



Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

Case Report

From impure to harmful: Asymmetric expectations about immoral agents[☆]Alek Chakroff^{a,*}, Pascale Sophie Russell^b, Jared Piazza^c, Liane Young^a^a Boston College, Chestnut Hill, MA, USA^b University of Surrey, Guildford, Surrey, UK^c Lancaster University, Lancaster, UK

ARTICLE INFO

Article history:

Received 16 March 2016

Revised 10 August 2016

Accepted 10 August 2016

Available online xxxx

Keywords:

Morality

Harm

Purity

Moral judgment

Behavior expectation

ABSTRACT

How does information about agents' past violations influence people's expectations about their future actions? We examined this question, with a focus on the contrast between past harmful and past impure actions. Participants' judgments reflected two independent influences: action consistency and expectation asymmetry. An expectation asymmetry was observed across seven studies, including two pilot studies and two supplemental studies: impure agents were judged as more likely to be harmful than harmful agents were judged likely to be impure. This expectation asymmetry is not due to an expectation that impure agents will be globally deviant, i.e., likely to commit all kinds of violations (Study 1), nor is it due to differences in the perceived wrongness or weirdness of harmful versus impure acts (Study 2). Study 3 demonstrated that this asymmetry is not attributable to the perceived harmfulness of impure actions; only impure agents, and not harmful agents, were expected to be more harmful than they were previously. These findings highlight an important asymmetry in the way people make predictions about future wrongdoing: immoral agents are expected to behave consistently, and are also expected to be harmful, regardless of their prior violation.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

In advocating the death penalty for homosexual acts, the psychologist Paul Cameron noted that “those who act on their homosexual desires or interests usually end up being parasites on society, and parasites that are very dangerous for society, not only because they take far more than they contribute to society, but they particularly injure children” (Gettys, 2014). This statement captures a tragically widespread sentiment that gay men and women are likely to act in harmful ways. Similar charges have been leveled at people who engage in other forms of non-normative sexual practices, such as zoophilia, despite evidence to the contrary (see Beetz, 2002, 2004). What is the psychology behind inferences like this? The present research investigates people's expectations about agents' future transgressions, based on information about agents' past behaviors.

1.1. The expected behaviors of others

A key function of social cognition is behavior prediction (Dennett, 1989; Saxe, 2012). A reasonable predictor of how people may behave

in the future is how they behaved in the past. People are expected to behave *consistently* over time (Kelley, 1967; Quoidbach, Gilbert, & Wilson, 2013). Thus, someone who has unjustly harmed others in the past may be predicted to be similarly harmful in other contexts and also across time. These inferences are not necessarily irrational. For example, the odds that property offenders will be rearrested for the same kind of crime are 2.7 times the odds that non-property offenders will be arrested for a property offense (Langan & Levin, 2002). Such statistics indicate that at least some offenders are likely to commit similar kinds of crimes over time, though we note there is a general tendency to overestimate the likelihood of reoffending within particular domains, and this is true for both violent and sexual crimes (Cunliffe & Shepherd, 2007).

Do people's inferences follow this logic across all kinds of past actions? Prior work has distinguished between two broad types of moral violations: *harmful* (e.g., theft, murder, abuse) versus *impure* (e.g., incest, bestiality, eating dog meat) (Graham, Haidt, & Nosek, 2009; Inbar, Pizarro, Knobe, & Bloom, 2009; Rozin, Lowery, Imada, & Haidt, 1999; Russell & Giner-Sorolla, 2013; Young & Saxe, 2011). When evaluating harmful acts, people weigh circumstantial evidence before passing judgment, including whether the act was intentional and whether there was a good reason for the agent to perform the act (Cushman, 2008; Sousa, Holbrook, & Piazza, 2009). By contrast, judgments of impure acts tend to be less affected by contextual factors, including the

[☆] The authors contributed equally to the writing and the research reported in this paper.

* Corresponding author at: Boston College Dept. of Psychology, 140 Commonwealth Avenue, Chestnut Hill, MA 02467, USA.

E-mail address: alekchakroff@gmail.com (A. Chakroff).

intentions of the agent (Chakroff, Dungan, & Young, 2013; Chakroff et al., 2015; Russell & Giner-Sorolla, 2011a; Young & Saxe, 2011) and the reasons and external circumstances leading to the act (Piazza, Russell, & Sousa, 2013; Russell & Giner-Sorolla, 2011b, 2011c). While harmful acts are often thought to be motivated by a combination of internal and external forces, impure acts are more likely to be attributed to causes internal to the agent rather than situational forces (Chakroff & Young, 2015; Russell & Piazza, 2015).

1.2. Pilot research: the expected behaviors of harmful and impure agents

Given this prior work, we predicted that impure behaviors would be more informative about an agent and the agent's likely future behavior, compared to harmful acts. Simply learning that an agent has acted impurely in the past might raise expectations that the agent will act impurely in the future, e.g., someone who committed incest yesterday may be likely to do so tomorrow. By contrast, learning that an agent was harmful in the past may be less informative: the harmful act could have been due to external circumstances and less indicative of the agent's stable, underlying dispositions, e.g., someone who punched a man yesterday may be unlikely to do so tomorrow. Two studies were independently conducted in two different labs (PSR + JP; AC + LY) investigating the expected behaviors of agents who committed harmful versus impure actions. As reported in Pilot Studies 1 and 2 (see Supplemental materials), both impure and harmful agents were largely expected to engage in future actions of a similar kind (though the precise means varied across studies). However, in addition to this consistency effect, we also observed an unanticipated finding, an *expectation asymmetry*: impure agents were expected to be harmful more than harmful agents were expected to be impure. Indeed, in at least one of the studies (Pilot Study 1), the impure agent was expected to be harmful just as much as they were expected to be impure. Notably, this expectation asymmetry (more harmful behavior anticipated from impure agents than impure behavior anticipated from harmful agents) was obtained when contrasting harmful acts with two different kinds of impure acts: sexual deviance (e.g., incest; Pilot Study 1) and actions involving contact with taboo substances (e.g., touching feces; Pilot Study 2). Furthermore, this expectation asymmetry occurred despite what might be predicted on the basis of a reverse or negative halo effect (Nisbett & Wilson, 1977; Thorndike, 1920)—namely, that both agents should be expected to behave badly in all respects.

1.3. The present research

Why would impure agents be expected to behave in harmful ways? One possibility is that, once someone has committed an impure act, observers regard the impure individual as globally corrupt and capable of anything. By contrast, harmful agents may not engender the same kind of attributions because harmful acts are generally more common and easier to justify in terms of external causes or reactions to circumstances (see earlier discussion). On this account, impure agents, but not harmful agents, should be judged as likely to engage in *all kinds* of violations. We refer to this account of the expectation asymmetry as the global deviance account.

