

# The Content-Dependence of Imaginative Resistance

Hanna Kim, Markus Kneer and Michael T. Stuart<sup>1</sup>

## 1. Introduction

Imaginative resistance “occurs when an otherwise competent imaginer finds it difficult to engage in some sort of prompted imaginative activity” (Gendler and Liao 2016). There are now many puzzles surrounding this phenomenon (see e.g., Gendler and Liao 2016, Kieran and Lopes 2003, Todd 2009, Walton 2006, Weatherson 2004). In this paper, we focus on the oldest and most widely discussed puzzle, which was baptized by Tamar Gendler in her (2000), and traces back to Kendall Walton, Richard Moran, and ultimately, David Hume.

This version of the puzzle of imaginative resistance involves “our inability or unwillingness to imagine counter-moral propositions in fiction” (Kind and Kung 2016: 23). In Hume’s words, there are many things in a fiction which we know are descriptively wrong, and that will “detract but little from the value of those compositions” (Hume 1757, para. 33). For example, we seem to have no problem with leprechauns, hobbits and vampires, which we know do not exist. But “a very violent effort is requisite to change our judgment of manners, and excite sentiments of approbation or blame, love or hatred, different from those to which the mind, from long custom, has been familiarized” (ibid.). For example, if Shakespeare had written all the circumstances of Duncan’s murder at the hands of Macbeth exactly as we know them, but added that Macbeth’s actions were morally praiseworthy, this would have been very difficult for us to imagine (Moran 1994). The puzzle, then, concerns how we can explain the difference in difficulty between imagining *counterevaluative* and *counterdescriptive* propositions.<sup>2</sup>

The empirical assumption underlying the puzzle is that people do indeed experience more imaginative resistance when they attempt to imagine scenarios that are evaluatively deviant rather than descriptively deviant. We apparently have a harder time imagining that something ugly is beautiful, that something morally wrong is right, that something dull is funny, or that something clumsy is elegant, than we do imagining something that is descriptively false.

This “curious asymmetry” (Kieran and Lopes 2003: 8; Matravers 2003: 91) is widely assumed in the literature, not very often argued for, and frequently restricted to *moral* deviance. For example, Moran asks,

Why can we not, as it seems, treat the judgments of morality and decency the same way we treat any other judgments, and accept as fictionally true what the story tells us (or implies) is true, and comfortably leave our genuine attitudes at the door? What happens to our sense of distance at that point, the distance between what we can imagine and what we actually believe? The suggestion here is that we cannot treat these as on a par with other fictional truths. (1994: 97)

Magdalena Balcerak Jackson writes that “[m]ost convincing examples of imaginative resistance involve requests to imagine situations where morally highly deviant behaviors and attitudes are endorsed” (2016: 47). Neil van Leeuwen wonders why “it seems easy to incorporate outlandish descriptive propositions into our understanding of the story, but our minds are far less flexible about incorporating outlandish moral propositions” (2016: 103-4).<sup>3</sup>

Like Hume and Moran, Walton uses introspection and thought experiments to motivate the claim that imagining counterevaluatives is more difficult than imagining counterdescriptives (1990: 154-5; 1994: 32 ff.), though he expands the claim beyond morality to other evaluatives. For

example,

If in a story a comedian tells [“a really dumb joke”] and the author simply writes explicitly in the text that it is hilariously funny, I expect that I would attribute a juvenile or an incomprehensible sense of humor to the narrator, and stick with my own judgment that the joke is not funny. I insist on applying my own sense of humor. (Walton 1994: 40)

This example of Walton’s shows that it isn’t just *moral* deviance that we resist, but counterevaluatives of many kinds (see also Yablo 2002: 485). As we said above, this was Hume’s original point. In this paper, then, we will speak of the puzzle of imaginative resistance in terms of counterevaluatives broadly conceived.

What sorts of reasons are there for believing that imaginative resistance is more powerful for counterevaluatives than for counterdescriptives? The most common strategy is to provide a few short narratives or narrative sketches to pump intuitions into agreement. “We happily go along with talking mice and time-travel tales, but we balk if recreational torture is endorsed, or presented as truly permissible within the story” (Driver 2008: 302). A second strategy differentiates between kinds of imagination: one kind draws on the faculty of sentiment, and another on (what Hume called) the faculty of understanding. Desire-like (Currie 2002) or value-like (Stokes 2006) imagination on the sentiment side is more difficult to control than cognitive imagination on the understanding side. If this is true, it would explain why we experience more resistance with counterevaluatives than counterdescriptives. A third reason is that while we often import real beliefs, desires and values into fictions (e.g., we assume Sherlock Holmes has at least one kidney because we believe all humans do), we can also export imaginings into real beliefs, desires and values (e.g., we might come to believe, desire and value things about 19<sup>th</sup> century London because

of Dickens's novels). Imaginative resistance is then argued to be about what we *ought* to imagine, given that we risk exporting what we imagine into what we believe, desire and value. We ought to resist imagining any counterevaluative that might "contaminate" (that is, negatively affect) our moral framework (Gendler 2000). A fourth reason might be that there are no counterdescriptives that cannot be imagined. One of the best contenders for counterdescriptives that we cannot imagine are conceptual impossibilities, which Gendler argues can be imagined after all (Gendler 2000). If we can imagine that the sum of five and seven both is and is not equal to twelve, then certainly we can imagine any descriptive claim in a fiction. And this is not so, Gendler claims, for counterevaluatives.

But there are also good reasons for doubting that it is more difficult to imagine counterevaluatives than counterdescriptives (see, e.g., Mothersill 2006, Tanner 1994, Todd 2009, Weatherston 2004, Yablo 2002). One is that Hume's distinction between sentiment and understanding might be cognitively crude. If there is no defensible distinction between the faculties of sentiment and understanding, or if such faculties are unmotivated, or do not exist, the asymmetry disappears. Another is that we might not believe there are moral facts, in which case, there can be no moral *counterevaluatives*, and again the distinction loses its bite. Or perhaps we could deny any strict distinction between facts and values (see Marchetti and Marchetti 2017). We might also target Gendler's claim that conceptual impossibilities can be imagined. Stock (2003), for example, argues that they cannot.

To help decide the issue, we could turn to experiment. Black and Barnes (2017) show that there is quite a bit of individual variability in the amount of imaginative resistance people experience, and also that imaginative resistance is positively correlated with moral features (e.g., moral authority, moral harm, disgust sensitivity, fear of moral contagion, etc.). If this is right,

people do experience some form of imaginative resistance when exposed to certain counterevaluatives. What we are interested in, however, is the assumption on which much of the work on this topic is premised: whether it really is *more difficult* to imagine a counterevaluative than a counterdescriptive.

In our study, we investigated whether the degree to which a proposition was counterevaluative or counterdescriptive influenced levels of imaginative resistance. Differently put, we explored whether imaginative resistance is sensitive not only to proposition *type* (evaluative vs. descriptive) but *content*, independently of type. We found that it was. A pilot study indicated that, in general, counterevaluative claims generate significantly more resistance than counterdescriptive claims. However, the degree of counterfactuality also had a significant effect on resistance judgments.<sup>4</sup> Differently put, imaginative resistance cannot be accounted for purely in terms of claim *type*. Instead, features of content (of which degree of counterfactuality is but one) seem to have an impact, too.

