Supplemental Materials for Chakroff, Russell, Piazza, and Young,

"From Impure to Harmful: Asymmetric Expectations about Immoral Agents"

Pilot Study 1: Expectations of Harmful versus Impure-Sexual Agents

Method

Participants. We recruited 73 adults located in the United States (33 female; M_{age} = 31.13 years, *SD* = 9.30) via Amazon Mechanical Turk (www.mturk.com). The aim was to recruit 35 participants per condition, 70 total, and data analysis did not commence until this recruitment goal was attained.

Materials and procedures. In a 2 (*agent type*: harmful vs. impure-sexual) x 2 (*expected behavior*: harmful vs. impure-sexual) mixed-measures design participants were randomly assigned to read about an agent who intentionally performed a *harmful* violation (assault) or *impure-sexual* violation (incest) (see Appendix SA for scenarios and measures). Scenarios did not contain information about the agent's motives, so participants could imagine the agent's motives and the circumstances surrounding the act. The two acts were selected from past research as being comparably wrong (Piazza et al., 2013). Nevertheless, we had our new sample of participants rate the wrongness of the behavior (1 = *Not at all wrong*; 7 = *Very wrong*). Participants also completed appraisals of the behavior and/or agent's *character*, and reported their feelings of anger and disgust in reaction to the scenarios. Participants then rated the likelihood that the agent would commit eight impure-sexual violations (α = .92) and eight harmful violations (α = .94). Finally, participants completed a standard demographic questionnaire before being debriefed and paid. No other measures were collected.

Results

The main effect of *agent type* was not significant, $F(1, 71) = 0.67$, $p = .42$, $\eta^2_p = .009$, but the main effect of *expected behavior* was significant, $F(1, 71) = 61.23$, $p < .001$, $\eta^2_p = .46$, such that agents were judged as more likely to be harmful versus impure overall. The interaction between *agent type* and *expected behavior* was significant, $F(1, 71) = 98.32$, $p < .001$, $\eta^2_p = .581$ (see Fig. S1 for means and standard errors). We rescored the expected behavior measures in terms of behavior performed in a *different* domain (e.g., how likely a harmful agent is to be impure); a one-way ANOVA on this new measure revealed a behavior expectation asymmetry: the impure agent was more likely to commit harmful violations than the harmful agent was to commit impure violations, $t(71) = 3.45$, $p = .001$, $d = .81$. Additionally, the harmful agent was seen as more likely to commit harmful versus impure-sexual violations, $t(36) = 11.75$, $p < .001$, $d = 3.92$, whereas the impure agent was seen as just as likely to commit harmful and impure-sexual violations, $t(35) = 1.60$, $p = .12$, $d = .54$. Finally, when comparing the expected same-domain behaviors (harmful→harmful vs. impure→impure), the harmful agent was expected to engage in same-domain behaviors more than the impure agent was, $t(71) = 2.05$, $p = .044$, $d = .48$.

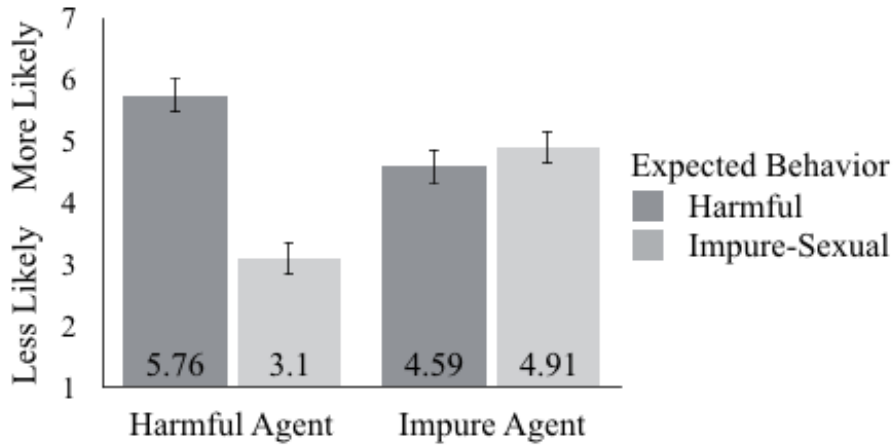

Figure S1. Pilot Study 1 expected behavior means by agent type (harmful versus impure). Error bars ± 1 S.E.

