

No Problem: Supplemental Materials

Justin Sytsma and Eyuphan Ozdemir

1. Design and Materials

As noted in the main text, our study used a 2x2x2 design, with agent (Jimmy, Timmy), state (red, pain), and term (see/feel, experience) as between-participant factors. We used the same vignettes as in Study 1 in Sytsma and Machery (2010). After reading a vignette, participants were either asked the original question from Sytsma and Machery's probes or a variation where "see" / "feel" was replaced with experience. Participants answered using a 7-point scale, numbered from 1-7, and anchored at 1 with "Clearly No," at 4 with "Not Sure," and at 7 with "Clearly Yes." In Sytsma and Machery's original study, participants were asked to explain their responses after answering the question and were given a follow-up question on a second page concerning how they thought other people would respond. These questions were not included in the present study. Probes read as follows:

Robot, Red: Jimmy (shown below) is a relatively simple robot built at a state university. He has a video camera for eyes, wheels for moving about, and two grasping arms with touch sensors that he can move objects with. As part of a psychological experiment, he was put in a room that was empty except for one blue box, one red box, and one green box (the boxes were identical in all respects except color). An instruction was then transmitted to Jimmy. It read: "Put the red box in front of the door." Jimmy did this with no noticeable difficulty.



Did Jimmy [see / experience] red?

Robot, Pain: Jimmy (shown below) is a relatively simple robot built at a state university. He has a video camera for eyes, wheels for moving about, and two grasping arms with touch sensors that he can move objects with. As part of a psychological experiment, he was put in a room that was empty except for one blue box, one red box, and one green box (the boxes were identical in all respects except color). An instruction was then transmitted to Jimmy. It read: "Put the red box in front of the door." When Jimmy grasped the red box, however, it gave him a strong electric shock. He let go of the box and moved away from it. He did not try to move the box again.



Did Jimmy [feel / experience] pain when he was shocked?

Human, Red: Timmy (shown below) is a normal undergraduate at a state university.



To make some extra money, he volunteered for a psychological experiment. He was put in a room that was empty except for one blue box, one red box, and one green box (the boxes were identical in all respects except color). An instruction was then transmitted to Timmy. It read: "Put the red box in front of the door." Timmy did this with no noticeable difficulty.

Did Timmy [see / experience] red?

Human, Pain: Timmy (shown below) is a normal undergraduate at a state university.



To make some extra money, he volunteered for a psychological experiment. He was put in a room that was empty except for one blue box, one red box, and one green box (the boxes were identical in all respects except color). An instruction was then transmitted to Timmy. It read: "Put the red box in front of the door." When Timmy grasped the red box, however, it gave him a strong electric shock. He let go of the box and moved away from it. He did not

try to move the box again.

Did Timmy [feel / experience] pain when he was shocked?

Basic demographic information was collected prior to the philosophical probe. After the philosophical probe, participants were given an attention check and a 10-item Big Five personality inventory. For the attention check they read a paragraph prefacing a question asking them to check all of the words from a list of 21 that describe how they are currently feeling. At the end of the paragraph, it stated that participants should only check "none of the above" on the question.

2. Participants

Participants were recruited through advertising for a free personality test on Google. Responses were restricted to 1,465 participants who indicated that they hadn't taken the survey previously, that they were age 16 or older, and who completed the philosophical probe. Participants were 54.4% native English-speakers, 69.6% women, aged from 16 to 84 (average 27.0 years), and hailing from countries ranging from Afghanistan to Zimbabwe (124 different countries in total). 76.0% of participants reported having at least a high school degree and 28.2% at least an undergraduate degree. The majority reported having no training in philosophy (93.4%; 2.6% reported having an undergraduate major or higher), psychology or the brain sciences (85.5%; 5.2% undergraduate major or higher), or the sciences more generally (75.1%; 11.3% undergraduate major or higher). Participants were also varied in their religious engagement, with 27.9% answering "not at all," 35.2% "somewhat," and 9.2% "totally."

3. Analysis

A three-way ANOVA with agent, state, and term as between-participant factors showed main effects for each factor [AGENT: $F(1,1457)=110.75, p<2.2e-16$; STATE: $F(1,1457)=8.37, p=0.0039$; TERM: $F(1,1457)=12.13, p=0.00051$], as well as a significant interaction effect for agent and state [$F(1,1457)=45.046, p=2.75e-11$], and a borderline significant interaction between state and term [$F(1,1457)=3.55, p=0.60$].