The pilot study raised a worry for imaginative resistance traditionally conceived: What if the detected overall difference in resistance between counterevaluative and counterdescriptive claims is not in fact due to claim *type*, but instead to features of content for which the experiment does not control? More precisely, resistance might be triggered by contents that represent states of affairs that are considered unlikely, astonishing, unusual, etc. For simplicity, we introduce a catch-all term for potential resistance-inducing properties of this sort: “weirdness.” It might turn out that the counterevaluatives invoked in the literature as those that produce imaginative resistance (e.g., the proposition that the practice of genocide or slavery is morally acceptable, or that it is evil to associate with people of other races (Walton, 1994: 28)) are simply “weirder” than the counterdescriptives with which they are standardly compared, (e.g., the proposition that there is a

ring that makes its wearer invisible, or that a village in Scotland appears and disappears every one hundred years (Walton, 1994: 31)). If this turns out to be correct, the difference in imaginative resistance triggered by different claims would thus not be due to claim *type*, but weirdness of claim *content*. Our study addressed this worry by exploring imaginative resistance across claim types while controlling for a variety of features that might render their content weird.

## 2. Experiment

### 2.1 Participants

1216 participants were recruited on Amazon Mechanical Turk to complete a paid online survey on Qualtrics. The IP location was restricted to the US. Subjects who failed an attention test, took less than 15 seconds to complete the main task, changed their response more than 10 times, or had native languages that were not English were excluded. The exclusion criteria were determined in advance of the experiment. 845 subjects remained, of whom 441 were female. The average age was 38.8 years ( $SD=12.3$  years).

### 2.2 Method and Materials

Participants were randomly assigned to one of twenty-four conditions in a 6 (claim type: evaluative 1 vs. evaluative 2 vs. evaluative 3 vs. descriptive 1 vs. descriptive 2 vs. descriptive 3) x 4 (degree of counterfactuality: low vs. medium-low vs. medium-high vs. high) design. Evaluative 1 involved a moral norm, evaluative 2 involved a humor norm, and evaluative 3 involved an aesthetic norm. Each participant responded to three questions regarding weirdness and three questions regarding imaginative resistance (truth, difficulty, and possibility). Appendix 2 in the Supplementary

Materials contains the complete text of all scenarios and questions.<sup>5</sup> We will use the aesthetic condition as an example. The prompt and questions (variations of degree in square brackets) read:

Adaleine, Picasso's greatest student, was a prolific painter, whose work was unfortunately lost to history - with the exception of her last painting: A 3 x 4' canvas, painted from edge to edge in the exact same shade of yellow as the McDonald's golden arches. It is [a somewhat beautiful painting<sub>i</sub> / a beautiful painting<sub>ii</sub> / a very beautiful painting<sub>iii</sub> / without doubt one of the most beautiful works ever made<sub>iv</sub>].

Q1 (weirdness 1): How unusual is it that Adaleine's painting is [a somewhat beautiful painting<sub>i</sub> / a beautiful painting<sub>ii</sub> / a very beautiful painting<sub>iii</sub> / without doubt one of the most beautiful works ever made<sub>iv</sub>]?

(1=completely ordinary; 7=completely unusual)

Q2 (weirdness 2): If there is a world where Adaleine's painting is [a somewhat beautiful painting<sub>i</sub> / a beautiful painting<sub>ii</sub> / a very beautiful painting<sub>iii</sub> / without doubt one of the most beautiful works ever made<sub>iv</sub>], how different would this world be from ours? (1=completely the same; 7=completely different)

Q3 (weirdness 3): How surprised would you be if you found out that Adaleine's painting is [a somewhat beautiful painting<sub>i</sub> / a beautiful painting<sub>ii</sub> / a very beautiful painting<sub>iii</sub> / without doubt one of the most beautiful works ever made<sub>iv</sub>] in our world? (1=completely unsurprised; 7=completely surprised)

Q4 (truth): If you were to find the previous scenario within a work of fiction, to what extent would you agree that the following statement is true in the fictional scenario? (1=completely agree, 7=completely disagree)

“Adaleine's painting is [a somewhat beautiful painting<sub>i</sub> / a beautiful painting<sub>ii</sub> / a very beautiful painting<sub>iii</sub> / without doubt one of the most beautiful works ever made<sub>iv</sub>].”

Q5 (difficulty): How difficult is it for you to imagine that Adaleine's painting is [a somewhat beautiful painting<sub>i</sub> / a beautiful painting<sub>ii</sub> / a very beautiful painting<sub>iii</sub> / without doubt one of the most beautiful works ever made<sub>iv</sub>]?

(1=very easy, 7=very difficult)

Q6 (possibility): To what extent do you deem it possible to imagine that Adaleine's painting is [a somewhat

beautiful painting<sub>i</sub> / a beautiful painting<sub>ii</sub> / a very beautiful painting<sub>iii</sub> / without doubt one of the most beautiful works ever made<sub>iv</sub>]?) (1=completely possible, 7=completely impossible)

Participants responded to all questions on Likert scales ranging from 1 to 7, which were anchored as specified above. The order of test questions was fixed. The experimental prompts were preceded by an attention check and followed by a demographic questionnaire.

## 2.3 Results

Since we were interested in the impact of claim type and degree, we aggregated across scenarios of the same claim type. We conducted a mixed ANOVA with Greenhouse-Geisser correction (within-subjects factor: measure – truth vs. difficulty vs. possibility; between-subjects factors: claim type – evaluative vs. descriptive, degree – low vs. medium-low vs. medium-high vs. high). There was a significant main effect for degree  $F(3,837)=36.18, p<.001, \eta_p^2=.115$ , a significant main effect for measure  $F(1.80,1504.26)=47.06, p<.001, \eta_p^2=.053$ , a significant main effect for claim type  $F(1,837)=37.68, p<.001, \eta_p^2=.043$ . The measure\*degree interaction was significant  $F(5.39, 1504.26)=9.63, p<.001, \eta_p^2=.033$ , and so were the measure\*claim type interaction  $F(1.80, 1504.26)=33.30, p<.001, \eta_p^2=.038$  and the degree\*claim type interaction  $F(3,837)=3.18, p=.023, \eta_p^2=.011$ . The three-way interaction was not significant  $F(5.40, 1504.26)=.78, p=.577, \eta_p^2=.003$ .

To further explore the effect of claim type and degree of counterfactuality on the three distinct measures of imaginative resistance, we ran an ANOVA each for truth, difficulty and possibility judgments. With truth as the dependent variable (Figure 7.1), claim type proved significant  $F(1,837)=93.38, p<.001, \eta_p^2=.10$ , and so did degree  $F(3,837)=5.62, p=.001, \eta_p^2=.02$ . The interaction was not significant  $F(3,837)=1.03, p=.38, \eta_p^2=.004$ .