The aims of the present research are threefold. First, we sought to replicate the findings from the pilot studies and garner further support for the expectation asymmetry: more harmful behavior should be anticipated from impure agents than impure behavior should be anticipated from harmful agents (Studies 1–2 and S1–S2). Second, we tested whether the expectation asymmetry could be explained by the global deviance account (Studies 1 and S1). Third, we sought to determine whether the expectation asymmetry can be explained at the level of action perception or action expectation, that is, by observers' perceiving initial impure actions as in fact harmful, or whether, as hypothesized, observers simply expect impure actors to act in ways that are more harmful than their initial impure act (Study 3).

2. Study 1. Impure agents and global deviance

According to the global deviance account, impure agents should be judged as highly likely to engage in all kinds of violations. A liberal test of this account would be to test whether agents who committed one kind of impure violation (e.g., sexually deviant acts, labelled “impure-sexual violations” below) are also expected to commit other kinds of impure violations (e.g., contact with impure substances, labelled “impure-substance violations” below). In Study 1, participants judged the likelihood that an agent who had committed one of three kinds of violations (harmful, impure-sexual, impure-substance) would engage in all three kinds of violations. We sought to replicate the findings of Pilot Studies 1–2 using new scenarios, while also testing the global deviance account of the expectation asymmetry. Our Pilot Study 1 reliably observed the expectation asymmetry using $n = 35$; thus, our recruitment strategy throughout aimed to recruit at least 35–50 participants per cell. Participants were only excluded if they did not complete the entire study. Because Study 1 and Studies S1–S2 involved the most complex designs, these studies required the largest samples.

2.1. Method

2.1.1. Participants

Four participants did not finish the survey and were excluded. We recruited a sample of 255 adults located in the United States (139 female; $M_{age} = 36.68$ years, $SD = 11.37$) via Amazon Mechanical Turk.

2.1.2. Materials and procedures

In a 3 (*agent type*: harmful vs. impure-sexual vs. impure-substance) \times 3 (*expected behavior*: harmful vs. impure-sexual vs. impure-substance) mixed-measures design, participants were randomly assigned to read about an agent who intentionally performed a harmful violation (e.g., assault), an impure-sexual violation (e.g., incest), or an impure-substance violation (e.g., eating dog meat) (see Appendix A for stimuli). For each subdomain of impure acts there were two scenarios, and for harmful acts there were four scenarios. Participants were randomly assigned to one of the eight scenarios. To standardize the agent and target of each action, the scenarios always involved two brothers, and participants were asked to imagine that one of the brothers “willfully and intentionally” engaged in the act. Participants then rated the likelihood that the agent would commit the eight harmful violations ($\alpha = 0.95$) and eight impure-sexual violations ($\alpha = 0.93$) used in Pilot Study 1, and the eight impure-substance violations ($\alpha = 0.94$) from Pilot Study 2. Participants then rated the wrongness of the behavior and completed additional measures,¹ including whether something internal or external to the agent was the likely cause of the action (see Appendix SE), and demographic questions² before being debriefed and paid.

2.2. Results

For brevity, wrongness measures for all studies are reported in Table S1, and internal vs. external attribution measures for Studies 1–2 are reported in Table S2. Three separate within-subjects ANOVAs of likelihood judgments are reported for each agent type: harmful, impure-sexual,

¹ In all studies, we tested ancillary hypotheses building on earlier work of the authors (e.g., Chakroff & Young, 2015; Russell & Piazza, 2015), regarding the emotional reactions elicited by harmful versus impure agents, and inferences about the motives and character of these agents. These measures were not the focus of the present paper and will be summarized only briefly here. Harmful acts tended to evoke anger, while impure acts (sexual and substance) tended to evoke disgust. Harmful actors were perceived to be cruel, while impure actors were perceived as perverse and motivated by sexual desires. Finally, harmful acts were more likely to be seen as having a victim and involving a lack of consent than impure acts (please contact the authors for more information).

² Participants reported their political orientation in all studies. In no analysis did social conservatism significantly predict participants' judgments of likelihood that agents would commit harmful or impure acts.

and impure-substance. See Fig. 1 for likelihood means and standard errors.

2.2.1. Harmful agents

The main effect of expected behavior was significant, $F(2, 132) = 107.52, p < 0.001, \eta^2_p = 0.62$. Post-hoc tests revealed that harmful agents were judged as more likely to commit harmful violations ($M = 5.18, SE = 0.17$) than impure-sexual violations ($M = 2.96, SE = 0.14$) or impure-substance violations ($M = 2.91, SE = 0.15$) ($ps < 0.001$, Bonferroni corrected), but were equally likely to perform the two kinds of impure violations ($p > 0.99$).

2.2.2. Impure-sexual agents

The main effect of expected behavior was significant, $F(2, 62) = 14.60, p < 0.001, \eta^2_p = 0.32$. Post-hoc tests revealed that impure-sexual agents were judged as equally likely to commit impure-sexual violations ($M = 3.86, SE = 0.22$) and harmful violations ($M = 3.91, SE = 0.21$) ($p > 0.99$), and were judged as more likely to commit harmful violations and impure-sexual violations than impure-substance violations ($M = 3.20, SE = 0.21$) ($ps < 0.001$, Bonferroni corrected).

2.2.3. Impure-substance agents

The main effect of expected behavior was significant, $F(2, 55) = 30.34, p < 0.001, \eta^2_p = 0.53$. Post-hoc tests revealed that impure-substance agents were judged as more likely to commit harmful violations ($M = 4.20, SE = 0.25$) than impure-substance violations ($M = 3.57, SE = 0.24$) ($p = 0.004$, Bonferroni corrected), and more likely to commit harmful violations and impure-substance violations than impure-sexual violations ($M = 2.94, SE = 0.21$) ($ps < 0.001$, Bonferroni corrected).

2.2.4. Expectation asymmetry

In a replication of Pilot Study 1, impure-sexual agents were judged more likely to commit harmful violations ($M = 3.91, SE = 0.21$) than harmful agents were to commit impure-sexual violations ($M = 2.96, SE = 0.14$), $F(1, 196) = 14.33, p < 0.001, \eta^2_p = 0.07$. Replicating Pilot Study 2, impure-substance agents were judged more likely to commit harmful violations ($M = 4.20, SE = 0.25$) than harmful agents were to commit impure-substance violations ($M = 2.91, SE = 0.15$), $F(1, 189) = 20.84, p < 0.001, \eta^2_p = 0.10$. Notably, there were no differences in the perceived likelihood that agents would commit non-harm, other-domain violations: that a harmful agent would commit an impure violation (either sexual or substance), that an impure-sexual agent would commit an impure-substance violation, or that an impure-substance agent would commit an impure-sexual violation, $F(2, 252) = 0.59, p = 0.56, \eta^2_p = 0.005$. These judgments highlight a “baseline” of violation expectedness ($M = 3.00, SD = 1.64$), albeit one that does not depend on agents’ past behaviors (e.g., harmful vs. impure). Finally, this exact pattern of results emerged in an independent replication study

with higher power ($N = 325$) and a balanced design (see Study S1 in Supplemental materials).

