Agent framing moderates concerns about moral contagion

Stephen J. Flusberg (stephen.flusberg@purchase.edu)
Carly LaPlace (carly.laplace@purchase.edu)
SUNY Purchase College, Department of Psychology
735 Anderson Hill Road, Purchase, NY 10577, USA

Abstract

Concerns about moral contamination shape people's attitudes towards the objects they encounter in daily life. For example, money seems less desirable when it comes from a robbery (Tasimi & Gelman, 2017). Drawing on the *theory of dyadic morality*, we hypothesized that increasing an individual's sense of agency would reduce the salience of moral contagion and make people feel less vulnerable to moral contamination. Across two experiments, we adapted the study design of Tasimi and Gelman (2017), asking participants how much they desired a $1 (Experiment 1) or $100 (Experiment 2) bill associated with different negative moral histories. We modified the stimulus language so that participants were framed as either the moral *agent* or *patient* for all scenarios. As predicted, participants in the agent language condition expressed nearly the same level of desire regardless of the bill's moral history, highlighting the role that feelings of agency play in moral decision-making.

Keywords: Moral contagion**,** economic judgments, dyadic morality, linguistic framing

Introduction

People all over the world spend good money to read books and watch films about the evilest men who ever lived, from Hitler to Dahmer to Voldemort. But what would you pay to wear their used trousers? Research on *moral contagion*—the intuitive, magical belief that an individual's moral qualities contaminate the people and objects they come in contact with—suggests *not very much* (e.g., Eskine, Novreske, & Richards, 2013; Nemeroff & Rozin, 1994; Newman & Bloom, 2014; Rozin et al., 1989). For example, one study found that people would bid more for a sweater worn by a disliked, morally transgressive celebrity if it were sterilized first, while the opposite was true for clothing belonging to a more popular celebrity (Newman & Bloom, 2014).

A number of recent studies have focused on how contagion beliefs influence how people value and spend *money* (e.g., Di Muro & Noseworthy, 2013; Galoni & Noseworthy, 2014; Kardos & Castano, 2012; Tasimi & Gelman, 2017). This is especially notable since money is typically thought to be a fungible resource: rationally speaking, a dollar should be seen as equally valuable no matter who has touched it or where it comes from. And yet, several experiments have shown that moral contagion shapes attitudes towards money. Kardos and Castano (2012), for example, asked participants to imagine they had won the lottery and indicate how much of the winnings they would spend on a personal vacation. Some participants were told to imagine they found the winning lottery ticket on the ground, while others were told to imagine they secretly picked up the winning ticket after the winner

accidentally dropped it. Those in the latter condition—where the money was acquired in a morally dubious fashion—experienced more (hypothetical) guilt and indicated they would spend less of their winnings on a vacation.

A recent paper by Tasimi and Gelman (2017; hereafter T&G) elegantly demonstrates that the moral history of a particular piece of currency significantly impacts how people value that money. Across eight experiments, participants were presented with a series of scenarios that asked them to rate how much they desired a piece of currency that was associated with different moral histories and physical properties. The scenarios included:

a. ***neutral-giver/neutral-money*** (e.g. "Henry has a dollar in his desk. Henry says you can have the dollar, if you want")
b. ***bad-giver/neutral-money*** (e.g. "Paul stole a dollar from another person. The dollar that he stole is in his pocket. Paul has another dollar that he did not steal, in his desk. Paul says you can have the dollar in his desk, if you want")
c. ***neutral-giver/bad-money*** (e.g. "Frank found a stolen dollar in his desk. Frank says you can have the dollar, if you want")
d. ***bad-giver/bad-money*** (e.g. "Brian stole a dollar from another person. Brian says you can have the dollar, if you want")
e. ***neutral-giver/dirty-money*** (e.g. "Marvin sneezed and used a dollar to wipe his nose. Marvin says you can have the dollar, if you want")

T&G consistently found that participants were less likely to want the money when it had a negative moral history (the dollar bill itself was stolen) or negative moral association (the bill was offered up by a thief). The effects of a negative moral history were slightly stronger than the effects of a negative moral association (i.e., people wanted the money less in scenario *c* than in scenario *b* above), suggesting people viewed the currency itself as especially contaminated when it originated from a crime. That participants wanted the money even less in scenario *d*, when it was a thief offering up his stolen dollar, suggests that negative moral histories and associations elicit independent, additive concerns about moral contagion.