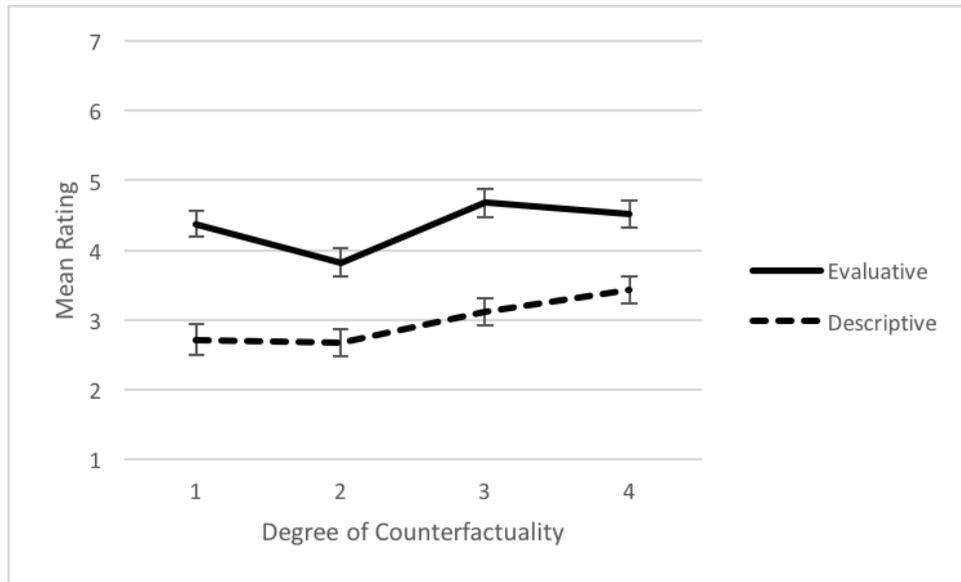


Figure 7.1. Mean ratings of imaginative resistance in terms of truth for evaluative and descriptive claims across degrees of counterfactuality. Error bars designate standard error of the mean.

For the dependent variable difficulty (Figure 7.2), claim type did not prove significant  $F(1,837)=3.25, p=.07, \eta_p^2=.004$ ), degree was significant  $F(3,837)=35.16, p<.001, \eta_p^2=.11$ , and the interaction was not significant  $F(3,837)=1.57, p=.20, \eta_p^2=.006$ .

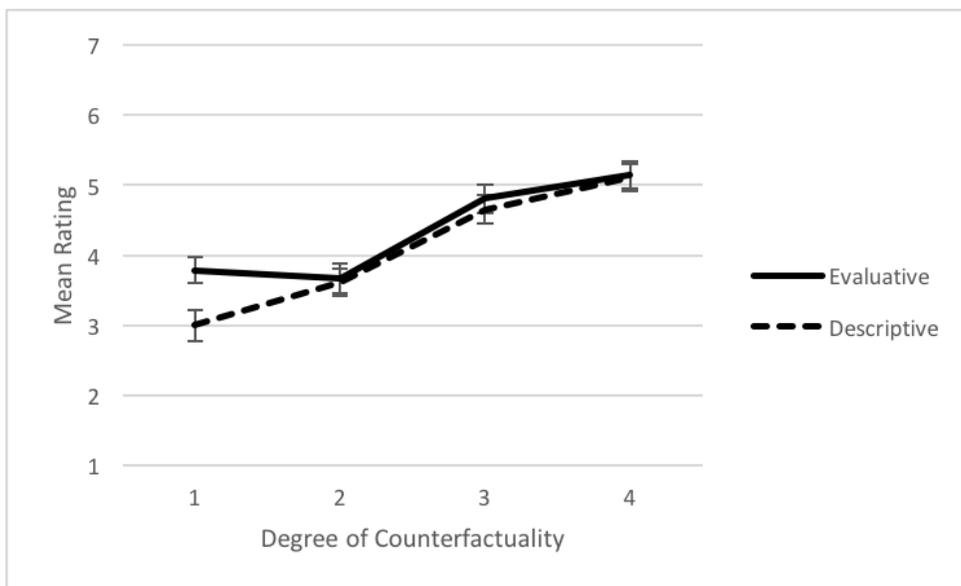


Figure 7.2. Mean ratings of imaginative resistance in terms of difficulty for evaluative and descriptive claims across degrees of counterfactuality. Error bars designate standard error of the mean.

For the dependent variable possibility (Figure 7.3), claim type proved significant  $F(1,837)=7.49$ ,  $p=.006$ ,  $\eta_p^2=.01$ , degree was significant  $F(3,837)=38.03$ ,  $p<.001$ ,  $\eta_p^2=.12$ , and the interaction was also significant ( $F(3,837)=4.45$ ,  $p=.004$ ,  $\eta_p^2=.02$ ).

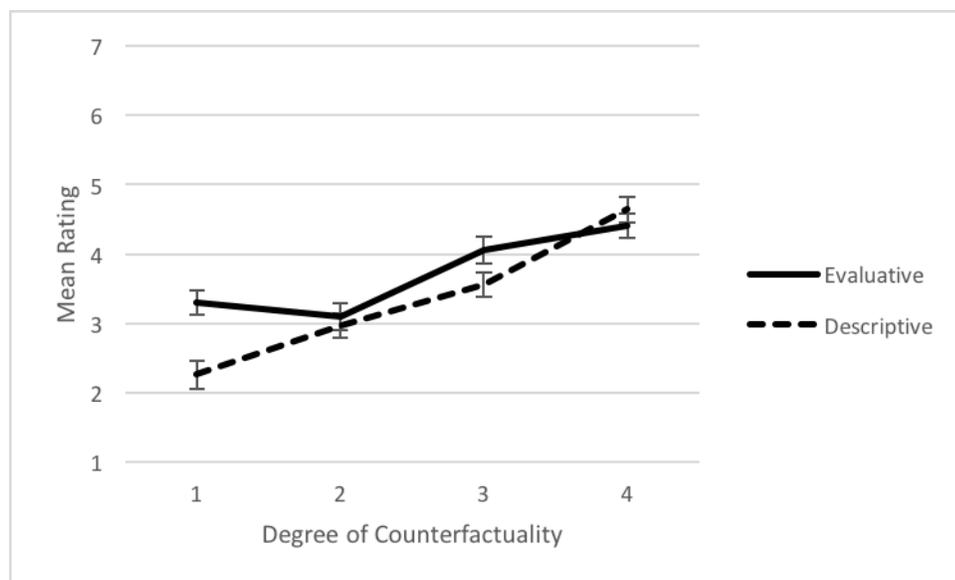


Figure 7.3. Mean ratings of imaginative resistance in terms of possibility for evaluative and descriptive claims across degrees of counterfactuality. Error bars designate standard error of the mean.

In short, while claim type has a considerable impact on imaginative resistance judgments regarding truth ( $\eta_p^2=.10$ , a medium effect), it has next to no effect on either difficulty or possibility judgments ( $\eta_p^2<.01$ , i.e. not even a small effect size). By contrast, degree of counterfactuality has a pronounced effect on difficulty ( $\eta_p^2=.11$ ) and possibility judgments ( $\eta_p^2=.12$ ), yet a much smaller

effect on truth judgments ( $\eta_p^2=.02$ ).