In sum, participants judged the impure agent as more likely to commit a harmful violation than they judged the harmful agent as likely to commit an impure-sexual violation. Also, while participants judged the harmful agent as more likely to commit harmful versus impure-sexual violations, they judged the impure agent as similarly likely to commit impure-sexual and harmful violations. Having established the behavior expectation asymmetry using one subdomain of impurity (sexual), we sought to replicate our findings using another subdomain (substance).

Pilot Study 2. Expectations of Harmful versus Impure-Substance Agents

Method

Participants. We recruited 165 adults located in the United States (57 female; M_{age} = 30.8 years, *SD* = 9.04) via Amazon Mechanical Turk. Four participants did not finish the survey and were excluded from analysis.

Materials and procedures. In a 2 (*agent type*: harmful vs. impure-substance) x 2 (*expected behavior*: harmful vs. impure-substance) mixed-measures design, participants read a sentence describing an agent committing either a harmful act (e.g., "John once cut someone with a knife when he was upset") or a impure-substance act (e.g., "John once found a dead dog and cooked it up for dinner") - see Appendix SB for scenarios and measures. Participants rated the likelihood that the agent would engage in various behaviors, consisting of nine harmful acts (α = .97) and nine impure-substance acts (α = .97). All likelihood judgments were collected on a single page, with order randomized across participants. Judgments of likelihood were made on 100-point slider scales with anchors "Not at all likely" to "Absolutely likely." Participants then rated the wrongness of the behavior (1 = *Not at all wrong*; 7 = *Very wrong*). Finally, participants read a sentence describing an agent engaging in a morally neutral action (e.g., Steven once cooked up a pot roast for dinner), and completed all above measures regarding the neutral agent.

Results

We found main effects of both *agent type*, $F(1,159) = 6.60$, $p = .01$, $\eta^2_p = .04$, and *expected behavior*, $F(1,159) = 6.59$, $p = .01$, $\eta^2_p = .04$. Consistent with Pilot Study 1, we found an interaction between *agent type* and *expected behavior*, $F(1,159) = 241.29$, $p < .001$, $\eta^2_p = .60$ (see Fig. S2 for means and standard errors). As in Pilot Study 1, we found a behavior expectation asymmetry: the impure agent was judged more likely to commit harmful violations than the harmful agent was to commit impure-substance violations, $F(1, 159) = 11.30$, $p = .001$, $\eta^2_p = .07$. Additionally, the harmful agent was seen as more likely to commit harmful versus impure-sexual violations, $t(159) = 5.34$, $p < .001$, whereas the impure agent was seen as more likely to commit impure-substance than harmful violations, $t(159) = -10.73$, $p < .001$ – a divergence from Pilot Study 1. The harmful agent was seen as no more likely to engage in same-domain behaviors more than the impure agent was, $t(159) = -.96$, $p = .34$. Notably, agents who committed morally neutral actions were judged no more likely to commit harmful than impure-substance violations, $F(1, 160) = 0.26$, $p = .61$, $\eta^2_p = .002$ (harmful: $M = 15.61$, $SE = 1.42$; impure-substance: $M = 15.25$, $SE = 1.43$), ruling out the possibility that people generally judge others (even non-transgressors) as more likely to be harmful than impure.

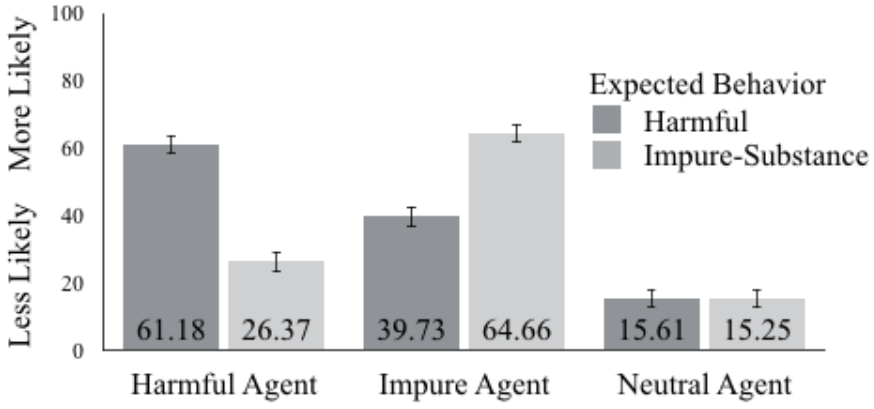

Figure S2. Pilot Study 2 expected behavior means by agent type (harmful versus impure). Error bars ± 1 S.E.