In sum, we replicated the expectation asymmetry observed in Pilot Studies 1–2. Both kinds of impure agents were judged more likely to be harmful than harmful agents were judged to be impure in either subdomain. However, impure agents were not seen as likely to be globally deviant. Impure agents were judged no more likely than harmful agents to act in a different subdomain of impurity. Put in terms of the stimuli: The man who kissed his brother on the mouth is judged more likely to start fires than to eat worms.

3. Study 2. Ruling out weirdness and wrongness

Because impure violations may be judged as weirder or more atypical than harmful violations (Chakroff & Young, 2015; Gray & Keeney, 2015), the expectation asymmetry may simply reflect participants’ reasonable assumptions that agents are more likely to perform less weird actions (harmful actions) than more weird actions (impure actions). Insofar as the impure acts used in Study 1 may be judged as weirder than the harmful acts, this difference may partially account for why impure agents are expected to engage in both harmful and impure acts. If behavior expectations simply track judgments of behavior typicality, then controlling for the weirdness of the harmful and impure actions (and hence the perceived deviance of the agent) should eliminate the expectation asymmetry.

Study 2 was conducted with the aim of experimentally controlling for the level of weirdness and wrongness in both sets of actions (original and expected behaviors). To this end, we conducted a pre-study to obtain harmful and sexual acts that were matched on both weirdness and wrongness. For the pre-study we developed an inventory of 37 harmful acts and 37 impure-sexual acts, and presented a random subset of roughly half ($n = 18$) the items to 224 undergraduate students recruited from the University of Pennsylvania (78 male; $M_{age} = 19.58$ years, $SD = 1.43$). Participants rated the acts on a number of dimensions, including how “wrong,” “harmful,” “weird,” “contaminating,” and disgusting (i.e., “How nauseous/queasy/grossed out does this act make you feel?”) on 1–9 scales. Mean comparisons for these variables can be found in Appendix B. As discussed above, if the perceived weirdness or moral wrongness of the act is what drives the asymmetry, then experimentally controlling for these factors should eliminate the asymmetry.

3.1. Method

3.1.1. Participants

Six participants did not complete the entire study. We recruited a new sample of 101 participants located in the United States (38 female, $M_{age} = 32.19, SD = 10.35$) via Amazon Mechanical Turk; the recruitment strategy was to obtain $n = 50$ per between-subjects cell.

3.1.2. Materials and procedures

We used a 2 (agent type: harmful vs. impure-sexual) \times 2 (expected behavior domain: harmful vs. impure-sexual) mixed-measures design. We selected from our pre-study three harmful acts and three sexual acts matched on weirdness and wrongness for our manipulation of agent type. We also selected 16 violations that could be matched on weirdness and wrongness to use as expected behaviors (see Appendix B for stimuli, weirdness and wrongness means, and mean comparisons across domains). Ratings from the pre-study confirmed that the harmful acts selected for expected behaviors were rated significantly more harmful than the selected impure acts (see Appendix B for means and mean comparisons). While the mean harmfulness ratings for the initial harmful acts were higher than the initial impure acts, the difference did not quite reach levels of significance due to low power (the pre-study means are based on item-level scores, rather than participant-level scores, since participants received only a subset of items). We also compared whether the initial impure acts were seen as more harmful than the initial harmful

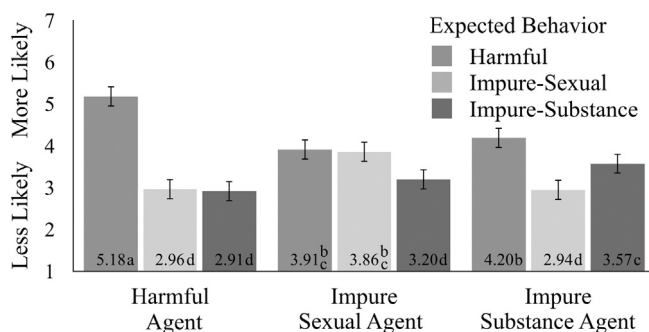


Fig. 1. Study 1 expected behavior means by agent type (harmful, impure-sexual, impure-substance). Means with different letter subscripts differed at $p < 0.01$. Error bars ± 1 SE.

acts were seen as impure, using our pre-study measures of contamination and disgust as indicators of impurity. The impure acts were rated equally harmful ($M = 5.29, SE = 0.50$) as the harmful acts were rated as contaminating ($M = 5.12, SE = 0.89$), $t(4) = 0.11, p = 0.92$, and as disgusting ($M = 5.54, SE = 0.90$), $t(4) = 0.16, p = 0.88$. This implies that any observed expectation asymmetry cannot be explained by initial differences in harmfulness and impurity.

Participants were randomly assigned to one of the six scenarios, and as in previous studies they were asked to imagine that the actor “willfully and intentionally” engaged in the act. As in Study 1, the likelihood indices were reliable: harmful ($\alpha = 0.94$), impure-sexual ($\alpha = 0.94$). In the appraisal block, three items assessed wrongness using the same items as Study 1 ($\alpha = 0.87$). The order of the likelihood and appraisal blocks was randomized.

3.2. Results

There was no difference in wrongness by agent type, confirming that the scenarios were properly matched on wrongness for this study (see Supplemental Materials). We then conducted a 2 (*agent type*) \times 2 (*expected behavior*) mixed-model ANOVA on likelihood judgments. Means and standard errors are reported in Fig. 2. Overall, agents were seen as more likely to be harmful than impure, in a main effect of expected behavior domain, $F(1, 99) = 13.52, p < 0.001, \eta^2_p = 0.12$. There was no main effect of agent type, $F(1, 99) = 0.91, p = 0.34, \eta^2_p = 0.009$. The interaction between agent type and expected behavior was significant, $F(1, 99) = 26.36, p < 0.001, \eta^2_p = 0.21$: harmful agents were seen as more likely to be harmful than impure, but impure agents were seen as equally likely to be harmful and impure. As in Study 1, we found an expectation asymmetry: impure agents were judged as more likely to be harmful than harmful agents were likely to be impure, $F(1, 99) = 4.08, p = 0.046, \eta^2_p = 0.04$.

In sum, when equating harmful and impure violations on wrongness and weirdness, both in terms of the initial actions and expected behaviors, the expectation asymmetry still emerged, suggesting that participants are not simply judging agents who performed weird acts as more likely to perform less weird acts than vice versa.

