



RECONSIDERING RESOLUTIONS

BY ALIDA LIBERMAN

JOURNAL OF ETHICS & SOCIAL PHILOSOPHY

VOL. 10, No. 2 | MAY 2016

URL: WWW.JESP.ORG

COPYRIGHT © ALIDA LIBERMAN 2016

Reconsidering Resolutions

Alida Liberman

MANY OF US ENGAGE IN THE FAMILIAR PROCESS of making resolutions, often at the start of a new year. We commit ourselves to following certain courses of action that we very much want to achieve but expect may be difficult, such as procrastinating less or exercising more. In *Willing, Wanting, Waiting*, Richard Holton lays out a detailed account of resolutions, arguing that they enable agents to resist temptation. Holton claims that temptation often leads to inappropriate shifts in judgment, and that resolutions are a special kind of first- and second-order intention pair that blocks such judgment shift. In this paper, I elaborate upon an intuitive but underdeveloped objection to Holton's view – namely, that his view does not enable agents to successfully block the transmission of temptation in the way that he claims, because the second-order intention is as equally susceptible to temptation as the first-order intention alone would be. I appeal to independently compelling principles – principles that Holton should accept, because they help fill an important explanatory gap in his account – to demonstrate why this objection succeeds. This argument both shows us where Holton's view goes wrong and points us to the kind of solution we need. In conclusion, I sketch an alternative account of resolutions as a first-order intention paired with a second-order *desire*; I argue that my account is not susceptible to the same objection because a temptation that cannot be blocked by an intention can be blocked by a desire.

1. Holton's View of Resolutions

Although other philosophers have offered philosophical accounts of resolutions, Holton's account is the best developed and most influential.¹ Holton

¹ For example, see McGuire (1961), who argues that you can *decide* to do something only if you know that you are able to do it, but can *resolve* to do something when you recognize that there is a chance of failure. McClennen (1990) outlines a notion of *resolute choice* in which the chooser commits herself to a plan that seems best at the time of initial choice, and then sticks with this plan regardless of whether it continues to seem best to her; his notion is elaborated upon and modified by Gauthier (1997). Cohen and Handfield (2010) argue that *resolve* is a distinct rational capacity to cease or suppress deliberation, and sketch a non-consequentialist, virtue-based account of resolutions.

It is a common idea among nonphilosophers that a resolution is a promise that you make to yourself. Several philosophers who defend the existence of promises to the self claim in passing that resolutions are a type of self-promise; for example, see Raz (1977: 210), Dannenberg (2015: 159), Habib (2009: 545n) and Rosati (2011: 124-25). However, the existence of genuine promises to the self is controversial. Some argue that self-promises do not exist because we have no moral obligations to ourselves at all; see Singer (1959). Others argue that we can have moral obligations to the self, but that the possibility of promisee release

understands resolutions as a special type of intention. He builds on Michael Bratman's (1987) prominent account of intention, according to which intentions are non-reducible mental states that aid in inter- and intrapersonal planning and coordination. Typically, intentions are effective at bringing agents to act.

Sometimes, though, temptations of various sorts threaten to sway an agent from her intended course of action. Holton argues that temptation "frequently works not simply by *overcoming* one's better judgement, but by *corrupting* one's judgement," noting that "as a matter of empirical fact, temptation normally induces judgment shift" (97). In other words, an agent who succumbs to the temptation to smoke another cigarette in spite of an intention not to do so does not typically maintain her judgment that she ought not smoke while akratically lighting up anyway. Rather, in many cases her judgment about smoking shifts; she *changes her mind* about whether to smoke, making smoking in line with her current judgment about what is best.

Holton argues that this sort of judgment shift leads to an agent's judgment becoming "either corrupted or powerless" in such a way that it "cannot be the motor of resistance" to temptation (98). He proposes that what can function as the source of resistance in such cases is a *resolution*, cashed out as a pair of intentions, one first order and the other second order. Resolutions consist in "both an intention to engage in a certain action, and a further intention not to let that intention be deflected" (11). So an agent who *resolves* to quit smoking both intends not to have any more cigarettes and intends not to reconsider the intention not to have any more cigarettes. This second-order intention prevents judgment shift. For if the agent does not allow herself to reconsider whether she should hang on to her intention to avoid cigarettes, her judgment will not have the opportunity to shift, and she will be likely to act as she originally intended.

An important and insufficiently appreciated insight in Holton's account is that temptation leads to judgment shift not merely through some brute causal process. Rather, a mental process is involved that, while not fully rational, allows the agent who succumbs to temptation to make sense of her action; as a term of art, I call this a *sense-making* process. Holton suggests that one major cause of judgment shift is that agents typically foresee that they are likely to succumb to temptation. They want to see themselves in a good light, and avoid the cognitive dissonance of believing that it would be best to refrain from Φ ing while believing that they will Φ anyway. An easy way to resolve this dissonance is by changing the belief that it would be best to refrain from Φ ing. This method of resolving cognitive dissonance by shifting judgment is not rational, but it at least allows the agent to make sense of her own action.

makes promissory obligations to the self implausible; for example, see Hills (2003) and Liberman (unpublished manuscript).

However, more needs to be said. Holton's appeal to the avoidance of cognitive dissonance leaves unanswered the more fundamental question of *why* agents predict that they will succumb to temptation in the first place. Why are we likely to act as we are tempted to, even when we have explicitly intended not to so act? This question is independently interesting; temptation is a frequent and frustrating part of our everyday experience, and we want a complete philosophical understanding of how temptation works. But the question is of particular importance for Holton, for the answer will help explain why temptation leads agents to revise their first-order intentions in the first place, and Holton's account of resolutions is motivated by the need to prevent such revision. I return to this question in Section 4.²

2. The Efficacy Objection

One might wonder why adding a second-order intention makes any difference for preventing judgment shift. We can call this the *efficacy objection*: why is a resolution more effective at resisting temptation than a first-order intention alone would be? Holton's answer is an appeal to *willpower* as the means by which agents resist temptation and achieve strength of will. For Holton, willpower is "substantial; it is at least a skill and perhaps a self-standing faculty, the exercise of which causally explains our ability to stick to a resolution" (112). Exerting willpower requires effort. Like other faculties, willpower can be built up by exercise or depleted by overuse.

But even if we granted that agents possess some faculty of willpower and we perfectly understood how this faculty works, a worry about efficacy would remain. Sarah Paul characterizes Holtonian temptation as "essentially involv[ing] the tantalizing action's petitioning of one's awareness with a practical proposal" (2011: 891). She argues that this seems to be a problem for Holton, for "surely this already suffices for the practical question to be raised" (891). And once the practical question is raised, the answer is obvious, for "once temptation has set in, *of course* the pleasure afforded by [for example] the wine outweighs all other considerations. Revision will be well-nigh automatic (891)." Paul goes on to suggest that, because temptation pre-

² I am presuming in this discussion that diachronic norms of rationality govern our resolutions, such that something goes wrong with you if you abandon your resolution without good reason. There is currently a live philosophical debate about whether irreducibly diachronic norms of (epistemic and/or practical) rationality exist (for example, see Christensen 2000, Ferrero 2012, Hedden 2015 and Moss 2015 for arguments that no such norms exist, and Bratman 1987, Holton 2009, Broome 2013, Hlobil 2015 and Carr 2016 for arguments that they do). However, I do not need to delve into this debate for present purposes, since one could translate Holton's (and my) claims about the stability of resolutions into a "time-slice rationality" framework that denies the existence of irreducible diachronic norms of practical rationality. For example, Hedden proposes that the stability of intentions is a causal fact, and that "if you have background evidence that you typically are rational in what intentions you form, and reconsidering intentions often stems from temptation, it will be rational not to re-open deliberation unless you have strong evidence that this case is different" (126).

sents the alternative action so vividly, it will push for revision of not only the first-order intention but also the entire resolution.

