Evidence of factive norms of belief and decision

John Turri[1]

Abstract According to factive accounts of the norm of belief and decision-making, you should not believe or base decisions on a falsehood. Even when the evidence misleadingly suggests that a false proposition is true, you should not believe it or base decisions on it. Critics claim that factive accounts are counterintuitive and badly mischaracterize our ordinary practice of evaluating beliefs and decisions. This paper reports four experiments that rigorously test the critic's accusations and the viability of factive accounts. The results undermine the accusations and provide the best evidence yet of factive norms of belief and decision-making. The results also help discriminate between two leading candidates for a factive norm: truth and knowledge. Knowledge is the superior candidate.

1 Introduction

We routinely make assertions, form beliefs and make decisions. These are ubiquitous and unavoidable in the course of ordinary human affairs. It is important to do these things correctly. Our individual and collective well being often depends on it. Unsurprisingly, then, philosophers are keenly interested in what such correctness consists in. Over the past decade, an enormous amount of work has been done on the norms of assertion, belief and decision (or practical reasoning).

✉ John Turri
    john.turri@gmail.com

[1]   Philosophy Department and Cognitive Science Program, University of Waterloo,
      200 University Avenue West, Waterloo, ON N2L3G1, Canada

According to the knowledge account of the norm of assertion, you should assert a proposition only if you know that it's true. A wealth of observational and experimental evidence supports the knowledge account (MacIver 1938; Unger 1975; Williamson 2000; Turri 2010, 2014a, b; under review a, b; Turri, Friedman and Keefner under review; for a comprehensive review; see Turri under review c; see also Benton 2014). But critics object that the knowledge account faces intuitively compelling counterexamples. The most common and persistent sort of example features a reasonable false assertion (e.g. Douven 2006; Hill and Schechter 2007; Lackey 2007; Levin 2008). The idea here is that it is sometimes reasonable to believe false propositions, and if it is reasonable for you to believe a false proposition, then intuitively there is no sense in which you should not assert it. But knowledge requires truth, so you can't know any such proposition and, hence, knowledge isn't the norm of assertion.

According to the knowledge account of the norm of belief, you should believe a proposition only if you know that it is true (Williamson 2000, pp. 255–256; see also Sutton 2007; Bach 2008, p. 77). Much less evidence supports this view about belief than its analog about assertion (for some discussion, see Huemer 2007, 2011; Bird 2007; Littlejohn 2013; Turri in press a). Critics object to the knowledge account of belief by, again, claiming that it is "completely implausible" (McGlynn 2013, p. 390) and that it faces intuitively compelling counterexamples. Foremost among these are cases of reasonable false belief (Conee 2007; Benton 2012, ch. 6).

According to the knowledge account of the norm of decision-making (or practical reasoning), you should act based on a proposition only if you know that it is true (Hawthorne 2004, p. 29–30; Hawthorne and Stanley 2008). Again, less evidence supports this view about decision than its analog about assertion (for some evidence, see Hawthorne and Stanley 2008; Montminy 2013). And, once again, critics claim that it faces obvious counterexamples and is at odds with our ordinary practice of evaluating decisions. Foremost among the complaints are, yet again, cases of decision based on reasonable false beliefs (Hill and Schechter 2007, p. 115; Douven 2008, pp. 106–107, n. 9; Brown 2008).

All three knowledge accounts—of assertion, of belief, and of decision-making—face the same objection. In each case, critics object that the account is highly counterintuitive and revisionary. They claim that it faces obvious counterexamples and coheres poorly with ordinary practice. The supposed counterexamples focus on reasonable false beliefs. The implication is that an account that respects ordinary practice must not require truth. In short, it must be *non-factive*. Critics have developed a variety of non-factive views.

Researchers investigating the norms of assertion, belief and decision-making agree that the project is, at least in large part, empirical. As one prominent critic of factive accounts puts it, "the project of determining which rule governs the practice of assertion is best conceived not as an *a priori* investigation into the nature of assertion but, rather, as an empirical project." This implies that "any proposal made in the course of" the investigation is "subject to the exact same standards of evaluation as are employed in the empirical sciences generally" (Douven 2006, p. 450). In short, an adequate theory "must face the linguistic data" (Douven 2006, p. 450; see also MacIver 1938; Unger 1975; Williamson 2000; Fantl and McGrath 2002, 2007, 2009, 2013; Hawthorne 2004; Stanley 2005; Weiner 2005; Lackey 2007; Hawthorne and Stanley

2008; Kvanvig 2009; Turri 2011; Smithies 2012; Smith 2012; Reed 2013; Turri 2014b; Coffman 2014; Wright 2014; Pritchard 2014). More specifically, researchers aim for accounts that provide charitable and psychologically plausible interpretations of relevant patterns in ordinary thought, talk, and behavior. The less one's theory requires reading error into these patterns, the better.

The underlying assumption here is that people are at least implicitly sensitive to the norms in question and, consequently, that their normative intuitions can be a source of evidence about the content of the norms. Put otherwise, people's intuitive judgments about cases tend to manifest their competence, resulting in detectable patterns. We can then use these patterns when theorizing about the content of the norms. The patterns will make some proposals much less likely than others, given what we expect from competent performance. This basic approach is shared by much influential work in philosophy (e.g. Locke 1690/1975, book 4.11; Reid 1764/1997; Moore 1959; Austin 1956; Goldman 1979, 1993; Dretske 1981; Stroud 1984; Vogel 1990; Cohen 1988, 2013; DeRose 1995, 2009; Greco 2010). It is also shared by social and cognitive scientists who appeal to competent speakers' intuitions to support theories of syntax, experimental investigations of the relationship between semantics and pragmatics, and other forms of psycholinguistic experimentation (e.g. Chomsky 1957; Noveck and Sperber 2004). It also shared by legal theorists attempting to justify or motivate changes to aspects of criminal law doctrines (Fletcher 1978; Robinson 1982; Robinson and Darley 1998).

A recent series of experiments tested whether critics have correctly described our ordinary practice of evaluating reasonable false assertions (Turri 2013; Turri and Blouw 2015). People in these experiments considered a simple story about Maria. Maria is a watch collector who owns so many watches that she can't keep track of them all by memory alone, so she maintains a detailed inventory of them. She knows that the inventory, although not perfect, is extremely accurate. One day someone asks Maria whether she has a 1990 Rolex in her collection. She consults the inventory and it says that she does have one. At the end of the story, one group of people was told that the inventory was right. Another group of people was told that the inventory was wrong. Everyone then answered the same question: should Maria say that she has a 1990 Rolex in her collection?

The results were absolutely clear. When the assertion would be true, virtually everyone said that Maria should make the assertion. But when the assertion would be false, the vast majority said that she should not make the assertion. This same basic pattern persisted when people were questioned in different ways, in high and low stakes settings, and for assertions with positive and negative content. One study in particular demonstrated how much subtlety and sophistication informs ordinary judgments about assertability. Instead of answering "yes" or "no" to whether Maria should make the assertion, or rating their agreement with the statement that Maria should make the assertion, people performed a much more open-ended task of identifying what Maria should say. When the assertion would be true, the vast majority of people answered that Maria should assert that she owns the watch. But when the assertion would be false, very few people answered that way. Instead the most most common response was that Maria should assert that she "probably" owns one, which, on the most natural interpretation of the case, is actually true because of Maria's evidence.

So critics have mischaracterized how reasonable false assertions are actually evaluated. This raises the prospect that critics have also overstated their objections to factive accounts of the norms of belief and decision-making.