4. Study 3. Appraisals of the harmfulness of initial and expected acts

In the previous studies, participants judged impure agents as more likely to perform harmful actions than harmful agents were to perform impure actions. This expectation asymmetry was present even when harmful and impure acts were equated on weirdness and wrongness (Study 2). A possible complication, however, is that according to past research impure acts can also be perceived as somewhat harmful (e.g., Gray,

Schein, & Ward, 2014). Thus, an alternative explanation may be that participants perceived the impure acts as harmful, and these perceptions of harm gave rise to expectations that impure (and therefore also harmful) agents would be harmful—that is, a consistency effect for both impure and harmful agents. We suggest that, while this interpretation is plausible for the impure agents, it fails to explain the complete pattern of results for harmful agents. If participants viewed impure acts as harmful, we would expect them to judge the harmful agents as likely to engage in impure (and therefore also harmful) acts, much like impure agents were expected to engage in harmful acts. But harmful agents were not expected to engage in impure acts to the same extent that impure agents were judged likely to commit harmful acts, reflecting the asymmetry. Overall, the pattern of results is more consistent with the idea that impure agents are expected to engage in harmful acts even when their initial actions are not perceived to be especially harmful. If impure acts were simply seen as harmful, there would be no asymmetry.

Here, we sought to obtain direct evidence that impure agents are expected to be harmful, as opposed to the alternative account that their original impure acts were perceived as harmful. In other words, we sought to show that the asymmetry emerges at the level of people's behavior expectations, rather than at the level of people's appraisals of impure actions. To tease apart these two accounts, we had participants directly appraise the harmfulness of all acts (original and expected), in addition to judging the likelihood of agents committing harmful and impure acts. Consistent with the view that participants expect impure agents to behave harmfully, rather than the view that participants perceive the original impure actions as harmful, we predicted that participants would expect impure agents to commit acts that were more harmful than the initial acts they performed.

4.1. Method

4.1.1. Participants

We recruited a new sample of 197 participants located in the United States (79 female, $M_{age} = 33.19, SD = 10.67$) via Amazon Mechanical Turk.

4.1.2. Materials and procedures

We selected 10 harmful behaviors and 10 impure-sexual behaviors previously used as either initial or expected behaviors in Study 2 (see Appendix C). Participants were randomly assigned to view 1 of the 20 acts as an initial behavior intentionally committed by an agent named John. Participants then rated the likelihood that the agent would intentionally commit each of the 20 acts (i.e., expectation of behavior); all items were presented in random order on a single page. Thus, in this study, each item from the set of 20 acts was at some point treated as the initial act, and appraisals were made over all 20 items, including the initial act. This approach enabled us to directly compare mean differences in appraisals of initial acts and expected behaviors. Participants indicated whether the actor was likely to engage in the expected behavior by selecting one of three options: 0 (“Not at all likely”), 1 (“Somewhat likely”), and 2 (“Very likely”). We focused our analysis on behaviors participants thought the actor was at least “Somewhat likely” to perform in the future. Note that we included three options to allow for gradation in participants’ “Likely” responses, yet the aim for analysis was to collapse together acts that were rated as either somewhat or very likely.

On the next page, participants judged “How harmful” each of the 20 acts was, using Likert scales ranging from 1 (“Not at all Harmful”) to 7 (“Extremely Harmful”). Subsequent pages contained measures of disgust, weirdness, and then moral wrongness, presented with the same design as the harmfulness measure. We included a measure of disgust as a proxy measure for impure actions. We included this measure with the aim of showing that, consistent with the expectation asymmetry, harmful agents are not expected to engage in more disgusting actions than performed previously, whereas impure agents are expected to engage in more harmful actions than performed previously. No

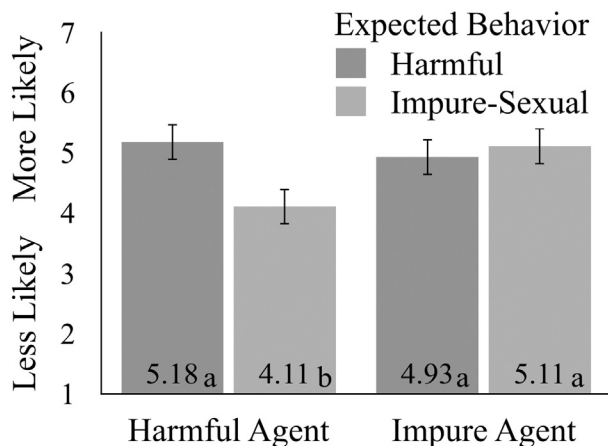


Fig. 2. Study 2 expected behavior means by agent type (harmful versus impure), collapsing across levels of extremeness for both agent type and expected behavior. Means with different letter subscripts differed at $p < 0.05$. Error bars ± 1 SE.

other measures were collected, aside from basic demographic information (e.g., age, gender, political orientation).

4.2. Results and discussion

4.2.1. Likelihood of harmful and impure acts

Distributions of likelihood judgments (i.e., not at all to very likely) by initial action and expected action domain are displayed in Fig. 3. For analysis, we compared the mean number of expected (i.e., “likely”) acts across domains. We used a non-parametric test (Mann-Whitney) as the mean range was restricted from 0 to 10. In a conceptual replication of Studies 1–2 (and Studies S1–S2), we found an expectation asymmetry: participants expected impure agents to perform more harmful acts ($M = 4.20, SE = 0.37$), than the number of impure acts they expected harmful agents to perform ($M = 3.34, SE = 0.34$) ($U = 3930.0, p = 0.045$) – in Fig. 3, the contrast of $H \rightarrow I$ and $I \rightarrow H$.

4.2.2. Appraisals

We conducted separate 2 (*agent type*) \times 2 (*expected behavior*) mixed-model ANOVAs on appraisals of harmfulness, disgust, weirdness, and wrongness. We found no significant main effects of, or interactions with, agent type (all p s > 0.14), and thus report only means from the main effects of expected behavior in Table 1. In brief, participants judged harmful behaviors as more harmful than impure behaviors, impure behaviors as more disgusting and weirder than harmful behaviors, and harmful and impure behaviors as similarly wrong. Thus, while the impure and harmful acts were judged equally wrong, harmful acts were rated significantly more harmful than the impure acts. This pattern of results is inconsistent with the perceived harm account. Focusing on the initial acts, as in Study 2, we examined whether impure acts were seen as more harmful than harmful acts were seen as impure, using disgust as an indicator of impurity. Quite the opposite, harmful acts were judged as slightly more disgusting ($M = 4.27, SE = 2.16$) than impure acts were judged as harmful ($M = 3.63, SE = 2.20$), $t(195) = 2.20, p = 0.045$, a result that notably goes in the opposite direction of the observed expectation asymmetry, and thus, if anything, works against our hypothesis.

