Reasons and Basing in Commonsense Epistemology: Evidence from Two Experiments *

John Turri   *University of Waterloo*

I accomplish two things in this paper. I explain the motivation for including experimental research in philosophical projects on epistemic reasons and the basing relation. And I present the first experimental contributions to these projects. The results from two experiments advance our understanding of the ordinary concepts of reasons and basing and set the stage for further research on the topics. More specifically, the results support a causal theory of the basing relation, according to which reasons are causes, and a dualist theory of epistemic reasons, according to which reasons include both psychological and non-psychological items.

Keywords: social cognition, theory of mind, reasons, causation, epistemology

Introduction

Images and methods

The most important development in philosophy this century has been the rapidly growing dissatisfaction with a primarily armchair methodology, and a corresponding growth in the use of empirical evidence and science to address philosophical questions. To appreciate some of the motivation for this change, consider Wilfrid Sellars' evocative and rightly famous distinction between the *manifest* and *scientific* images. The manifest image consists of our ordinary ways of understanding the world and our place in it, our "pre-reflective orientation and common heritage" that "took shape in the mists of prehistory" (Sellars, 1963, pp. 3, 5). The scientific image consists of the deliverances of organized scientific inquiry, which can diverge radically from commonsense. Sellars proposed that contemporary philosophy's distinctive objective is to unify these different images into a coherent whole, to transcend this "duality" of images. While it is immediately obvious that getting straight on the scientific image requires empirical evidence and, well, science, it turns out that the same is true for the manifest image. For it is an "essentially social" phenomenon, marked by "common standards of correctness and relevance," that "transcends" individual thinkers and has "an objective existence . . . in human thought generally" (Sellars, 1963, p. 14). Objective social phenomena are objects of scientific investigation in their own right.

Setting aside the wonderful metaphor, which I find deeply fascinating and insightful, the basic idea can be stated as follows. There is what commonsense says the world is like, and then is what our best evidence says the world is like. Empirical science provides our best evidence

about what the world is like, so this side of the ledger must involve empirical scientific evidence. It is tempting to think that the other side of the ledger needn't involve empirical scientific evidence. After all, if it's *common*sense, then we all share it and one could successfully characterize it through careful and systematic reflection, perhaps supplemented by comparing notes occasionally with others engaged in reflection of similarly high quality. Sellars gives voice to this temptation when he claims that many philosophical systems have "more or less adequate[ly]" characterized commonsense, and that twentieth-century anglophone philosophy has captured it "in something like its pure form" (Sellars, 1963, p. 15). It is true that social observation and reflection provide insight into our shared standards and practices. Life as we know it would be impossible otherwise. But the insight afforded, albeit impressive, is far from complete, and systematic scientific inquiry is needed to begin filling in the gaps, as Sellars went on to acknowledge (Sellars, 1963, p. 15).

Generations of social and cognitive scientists have made progress on this front, including work ranging from intuitive physics (e.g. McCloskey, Washburn and Felch, 1983), to folk biology (e.g. Berlin, Breedlove and Raven, 1973), to commonsense psychology (e.g. Kelley, 1992). They have also begun the daunting task of reconciling commonsense and the deliverances of science in ways that are meaningful and useful to us. They do this by identifying the limitations of our pretheoretical understanding and developing techniques to integrate hard-won theoretical knowledge to improve judgment and decision-making (e.g. Schtulman, 2017; Gigerenzer, 2014; Kosko, 1993). For example, people can learn to respect base rates and remedy their ignorance about how diseases are actually transmitted. Decades into these challenging pursuits, there is much left to learn. But if Sellars is correct that reconciling science and commonsense is philosophy's distinctive task, then it turns out that these impressive philosophical accomplishments are due to scientists.

Another motivation for the increasing role of scientific methods in philosophy is less inspiring, what we might call *the decades long false start*. Philosophers have made numerous false claims about ordinary thought and speech. Of course, researchers in every field make mistakes, so in itself, this is unremarkable. What makes these mistakes remarkable, regardless of the field they occur in, is that they follow a pattern that teaches a valuable lesson to those willing to learn (for reviews, see Turri (2016), Buckwalter and Turri (2018), Turri (2019)). The pattern begins with one philosopher making an empirical claim, based on introspection or anecdotal social observation, about how we ordinarily speak or classify objects. The philosopher relies on this claim to motivate or defend a theory of some ordinary concept of philosophical interest. Other philosophers accept the claim but dispute its theoretical significance and propose alternative explanations. An entire literature grows up around the dispute, which enters the annals via textbooks, handbooks, and encyclopedia entries. A central tendency in this literature is to treat the original empirical claim as a datum to be explained, or, at worst, explained away. Many years later, someone decides to go back and check whether the original empirical claim is actually true. A few simple experiments provide clear evidence that it isn't. Thus the debate was predicated on a false assumption and all the time and energy devoted to it would have been better spent theorizing about actual facts. In addition to opportunity cost for the field as a whole, let's not lose sight of the human element.

This sort of outcome can be especially disappointing and frustrating for people invested in the debate. For many of us, it's not easy to let go of something that has been a source of personal satisfaction and professional success. This can create additional obstacles to advancing collective understanding of the underlying issues.

We should do our best to avoid unforced errors like this, and normalizing experimentation within the discipline has helped. Again, philosophy is not special in this regard. Researchers in all fields should be quick to rely, when possible, on empirical evidence and controlled experiments to separate wheat from chaff, especially early in a research program when formative clues can put researchers on more promising paths.

Reasons and basing

Epistemic reasons and the basing relation are a closely related pair of philosophical topics that have not, as far as I'm aware, been directly subjected to empirical scrutiny. My goal in the present paper is to begin filling that gap.