The present paper reports four experiments designed to investigate this possibility. The experiments are modeled after the earlier work on assertion. Experiment 1 focuses on evaluations of belief. Experiment 2 focuses on evaluations of decision. In each case, the findings demonstrate that critics have indeed overstated their objections. Factive accounts fit the data very well, but non-factive alternatives do not. Experiment 3 proposes one explanation for why critics were so badly mistaken. The explanation is that critics falsely assume that the quality of someone's evidence for a proposition is relevantly insensitive to the proposition's truth value. Experiment 4 replicates the findings from the first three experiments and provides initial evidence that a knowledge account of the relevant norms better fits the behavioral data than a mere truth account does.

2 Experiment 1: belief

2.1 Method

2.1.1 Participants

Two hundred and one U.S. residents were tested (aged 18–71 years, mean age = 32 years; 96 % reporting English as a native language; 77 female). Participants were recruited and tested online using Amazon Mechanical Turk and Qualtrics and compensated $0.30 for approximately 2 min of their time. Repeat participation was prevented.

2.1.2 Materials and procedure

Participants were randomly assigned to one of four conditions in a 2 (Cover Story: Watch/Name) × 2 (Truth Value: True/False) between-subjects design. Participants read a single story about an agent who keeps imperfect but extremely accurate inventory of some items of interest. The Cover Story factor varied which story participants read. Some participants read about Maria, a watch collector with over ten thousand watches; the question arises whether she owns a certain type of watch. Other participants read about Mario, a human resource manager for a company with over ten thousand employees; the question arises whether he has an employee with a certain name. In each case, Maria or Mario consults the inventory to answer the question. Maria reads that she does own a watch of that type; Mario reads that he does have an employee by that name. I didn't expect an effect of Cover Story. I included it as a robustness check to guard against effects being driven by superficial features of any one storyline. The Truth Value factor varies whether the inventory is right or wrong on this occasion. The Appendix includes the complete text of the stories.

After reading the story, participants answered a single test question: "What should Maria/Mario believe?" Responses were collected on a standard 7-point Likert scale

anchored with "I definitely do have one" (=+3), "I do have one," "I probably do have one," "It's unclear" (=0), "I probably don't have one," "I don't have one," and "I definitely don't have one" (=−3), appearing in that order left-to-right across the participant's screen. Participants never saw the numerical values, only the qualitative anchors. After testing, participants completed a brief demographic questionnaire.

2.2 Results

Preliminary analysis revealed that neither participant age nor participant gender affected response to test question. Neither demographic factor entered into any interactions either. Accordingly, the analyses that follow collapse across these factors. The same is true for Experiment 2 below.

A one-way analysis of variance revealed no effect of Cover Story on response to the test question, $F(1, 197) = 1.34$, $p = .248$, no interaction between Cover Story and Truth Value, $F(1, 197) = 1.60$, $p = .208$, and a main effect of Truth Value, $F(1, 197) = 117.88$, $p < .001$. Follow-up analysis revealed that mean response was significantly higher in True conditions ($M = 2.03$, $SD = 1.07$) than False condition ($M = 0.01$, $SD = 1.53$), independent samples t test, $t(181) = 10.89$, $p < .001$. The magnitude of the mean difference ($MD = 2.02$, 95 % CI 1.65–2.39) was extremely large, $d = 1.62$. One sample t tests revealed that mean response was significantly above the neutral midpoint (=0) in True conditions, $t(98) = 18.82$, $p < .001$, $MD = 2.03$, 95 % CI 1.81–2.24, $d = 1.9$, whereas it didn't differ from midpoint in False conditions, $t(101) = 0.07$, $p = .948$, $MD = .01$, 95 % CI −0.29 to 0.31.

In True conditions, 82 % of participants said that the agent should believe "I definitely do have one" or "I do have one," compared to only 18 % of participants in False conditions. A χ^2 test for independence (with Yates continuity correction) indicated that the magnitude of the difference in frequencies was extremely large, $\chi^2(1, 201) = 80.22$, Cramer's V = .642. The mode response in True Watch and True Name was "I do have one." In False Watch it was "I probably do have one." In False Name it was "It's unclear." (See Fig. 1).

2.3 Discussion

The results show that critics of factive accounts have mischaracterized the natural and intuitive view of reasonable false belief. People considered a case where someone forms a belief based on information from a reliable source. Changing just the truth value of the source's information on the present occasion radically changes what people think the agent should believe. When the proposition is true, people overwhelmingly say that the agent should believe that the proposition is true or definitely true. But when the proposition is false, very few people say that the agent should believe that the proposition is true. Instead the most common response is that the agent should believe that the proposition is probably true, or even that it's unclear whether the proposition is true.
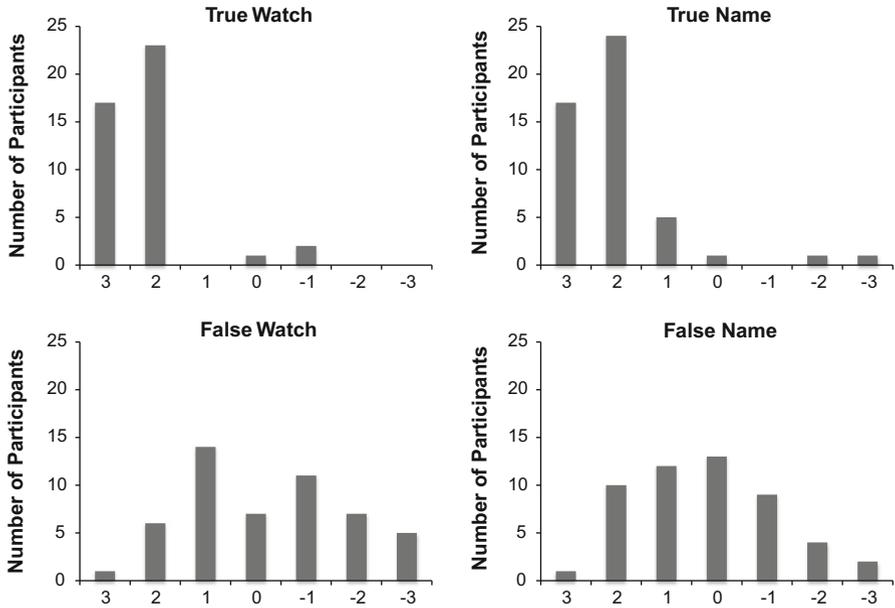
Fig. 1 Experiment 1. Distribution of responses to the test question. 3 = "I definitely do have one," 2 = "I do have one," 1 = "I probably have one," 0 = "It's unclear," −1 = "I probably don't have one," −2 = "I don't have one," −3 = "I definitely don't have one"

3 Experiment 2: decision

3.1 Method

3.1.1 Participants

Two hundred new participants were tested (aged 18–66 years, mean age = 32 years; 96 % reporting English as a native language; 84 female). Participants were recruited and tested the same way as in Experiment 1.

3.1.2 Materials and procedure

Participants were randomly assigned to one of four conditions in the same 2 (Cover Story: Watch/Name) × 2 (Truth Value: True/False) between-subjects design used in Experiment 1. The stories were very similar to those used in Experiment 1. The narrative changed slightly so that it lead more naturally to a question about what the agent should decide. In the Watch story, Maria learns that if she owns a certain watch, then she needs to make an appointment to revise their insurance policy, which will take several hours; but if she doesn't own one, then she doesn't need to make an appointment. In the Name story, Mario learns that if he has an employee with a certain name, then he needs to make an appointment to revise paperwork on file with the immigration office, which will take several hours; but if he doesn't have one, then he doesn't need to make an appointment. The Appendix includes the complete text of the stories.

