

## **Weakness of Will, Reasonability, and Compulsion\***

James R. Beebe (University at Buffalo)

Forthcoming in *Synthese*

Experimental philosophers have recently begun to investigate the folk conception of weakness of will (e.g., Mele 2010; May and Holton 2012; Beebe forthcoming; Sousa and Mauro forthcoming). Their work has focused primarily on the ways in which akrasia (i.e., acting contrary to one's better judgment), unreasonable violations of resolutions, and variations in the moral valence of actions modulate folk attributions of weakness of will. A key finding that has emerged from this research is that—contrary to the predominant view in the history of philosophy—ordinary participants do not think of weakness of will solely in terms of akrasia but see resolution violations and moral evaluations as playing equally important roles. The present article extends this line of research by reporting the results of four experiments that investigate (i) the interplay between hastily revising one's resolutions and the degree of reasonableness of the actions one had resolved to undertake, (ii) whether ordinary participants are willing to ascribe weakness of will to agents whose actions stem from compulsion or addiction, and (iii) the respects in which akratic action, resolution violations, and the seriousness of an addiction impact attributions of weakness of will to agents acting in accord with their addictions.

keywords: weakness of will, akrasia, intention, compulsion, addiction, experimental philosophy, folk psychology

---

\* Thanks to Frank Hindriks, Al Mele, Kevin Timpe, and two anonymous reviewers from *Synthese* for helpful comments on earlier drafts of this paper. Thanks also to Mark Alfano and Mark Phillips for assistance in setting up the experiments.

## 1. Introduction

Historically, the most influential and widely endorsed perspective concerning the nature of weakness of will has been the view that weak-willed action consists primarily in akrasia—i.e., acting in a manner contrary to what one judges to be the best course of action, all things considered.<sup>1</sup> In a recent series of experiments Joshua May and Richard Holton (2012), Beebe (forthcoming), and Paulo Sousa and Carlos Mauro (forthcoming) have found that ordinary participants attribute weakness of will to agents who act contrary to their better judgment but also to agents who revise their contrary-inclination-defeating resolutions too readily or who act contrary to what morality requires. Thus, in addition to finding that the notion of akratic action fails to fully capture the folk conception of weakness of will, these studies also reveal that resolution violations and moral judgments figure more prominently in folk thinking about weakness of will than philosophers might have expected—at least in light of the historical prominence of the identification of weakness of will with akrasia and the commitment of philosophers to action to the explication of the ordinary notion of weakness of will.

The recent turn toward using empirical data to adjudicate conflicting claims about the nature of the folk conception of weakness of will began when Al Mele (2010) employed data from surveys and experiments he constructed to support his own view of the folk conception and to argue against the view proposed by Holton (1999; 2003; 2009). Holton argued that understanding weak-willed action as the violation of contrary-inclination-defeating intentions (or ‘resolution violations’) better captured the folk conception of weakness of will than the

---

<sup>1</sup> For example, describing what he takes to be the common view of weakness of will, Plato writes, “many people who know what it is best to do are not willing to do it, though it is in their power, but do something else” (*Protagoras* 352d). Davidson (1970, 93) writes, “An agent’s will is weak if he acts, and acts intentionally, counter to his own best judgement; in such cases we sometimes say he lacks the will power to do what he knows, or at any rate believes, would, everything considered, be better.” Several philosophers have questioned whether weakness of will is really possible—since it requires acting contrary to what one believes one ought to do—while others have wondered whether weak-willed actions can be free, if the agents who perform them are overcome by their passions or desires (cf. *Protagoras* 358b-c, Hare 1952, Hare 1963, Watson 1977, Audi 1979, Pugmire 1982, and Mele 1986).

traditional view that identifies it with akrasia. He contended that “actors show weakness of will when they revise an intention as a result of a reconsideration that they should not have performed; that is, when their reconsideration exhibits tendencies that it is not reasonable for the agent to have” (1999, 248).<sup>2</sup> Holton (1999, 250) explained that the kind of intention in question is a future-directed one that is designed to defeat inclinations to act in a contrary manner.

In an effort to undermine Holton’s claims about resolution violations and weakness of will, Mele (2010, 395ff.) asked participants “What is weakness of will? Please answer this question and briefly describe one example of weakness of will.” About 15% of participants gave answers that cited typical features of akrasia, but only 1.4% gave an answer that mentioned failing to do what one resolved or intended to do. In a second study Mele asked participants which of the following descriptions of weakness of will they thought was more accurate:

- (A) Doing something you believed or knew you shouldn’t do (for example, going to a party even though you believed it would be better to stay home and study).
- (B) Doing something you decided or intended not to do (for example, going to a party even though you decided to stay home and study).
- (C) Neither. The descriptions are equally accurate or inaccurate.

49% of participants chose option (A), which concerned acting contrary to one’s better judgment, while only 33% chose (B), which concerned acting contrary to one’s intentions. Mele (2010,

---

<sup>2</sup> Holton (1999, 241) wrote: “I do not agree that this is the untutored view. Whenever I have asked non philosophers what they take weakness of will to consist in, they have made no mention of judgments about the better or worse course of action. Rather, they have said things like this: weak-willed people are irresolute; they do not persist in their intentions; they are too easily deflected from the path that they have chosen.” Holton (1999) credits Wiggins (1987) with defending a similar view.