With respect to the basing relation, I will focus on the lead premise in the principal argument for a causal theory of the basing relation. According to this theory, to believe something for a reason is for the reason to non-deviantly cause the belief (Turri, 2011). The principal argument for this view begins from the premise that, intuitively, reasons are "difference-makers". Simplifying, the argument then proceeds to claim that the best explanation for why reasons are difference-makers is that reasons are causes. The argument relies on intuition or verdicts about cases at more than one juncture, all of which could be worth exploring. My focus here will be the most important juncture, which is whether reasons are intuitively viewed as difference-makers. If they aren't, then that poses a serious problem for the argument. I am not aware of existing findings on the concepts of "difference-making," reasons, or inference that would lead us to expect one outcome or another.

With respect to epistemic reasons, I will focus on three competing theories about *the kind of thing reasons are* (Turri, 2009). One view is that all reasons are psychological items of the agent, such as beliefs, knowledge, emotions, and perceptual and sensory states. A second view is that all reasons are non-psychological items, such as physical objects, external facts, or states of affairs. A third view is that reasons include both psychological and non-psychological items. Call these views *psychologism*, *anti-psychologism*, and *dualism*, respectively.

The principal case for psychologism and against anti-psychologism and dualism relies on the premise that being given a description of the agent's mental condition "is sufficient for us to understand" his reasons, in which case "it is superfluous" to identify additional, non-psychological items as reasons (Turri, 2009, p. 504). This appears closely related to the following assumption about radically deceived agents:

> Ideally a theory should respect the following intuition. Your victimized twin, Vic, recently ensnared by an evil genius, undergoes an indefinite succession of experiences

indistinguishable from yours. Responding to the very same kinds of sensory experiences as you, Vic believes that he or she is reading an epistemology paper. Intuitively you and your twin believe what you do *for the same reasons*. (Turri, 2009, p. 493).

With respect to this assumption, I believe existing findings should lead us to expect that it is false. In particular, research on the mistaken "extramission" folk theory of perception counts against it (Cottrell and A, 1994; Winer et al., 1996, 2002). According to this theory, when people perceive, it is because a force emanates from their sense organs and interacts with the perceived object. For example, one sees an object because rays emanate from one's eyes and strike the object. Vision science shows that this is exactly backwards: one sees because rays emanating from the object strike one's eye. The naive extramission theory is present from early in childhood, persists into adulthood, shows up on a variety of measures, and is surprisingly resistant to explicit instruction. One psychologist characterizes the ordinary concept of perception in a way consistent with the extramission theory: "Perceiving is experienced as a direct contact with the environment; it is a means whereby objective facts enter the life space" (Heider, 1958, p. 15). In light of existing evidence, then, we should expect that psychologism is false as a theory of the ordinary reason concept. The remaining question is whether a direct test of the matter will favor anti-psychologism or dualism.

Experiment 1 focuses on reasons and difference-making. Experiment 2 focuses on the ontology of reasons. To preview the main findings, the results of experiment 1 provide initial support for the view that reasons are difference-makers. The results of experiment 2 provide strong initial support for a dualist theory of reasons.

General methods

The following statements are true of all studies reported here. I report all manipulations and measures used. All participants were adult residents of the United States. No participants were excluded from analysis. I recruited and tested people using an online platform of Amazon Mechanical Turk (https://www.mturk.com), TurkPrime (Litman, Robinson and Abberbock, 2017), and Qualtrics (https://www.qualtrics.com). Participants completed a brief demographic questionnaire after testing. I used R 3.5.1 for all analyses (R Core Team, 2018).

Experiment 1: difference making

Method

No research on the topic existed to inform an *a priori* power analysis regarding sample size, so I defaulted to the conventional "rule of thumb" of 30 participants per condition, plus a few extra as a precaution against participants dropping out or refusing to sign the consent form after initially clicking to participate.

Participants

Participants' mean age was 35.29 years (range = 19-76, sd = 10.69), and 56% (172 of 305) were female.

Materials and procedure

Participants were randomly assigned to one of ten conditions in a 2 (Status: none, some) × 5 (Scenario) mixed experimental design. Participants first read a brief scenario about an agent considering a question, then responded to two test statements (within-subjects). The Status factor manipulated whether a potential basis (reason) was described as making no difference to the agent's thinking (none), or as making a difference (some). The Scenario factor manipulated what question the agent was considering, and other features of her situation. I included this factor to support generalization of the results beyond the specific stimuli studied here (Judd, Westfall and Kenny, 2012; Clark, 1973; Baayen, Davidson and Bates, 2008). All stimuli used in this study are included in an appendix. To give readers a sense of the materials, I include one scenario and the test statements here, with the Status (none/some) manipulation shown in brackets.

> (Juror) Lauren is on a jury hearing a criminal case. After listening carefully throughout the trial, Lauren concludes that the defendant was innocent. A grainy photograph seemed to place the defendant a mile away from the scene of the crime. This photograph made [absolutely no/a] difference to Lauren's thinking on the matter.
>
> 1. Lauren's conclusion was partly based on the photograph. (basing)
> 2. The photograph made a difference to Lauren's thinking. (difference)

The basing statement appeared below the scenario on the first screen of the study. Participants then advanced to a new screen (they could not go back) and responded to the difference statement from memory. Responses were collected on a standard 7-point Likert scale, 1 ("strongly disagree") - 7 ("strongly agree"), arranged left-to-right on the participant's screen.

Data analysis

One principal research question is whether Status would affect response to the basing statement. If it doesn't, then the results would undermine a primary argument for a causal theory of the basing relation. If it does, then the results would support the causal theory. Relatedly, it will be informative to discern whether such a Status effect is robust across the scenarios and the demographic variables of sex and age. A second research question pertains to the relationship between the basing and difference statements. To the extent that participants respond differently to them, it puts pressure on the causal theory of the basing relation. To the extent that participants respond similarly to them, it supports the causal theory. To address these questions, I conducted a linear mixed effects analysis on participant response and followed up with appropriate correlation analysis and t-tests.