After reading the story, participants responded to a single decision probe: "Maria/Mario _____ make an appointment." Responses were collected on a standard 6-point Likert scale anchored with "definitely should" ($=+3$), "should," "probably should," "probably should not," "should not," "definitely should not" ($=-3$), appearing in that order left-to-right across the participant's screen. Participants never saw the numerical values, only the qualitative anchors. After testing, participants completed a brief demographic questionnaire.

3.2 Results

A one-way analysis of variance revealed no effect of Cover Story on response to the test question, $F(1, 196) = 1.16$, $p = .282$, no interaction between Cover Story and Truth Value, $F(1, 196) = 0.55$, $p = .457$, and a main effect of Truth Value, $F(1, 196) = 205.11$, $p < .001$. Follow-up analysis revealed that mean response was significantly higher in True conditions ($M = 2.50$, $SD = 0.89$) than False conditions ($M = -0.53$, $SD = 1.91$), independent samples t test, $t(141.63) = 14.34$, $p < .001$. The magnitude of the mean difference ($MD = 3.03$, 95 % CI 2.61–3.45) was extremely large, $d = 2.43$. One sample t tests revealed that mean response was significantly above the neutral midpoint ($=0$) in True conditions, $t(98) = 28.05$, $p < .001$, $MD = 2.5$, 95 % CI 2.32–2.67, $d = 2.82$, whereas it was significantly below the midpoint in False conditions, $t(100) = -2.81$, $p = .006$, $MD = -0.53$, 95 % CI -0.91 to -0.16, $d = 0.28$.

In True conditions, 88 % of participants said that the agent "should" or "definitely should" make the appointment, compared to only 22 % of participants in False conditions. A χ^2 test for independence (with Yates continuity correction) indicated that the magnitude of the difference in frequencies was extremely large, $\chi^2(1, 200) = 88.1$, Cramer's V $= .664$. The mode response in True Watch and True Name was "should." In False Watch it was "definitely should not." In False Name it was "probably should not." (See Fig. 2).

3.3 Discussion

The results show that critics of factive accounts have badly mischaracterized the natural and intuitive view of decision-making. People considered a case where someone has to make a decision based on information from a reliable source. Changing just the truth value of the source's information on the present occasion radically changes what people say the agent should decide. When the proposition is true, people overwhelmingly say that the agent should or definitely should perform a certain action, in this case making an appointment. But when the proposition is false, very few people say that the agent should perform that action. Instead the most common response is that the agent should not perform that action.

4 Experiment 3: evidence

When explaining their interpretation of cases like the ones tested above, critics of factive accounts appeal heavily to the fact that the agent has "excellent reasons" (Douven 2008, p. 106, n. 9) or is "epistemically justified" (Hill and Schechter 2007, p. 115) or
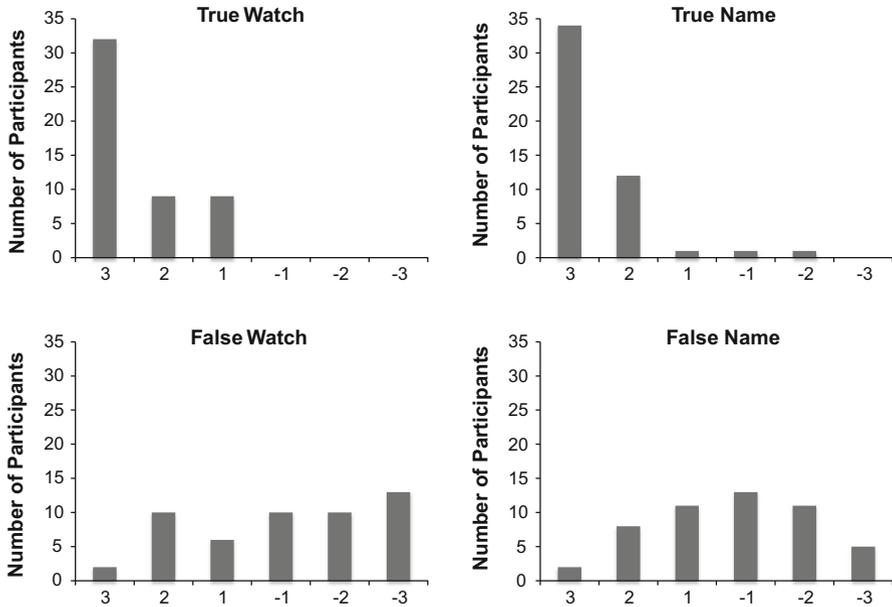
Fig. 2 Experiment 2. Distribution of responses to the test question. 3 = "definitely should," 2 = "should," 1 = "probably should," −1 = "probably should not," −2 = "should not," −3 = "definitely should not"

has sufficient "evidence" (Kvanvig 2009, p. 145) to believe certain things. This seems to assume something important about evidence. In particular, it seems to assume that, on any particular occasion, the quality of evidence is insensitive (enough) to the truth of the matter. For example, suppose that Maria has a highly reliable source of information, such as an extremely accurate inventory of watches she owns. When Maria consults the inventory, it says that she owns a certain watch. The evidence this provides Maria is equally good regardless of whether the inventory is right on this particular occasion. Or so critics assume.

Something like this view of evidence is popular among contemporary philosophers (e.g. Chisholm 1989, p. 76; Feldman 2003, p. 29). For instance, consider the notorious thought experiment involving a radically deceived brain-in-a-vat:

> Could you not be floating in a tank while super-psychologists stimulate your brain electrochemically to produce exactly the same experiences you are now having, or even to produce the whole sequence of experiences you have had in your lifetime thus far? (Nozick 1981, p. 167)

Commenting on cases like this, one leading epistemologist says that such victims are "justified" because they "possess *good reason* for their beliefs—indeed, precisely the same reasons that we ourselves possess for our beliefs about the world." The super-psychologist's "machinations" cause the victims to have highly misleading evidence, "but this has no tendency to show that those reasons are not good ones." This verdict on the case, we're told, is "beyond question at an intuitive level" (BonJour and Sosa 2003, pp. 185–186).

Focusing on less extravagant examples, another leading epistemologist makes the same basic point. Consider a "typical case" where there is "nothing odd" and "things are exactly as the person believes them to be." Compare that to an "unusual case" where "the person has that very same evidence, but the proposition in question is nevertheless false." The "key thing to note" here is that in each case the person "has exactly the same reasons for believing exactly the same thing." Consequently, if the person has a good reason to believe the proposition in either case, then the person has a good reason in both cases (Feldman 2003, p. 29).

The present experiment tests whether this is the natural and intuitive view of evidence.

4.1 Method

4.1.1 Participants

One hundred fifty-three new participants were tested (aged 18–66 years, mean age = 30 years; 96 % reporting English as a native language; 57 female). Participants were recruited and tested the same way as in earlier experiments.

4.1.2 Materials and procedure

Participants were randomly assigned to one of three conditions, Control/True/False, in a between-subjects design used in Experiment 1. Participants read a version of the Watch story used in Experiment 2. The story for the True condition was the True Watch story from Experiment 2; the story for the False condition was the False Watch story from Experiment 2. The story for the Control condition was exactly the same as the others, except that it does not specify at the end whether the inventory is right or wrong (i.e. it omits the final sentence).