400) concludes from this that the ordinary notion of weakness of will “is much closer to a relatively standard conception of akrasia than Holton would have us believe.”<sup>3</sup>

Beebe (forthcoming), however, argues that Mele was able to obtain this result only because (B) failed to represent the relevant kind of resolution violation as one that is unreasonable for the agent or otherwise in violation of some normative requirement on the agent. When Beebe (forthcoming) replaced ‘decided or intended’ in Mele’s answer (B) with ‘resolved’—which seems to have normative implications that the former lacks—significantly more participants were found to choose (B) than (A). This result lends support to Holton’s claim that the unreasonable violation of resolutions figures prominently in folk thinking about weakness of will.<sup>4</sup>

Motivated by Mele’s (2010) empirical critique of his claims, Holton teamed up with Joshua May to formulate a series of experiments that showed that resolution violations do indeed play a role in the folk conception of weakness of will. Holton, however, abandoned his previous claim that weakness of will was solely a matter of resolution violation. May and Holton (2012, 342) describe their resulting prototype model of the folk notion as follows:

Indeed, our findings suggest that no simple account phrased in terms of necessary and sufficient conditions will do the job. The ordinary notion of weakness of will is more like a prototype or cluster concept. There are core cases that possess a number of features. As these features are removed, people are less inclined to describe the resulting cases as ones of weakness of will. Akrasia and resolution-violation are indeed among these features.

---

<sup>3</sup> Although Mele (2010, 394, 397) believes that akratic actions constitute the “traditional” or “orthodox” instances weakness of will, he allows that there may be some “untraditional” or “unorthodox” cases of weakness of will that do not involve akrasia.

<sup>4</sup> Mele (2010) also carried two other studies that are not undermined by the present considerations.

However, neither is sufficient on its own for an ascription of weakness of will; and other features also play a role, such as the moral valence of the action.

Beebe (forthcoming) followed the lead of May and Holton but showed that akrasia and the violation of resolutions can each be sufficient to motivate folk attributions of weakness of will.

Because of the somewhat unexpected role that resolution violations have been shown to play in the literature described above, the present article reports three experiments that were designed to investigate the relation between resolution violations and folk attributions of weakness of will in further detail. Since, according to Holton (1999), when an agent revises an intention not to do *A* too readily and subsequently performs *A*, it is not so much the doing of *A* as it is the unreasonable and hasty setting aside of the resolution that is the instance of weakness of will, the first vignette-based experiment featured protagonists who make resolutions, set them aside, renew their original resolutions, and then act in accord with those resolutions. As we will see in Section 2 below, participants' attributions of weakness of will did not appear sensitive to the fact that the protagonists flip-flopped between making a resolution, setting it aside, and renewing the resolution. What did affect attributions of weakness of will, however, were the respects in which their resolutions and subsequent actions were reasonable or otherwise normatively well-founded.

Section 3 reports the results of a second experiment that features a protagonist who revises his resolution in light of new information about his future course of action and does so in a reasonable manner. Participants were found to be disinclined to view the protagonist as weak-willed. Experiment 3 describes an agent who resolves to undertake an unreasonable and dangerous course of action but then sets this resolution aside after engaging in varying degrees of

reflection upon the potential danger involved. Participants were again disinclined to view this agent as weak-willed, in light of the reasonableness of his resolution violation.

Finally, a fourth experiment was constructed that investigated whether experimental subjects would be willing to ascribe weakness of will to agents whose actions stem from compulsion or addiction and the respects in which akrasia, resolution violations, and the seriousness of an agent's addiction might impact attributions of weakness of will in such contexts. Contrary to much philosophical thinking on the matter, participants were happy to attribute weakness of will to kleptomaniacs, smokers, and cocaine addicts who were acting in accord with their addictions. In these cases, whether an agent violated a resolution or acted akratically had less of an impact on attributions of weakness of will than in previous experiments.

## **2. Experiment 1**

Holton (1999, 248) claimed that “actors show weakness of will when they revise an intention as a result of a reconsideration that they should not have performed; that is, when their reconsideration exhibits tendencies that it is not reasonable for the agent to have.” Holton (1999, 249) suggests it can be reasonable to revise one's resolutions if (i) one believes that circumstances have changed in such a way that they defeat the purpose of having them, (ii) one believes the intentions can no longer be carried out, or (iii) “one believes that they will lead one to great suffering when that suffering was not envisaged at the time of forming the intention.” However, he maintains that it is not reasonable “to have a tendency to reconsider intentions that were expressly made in order to get over one's later reluctance to act” (ibid.).

In order for there to be weak-willed action, an agent must obviously perform an action of some kind. However, according to Holton, an agent who has hastily and unreasonably revised a resolution has already displayed weakness of will—even if they have not yet performed a weak-willed action and indeed even if they never will. Therefore, in Experiment 1 three vignette sets were constructed that described agents as making resolutions, setting aside their resolutions, renewing the original resolutions, and then acting in accord with those resolutions. Since the actions in question were in accord with each agent’s most recent resolution and half the time were also in accord with the agent’s considered judgments, the vignettes allow us to observe what effect (if any) that hasty revisions and re-revisions of resolutions might have on participants’ attributions of weakness of will, independently of other features of an agent’s action.

Building upon the cases used by Beebe (forthcoming) the first vignette features an agent who performs a morally bad action:

Franz, a Nazi soldier, has been charged with the task of searching every flat in a certain Jewish ghetto and imprisoning every Jewish intellectual he finds there. Franz, however, has been lax in carrying out his duties and has not imprisoned every Jewish intellectual he has found. Franz *believes / has not decided whether he believes* that it would be best for him if he strictly carried out his duties, *so / but* one day he resolves to start imprisoning every Jewish intellectual he finds. The next morning, however, Franz decides to set aside this resolution and to continue being lax in carrying out his duties. Yet as he begins his workday, Franz resolves once again to start imprisoning every Jewish intellectual he finds. Later that day Franz imprisons a Jewish intellectual he encounters in the ghetto.

The italicized phrases represent the sole differences between two versions of the vignette. In the first version the resolution Franz makes is described as being in accord with his considered judgment (CJ), but in the second he is described as not having made such a judgment (not-CJ). Participants were then asked a question about Franz's resolution revisions—not about his ultimate action. They were asked to indicate the extent to which they agreed or disagreed with the following claim: “When Franz changes his resolution the second time, he displays weakness of will.” Responses were recorded on a five-point Likert scale, with ‘1’ marked ‘Strongly Disagree’ and ‘5’ marked ‘Strongly Agree.’