In sum, when comparing harmful to *impure-substance* violations, a second subdomain of impurity, we found that impure agents were judged as more likely to be harmful than harmful agents were likely to be impure. We also found that harmful agents were judged as more likely to be harmful than impure, and impure agents were judged as more likely to be impure than harmful. Importantly, the behavior expectation asymmetry was not found for morally neutral, non-deviant agents.

Study S1: Replication

Method

Participants. We recruited a sample of 325 adults located in the United States (126 female; M_{age} = 34.42 years, SD = 12.11) via Amazon Mechanical Turk.

Materials and procedures. In a 3 (*agent type*: harmful vs. impure-sexual) x 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). For each agent type, there were two scenarios; therefore, participants were randomly assigned to one of the six action domain scenarios (see Appendix SC for full list of scenarios). To standardize the agent and target of each action, the scenarios were worded so that they 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 24 violations used in Study 1. Participants then completed demographic questions before being debriefed and paid.

Results

Separate 3 (*expected behavior*) within-subjects ANOVAs on likelihood judgments are reported for each agent type: harmful, impure-sexual, and impure-substance. See Fig. S3 for likelihood means and standard errors.

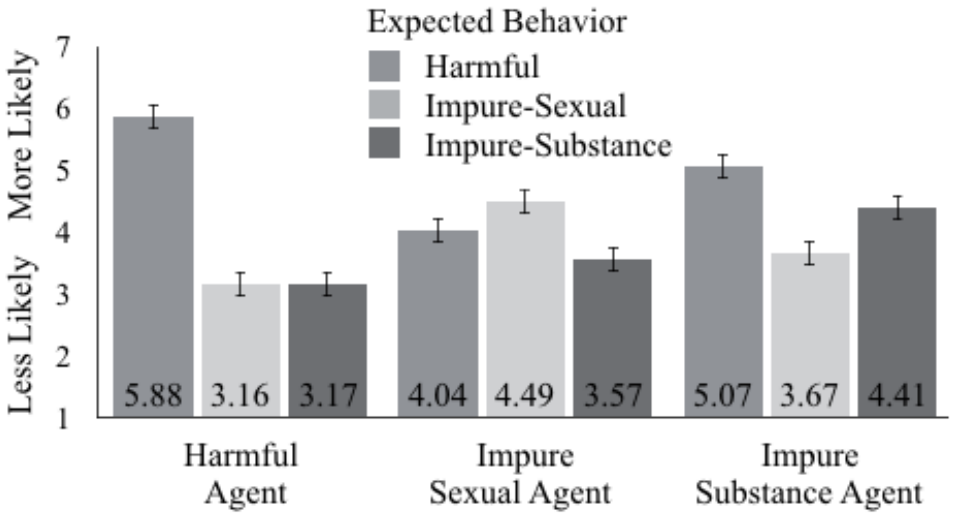

Figure S3. Study S1 expected behavior means by agent type (harmful, impure-sexual, impure-substance). Error bars ± 1 SE.

Harmful agents. The main effect of *expected behavior* was significant, $F(2, 104) = 146.89$, $p < .001$, $\eta^2_p = .74$. Post-hoc tests revealed that harmful agents were judged as more likely to commit harmful violations ($M = 5.88$, $SE = 0.18$) than impure-sexual violations ($M = 3.16$, $SE = 0.16$) or impure-substance violations ($M = 3.17$, $SE = 0.16$) ($ps < .001$, Bonferroni corrected), but were equally likely to perform the two kinds of impure violations ($p > .99$).

Impure-sexual agents. The main effect of *expected behavior* was significant, $F(2, 110) = 32.42$, $p < .001$, $\eta^2_p = .37$. Post-hoc tests revealed that impure-sexual agents were judged as more likely to commit impure-sexual violations ($M = 4.49$, $SE = 0.18$) than harmful violations ($M = 4.04$, $SE = 0.16$) ($p = .003$, Bonferroni corrected), and were judged as more likely to commit harmful violations and impure-sexual violations than impure-substance violations ($M = 3.57$, $SE = 0.17$) ($ps < .001$, Bonferroni corrected).