4.2.3. Comparing initial and expected acts

Fig. 4 depicts difference scores for expected – initial acts, separately for four DVs (harmful, disgusting, weird, wrong), and two types of initial acts (harmful vs. impure). Full statistics for each comparison are reported in Table 2. In general, aggregating across DVs, agents were expected

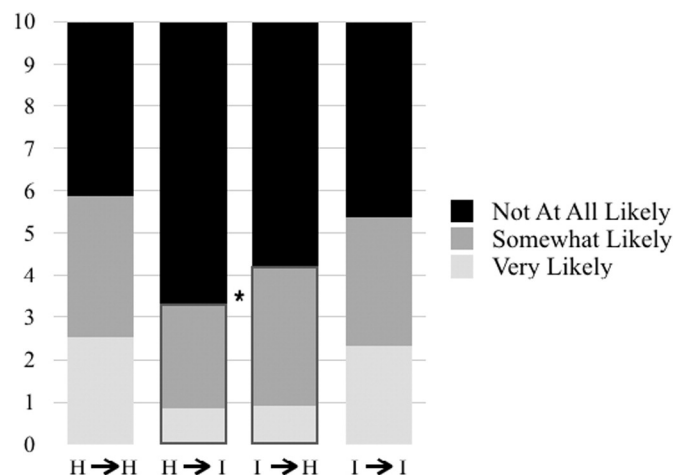


Fig. 3. Distributions of likelihood judgments for each kind of initial action and expected behavior (Harmful versus Impure) for Study 3. For example, column $H \rightarrow I$ reflects the average number of impure acts, out of 10, that harmful agents are expected to perform: Not at all likely (black), somewhat likely (darker gray), or very likely (lighter gray). Participants expected impure agents to perform more harmful acts ($I \rightarrow H$) than they expected harmful agents to perform impure acts ($H \rightarrow I$).

Table 1
Appraisal and judgment means, standard errors, and statistics for Study 3.

Measure	Expected behaviors		F	η^2_p
	Harmful	Impure		
Harmful	5.07 (0.07)	4.17 (0.11)	107.74*	0.36
Disgusting	4.24 (0.10)	5.25 (0.08)	156.26*	0.45
Weird	4.78 (0.09)	5.57 (0.07)	131.41*	0.40
Wrong	5.08 (0.08)	5.18 (0.09)	2.25 ^{ns}	0.01

* $p < 0.001$.

^{ns} $p > 0.05$.

to perform acts that are less “intense” (i.e., harmful, disgusting, weird, and wrong) than their initial acts, $F(1184) = 8.400, p = 0.004, \eta^2_p = 0.04$.

However, there was one exception to this general pattern. As can be seen in Fig. 4, under column $I \rightarrow H$ (i.e., contrasts of Impure initial actions with expected Harmful actions), agents who originally committed impure acts were expected to engage in harmful actions rated as significantly *more harmful* than the original acts they performed. This was the only comparison that was significantly positive ($M = 0.52, SE = 0.21$), $t(69) = 2.26, p = 0.027$, indicating that, against the trend, impure agents were expected to perform acts that were significantly more harmful than their initial act. In contrast, see column $H \rightarrow I$. Harmful agents were expected to engage in harmful acts that were *less harmful* than their initial act ($M = -0.35, SE = 0.15$), $t(106) = -2.39, p = 0.019$. Also, harmful agents were not expected to engage in impure acts that were more disgusting than their initial harmful act ($M = -0.07, SE = 0.24, t(84) = 0.2839, p = 0.78$). This point is worth emphasizing. The impure acts that harmful agents were expected to perform were judged as no more disgusting than the initial harmful act. In contrast, the harmful acts that impure agents were expected to perform were significantly more harmful than the initial impure act.

These results reveal that impure agents are expected to be more harmful than they were previously, as explicitly indicated by the participants. Furthermore, this inference goes against the grain of all other inferences. Harmful agents are not expected to be more harmful than they were previously, and no agent is expected to be more disgusting, weird, or wrongful than they were previously. Thus, this study isolates expectations of harmfulness as the main driver of the expectation asymmetry. Furthermore, these results support our claim that the expectation asymmetry results from an expectation that impure agents will be harmful—indeed more harmful than past behavior might suggest.

5. General discussion

Across seven studies (including two pilot studies, one direct replication [S1], and another study that examined self- vs. other-directed acts [S2]), we found an asymmetry in participants' expectations of immoral agents' behavior: impure agents were judged more likely to be harmful than harmful agents were expected to be impure. Interestingly, impure agents were not seen as globally deviant, i.e., likely to engage in all kinds of immoral behavior (see Study 1 and its replication Study S1). Although slight variations emerged across studies, we found broad evidence for a consistency effect—impure agents (and harmful agents) were expected to behave in a manner similar to their prior behavior (e.g., impure-sexual \rightarrow impure-sexual). However, impure agents were not expected to be impure in a qualitatively different manner (e.g., impure-sexual \rightarrow impure-substance). This finding was replicated in Study S1. Furthermore, the expectation asymmetry was not accounted for by differences in the perceived wrongness, harmfulness, or weirdness of the harmful versus impure acts (Studies 2–3). The asymmetry emerged when the initial harmful and impure acts were experimentally matched on wrongness and weirdness (Study 2) and when the impure acts were appraised as much less harmful than the harmful acts (Study 3).

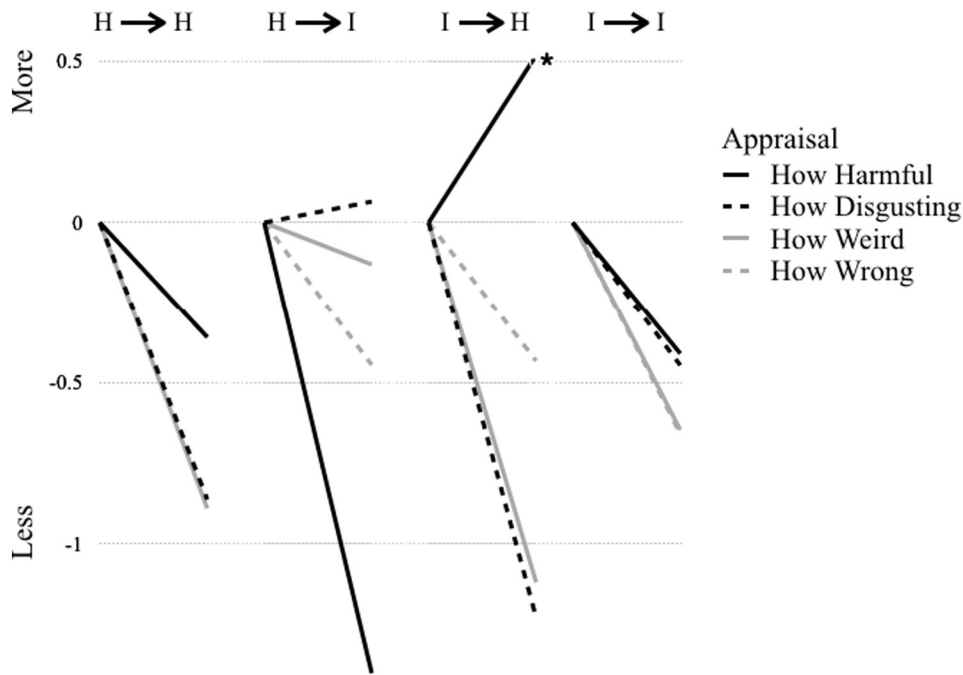


Fig. 4. Appraisal measures for initial acts versus expected acts for Study 3, zeroed to the magnitude of the initial act. Harmfulness plotted in solid black; Disgust plotted in dotted black; Weirdness plotted in solid gray; Wrongness plotted in dotted gray. Column labels (e.g., H → I) indicate the kind of act initially performed, and following the →, the kind of act that was expected (H = Harmful, I = Impure). The positive sloped solid black line compares the harmfulness of initial impure acts to the harmfulness of impure agents' expected harmful acts, the only comparison which revealed greater magnitude for expected > initial appraisals.