A second vignette set featured an agent performing a morally neutral action:

John has been reading a book on Buddhism and the benefits of meditating every day. John *believes / has not decided whether he believes* that it would be best for him if he meditated thirty minutes every night before going to bed, *so / but* one day he resolves to meditate for thirty minutes every night. The next day, however, John decides to set aside this resolution and to continue with his ordinary bedtime routine. Yet when bedtime rolls around, John resolves once again to meditate for thirty minutes every night. That night John meditates before going to bed.

Participants were asked whether they agreed or disagreed with the statement “When John changes his resolution the second time, he displays weakness of will,” and their responses were recorded as before.

The final set of vignettes used in Experiment 1 involved a protagonist who performed a morally good action:

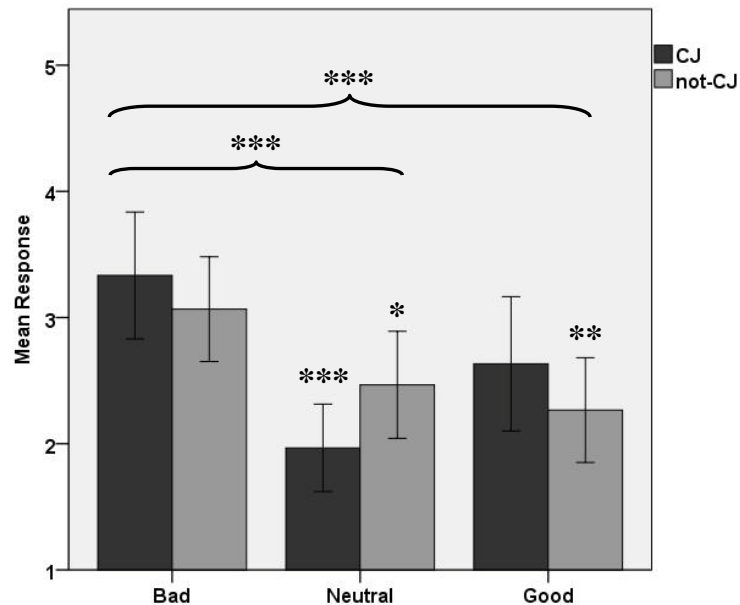
John cannot afford People Magazine, so he often steals one from a street vendor while walking to work. John *believes / has not decided whether he believes* that it would be best



for him if he did not steal, *so / but* one day he resolves to stop stealing People Magazine from street vendors. The next day, however, John decides to set aside this resolution and to continue his habit of stealing magazines. Yet when John begins his walk to work, he resolves once again to stop stealing People Magazine from street vendors. Later that day John walks by a street vendor's stall and does not steal a People Magazine.

John's action is morally good simply because he does not perform a bad action he had previously considered and habitually performs.

Using Amazon's Mechanical Turk (<http://www.mturk.com>), 180 participants (average age = 36, 66% female, 76% Anglo-American) were each given one of the six vignettes described above. Summary results are displayed in Figure 1.



*Figure 1.* Mean participant responses in the Bad (CJ = 3.33, not-CJ = 3.07), Neutral (CJ = 1.97, not-CJ = 2.47), and Good (CJ = 2.63, not-CJ = 2.27) conditions of Experiment 1. An ‘\*’, ‘\*\*’, or ‘\*\*\*’ by itself indicates that the mean differs significantly from the neutral midpoint at either the .05, the .01, or the .001 level. An ‘\*’, ‘\*\*’, or ‘\*\*\*’ with a bracket indicates a statistically significant difference between pairs of conditions at either the .05, the .01, or the .001 level. In each figure error bars represent 95% confidence intervals.

In contrast to previous experiments (e.g., May and Holton 2012; Beebe forthcoming; Mauro and Sosa forthcoming), mean attributions of weakness of will were much lower because the protagonists in the vignettes above acted in accord with rather than in violation of their resolutions and never acted contrary to their considered judgments. Participants were in general disinclined to attribute weakness of will to John the (neutral) meditator for meditating or to John the (good) magazine thief for not stealing, but they were evenly split about (bad) Franz who imprisoned a Jewish intellectual. No main effect was found for the CJ/not-CJ variable—i.e., it did not matter significantly whether or not the protagonists acted akratically, and there was no significant interaction between the CJ/not-CJ variable and moral valence.<sup>5</sup> A medium-sized main effect was, however, found for moral valence, which reflects the fact that mean responses in the bad condition were significantly higher than in the good and neutral conditions.<sup>6</sup>

In short, the fact that protagonists made a resolution, set it aside, and then renewed the resolution did not in general lead participants to think that they displayed weakness of will. More important to participants seems to be the question of how reasonable or otherwise normatively well-founded a given course of action was. When John the magazine thief resolves not to steal magazines, sets this resolution aside, resolves anew to refrain from this immoral and illegal activity, and then acts in accord with this resolution, he winds up doing what he morally and legally ought to do. So, even if he displays a lack of internal fortitude when he wavers with respect to his resolution, the fact that he ultimately steels himself to resist the compulsion to steal and then follows through appears to keep his action from being viewed as weak-willed. The

---

<sup>5</sup> Two-way ANOVA  $F$ -ratio for akratic action variable: ( $F(1, 174) = .063, p > .05$ , partial eta squared = 0. Interaction:  $F(2, 174) = 2.387, p > .05$ , partial eta squared = .027.

<sup>6</sup>  $F(2, 174) = 11.212, p < .001$ , partial eta squared = .114. Planned contrasts and a post-hoc Tukey's HSD test revealed a significant difference (at the .001 level) between the bad condition and the good and neutral conditions. No gender effects were found.

resolution of Franz the Nazi, however, is another matter. He resolves to obey the morally reprehensible orders of his superiors and to imprison every Jewish intellectual he finds. Even if (in the CJ condition) he sincerely believes this is the best thing for him to do, the fact remains that he is violating a very serious normative requirement in doing so. Ordinary participants appear to be quite sensitive to the normative aspects of the cases and thus apportion their attributions of weakness of will accordingly.