After reading the story, participants responded to a single question: "How would you rate the quality of Maria's evidence for thinking that she has one in her collection?" Responses were collected on a standard 6-point Likert scale anchored with "very good" (=+3), "good," "somewhat good," "somewhat bad," "bad," "very bad" (=−3), appearing in that order left-to-right across the participant's screen. Participants then went to a new screen where they responded to an open knowledge probe: "Maria _____ that she has one in her collection." Responses were collected on a standard 6-point Likert scale anchored with "definitely knows" (=+3), "knows," "probably knows," "probably doesn't know," "doesn't know," and "definitely doesn't know" (=−3). Participants never saw the numerical values, only the qualitative anchors. After testing, participants completed a brief demographic questionnaire.

The point of including the knowledge probe is to check how well people's judgments about evidential quality correlate with knowledge judgments. Critics of factive accounts often claim that having knowledge-grade justification or something closely related, but not truth or knowledge itself, is the norm of assertion, belief or decision (Lackey 2007, p. 611; Kvanvig 2009, p. 156; Smithies 2012; Coffman 2014, p. 37). Thus it would be good to know if people's judgments about evidential quality were tracking a conception of quality closely connected to their knowledge judgments.
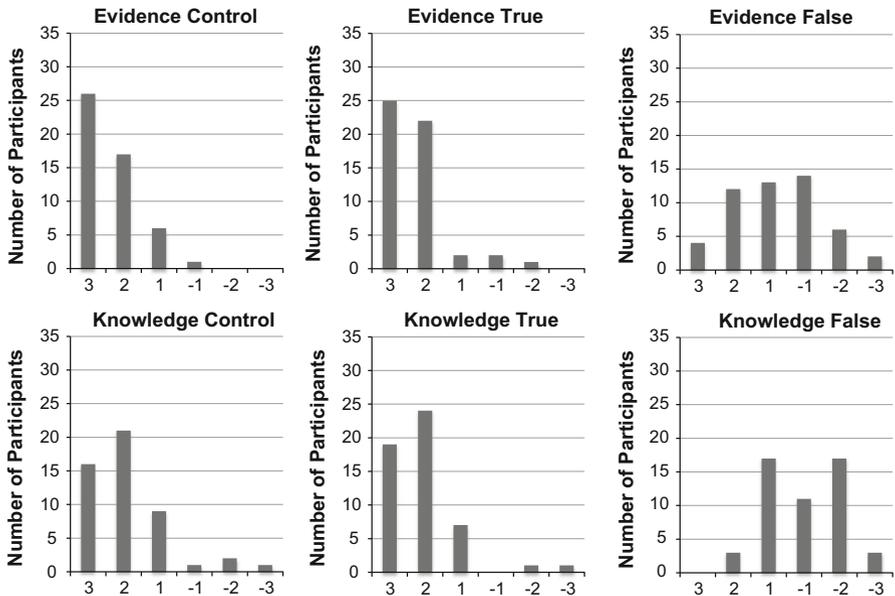
Fig. 3 Experiment 3. *Top panel* distribution of responses for the evaluation of evidence: 3 = "very good," 2 = "good," 1 = "somewhat good," −1 = "somewhat bad," −2 = "bad," −3 = "definitely should not." *Bottom panel* distribution of responses to the knowledge probe: 3 = "definitely knows," 2 = "knows," 1 = "probably knows," −1 = "probably doesn't know," −2 = "doesn't know," −3 = "definitely doesn't know"

4.2 Results

A one-way analysis of variance revealed an effect of condition on the evaluation of evidence, $F(1, 150) = 40.86$, $p < .001$. Follow-up analysis revealed that mean evaluation was significantly higher in Control ($M = 2.34$, $SD = 0.85$) than False ($M = 0.33$, $SD = 1.72$), independent samples t test, $t(73.3) = 7.47$, $p < .001$. The magnitude of the mean difference ($MD = 2.01$, 95% CI 1.47–2.54) was extremely large, $d = 1.75$. Mean evaluation was significantly higher in True ($M = 2.25$, $SD = 1.06$) than in False, $t(83.2) = 6.80$, $p < .001$. The magnitude of the mean difference ($MD = 1.92$, 95% CI 1.36–2.47) was extremely large, $d = 1.49$. Mean evaluation did not differ between Control and True, $t(100) = 0.47$, $p = .638$. One sample t tests revealed that mean evaluation was significantly above the midpoint ($=0$) in Control, $t(49) = 19.52$, $p < .001$, $MD = 2.34$, 95% CI 2.10–2.58, $d = 2.70$; it was significantly above the midpoint in True, $t(51) = 15.25$, $p < .001$, $MD = 2.25$, 95% CI 1.95–2.55, $d = 2.12$; and it didn't differ significantly from midpoint in False, $t(50) = 1.39$, $p = .172$.

In Control, 98% of participants agreed that the evidence was good to some degree or other (i.e. "very good," "good," or "somewhat good"), $\chi^2(1, 50) = 46.08$, $p < .001$; in True, 94% of participants answered that way, $\chi^2(1, 52) = 40.69$, $p < .001$; in False, 57% did, $\chi^2(1, 51) = 0.96$, $p = .327$. The mode response in Control and True was "very good." In False it was "somewhat bad." (See Fig. 3).

Knowledge judgments and evidential evaluations were strongly positively correlated, Spearman's correlation, $r_s(151) = .622$, $p < .001$. Still, overall, mean eviden-

tial evaluation (M = 1.64, SD = 1.56) was higher than mean knowledge judgment (M = 1.09, SD = 1.80), paired samples *t* test, t(152) = 4.40, p < .001.

4.3 Discussion

In this experiment, people were asked to evaluate the evidence someone gets from a source known to be reliable. Changing the truth value of the proposition radically changed how people rated the evidence's quality. When the proposition was true, people overwhelmingly said that the evidence was very good. But when the proposition was false, very few people said that the evidence was very good. Instead the central tendency was one of ambivalence about the evidence's quality. These results support the hypothesis that many theorists—including prominent critics of factive accounts of the norms of assertion, belief and decision—have mischaracterized the natural and intuitive view of evidence. In particular, they have falsely assumed that "it is beyond question at an intuitive level" that the quality of someone's evidence for a proposition is insensitive to the proposition's truth value (BonJour and Sosa 2003, p. 186).

5 Experiment 4: putting it all together

This experiment begins looking more closely at the relationship between truth and the evaluation of beliefs and decisions, on the one hand, and judgments about evidence and knowledge, on the other. Until now I have simply evaluated the effect of manipulating truth value on the evaluation of belief, decision and evidence. In each case I found that truth value enormously influenced evaluations. We now know *that* truth value influences them. However, the earlier experiments weren't designed to estimate *how* truth value accomplishes this. This experiment is designed to help answer this question. To do so, I will divide participants into conditions that manipulate truth value and have them evaluate a belief or decision, just like earlier experiments. But I will also have each participant rate how good the agent's evidence is and whether the agent has knowledge. This enables me to use powerful statistical techniques—linear regression and bootstrap mediation analysis—to estimate how truth value affects the evaluation of belief and decision. In particular, it enables me to determine whether judgments about evidence or knowledge affect the evaluations when controlling for the influence of truth value. Similarly, it enables me to determine whether people's judgments about evidence and knowledge mediate the effect of truth value on evaluations of belief and decision.

5.1 Method

5.1.1 Participants

Three hundred new participants were tested (aged 18–71 years, mean age = 32 years; 95 % reporting English as a native language; 112 female). Participants were recruited and tested the same way as in earlier experiments.

