Carter, E., Kofler, L., Forster, D., & McCullough, M. (2015). A Series of Meta-Analytic Tests of the Depletion Effect: Self-Control Does Not Seem to Rely on a Limited Resource. *Journal of Experimental Psychology: General, 144(3).*

Davidson, D. (1963/2001). Actions, Reasons, & Causes. *The Journal of Philosophy, 60(23), 685-700.* Reprinted in: Davidson 2001.

Davidson, D. (1970/2001). How is Weakness of the Will Possible? In *Moral Concepts*, ed. Joel Feinberg. Oxford: Oxford University Press. Reprinted in: Davidson 2001.

Davidson, D. (1982). Paradoxes of Irrationality. In *Philosophical Essays on Freud*, ed. Richard Wollheim & James Hopkins. Cambridge: Cambridge University Press.

Davidson, D. (1985a/2004). Incoherence & Irrationality. *Dialectica, 39(4), 345–354.* Reprinted in: Davidson 2004.

Davidson, D. (1985b/2004). Deception & Division. In *The Multiple Self*, ed. Jon Elster. Cambridge University Press. Reprinted in: Davidson 2004.

Davidson, D. (2001). *Essays on Actions & Events*, 2nd Ed. Oxford: Clarendon Press.

Davidson, D. (2004). *Problems of Irrationality*. Oxford: Clarendon Press.

Evans, J., & Stanovich, K. (2013). Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspectives on Psychological Science, 8(3), 223–241.*

Frankish, K. (2012). Dual systems and dual attitudes. *Mind & Society, 11, 41–51.*

Greene, J. (2008). The Secret Joke of Kant’s Soul. In *Moral Psychology, Vol. 3*, (ed.) W. Sinnott-Armstrong. Cambridge, MA: MIT Press.

Hagger, M. & Chatzisarantis, N. L. D. (2016). A multilab preregistered replication of the ego-depletion effect. *Perspectives on Psychological Science, 11(4), 546–573.*

Haidt, J. (2001). The Emotional Dog & Its Rational Tail: A Social Intuitionist Approach to Moral Judgment. *Psychological Review, 108(4), 814–834.*

Haidt, J. (2003). The emotional dog does learn new tricks: A reply to Pizarro and Bloom (2003). *Psychological Review, 110(1), 197–198.*

Holton, R. (2003). How is strength of the will possible? In *Weakness of Will & Practical Irrationality,* (eds.) C. Tappolet & S. Stroud. Oxford: Clarendon Press.

Korsgaard, C. (1992). *The Sources of Normativity.* Presented the Tanner Lectures on Human Values at Clare Hall, Cambridge University.

Korsgaard, C. (1997). The Normativity of Instrumental Reason. In *Ethics & Practical Reason*, (eds.) G. Cullity & B. Gaut. Oxford: Clarendon Press.

Marks, D. (1973). Visual Imagery Differences in the Recall of Pictures. *British Journal of Psychology, 64(1), 17–24.*

McClure, S. & Bickel, W. (2014). A dual-systems perspective on addiction: contributions from neuroimaging and cognitive training. *Annals of the New York Academy of Sciences, 1327(1), 62–78.*

Mele, A. (1987). *Irrationality.* New York City, NY: Oxford University Press.

Mercier, H. (2011). What good is moral reasoning? *Mind & Society, 10(2), 131–148.*

Mercier, H. & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences, 34(2), 57–74.*

Metcalfe, J. & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: dynamics of willpower. *Psychological Review, 106(1), 3–19.*

Nagel, Thomas (1970). *The Possibility of Altruism*. Oxford: Clarendon Press.

Peacocke, C. (1985a). Imagination, Experience, & Possibility. In *Essays on Berkeley: A Tercentennial Celebration*,ed. John Foster & Howard Robinson. Oxford: Clarendon Press.

Peacocke, C. (1985b). Intention & Akrasia. In *Essays on Davidson*, ed. Bruce Vermazen & Merrill Hintikka. Oxford: Oxford University Press.

Pizzaro, D. & Bloom, P. (2003). The Intelligence of the Moral Intuitions: Comment on Haidt (2001). *Psychological Review, 110(1), 193–196.*

Selman, R. (1971). The relation of role taking to the development of moral judgment in children. *Child Development, 42(1), 79–91.*

Street, S. (2008). Constructivism about Reasons. *Oxford Studies in Metaethics, 3, 208–245.*

Street, S. (2012). Coming to Terms with Contingency: Humean Constructivism about Practical Reason. In *Constructivism in Practical Philosophy*, (eds.) J. Lenman & Y. Shemmer. Oxford: Oxford University Press.

Thompson, V. A. (2009). Dual-process theories: A metacognitive perspective. In *In Two Minds: Dual Processes and Beyond*, (eds.) J. Evans & K. Frankish. Oxford: Oxford University Press.