The important role that normative considerations play in Experiment 1 provides further support for Sousa and Mauro's (forthcoming) claim that the folk concept of weakness of will is more of a normative and less of a purely descriptive concept than philosophers have traditionally assumed. In a series of experiments they found that when factors such as akratic action and resolution violation are held constant, (i) participants attributed weakness of will to agents who performed immoral actions but not to agents who performed moral ones, and (ii) participants attributed strength of will to agents who performed moral actions but not to agents who performed immoral ones. Because participants applied the concept of weakness of will when blame was appropriate and not otherwise, and they applied the concept of strength of will when credit was appropriate and not otherwise, Sousa and Mauro conclude that a moral evaluation of an agent's action may be built into the concept of weakness of will.

Beebe (forthcoming) proposes a 'normative reasons account' of the folk conception of weakness of will, according to which weakness of will consists in actions performed in violation of undefeated or overriding normative reasons that agents have for acting in certain ways. This account appears capable of explaining the data from Experiment 1, since Franz has normative reasons for not imprisoning Jewish intellectuals, even if he does not appear to recognize or act upon those reasons.

### 3. Experiment 2

As we noted above, Holton (1999, p. 249) suggests that it can be reasonable to reconsider one's intentions "if one believes that they will lead one to great suffering when that suffering was not envisaged at the time of forming the intention." In a brief experiment, I took the two versions of the Franz case described above and modified them so that Franz obtains new information about the consequences of following his orders:

Franz, a Nazi soldier, has been charged with the task of searching every flat in a certain Jewish ghetto and imprisoning every Jewish intellectual he finds there. Franz, however, has been lax in carrying out his duties and has not imprisoned every Jewish intellectual he has found. Franz *believes / has not decided whether he believes* that it would be best for him if he strictly carried out his duties, *so / but* one day he resolves to start imprisoning every Jewish intellectual he finds. The next morning, however, Franz learns that the Jewish intellectuals he has been charged with imprisoning are innocent people who will be executed by prison officials. Consequently, he decides to set aside his earlier resolution and to continue being lax in carrying out his duties. Later that day Franz fails to imprison a Jewish intellectual he encounters in the ghetto.

Participants were then each asked the extent to which they agreed or disagreed with the following two statements:

(Q1) When Franz changes his resolution, he displays weakness of will.

(Q2) When Franz fails to imprison a Jewish intellectual he encounters in the ghetto, he displays weakness of will.

Using Amazon's Mechanical Turk, 60 participants (average age = 35, 53% female, 82% Anglo-American) received either the considered judgment (CJ) or the not considered judgment (not-CJ) version of the case. Results from Experiment 2 are displayed in Figure 2.

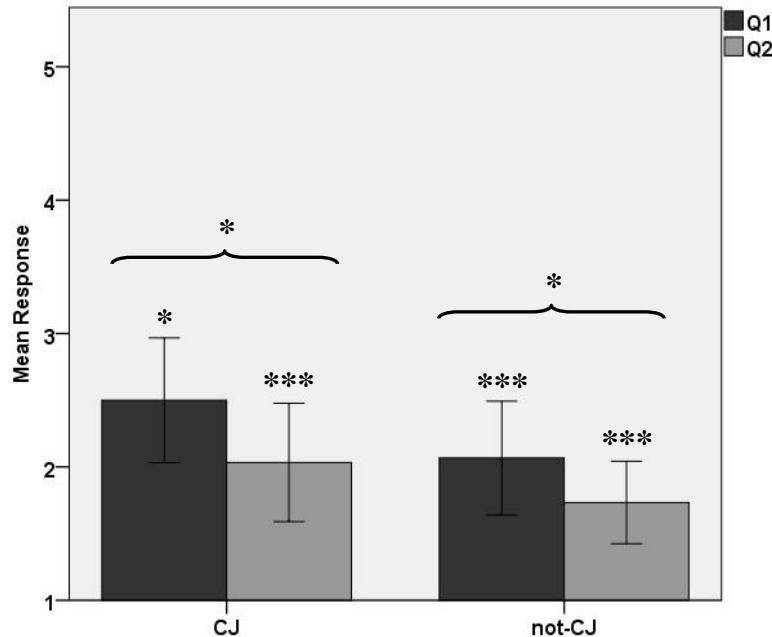


Figure 2. Mean participant responses in the CJ (Q1 = 2.50, Q2 = 2.03) and not-CJ (Q1 = 2.07, Q2 = 1.73) conditions of Experiment 2.<sup>7</sup>

Unsurprisingly, participants were much less inclined to attribute weakness of will to Franz than in previous cases, since he is presently depicted as (i) acting in accord with what he now seems to believe is the best course of action, (ii) acting in accord with the decision he made to set aside his previous resolution, and (iii) revising his previous resolution in a reasonable manner. There was no main effect for the CJ/not-CJ variable (perhaps due to the relatively small sample size), but there was a large effect concerning the difference between the answers

<sup>7</sup> Technically, there should be only one bracket designating a significant main effect for the difference between participant responses to Q1 and Q2. Two brackets seemed preferable to none, given how the data is organized in Figure 16.

participants gave to Q1 and the answers they gave to Q2.<sup>8</sup> This latter fact provides support for Holton's (1999) view that the revision of a resolution can be the locus of weakness of will, independently of whatever subsequent action an agent may perform. The results also provide modest support for the claim that the reasonableness of a resolution revision is an important determinant of whether it will count as weak-willed.