This pattern of results replicated whether the currency involved was a $1 bill, four quarters, or a $100 bill. Interestingly, in the latter case, when the money at stake was relatively valuable, participants didn't seem to mind if it was covered in snot, though in the case of the $1 bill they actually

had the lowest levels of desire for the physically dirty cash. Finally, follow-up experiments attempted to rule out the possibility that participants responded the way they did because of concerns about getting in trouble for dealing in stolen goods, and also revealed that participants were more likely to spend the money they were offered on charity when it was morally contaminated.

One noteworthy feature of this experiment is that the stimulus language always framed the participant as a passive observer (or *patient*), with the protagonist described in the vignette serving the role of the active *agent*. For example, in scenario *c* above, the character "Frank" both finds the stolen dollar and then offers it up to the participant. Is this particular framing important for eliciting effects of moral contagion, or does the moral history of money affect people's desire regardless of how it is talked about?

One reason to think that this *"patient"* framing might interact with concerns about moral contagion comes from the broader literature on moral reasoning; specifically, from the work of Kurt Gray and colleagues on the *theory of dyadic morality* (TDM. e.g., Gray & Wegner, 2009; Schein, Goranson, & Gray, 2015; Schein & Gray, 2018). According to TDM, moral judgments are made by comparing the situation at hand to an idealized cognitive template or schema of an immoral action, which involves two individuals (hence *dyadic*): an intentional *agent* (the perpetrator) inflicting harm on a passive *patient* (the victim of harm). This is clearly illustrated in prototypical immoral acts like murder and assault, but Gray and colleagues have argued this same harm-based template is deployed even for moral judgements that appear not to involve harm (e.g., for those who judge masturbation to be immoral. See Schein & Gray, 2018, for a recent review of this theory and the empirical evidence that supports it).

Importantly, studies suggest that once someone has been categorized as a moral agent or patient, they become *typecast* in that role, viewed through the lens of agent or patient in other contexts (Gray & Wegner, 2009). In other words, once a person is judged to be a moral patient, it is less likely that they will be perceived as being capable of *performing* good or evil actions (i.e., as becoming a moral agent), and vice versa. What's more, individuals typecast as moral patients are viewed as more *vulnerable* to the moral actions of others, and to harm in general, while those typecast as moral agents are viewed as less vulnerable and more resistant to harm.

One consequence of this theory is that feelings of agency might alleviate concerns about moral contamination. Consider that in every scenario presented to participants in T&G, the protagonist in the vignette was framed as the agent while the participant was framed as the patient. TDM suggests that by taking on the role of the patient, participants would feel especially vulnerable to the negative moral actions of others, which would include the moral contamination associated with handling stolen money. On the other hand, because TDM argues that moral agents are conceived of as the locus of good and bad acts who are relatively resistant to the harmful actions of other, making the participant feel like more of an agent should *reduce* the feeling that the negative actions of others might somehow impact them (via moral contagion). In other words, if you are the one who is actively *taking*—rather than *accepting*—(metaphorically) dirty money from a bad guy, then this might not seem so bad (just ask Robin Hood).

We investigated this possibility in two experiments. Participants read a series of five scenarios adapted from T&G and indicated how much they desired a $1 (Experiment 1) or $100 (Experiment 2) bill associated with different moral histories. Critically, we modified the stimulus language so that the participant was framed as either the moral *agent* or *patient* for all scenarios. As predicted, the impact of the moral history of money was attenuated for those in the agent language condition, suggesting that researchers interested in moral contagion should carefully consider the role that feelings of agency and other moderating factors might play in this domain.

Experiment 1

Methods

Participants 204 participants were recruited through Amazon's Mechanical Turk crowdsourcing platform. (Buhrmester, Kwang, & Gosling, 2011), using the TurkPrime platform (Litman, Robinson, & Abberbock, 2017). Following T&G, we aimed for a sample size of approximately 100 participants in each of the two language framing conditions. All participants were located in the United States, had a good performance record on MTurk (at least a 90% rating), and were at least 18 years old. See Table 1 for participant demographic data.