Next, we explored whether claim type still has an impact on imaginative resistance once weirdness of content is controlled for. Recall that for each target claim  $c$ , participants rated how unusual they consider  $c$ , to what extent a world in which  $c$  is true would differ from ours, and how surprised they would be to learn that  $c$  is actually the case. Averaging across the results, we calculated a weirdness composite score. Across conditions, reliability analyses produced a Cronbach's alpha of .75 among the questions used to compute the weirdness score, indicating that the composite score was strongly internally consistent. On average, weirdness scores for the counterdescriptive claims ( $M=5.31$ ,  $SD=1.45$ ) were significantly higher than those for counterevaluative claims ( $M=4.77$ ,  $SD=1.60$ ),  $t(843)=-5.19$ ,  $p<.001$ , Cohen's  $d=.36$ , a small effect.<sup>6</sup>

We also conducted hierarchical multiple regressions (HMRs) in order to explore whether claim type had an impact on the three measures of imaginative resistance, once weirdness of content had been controlled for. Table 7.1 summarizes the HMR results for all three types of imaginative resistance judgments.

<i>Predictor</i>	<u><i>Truth</i></u>				<u><i>Difficulty</i></u>				<u><i>Possibility</i></u>			
	<i>B</i>	<i>SE b</i>	<i>B</i>	<i>p</i>	<i>B</i>	<i>SE b</i>	$\beta$	<i>p</i>	<i>b</i>	<i>SE b</i>	$\beta$	<i>p</i>
<i>Step 1</i>												
Constant	1.965	0.245		<.001	0.591	0.218		<.01	0.382	0.204		Ns
Weirdness	0.344	0.047	0.247	<.001	0.73	0.041	0.519	<.001	0.636	0.039	0.492	<.001
<i>Step 2</i>												
Constant	2.293	0.229		<.001	0.716	0.217		<.01	0.517	0.202		<.001
Weirdness	0.435	0.044	0.312	<.001	0.764	0.041	0.544	<.001	0.673	0.039	0.521	<.001

Claim Type   -1.595   0.136   -0.369   <.001   -0.609   0.129   -0.14   <.001   -0.657   0.12   -0.164   <.001

Table 7.1. Hierarchical regression analyses for truth ( $R^2=.061$  for Step 1;  $\Delta R^2=.132$  for Step 2, all  $ps<.001$ ), difficulty ( $R^2=.270$  for Step 1;  $\Delta R^2=.019$  for Step 2, all  $ps<.001$ ) and possibility ( $R^2=.242$  for Step 1;  $\Delta R^2=.026$  for Step 2, all  $ps<.001$ ).

*Truth:* The first hierarchical multiple regression revealed that at stage one, weirdness contributed significantly to the regression model,  $F(1,844)=54.71$ ,  $p<.001$ , and accounted for 6.1% of the variation in truth judgments. Introducing claim type explained an additional 13.2% of variation in truth judgments and this change in  $R^2$  was significant,  $F(2,844)=100.67$ ,  $p<.001$ . Together the two independent variables accounted for 19.3% of the variance in truth judgements. Furthermore, the positive value of  $b=.435$  for weirdness shows that weirdness was a positive predictor of truth – i.e., the weirder the claim, the less likely one was to accept it as true in the fiction. The negative value of  $b$  for claim type (i.e., -1.595) indicates that evaluative claims (coded as 0) were less likely to be accepted as true in the fiction than descriptive ones (coded as 1).

*Difficulty:* The second hierarchical multiple regression revealed that at stage one, weirdness contributed significantly to the regression model,  $F(1,844)=311.43$ ,  $p<.001$ , and accounted for 27.0% of the variation in difficulty judgments. Introducing claim type only explained an additional 1.9% of variation in difficulty judgments and this change in  $R^2$  was significant,  $F(2,844)=170.90$ ,  $p<.001$ . Together the two independent variables accounted for 28.9% of the variance in difficulty judgments. As above, the weirder the claim, the more difficult it was to imagine it ( $b>0$ ). Furthermore, descriptive claims were deemed less difficult to imagine than evaluative ones ( $b<0$ ).

*Possibility*: The third hierarchical multiple regression revealed that at stage one, weirdness contributed significantly to the regression model,  $F(1,844)=269.14$ ,  $p<.001$ , and accounted for 24.2% of the variation in possibility judgments. Introducing claim type only explained an additional 2.6% of variation in possibility judgments and this change in  $R^2$  was significant,  $F(2,844)=154.18$ ,  $p<.001$ . Together the two independent variables accounted for 26.8% of the variance in possibility judgments. Again, the weirder the claim, the more it was considered impossible to imagine it ( $b>0$ ). Furthermore, descriptive claims were deemed less impossible to imagine than evaluative ones ( $b<0$ ).

In a final analysis we double-checked the results of the HMR analyses. The six scenarios, each of which invoked target claims differing in terms of degree of counterfactuality made for 24 individual conditions. We conducted a Tukey post-hoc test for the weirdness scores of the 24 conditions, and used the largest homogenous subset to explore the impact of claim type on the three measures of imaginative resistance. Differently put, we explored whether claim type still has an impact on judgments of imaginative resistance (truth, difficulty, and possibility), when the claims at stake are deemed weird to similar extents. Consistent with the HMR results, imaginative resistance conceived in terms of truth in fiction is significantly higher for evaluative claims than for descriptive claims ( $p<.001$ ), though no significant difference could be detected across claim types for difficulty judgments ( $p=.32$ ) or possibility judgments ( $p=.63$ ). For the full analysis, cf. Appendix 3 in the Supplementary Materials.<sup>7</sup>

## **2.4. Discussion**

So far, we have proceeded under the assumption that judgments regarding truth in fiction, difficulty and the possibility of imagining a non-actual state of affairs all belong to a single,

uniform category. They all capture different, yet not unrelated, aspects of imaginative resistance (where imaginative resistance is understood in a broad sense), just as judgments regarding wrongness, blame, permissibility and punishment are standardly conceived of as different types of moral judgments. However, given the results, it might be helpful to draw more attention to the distinction between imaginative resistance conceived in terms of truth judgments on the one hand, and in terms of difficulty and possibility on the other. Our findings, we would like to suggest, might support a *Dual Process Model of Imaginative Resistance*, a model which is structurally (and structurally only) similar to Cushman’s (2008, 2013) *Dual Process Model of Moral Judgment*. According to the latter, wrongness and permissibility judgments are principally sensitive to mental states, whereas blame and punishment judgments are sensitive both to mental states and causal factors and outcomes. In the case of imaginative resistance, we found truth in fiction (or “fictionality”) judgments to be sensitive both to claim type (evaluative vs. descriptive) and features of content (degree of counterfactuality, or weirdness); difficulty and possibility judgments, by contrast, were found to be sensitive to features of content only (Figure 7.4).<sup>8</sup>

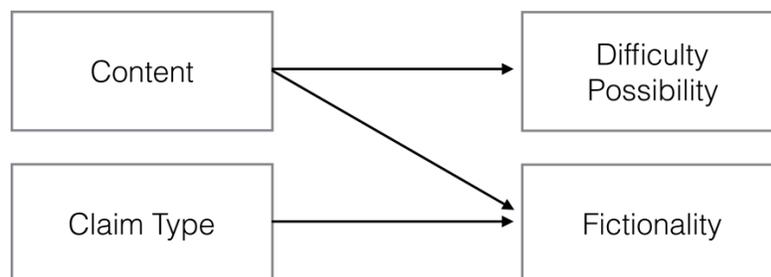


Figure 7.4. The Dual Process Model of Imaginative Resistance

### 3. General discussion

There are many questions to ask concerning imaginative resistance. Some relate to its scope. For

instance: Does imaginative resistance arise from certain moral propositions, evaluative propositions more generally, or also for counterdescriptive propositions? Do we resist any of these types of proposition more than others? And what explains the varying amounts and types of resistance? In the following, we will briefly trace out the implications of our findings for these questions. In closing we will address some possible directions for future research.