#### 4. Experiment 3

In order to better understand the respects in which the degree of reasonableness of a resolution revision affects participants' willingness to attribute weakness of will, a third experiment was designed that featured a protagonist who backs out of a rash and dangerous act. The cases were inspired by the following reflections from Holton (1999, 249-50):

Suppose I intend to perform a rash and dangerous act: leaping from a high cliff on an untested homemade hang-glider. Would I show weakness of will in revising my intention? That rather depends on why I revise it. If I reassess the chances that the hang-glider will fail, and decide on a program of more cautious preliminary testing, that is surely a reasonable revision. If, on the other hand, I simply suffer a failure of nerve, backing off from the edge of the cliff at the sight of the drop before me, a failure I would have suffered no matter how good the hang-glider, then that is plausibly a case of weakness of will.

The first four cases used in Experiment 3 began with the following description:

---

<sup>8</sup> CJ/not-CJ:  $F(1, 58) = 2.050, p > .05$ , partial eta squared = .034. Q1/Q2:  $F(1, 58) = 9.259, p < .01$ , partial eta squared = .138. If the latter effect does not appear to be large, readers should keep in mind that when two data points are taken from the same participant, the difference between them does not need to be as large as the difference between data points taken from two different participants in order to count as statistically significant or large. No gender effects were found.

Richard is nervous about the prospect of leaping from a high cliff on an untested homemade hang glider.

Each of them ended with:

On the morning of the jump, the sight of the steep drop below the cliff causes Richard to be unable to follow through with his plan.

In between these two sentences appeared descriptions of whether Richard's actions were (i) akratic and violated a resolution (AA/RV), (ii) akratic but did not violate a resolution (AA/not-RV), (iii) not akratic but violated a resolution (not-AA/RV), or (iv) neither akratic nor violating a resolution:

(AA/RV) However, because Richard believes that this dangerous act will earn him the respect of his buddies and will impress an attractive young woman he met recently, he thinks it would be best for him if he undertook this course of action. Richard resolves that he will not let the dizzying height of the cliff cause him to back out of the jump.

(AA/not-RV) However, because Richard believes that this dangerous act will earn him the respect of his buddies and will impress an attractive young woman he met recently, he thinks it would be best for him if he undertook this course of action. But Richard has not resolved that he will not let the dizzying height of the cliff cause him to back out of the jump.

(not-AA/RV) Although Richard believes that this dangerous act will earn him the respect of his buddies and will impress an attractive young woman he met recently, he has not decided whether he believes it would be best for him to undertake this course of action. Nevertheless, Richard resolves that he will not let the dizzying height of the cliff cause him to back out of the jump.

(not-AA/not-RV) Although Richard believes that this dangerous act will earn him the respect of his buddies and will impress an attractive young woman he met recently, he has not decided whether he believes it would be best for him to undertake this course of action. Richard has also not resolved that he will not let the dizzying height of the cliff cause him to back out of the jump.

Participants were then asked to indicate the extent to which they agreed or disagreed with the following claim: “Richard displayed weakness of will in this case.” Responses were recorded on a five-point Likert scale, with ‘1’ marked ‘Strongly Disagree’ and ‘5’ marked ‘Strongly Agree.’

For the sake of comparison, a second set of four vignettes was constructed that was exactly like the previous one, except for the following modification in the way the morning of the jump was described:

On the morning of the jump, Richard reconsiders the extreme danger associated with the jump and decides that it would be better to subject the homemade hang glider to a series of safety tests before attempting to use it to leap from the high cliff.

For ease of reference, I call the second set of cases the ‘reconsiders’ (R) cases and the first set the ‘not reconsiders’ (not-R) cases.

Using Amazon’s Mechanical Turk, 240 participants (average age = 36, 63% female, 71% Anglo-American) received one of the 8 resulting vignettes. Mean responses are represented in Figure 3.



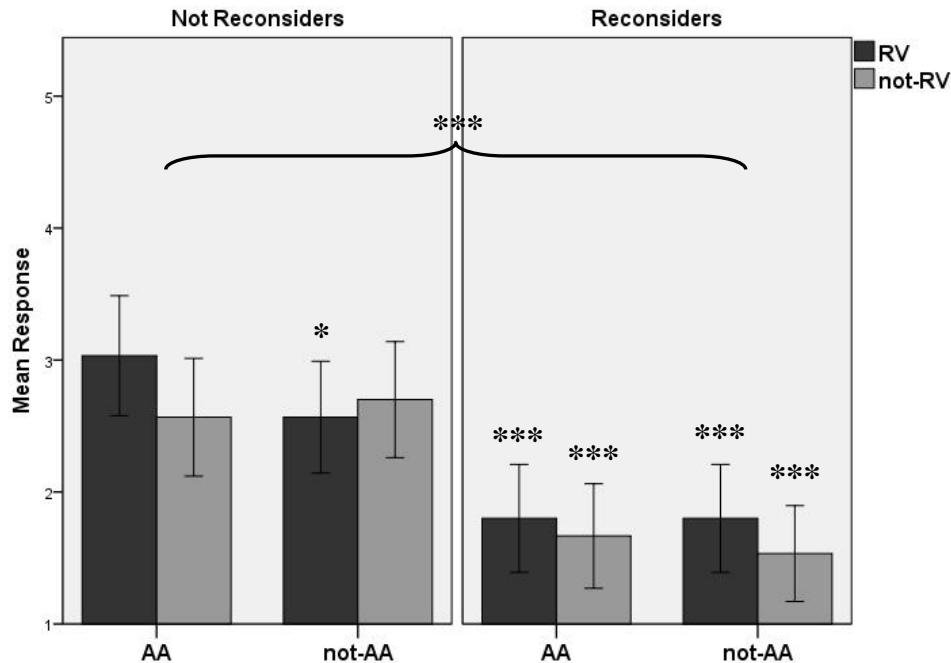


Figure 3. Mean participant responses in the AA/RV/not-R (3.03), AA/not-RV/not-R (2.57), not-AA/RV/not-R (2.57), not-AA/not-RV/not-R (2.70), AA/RV/R (1.80), AA/not-RV/R (1.67), not-AA/RV/R (1.80), and not-AA/not-RV/R (1.53) conditions of Experiment 3.