Table 1: **Demographic data for Experiments 1 & 2**

	Experiment 1	Experiment 2
N	204	203
% female	43.1%	41.4%
Mean age (SD)	36.7 (11.47)	37.1 (11.3)
% white	76.5%	77.3%
% democrat/republican	43.6% / 22.1%	43.8% / 24.6%
Mean pol. ideology (SD)	39.3 (28.2)	41.9 (30.6)

Materials & Procedure The experiment was created using Qualtrics online survey software. Participants were randomly assigned to either a *patient language* condition, which was directly adapted from T&G Experiment 1, or an *agent language* condition, which consisted of slightly modified scenario language designed to frame the participant as having relatively more agency in the task.

Every participant responded to five distinct scenarios that required them to rate how much they desired a $1 bill associated with a particular moral history, presented in a randomized order, using a 1-7 rating scale (1 = not at all, 7 = very much). The language used in each scenario for each condition is presented in Table 2. Note that the key differences in the *agent language* condition are (1) that the

dollar is always referred to as "up for grabs," rather than the character in the scenario saying the participant can "have the dollar, if [they] want," and (2) the question specifically asked participants how much they wanted to "take" the dollar, rather than simply asking how much they "want" the dollar.

Finally, participants completed the *Social Desirability Scale-17* (SDS-17; Stöber, 2001) and a series of basic demographic questions. The SDS-17 measures the degree to which people tend to respond to surveys by providing socially desirable answers. It consists of 17 true/false statements (e.g., *"In traffic I am always polite and considerate of others"*) that participants can respond to (a) in a socially desirable way that is unlikely to be true for most people (i.e., responding "true" to the example statement above, coded as a 1), or (b) in a less socially desirable way that is more likely to be accurate (i.e., responding "false" to the example statement above, coded as a 0). In the current sample, the mean social desirability score was 7.74/17 (*SD* = 2.22). We included this measure to help control for the possibility that any observed effects were being driven by a desire to give socially desirable ratings rather than concerns over moral contagion.

Table 2. Stimulus language used in Experiment 1.

Scenario	Patient Language	Agent Language
neutral-giver/ neutral money	Henry has a dollar in his desk. Henry says you can have the dollar, if you want. How much do you want the dollar?	Henry has a dollar in his desk that is now up for grabs. How much do you want to take the dollar?
bad-giver/ neutral-money	Paul stole a dollar from another person. The dollar that he stole is in his pocket. Paul has another dollar that he did not steal, in his desk. Paul says you can have the dollar in his desk, if you want. How much do you want the dollar?	Paul has a dollar he stole from another person. The dollar that he stole is in his pocket. Paul has another dollar that he did not steal, in his desk, which is now up for grabs. How much do you want to take the dollar in his desk?
neutral-giver/ bad-money	Frank found a stolen dollar in his desk. Frank says you can have the dollar, if you want. How much do you want the dollar?	Frank found a stolen dollar in his desk that is now up for grabs. How much do you want to take the dollar?
bad-giver/ bad money	Brian stole a dollar from another person. Brian says you can have the dollar, if you want. How much do you want the dollar?	Brian stole a dollar that is now up for grabs. How much do you want to take the dollar?
neutral-giver/ dirty-money	Marvin sneezed and used a dollar to wipe his nose. Marvin says you can have the dollar, if you want. How much do you want the dollar?	Marvin sneezed on a dollar that is now up for grabs. How much do you want to take the dollar?

Results

Data were analyzed using jamovi open source statistics software (jamovi project, 2018). We ran a repeated-measures ANOVA with *scenario* as the within-subjects factor and *language condition* as a between-subjects factor, with desire for money as the dependent variable. We included social desirability as a covariate in the analysis, though excluding this factor did not meaningfully affect our findings.