Impure-substance agents. The main effect of *expected behavior* was significant, $F(2, 105) = 49.75$, $p < .001$, $\eta^2_p = .49$. Post-hoc tests revealed that impure-substance agents were judged as more likely to commit harmful violations ($M = 5.07$, $SE = 0.18$) than impure-substance violations ($M = 4.41$, $SE = 0.20$), and more likely to commit harmful violations and impure-substance violations than impure-sexual violations ($M = 3.67$, $SE = 0.18$) (all ps < .001, Bonferroni corrected).

Behavior Expectation Asymmetry. Impure-sexual agents were judged more likely to be harmful ($M = 4.04$, $SE = 0.16$) than harmful agents were to commit impure-sexual violations ($M = 3.16$, $SE = 0.16$), $F(1, 216) = 15.46$, $p < .001$, $\eta^2_p = .07$. Impure-substance agents were judged more likely to be harmful ($M = 5.07$, $SE = 0.18$) than harmful agents were to commit impure-substance violations ($M = 3.17$, $SE = 0.16$), $F(1, 211) = 62.45$, $p < .001$, $\eta^2_p = .23$. Notably, there were no significant differences in the perceived likelihood that harmful agents would be impure, or that impure agents would commit novel kinds of impure violations, $F(2, 322) = 2.53$, $p = .08$, $\eta^2_p = .02$. These judgments highlight a violation likelihood "baseline" ($M = 3.47$, $SD = 1.75$), albeit one that does not depend on agents' past behaviors (e.g., harmful vs. impure).

Study S2. Agents Who Act on Themselves versus Others

When evaluating immoral acts, people are sensitive to both the kind of violation (e.g., harmful versus impure) and the target of the action (e.g., oneself versus another). For example, when collapsing across action domains, Chakroff et al. (2013) found self-directed acts to be judged as relatively more impure than harmful and other-directed acts as relatively more harmful than impure. Other-directed impure acts may be seen as particularly victimizing and harmful, while self-directed impure acts may not be judged as wrong because there is no victim beside the agent themselves (Chakroff et al., 2013). The primary aim of this study was to see if the expectation asymmetry would still emerge when comparing self-

directed impure agents with self-directed and other-directed harmful agents. Since self-directed impure actions do not involve harm to another we would not expect them to be judged very wrong, compared to other-directed harmful actions (Gray et al., 2012). Therefore, the expectation asymmetry should weaken when comparing between self- and other-directed agents.

Method

Participants. We recruited a new sample of 320 participants (35% female; M_{age} = 29.7) via Amazon Mechanical Turk. Five participants did not finish the survey and were excluded from analysis, leaving 315 participants in total.

Materials and procedures. In a 2 (*agent type*: harmful vs. impure-substance) x 2 (*target*: self vs. other) between-subjects design, participants read a sentence describing an individual committing either a harmful violation or an impure-substance violation, that was either self-directed or other-directed (e.g., "John cut himself [someone] with a sharp knife"), as in Chakroff et al. (2013) (see Appendix SD for stimuli). Participants then judged the likelihood that the agent would commit harmful violations (α = .96) and impure-substance violations (α = .95). Participants also rated the extent to which the act was wrong, weird, and their overall confidence in their likelihood judgments (see Supplemental Materials and Tables S1 and S3 for results for these measures). Judgments were collected in separate random blocks.

Results

A 2 (*agent type*) x 2 (*target*) x 2 (*expected behavior*) mixed-model ANOVA revealed main effects of agent type, $F(1, 311) = 34.94$, $p < .001$, $\eta^2_p = .10$, target, $F(1, 311) = 4.42$, $p = .04$, $\eta^2_p = .01$, and expected behavior, $F(1, 311) = 11.09$, $p = .001$, $\eta^2_p = .03$ (see Fig. S4 for means and standard errors). The two-way interaction between agent type and expected behavior domain was significant, $F(1, 311) = 91.14$, $p < .001$, $\eta^2_p = .23$, as in previous

studies. The interaction between target and expected behavior domain was also significant, $F(1, 311) = 69.12$, $p < .001$, $\eta^2_p = .18$, revealing that agents who acted on others were judged as more likely to be harmful than impure, while agents who acted on themselves were judged as more likely to be impure than harmful (see Fig. S4). The interaction between agent type and target was not significant, $F(1, 311) = 0.24$, $p = .63$, $\eta^2_p = .001$, nor was the three-way interaction, $F(1, 311) = 0.46$, $p = .50$, $\eta^2_p = .001$.