5.1.2 Materials and procedure

Participants were randomly assigned to one of four conditions in a 2 (Type: Believe/Decide) × 2 (Truth Value: True/False) between-subjects design. Participants read a version of the Watch story used in Experiment 2. The story for the True conditions was the True Watch story from Experiment 2; the story for the False conditions was the False Watch story from Experiment 2.

After reading the story, participants in Believe conditions responded to the same question about what Maria should believe as in Experiment 1: "What should Maria believe?" Responses were collected on a standard 6-point Likert scale anchored with "I definitely do have one" (=+3), "I do have one," "I probably do have one," "I probably don't have one," "I don't have one," "I definitely don't have one" (=−3), appearing in that order left-to-right across the participant's screen. Participants in Decide conditions responded to the same open probe about what Maria should do as in Experiment 2: she "definitely should" (=+3), "should," "probably should," "probably should not," "should not," "definitely should not (=−3). Participants then went to a new screen and rated the quality of Maria's evidence, "How would you rate the quality of Maria's evidence for thinking that she has one in her collection?" Responses were collected on a standard 6-point Likert scale anchored with "very good" (=+3), "good," "somewhat good," "somewhat bad," "bad," "very bad" (=−3). Participants then went to a new screen and responded to an open knowledge probe: "Maria _____ that she has one in her collection." Responses were collected on a standard 6-point Likert scale anchored with "definitely knows" (=+3), "knows," "probably knows," "probably doesn't know," "doesn't know," and "definitely doesn't know" (=−3). Finally, participants then went to a new screen and answered a manipulation check from memory: "What was objectively true in the story? Maria _____ have one in her collection" (does/does not). The story remained at the top of the screen for all questions except for the manipulation check. Participants never saw numerical values, only qualitative anchors. After testing, participants completed a brief demographic questionnaire.

5.2 Results

The Truth Value manipulation was extremely effective; 94 % of participants correctly identified from memory whether the proposition was true, $\chi^2(1, 300) = 225.33$, $p < .001$. As can be seen by visually inspecting Fig. 4 below, I observed the same basic pattern as in earlier experiments for all of the dependent measures. A detailed analysis of the data can be found in an appendix. In the main text, I will focus on regression and mediation analyses. A multivariate analysis of variance revealed main and interaction effects of Type and Truth Value on the dependent variables. So I analyzed the data for Believe and Decide conditions separately, beginning with the Believe conditions.

I conducted a standard multiple linear regression with response to the "should believe" question ("belief evaluation") as the outcome and Truth Value, evidential evaluation, and knowledge score as predictors. The purpose of this test is to estimate the unique contribution that each of these predictors makes to predicting belief evaluation. Preliminary analyses revealed no violation of the assumptions of
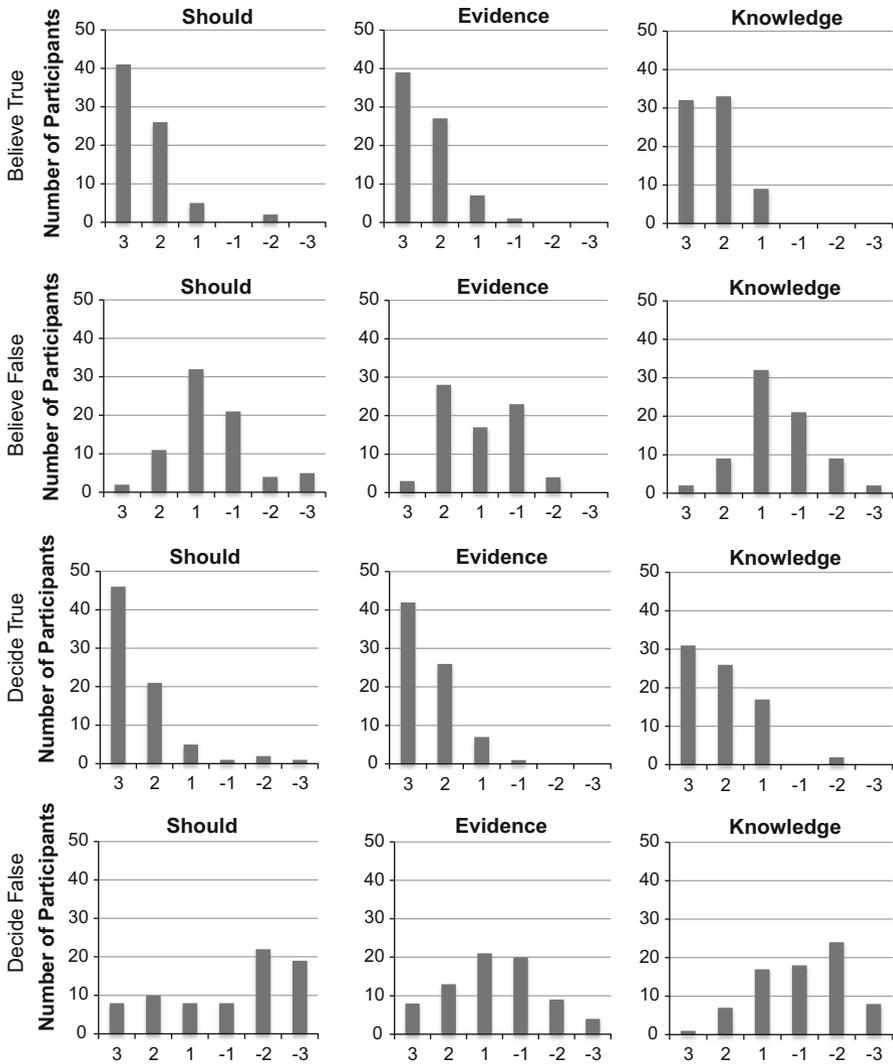
Fig. 4 Experiment 4. Distribution of responses across the four conditions

normality, linearity, multicollinearity and homoscedasticity. The overall model was significant, $F(3, 145) = 61.1$, $p < .001$, $R^2 = .558$. Belief evaluation was significantly predicted by Truth Value, $t(148) = 3.73$, Beta $= .305$, $p < .001$, and knowledge scores, $t(148) = 5.32$, Beta $= .415$, $p < .001$. But belief evaluation was not significantly predicted by evidential evaluations, $t(148) = 1.82$, Beta $= .127$, $p = .072$. Knowledge scores were the strongest predictor of belief evaluation. A multiple-mediators bootstrap analysis (Hayes 2013) (independent variable: Truth Value; outcome: belief evaluation; potential mediators: evidential evaluation and knowledge score) revealed that evidential evaluation did not mediate the effect of Truth Value on belief evaluation, 95 % CI for the indirect effect $= -0.09$ to 0.63. But knowledge scores did significantly mediate
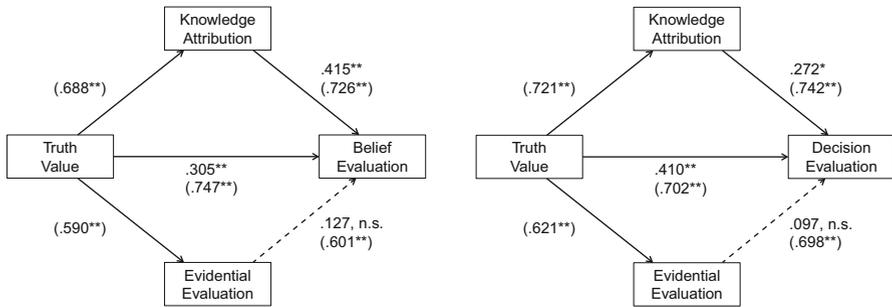
Fig. 5 Experiment 4. Mediation results from Experiment 4. *Left panel* belief evaluation. *Right panel* decision evaluation. Parenthetical values represent the strength of a simple regression between the two variables; values outside the parentheses represent the strength of the relationships in a model used to test for mediation. *p < .01, **p < .001

the effect of Truth Value on belief evaluation, 95 % CI for the indirect effect = 0.47 to 1.54. The ratio of the indirect effect through knowledge to the total effect of Truth Value was surprisingly large (=.43) (Fig. 5).