There was no main effect of social desirability ($F < 0.5$, p = .82), and social desirability did not interact with *scenario* ($F < .05$, $p = 0.86$). Neither was there a main effect of *language condition*, $F(1, 201) = 0.03$, $p = 0.86$. There was, however, a main effect of *scenario*, $F(4, 804) = 3.09$, $p = 0.015$, $\eta_p^2 = .015$[1], replicating the overall pattern of results observed by T&G. Crucially, this effect was qualified by a significant *scenario* by *language condition* interaction, $F(4, 804) = 13.71$, $p < 0.001$, $\eta_p^2 = .064$. We therefore conducted a series of Bonferroni-corrected post-hoc t-tests to compare mean desire for money between each scenario, both within and between language conditions.
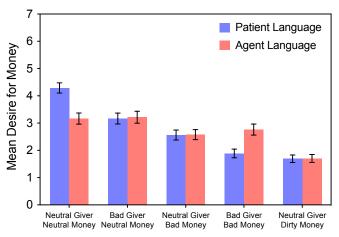

Figure 1: Mean monetary desire for every scenario and both language conditions for Experiment 1. Error bars represent SEMs.

The results of the *patient language* condition, which closely mirrored the methods of T&G Experiment 1, replicated the basic findings from the original study. First, participants desired the money significantly more in the *neutral-giver/neutral-money scenario* than in all other scenarios (all *p*'s < 0.001), suggesting that both the moral history and disgusting physical properties of money affect

[1] Though Mauchly's test indicated that the assumption of sphericity was violated in this analysis, as well as the parallel analysis in Study 2, the *F* and *p*-values remain nearly identical under both Greenhouse-Geisser and Huynh-Feldt corrections in all cases. For simplicity, we present the uncorrected values.

desirability. Second, participants desired the money more in the *bad-giver/neutral-money* scenario than in the *neutral-giver/bad-money* scenario, suggesting that the moral history of an actual dollar bill matters more than its moral associations (though this effect was only marginal, $p = 0.053$). Finally, participants wanted the money more in the *bad-giver/neutral-money* scenario than in the *bad-giver/bad-money* scenario ($p < 0.001$), supporting the idea that a particular piece of currency can be selectively *contaminated* by its moral history.

The results were substantially different in the *agent language* condition. Here, there were no significant differences in mean desire for the money between the *neutral-giver/neutral-money scenario* and any of the three scenarios with a negative moral history or association (all p's > 0.13). The only lingering hint of moral contagion was that participants did desire the money marginally more in the *bad-giver/neutral-money* scenario than in the *neutral-giver/bad-money* scenario ($p = 0.059$). Finally, participants showed the lowest levels of desire for the sneezed-on dollar compared to all other scenarios (all p's < 0.001).

The between-subjects post-hoc tests revealed that participants in the *agent language* condition desired the money *less* in the *neutral-giver/neutral-money* scenario ($p < 0.001$) and *more* in the *bad-giver/bad-money* scenario ($p = 0.038$) than those in the *patient* language condition. No other between-group differences were statistically significant.

Discussion

In this experiment, participants responded to five different scenarios that asked them about their desire for a $1 bill that was associated with different moral histories. In the *patient language* condition—which closely matched Experiment 1 in T&G and placed the protagonist described in each scenario into the agentive role—we replicated the findings from the original T&G study. Specifically, participants expressed less desire for the dollar when it was offered up by a thief, even less desire when the dollar itself was stolen, even less desire when it was stolen *and* offered by a thief, and the least amount of desire when it was sneezed on.

In support of our original hypothesis, however, these effects of moral contagion were greatly attenuated in the *agent language* condition, where the stimulus language was modified to place the participants themselves into a more agentive role. In this condition, participants expressed roughly the same amount of desire for the money when it was not stolen or associated with a thief as when it *was* stolen and/or associated with a thief (though they still wanted little to do with the disgusting, snot-covered bill).

This marked difference between the two conditions is most apparent in the diverging responses for two scenarios: (1) the *neutral-giver/neutral money* scenario, where participants in the patient language condition expressed more desire for the dollar than those in the agent language condition, and (2) the *bad-giver/bad-money* scenario, where participants in the agent language condition expressed more desire for the dollar. This pattern of responding is well captured by appealing to TDM. Recall that, according to this theory, moral agents are viewed as having the power to commit moral and immoral acts and are relatively resistant to the actions of others, while moral patients are viewed as being more passively vulnerable to harms. In the *neutral-giver/neutral money* scenario, participants in the agent language condition may have felt that "taking" money from a neutral protagonist would be something of a moral violation, even when that money was described as "up for grabs"; taking money from strangers, even when it is not an act of robbery, is generally frowned upon in the United States, where individual responsibility is a prized virtue. That participants in this condition responded similarly to *all* of the scenarios (except the dirty money scenario, which posed a potential health risk), expressing a very moderate desire for the dollar, suggests that they viewed each scenario through this same lens: it feels a bit wrong to *take* a dollar from another person, no matter what they did or where the dollar came from.