Results

I conducted a linear mixed effects analysis of participant response. I included as fixed effects Status (none, some), the type of judgment (basing, difference), an interaction between Status and type of judgment, and participant age and sex. I included random intercepts for Scenario (five variants) and participant nested within Scenario. Step-wise elimination of non-significant effects (Kuznetsova, Brockhoff and Christensen, 2017) resulted in a reduced model including only Status and participant, which did not differ significantly from the full model, $\chi^2(5) = 3.27$, p = .659. The reduced model explained 92% of variance in participant response (Nakagawa and Schielzeth, 2012). The difference between Status conditions explained 54%, and the difference among participants explained 38%.

(A) Means of Difference-making and Basing Judgments

(B) Correlation of Difference-making and Basing Judgments
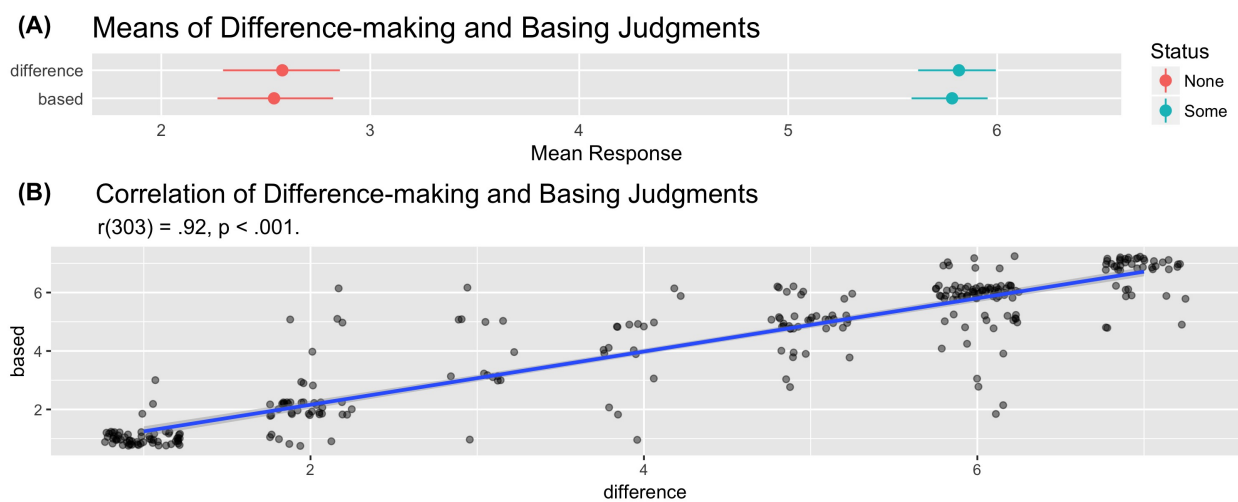
r(303) = .92, p < .001.

Figure 1: Experiment 1. (A) Mean response to the basing and difference-making statements (within-subjects) across the two Status conditions (none, some) (between-subjects). Scales ran 1 ("strongly disagree") – 7 ("strongly agree"). Error bars show 95% bootstrapped confidence intervals. (B) Scatterplot showing the correlation between difference-making and basing judgments. Points are jittered to avoid overplotting.

A bivariate correlation analysis on the basing and difference statements found that they were extremely strongly correlated, r(303) = 0.92 [0.9, 0.93], p < .001. (See Figure 1.) A paired samples t-test on the two statements detected no difference, t(304) = -0.7, MD = -0.04 [-0.14, 0.07], p = .483. Independent samples t-tests revealed that mean response to the basing statement was significantly lower in the none condition (M = 2.54, SD = 1.75) than in the some condition (M = 5.78, SD = 1.18), t(264.79) = -18.98, MD = -3.24 [-3.58, -2.91], p < .001, d = -2.18. Similarly, mean response to the difference statement was significantly lower in the none condition (M = 2.58, SD = 1.8) than in the some condition (M = 5.82, SD = 1.17), t(259.04) = -18.58, MD = -3.24 [-3.58, -2.89], p < .001, d = -2.18. The median response to both statements was "agree" (=6) in the some condition and "disagree" (=2) in the none condition; the modal response to both statements was "agree" in the some condition and "strongly disagree" (=1) in the none condition.

Discussion

This experiment began looking at the relationship between judgments about what *makes a difference* to whether an agent draws a conclusion, and the reasons an agent's conclusion *is based on*. The two judgments were strongly correlated and statistically indistinguishable by their mean, median, and mode. This same pattern was robust across a range of stimuli and the demographic variables of biological age and sex. The results support the principal argument for a causal theory of the basing relation, which begins from the premise that reasons are difference-makers.

Experiment 2: reason comprehension and attribution

Method

No research on the topic existed to inform an *a priori* power analysis regarding sample size. Because this experiment used less sensitive measures than experiment 1, I recruited more participants per condition, plus a few extra for the same reason described in experiment 1.

Participants

Participants' mean age was 36.58 years (range = 19-69, sd = 11.85), and 55% (101 of 184) were female.

Materials and procedure

Participants were randomly assigned to one of three conditions (Normal, False, Brain) in a mixed experimental design.[1] Participants first read a brief scenario about an agent who draws a conclusion, then they rated whether they had enough information to understand the agent's reasons, and finally they rated whether ten items (within-subjects) were among the agent's reasons for drawing the conclusion. The scenarios were closely adapted from previous research on belief evaluation (Turri, 2015). In all conditions, the agent, Victor, concludes that there is a fox nearby. The conditions differed in whether Victor is a normal human whose conclusion is true (Normal), a normal human whose conclusion is false (False), or a human brain-in-a-vat whose conclusion is false (Brain).