In other words, it seems that the second-order intention should succumb to the same temptation to which the first-order intention is susceptible. John Maier (2010) articulates this worry as follows:

For any considerations in favor of revising one's first-order intention would presumably *also be grounds for revising one's second-order intention to not let that intention be deflected*. But then it does not seem that a higher-order intention of this sort will be guaranteed to be any more rationally robust than a first-order intention – it is, as it were, just another intention (my emphasis; 361).

It might seem that not even willpower can prevent the transmission of temptation from the first-order intention to the second-order intention. Why do the very same considerations that tempt you toward watching yet another episode of your favorite TV show – say, your burning desire to find out what happens next, and your aversion to working – not *also* tempt you to reconsider your resolution to turn off the TV and get to work on your paper, regardless of whether you possess willpower?

Something is intuitively compelling about this objection; philosophers before me have noticed it, and it seems *prima facie* plausible. But as of yet, the objection is underdeveloped. For we have not seen *why* the temptation that pushes on the first-order intention pushes on the second-order intention as well. What causes the temptation to transmit in this way? I will answer that question in the next two sections of this paper, thereby making what was previously merely an intuitive objection more rigorous. In doing so, I will fill in the explanatory gap mentioned at the end of Section 1, above, and explain why agents are likely to abandon their first-order intentions in light of temptation.

When thinking about how to precisify the efficacy objection, it is important to realize that the first- and second-order intentions involved in a resolution are about different things: one is an intention *to* Φ , and the other is an intention *to refrain from reconsidering* the intention *to* Φ . We cannot simply assume that a temptation to refrain from Φ ing will also be a temptation to reconsider an intention *to* Φ . For, at least in principle, these two different intentions could be susceptible to entirely different temptations. And if this were so, a temptation that leads to judgment shift at the first-order level would not automatically lead to judgment shift at the second-order level as well.

Rather, I will construct a *Temptation Transmission* argument to show that a temptation to abandon a first-order intention will automatically *also* yield an equally strong temptation to abandon any second-order intention about maintaining that first-order intention. The argument relies on certain background principles, which I lay out in the Section 3 before constructing the argument in Section 4. The first stage of *Temptation Transmission* explains why

the tempted agent feels pressure to act contrary to her first-order intention in the first place, thereby answering the general question about temptation posed at the end of Section 1. Holton must accept an explanation like this one in order to motivate his account. But the second stage of the argument entails that any temptation that puts pressure on an agent's first-order intention necessarily *also* puts pressure on her second-order intention – which means that the second-order intention will not reliably block the very temptation it was designed to circumvent, and the resolution will be no more effective at preventing judgment shift than a first-order intention alone would be. Since both stages of the argument employ similar reasoning and draw on the same plausible principles, Holton cannot easily resist the second stage without also rejecting the first. And Holton has good reason to accept the first, as it fills in a gap in his account.

3. Background Principles for the Temptation Transmission Argument

3.1. Temptation Claim

The starting point for *Temptation Transmission* is the following empirical claim:

Temptation Claim: Temptation works by altering the appearances in favor of there being a reason to do the tempting thing, from the agent's perspective.

Temptation Claim is a general formulation of the empirical understanding of temptation to which Holton is committed. Holton claims that what causes people to yield to temptation is desire in a certain sense. The sense of desire he has in mind is very like Scanlon's (1998) concept of *desire in the directed-attention sense*, according to which a "person's attention is directed insistently toward considerations that present themselves as counting in favor of" the desired object or action (Holton 2009: 39). Assuming (as Scanlon does) that a reason for X is a consideration that counts in favor of X, desire in this sense directs the agent's attention toward apparent reasons for X. Since temptation leads to desire in this sense, temptation draws the agent's attention to apparent reasons for the tempting thing.³

Temptation Claim captures the phenomenology of temptation well. Sometimes temptation clues the agent in to genuine reasons that she has to

³ Holton notes that the features of desire that Scanlon points out "may be necessary for desire to arise; but they do not constitute it" (102). He suggests that "what is missing in Scanlon's characterization is the idea that desire *pulls* me to a certain course of action," concluding that the sense of desire he is after "is a state that preoccupies an agent's attention with an urge to perform a certain action" (102). Temptation Claim makes sense of this: the agent's attention is directed toward an apparent reason, and the apparent reason is what creates the urge to act. Since agents (typically) act for reasons, this is a natural picture. And if we did not cash out the urge in terms of reasons, it is unclear what else could be motivating or driving the agent.

do the tempting thing; an agent who is tempted to abandon her resolution to go to the gym at 7 a.m. because she feels extremely sleepy when her alarm goes off really does have a reason to sleep in, albeit likely a weaker reason than the reason to go as planned. In such a case, the temptation might make the reasons to stay in bed seem stronger to the sleepy agent than she takes them to be upon reflection, when not being tempted.

But temptation can also make it seem to the agent like there is a reason for her to do the tempting thing when by her own lights she has none. For example, consider a recovering alcoholic who is committed to maintaining his sobriety. He is feeling tempted by a bottle of whiskey. He does not take himself to actually have a reason to drink. Rather, he takes himself to have strong reasons not to. But in the moment of temptation, it nevertheless *seems* to him as if he has a reason to drink. The appearances of what he has reason to do have altered, even though his actual reasons have not.

Moreover, it seems likely that agents who foresee the need to form a resolution in anticipation of a certain temptation also foresee that the temptation might alter the appearances in favor of there being a reason for them to do the tempting thing. For if the temptation did not seem to provide them with reasons, then they would not have needed to form resolutions in the first place; merely intending to act would have sufficed. A temptation that is serious enough to motivate an agent to form a resolution will seem to present that agent with reasons for acting.

3.2. Rational Means Reasons Transmission

Temptation Transmission relies in large part on the philosophically commonplace and highly intuitive observation that reasons for ends transmit to the necessary means to those ends.⁴ In particular, I rely on a subjective version of the principle, which concerns not what reasons there are, but what reasons a rational agent will take there to be. The transmission principle I rely on is also broader than most, in a way that will play an important role in the arguments to come:

Rational-Means Reasons Transmission (RMRT): Where E is an intentional action, if it appears to a rational agent A that (1) there is a reason of strength X for an agent A to attain end E, and (2) M is the only rationally permissible way to attain E, then there will appear to A to be a reason of strength X for A to do M.

RMRT states that when a rational agent takes herself to have a reason to take some end – say, a reason to travel from L.A. to New York for a long week-

⁴ Versions of a necessary-means reasons-transmission principle have been formulated and defended in slightly different ways by Nagel (1970: 34), Darwall (1970: 16), Kolodny (2007: 251), Schroeder (2009: 245), Bratman (2009: 424) and Way (2012: 494), among others, none of whom formulate the principle in the way I do here.

end – then she will also take herself to have an equally strong reason to take whatever means are necessary for attaining that end in a rationally permissible way – say, buying a plane ticket. Several special features of this principle are worth highlighting. First, RMRT includes a commonly accepted constraint that the strength of the reason transmitted from end to means remains constant, such that an apparent reason of strength X for the end yields an apparent reason of strength X for the sole rationally permissible means.⁵

Second, RMRT allows reasons to transmit in a broader range of cases than do typical necessary-means reasons-transmission principles. Under RMRT, a means *counts as necessary* for the purposes of reasons transmission if the means is the only rationally permissible option. Buying a plane ticket is not *strictly* necessary to get from L.A. to New York for a long weekend – you could drive, or board the plane with a stolen or counterfeit ticket. But neither of these other options makes any sense; driving takes far too long to make a weekend trip worthwhile, and flying with a stolen or counterfeit ticket is far too difficult and risky. RMRT is broader than standard means-end transmission principles, which state that reasons transmit from end to means only if the means is strictly necessary. But RMRT is not overly broad, for it restricts the reasons transmission to the *sole* rationally permissible way to attain an end. For example, suppose you have a reason of strength X to pay for an item at the store. You can pay with cash, credit card or check. Your credit card is maxed out, leaving paying with cash or check as rationally permissible options for you. RMRT does not predict that you have a reason of strength X to pay by check *and* a reason of strength X to pay with cash. Rather, RMRT strikes a balance between being more permissive than *necessary-means* reasons-transmission principles, but less permissive than *sufficient-means* reasons-transmission principles.⁶

Third, unlike most reasons-transmission principles, RMRT is concerned not with what (objective) reasons there are, but with what (subjective) reasons agents take there to be. I have already argued that temptation provides agents not necessarily with actual reasons to act, but with the *appearance of* reasons to act. It follows that if we are to employ RMRT in the *Temptation Transmission* argument, RMRT must apply to the transmission of the appearance of reasons. Accordingly, RMRT tells us that if agents are rational, reasons will *appear to them* to transmit from means to sole permissible end.⁷ For

⁵ Kolodny (forthcoming) argues that reasons will not always transmit full strength to necessary means. But his counterexample cases turn on features that do not matter for the ways in which I will be applying RMRT. So while this complication is important for a general account of reasons transmission, I may gloss over it here without problem.