I next conducted a standard multiple linear regression with response to the "should decide" question ("decision evaluation") as the outcome and Truth Value, evidential evaluation, and knowledge score as predictors. The purpose of this test is to estimate the unique contribution that each of these predictors makes to predicting decision evaluation. Preliminary analyses revealed no violation of the assumptions of normality, linearity, multicollinearity and homoscedasticity. The overall model was significant, $F(3, 147) = 47.69$, $p < .001$, $R^2 = .493$. Decision evaluation was significantly predicted by Truth Value, $t(147) = 4.57$, Beta = .410, $p < .001$, and knowledge scores, $t(147) = 3.05$, Beta = .272, $p = .003$. But decision evaluation was not significantly predicted by evidential evaluations, $t(147) = 1.23$, Beta = .097, $p = .219$. Truth Value was the strongest predictor of action evaluation. A multiple-mediators bootstrap analysis (independent variable: Truth Value; outcome: decision evaluation; potential mediators: evidential evaluation and knowledge score) revealed that evidential evaluation did not mediate the effect of Truth Value on decision evaluation, 95 % CI for the indirect effect $=-0.27$ to 0.85. But knowledge scores did significantly mediate the effect of Truth Value on decision evaluation, 95 % CI for the indirect effect = 0.15 to 1.65. The ratio of the indirect effect through knowledge to the total effect of Truth Value was smaller than for belief evaluation (=.30).

It is worth briefly noting some other interesting aspects of the results. A multivariate analysis of variance revealed interesting interaction effects of Type and Truth Value on the dependent variables. Follow-up independent samples *t* tests revealed that in True conditions, Type of evaluation (i.e. belief versus decision) did not affect mean response to the "should" question, the evaluation of evidence, or knowledge scores (all ps > .10). But things were much different in False conditions. Mean evaluation of belief (M = 0.21, SD = 1.53) was significantly higher than mean evaluation of action (M = −0.76, SD = 2.16), $t(133.35) = 3.19$, $p = .002$, MD = 0.97, 95 % CI 0.37–1.58, d = 0.55. Knowledge scores were significantly higher in Believe (M = 0.15, SD = 1.47) than Decide (M = −0.75, SD = 1.65), $t(148) = 3.52$, $p < .001$, MD = 0.89, 95 % CI

0.39–1.40, d = 0.58. Evidential evaluation was non-significantly higher in Believe (M = 0.68, SD = 1.47) than Decide (M = 0.28, SD = 1.77), t(143.3) = 1.51, p = .134.

5.3 Discussion

This experiment replicated all the main findings from the first three experiments. Once again, changing just the truth value of a reliable source's information radically changed how people evaluate an agent's beliefs, decisions and evidence. More importantly, this experiment began digging deeper into the underlying causal relationships among these judgments. Even controlling for the influence of evaluations of evidence and knowledge judgments, the proposition's truth value affected the evaluation of beliefs and decisions. Controlling for the influence of truth and knowledge, evidential evaluations did not affect the evaluation of beliefs and decisions. By contrast, controlling for the influence of truth and evidence, knowledge judgments still affected the evaluation of beliefs and decisions. Moreover, for the evaluation of both belief and decision, knowledge judgments mediated the effect of truth. Overall, these results support the view that truth and knowledge are both integrally connected to the ordinary evaluation of belief and decision; the results do not suggest that evidence is similarly connected.

6 Conclusion

According to factive accounts, the norms of assertion, belief, and decision-making require at the very least truth and perhaps even knowledge. If these views are correct, then you should not believe, assert, or make decisions based on a falsehood, even when the evidence suggests that it is not a falsehood. Critics claim that factive accounts badly mischaracterize our actual practice of evaluating assertions, beliefs and decisions. In short, factive accounts are said to be counterintuitive and revisionary. To support these claims, critics appeal to intuitions about cases where an agent forms a false belief based on highly reliable evidence or bases a decision on such a belief. If these claims are correct and cannot be explained away, then they count against the hypothesis that assertion, belief,and decision-making have factive norms.

But when I tested these claims, a dramatically different picture of our evaluative practices emerged. I tested cases where a known highly reliable source provides an agent with evidence that a certain proposition was true. When the proposition was true, the vast majority of people said that the agent should believe the proposition. But when the proposition was false, only a small minority said that the agent should believe the proposition. I also tested very similar cases where the agent will decide whether to pursue a course of action. When the proposition was true, the vast majority of people said that the agent should pursue the course of action. But when the proposition was false, only a small minority said that the agent should pursue the course of action.

Overall the results disprove the accusation that factive accounts are counterintuitive or revisionary. Instead the results cohere well with the hypothesis that the norms of belief and decision are factive; indeed, they are arguably the best evidence to date for factive norms of belief and decision-making. The effect of truth on these normative judgments was not only statistically significant but also extremely large. I

observed the same basic pattern of results across multiple scenarios, for men and for women, and for younger and older participants. These results parallel recent findings which support a factive norm of assertion (e.g. Turri 2013). They are also consistent with existing findings on people's intuitions about responsibility and punishment, where objective facts matter beyond what is reflected in the agent's evidence about the situation (Robinson and Darley 1998).

Like prior work criticizing factive norms, prior work supporting factive norms has relied largely on appeals to introspective and social observation (e.g. Hawthorne 2004; Stanley 2005; Hawthorne and Stanley 2008; Smithies 2012). Social and introspective observation are natural places to start when investigating the norm of an important activity that we're all familiar with. But some caution is warranted because social and introspective observation have some well known limitations. As social and cognitive psychologists have shown, not infrequently we misunderstand the dynamics of social interaction and the source of our own actions and reactions (Milgram 1974, pp. 103–104; Schwitzgebel 2008; Ross and Nisbett 2011; Lieberman 2013, pp. 4–5). Moreover, researchers have persistently disagreed over what is intuitive or characteristic of ordinary practice. And there is no hope that even the most astute introspection or social observation can discern causal mediation or precisely estimate the unique contributions that multiple factors make to an outcome.

Fortunately we can also systematically investigate the matter by established methods of experimental cognitive and social science. Controlled experimentation supplements introspection and social observation; it simultaneously builds on the insights they afford and overcomes their limitations, thereby increasing confidence that we have accurately identified the norm.

I also found evidence for an explanation of why critics of factive accounts have gotten things wrong. They seem to assume that the quality of someone's evidence for a proposition is insensitive to the proposition's truth value. They do not argue for this assumption; instead they claim that this is the natural and intuitive view. But the intuition is not widely shared. I asked people to evaluate the evidence someone got from a reliable source indicating that a certain proposition is true. When the proposition was actually true, people overwhelmingly said that the evidence was very good. But when the proposition was false, almost no one said that the evidence is very good; instead they were overall ambivalent on the evidence's quality. This finding has implications beyond the debate over factive accounts of the norms of assertion, belief and decision-making. For example, it also bears on the dispute between internalists and externalists in epistemology. In particular, it suggests that internalists have been relying on idiosyncratic intuitions about evidential quality when discussing, for example, notorious thought experiments about radically deceived subjects such as brains-in-vats (Lehrer and Cohen 1983; BonJour 2002; Feldman 2003; for further discussion and relevant results, see Turri in press b).