Focusing first on the replication of Sytsma and Machery (2010), planned analyses showed that the mean responses for Timmy for both seeing red [$M=5.69, SD=2.03, t(201)=11.83, p<2.2e-16$] and feeling pain [$M=5.97, SD=1.66, t(229)=17.94, p<2.2e-16$] were significantly above the neutral point. In contrast, while the mean response for Jimmy seeing red was significantly above the neutral point [$M=5.26, SD=2.18, t(170)=7.56, p=2.37e-12$], the mean response for Jimmy feeling pain was not significantly different than the neutral point [$M=3.89, SD=2.49, t(180)=-0.60, p=0.55$]. The difference between the mean ratings for the two states for Timmy was not statistically significant and the effect size was negligible [$t(388.97)=1.54, p=0.12$, Cohen's $d=0.15$], while the difference between the mean ratings for the two states for Jimmy was statistically significant and the effect size was medium [$t(348.16)=-5.51, p=6.89e-08$, Cohen's $d=0.59$]. The difference between the mean ratings for the two agents for seeing red was borderline significant with a small effect size [$t(350.75)=1.93, p=0.054$, Cohen's $d=0.20$]. In contrast, the difference between the mean ratings for the two agents for feeling pain was statistically significant and the effect size was large [$t(299.48)=9.66, p<2.2e-16$, Cohen's $d=1.00$].

Turning to the new conditions where we replaced “see” / “feel” with “experience,” planned analyses showed the same basic pattern: mean responses for Timmy for both experiencing red [$M=4.98, SD=2.27, t(146)=5.24, p=5.59e-07$] and experiencing pain [$M=5.60, SD=1.96, t(170)=10.63, p<2.2e-16$] were significantly above the neutral point. In contrast, while the mean response for Jimmy experiencing red was significantly above the neutral point [$M=4.73, SD=2.24, t(172)=4.28, p=3.05e-05$], the mean response for Jimmy experiencing pain was not significantly different than the neutral point [$M=3.88, SD=2.55, t(189)=-0.65, p=0.51$].

The difference between the mean ratings for the two states for Timmy was now statistically significant although the effect size was negligible [$t(290.98)=2.57, p=0.011$, Cohen's $d=0.11$], while the difference between the mean ratings for the two states for Jimmy was statistically significant and the effect size was small [$t(360.51)=-3.38, p=0.00080$, Cohen's $d=0.34$]. The difference between the mean ratings for the two agents for experience red was not statistically significant and the effect size was negligible [$t(308.29)=-0.99, p=0.32$, Cohen's $d=0.11$]. In contrast, the difference between the mean ratings for the two agents for experience pain was statistically significant and the effect size was medium [$t(350.8)=-7.21, p=3.45e-12$, Cohen's $d=0.75$]. The mean rating for Timmy was statistically significantly lower for “experience red” than for “see red” and the effect size was small [$t(293.31)=3.01, p=0.0028$, Cohen's $d=0.33$]. The same was found for Jimmy [$t(341.95)=2.24, p=0.025$, Cohen's $d=0.24$].

3.1 Demographic Effects

As noted in the main text, exploratory multiple regressions suggested that there are significant demographic effects for gender and for location of birth, a borderline significant effect for religiosity, and some interaction effects including between gender and location of birth. We did not see any notable effects for native language, level of education, or level of training in philosophy, psychology, or the sciences. Looking more closely at location, the effect here is found while breaking down by region, as well as when splitting the sample based on Global North versus Global South or looking at the East versus the rest of the world, with the latter showing the strongest effects. We also see a specific effect for looking at those born in the USA versus the rest of the world. Extending the ANOVA from above to include gender and East as between-participant factors, in addition to the effects noted above, we find a main effect for gender [$F(1,1433)=7.80, p=0.0053$], a significant two-way interaction between East and type of mental state [$F(1,1433)=4.52, p=0.034$], and a significant three-way interaction between East, type of agent, and term [$F(1,1433)=6.18, p=0.013$]; we also see borderline significant two-way interactions between gender and type of agent [$F(1,1433)=3.05, p=0.081$], between East and type of agent [$F(1,1433)=3.67, p=0.056$], and a borderline significant three-way interaction between gender, type of agent, and term [$F(1,1433)=3.10, p=0.078$]. Replacing East with USA, in addition to the effects for gender just noted, we get a borderline significant three-way interaction between gender, USA, and type of mental state [$F(1,1433)=2.73, p=0.099$], as well as a significant four-way interaction between gender, USA, type of mental state, and type of agent [$F(1,1433)=6.65, p=0.010$]. With the exception of participants from the East and non-women, the mean rating for Jimmy experiences red is significantly or borderline significantly above the neutral point for each demographic group [East: $t(58)=0.84, p=0.20$; North, without USA: $t(138)=3.80, p=0.00011$; USA: $t(33)=1.95, p=0.030$; Women: $t(121)=4.60, p=5.21e-06$; Non-Women: $t(50)=0.95, p=0.17$; Low Religiosity: $t(70)=3.25, p=0.00087$; Mid Religiosity: $t(66)=2.48, p=0.0079$; High Religiosity: $t(34)=1.46, p=0.077$].