⁶ For arguments that reasons transmit to sufficient means as well as necessary means, see Bedke (2009: 687n) and Way (2010: 224). For a detailed discussion of whether means-end reasons-transmission principles should be necessary or sufficient, see Kolodny (forthcoming), who concludes that the means must be non-superfluous, which rules out merely sufficient means.

⁷ My claim here is narrow, and concerns only whether an apparent reason to attain an end transmits to the means necessary for rationally attaining that end. The truth of this claim has

RMRT to apply, the agent in question need not be a rationally perfect agent; few (if any) of us can meet that demanding criterion. Rather, RMRT will hold if the agent is rational with regard to the means-end reasoning process in which she is engaged. That is, RMRT will apply to agents who are capable means-end reasoners – even if they have not perfectly proportioned all of their credences to their evidence, or if they hold one pair of incompatible beliefs, or if they fail to be perfectly rational in some other way. As I will argue in the next section, *Temptation Transmission* applies to agents involved in a sense-making process regarding their own behavior, including their own means-end reasoning. In order to participate in such a process, agents must be minimally capable means-end reasoners, as required by RMRT.

To highlight the independent plausibility of RMRT, consider a series of examples. First, suppose a just country requires all of its citizens to either join the military for two years or perform alternative service. Smith correctly takes herself to have a reason to satisfy this legal requirement. Because Smith is a principled and committed pacifist, joining the military is rationally impermissible for her. Accordingly, the only rationally permissible way for Smith to attain the end of satisfying her legal requirement is by performing alternative service. It follows that, so long as she is a capable means-end reasoner, Smith will take herself to have a reason to perform alternative service.⁸ Second, suppose that Smith's country is unjust, and that she does not in fact have good reason to obey her country's laws – or, suppose that she has good reason to obey the law, but that performing alternative service is not in fact the only rationally permissible way for her to do so (say, because she mistakenly believes that her religion commits her to pacifism when in fact it commands

no impact on broader debates in epistemology about whether all of our reasons are always apparent to us, or whether all of our mental states are luminous (i.e., such that we are always in a position to know what they are). The transmission of apparent reasons from end to means is compatible with the existence of (subjective or objective) reasons that are not apparent.

⁸ The case of pacifist Smith choosing between joining the military or performing alternative service has been discussed by Horty (1993: 73), Goble (2004: 80) and Nair (2014: 2). Following Horty, these authors all formulate the example as an instance of a pattern of good or valid reasoning involving disjunction: if you have reason to do A vs. B, and you have reason not to do A, then you can conclude that you have reason to do B. RMRT as formulated restricts reasons transmission only to *means*. But the principle could instead be formulated as an instance of the following broader disjunctive principle – call it Disjunctive Reasons Transmission (DRT): if (1) you take yourself to have a reason of strength X to do (A vs. B), and (2) you take A to be rationally impermissible, then (C) if you are rational, you will take yourself to have a reason of strength X to do B.

The objective version of DRT faces a worry about reasons explosion: if it is possible for there to be a reason to do A *and* for A to be rationally impermissible, we can generate that you have a reason to do B, for any B. The same worry will apply to the subjective version of DRT if it is possible for an agent to rationally take there to be a reason for her to do A, and for her to rationally take A to be rationally impermissible. However, even if this is possible, I avoid explosion by restricting DRT to range only over means, thereby preventing *any* B from being included in the disjunction.

her to fight in just wars). Nevertheless, if it *seems to Smith* that she has good reason to obey the laws of her country, and it *seems to her* that the only rationally permissible way to do so is performing alternative service, then (assuming Smith is a capable means-end reasoner) it will also *seem to her* that she has a good reason to serve.

Similarly, suppose a recovering alcoholic is clear-sighted enough to recognize that he does not really have a good reason to drink whiskey. But the tempting scotch in front of him makes it appear to him that he does. If he believes that opening the bottle is the only rationally permissible way for him to have a glass of scotch, it will also seem to him that he has a good reason to open the bottle. It is not clear how we would explain this if we denied that RMRT applied to the appearance of reasons. These cases show us that the appearance of reasons transmits from end to sole permissible means in rational agents, even if the agent is mistaken about what reasons there in fact are.

Temptation Transmission starts with Temptation Claim, and then applies RMRT to a series of independently compelling claims about which mental states are the only rationally permissible means to rationally holding other mental states. Agents who follow the steps laid out in *Temptation Transmission*, while not fully rational, engage in a sense-making process that enables them to see themselves in a good light. As we shall see in the next section, the argument concludes that temptation necessarily transmits – not only to first-order intentions, as Holton claims it does, but to resolute second-order intentions as well. Moreover, seeing the exact way in which temptation transmits will show us what sort of view we need to stop the transmission.

4. The Temptation Transmission Argument

The following three claims form the basis of *Temptation Transmission*:

- 1) When Φ ing is an action about which an agent has formed a resolution, an apparent reason to Φ (stemming from temptation) necessarily leads to an apparent reason to intend to Φ .
- 2) An apparent reason to intend to Φ necessarily leads to an apparent reason to reconsider the intention to avoid Φ ing (call this intention “Intention 1”).
- 3) An apparent reason to reconsider Intention 1 necessarily leads to an apparent reason to reconsider the intention not to reconsider Intention 1 (call this intention “Intention 2”).

If these three claims are true, Holton’s account of resolutions is in trouble, for Intention 2 will not successfully block temptation and prevent reconsideration after all. The first two claims by themselves are not problematic for Holton. To the contrary, Holton’s argument needs claims of this sort: they answer the question from Section 1, and explain why temptation is likely to

lead to judgment shift and abandonment of the first-order intention in the first place, such that we need something additional to prevent such shift and abandonment. The third claim is what puts pressure on Holton's account – but, as we shall see, this claim is difficult to resist while accepting the first two, as it invokes similar reasoning and draws on principles that Holton accepts.

The three claims above can be defended by appealing to RMRT and to the relationships between means and rationally permissible ends among various sorts of mental states. I will work through these claims one at a time, before putting the entire argument together. (While I lay out my discussion in terms of intending to Φ in spite of a resolution to avoid Φ ing – e.g., resolving to avoid eating sweets and being tempted to eat a donut – we could just as easily cash things out in terms of a positive resolution and a temptation to do something else – e.g., resolving to eat only healthy food and being tempted to eat a donut.)

4.1. Defending Claim 1

The empirical claim about temptation states that a temptation to eat a donut makes it seem to the agent like she has a reason to eat the donut. RMRT states that reasons transfer from ends to sole, rationally permissible, sufficient means. If it appears to the agent that intending to violate her resolution by eating a donut is the only rationally permissible way to do so, we will have established Claim 1: that an apparent reason to abandon one's resolution by eating the donut will always lead to an apparent reason to *intend to eat* the donut as well.

Surely intending to Φ is not always *necessary* for Φ ing; many of our actions are unintentional. Nor is it always the case that intending to Φ is necessary for *rationally* Φ ing; habitual actions such as using a turn signal while driving can be rationally performed unintentionally. But it makes sense to form resolutions only about actions that are in your direct intentional control; if you know that you will automatically reach out to deflect a ball that is thrown at you, it is silly for you to resolve not to deflect it, and superfluous for you to resolve to deflect it. Claim 1 states that when Φ ing is an action that is in your intentional control, intending to refrain from Φ ing is the only rationally permissible way to refrain from Φ ing. In other words, you can *rationally* fail to satisfy a resolution only if you do so *intentionally*.