Regarding scope, our pilot data showed that imaginative resistance does exist for non-moral counterevaluative claims, as well as for counterdescriptive claims (see also Black and Barnes 2016). In other words, imaginative resistance has extremely wide scope. Do the different kinds of content face different levels of resistance in principle? We found that resistance can be stronger for non-moral counterevaluatives than for moral counterevaluatives. Thus, if Hume's asymmetry did obtain, Hume, Walton and Yablo would be correct that it holds for counterevaluatives of all kinds, rather than merely moral counterevaluatives. The overwhelming focus in the literature on specifically *moral* counterevaluatives might therefore be somewhat inappropriate. We have also shown that resistance can be more pronounced for certain counterdescriptives than for certain counterevaluatives. Therefore, the strong claim that counterevaluatives always produce more imaginative resistance than counterdescriptives is false.

More importantly, the results from our experiment show that out of the three measures of imaginative resistance tested (truth, difficulty, possibility), claim type matters only for truth judgments (that is, whether something is judged to be true in the fiction). And even in this case, claim type is neither the only significant factor driving resistance (degree is also significant:  $(F(3,837)=36.18, p<.001, \eta_p^2=.115)$ ), nor is its impact particularly pronounced (the effect size of the main effect of claim type on resistance was medium-small,  $\eta_p^2=.043$ ). This suggests that, with respect to any of the three measures tested here, there is no principled asymmetry we can draw

between counterevaluatives and counterdescriptives. Hence the alleged puzzle, according to which a difference in imaginative resistance is due *exclusively* or, at the very least, *predominantly* to claim type, simply does not exist. On the other hand, if Hume's original puzzle is taken to refer only to imaginative resistance *in the sense of resistance to truth in fiction judgments*, and his observation is merely that claim type *partially explains* such resistance, the puzzle remains, though in a much weaker form than usually assumed.

Turning to explanations of imaginative resistance, we found that the “weirdness” of a claim was the main explanatory factor for resistance conceived in terms of difficulty and possibility, and a partial predictor of resistance conceived in terms of judgments of truth in fiction. As long as one wants to postulate a general concept of imaginative resistance that captures a variety of distinct, though related phenomena, one might want to adopt a dual process model: difficulty and possibility judgments are influenced principally by features of content (as measured by degree of weirdness), whereas truth in fiction judgments are influenced by both content and claim type. A number of questions now stand out. First, what is weirdness? In our study it was an amalgamation of (i) unusualness, (ii) difference from the actual world, and (iii) surprisingness. But how exactly do these factors contribute to imaginative resistance? Second, what is it about truth in fiction judgments that makes them turn on content *and* claim type, while difficulty and possibility judgments are only tied to weirdness? We take up these questions in turn.

### **3.1 Unusualness**

“We happily go along with talking mice and time-travel tales, but we balk if recreational torture is endorsed, or presented as truly permissible within the story” (Driver 2008: 302). Perhaps this is because fictions endorsing recreational torture are more unusual than talking mice tales, in the

sense that most of us have some experience with fables, movies and cartoons that portray non-human creatures that talk, but we don't typically have as much experience with stories that involve morally permissible recreational torture. If imaginative resistance is a function of unusualness, and unusualness is a function of previous experience, then, a fortiori, imaginative resistance is a function of previous experience. This is just as Hume said: we resist imagining things "different from those to which the mind, from long custom, has been familiarized."

### **3.2 Difference from the Actual World**

Our reaction to time travel vs. torture might not rely on the perceived difference between the fictional and actual world; it would be very difficult to work out whether the time travel world is *farther* from the actual world than the morally permissible torture world. But the difference between the fictional and the actual world is very helpful in explaining the *degree* of imaginative resistance we experience. A story with talking mice is closer to the actual world than a story with talking mice who have also mastered time travel, which is closer still than a world in which those mice are morally right to travel through time in order to torture recreationally.

Also, closeness of a fictional world to the actual world might contribute to an explanation of the *political* sense of imaginative resistance. This kind of imaginative resistance results any time we are asked to imagine something that we recognize might poison our actual beliefs, values and decision-making behavior. And it should hold equally for counterdescriptives as for counterevaluatives; as it does, for example, with imagining there to be important cognitive differences between human races, or imagining that the refugees in our countries have come purposely to cause damage to our societies. In other words, perhaps distance from our own doxastic-evaluative framework (or distance from the set of evaluative and descriptive truths) is a

type of perceived distance from the actual world, which could explain the political sense of imaginative resistance. We refuse to imagine recreational torture as morally permissible because it would not be morally right in this world, whatever the case in other worlds (cf. Gendler, 2000). And we should not spend too much time in those worlds, lest we get any strange ideas.

### **3.3 Surprisingness**

Surprise is a response to a violated expectation, and that expectation can be built upon several things, including previous experience, an inference, or convention. When the expectation violated is based on previous experience, surprise becomes related to unusualness, as often-experienced correlations that suddenly cease to hold will be judged unusual and surprising. When the expectation violated is built upon an inference about what will happen in a fiction, surprise becomes related to perceived nearness of a fictional world to the actual world, since the worlds about which our inferences fail in surprising ways will be judged as distant worlds.

Expectations can also be set up by context, including pragmatic context and genre. Liao et al. (2014) show experimentally that imaginative resistance diminishes or vanishes when dealing with less realistic genres (see also Black et al. in press). Stock (2017: chapter 4) provides examples where children murder innocent parents, innocent people are tortured or killed, and women are degraded, and there is very little or no resistance because of the genre, whether gothic, erotic, horror or black humour. In these cases we do not balk, presumably because the pragmatic context (including genre) can change what we expect to happen, by changing our dispositions to characterize and frame certain objects and events (Camp forthcoming). The pragmatic context tells us “what to do” when confronted with a fiction. When we are confronted with a fiction about talking mice we know what to do: imagine the talking mice roughly as tiny people who like cheese

and are afraid of cats. We can experience resistance if those mice do something talking mice don't normally do in their genre. When we're confronted with a story in which recreational torture is morally permissible, we will experience resistance only if that story isn't embedded in a pragmatic context that makes recreational torture morally permissible. Because genre sets up expectations that are violated, genre could certainly be included in weirdness as an explanatory factor for imaginative resistance.