There are two main candidates for a factive norm of belief and decision-making: truth and knowledge. My findings help to discriminate between the two (Experiment 4). Two findings are especially relevant here. First, even when controlling for the influence of truth and evidence, knowledge judgments affect the evaluation of beliefs and decisions. Second, while truth has a large effect on these evaluations, knowledge judgments mediate its influence. A knowledge account of the norms of belief and decision

can explain the data. Knowledge requires truth, so the knowledge account predicts that truth will influence the evaluations and, moreover, it explains why knowledge judgments mediate truth's influence. But the truth account cannot easily explain the powerful independent influence that knowledge exerts on these evaluations. Neither can the truth account easily explain why knowledge mediates truth's influence. The knowledge account is a much better fit for the evidence on hand.

More general theoretical considerations also support knowledge accounts. It has been established that knowledge is the norm of assertion (for a review, see Turri under review c). If, as many researchers suspect, belief is best understood as assertion to oneself (e.g. Sellars 1963, p. 180; Dummett 1981, p. 362; see also Williamson 2000; Adler 2002), then the knowledge account of assertion entails a knowledge account of belief. Alternatively, suppose, as many other researchers expect, that belief is prior to assertion in the order of explanation and assertion is best understood as the expression of belief (e.g., Searle 1979; Bach and Harnish 1979). In that case, the knowledge account of belief could help explain why the knowledge account of assertion is true (Bach 2008). Either way, knowledge accounts of belief and assertion form a natural pair. Moreover, for my part, I am suspicious of the idea that one could be adequately positioned to properly believe and assert a proposition while, simultaneously, being inadequately positioned to make decisions based on it. That is, I find the following combination of claims counterintuitive: with respect to a certain proposition, you should believe it, and you should assert it, but you should not make decisions based on it. Further investigation of the matter is obviously required. But if my suspicion is well founded, then the knowledge account of decision-making comes along for the ride.

Nevertheless, however theoretically elegant and satisfying such a unified normative pantheon might be, the world is under no obligation to cooperate in such matters. Much more relevant is the empirical evidence presented here.

It might be argued that all my results show is that ordinary practice is deeply confused about how we *should* evaluate beliefs, decisions and evidence. On this approach, the results merely show that people are strongly disposed to incompetently evaluate beliefs, decisions and evidence. In response, this argument differs fundamentally from the criticism I address in the paper. The criticism I address is that factive accounts are revisionary, which is interpreted as evidence against factive accounts. The argument also rejects an assumption widely accepted in the literature: we should prefer theories that minimize the amount of error we read into ordinary behavioral patterns. Of course, critics might propose new arguments to think that factive accounts are very bad and, thus, that we should radically revise ordinary practice. Such attempts are to be welcomed and should be judged fairly on their merits. But lazy appeals to intuition and armchair generalizations about ordinary practice should be excluded from serious discussion and, indeed, actively discouraged.

In conclusion, the results from four experiments accomplished three main things. First, they undermined the most persistent objections to factive norms of belief and decision. Second, they provided the best evidence to date that the norms of belief and decision-making are factive. Third, they provided the best evidence to date that knowledge specifically is the norm of belief and decision-making.

Appendix 1: stimuli

Experiment 1 materials

(Watch: True/False) Maria is a watch collector who owns over ten thousand watches. She cannot keep track of all her watches by memory alone, so she maintains a detailed inventory of them. She keeps the inventory up to date. Maria knows that the inventory isn't perfect, but it is extremely accurate. ¶ [1] Today someone asked her, "Maria, I was wondering, do you have a 1990 Rolex Submariner in your collection?" ¶ Maria consults her inventory. It says that she does have a 1990 Rolex Submariner. [And this is just another case where the inventory is exactly right: she does have one in her collection. / But this is one of those rare cases where the inventory is wrong: she does not have one in her collection.].

(Name: True/False) Mario manages human resources for a company with over ten thousand employees. He cannot keep track of all their names by memory, so he maintains a detailed inventory of them. He keeps the inventory up to date. He knows that the inventory isn't perfect, but it is extremely accurate. ¶ Today someone asked him, "Mario, do you have an employee working for you named Rosanna Winchester?" ¶ Mario consults his inventory. It says that he does have an employee by that name (And this is just another case where the inventory is exactly right: he does have an employee by that name. / But this is one of those rare cases where the inventory is wrong: he does not have an employee by that name.).

Experiment 2 materials

The stories were exactly the same as in Experiment 1, except for the second paragraph. Below are the new second paragraphs.

(Watch) Today her partner informed her, "Maria, I just talked to the bank. If you have a 1990 Rolex Submariner in your collection, then you need to make an appointment to revise our insurance policy, which will take several hours. But if you don't have one, then you don't need to make an appointment."

(Name) Today his advisor informed him, "Mario, I just got a call from the immigration office. If we have an employee named Rosanna Winchester, then you need to make an appointment to revise the paperwork we filed, which will take several hours. But if we don't have one, then you don't need to make an appointment."

---

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

Appendix 2: experiment 4 analyses

A multivariate analysis of variance revealed main and interaction effects of Type and Truth Value on the dependent variables. So I analyzed the data for Believe and Decide conditions separately.

Believe conditions

In response to the question of what Maria should believe, mean response was significantly higher in True (M = 2.41, SD = 0.84) than False (M = 0.21, SD = 1.53), independent samples t test, t(115.56) = 10.87, p < .001. The magnitude of the mean difference (MD = 2.91, 95 % CI 1.79–2.60) was extremely large, d = 2.02. One sample t tests revealed that mean response was significantly above the neutral midpoint (= 0) in True, t(73) = 24.56, p < .001, MD = 2.41, 95 % CI 2.21–2.60, d = 2.87, whereas it did not differ from the midpoint in False, t(74) = 1.21, p = .230, MD = 0.21, 95 % CI −0.14 to 0.56. In True, 83 % of participants said that the agent should believe "I definitely do have one" or "I do have one," compared to only 17 % of participants in False. A χ^2 test for independence indicated that the magnitude of the difference in frequencies was extremely large, $\chi^2(1, 149) = 80.29$, Cramer's V = .734. The mode response in True was "I definitely do have one." In False it was "I probably do have one."

Mean evaluation of Maria's evidence was significantly higher in True (M = 2.39, SD = 0.77) than False (M = 0.68, SD = 1.47), independent samples t test, 112.27) = 8.90, p < .001. The magnitude of the mean difference (M = 1.71, 95 % CI 1.33–2.10) was extremely large, d = 1.68. One sample t tests revealed that mean evaluation was significantly above the neutral midpoint in True, t(73) = 26.61, p < .001, MD = 2.39, 95 % CI 2.21–2.57, d = 3.10, and also in False, t(74) = 4.0, p < .001, MD = .68, 95 % CI 0.34–1.02, d = 0.46. In True, 99 % of participants agreed that the evidence was good to some degree or other (i.e. "very good," "good," or "somewhat good"; in False, 64 % did. A χ^2 test for independence indicated that the magnitude of the difference in frequencies was medium, $\chi^2(1, 149) = 29.30$, Cramer's V = .443. The mode response in True was "very good." In False it was "good."