As in previous studies, we found an expectation asymmetry: the impure agent was judged more likely to be harmful than the harmful agent was judged to be impure, $F(1, 312) = 44.88$, $p < .001$, $\eta^2_p = .13$. Investigating the effect separately by target, the expectation asymmetry was found for agents who acted on others, $F(1, 146) = 68.63$, $p < .001$, $\eta^2_p = .32$. However, the asymmetry was only marginal for agents who acted on themselves, $F(1, 165) = 3.56$, $p = .06$, $\eta^2_p = .02$. The reduced asymmetry for self-directed agents is likely due in large part to the perception that self-directed harmful acts are in some sense impure, and that other-directed impure acts are in some sense harmful (Chakroff et al., 2013). Thus, the cleanest test of the expectation asymmetry in this study is to compare the expected harmful behavior of impure agents who acted on themselves (not seen as very harmful) to the expected impure behavior of agents who harmed others (not seen as very impure). Critically, this comparison revealed a significant expectation asymmetry, $F(1, 157) = 18.92$, $p < .001$, $\eta^2_p = .11$

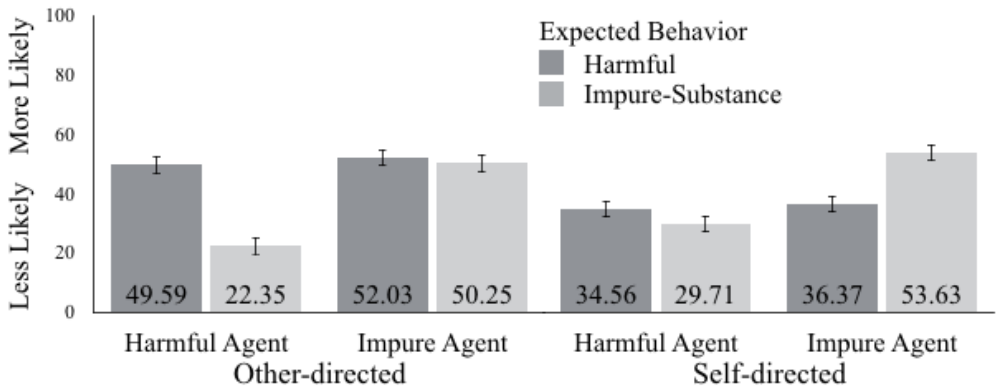

Figure S4. Study S2 expected behavior means by agent type (harmful versus impure) and target (other-directed versus self-directed). Error bars ± 1 S.E.

References for Supplements

Chakroff, A., Dungan, J., & Young, L. (2013). Harming ourselves and defiling others: What determines a moral domain? *PloS One, 8(9),* e74434.

Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry, 23(2),* 101-124.

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.

Table S1

Wrongness judgments for all studies: F-statistics, estimated effect sizes, means, and standard errors.

Study	***F***	η^2_p	**Harmful**	**Impure**
Pilot Study 1	4.78*	.06	6.40(.25)	5.63(.25)
Pilot Study 2	105.82**	.40	74.31(2.89)	32.17(2.91)
Study 1	15.88**	.06	6.33(.19)	5.21(.20)
Study 2	0.04	< .001	6.47(.33)	6.38(.32)
Study S2	0.01	< .001	43.19(2.23)	43.49(2.08)

	F	η^2_p	**Other-directed**	**Self-directed**
Study S2	145.46**	.32	61.72(2.22)	24.96(2.09)

Note. * $p < .01$, ** $p < .001$. In Pilot Study 2 and Study 1 wrongness judgments were a composite of 3 items (see Appendix SE).

Table S2

Proportion of internal vs. external cause judgments for Studies 1-2. P-values represent binomial test results.

Study	*p*	Agent Type	Internal	External
Study 1				
	.030	Harmful	.40	.60
	.001	Impure-Sexual	.72	.28
	.597	Impure-Substance	.54	.46
Study 2				
	.480	Harmful	.44	.56
	< .001	Impure-Sexual	.88	.12

Table S3

Study S2: Statistics for measures of weirdness and confidence.