> (Normal/False) Victor is a healthy human adult sitting on his patio in a fine neighborhood. Victor is currently enjoying a variety of perfectly vivid sensory experiences, thanks to a team of scientists who helped save his life with a supercomputer that detected a heart condition. Victor was unaware that scientists could do that, just as he was unaware that he had a bad heart condition in the first place. Everything seems perfectly normal to him now. The scientists monitor him regularly.¶[2] As Victor sits

---

[1] I capitalize the names of conditions in this experiment in order to avoid confusion between the names of some experimental conditions ("Normal" and "Brain") and dependent measures ("normal" and "brain").

[2] Indicates a paragraph break on the participant's screen.

there on his patio, it seems like a reddish four-legged animal with pointy ears and a bushy tail is walking through a nearby flowerbed. Victor's experiences seem entirely natural and normal. He concludes that it is a fox. [And/But] things [are/are not] exactly as they seem to Victor: as he sits there on his patio, he [is/is not] looking at a fox.

(Brain) Victor is a healthy human brain sitting in a vat of fluid in a fine laboratory. Victor is currently enjoying a variety of perfectly vivid sensory experiences, thanks to a team of scientists creating them through a supercomputer that is hooked up to Victor. Victor was unaware that scientists could do that, just as he was unaware that his body died and he was put in a vat in the first place. Everything seems perfectly normal to him now. The scientists monitor him regularly.¶ As Victor sits there in his vat, it seems like a reddish four-legged animal with pointy ears and a bushy tail is walking through a nearby flowerbed. Victor's experiences seem entirely natural and normal. He concludes that it is a fox. But things are not as they seem to Victor: as he sits there in his vat, he is not looking at a fox.

Participants were then asked to rate the agreement or disagreement with this statement:

Based on the scenario described above, I have enough information to understand Victor's reason(s) for concluding that it is a fox. (comprehension)

Responses were collected on a standard 7-point Likert scale, 1 ("strongly disagree") - 7 ("strongly agree"), arranged left-to-right on the participant's screen. Participants then went to a new screen and, while the story remained visible atop the screen, read these instructions,

Victor concluded that it is a fox. What are Victor's reasons for drawing that conclusion? For each item, please rate whether it is ("Yes") or isn't ("No") one of Victor's reasons for concluding that it is a fox.

Below the instructions was a matrix table with ten items arranged vertically (order randomly rotated). To the right were three possible response options for each item, in the following order, left-to-right on the participant's screen: "Yes", "Unclear", "No". Participants selected their response by clicking a radio button. These were the ten items (all presented within-subjects):

1. Sensory experiences (sense)
2. The situation seems normal (normal)
3. Background knowledge of what foxes looks like (background)
4. The scientists' activities (scientists)
5. The supercomputer's operations (computer)
6. The physical animal itself (animal)
7. Consciousness (conscious)

8. Brain chemistry (brain)
9. It seems like there is a fox (seems)
10. The fact that there is a fox (fact)

Data analysis

One principal research question is whether participants would agree that they had enough information to comprehend Victor's reasons, and whether they would do so at similar rates across conditions. A second question is whether assignment to condition would affect which reasons participants attributed to Victor. To address these questions, I conducted an analysis of variance on response to the comprehension statement, followed up by appropriate t-tests. I then conducted a linear mixed effects analysis on reason attributions from the list of items, followed up by appropriate analyses of variance.

Results

I conducted an analysis of variance on response to the comprehension statement, with assignment to condition as the independent variable. This revealed a significant difference across conditions, $F(2, 181) = 9.65$, $p < .001$, $\eta^2 = 0.096$. (See Figure 2.) Follow-up independent samples t-tests showed that mean respone was higher in the Normal condition ($M = 5.82$, $SD = 1.02$) than in the False condition ($M = 4.64$, $SD = 1.71$), $MD = 1.18$ [0.68, 1.69], $t(97.3) = 4.65$, $p < .001$, $d = 0.84$. Mean response was higher in the Normal condition than in the Brain condition ($M = 5.15$, $SD = 1.67$), $MD = 0.68$ [0.18, 1.17], $t(98.75) = 2.7$, $p = .008$, $d = 0.49$. Mean response was nonsignificantly higher in the Brain condition than in the False condition, $MD = 0.51$ [-0.1, 1.11], $t(119.93) = 1.66$, $p = 0.1$, $d = 0.3$. Despite the differences, one sample t-tests showed that mean response was above the neutral midpoint (= 4) in all three conditions: Normal, $t(61) = 14.12$, $p < .001$, $d = 1.79$; Brain, $t(60) = 5.36$, $p < .001$, $d = 0.69$; False, $t(60) = 2.91$, $p = .005$, $d = 0.37$. Median response "agree" (= 6) in the Normal and Brain conditions and "slightly agree" (= 5) in the False condition. Modal response was "agree" in all three conditions.