Certainly, it is *possible* to eat the donut and fail to satisfy your resolution without intending to do so. We often fail to keep our resolutions in spite of ourselves by acting mindlessly and not forming any intentions about what we do, or by acting akratically in a way that is contrary to our intentions. But these ways of failing to keep a resolution are not *rational*. To fail to keep your resolution in a way that allows you to make sense of your own behavior, you must choose to abandon the resolution; that is, you must do so on purpose,

or intentionally. Mindless or akratic behavior does not enable you to tell a coherent story about your own actions in the way that intentional behavior does.

To illustrate with another example: an agent who resolves not to watch yet another episode of her favorite show might let the next episode play because of apathy or mindlessness. But if she is to watch another episode in a way that allows her to make rational sense of her own actions, she must *intend* to let another episode play. Temptation makes it seem to the agent like there is a reason to watch another episode. It also seems to the agent that *intending* to watch another episode is the only rationally permissible way for her to give up on her resolution and watch the next episode, because this is the only way she can make sense of her own behavior. And so, by RMRT, there will appear to the agent to be a reason to intend to watch another episode.

4.2. Defending Claim 2

Claim 2 states that an apparent reason to intend to Φ necessarily leads to an apparent reason to reconsider the intention to avoid Φ ing. The argument for this claim proceeds in two steps. First, I argue that rational agents will believe that giving up any intention that you have not to Φ is the only rationally permissible way to intend to Φ . By RMRT, an apparent reason to intend to Φ will therefore lead to an apparent reason to abandon the intention not to Φ . Second, I argue that rational agents will also believe that reconsidering an intention is the only rationally permissible way to abandon it. And so again, by RMRT, an apparent reason to abandon the intention not to Φ will therefore lead to an apparent reason to reconsider that intention.

The first means-end claim is true because of the following commonly accepted and intuitively obvious constraint on intentions:

Intention Consistency Constraint: It is not rational for an agent to intend to Φ at time t and to simultaneously intend not to Φ at time t .⁹

When intentions work properly, they bring agents to action in accordance with them. Since it is not possible to simultaneously eat a donut and not eat a donut, it does not make sense to simultaneously intend to do both of these things. An agent who simultaneously intends to eat a donut and intends not to eat a donut must therefore abandon at least one of these intentions if she is to be rational. It follows from this that giving up any intention to avoid eating a donut is the only rationally permissible means to intending to eat the

⁹ Philosophers ground this constraint differently; for example, cognitivists about intention (such as Gilbert Harman and David Velleman) offer different explanations than those who endorse planning theories of intention, like Bratman. But there is a general consensus that some such consistency constraint exists. See Bratman (2009a) for a detailed discussion of this.

donut – and agents who are capable means-end reasoners will recognize this. Giving up the intention to avoid the donut might not be something that the agent must do in order to intend to eat the donut at all; it might be possible to hold conflicting intentions. But since it is not possible to *rationally* do so, we can apply RMRT to conclude that a reason to intend to Φ yields an equally good reason to abandon the intention not to Φ .

The second means-end claim states that capable means-end reasoners will believe that agents cannot rationally revise or abandon intentions unless they first *reconsider* those intentions. This is true to our everyday experience. Suppose that, when I wake up, I intend to go to a yoga class at 8 p.m. But I procrastinate all day, and do not finish the work that I hoped to complete. It would not be rational for me to go *immediately* from being in this state to being in the state of abandoning my intention to go to yoga. To the contrary, I have to think about the new state of affairs, temporarily shelving the intention to go to yoga and deliberating anew about what to do. Do I go to yoga as planned and not get any work done today? Or do I skip yoga to do work? If I did not go through some sort of deliberative process like this, my revision would be entirely out of the blue. If I am lucky, this revision will be in accordance with the rational course of action. But since I will not always be lucky, this is not a rational disposition to have in general. In general, some sort of reconsideration must always occur, lest agents revise inappropriately.

Reconsideration of an intention need not be a long or drawn-out process, and it need not involve difficult deliberation. If I intend to bake a cake, and my oven catches fire while it is preheating, I need not *deliberate* about whether to attempt to bake a cake in the flaming oven; I can decide immediately that baking a cake no longer makes sense. But I will go through some process of revision, even if very quick. One might worry that it is possible to rationally abandon an intention without first reconsidering it. Sometimes our intentions simply disappear or fall away; for example, you might intend to go to a concert at the end of the week and forget about it before the day arrives, without any irrationality. But when we are dealing with an intention about which you have made a resolution, this is not the case. If you *resolve* to go to the concert at the end of the week, something does go wrong if you simply forget about it. So we can apply RMRT to conclude that a reason to abandon the intention to Φ leads to a reason to reconsider the intention to Φ .

So far, *Temptation Transmission* has explained why a temptation to Φ necessarily leads, via a sense-making process, to the appearance of a reason to reconsider the intention not to Φ . This answers the question from Section 1 about why agents predict that they will be likely succumb to temptation. When an agent intends not to Φ and is tempted to Φ , there appears to her to be a reason to abandon her intention. She predicts that she will succumb to temptation because she predicts that she will respond to this apparent reason. *Temptation Transmission* starts with an account of temptation that Holton accepts, and then relies on plausible and widely accepted general principles to

establish an explanation of the sort Holton needs to motivate his account of resolutions. Up until this point, then, Holton should be sympathetic to *Temptation Transmission*.

4.3. Defending Claim 3

The core of the defense of Claim 3 is the claim that rationally reconsidering an intention not to Φ (what I will call Intention 1) requires abandoning the second-order intention not to reconsider that intention (what I will call Intention 2). From this claim, we apply RMRT to establish that an apparent reason to reconsider Intention 1 will always lead to an apparent reason to abandon Intention 2. As we saw in Section 4.2, abandoning an intention in a rationally permissible way requires first reconsidering it. This will establish Claim 3: that an apparent reason to reconsider Intention 1 necessarily leads to an apparent reason to reconsider Intention 2.

Why should capable means-end reasoners believe that the only way to rationally abandon Intention 1 is to reconsider the second-order intention not to reconsider Intention 1? We can best answer this question by thinking about what it would be like for an agent to reconsider an intention *without* first giving up her intention not to do so. Suppose that I intend not to eat any dairy products, and I intend not to reconsider this intention. As I am grocery shopping, I see that my favorite cheese is on sale. While actively maintaining my second-order intention not to reconsider whether to intend to avoid dairy, I go ahead and reconsider anyway. That is, even though I intend not to reconsider the intention to avoid dairy, I begin questioning whether I really want to maintain this intention after all.

Such a state seems rationally impermissible, and with good reason – for it is an example of a paradigmatic sort of akratic irrationality. Akrasia can happen at two stages.¹⁰ First, and most frequently discussed, there can be an akratic break between what an agent judges best and what she intends to do. Second, there can be an akratic break between what an agent intends to do and what she actually does.¹¹ If I intend not to have a second glass of wine

¹⁰ See Tappolet (2013) for a helpful taxonomy of philosophical views on weakness of will and akrasia. Rorty (1980) notes that an akratic break can occur at multiple places, including between the formation of an intention and action. Wiggins (1978) argues that weakness of will has to do with failing to persist with one's intentions. Mele (1987) treats acting against one's intentions as one form of weakness of will, cashing it out as a special case of acting against one's best judgment.