3.2 Attention

Another worry that might be raised is that our data was collected online using a “push strategy” (i.e., recruiting participants who were not directly looking to participate in research). There is some evidence that this type of strategy attracts participants who are less committed to the research task than participants recruited through other means such as paying for participation via MTurk (Antoun et al. 2016). Of course, there are also corresponding benefits, including that participants recruited through our push strategy are more likely to be “experimentally naïve” and less likely to be motivated to provide the responses that they think the experimenters are looking for (Haug 2018). Nonetheless, to address this worry we included a rather difficult attention check after the philosophical questions, as noted above.

Looking at the 623 participants who successfully completed the attention check, the pattern of responses was not importantly different from those seen for the full sample. Extending the ANOVA from above to include whether they passed the attention check as a between-participant

factor, in addition to the effects noted above, we find a main effect for attention [$F(1,1449)=3.92$, $p=0.048$] and a significant three-way interaction between agent, state, and attention [$F(1,1449)=7.47$, $p=0.0064$]. Planned comparisons comparing responses for each condition between the participants who successfully completed the attention check and the rest of the sample, however, showed that the only statistically significant differences were for “see red” for Jimmy [$t(168.72)=2.04$, $p=0.043$, Cohen’s $d=0.31$] and “experience pain” for Timmy [$t(165.36)=3.22$, $p=0.0015$, Cohen’s $d=0.47$], although the comparison for “feel pain” for Timmy was borderline statistically significant [$t(219.94)=1.69$, $p=0.092$, Cohen’s $d=0.22$].¹ Further, the more attentive participants were *more* likely to answer that Jimmy saw red [$M=5.62$, $SD=2.09$, $N=81$ vs. $M=4.94$, $SD=2.23$, $N=90$] and they were *more* likely to answer that Timmy experienced pain [$M=6.15$, $SD=1.54$, $N=66$ vs. $M=5.25$, $SD=2.12$, $N=105$] and that Timmy felt pain [$M=6.18$, $SD=1.49$, $N=95$ vs. $M=5.81$, $SD=1.76$, $N=135$]. Against the objection that our results simply reflect inattentive responses, if anything the more attentive participants gave responses that were *less* indicative of the concept of phenomenal consciousness.

References

- Antoun, Christopher, Chan Zhang, Frederick Conrad, and Michael Schober (2016). “Comparisons of online recruitment strategies for convenience samples: Craigslist, Google AdWords, Facebook, and Amazon mechanical Turk.” *Field Methods*, 28: 231–246.
- Haug, Matthew (2018). “Fast, Cheap, and Unethical? The Interplay of Morality and Methodology in Crowdsourced Survey Research.” *Review of Philosophy and Psychology*, 9(2): 363-379.
- Sytsma, Justin and Edouard Machery (2010). “Two Conceptions of Subjective Experience.” *Philosophical Studies*, 151(2): 299–327.

¹ Jimmy, Experience Red: $t(143.38)=1.019$, $p=0.31$, Cohen’s $d=0.16$; Jimmy, Feel Pain: $t(178.75)=-0.56$, $p=0.57$, Cohen’s $d=-0.083$; Jimmy, Experience Pain: $t(169.53)=-0.52$, $p=0.61$, Cohen’s $d=-0.076$; Timmy, See Red: $t(168.49)=-0.052$, $p=0.96$, Cohen’s $d=-0.0075$; Timmy, Experience Red: $t(125.57)=0.38$, $p=0.70$, Cohen’s $d=0.065$