   I conducted a linear mixed effects analysis on reason attributions. (See Figure 2.) I included as fixed effects Condition, item (i.e. the various items that could be attributed as reasons), an interaction between Condition and item, and participant sex and age. I also included a random intercept for participant. Step-wise elimination of non-significant effects resulted in a reduced model that excluded sex and age but included Condition, item, a Condition by item interaction, and participant. This reduced model did not differ significantly from the full model, $\chi^2(2) = 0.41$, $p = .815$. The reduced model explained 34% of variance in reason attributions (Nakagawa and Schielzeth, 2012). The fixed effects explained 29%, and the difference among participants explained 5%.
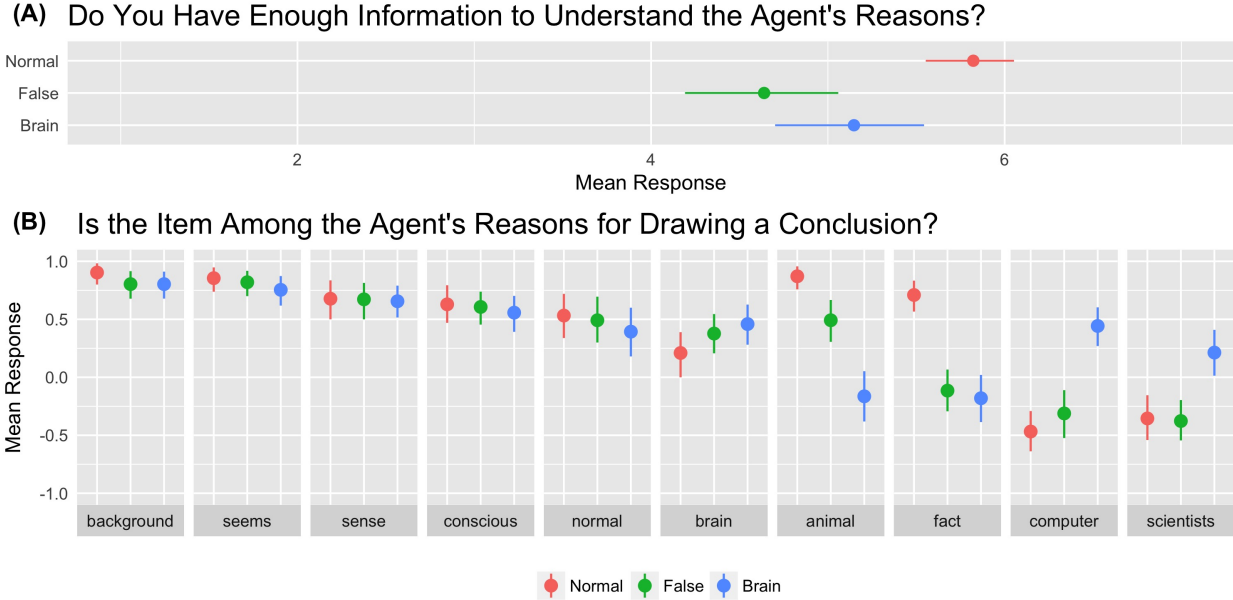
Figure 2: Experiment 2. (A) Mean agreement that participants had enough information to understand the agent's reasons. Scales ran 1 ("strongly disagree") – 7 ("strongly agree"). (B) Mean agreement that the item was among the agent's reasons for drawing a conclusion. Options were "Yes" (= 1), "Unclear" (= 0), and "No" (= -1). Error bars show 95% bootstrapped confidence intervals.

Table 1: Experiment 2. Analysis of variance for attribution of reasons across the three conditions.

Item	Df1	Df2	Sum Sq	F value	p value	eta^2
background	2	181	0.411	1.137	.323	0.012
seems	2	181	0.321	0.806	.448	0.009
sense	2	181	0.016	0.021	.980	0.000
conscious	2	181	0.165	0.207	.813	0.002
normal	2	181	0.625	0.518	.597	0.006
brain	2	181	1.990	1.962	.143	0.021
animal	2	181	33.665	36.896	< .001	0.290
fact	2	181	30.339	30.512	< .001	0.252
computer	2	181	29.037	26.934	< .001	0.229
scientists	2	181	13.684	11.387	< .001	0.112

Of particular interest is the interaction between Condition and item. To follow up on this, I conducted an analysis of variance for each item with Condition as the independent variable. (See Table 1.) These analyses found no difference across the three conditions for these items: "background knowledge of what foxes look like," "it seems like a fox," "sensory experiences," "consciousness," "the situation seems normal," and "brain chemistry." All of those items are, broadly speaking, internal to the agent. By contrast, the analyses revealed significant variation across the three conditions for these items: "the physical animal itself," "the fact that there is a fox," "the supercomputer's operations," and "the scientists' activities." All of those items are external to the agent.

Inspection of Figure 2 reveals two interesting patterns associated with the interaction observed in the mixed model and the follow-up analyses of variance. First, in the Normal condition, mean agreement was high that *the physical animal* and *the fact that there is a fox* were among Victor's reasons. By contrast, in the False condition, mean agreement was high for *the physical animal* but not for *the fact that there is a fox*. And in the Brain condition, mean agreement was low for both of those items. Second, participants tended to disagree that *the supercomputer's operations* and *the scientists' activities* were among Victor's reasons in the Normal and False conditions, whereas they tended to agree in the Brain condition.

Discussion

This experiment began investigating judgments about an agent's reasons for drawing a conclusion. Across three conditions, participants agreed that they had enough information to comprehend the agent's reasons, as reflected in mean, median, and modal judgments. Participants also rated whether specific items were among the agent's reasons. The items most strongly and consistently rated as reasons were all psychological items, such as sensory experiences. But there were critical differences across conditions when it came to non-psychological objects and facts. Crucially, in some conditions, participants included these non-psychological items among the agent's reasons. Overall, these results provide strong initial support for a dualist theory of the ordinary reason concept, according to which reasons include both psychological and non-psychological items.

General Discussion

In this paper, I began the project of experimentally investigating the ordinary concepts of *reasons* and *basing*. In experiment 1, I studied the relationship between judgments about what *makes a difference* to whether an agent draws a conclusion, and the reasons an agent's conclusion *is based on*. I found that the two judgments were strongly correlated and statistically indistinguishable by their mean, median, and mode. This pattern persisted across a range of stimuli and the demographic variables of biological sex and age. The results are relevant to evaluating the principal argument for a causal theory of the basing relation. This argument assumes that reasons are in-

tuitively difference-makers. The present findings provide initial support for this critical assumption. However, the results do not prove that the assumption is true. Further work is required to test whether, and to what extent, judgments about basing can be pried apart from judgments of difference-making.