What can account for the expectation asymmetry? In particular, why are impure agents expected to be harmful? One possibility is that people perceive impure acts as causing harm in addition to being impure (see Gray & Keeney, 2015; Gray et al., 2014). While there are certainly instances in which people perceive impure actions as causing harm, this explanation does not account for the full pattern of the present results. First, if participants saw impure acts as harmful, they should have also judged harmful agents as likely to engage in harmful and impure acts at similarly high levels, particularly when all acts were matched for weirdness and wrongness. However, we observed the expectation asymmetry even when the weirdness and wrongness of the harmful and impure acts were equated experimentally (Study 2). Perhaps more importantly, in Study 3, the asymmetry persisted despite clear differences in the perceived harmfulness of the harmful and impure actions.

Table 2
Appraisals and judgments difference scores for expected acts > initial acts for Study 3, reported separately for harmful versus impure initial acts, and harmful versus impure expected acts.

Label	Initial	Expected	DV	Expected > initial	t	p
H → H	Harmful	Harmful	Harmful	-0.35	-2.39	0.019
			Disgusting	-0.86	-4.87	<0.001
			Weird	-0.90	-4.96	<0.001
			Wrong	-0.36	-1.84	0.069
H → I	Harmful	Impure	Harmful	-1.40	-6.45	<0.001
			Disgusting	0.07	0.28	0.781
			Weird	-0.14	-0.59	0.559
			Wrong	-0.44	-1.77	0.080
I → H	Impure	Harmful	Harmful	0.52	2.26	0.027
			Disgusting	-1.29	-5.31	<0.001
			Weird	-1.15	-4.91	<0.001
			Wrong	-0.46	-1.85	0.068
I → I	Impure	Impure	Harmful	-0.40	-1.76	0.083
			Disgusting	-0.46	-2.07	0.042
			Weird	-0.64	-2.70	0.009
			Wrong	-0.66	-2.83	0.006

Another possibility is that differences in people's causal attributions for impure versus harmful acts may contribute to the expectancy asymmetry. As mentioned in the Introduction, past research has shown that harmful acts are generally thought to stem from both internal and external causes, while impure acts are thought to be largely internally generated (Chakroff & Young, 2015; Russell & Piazza, 2015). Indeed, in Studies 1–2 (see Table S2 in Supplemental materials), we found that the impure acts (especially sexually impure acts) were largely attributed to internal causes, whereas harmful acts were attributed either more to external than internal causes, or equally to each. Thus, insofar as impure acts draw attention to the relative contribution of internal causes, this focus on internal causes may increase the expectation that the impure agent will behave immorally in other respects. However, in Study 1 and its replication (see Supplemental materials), we found that impure agents were not expected to behave immorally in *all* respects, only in similarly impure and harmful ways. Thus, differences in causal attributions do not help explain why harmful actions are given priority over other kinds of other-domain actions when people estimate the likelihood of immoral behaviors.

Another plausible account is that people conceptually equate wrongdoing with harming others (Gray & Schein, 2012; Gray, Young, & Waytz, 2012; Gray et al., 2014). Thus, insofar as impure agents are conceptualized as wrongdoers, as they were in many of our studies (Studies 2–3 used impure and harmful actions matched on measures of wrongdoing), then observers may use a harm-based conceptual framework as the basis for forecasting how agents may behave in the future. Because conceptualizations of wrongdoing and harm are closely tied together, at least more so than wrongdoing and impurity (see Gray & Schein, 2012; Gray et al., 2014), then intuitions of harm may be accessible for use when judging the likely behavior of impure agents. The inverse however may not be the case when predicting the behavior of harmful agents. We believe this explanation deserves further empirical scrutiny (see also Supplemental Study S2 for additional support). Future studies should continue to probe the likely complex factors that contribute to this expectation asymmetry. Indeed, in Study 3 we found that people expected impure agents to act in ways that were more harmful, but also less wrong (see Fig. 4, I → H). Thus, in at least

one of our studies, expectations about harmfulness and wrongdoing were clearly dissociated, contrary to what a harm-centric view of wrongdoing would predict.

5.1. Theoretical implications and limitations

Our studies have important implications for current debates surrounding how people think about impure actions and the reasons why impure actors are often treated harshly. One account that we tested is that people cannot help but perceive impure actions as harmful, much like people cannot help but see the absent triangle within the Kanisza triangle perceptual illusion (see Gray et al., 2014). However, in Study 3 we used impure actions that were explicitly appraised by our participants as less harmful than the harmful actions, yet matched on wrongness. Despite these differences in perceived harm, the expectation asymmetry still emerged. People still expected impure agents to perform acts that were more harmful than their initial acts – a pattern not seen for judgments of harmful agents who were expected to engage in somewhat less harmful actions. Thus, people do not need to think that impure agents (e.g., zoophiles) are actually causing harm to anticipate that they will engage in harmful practices (e.g., torturing animals) in the future. It is sufficient to simply appraise the actor as engaging in impure forms of wrongdoing for expectations of harmfulness to emerge. This finding is important because it suggests that addressing the perceived harmfulness of impure actions may not be an adequate intervention to deterring expectations of future harmfulness. Rather, what may need to be redressed is the construal of the impure action as a form wrongdoing.

Nonetheless, our studies had a number of limitations that should be addressed in future research. First, in our test of the global deviance hypothesis (Study 1), we relied on two sub-domains of impurity. Future research could test the scope of this account by examining other kinds of violations, such as acts of disobedience or disloyalty towards one's group (e.g., washing a toilet with the national flag) which, like impure acts, are not overtly harmful yet are condemned by many (Graham et al., 2009; Gray & Schein, 2012; Wright & Baril, 2011). If disobedient and disloyal agents are also expected to be harmful, but not immoral in other respects (e.g., impure), this would provide further evidence for the account that harmful acts dominate people's expectations of immoral actors (Gray & Schein, 2012; Gray et al., 2014). Second, we did not explicitly specify whether recipients of sexual-impure acts gave his or her consent. Since consensual sexual acts are often seen as permissible (Russell & Piazza, 2015), future work should examine whether agents who commit impure acts with the consent of the recipient are just as likely to be judged as harmful as those who act without consent.