Weirdness		
Expected Behavior 52.04**, η^2_p = .14	Harmful 62.73; *SE* = 1.89	Impure *M* = 81.31; *SE* = 1.75
Target 1.38*ns*, η^2_p = .004	Other-directed 73.54; *SE* = 1.86	Self-directed *M* = 70.51; *SE* = 1.77
Expected Behavior x Target 0.02*ns*, η^2_p < .001	Harm Self *M* = 61.39*; SE* = 2.59 Impure Self *M* = 79.62*; SE* = 2.40	Harm Other *M* = 64.07*; SE* = 2.74 Impure Other *M* = 83.00*; SE* = 2.56
Confidence		
Expected Behavior 0.06*ns*, η^2_p < .001	Harmful *M* = 38.6; *SE* = 2.27	Impure 37.86; *SE* = 2.11
Target 7.48*, η^2_p = .02	Other-directed *M* = 42.47; *SE* = 2.25	Self-directed *M* = 34.00; *SE* = 2.12
Expected Behavior x Target 1.60*ns*, η^2_p = .005	Harm Self *M* = 32.40*; SE* = 3.12 Impure Self *M* = 35.59; *SE* = 2.88	Harm Other *M* = 44.80*; SE* = 3.29 Impure Other *M* = 40.14; *SE* = 3.08

*Note. * p < .01; ** p < .001*

Appendix SA

Pilot Study 1 Scenarios

Harmful

Frank is at his local bar. He walks over to a man sitting at the bar and punches him off his bar stool.

Impure-Sexual

Frank has a sister. When nobody is around, Frank and his sister find a secret hiding place and kiss each other passionately on the mouth.

Pilot 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 a chicken or pork tenderloin.

Appendix SB

Pilot Study 2 Scenarios

Harmful

John once cut someone with a knife when he was upset.
Robert once closed a door on someone's fingers when he was upset.

Impure-Substance

John once found a dead dog on the road and cooked it up for dinner.
Robert once smeared cat poop on his face for fun.

Neutral

Ben took a slow, scenic drive for fun.
Steven once cooked up a pot roast for dinner.

Pilot Study 2 Expected Behaviors

Harmful

Kick someone in the shin.
Scratch someone on the arm.
Prick someone's hand with a needle.
Poke someone in the eye.
Burn someone on the arm.
Hit someone's finger with a hammer.
Cut someone's cheek with a razor.
Pinch someone hard on the arm.
Whip someone with a belt.

Impure-Substance

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

Appendix SC

Study 1: Replication Scenarios

Harmful

John cuts his brother with a sharp knife.
John slaps his brother's face.

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.

Appendix SD

Study S2 Scenarios

Other-directed Harmful

John punched someone in the ribs.
John cut someone with a sharp knife.
John closed a door on someone's fingers.
John poured a cup of painfully hot water on someone's lap.

Other-directed Impure-Substance

John poured a cup of urine on someone's lap.
John smeared cat poop on someone's arm.
John bought dog meat at a foreign grocery store and served it to someone.
John ordered stir-fried rat for someone at a small restaurant in the city.

Self-directed Harmful

John punched himself in the ribs.
John cut himself with a sharp knife.
John closed a door on his own fingers.
John poured a cup of painfully hot water on his own lap.

Self-directed Impure-Substance

John poured a cup of urine on his own lap.
John smeared cat poop on his own arm.
John bought dog meat at a foreign grocery store and ate it.
John ordered stir-fried rat for himself at a small restaurant in the city.

Appendix SE: Additional Measures

Pilot Study 1: Wrongness

I think Frank's behavior is… [1 = Not at all wrong; 4 = Moderately wrong; 7 = Very wrong]

Pilot Study 2: Wrongness

Were these actions morally blameworthy?
Were these actions deserving of punishment?
Were these actions immoral?

Study 1: *Wrongness*

John's behavior is wrong
John's behavior is immoral
John's behavior is deserving of condemnation

Studies 1-2: Internal / External Causes

Which is the more likely cause of this action? [Forced choice]
- Something internal to the person (e.g., personal desires, appetites)
- Something external to the person (e.g., circumstances, provocation)

Study S2: Disgust

Is this behavior gross?
Is this behavior physically disgusting?

Study S2: Weirdness

Is this behavior common in the general population? [Reverse scored]
Is this behavior abnormal or weird?

Study S2: Confidence

Is John's likelihood behavior easy to predict?
Are you confident in predicting John's behavior?