In experiment 2, I studied judgments about an agent's reasons for drawing a conclusion. On the one hand, I collected judgments about whether people thought they had enough information to understand the agent's reasons. Across three conditions that differed in whether the agent was a normal human drawing a true conclusion, a normal human drawing a false conclusion, or a brain-in-a-vat drawing a false conclusion, participants agreed that they had enough information to comprehend the agent's reasons, as reflected in mean, median, and modal judgments. Nevertheless, mean agreement was significantly higher for the normal human inferring a true conclusion. On the other hand, I collected judgments about which items were among the agent's reasons. The items most strongly and consistently rated as reasons were all psychological items, such as background knowledge, how things seem, and sensory experiences. But there were critical differences across conditions when it came to items external to the agent, including objects and facts. When a normal human drew a true conclusion that there was a fox nearby, participants included among his reasons the physical animal and the fact that it was a fox. By contrast, when a brain-in-a-vat drew a false conclusion, participants included neither the animal nor the fact among his reasons. Yet again by contrast, when a normal human drew a false conclusion that there was a fox nearby, participants included the animal but not the fact among his reasons. Another interesting contrast occurred for the activities of a supercomputer and team of scientists, which were mentioned in different capacities across the three conditions. Participants included the supercomputer's and scientists' activities among the brain-in-a-vat's reasons, but they did not do so for the normal human, regardless of whether he drew a true or false conclusion.

The results of experiment 2 provide strong initial support for a dualist theory of the ordinary reason concept. According to this theory, reasons include both psychological and non-psychological items. By the same token, the results from experiment 2 undermine two competing theories, psychologism, according to which all reasons are psychological items, and anti-psychologism, according to which no reasons are psychological items. The results also undermine a principal assumption used to argue for psychologism. This assumption is that, intuitively, a normally embodied agent and his victimized mental "twin" form beliefs *for the same reasons*. The present results provide powerful initial evidence against this assumption. In particular, for a normally embodied agent, people clearly counted a physical object and an external fact among his reasons. Moreover, they did so at rates comparable to the most highly rated psychological items.

This interpretation of experiment 2 is limited in at least two important ways. First, it assumes that participants conceive of psychological states in a certain way. More specifically, it assumes that they do *not* view psychological states as extending out into the world to *include* what I am calling "external" physical objects or facts. One reason to suspect that this assumption might be wrong comes from the psychologist Fritz Heider, quoted in the introduction, who claimed that, on

the ordinary understanding, perception is "a means whereby objective facts enter the life space" (Heider, 1958, 15). If they enter the life space, then perhaps they do so by entering the "mind space," in which case the results from experiment 2 would, surprisingly, support psychologism. Some philosophers have discussed similar views (McDowell, 1994, lecture II; Turri, 2009, p. 507, n. 5). This is a worthy topic for future investigation. Of course, even if the alternative *world-involving* conception of psychological states turned out to be correct, this would leave untouched the finding that people don't attribute the same reasons to a normally embodied agent and his victimized twin. Indeed, it would likely help explain why people attribute different reasons to the two agents.

Second, experiment 2 involved a between-subjects comparison in order to assess whether people attributed the same reasons to a normal agent and his victimized twin. That is, one group of people rated the normal agent's reasons, and a different group of people rated the victimized twin's reasons. In principle, there is nothing problematic with this and the results are informative. However, it is also worth studying this question within-subjects. That is, it would be informative to also see whether one and the same group of people would attribute different reasons to the normal agent and his victimized twin. This might get closer to what philosophers who originally made the claim of parity had in mind. But even if things turned out that way, such a finding's significance should be considered carefully. Future research could profitably investigate this possibility directly. In the meantime, the hypothesis that a within-subjects comparison would turn out differently should be treated as speculative.

Proponents of psychologism might object that participants who counted physical objects and external facts as reasons were perspective-taking rather than literally attributing reasons. In other words, by counting these non-psychological items as reasons, participants were indicating *how things seemed to the agent*. If this is correct, then the results do not ultimately support dualism. In response, on the one hand, the objection is not well motivated. It is plausible that participants would want to express their understanding of the agent's perspective, and some of the options were well suited to the task. But it is unclear why they would continue perspective-taking for options that were comparatively poorly suited for doing so. On the other hand, the objection cannot explain the observed results. In particular, it cannot explain why participants attributed non-psychological reasons in some conditions but not others. The agent in all three conditions concludes that it was a fox, and participants in all three conditions overwhelmingly counted *it seems like a fox* among the agent's reasons. Thus the objection predicts that perspective-taking, and consequently the attribution of the relevant reasons, will occur in all conditions. But this did not occur. Instead particants selectively attributed the relevant reasons only for normally embodied agents, and even then there were further differences, depending on whether the agent was deceived.

Proponents of psychologism might object that participants counted physical objects and external facts as reasons in order to convey a positive assessment of the agent's reasons, all of which they judged were psychological. For example, when participants counted the physical animal or

external fact as reasons, perhaps they were indicating that the agent knows that it is a fox, or that it was a reasonable conclusion in the circumstances. Nothing in the data rules out this hypothesis. Unlike the objection based on perspective-taking, the present objection can explain the differences across conditions, because people ordinarily view truth as relevant to knowledge (Starmans and Friedman, 2012; Buckwalter, 2014), and the truth-value of the agent's conclusion varies appropriately across conditions. However, the present objection suffers in at least two ways. On the one hand, it is more complicated than the interpretation I offered above, according to which people attributed reasons in order to attribute reasons. On the other hand, it appears ad hoc. If there were independent evidence to think that psychologism was the correct account of the ordinary reason concept, then it would not be ad hoc. But I am aware of no such evidence. Notwithstanding these concerns, I think that further investigation of the objection could be informative. One possibility would be to first ask participants to rate a knowledge attribution, and then ask them to identify the agent's reasons. The basic idea here is that if the objection is correct, then participants who first attribute knowledge will no longer be motivated to use the reason attribution to convey that the agent knows. Accordingly, if participants no longer counted external objects and facts as reasons using this new procedure, then it would support the objection. By contrast, if participants continued to count external objects and facts as reasons, then it would support the intepretation I offered above.