Thompson, V., Turner, J., & Pennycock, G. (2011). Intuition, Reason and Metacognition. *Cognitive Psychology*, *63*, 107–140.

Trope, Y. & Liberman, N. (2003). Temporal Construal. *Psychological Review, 110, 403–421.*

Wilson, T., Lindsay, S., & Schooler, T. (2000). A Model of Dual Attitudes. *Psychological Review, 107(1), 101–126.*

1. Davidson elaborates on his endorsement of this criterion in many later papers (e.g. 1982, 1985a, 1985b). [↑](#footnote-ref-1)
2. Katsafanas (2013) argues that, since most of our behavior is caused by System 1, we need to redraw the line between what counts as agentially active and agentially passive. Otherwise, we are forced to make the strange conclusion that we *qua* agents are rarely active. He proposes that agential activity involves acting on motives that one would endorse if one were to reflect on them (and agential passivity as the opposite of that). [↑](#footnote-ref-2)
3. It is often mistakenly believed that System 2 is immune to error (Evans & Stanovich 2013). [↑](#footnote-ref-3)
4. Having said that, I should note that there is a related discussion in the literature on dual attitudes (e.g. Wilson, Lindsay, & Schooler 2000; Carruthers 2009; Frankish 2012). [↑](#footnote-ref-4)
5. In moral psychology, for example, a significant part of the discussion is whether System 1 processes and System 2 processes are “emotional” and “rational”, respectively (Haidt 2001, 2003; Pizarro & Bloom 2003), or deontological and consequentialist, respectively (Greene 2008, Mercier 2011). [↑](#footnote-ref-5)
6. This question actually concerns every non-propositional conative attitude (not just desires) but I will refer to that motley bunch simply as “desires” for simplicity. [↑](#footnote-ref-6)
7. As a disclaimer, I should emphasize Street (2008, 2012) is not clear about whether she means this. [↑](#footnote-ref-7)
8. Since Davidson (1970) suggests that the ATCJ is theoretical in issue, he would probably want to say that this construction process eliminates the motivational power of System 2 value judgments (that is, if he were to get on board with the general framework that I have proposed). I do not want to make such a strong claim, though, for reasons that I will discuss in §4. [↑](#footnote-ref-8)
9. Another advantage of logical construction is that it allows the agent to appeal to deeply-held attitudes in System 1 when providing logically consistent and rationally constrained justifications for action to other agents (Haidt 2001, Baumeister & Masicampo 2010, Mercier & Sperber 2011). [↑](#footnote-ref-9)
10. Often, the values might stabilize with different relative strengths than the original values. This makes sense of Mele’s (1987) claim that “the motivational force of a want may be out of line with the agent’s evaluation of the object of that want [i.e. the corresponding value],” (p. 37). [↑](#footnote-ref-10)
11. This makes synchronic reduction impossible: the value judgments in System 2 at any given time can only be reduced to desires in System 1 over some interval of time. For example, I might consider vegan lemon meringue pie to have value for me (in a passive way) even though I might not be desiring them at this particular moment (in an active way). [↑](#footnote-ref-11)
12. There is some evidence that suggests that there is another mechanism that can directly “boost” the motivational strength of value judgments in System 2 (cf. Carter, Kofler, Forster, & McCullough 2015; Hagger & Chatzisarantis 2016). Holton (2003) argues that this mechanism is best thought of as a metaphorical “muscle” that intervenes by overpowering our desires (in System 1) but which consumes a great deal of energy, is rapidly fatigued, requires a significant amount of time to recover, and can gain strength and endurance through practice. In my view, these characteristics make this willpower mechanism suitable for emergency interventions, but not suitable for the quotidian continence that I am interested in. [↑](#footnote-ref-12)
13. I describe these steps in deliberative terms as a matter of clarity rather than commitment. After all, I do believe that imaginative construction happens automatically and even subconsciously. [↑](#footnote-ref-13)
14. Note that I do *not* claim that just any kind of supposing can cause S-imagining to activate the valenced, sensorimotor perceptual processing that gives rise to strongly motivating desires. I only claim that supposing about what one takes to be a possible or actual future can activate that perceptual processing that gives rise to strongly motivating desires. [↑](#footnote-ref-14)
15. This definition bears many similarities to the notion of “agential activity” by Katsafanas (2013). [↑](#footnote-ref-15)
16. Note that motivation is not constrained by instrumental rationality so we can know that this question is answerable, at least in principle. I am grateful to one of my anonymous reviewers for encouraging me to answer this second question. [↑](#footnote-ref-16)
17. Note that this raises questions about what operating context second-level deliberation occurs in. [↑](#footnote-ref-17)