Finally, future studies should examine whether this expectation asymmetry is limited to transgressive actions or may apply in an analogous way to non-transgressive actions. For example, in the domain of recreational activities, we might examine expectations about a person who has tried bungee jumping for the first time. Do people expect this person to next attempt another non-traditional sport, like hang gliding or water polo, or judge them more likely to perform a more prototypical sport like tennis or swimming? Such a study would help elucidate the underlying conceptual processes operating within the expectation asymmetry by determining whether a similar asymmetry may occur outside of the scope of morality.

5.2. Conclusion

The present work has revealed a robust asymmetry, within the moral domain, in people's expectations about immoral agents: harmful agents are judged as likely to behave in similarly harmful ways, while impure agents are judged as likely to behave in both impure and harmful ways. This asymmetry is a tragic quirk of our minds, one that may underlie the reactive and punitive measures taken against individuals whose behaviors are largely harmless but are nonetheless perceived to defy normative standards of propriety. Our findings highlight the

need for increased public awareness of the natural tendency to stigmatize individuals as dangerous, for example, when individuals have engaged in harmless, but to the minds of some, unacceptable actions.

Acknowledgments

Many thanks to Joshua Greene, Fiery Cushman, Justin Landy, and Roger Giner-Sorolla for helpful discussion and comments. Thanks also to Daniel Noon for help with coding. This project was made possible by a grant from the John Templeton Foundation and also by the University of Surrey pump priming scheme.

Appendix A

Study 1 scenarios

Harmful

John cuts his brother with a sharp knife.
John pours a cup of painfully hot water on his brother's lap.
John slaps his brother's face.
John hangs his brother by a rope for a few seconds then releases him.

Impure-sexual

John kisses his brother on the mouth.
John strokes his brother's bare inner thigh.

Impure-substance

John buys dog meat from a foreign grocery store and serves it to his brother.
John orders stir-fried rat for his brother at a small restaurant in a city in the USA.

Study 1 expected behaviors

Harmful

Vandalize property
Start fires
Harm innocent animals
Tell lies to get ahead in his career
Get into fist fights
Drive under the influence of drugs or alcohol
Steal from a corporate organization
Cheat on an exam

Impure-sexual

Have sex with an elderly man/woman
Have sex with a dead animal, such as a dog or deer
Have an incestuous sexual relationship
Masturbate in public
Expose himself in public
Have sex with a corpse
Have sex with someone of the same gender⁺
Have sex with roast meat such as chicken or pork tenderloin
⁺ While this item may not be construed by many individuals as an impure action, empirically this item factor loaded with the other impure-sexual items (e.g., at 0.644 for Pilot Study 1), and thus was retained.

Impure-substance

Eat a worm
Pick up a used band-aid
Drink cow blood
Taste his own earwax
Step in vomit barefoot
Pick up dog poop barehanded
Pour urine on himself
Lick someone's shoe

Appendix B

Study 2 initial acts

Harmful ($M_{\text{wrong}} = 6.95$; $M_{\text{weird}} = 7.61$; $M_{\text{harmful}} = 7.65$; $M_{\text{disgusting}} = 5.55$; $M_{\text{contaminating}} = 5.12$).

- Mark cuts off his friend's nipple.
- Mark sticks the tip of a pencil in his friend's ear while he sleeps.
- Mark hits his friend over the head with a textbook.

Impure-sexual ($M_{\text{wrong}} = 6.92$; $M_{\text{weird}} = 7.84$; $M_{\text{harmful}} = 5.28$; $M_{\text{disgusting}} = 7.06$; $M_{\text{contaminating}} = 5.49$).

- Mark sticks his penis in his dead girlfriend's body.
- Mark sucks on his sister's fingers one at a time.
- Mark kisses his sister on the mouth.

Mean comparisons

Wrongness: 6.95 ($SD = 1.07$)–6.92 ($SD = 1.44$), $t(4) = 0.024$, $p = 0.982$; Weirdness: 7.61 ($SD = 1.10$)–7.84 ($SD = 1.14$), $t(4) = -0.25$, $p = 0.817$; Harmfulness: 7.65 ($SD = 0.95$)–5.28 ($SD = 1.74$), $t(4) = 2.06$, $p = 0.108$; Disgusting: 5.55 ($SD = 2.21$)–7.06 ($SD = 1.60$), $t(4) = -0.96$, $p = 0.39$; Contaminating: 5.12 ($SD = 2.20$)–5.49 ($SD = 2.37$), $t(4) = -0.20$, $p = 0.85$.

Study 2 expected behaviors

Harmful ($M_{\text{wrong}} = 6.28$; $M_{\text{weird}} = 6.28$; $M_{\text{harmful}} = 6.88$; $M_{\text{disgusting}} = 5.37$; $M_{\text{contaminating}} = 4.14$).

- Wring his sister's arm until it gives her an "Indian brush burn."
- Elbow his friend in the stomach.
- Spank his brother on his bottom.
- Pierce his partner's belly button with a needle.
- Superglue pennies and other small objects to a stray cat.
- Pierce the webbed feet of his pet frog with earrings.
- Smash a lizard with a hammer.
- Choke his girlfriend for a few seconds then release her.

Impure-sexual ($M_{\text{wrong}} = 5.86$; $M_{\text{weird}} = 6.36$; $M_{\text{harmful}} = 4.73$; $M_{\text{disgusting}} = 5.93$; $M_{\text{contaminating}} = 4.46$).

- Caress the genitals of a transgender friend.
- Grope his friend's bottom.
- Have sex with his girlfriend in a public bathroom stall.
- Have anal sex with his romantic partner who had gender reassignment surgery.
- Fondle a goat's genitals.
- Tickle his son's genitals with a feather.
- Stroke his mother's inner thigh.
- Let his stepbrother fondle his genitals.

Mean comparisons

Wrongness: 6.28 ($SD = 1.48$) – 5.86 ($SD = 1.84$), $t(14) = 0.50$, $p = 0.621$; Weirdness: 6.48 ($SD = 1.66$) – 6.36 ($SD = 1.88$), $t(14) = 0.14$, $p = 0.895$; Harmfulness: 6.88 ($SD = 0.90$) – 4.73 ($SD = 1.25$), $t(14) = 3.92$, $p = 0.002$; Disgusting: 5.37 ($SD = 1.52$) – 5.93 ($SD = 1.60$), $t(14) = -0.71$, $p = 0.49$; Contaminating: 4.14 ($SD = 1.50$) – 4.46 ($SD = 1.12$), $t(4) = -0.47$, $p = 0.64$.