Two aspects of the findings are worth remarking on, even though they do not relate directly to the central theoretical questions that the experiments were designed to address. The first pertains to the "false" condition of experiment 2, which featured a normally embodied agent who drew a false conclusion that there is a fox nearby. When asked to rate this agent's reasons, participants included "the physical animal itself," but not "the fact that it is a fox." In other words, when it was false that it was a fox, participants did not count the non-existent "fact" among the agent's reasons. But they still counted the animal among his reasons. Presumably this is because they interpreted him as looking at an actual animal, which he mistook for a fox. The animal exists and causes him to draw a false conclusion, so it's rated as one of his reasons. Whatever the explanation, I find this array of subtly discerning judgments to be intriguing and potentially worth exploring in its own right.

The second pertains to the "brain in a vat" condition of experiment 2. When asked to rate the brain's reasons for concluding that there was a fox nearby, participants included "the scientists' activities" and "the supercomputer's operations." It is implausible, one might argue, that these are the brain's reasons *in the same sense* as its sensory experiences or background knowledge about foxes are. Instead of being among the brain's *reasons for* drawing its conclusion, the supercomputer's operations and scientists' activities arguably belong in the broader category of *reasons why* the brain drew its conclusion. When I say that they do not "belong" in the *reasons for* category, I do not mean that they aren't classified in that category as ordinarily understood. Instead, I mean that there seems to be a good argument for ultimately concluding that they are not in that category. The argument is that if the brain drew its conclusion based on those reasons, then it would be fool-

14

ish; but the brain is not being foolish; therefore, its conclusion is not based on those reasons. The fact that a supercomputer is making it seem to me that I'm looking at a fox, I submit, is a foolish reason to believe that I'm looking at a fox. Similar remarks apply to the activities of scientists who operate the supercomputer so as to present me with such an appearance. The brain is making a mistake, but it's not being foolish.

One possiblity raised by this discussion is that the ordinary reason concept doesn't distinguish between what I called *reasons why* and *reasons for*. Instead, it might indiscriminately include them both, perhaps separated by a latent "fault line" apt to fracture and create a distinction if one undergoes appropriate experience or instruction. Another hypothesis allows that there are distinct ordinary concepts for each of those categories and people adequately distinguish between them. Instead, people are performing a different task when they include the supercomputer's or scientists' activities among the brain's reasons. Perhaps they are thinking ahead to how the brain might defend itself if challenged on its false belief. Put simply, participants might be anticipating how to *excuse the failure*, effectively substituting one task for another (for precedents, see Turri (2013) and Turri and Blouw (2015)). I emphasize that this "task substitution" hypothesis, while perhaps not implausible, is merely speculative at this point. Only further research designed specifically to address these issues can begin providing definite answers.

Finally, nothing I have said should be taken to imply that commonsense has the last word on what *reasons* and *basing* actually are. My focus has been explicitly on the ordinary concepts of these categories, on characterizing how they appear in the manifest image. A different project is required to study the things themselves. Theorizing and reflection will play a role, just as they do in all serious scientific endeavors, but such a project will involve the tools of cognitive psychology and neuroscience, among other fields. To claim that we are studying the things themselves by reflection alone will come across as a bit of self-serving naiveté, at best.

In conclusion, I would like to emphasize that while the present experiments are an interesting first step in understanding the ordinary concepts of basing and reasons, they are but one step. The present findings were robust across biological sex and age, but the population sampled was limited to adult residents of the U.S. Children and people from different cultures might yield importantly different results. Moreover, I am just one researcher with his own perspective, background assumptions, and interests, which influenced my choice of stimuli, dependent measures, and testing procedures. Other researchers might make different choices that shed different light on the questions addressed here. That is to be welcomed as part of the cumulative and collaborative process of scientific, and philosophical, discovery.

Appendix A: Stimuli for experiment 1

(Juror) Lauren is on a jury hearing a criminal case. After listening carefully throughout the trial, Lauren concludes that the defendant was innocent. A grainy photograph seemed to place the defendant a mile away from the scene of the crime. This photograph made [absolutely no/a] difference to Lauren's thinking on the matter.
1. Lauren's conclusion was partly based on the photograph.
2. The photograph made a difference to Lauren's thinking.

(Vitamins) Lauren has been researching the health effects of vitamin supplements. After reading through several recent publications, Lauren concludes that taking vitamins is mildly beneficial. One experiment suggested that taking vitamins could prevent cancer. This particular experiment made [absolutely no/a] difference to Lauren's thinking on the matter.
1. Lauren's conclusion was partly based on the experiment.
2. The experiment made a difference to Lauren's thinking.

(Winner) Lauren has been wondering which team will win the tournament. After considering information from many sources, Lauren concludes that Springfield will win. One social media post said that Springfield's star player was least likely to be injured. This particular post made [absolutely no/a] difference to Lauren's thinking on the matter.
1. Lauren's conclusion was partly based on the social media post.
2. The social media post made a difference to Lauren's thinking.

(Loyalty) Lauren has been questioning her husband's loyalty. After considering the matter from many angles, Lauren concludes that her husband is loyal. Her husband never speaks to her in a disrespectful tone. His tone made [absolutely no/a] difference to Lauren's thinking on the matter.
1. Lauren's conclusion was partly based on her husband's tone.
2. The husband's tone made a difference to Lauren's thinking.

(Mammals) Lauren is studying the evolution of mammals. After evaluating evidence from many different fields, Lauren concludes that the first mammals were marsupials. A comparative anatomist from Australia previously discussed the same view. His discussion made [absolutely no/a] difference to Lauren's thinking on the matter.
1. Lauren's conclusion was partly based on the anatomist's discussion.
2. The anatomist's discussion made a difference to Lauren's thinking.

Appendix B: Statistical tables

Table 2: Experiment 1. Likelihood ratio tests for the random effects and their order of elimination from the fully specified linear mixed effects model.