### **3.4 Other Factors**

Our weirdness measure draws on unusualness, distance from the actual world and surprisingness. But presumably, other factors could figure into weirdness that would make it an even better predictor of imaginative resistance. Another possibility is that a fictional claim is weirder when its assertions contravene commitments that are more deeply held (following Todd 2009). That is, they are "weird" because we cannot work out how to tell a consistent story given the claims made in the fiction plus our existing commitments. We either don't know which claims need to be altered, or how to alter them. Since moral beliefs are plausibly among those things to which we are most strongly committed, this would explain why we often experience imaginative resistance to counterevaluatives: because many of us are strongly committed to our values. But it also explains why we can find counterdescriptives that produce *more* resistance (Black and Barnes 2016), because to deny these we would have to make equally great (or greater) adjustments to our webs of commitment. Of course, all of this holds only when the reader knows that the author intends her to believe that the weird claim does follow from the narrative in the sense that if the scenario were actual, the weird claim would actually follow (Stock 2017: chapter 4). This isn't always the case: sometimes authors simply want their readers to entertain weird claims for fun, puzzlement, or some other effect, in which case the resistance disappears. We can set aside these cases, since our studies

follow the philosophical literature in presenting cases that do seem to ask the reader to assent that  $p$  follows from  $q$  in the scenario (e.g. that shooting Jack and Jill was the right thing to do given the other features of the case). If the centrality of the commitments countered by the scenario does play a role in determining weirdness, imaginative resistance can then be used as a partial measure for how central a commitment is. The suggestion would be: the more difficult it is to see a counterfactual inference through, the more likely it is that the inference requires denying a central commitment.

Another potential source of weirdness is the epistemological authority of the author (Matravers 2003). According to Matravers, we will trust journalists to tell us which events took place, but not what is morally right or hilariously funny. In other words, it isn't the content that drives imaginative resistance, but the perceived epistemological status of the author. While our results show that imaginative resistance is not unique to evaluatives, we have not explored an idea along Matravers's line. It is certainly possible that we might resist imagining when an author's authority breaks down. We merely urge caution in thinking that such authority only breaks down when it comes to values. People can and do accept value authorities, from advice columnists to religious leaders (Stock 2005), and likewise when it comes to descriptive claims, people can and do reject the authority of certain sources, for example, Breitbart News and Donald Trump. In other words, claims deriving from sources whose authority we do not trust might strike us as weird. The reverse also holds: when we encounter a weird headline, we often inquire about its author or source.

Another possible weirdness-factor can be inferred from Weatherson (2004). Weatherson asserts that a fictional claim will be resisted in imagination when it violates the *Virtue* principle, which is premised on the idea that some fictional claims are the kind that must be true in virtue of

lower level facts. The principle states that it must be possible to work out (from what's given in the story) what  $p$  is true in virtue of, when the story is about those lower level facts (Weatherson 2004: 18).<sup>9</sup> When *Virtue* is not satisfied, we experience imaginative resistance. For example, recreational torture is morally wrong in virtue of the harm that torture does. If it were possible to torture people recreationally in a way that caused no harm, perhaps it would be morally permissible. It is possible therefore that violating *Virtue* is a further source of weirdness.

In a somewhat similar vein, Kathleen Stock argues that we do not usually need to make recourse to lower level facts. Rather, when we fail to imagine something, it is because we do not understand it (2005). That is, we fail to find a context that makes sense of the scenario we are asked to imagine. Stock and Weatherson will therefore agree that whether we can find a way to understand the fiction in terms of its context or lower level facts will be relative to the individual. In our view, Stock and Weatherson (and others, like Nanay 2010) identify another possible source of weirdness: we cannot imagine something when we do not understand it (cannot make sense of it), whether because we cannot work out how it could obtain in terms of lower level facts or because of the context, including the presumed intentions of the author. This notion of weirdness as a lack of confidence in the quality of one's epistemological position with respect to drawing certain inferences might be resolved into unusualness and distance from the actual world, but perhaps not.

#### 4. Conclusion

With the work of Black and Barnes (2016, 2017), Black et al. (in press), Liao et al. (2014), and Phelan (2017), we're still just beginning to bring empirical data to bear on philosophical issues of

imaginative resistance. With respect to further avenues of research, we think it would be fruitful to follow up on the factors contributing to weirdness. We should like to know how many there are, what the nature of each is, and how each contributes to imaginative resistance.

Another issue worth pursuing is the discrepancy we found between judgments concerning truth on the one hand, and difficulty and possibility on the other. Both the claim type and the content seem to affect whether people judge something to be true in a fiction, while only content affects whether people judge something fictional to be difficult or (im)possible to imagine. As we mentioned above, this might be congenial to a dual process model of imaginative resistance. The main question remains, however. Why are judgments concerning truth in fiction sensitive to claim type? One could attempt to avoid this question by saying that judgments about truth in fiction are not relevant to imaginative resistance. But we have tried to maintain a broad notion of imaginative resistance throughout, and judgments about truth in fiction are important for several authors in the literature. Still, the truth in fiction puzzle is distinct from the imaginability puzzle (Liao et al. 2014: 342), and perhaps different philosophical accounts will be differently suited to explain these different puzzles. For example, those that focus on the authority of an author might be in a better position to explain our unwillingness to claim that a certain fictional statement is true in the fiction. While those who focus on our inability to process certain claims might be in a better position to explain our experienced difficulty in trying to imagine those claims. In any case, a complete account of imaginative resistance (in the broad sense including truth in fiction judgments) must make reference to content, and not just claim type. And if we want to preserve the asymmetry upon which the original puzzle of imaginative resistance is based, this can only be done, and only partially, in the domain of truth in fiction judgments.

Summing up, in terms of difficulty and possibility, imaginative resistance depends largely

on the weirdness of the content, not on whether the claim is evaluative or descriptive. Truth judgments, by contrast, are sensitive to features of content and claim type. In discussion, we have suggested that the weirdness of a scenario is relative to the individual, and might depend on its (i) unusualness, e.g., on how often a subject has experienced the content of the scenario; (ii) the perceived distance between the fictional and actual world, resulting from, e.g., the distance between the worldviews that are appropriate for each world; and (iii) surprisingness, resulting from violated expectations generated by previous experience, inference, or fictional context. We are left with many interesting questions, including finding other factors that contribute to weirdness. These might include the level of personal commitment to what is violated in the scenario, the epistemological authority of the author presenting the scenario, and how easy it is for the subject to understand the scenario or to work out the lower level facts that justify ascriptions of higher level properties in the scenario. These should be tested experimentally, where possible. Whatever weirdness turns out to be in the final analysis, it is *not* a property exclusive to counterevaluatives, and there is no curious asymmetry of imaginative resistance in the sense of a) difficulty imagining something or b) difficulty judging something to be (im)possible. There does, however, seem to be a version of the puzzle that survives. Namely: why is it that when it comes to deciding whether something is true in a fiction, we need to take into account whether that content is evaluative or descriptive?