¹¹ Holton uses "akrasia" in a narrow sense to refer only to judging that Φ ing is best, and then failing to intend to Φ . What I am calling second-stage akrasia will not always count as akratic in Holton's narrow sense, for it might not involve going against your best judgment: you might judge that Φ ing is best, and then intend to \square and fail to do so. However, it does not matter whether we refer to this gap between intention and action as akratic. What matters is whether a gap like this is problematic, and (as I shall argue below) Holton seems to presume that it is.

with dinner but I do so anyway, I am irrational by my own lights, for I am not successfully carrying out my own plans. When I intend not to reconsider my intention to avoid dairy but then reconsider this intention anyway, I am akratic in precisely the way that I am when I intend not to have a second glass of wine but then have one anyway.¹²

If I am to reconsider my intention to avoid dairy without being guilty of akratic irrationality, I must first *abandon* any intention I have to avoid engaging in such reconsideration, just as I would have to abandon my intention to refrain from a second glass of wine in order to rationally indulge in another glass. And as we saw in the previous section, rationally abandoning an intention requires reconsidering that intention. This shows us that rationally reconsidering the intention not to Φ requires reconsidering the second-order intention not to engage in such reconsideration.

It would be very difficult for Holton to reject this appeal to anti-akrasia norms, because such norms form an important part of his own argument. Recall that Holton claims that a major cause of judgment shift is that agents foresee that they are likely to succumb to temptation, and want to avoid the cognitive dissonance of believing that it would be best to refrain from Φ ing while believing that they will Φ anyway. The fact that agents feel dissonance when they predict that they will Φ while intending not to is evidence that they recognize that something is *irrational* about being in this state. Were it rationally permissible to intend to Φ and not act on that intention, agents would not feel any cognitive dissonance. If they did not feel dissonance, they would not feel pressure to shift their judgments, and we would not need resolutions to block such shift.

For ease of reference, I summarize the *Temptation Transmission* argument in the following table; recall that Intention 1 refers to the agent's intention to avoid Φ ing, while Intention 2 refers to the agent's second-order intention not to abandon Intention 1.

¹² Some important qualifications are necessary for a proper understanding of the general norm forbidding akasia between what an agent intends to do and what she actually does. We do not want to predict that an agent who intends to Φ at time $t+1$ is akratic because she fails to Φ at some earlier time t . So the proper formulation of the general norm will have to include some reference to tense, or perhaps be formulated as a synchronic rather than diachronic norm. Nor do we want to predict that an agent who intends to Φ but is prevented from Φ ing by entirely external circumstances is guilty of akratic irrationality, which means that the general norm will have to be formulated so as to rule out such cases. Hammering out this general formulation is beyond my present scope. However, we need not worry about ruling out external circumstances for my present purposes. I employ the intention/action anti-akasia norm with regard only to higher-order mental states: the agent intends not to reconsider her intention, and then reconsiders that intention anyway. Such higher-order mental states are internal, and – at least barring the interference of nefarious neuroscientists – are therefore in the agent's control.

	First stage	Transmission	Second stage	Why?
1	Temptation to Φ	Necessarily leads to ...	Appearance of a reason to Φ	<i>Empirical claim</i>
2	Appearance of a reason to Φ	Necessarily leads to ...	Appearance of a reason to intend to Φ	<i>Intentional action requires an intention</i>
3	Appearance of a reason to intend to Φ	Necessarily leads to ...	Appearance of a reason to give up Intention 1	<i>Intention consistency constraint</i>
4	Appearance of a reason to give up Intention 1	Necessarily leads to ...	Appearance of a reason to reconsider Intention 1	<i>Abandoning an intention requires reconsidering it</i>
5	Appearance of a reason to reconsider Intention 1	Necessarily leads to ...	Appearance of a reason to give up Intention 2	<i>Anti-akrasia norm</i>
6	Appearance of a reason to give up Intention 2	Necessarily leads to ...	Appearance of a reason to reconsider Intention 2	<i>Abandoning an intention requires reconsidering it</i>

The first four stages of the argument answer an independently important question about how temptation leads via a sense-making process to the appearance of a reason to abandon one’s first-order intention. This is a claim that Holton can accept. The last two stages, boxed in black, draw on an anti-akrasia norm that Holton in fact accepts to form an argument that Holton cannot accept: that temptation leads via a sense-making process to the appearance of a reason to abandon one’s second-order intention as well. The efficacy objection is now back in force: since the temptation to Φ leads to the appearance of an *equally strong* reason to abandon the first-order intention to Φ *and* to abandon the resolute second-order intention, the second-order intention cannot do any meaningful work in blocking temptation and preventing judgment shift.

5. Blocking the Efficacy Objection Another Way?

It is important to note the limits of *Temptation Transmission*. I am not claiming that agents who have formed Holtonian resolutions will *never* be able to resist temptation and prevent judgment shift. I am simply claiming that the method of resistance that Holton proposes – a second-order intention, specifically aimed at preventing reconsideration of the first-order intention – is susceptible to the very same pressure that it was designed to resist and is therefore unreliable. This is compatible with the claim that, in certain cases, agents will have *other* sources of resistance that enable them to avoid reconsidering the second-order intention and thereby prevent inappropriate judgment shift. Holton’s argument will still be in trouble if such other means of resistance are sometimes possible on a case-by-case basis, so long as no means of resistance bolsters second-order intentions across the board.

What might such a broad boost to our second-order intentions consist in? I have argued that a temptation that pushes with strength X on Intention 1 will also push with strength X on Intention 2, such that, all else being equal, Intention 2 will be no more effective at resisting judgment shift than Intention 1 alone would be. But if all else is not equal, there might be general, independent reasons for the agent to maintain Intention 2 that do not also apply to Intention 1. This defense requires that these reasons to maintain Intention 2 are *not ipso facto also* reasons to maintain Intention 1, lest Intentions 1 and 2 be equally well supported, and therefore equally susceptible to judgment shift. Are there any such reasons?

These reasons would have to be apparent to the agent, for reasons to maintain Intention 2 that the agent does not know about will not help the agent maintain Intention 2. What apparent reasons might an agent have for maintaining an intention to avoid reconsidering another intention? One answer is whatever reasons she has for holding that other intention in the first place. That smoking is unhealthy is a strong reason to intend to quit smoking. The unhealthiness of smoking seems to be an equally strong reason not to reconsider the intention to quit. But this strategy of deriving the reasons for Intention 2 from the reasons for Intention 1 will not help, for the reasons supporting Intention 2 would be of the same strength and type as the reasons supporting Intention 1, making the two intentions equally well supported. The reasons for Intention 2 cannot be so general that they necessarily support Intention 1 as well, lest the two intentions remain equally resistant to judgment shift.

Nor can the reasons for Intention 2 be too specific. We are seeking reasons that are general enough that they always, or at least very often, apply to Intention 2. For if agents occasionally had strong enough independent reasons to maintain their second-order intentions, Holtonian resolutions would succeed at resisting temptation only occasionally, and Holton's account would remain in trouble. To attain the requisite generality, the reasons for Intention 2 should be independent of the particular content of Intention 1. One candidate for such a reason is the claim that an agent's having an intention not to reconsider Intention 1 makes her maintaining Intention 1 more likely than if she did not have such a second-order intention. So long as the agent had a general reason to maintain her first-order intentions, she would have reason to form second-order intentions that bolster those first-order intentions. But this proposal is question-begging, for the candidate reason – that the presence of Intention 2 makes maintaining Intention 1 more likely – is precisely what is up for debate.

Second, even if this proposal were not question-begging, it is not obvious that agents really do have broad reasons to maintain their first-order intentions that are completely independent of the content of those intentions. In general, we have reason to be resolute in our particular intentions only to the extent that those intentions are good to have. Sometimes, particular resolutions ought to be abandoned, as when new information comes to light that

entails that the resolution no longer makes sense. If agents had very a strong general reason to remain resolute independently of the content of their particular resolutions, they might not be sufficiently responsive to situational factors that override or undermine particular resolutions. We do not seem to have a strong candidate for a reason for Intention 2 that is narrow enough so as not always to support Intention 1 as well, but general enough to support Intention 2 across the board.

If we cannot find general reasons that support Intention 2 but not Intention 1, perhaps we can bolster our second-order intentions by appealing to willpower. It is not entirely clear exactly how Holton takes the willpower mechanism to work. But suppose willpower functions as a source of friction or stickiness once a decision has been made and an intention formed: the resolute agent intends to Φ , intends not to reconsider this intention, and employs willpower to buckle down and follow through on each of these intentions, even if buffeted by temptation or other barriers to action. Might willpower enable an agent to resist reconsidering Intention 2, even when temptation presents a reason to reconsider it that is just as strong as the reason to reconsider Intention 1?