	Eliminated	logLik	AIC	LRT	Df	p value
		-940.078	1898.157			
Scenario	1	-940.078	1896.157	0.000	1	> .999
ID:Scenario	0	-1110.604	2235.209	341.052	1	< .001

Table 3: Experiment 1. F tests for the fixed effects and their order of elimination from the fully specified linear mixed effects model.

	Eliminated	Sum Sq	NumDF	DenDF	F value	p value
Status:Judgment	1	0.002	1	305	0.004	0.947
Judgment	2	0.198	1	305	0.495	0.482
sex	3	0.380	1	305	0.945	0.332
age	4	0.735	1	305	1.830	0.177
Status	0	156.865	1	305	390.561	0.000

Table 4: Experiment 2. Likelihood ratio tests for the random effects and their order of elimination from the fully specified linear mixed effects model.

	Eliminated	logLik	AIC	LRT	Df	p value
		-1817.230	3702.460			
ID	0	-1831.993	3729.985	29.526	1	< .001

Table 5: Experiment 2. F tests for the fixed effects and their order of elimination from the fully specified linear mixed effects model.

	Eliminated	Sum Sq	NumDF	DenDF	F value	p value
sex	1	0.006	1	184	0.014	0.906

	Eliminated	Sum Sq	NumDF	DenDF	F value	p value
age	2	0.158	1	184	0.397	0.530
Condition:Item	0	106.466	18	1656	14.814	0.000

References

Baayen, R H, D J Davidson and D M Bates. 2008. "Mixed-effects modeling with crossed random effects for subjects and items." *Journal of Memory and Language* 59(4):390–412.

Berlin, Brent, Dennis E Breedlove and Peter H Raven. 1973. "General principles of classification and nomenclature in folk biology." *American Anthropologist* 75(1):214–242.

Buckwalter, Wesley. 2014. "Factive verbs and protagonist projection." *Episteme* 11(4):391–409.

Buckwalter, Wesley and John Turri. 2018. Moderate scientism in philosophy. In *Scientism: prospects and problems*, ed. J R Ridder, R Peels and R van Woudenberg. Oxford University Press.

Clark, Herbert H. 1973. "The language-as-fixed-effect fallacy: a critique of language statistics in psychological research." *Journal of verbal learning and verbal behavior* 12:335–359.

Cottrell, Jane E and Winer Gerald A. 1994. "Development in the understanding of perception: the decline of extramission perception beliefs." *Developmental Psychology* 30(2):218–228.

Gigerenzer, Gerd. 2014. *Risk savvy: how to make good decisions*. New York: Viking.

Heider, Fritz. 1958. *The psychology of interpersonal relations*. Mansfield Centre, CT: Martino Publishing.

Judd, Charles M, Jacob Westfall and David A Kenny. 2012. "Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem." *Journal of personality and social psychology* 103(1):54–69.

Kelley, H H. 1992. "Common-sense psychology and scientific psychology." *Annual Review Of Psychology* 43(1):1–23.

Kosko, Bart. 1993. *Fuzzy thinking: the new science of fuzzy logic*. New York: Hyperion.

Kuznetsova, Alexandra, Per B Brockhoff and Rune H B Christensen. 2017. "lmerTestPackage: Tests in Linear Mixed Effects Models." *Journal of Statistical Software* 82(13).

Litman, Leib, Jonathan Robinson and Tzvi Abberbock. 2017. "TurkPrime.com: A versatile crowd-sourcing data acquisition platform for the behavioral sciences." pp. 1–10.

McCloskey, Michael, Allyson Washburn and Linda Felch. 1983. "Intuitive physics: The straight-down belief and its origin." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 9(4):636.

McDowell, John. 1994. *Mind and world*. Cambridge, Mass.: Harvard University press.

Nakagawa, Shinichi and Holger Schielzeth. 2012. "A general and simple method for obtaining R2from generalized linear mixed-effects models." *Methods in Ecology and Evolution* 4(2):133–142.

R Core Team. 2018. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Schtulman, Andrew. 2017. *Scienceblind: why our intuitive theories about the world are so often wrong*. New York: Basic Books.

Sellars, Wilfrid. 1963. *Science, perception and reality*. Atascadero, CA: Ridgeview Publishing Company.

Starmans, Christina and Ori Friedman. 2012. "The folk conception of knowledge." *Cognition* 124(3):272–283.

Turri, John. 2009. "The ontology of epistemic reasons." *Nous* 43(3):490–512.

Turri, John. 2011. "Believing for a reason." *Erkenntnis* 74(3):383–397.

Turri, John. 2013. "The test of truth: An experimental investigation of the norm of assertion." *Cognition* 129(2):279–291.

Turri, John. 2015. "The radicalism of truth-insensitive epistemology: truth's profound effect on the evaluation of belief." *Philosophy and Phenomenological Research* 93(2):348–367.

Turri, John. 2016. How to do better: toward normalizing experimentation in epistemology. In *Advances in experimental philosophy and philosophical methodology*, ed. Jennifer Nado. London: Bloomsbury Academic pp. 35–51.

Turri, John. 2019. Experimental epistemology and "Gettier" cases. In *The Gettier Problem*, ed. Stephen Hetherington. Cambridge University press.

Turri, John and Peter Blouw. 2015. "Excuse validation: a study in rule-breaking." *Philosophical Studies* 172(3):615–634.

Winer, Gerald A, Jane E Cottrell, Kiriaki D Kareffilaki and Matthew Chronister. 1996. "Conditions affecting belief about visual perception among children and adults." *Journal of Experimental Child Psychology* 61(2):93–115.

Winer, Gerald A, Jane E Cottrell, Virginia Gregg, Jody S Fournier and Lori A Bica. 2002. "Fundamentally misunderstanding visual perception: Adults' belief in visual emissions." *American Psychologist* 57(6-7):417–424.