Those in the patient language condition, on the other hand, were made to feel vulnerable to the negative actions of others, but relatively less capable of committing moral or immoral acts themselves. In addition, the patient language stimuli specifically noted that the protagonist says the participant can have the dollar if they want. Taken together, this would make the desire for the dollar more acceptable in the *neutral-giver/neutral-money* scenario, since the participants-as-patients do not view themselves as the direct locus of moral actions. In the *bad-giver/bad-money* scenario, on the other hand, those in the patient language condition (and in T&G) seem to feel *especially* vulnerable to moral contagion since the dollar in question was both stolen *and* offered up by a thief. As we hypothesized, however, being made to feel more agentive in this situation led those in the agent language condition to feel relatively invulnerable to the moral contamination associated with the negative actions of others, and so they responded similarly to this scenario as they did to the others.

One limitation of Experiment 1 was that it only considered the desire people expressed for a very small amount of money. Following T&G, therefore, we replicated this study in Experiment 2 but replaced the $1 bill in each scenario with a $100 bill. T&G found that this shift to a higher valued currency did not ameliorate participant fears of moral contagion: the overall pattern of results they observed was largely the same in this case (though participants were more willing to accept a $100 bill that was sneezed on than a stolen bill or one offered by a thief). Thus, a reticence to accept morally contaminated money is not restricted to small amounts of cash. We expected, therefore, that participants in the agent language condition would be similarly un-affected when confronted with a larger amount of money. That is, we hypothesized that the results of Experiment 2 would largely mirror the results of Experiment 1.

Experiment 2

Methods

Participants 203 participants were recruited online through MTurk using TurkPrime, using the same exclusion criteria as in Experiment 1. See Table 1 for demographic data.

Materials & Procedure Experiment 2 was nearly identical to Experiment 1, with one critical exception: in each scenario, it was a $100 bill that was now under consideration instead of a $1 bill.

Results

We ran the same analysis on the data as we did in Experiment 1. The overall pattern of results was quite similar, though average desire for the money was higher across the board (see Figure 2). As in Experiment 1, we observed a main effect of *scenario*, $F(4, 800) = 3.77$, $p = 0.005$, $\eta_p^2 = .019$, and a significant interaction between *scenario* and *language* condition, $F(4, 800) = 16.23$, $p < 0.001$, $\eta_p^2 = .074$. We therefore conducted a series of Bonferroni-corrected post-hoc t-tests to compare mean desire for money between each *scenario*, both within and between *language* conditions.

Once again, the results of the *patient language* condition closely mirrored the findings of T&G, with one exception. Participants desired the money significantly more in the *neutral-giver/neutral-money scenario* than in all other scenarios (all p's < 0.001), and participants desired the money more in the *bad-giver/neutral-money* scenario than in the *neutral-giver/bad-money* scenario ($p < 0.001$). However, in Experiment 2 there was no difference in desire for the money between the *bad-giver/neutral-money* and *bad-giver/bad-money* scenarios ($p > 0.99$). Finally, like T&G, in this $100 version of the task participants desired the money *more* when it was physically dirty than when it had a negative moral history (i.e. compared to the *neutral-giver/bad-money* and *bad-giver/bad-money scenarios*; p's < 0.001), though there was no difference in desire for the money between the dirty money and *bad-giver/neutral-money* scenarios ($p > 0.99$).

The pattern of results in the *agent language* condition were again quite different and largely replicated what we found in Experiment 1. Once again, there were no significant differences in mean desire for the money between the *neutral-giver/neutral-money scenario* and any of the three scenarios with a negative moral history or association (all p's > 0.18). However, participants did desire the money marginally more in the *bad-giver/neutral-money* scenario than in the *neutral-giver/bad-money* scenario ($p = 0.054$). There was no difference in desire for the money between the *bad-giver/neutral-money* and *bad-giver/bad-money* scenarios ($p > 0.99$). Finally, participants desired the money less in the *dirty money* scenario compared to the *neutral-giver/neutral money* ($p = 0.042$) and *bad-giver/neutral money* ($p = 0.011$) scenarios, but not compared to the *neutral-giver/bad-money* and *bad-giver/bad-money* scenarios (p's > 0.99). less than the

The between-subjects post-hoc tests revealed that participants in the *agent language* condition desired the money *less* in the *neutral-giver/neutral-money* scenario than

those in the *patient* language condition ($p < 0.001$), but no other differences were statistically significant.