Suppose that an agent who intends to Φ employs willpower to degree X to sustain her intention. If she is subject to a temptation-based reason of strength $X+10$ to reconsider this intention, her willpower may fail her. To prevent this, she can form a second-order intention and employ willpower to degree X to sustain that intention. If she is not faced with a similar reason of strength $X+10$ to reconsider, this second-order intention is likely to remain effective where the first-order intention failed. But what if temptation transmits, and the agent is faced with a reason of strength $X+10$ to reconsider the second-order intention as well? Then the second-order intention is likely to fail as well.

We have no good reason to suppose that an agent's degree of willpower will as a rule be any greater at the second-order level than it is at the first-order level. For both intentions are aimed at the same goal, and the degree of willpower the agent is willing and able to exert for any given intention is presumably a function of how important the goal is to her. If the amount of temptation-based pressure and the degree of willpower employed are the same at each level, then the Intention 2/Intention 1 pair will not be any more effective than Intention 1 alone would be. However, *Temptation Transmission* will remain forceful even if we grant that agents are occasionally able to increase their second-order willpower, and are therefore on a case-by-case basis able to remain firm in their Holtonian resolutions. For this further source of resistance to judgment shift is neither systematic nor fully general. In the next section, I offer an account of resolutions that defuses *Temptation Transmission* entirely, rather than circumventing it on a case-by-case basis.

6. An Alternative Account of Resolutions

If temptation cannot always and easily be blocked by a second-order intention, perhaps it can be blocked in another way. One strategy can be dismissed immediately. Adding a third-order intention to the resolution – an intention not to reconsider your second-order intention not to reconsider your first-order intention not to Φ – will not suffice. This is because we can draw on RMRT and the anti-akrasia norm to apply *Temptation Transmission* mutatis mutandis to a third-order intention, as well as to intentions of progressively higher orders. Nor do we want to give up on the plausible fundamental background principles that underlie *Temptation Transmission*.

Rather, the *Temptation Transmission* argument shows us that we must appeal to a mental state of a *different sort* to block temptation – specifically, to a mental state that is not subject to an anti-akrasia norm. Desires can successfully play this role:

Second-Order Desire Account (SODA): Resolving to Φ involves intending to Φ , and *desiring* not to reconsider the intention to Φ .¹³

¹³ Earlier discussions of resolution or of strength of will make reference to second-order desires; however, SODA is distinct from each in important ways. Bigelow, Dodds and Pargetter (1990) argue that temptation is best understood as a first-order desire to do X, and a second-order desire that the first-order desire to do X not be causally operative: “to be tempted is to have a desire which you want *not* to be your strongest desire” (44). In their account, strength of will consists in acting in line with the second-order desire, rather than the first. SODA proposes that an agent is resolute when she acts in accordance with a second-order desire, which at first glance appears to be a similar view. However, my account differs from theirs in two main ways. First, (with Holton) I construe temptation as the (appearance of a) reason to act, rather than as a pair of desires; this is because I believe that succumbing to temptation usually involves engaging in a sense-making process that involves responding to (apparent) reasons, rather than merely acting akratically. Second, the content of the second-order desire in SODA is an *intention*, rather than a first-order desire. The presence of the intention is essential, as the agent forms a second-order desire because she cares about being an effective agent, and carrying out her (general or specific) plans and goals (see the discussion of fetishism in Section 6.3 below).

Kovach and Fitzpatrick (1999) argue that a necessary criterion for an account of resolutions is the ability to accommodate both resolutions that aim at *self-control* and those that aim at *self-transformation*. They assess whether resolutions might be simple (or “narrow”) second-order desires, and argue that desires of this sort do not achieve self-control (because I might desire that I avoid donuts, and desire to have this desire, without actually avoiding donuts) or self-transformation (which requires that the desire “flow naturally and without conflict from values” (166)). They then assess whether resolutions might be second-order desires about the efficacy of a desire as Bigelow et al. propose: they argue that the proposal is on the right track, but needs to be supplemented. Kovach and Fitzpatrick offer a *process-specific* desire account, according to which resolutions are second-order desires in which “the kinds of processes by which they are effected are included in their content” (167). These processes are those that “are constitutive of the projects of self-control and self-transformation” (168). For example, an agent who aims at self-control might engage in a process of vividly imagining the negative consequences of giving in, or imagining the tempting thing as distasteful, or encouraging friends to chastise him if he fails; an agent who aims

SODA is meant to lay out necessary conditions on resolving to Φ , and does not purport to lay out a full set of sufficient conditions. According to SODA, an agent who resolves to Φ must at least intend to Φ , and must desire that she carry out that intention, even in spite of contrary inclinations – she must *care about* whether she maintains her intention to Φ .¹⁴

In the rest of this section, I explain why SODA is appealing as an alternative to Holton’s view. First, SODA does not fall prey to *Temptation Transmission*. Second, it can explain data about resolutions that intention-based views have a hard time accommodating. Finally, I address worries that Holton raises about whether a view like SODA can adequately capture the phenomenology of making and keeping resolutions.

6.1. Why SODA Avoids the Temptation Transmission Argument

SODA will not fall prey to the same *Temptation Transmission* argument that troubles Holton’s view. For the argument to be effective against SODA, we would need to employ an analogue of Claim 3 that was about desire rather than intention:

at self-transformation might attempt to consciously change her self-conception.

Kovach and Fitzpatrick’s account differs from SODA in that the content of the second-order desire that makes up the resolution is (1) a desire, rather than an intention, and (2) necessarily about specific processes having to do with engaging in self-control or self-transformation. I have already addressed why the content of the second-order desire must be an intention rather than a desire. I find the process component of their account overly restrictive. While resolutions generally do aim at self-control or self-transformation, I do not think that the processes by which these aims are attained must be part of the object of the desire. If I intend to avoid donuts and desire that this intention be effective, I have successfully resolved to avoid donuts – even if I do not have any desires about the particular processes that I will engage in to make this intention effective.

¹⁴ One might worry that this proposal is implausible if we can always form resolutions at will, since we cannot form desires at will and resolutions require desires. However, I do not think we can form resolutions at will when we lack the necessary desires. For example, imagine that you do not care at all about whether your intention to avoid smoking remains intact. That is, you intend to quit smoking, but could not care less about whether you maintain or abandon this intention. How could you form a resolution in such a case? It seems plausible that an agent who lacks all desire to remain consistent in her intention to Φ will be unable to genuinely resolve to Φ .

In the majority of cases, an agent who attempts to form a resolution to Φ will desire to maintain the intention to Φ , for she would not be attempting to form a resolution about Φ ing in the first place if she did not have such a desire. Cases in which an agent attempts to form a resolution but cannot do so because she lacks the requisite desire may be possible; perhaps an agent suffering from depression intends to Φ and aims to turn this intention into a resolution, but cannot bring herself to care about whether the intention is effective. But we do not expect agents in such situations to be able to successfully form resolutions. Nor are such cases common in typical agents.

Claim 3*: Reconsidering any desire to avoid reconsidering the intention not to Φ is necessary for rationally reconsidering the intention not to Φ .

If Claim 3* were true, we could use RMRT to argue that the appearance of a reason to reconsider the intention not to Φ necessarily leads to the appearance of a reason to reconsider your desire not to engage in such reconsideration. But Claim 3* is not true, because desires are not subject to the same rational norms forbidding akrasia that intentions are subject to.