## Acknowledgments

The authors would like to thank Edouard Machery for his unwavering support and the Pittsburgh Empirical Philosophy Lab for funding. We would also like to thank Shen-yi Liao, Wesley

Buckwalter, Florian Cova, and the members of the Pittsburgh Empirical Philosophy Lab for extensive and extremely helpful comments on earlier drafts of this paper. Mike Stuart would also like to thank audiences at the University of Graz Evidence and Imagination conference, the New College of the Humanities Mind and Brain conference, the Southern Society of Philosophy and Psychology meeting in Louisville, as well as the University of Pittsburgh and the Social Sciences and Humanities Research Council of Canada for funding.

## Notes

- 1 Names are in alphabetical order, the three authors contributed equally to this work.
- 2 We define a “counterevaluative” as a statement that overturns a held evaluative belief, and a “counterdescriptive” as a statement that overturns a held non-evaluative belief. When we want to be less specific, we will use “counterfactual” to mean a claim that is either counterevaluative or counterdescriptive.
- 3 Others who focus their attention on this assumption in its moral form include Brock (2012), Carruthers (2006), Currie (2002, 217), Dorsch (2016, 50), Driver (2008, 302), Kind (2016a, 8; 2016b, 167), Kung (2016, 234), Mahtani (2012), Matravers (2003), Modrack (2016, 25), Nichols (2006, 3-4), Rosenbaum (2016), Liao (2016), Sauchelli (2016), Stear (2015), Stock (2005), Stokes (2006), and Stueber (2016: 375).
- 4 Appendix 1 in the Supplementary Materials (<https://osf.io/n5cpd/>) contains the complete text of all scenarios and questions from the pilot study.
- 5 See <https://osf.io/n5cpd/> for Appendix 2.
- 6 A potential worry: If weirdness correlated very strongly with the measures of imaginative

resistance, one might be concerned that it captures the exact same phenomenon. A lack of correlation, by contrast, would cast doubt on the hypothesis that features related to the content of the target claim could be meaningful predictors of imaginative resistance. Weirdness correlated positively with judgments of truth in fiction ( $r=0.23$ ,  $p<.001$ ), difficulty ( $r=.52$ ,  $p<.001$ ), and possibility ( $r=.49$ ,  $p<.001$ ). With effect sizes just at the border between medium and large for difficulty and possibility, and a small effect size for truth in fiction, so we do not see any reason for concern.

- 7 See <https://osf.io/n5cpd/> for Appendix 3.
- 8 One might object that a Dual Process Model of Imaginative Resistance cannot be inferred from the fact that claim type has a significant effect on truth but a non-significant effect on difficulty and possibility. After all, the difference between “significant” and “not significant” is not itself statistically significant (Gelman and Stern, 2006). However, the original mixed ANOVA (within-subjects factor: measure – truth vs. difficulty vs. possibility; between-subjects factors: claim type – evaluative vs. descriptive, degree – low vs. medium-low vs. medium-high vs. high) showed that the measure\*claim type interaction  $F(1.80, 1509.02)=27.60$ ,  $p<.001$ ,  $\eta^2=.032$  was significant, which means that claim type had a different effect on the three measures. To see whether it was claim type’s effect on truth that differed significantly from claim type’s effect on difficulty and possibility, we further conducted two 2x2 ANOVAs: (i) 2 (evaluative x descriptive) x 2 (truth vs difficulty), and (ii) 2 (evaluative x descriptive) x 2 (truth vs possibility). For (i) truth vs difficulty, we found the claim type\*measure interaction  $F(1, 842)=34.06$ ,  $p<.001$ ,  $\eta^2=.039$  was significant; claim type affected truth ( $M_{eval}=4.48$ ,  $M_{desc}=2.89$ ) significantly more than it affected difficulty ( $M_{eval}=4.56$ ,  $M_{desc}=3.95$ ). For (ii) truth vs

possibility, we found the claim type\*measure interaction  $F(1, 842)=35.50$ ,  $p<.001$ ,  $\eta^2=.040$  was significant; claim type affected truth ( $M_{eval}=4.48$ ,  $M_{desc}=2.89$ ) significantly more than it affected possibility ( $M_{eval}=3.91$ ,  $M_{desc}=3.25$ ). Hence, the effect of claim type on truth judgments differed significantly from the effect of claim type on difficulty and possibility judgments, which supports the Dual Process Model.

- 9 This extends Walton's account (1994), which also presents a supervenience-based account of imaginative resistance. Walton claims that moral facts supervene on natural facts, and so if we are told to imagine a different moral fact with the same natural facts, we cannot. Weatherston extends this from evaluatives to all facts.

## References

- Balcerak-Jackson, M. (2016), 'On the Epistemic Value of Imagining, Supposing, and Conceiving', in A. Kind and P. Kung (Eds), *Knowledge through Imagination*, pp. 41-60. Oxford: Oxford University Press.
- Baron, R. M. and D. A. Kenny (1986), 'The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations', *Journal of Personality and Social Psychology*, 51: 1173-82.
- Black, J. E. and J. L. Barnes (2016), 'Impossible or Improbable: The Difficulty of Imagining Morally Deviant Worlds', *Imagination, Cognition and Personality*, 36 (1): 27-40.
- Black, J. E. and J. L. Barnes (2017), 'Measuring the unimaginable: Imaginative Resistance to Fiction and Related Constructs', *Personality and Individual Differences*, 111: 71-79.
- Black, J. E., S. C. Capps and J. L. Barnes (in press), 'Fiction, Genre Exposure, and Moral Reality',

*Psychology of Aesthetics, Creativity, and the Arts.*

- Brock, S. (2012), 'The Puzzle of Imaginative Failure', *Philosophical Quarterly*, 62: 443-463.
- Camp, E. (forthcoming), 'Imaginative Frames for Scientific Inquiry: Metaphors, Telling Facts, and Just-So Stories', in P. Godfrey-Smith and A. Levy (Eds), *The Scientific Imagination*. Oxford: Oxford University Press.
- Carruthers, P. (2006), 'Why Pretend?', in S. Nichols (Ed), *Architecture of Imagination*, pp. 89-110. Oxford: Oxford University Press.
- Currie, G. (2002), 'Desire in Imagination', in T.S. Gendler and J. Hawthorne (Eds), *Conceivability and Possibility*, pp. 201-221. Oxford: Oxford University Press.
- Cushman, F. (2008), 'Crime and Punishment: Distinguishing the Roles of Causal and Intentional Analyses in Moral Judgment', *Cognition*, 108 (2): 353-380.
- Cushman, F. (2013), 'Action, Outcome, and Value a Dual-System Framework for Morality', *Personality and Social Psychology Review*, 17 (3): 273-292.
- Dorsch, F. (2016), 'Hume', in A. Kind (Ed), *The Routledge Handbook of Philosophy of Imagination*, pp. 40-55. Oxford: Oxford University Press.
- Driver, J. (2008), 'Imaginative Resistance and Psychological Necessity', *Social Philosophy and Policy*, 25: 301-313.
- Gelman, A. and H. Stern (2006), 'The Difference between 'Significant' and 'Not Significant' Is Not Itself Significant', *The American Statistician*, 60 (4): 328-331
- Gendler, T. S. (2000), 'The Puzzle of Imaginative Resistance', *Journal of Philosophy* 97 (2): 55-81.
- Gendler, T. S. (2006), 'Imaginative Resistance Revisited', in S. Nichols (Ed), *The Architecture of the Imagination*, pp. 149-173. Oxford: Oxford University Press.