A three way ANOVA reveals a large main effect for the reconsiders/not reconsiders variable but no main effect for the AA or RV variables.<sup>9</sup> There were also no significant interactions between the variables. Mean responses in the reconsiders conditions all fell significantly below the midpoint, whereas participant responses in the ‘not reconsiders’ conditions generally fell numerically but not significantly below the midpoint.

Participants were thus (i) strongly disinclined to think Richard displayed weakness of will when he was described as reconsidering the extreme danger associated with the jump, and (ii) significantly less inclined to ascribe weakness of will to Richard when he reconsidered the danger than when he was merely described as being caused to be unable to follow through with

<sup>9</sup> Reconsiders:  $F(1, 232) = 49.319, p < .001$ , partial eta squared = .175. AA:  $F(1, 232) = .649, p > .05$ , partial eta squared = .003. RV:  $F(1, 232) = 1.604, p > .05$ , partial eta squared = .007. No gender effects were found.

his plan by the sight of the steep drop. The foregoing results thus provide empirical support for the following claim from Holton (1999, 241):

Not every case of a failure to act on one's intentions is a case of weakness of will. Sometimes, we realize that our intentions were ill-judged, or that circumstances have changed to make them inappropriate.

When we consider that May and Holton (2012), Beebe (forthcoming), and Sousa and Mauro (forthcoming) found that participants were quite willing to attribute weakness of will to agents who acted akratically or violated resolutions, it seems that the fact that the rejected action in Experiment 3 was rash or dangerous had a significant dampening effect on whether participants thought Richard displayed weakness of will in rejecting it. When the rejection was based upon further deliberation, this was seen to be true to an even greater degree. The results of Experiment 3 may also provide further support for Beebe's (forthcoming) normative reasons account of weakness of will because we see that extreme danger is viewed as giving agents normative practical reasons not to perform certain actions and that backing out of an action one has committed oneself to perform is not seen as weak-willed when backing out accords with an agent's normative reasons.

## **5. Experiment 4**

Folk attributions of weakness of will have thus been shown to be sensitive to whether agents (i) acted contrary to or in accord with their better judgments, (ii) reasonably or unreasonably set aside their resolutions to act in certain ways, (iii) acted contrary to or in accord with what morality and their practical reasons require, and (iv) engaged in rational deliberation about the potential consequences of their actions. According to philosophical tradition, addicts or agents

otherwise acting under compulsion should not be viewed as weak-willed regardless of whether they have acted akratically, unreasonably set aside their resolutions, acted contrary to what morality requires, or engaged in rational deliberation about the consequences of their actions. The reason is their lack of freedom. Aristotle, for example, claims that “when nature is the cause, no one would call the people akratic.”<sup>10</sup> Mele (2010, 402-403) concurs:

One necessary condition for an action’s being akratic, on a standard conception, is that it is uncompelled. A heroin addict who judges it best not to use the drug now but uses it now because he has an irresistible desire to do so does not count as akratically using the drug.

Holton (1999, 261), however, expresses the following doubts about this traditional view: First, it is not obvious to me that it would be a disaster if the account [that Holton originally defended] did classify compulsive acts as weak willed; for it is not obvious to me that they are not. Certainly, we would not normally *say* that a compulsive person was weak willed, but that could be because it would be inappropriate rather than false to say so—in the way that it would be inappropriate to say that sadistic torturers are unkind. Are we really averse to saying that compulsives are *pathologically* weak willed?

In another passage Holton (1999, 254) writes:

We surely can ascribe weakness of will to a person who has vowed to give up smoking, and who blithely starts up again straight away, saying that they have changed their mind (again, we would need to fill in the details in the right way). We can make perfect sense of this on the view proposed here: we simply say that they gave up their intention too easily. We need to be sure that the initial intention really was formed, and too sudden a change might make us doubt that; and the revision itself might be seen as involving a

---

<sup>10</sup> *Nicomachean Ethics* (trans. T. Irwin; Indianapolis: Hackett Publishing Co., 1985; 186), 1148b30.

kind of conflict. But there need be no internal conflict at all at the point at which they smoke the cigarette.

Mele (2012, 33) maintains that ‘core weak-willed action’ must be free of compulsion. However, he makes the following prediction about folk intuitions and cases of compulsion:

I would not be surprised if many—and even most—lay respondents to a story in which it is made clear that an agent is moved by an irresistible desire would, if asked, say that he displays some weakness of will. Some respondents who give this response may think that if the agent’s will were not weak, he would be able to resist the desire and that the desire’s being irresistible by him entails some weakness of will. (Mele 2010, 403)

Mele’s prediction can, of course, be straightforwardly tested.

A fourth experiment was thus performed to address the question of whether addicts or others acting under strong compulsions would be taken to display weakness of will when they act in accord with their addictions or compulsions. In Experiment 4 three sets of cases were constructed that combined the variables of akratic action (AA) / no akratic action (not-AA) and resolution violation (RV) / no resolution violation (not-RV) using vignettes about a kleptomaniac, a cigarette smoker, and a cocaine addict. The vignettes about the kleptomaniac each began with ‘Mark is a kleptomaniac, which means he has an uncontrollable urge to steal trivial things’ and ended with ‘The next day, Mark feels an overwhelming compulsion to steal. Unable to resist, Mark steals a magazine from a street vendor.’ Each included one of the following intermediate descriptions:

(AA/RV) Mark believes that it would be best for him if he did not steal, so one day he resolves that he will not steal any more.

(AA/not-RV) Mark believes that it would be best for him if he did not steal, but he has not resolved that he will not steal any more.

(not-AA/RV) Mark has not decided whether he believes that it would be best for him if he did not steal, but one day he resolves that he will not steal any more.

(not-AA/not-RV) Mark has not decided whether he believes that it would be best for him if he did not steal, and he has not resolved that he will not steal any more.