It is not necessarily irrational to desire that you Φ while nevertheless failing to Φ , or to desire that you refrain from Φ ing while Φ ing anyway. We all have very many desires that we do not act on, and we display no irrationality in doing so. I can desire to spend the day at the beach and rationally fail to do so because I have to spend the day at work instead. Similarly, I might desire to avoid an unpleasant task such as filing my taxes, but I display no irrationality in filing them anyway. It can be rational to do what you desire not to do in a way that it is *not* rational to do what you intend not to do. Structurally speaking, making the move from second-order intention to second-order desire is exactly what we need to avoid *Temptation Transmission*. This is because the argument's success depends on the existence of a second-order intention that is subject to an anti-akrasia norm requiring the agent to avoid the state of intending to Φ while not Φ ing. If we replace this intention with a mental state like desire that is *not* subject to the same anti-akrasia norm, we cannot generate the same argument.

6.2. How SODA Explains Variations in the Strength of Resolutions

Resolutions come in degrees; some resolutions are stronger than others. A resolution is stronger to the extent that an agent takes it more seriously, and a greater excuse is needed to license breaking it. My resolution to read *Don Quixote* is not very strong because this is a nice but inessential goal, whereas my resolution to pursue my research project is very strong because this is an important part of my career plan. Appealing to desire as a necessary component of a resolution gives us an easy and efficient explanation of how resolutions can vary in strength. For desires clearly come in degrees; some desires are stronger than others. We can explain the strength of a resolution as a direct result of the strength of the agent's desire to avoid reconsideration: the stronger the desire, the stronger the resolution. Since intentions do not come in fine-grained degrees, Holton's intention-based account cannot offer the same sort of explanation.¹⁵

¹⁵ In chapter 2 of *Willing, Wanting, Waiting*, Holton outlines a notion of "partial intention." This is not a fine-grained strength or degree of intention in the way that there are fine-grained strengths of desire. Rather, "an intention to F is partial iff it is designed to achieve a given end E and it is accompanied by one or more alternative intentions also designed to achieve E" (36).

Explaining the strength of resolutions in terms of the strength of the agent's desire to avoid reconsideration makes intuitive sense as well. In general, the degree to which an agent is resolute in Φ ing seems to depend not on how strong the temptation to Φ is, but on how much the agent cares about whether she Φ s. Suppose I resolve not to drink any beer at a party tonight, but I really do not care very much about whether I keep this resolution. Even beer that is not very tempting (say, mass-produced American lager) will be likely to sway me to break my resolution. But suppose I resolve not to drink any beer at a party tonight, and I care very much about whether I keep this resolution. In such a case, it seems that even extremely tempting beer (say, an unusual, local, craft IPA) will not be very likely to sway me to break my resolution. What explains the difference between these two cases is that the resolution is stronger in the second case, when my desire to avoid reconsideration is stronger.

This variation in strength of desires gives SODA another explanatory advantage over Holton's view. One might be worried that temptation puts pressure on an agent's resolution not just by making it appear that there is a *reason* to reconsider her resolute intention, but also by creating a *desire* in her to reconsider this intention. After all, desire aims at the good, and the tempting thing presents itself as good. SODA can explain when resolutions are able to resist desire-based temptations of this sort. If temptation ever works in this way, agents are faced with conflicting desires: a temptation-based desire to reconsider, and a resolution-based desire not to reconsider. Resolutions sometimes fail; neither Holton nor I want to claim that our respective accounts of resolutions enable agents to resist in every instance. According to SODA, whether a given resolution is effective when faced with a conflicting temptation-based desire will depend in part on how strong the resolution-based desire is, and whether it is strong enough to outweigh the temptation-based desire. It would not be as easy and straightforward to weigh the strength of a second-order intention against that of the temptation-based desire.

6.3. Why SODA Can Capture the Phenomenology of Resolutions

SODA is a natural option to consider once the possibility of a second-order account of resolutions has been raised. Indeed, in *Willing, Wanting, Waiting*, Holton considers and rejects a simple version of a higher-order desire account. Does Holton's rejection of higher-order desire views bring to light problems of comparable gravity for SODA as the efficacy objection creates for Holton's view? Holton considers the following higher-order desire view: suppose you resolve to Φ and "you have a *strong desire to be resolute*: a strong desire to stick to your resolutions" (116). He grants that this could enable the agent to resist temptation, for "when the date for implementing the resolutions comes, provided that your desire to be resolute is stronger than your

desire to [give up], you have a desire-driven way to give up” (116). Holton calls this the *further desire* approach, presenting a problem for it in a footnote:

The *further desire* approach seems to involve attributing to the strong-willed agent a desire for resoluteness that approaches a fetish. It is surely crazy to want to be resolute for its own sake, especially if, as a result of judgement shift, the agent comes to believe that the resolute course of action is the less desirable (118).

First, I’m not convinced that a desire for resoluteness is in fact as fetishistic as Holton makes it out to be. An agent who desires resoluteness need not care about resoluteness *simpliciter*, but might instead care about *effectively carrying out all of her endorsed plans*; being resolute helps her to carry out these plans. Fetishistic desires are those that you hold on to for no good reason, or even in spite of reasons to the contrary – but in general, we have good reason to be effective planning agents regarding our own endorsed plans, and to have resolute dispositions that help us attain this goal.

Second, suppose Holton is right that desires to be resolute are problematically fetishistic. SODA need not require that agents desire to be resolute in general. Rather, they could desire to be resolute with regard to *specific* intentions, which we can assume they take themselves to have independent reasons to act on. Such desires would not be fetishistic, because agents would have independent reasons to maintain them.

One might worry that second-order desires about specific intentions would not be very effective, because they are not distinct enough from first-order desires for the objects of those intentions.¹⁶ So an agent might desire that her intention not to eat the donut be effective *because* she desires to eat the donut. But if this desire not to eat the donut persists throughout the temptation, it is not clear that the resolution needs to do any extra work. And if the desire not to eat the donut does not persist, why should we expect the desire to remain resolute in the intention not to eat the donut to persist?

Neither of these possibilities is a problem for SODA. First, if the agent’s first-order desire to refrain from eating the donut persists throughout the temptation, then no resolution is needed, and so it does not matter whether the resolution does any extra work. A first-order desire to Φ can motivate you to carry out an intention to Φ . You need to *resolve* to Φ if you are not sure whether this first-order desire will remain strong in the face of temptation. If the first-order desire does remain firm, it can do the motivational work in bringing you to carry out the intention. But if the first-order desire lessens or disappears, the second-order desire can do the motivational work instead.

Second, the second-order desire can persist when the first-order desire does not because the second-order desire is held for additional reasons. I might desire to avoid eating a donut because I do not want to ruin my sup-

¹⁶ Thanks to an anonymous referee for raising this worry.

per, or because I do not want to get powdered sugar on my shirt, or because I want to heed my doctor's advice to consume less sugar, etc. I desire to remain firm in my intention to avoid eating donuts for another reason: because I care about carrying out my donut-avoidance plan and being an effective agent regarding the baked goods I consume. Granted, I care about carrying out my donut-avoidance plan *because* I am concerned with heeding my doctor's advice, etc. But my desire to carry out my plan is a desire about what kind of agent I want to be; this sort of desire is resistant to temptations that press on the content of the plan itself.

A more serious problem Holton finds with the further-desire approach is that it "completely misrepresents the phenomenology of strength of will" (118):

If these accounts were right, then sticking to a resolution would consist in the triumph of one desire (the stronger) over another. But that isn't what it feels like. It typically feels as though there is a *struggle*. One maintains one's resolutions by dint of effort in the face of the contrary desire ... by and large, maintaining strength of will requires effort (118).

There is an important insight here. Sticking to a resolution in the face of temptation indeed feels like a struggle, and maintaining strength of will requires effort. But rather than lead us to dismiss SODA completely, this worry should help us realize that desire-based accounts of resolutions need to be more nuanced than they might first appear.

An account like SODA need not be committed to a simple Humean picture in which desires do all of the motivating work and the agent is simply along for the ride, at the mercy of whichever desire is strongest. For SODA does not claim that a second-order desire is by itself *sufficient* to bring an agent to act resolutely. Rather, it claims only that the second-order desire is *necessary*. This is entirely compatible with it being the case that serious mental effort on the part of the agent is *also* necessary for resolute action. Even a distinct faculty of willpower of the kind Holton describes is compatible with SODA, so long as desire plays an essential role in triggering the agent to employ that willpower.