- Gendler, T. S. and S. Liao (2016), 'The Problem of Imaginative Resistance', in J. Gibson and N. Carroll (Eds), *The Routledge Companion to Philosophy of Literature*, pp. 405-418. London: Routledge.
- Hume, D. (1757), 'Of the Standard of Taste', in *Essays: Moral, Political and Literary*. Liberty Classics.
- Judd, C. M., D. A. Kenny and G. H. McClelland (2001), 'Estimating and Testing Mediation and Moderation in Within-Subject Designs', *Psychological Methods*, 6: 115–134.
- Judd C.M. and D. A. Kenny (1981), *Estimating the Effects of Social Interventions*. Cambridge University Press: New York.
- Kieran, M. and D. I. Lopes. (2003), *Imagination, Philosophy, and the Arts*. London: Routledge.
- Kind, A. (2016a), 'Introduction: Exploring imagination' in A. Kind (Ed), *The Routledge Handbook of Philosophy of Imagination*, pp. 1-12. Oxford: Oxford University Press.
- Kind, A. (2016b), 'Desire like Imagination', in A. Kind (Ed), *The Routledge Handbook of Philosophy of Imagination*, pp. 163-76. Oxford: Oxford University Press.
- Kind, A. and P. Kung (2016), 'Introduction: The Puzzle of Imaginative Use', in A. Kind and P. Kung (Eds), *Knowledge through Imagination*, pp. 1-3. Oxford: Oxford University Press.
- Kung, P. (2016), 'Thought Experiments in Ethics', in A. Kind and P. Kung (Eds), *Knowledge through Imagination*, pp. 227-246. Oxford: Oxford University Press.
- Liao, S. (2016), 'Imaginative Resistance, Narrative Engagement, Genre', *Res Philosophica*, 93: 461–482.
- Liao, S., N. Strohminger and C. S. Sripada (2014), 'Empirically Investigating Imaginative Resistance', *British Journal of Aesthetics*, 54: 339–355.

- Mahtani, A. (2012), 'Imaginative Resistance without Conflict', *Philosophical Studies*, 158: 415-429.
- MacKinnon, D. P., J. L. Krull and C. M. Lockwood (2000), 'Equivalence of the Mediation, Confounding and Suppression Effect', *Prevention Science*, 1(4): 173-181.
- MacKinnon, D. P., C. M. Lockwood, J. M. Hoffman, S. G. West and V. Sheets (2002), 'A Comparison of Methods to Test Mediation and Other Intervening Variable Effects', *Psychological Methods*, 7: 83-104.
- Marchetti, G. and S. Marchetti (2017), *Facts and Values: The Ethics and Metaphysics of Normativity*. London: Routledge.
- Matravers, D. (2003), 'Fictional Assent and the (So-Called) 'Puzzle of Imaginative Resistance'', in M. Kieran and D. I. Lopes (2003), *Imagination, Philosophy, and the Arts*, pp. 91-106. London: Routledge.
- Modrack, D. K. W. (2016), 'Aristotle on *Phantasia*', in A. Kind (Ed), *The Routledge Handbook of Philosophy of Imagination*, pp. 15-26. Oxford: Oxford University Press.
- Moran, R. (1994), 'The Expression of Feeling in Imagination', *Philosophical Review*, 103 (1): 75-106.
- Mothersill, M. (2006), 'Make-Believe Morality and Fictional Worlds', in J. L. Bermúdez and S. Gardner (Eds), *Arts and Morality*, pp. 74-94. London: Routledge.
- Nanay, B. (2010), 'Imaginative Resistance and Conversational Implicature', *Philosophical Quarterly*, 60: 586-600.
- Nichols, S. (2006), 'Introduction', in S. Nichols (Ed), *Architecture of Imagination*, pp. 1-18. Oxford: Oxford University Press.

- Phelan, M. (2017), *XPhi Replicability Project - Replication of 'LIAO, S., STROHMINGER, N., and SRIPADA, C. S. (2014), Empirically Investigating Imaginative Resistance. British Journal of Aesthetics 54: 339–355'*. Open Science Framework. July 28. [osf.io/7e8hz](https://osf.io/7e8hz).
- Rosenbaum, J. W. (2016), 'Poetic License: Learning Morality from Fiction in light of Imaginative Resistance', *Teorema: International Journal of Philosophy*, 35:165-183.
- Sauchelli, A. (2016), 'Gendler on the Puzzle(s) of Imaginative Resistance', *Acta Analytica*, 31: 1–9.
- Stear, N-H. (2015), 'Imaginative and Fictionality Failure: A Normative Approach', *Philosopher's Imprint*, 15: 1-18.
- Steuber, K. (2016), 'Empathy and the Imagination', in A. Kind (Ed), *The Routledge Handbook of Philosophy of Imagination*, pp. 368-79. Oxford: Oxford University Press.
- Stock, K. (2003), 'The Tower of Goldbach and Other Impossible Tales', in M. Kieran and D. I. Lopes (Eds), *Imagination, Philosophy, and the Arts*. London: Routledge.
- Stock, K. (2005), 'Resisting Imaginative Resistance', *Philosophical Quarterly*, 55: 607-24.
- Stock, K. (2017), *Only Imagine: Fiction, Interpretation and Imagination*. Oxford: Oxford University Press.
- Stokes, D. (2006), 'The Evaluative Character of Imaginative Resistance', *British Journal of Aesthetics*, pp. 46: 387-405.
- Tanner, M. (1994), 'Morals in Fiction and Fictional Morality (II)', *Supplement to the Proceedings of the Aristotelian Society*, pp. 68: 51–66.
- Todd, C. S. (2009), 'Imaginability, Morality, and Fictional Truth: Dissolving the Puzzle of 'Imaginative Resistance'', *Philosophical Studies*, 143: 187–211.

- Van Leeuwen, N. (2016), "Imagination and Action" in A. Kind (ed) *The Routledge Handbook of Philosophy of Imagination*, pp. 286-99. Oxford: Oxford University Press.
- Walton, K. (1990), *Mimesis as Make Believe*. Cambridge: Harvard University Press.
- Walton, K. (1994), 'Morals in Fiction and Fictional Morality', *Proceedings of the Aristotelian Society*, 68: 27-50. Reprinted in Walton, K. (2008), *Marvelous Images: On Values and the Arts*, pp. 27-45. Oxford: Oxford University Press.
- Weatherson, B. (2004), 'Morality, Fiction, and Possibility', *Philosophers' Imprint*, 4: 1-27.
- Yablo, S. (2002), 'Coulda, Woulda, Shoulda', in T. S. Gendler and J. Hawthorne (Eds), *Conceivability and Possibility*, pp. 441-492. Oxford: Oxford University Press.