Participants were asked the extent to which they agreed or disagreed with the statement “Mark displayed weakness of will in this case.” Participants were prompted to respond on a five-point Likert scale, with ‘1’ marked ‘Strongly Disagree’ and ‘5’ marked ‘Strongly Agree.’

The components of the cigarette vignette set were the following:

(Intro) Jim smokes 20 cigarettes per day.

(AA/RV) Jim believes that it would be best for him if he did not smoke, so one day he resolves to give up smoking.

(AA/not-RV) Jim believes that it would be best for him if he did not smoke, but he has not resolved to give up smoking.

(not-AA/RV) Jim has not decided whether he believes it would be best for him if he did not smoke, but one day he resolves to give up smoking.

(not-AA/not-RV) Jim has not decided whether he believes it would be best for him if he did not smoke, and he has not resolved to give up smoking.

(Ending) The next day, Jim feels an overwhelming compulsion to smoke. Unable to resist, Jim smokes a cigarette.

Participants were asked the extent to which they agreed or disagreed with the statement “Jim displayed weakness of will in this case.”

The following elements were used to construct the cocaine addict vignette set:

(Intro) Eric is addicted to cocaine.

(AA/RV) Eric believes that it would be best for him if he did not use cocaine, so one day he resolves to stop using the drug.

(AA/not-RV) Eric believes that it would be best for him if he did not use cocaine, but he has not resolved to stop using the drug.

(not-AA/RV) Eric has not decided whether he believes that it would be best for him if he did not use cocaine, but one day he resolves to stop using the drug.

(not-AA/not-RV) Eric has not decided whether it would be best for him if he did not use cocaine, and he has not resolved to stop using the drug.

(Ending) The next day, Eric feels an overwhelming compulsion to use cocaine. Unable to resist, Eric uses cocaine.

Participants were asked to indicate the extent to which they agreed or disagreed with the following claim: “Eric displayed weakness of will in this case.” Note that the moral valence of each of the vignette sets is bad according to common folk evaluations of these practices, with the third set being the worst.

Using Amazon’s Mechanical Turk, 360 participants (average age = 36, 63% female, 71% Anglo-American) received one of the 8 resulting vignettes. Mean responses are represented in Figure 4. Mean participant attributions of weakness of will fell significantly above the midpoint in all but the not-AA/not-RV conditions.

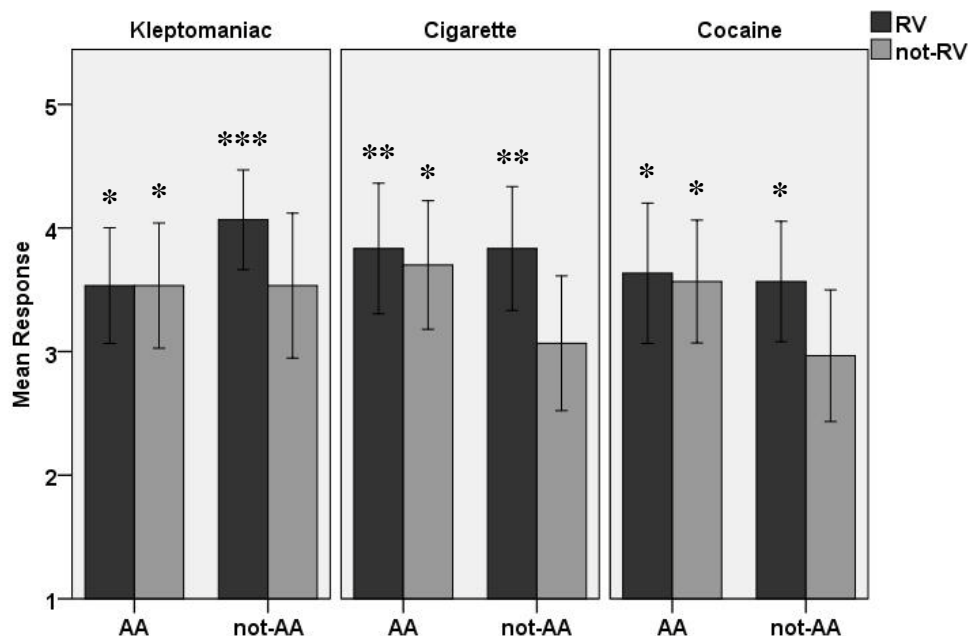


Figure 4. Mean participant responses in the three sets of compulsion cases of Experiment 4.

No main effect was found for the type of compulsion (kleptomaniac, cigarette, cocaine) variable or the AA variable, and only a small main effect was found for the RV variable.<sup>11</sup> There were no significant interactions between these variables, although the interaction between the AA and RV variables approached significance.<sup>12</sup>

The most theoretically significant aspect of the results from Experiment 4 is that in nine of the twelve conditions participants were more inclined to ascribe weakness of will than to remain neutral or to disagree that it should be ascribed. This finding runs contrary to the traditional philosophical view that says agents acting under compulsion cannot properly be said to display weakness of will—at least insofar as that view purports to represent or accord with ordinary thinking on the matter. These findings are also consistent with Beebe's (forthcoming)

<sup>11</sup> Compulsion:  $F(2, 348) = .934, p > .05$ , partial eta squared = .005. AA:  $F(1, 348) = .775, p > .05$ , partial eta squared = .002. RV:  $F(1, 348) = 5.818, p < .05$ , partial eta squared = .016.

<sup>12</sup>  $F(1, 348) = 3.813, p = .052$ , partial eta squared = .011. No gender effects were found.

normative reasons account of weakness of will, which does not require weak-willed agents who act in violation of undefeated or overriding normative reasons for acting to do so freely or in the absence of compulsion. It is neutral on the questions of whether addicts' actions are free and whether akratic action can be distinguished from acting under compulsion.<sup>13</sup> However, it predicts that ordinary subjects will take addicts' actions to be weak-willed when they violate normative practical reasons, which can stem from evaluative judgments and inclination-defeating resolutions the addicts have made.