Appendix C

Study 3 acts

When presented as an initial act, each item followed the agent's name "John," for example, "John elbowed his friend in the stomach." When presented as a expected act, the act was phrased in present tense and presented without an agent, for example, "Elbow his friend in the stomach".

Harmful

- John wrung his sister's arm until it gave her an "Indian brush burn."
- John elbowed his friend in the stomach.
- John spanked his brother on his bottom.
- John pierced his partner's belly button with a needle.
- John superglued pennies and other small objects to a stray cat.
- John pierced the webbed feet of his pet frog with earrings.
- John choked his girlfriend for a few seconds then release her.
- John cut off his friend's nipple.
- John stuck the tip of a pencil in his friend's ear while he sleeps.
- John hit his friend over the head with a textbook.

Impure-sexual

- John caressed the genitals of a transgender friend.
- John groped his friend's bottom.
- John had sex with his girlfriend in a public bathroom stall.
- John had anal sex with his romantic partner who had gender reassignment surgery.
- John fondled a goat's genitals.
- John tickled his son's genitals with a feather.
- John stroked his mother's inner thigh.
- John stick his penis in his dead girlfriend's body.
- John sucked on his sister's fingers one at a time.
- John kissed his sister on the mouth.

Appendix D. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.jesp.2016.08.001>.

References

- Beetz, A. M. (2002). *Love, violence, and sexuality in relationships between humans and animals*. Aachen: Shaker Verlag.
- Beetz, A. M. (2004). Bestiality/zoophilia: A scarcely investigated phenomenon between crime, paraphilia, and love. *Journal of Forensic Psychology Practice*, 4(2), 1–36.
- Chakroff, A., & Young, L. (2015). Harmful situations, impure people: An attribution asymmetry across moral domains. *Cognition*, 136, 30–37.
- Chakroff, A., Dungan, J., & Young, L. (2013). Harming ourselves and defiling others: What determines a moral domain? *PLoS One*, 8(9), e74434.
- Chakroff, A., Dungan, J., Koster-Hale, J., Brown, A., Saxe, R., & Young, L. (2015). When minds matter for moral judgment: intent information is neurally encoded for harmful but not impure acts. *Social Cognitive and Affective Neuroscience* (nsv131).
- Cunliffe, J., & Shepherd, A. (2007). Re-offending of adults: Results from the 2004 cohort. *Home Office Statistical Bulletin 06/04*. London: Home Office.
- Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analysis in moral judgment. *Cognition*, 108, 353–380.
- Dennett, D. C. (1989). *The intentional stance*. Cambridge, MA: MIT Press.
- Gettys, T. (2014). Anti-LGBT researcher Paul Cameron open to death penalty for 'dangerous' gay 'parasites'. *The raw story* Retrieved from <http://www.rawstory.com/rs/2014/02/anti-lgbt-researcher-paul-cameron-open-to-death-penalty-for-dangerous-gay-parasites/> (Oct. 2014)
- Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, 96(5), 1029.
- Gray, K., & Keeney, J. E. (2015). Impure or just weird? Scenario sampling bias raises questions about the foundation of morality. *Social Psychological and Personality Science*. <http://dx.doi.org/10.1177/1948550615592241>.
- Gray, K., & Schein, C. (2012). Two minds vs. two philosophies: Mind perception defines morality and dissolves the debate between deontology and utilitarianism. *Review of Philosophy and Psychology*, 3(3), 405–423.
- Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, 23(2), 101–124.
- Gray, K., Schein, C., & Ward, A. F. (2014). The myth of harmless wrongs in moral cognition: Automatic dyadic completion from sin to suffering. *Journal of Experimental Psychology: General*, 143(4), 1600–1615.
- Inbar, Y., Pizarro, D. A., Knobe, J., & Bloom, P. (2009). Disgust sensitivity predicts intuitive disapproval of gays. *Emotion*, 9, 435–439.
- Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska symposium on motivation*. University of Nebraska Press: Lincoln.
- Langhinrichsen, P. A., & Levin, D. J. (2002). *Recidivism of prisoners released in 1994*. Bureau of Justice Statistics Special Report. U.S. Department of Justice <http://www.bjs.gov/content/pub/pdf/rpr94.pdf>
- Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology*, 35, 250–256.
- Piazza, J., Russell, P. S., & Sousa, P. (2013). Moral emotions and the envisaging of mitigating circumstances for wrongdoing. *Cognition and Emotion*, 27(4), 707–722.

- Quoidbach, J., Gilbert, D. T., & Wilson, T. D. (2013). The end of history illusion. *Science*, 339(6115), 96–98.
- Rozin, P., Lowery, L., Imada, S., & Haidt, J. (1999). The CAD triad hypothesis: A mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *Journal of Personality and Social Psychology*, 76, 574–586.
- Russell, P. S., & Giner-Sorolla, R. (2011a). Moral anger, but not moral disgust, responds to intentionality. *Emotion*, 11, 233–240. <http://dx.doi.org/10.1037/a0022598>.
- Russell, P. S., & Giner-Sorolla, R. (2011b). Social justifications for moral emotions: When reasons for disgust are less elaborated than for anger. *Emotion*, 11, 637–646. <http://dx.doi.org/10.1037/a0022600>.
- Russell, P. S., & Giner-Sorolla, R. (2011c). Moral anger is more flexible than moral disgust. *Social Psychological and Personality Science*, 2, 360–364.
- Russell, P. S., & Giner-Sorolla, R. (2013). Bodily moral disgust: What it is, how it is different from anger, and why it is an unreasoned emotion. *Psychological Bulletin*, 139, 328–351. <http://dx.doi.org/10.1037/a0029319>.
- Russell, P. S., & Piazza, J. (2015). Consenting to counter-normative sexual acts: Differential effects of consent on anger and disgust as a function of transgressor or consenter. *Cognition and Emotion*, 29, 634–653. <http://dx.doi.org/10.1080/02699931.2014.930420>.
- Saxe, R. (2012). The happiness of the fish: Evidence for a common theory of one's own and others' actions. In K. D. Markman, W. M. Klein, & J. A. Suhr (Eds.), *The handbook of imagination and mental simulation*. East Sussex, UK: Psychology Press.
- Sousa, P., Holbrook, C., & Piazza, J. (2009). The morality of harm. *Cognition*, 113, 80–92.
- Thorndike, E. (1920). A constant error on psychological rating. *Journal of Applied Psychology*, 4, 25–29.
- Wright, J. C., & Baril, G. (2011). The role of cognitive resources in determining our moral intuitions: Are we all liberals at heart? *Journal of Experimental Social Psychology*, 47(5), 1007–1012.
- Young, L., & Saxe, R. (2011). When ignorance is no excuse: Different roles for intent across moral domains. *Cognition*, 120(2), 202–214.