Interestingly, we also observed a significant main effect of social desirability in the main analysis, $F(1, 200) = 5.89$, $p = 0.016$, $\eta_p^2 = 0.026$, though this variable did not interact with *scenario*. Collapsing across all scenarios and language conditions, we observed a negative correlation between social desirability and desire for money, $r(201) = -0.159$, $p = 0.023$. This suggests that when more money is at stake, those who are motivated to give socially desirable responses are less likely overall to express a desire to acquire those funds, perhaps because of social norms against accepting money from others in exchange for nothing.
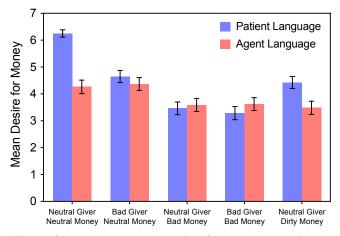

Figure 2: Mean monetary desire for every scenario and both language conditions for Experiment 2. Error bars represent SEMs.

Discussion

As predicted, the results of Experiment 2 largely mirrored the overall pattern of findings from Experiment 1, with a couple notable differences. Not surprisingly, average desire ratings were higher across the board in this version of the experiment, as one hundred times as much money was on the table in Experiment 2. Like T&G, we found that this was especially evident for the *dirty-money* scenario, suggesting the health risk posed by handling another person's effluvia appears to be worth $100.

Other than that, we observed most of same effects as in Experiment 1. Participants in the *patient language* condition showed evidence that fears of moral contagion lowered their desire for the money when it was offered by a thief, and even more so when the cash itself was stolen, though there was little evidence for an additive effect of these negative moral variables in this version of the experiment. Those in the *agent language* condition, on the other hand, were relatively immune to the possibility of moral contagion, responding with nearly the same level of desire for the money in all of the scenarios.

Interestingly, we did find a main effect of social desirability in this version of the experiment, such that participants with higher social desirability scores tended to express less desire overall for the money. This is consistent

with the view that Americans tend to think of taking money from others as a violation of social norms, which further supports the conclusion that the agentive framing highlights this value in participants.

General Discussion

Research finds that fears of moral contamination can dramatically shape our attitudes towards the people and objects we come in contact with in our daily lives (Eskine, Novreske, & Richards, 2013; Kardos & Castano, 2012; Newman & Bloom, 2014; Rozin et al., 1989; Tasimi & Gelman, 2017). T&G, for example, found that people express less of a desire for free cash when it comes from a robbery or is offered up by a thief. At the outset, however, we suggested that people may feel more or less vulnerable to moral contagion depending on whether they are currently thinking of themselves as more of a moral agent or patient (Gray & Wegner, 2009; Schein & Gray, 2018).

To test this idea, we adapted the study design of T&G in two experiments. The linguistic stimuli used in the original study always highlighted the protagonist in each scenario as more of the agent and the participant as more of the patient in the context of the experimental task. We hypothesized that by framing participants as the patient, the researchers may have made participants feel especially vulnerable to the negative moral actions of others (i.e., to moral contagion; Gray & Wegner, 2009; Schein & Gray, 2018). Therefore, we included a second condition in each experiment that modified the stimulus language to place participants in more of an agentive role. We predicted that the effects of moral contagion would be attenuated for participants in this *agent language* condition.

The results of both of our experiments, which were largely consistent with one another, supported our hypothesis. Participants in the *patient language* condition, which was adapted directly from T&G, showed the same effects of moral contagion as in the original set of studies: they expressed less desire for the money when it was offered by a thief, even less desire for the money when it was stolen, and even less desire for the money when it was stolen *and* offered by a thief (at least in Experiment 1).