Sousa and Mauro's (forthcoming) perspective also seems relevant to the present results. As we noted above, they argue that moral evaluation of an agent's action may be built into the concept of weakness of will on the basis of finding that participants attributed weakness of will to agents who performed immoral actions but not to agents who performed moral ones and attributed strength of will to agents who performed moral actions but not to agents who performed immoral ones. The results of Experiment 4 comport well with their perspective, insofar as each of the addicts' actions are bad or blameworthy in some respect. However, Sousa and Mauro's suggestion that the effects of akratic action and resolution violation variables on folk attributions of weakness of will may be negligible runs contrary to many of the results described above, in which participants display significant sensitivity to the variables of akratic action and resolution violation.

#### **4. Conclusion**

Commenting on the relevance of folk judgments to philosophical theorizing, Mele (2001, p. 27) wrote:

---

<sup>13</sup> Cf. Buss (1997), Tenenbaum (1999), and Mele (1986; 2002) for further discussion.



It is also worth noting that if there is a widely shared concept of intentional action, such [folk] judgments provide evidence about what that concept is, and a philosophical analysis of intentional action that is wholly unconstrained by that concept runs the risk of having nothing more than a philosophical fiction as its subject matter.

In the philosophical literature on weakness of will, philosophers have maintained a tacit commitment to ensuring that their accounts of weakness of will are consistent with ordinary judgments about particular cases. In accordance with this commitment, experimental philosophers have begun to investigate folk judgments more systematically in order to inform the philosophical debate. The present studies attempt to make such a contribution.

In Experiment 1 we failed to find evidence for Holton's claim that in some circumstances agents should be viewed as displaying weakness of will because of their unreasonable and hasty setting aside of their resolutions, even if they actually perform no weak-willed actions. Further empirical investigation of this claim is required before we can be confident about the extent to which it may or may not be correct. Experiment 1 did, however, uncover further evidence of the importance of the moral valence of actions to folk attributions of weakness of will.

In accord with Holton's predictions, in Experiment 2 we found that a protagonist who reasonably revises his resolution in light of new information about the consequences his actions will not be viewed as weak-willed. Also in accord with Holton's predictions, Experiment 3 showed that an agent who violates a resolution to undertake an unreasonable and dangerous course of action will not be viewed as weak-willed, especially if the agent explicitly reflects upon the potential danger involved.

Experiment 4 tested folk intuitions about whether agents who act under compulsion or in accord with addictions can display weakness of will. Contrary to the received philosophical view

on the matter, participants were willing to attribute weakness of will to kleptomaniacs, smokers, and cocaine addicts, and seriousness of the addictions did not affect their willingness to do so.

Promising topics for future investigation of the folk conception of weakness of will include the question of whether the folk would count as weak-willed agents who (i) akratically change their minds about the best course of action (Mele 1986, 674), (ii) akratically fail to apprehend the conclusion that is best supported by their practical premises (Mele 1986, 674), (iii) misinterpret the significance of their present situation due to akrasia (Mele 1986, 674), (iv) “act contrary to better judgments they made some time ago, did not revise, and are not aware of now” (Mele 2012, 7), (v) are temporarily insane (Mele 2012, 8), (vi) suffer from extreme listlessness (Mele 2012, 9), or (vii) act against a judgment about which course of action is best when that judgment is not rationally made or held (Mele 2012, 10).

**References**

- Beebe, J. R. forthcoming. The folk conception of weakness of will.
- Audi, R. 1979. Weakness of will and practical judgment. *Noûs* 13:173-96.
- Buss, S. 1997. Weakness of will. *Pacific Philosophical Quarterly* 78: 13-44.
- Davidson, D. 1970. How is weakness of will possible? In Joel Feinberg (ed.), *Moral Concepts*. Oxford: Oxford University Press, pp. 93-113.
- Hare, R. M. 1952. *The Language of Morals*. Oxford: Clarendon Press.
- Hare, R. M. 1963. *Freedom and Reason*. Oxford: Clarendon Press.
- Holton, R. 1999. Intention and weakness of will. *Journal of Philosophy* 96: 241-62.
- Holton, R. 2003. How is strength of will possible? In C. Tappolet & S. Stroud (Eds.), *Weakness of Will and Practical Irrationality*. Oxford: Oxford University Press, pp. 39-67.
- Holton, R. 2009. *Willing, Wanting, Waiting*. Oxford: Oxford University Press.
- May, J., and R. Holton. 2012. What in the world is weakness of will? *Philosophical Studies* 157: 341-60.
- Mele, A. 1986. Is akratic action unfree? *Philosophy and Phenomenological Research* 46: 673-9.
- Mele, A. 2001. Acting intentionally: Probing folk notions. In B. Malle, L. Moses, and D. Baldwin (eds.), *Intentions and Intentionality: Foundations of Social Cognition*. Cambridge, MA: MIT Press, pp. 27-43.
- Mele, A. 2002. Akratics and addicts. *American Philosophical Quarterly* 39:153-67.
- Mele, A. 2010. Weakness of will and akrasia. *Philosophical Studies* 150: 391-404.
- Pugmire, D. 1982. Motivated irrationality. *Proceedings of the Aristotelian Society*, Suppl. Vol. 56, pp.179-196.

Sousa, P., and C. Mauro. forthcoming. The evaluative nature of the folk concepts of weakness and strength of will.

Tenenbaum, S. 1999. The judgment of a weak will. *Philosophy and Phenomenological Research* 59: 875-911.

Watson, G. 1977. Skepticism about weakness of will. *Philosophical Review* 86: 316-39.

Wiggins, D. 1987. Weakness of will, commensurability, and the objects of deliberation and desire. In *Needs, Values, Truth*. Cambridge: Basil Blackwell, pp. 239-67.