Holton grants that this is often the case. In a discussion about willpower, he notes that whether an agent revises a resolution "will depend on, amongst other things, the strength of their desire to maintain those resolutions in particular, and the strength of their desire to maintain their resoluteness in general" (133). He also notes that "to be effective, these desires need not be the strongest. If the agent's willpower is sufficiently strong, a weak desire to be resolute might be all that is needed to keep it in place when it wavers in the face of a strong contrary desire" stemming from temptation (133).

SODA does not require that the desire to be resolute is the strongest desire that the agent has, capable of overwhelming all of her other desires. Only one claim about desire is essential for SODA: that the desire present in reso-

lutions is not subject to the anti-akrasia norm that drives *Temptation Transmission*. Any account of desire that meets this constraint can block the transmission of temptation. Since intentions of any order are subject to anti-akrasia norms, they cannot block the transmission of temptation all by themselves. Something more is needed; I have suggested that something is desire.

The primary point of disagreement between Holton's view and SODA is not about whether desire *can* play a role in helping agents to remain resolute; Holton grants that it can. Rather, SODA and Holton's view differ over the role played by second-order intentions. Holton claims that they are both necessary and sufficient for maintaining resolutions. I have argued that they are insufficient because they leave resolutions vulnerable to the efficacy objection. By more thoroughly developing this objection, we have not only seen why temptation transmits in such a way that Holton's view does not succeed. We have also fleshed out the previously under-explained mechanism by which agents succumb to temptation, and discovered that what is needed to avoid the objection is a mental state like desire that is not subject to anti-akrasia norms.¹⁷

Alida Liberman
Rotman Institute of Philosophy
Western University
aliberm3@uwo.ca

¹⁷ Thanks to Joshua Crabill, Steve Finlay, Ben Lennertz, Hanti Lin, Shyam Nair, Jake Ross, Mark Schroeder, Julia Staffel, Gary Watson, two anonymous referees and the audience at the 2014 Annual Meeting of the Society for Exact Philosophy for valuable comments on and discussion of this paper.

References

- Bedke, M. S. (2009) "The Iffiest Oughts: A Guise of Reasons Account of End-Given Conditionals," *Ethics* 119(4): 672-98.
- Bigelow, J., Dodds, S. M. and Pargetter, R. (1990) "Temptation and the Will," *American Philosophical Quarterly* 27(1): 39-49.
- Bratman, M. (2009a) "Intention, Belief, Practical, Theoretical," in E. S. Robertson, ed., *Spheres of Reason: New Essays in the Philosophy of Normativity*, Oxford: Oxford University Press, pp. 29-61.
- ____ (2009b) "Intention, Practical Rationality, and Self Governance," *Ethics* 119(3): 411-43.
- ____ (1987) *Intention, Plans, and Practical Reason*, Cambridge, MA: Harvard University Press.
- Broome, J. (2013) *Rationality Through Reasoning*, Oxford: Oxford University Press.
- Carr, J. (2016) "Don't Stop Believing," *Canadian Journal of Philosophy* 45(5-6): 744-66.
- Christensen, D. (2000) "Diachronic Coherence Versus Epistemic Impartiality," *Philosophical Review* 109: 349-71.
- Cohen, D. and Handfield, T. (2010) "Rational Capacities, Resolve, and Weakness of Will," *Mind* 119(476): 907-32.
- Dannenberg, J. (2015) "Promising Ourselves, Promising Others," *The Journal of Ethics* 19(2): 159-83.
- Darwall, S. (1983) *Impartial Reason*, Ithaca, NY: Cornell University Press.
- Ferrero, L. (2012) "Diachronic Constraints of Practical Rationality," *Philosophical Issues* 22(1): 144-64.
- Gauthier, D. (1997) "Resolute Choice and Rational Deliberation: A Critique and a Defense," *Noûs* 31(1): 1-25.
- Goble, L. (2004) "A Proposal for Dealing with Deontic Dilemmas," *Deontic Logic in Computer Science*: 74-113.
- Habib, A. (2009) "Promises to the Self," *Canadian Journal of Philosophy* 39(4): 537-57.
- Hedden, B. (2015) *Reasons Without Persons: Rationality, Identity, and Time*, Oxford: Oxford University Press.
- Hills, A. (2003) "Duties and Duties to the Self," *American Philosophical Quarterly* 40: 131-42.
- Hlobil, U. (2015) "There Are Diachronic Norms of Rationality," *Thought: A Journal of Philosophy* 4(1): 38-45.
- Holton, R. (2009) *Willing, Wanting, Waiting*, Oxford: Oxford University Press.
- Horty, J. (1993) "Deontic Logic as Founded on Nonmonotonic Logic," *Annals of Mathematics and Artificial Intelligence* 9: 69-91.
- Kolodny, N. (forthcoming) "Instrumental Reasons," in D. Star, ed., *The Oxford Handbook of Reasons and Normativity*, Oxford: Oxford University Press.
- ____ (2007) "How Does Coherence Matter?" *Proceedings of the Aristotelian Society* 107(1.3): 229-63.
- Kovach, V. and Fitzpatrick, J. (1999) "Resolutions," *Australasian Journal of Philosophy* 77(2): 161-73.
- Liberman, A. (unpublished manuscript) "Understanding Vows."
- Maier, J. (2010) "Review of *Willing, Wanting, Waiting*," *Australasian Journal of Philosophy* 89(2): 361-64.
- McClellan, E. F. (1990) *Rationality and Dynamic Choice: Foundational Explorations*, Cambridge: Cambridge University Press.
- McGuire, M. C. (1961) "Decisions, Resolutions, and Moral Conduct," *Philosophical Quarterly*, 11(42): 61-67.
- Mele, A. (1987) *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*, Oxford: Oxford University Press.
- Moss, S. (2015) "Time-Slice Epistemology and Action under Indeterminacy," in J. Hawthorne and T. Gendler, eds., *Oxford Studies in Epistemology, Vol. 5*, Oxford: Oxford University Press, pp. 172-94.
- Nagel, T. (1970) *The Possibility of Altruism*, Oxford: Clarendon Press.

- Nair, S. (2014) "Consequences of Reasoning with Conflicting Obligations," *Mind* 123(491): 753-90.
- Paul, S. (2011) "Review of *Willing, Wanting, Waiting*," *Mind* 120(479): 889-92.
- Raz, J. (2005) "The Myth of Instrumental Rationality," *Journal of Ethics and Social Philosophy* 1: 1-28.
- ____ (1977) "Promises and Obligations," in P. M. S. Hacker and J. Raz, eds., *Law, Morality and Society: Essays in Honor of H. L. A. Hart*, Oxford: Clarendon Press, pp. 210-28.
- Rorty, A. (1980) "Where Does the Akratic Break Take Place?" *Australasian Journal of Philosophy* 58: 333-46.
- Rosati, C. (2011) "The Importance of Self-Promises," in H. Sheinman, ed., *Promises and Agreements*, Oxford: Oxford University Press.
- Scanlon, T. M. (1998) *What We Owe to Each Other*, Cambridge, MA: Harvard University Press.
- Schroeder, M. (2009) "Means-End Coherence, Stringency, and Subjective Reasons," *Philosophical Studies* 143(2): 223-48.
- Singer, M. (1959) "On Duties to Oneself," *Ethics* 69(3): 202-5.
- Tappolet, C. (2013) "Weakness of Will," in H. LaFollette, ed., *The International Encyclopedia of Ethics*, Hoboken, NJ: Wiley-Blackwell.
- Way, J. (2012) "Transmission and the Wrong Kind of Reason," *Ethics* 122(3): 489-515.
- ____ (2010) "Defending the Wide Scope Approach to Instrumental Reason," *Philosophical Studies* 147(2): 213-33.
- Wiggins, D. (1987) "Weakness of Will, Commensurability, and the Objects of Deliberation and Desire," in *Needs, Values, Truth: Essays in the Philosophy of Values*, Oxford: Blackwell Publishers, pp. 239-67.