Participants in the *agent language* condition, on the other hand, expressed nearly the same level of desire for the money in each of these scenarios. In other words, they did not seem to feel vulnerable to moral contagion in the same way that participants in the *patient language* condition did. Interestingly, the average level of desire expressed by those in the *agent language* condition was relatively moderate overall, at around the level that participants in the *patient language* condition expressed for stolen cash. This could be due to the fact that, as individuals taking on the role of moral agents in this condition, these participants felt that *taking* money from a stranger would itself be a minor moral violation, which may have dampened down their overall levels of desire for the cash. However, future work is needed to directly test this possibility, and to fully account for these findings.

Finally, the astute observer will note that participants in the agent language condition did show some hints of being affected by the moral history of the money, especially when the cash itself was stolen. This is reflected in the marginal difference in desire levels between the *bad-giver/neutral-money* and *neutral-giver/bad-money* scenarios in both experiments. This may reflect the fact that the stimulus language in this condition still presented the protagonist in each scenario in relatively agentive terms, using active verbs to describe their activities (e.g., "Frank *found* a stolen dollar…"). In creating the stimuli, we found that we could not easily eliminate this level of agency associated with the protagonists without making things sound awkward and unnatural. Therefore, our efforts to make the participant view themselves as more of an agent involved *re*-framing the narrative right after we described the protagonist, indicating that the cash was "up for grabs" and asking participants how much they wanted to "take" the money. Both pilot testing and our overall pattern of results suggest that we succeeded in making participants feel less like passive patients and more like active moral agents in this condition, but it seems possible that participants still felt somewhat vulnerable to moral contamination from the stolen money. Future work is needed to full tease out these possibilities.

In sum, our findings indicate that people who feel more agentive in the context of a morally fraught situation may feel less vulnerable to the effects of moral contamination. In addition to providing some additional support for TDM, this work also suggests that researchers interested in moral contagion should carefully consider the effects of context and feelings of agency. Future research in our lab will continue to investigate these and other moderating factors in the domain of moral contagion.

References

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk a new source of inexpensive, yet high-quality, data? *Perspectives on psychological science*, 6(1), 3-5.

Di Muro, F., & Noseworthy, T. J. (2013). Money isn't everything, but it helps if it doesn't look used: How the physical appearance of money influences spending. *Journal of Consumer Research, 39*(6), 1330-1342.

Eskine, K. J., Novreske, A., & Richards, M. (2013). Moral contagion effects in everyday interpersonal encounters. *Journal of Experimental Social Psychology, 49(5),* 947–950.

Galoni, C., & Noseworthy, T. J. (2015). Does dirty money influence product valuations? *Journal of Consumer Psychology, 25*(2), 304-310.

Gray, K., & Wegner, D. M. (2009). Moral typecasting: Divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology, 96*(3), 505-520.

jamovi project (2018). jamovi (Version 0.9) [Computer Software]. Retrieved from https://www.jamovi.org

Kardos, P., & Castano, E. (2012). Money doesn't stink. or does it? the effect of immorally acquiring money on its spending. *Current Psychology, 31*(4), 381-385.

Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior research methods*, *49*(2), 433-442.

Nemeroff, C., & Rozin, P. (1994). The contagion concept in adult thinking in the United States: Transmission of germs and of interpersonal influence. *Ethos*, *22*(2), 158-186

Newman, G. E., & Bloom, P. (2014). Physical contact influences how much people pay at celebrity auctions. *Proceedings of the National Academy of Sciences*, *111*(10), 3705-3708

Rozin, P., Nemeroff, C., Wane, M., & Sherrod, A. (1989). Operation of the sympathetic magical law of contagion in interpersonal attitudes among Americans. *Bulletin of the Psychonomic Society*, *27*(4), 367-370

Schein, C., Goranson, A., & Gray, K. (2015). The uncensored truth about morality. *Psychologist*, *28*(12), 982-985

Schein, C., & Gray, K. (2018). The theory of dyadic morality: Reinventing moral judgment by redefining harm. *Personality and Social Psychology Review*, *22*(1), 32-70

Stöber, J. (2001). The social desirability scale-17 (SDS-17): Convergent validity, discriminant validity, and relationship with age. *European Journal of Psychological Assessment*, 17(3), 222-232.

Tasimi, A., & Gelman, S. A. (2017). Dirty money: The role of moral history in economic judgments. *Cognitive Science, 41*, 523-544.