Mean knowledge score was significantly higher in True (M = 2.31, SD = 0.68) than False (M = 0.15, SD = 1.47), independent samples t test, t(104.76) = 11.57, p < .001. The magnitude of the mean difference (M = 2.164, 95 % CI 1.79–2.54) was extremely large, d = 2.25. One sample t tests revealed that mean knowledge score was significantly above the neutral midpoint in True, t(73) = 29.19, p < .001, MD 2.31, 95 % CI 2.15–2.47, d = 3.40, whereas it did not differ from the midpoint in False, t(74) = 0.87, p = .39, MD = .15, 95 % CI −0.19 to 0.48.

Decide conditions

In response to the question of what Maria should decide to do, mean response was significantly higher in True (M = 2.33, SD = 1.20) than False (M = −0.76, SD = 2.16, SEM = 0.25), independent samples t test, t(115.78) = 10.85, p < .001. The magnitude

of the mean difference (MD = 3.09, 95 % CI 2.53–3.65) was extremely large, d = 2.02. One sample t tests revealed that mean response was significantly above the neutral midpoint in True, t(75) = 16.86, p < .001, MD = 2.33, 95 % CI 2.05–2.60, d = 1.94, and it was significantly below the midpoint in False, t(74) = −3.05, p = .003, MD = −0.76, 95 % CI −1.26 to −0.26, d = 0.35. In True, 88 % of participants said that the agent "definitely should" or "should" make the appointment, compared to only 24 % in False. A χ^2 test for independence indicated that the magnitude of the difference in frequencies was extremely large, $\chi^2(1, 151) = 63.15$, Cramer's V = .647. The mode response in True was "definitely should." In False it was "should not."

Mean evaluation of Maria's evidence was significantly higher in True (M = 2.42, SD = 0.77, SEM = 0.09) than False (M = 0.28, SD = 1.77, SEM = 0.20), independent samples t test, t(100.87) = 9.63, p < .001. The magnitude of the mean difference (MD = 2.14, 95 % CI 1.7–2.58) was extremely large, d = 1.92. One sample t tests revealed that mean evaluation was significantly above the neutral midpoint in True, t(75) = 27.39, p < .001, MD = 2.42, 95 % CI 2.24–2.60, d = 3.14, but it did not differ from the midpoint in False, t(74) = 1.37, p = .174, MD = 0.28, 95 % CI −0.13 to 0.69. In True, 99 % of participants agreed that the evidence was good to some degree or other; in False, 56 % did. A χ^2 test for independence indicated that the magnitude of the difference in frequencies was large, $\chi^2(1, 151) = 39.42$, Cramer's V = .511. The mode response in True was "very good." In False it was "somewhat good."

Mean knowledge score was significantly higher in True (M = 2.08, SD = 1.03, SEM = 0.12) than False (M = −0.75, SD = 1.64, SEM = 0.19), independent samples t test, t(124.39) = 12.68, p < .001. The magnitude of the mean difference (MD = 2.83, 95 % CI 2.39–3.68) was extremely large, d = 2.27. One sample t tests revealed that mean knowledge score was significantly above the neutral midpoint in True, t(75) = 17.60, p < .001, MD = 2.08, 95 % CI 1.84–2.31, d = 2.02, and it was significantly below the neutral midpoint in False, t(74) = −3.95, p < .001, MD = −0.75, 95 % CI −1.12 to −0.37, d = 0.46.

References

Adler, J. E. (2002). *Belief's own ethics*. Cambridge, MA: MIT Press.

Austin, J. L. (1956). A plea for excuses. *Proceedings of the Aristotelian Society*, *57*, 1–30.

Bach, K. (2008). Applying pragmatics to epistemology. *Philosophical Issues*, *18*(1), 68–88.

Bach, K., & Harnish, R. M. (1979). *Linguistic communication and speech acts*. Cambridge, MA: MIT Press.

Benton, M. A. (2012). *Knowledge norms: Assertion, belief, and action*. New Brunswick: Rutgers University.

Benton, M. A. (2014). Knowledge norms. *Internet Encyclopedia of Philosophy*. Retrieved January 2014, from http://www.iep.utm.edu/kn-norms/

Bird, A. (2007). Justified judging. *Philosophy and Phenomenological Research*, *74*(1), 81–110. doi:10.1111/j.1933-1592.2007.00004.x.

BonJour, L. (2002). *Epistemology: Classic problems and contemporary responses*. Lanham: Rowman & Littlefield.

BonJour, L., & Sosa, E. (2003). *Epistemic justification: Internalism vs. externalism, foundations vs. virtues*. Malden, MA: Blackwell.

Brown, J. (2008). Knowledge and practical reason. *Philosophy Compass*, *3*(6), 1135–1152. doi:10.1111/j.1747-9991.2008.00176.x.

Chisholm, R. (1989). *Theory of Knowledge* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.

Coffman, E. J. (2014). Lenient accounts of warranted assertability. In C. Littlejohn & J. Turri (Eds.), *Epistemic norms: New essays on action, belief and assertion* (pp. 33–59). Oxford: Oxford University Press.

Cohen, S. (1988). How to be a fallibilist. *Philosophical Perspectives*, *2*, 91–123.

Cohen, S. (2013). Contextualism defended. In M. Steup, J. Turri, & E. Sosa (Eds.), *Contemporary debates in epistemology* (2nd ed., pp. 69–75). Malden, MA: Wiley.

Conee, E. (2007). Review of Jonathan Sutton, without justification. *Notre Dame Philosophical Reviews*, *12*. Retrieved, from http://ndpr.nd.edu/review.cfm?id=11803

DeRose, K. (1995). Solving the skeptical problem. *The Philosophical Review*, *104*(1), 1–52.

DeRose, K. (2009). *The case for contextualism*. Oxford: Oxford University Press.

Douven, I. (2006). Assertion, knowledge, and rational credibility. *Philosophical Review*, *115*(4), 449–485. doi:10.1215/00318108-2006-010.

Douven, I. (2008). Knowledge and practical reasoning. *Dialectica*, *62*(1), 101–118. doi:10.1111/j.1746-8361.2008.01132.x.

Dretske, F. I. (1981). *Knowledge and the flow of information*. Cambridge, MA: MIT Press.

Dummett, M. (1981). *Frege: Philosophy of language* (2nd ed.). Cambridge, MA: Harvard University Press.

Fantl, J., & McGrath, M. (2002). Evidence, pragmatics, and justification. *The Philosophical Review*, *111*(1), 67–94.

Fantl, J., & McGrath, M. (2007). On pragmatic encroachment in epistemology. *Philosophy and Phenomenological Research*, *75*(3), 558–589.

Fantl, J., & McGrath, M. (2009). *Knowledge in an uncertain world*. Oxford: Oxford University Press.

Fantl, J., & McGrath, M. (2013). Practical factors affect whether you know. In M. Steup, J. Turri, & E. Sosa (Eds.), *Contemporary debates in epistemology* (2nd ed., pp. 84–95). Hoboken: Wiley.

Feldman, R. (2003). *Epistemology*. Upper Saddle River, NJ: Prentice Hall.

Fletcher, G. P. (1978). *Rethinking criminal law*. Boston: Little Brown & Co.

Goldman, A. I. (1979). What is justified belief? In G. Pappas (Ed.), *Justification and knowledge*. Dordrecht: Reidel.

Goldman, A. I. (1993). Epistemic folkways and scientific epistemology. *Philosophical Issues*, *3*, 271–285.

Greco, J. (2010). *Achieving knowledge: A virtue-theoretic account of epistemic normativity*. Cambridge: Cambridge University Press.

Hawthorne, J. (2004). *Knowledge and lotteries*. Oxford: Oxford University Press.

Hawthorne, J., & Stanley, J. (2008). Knowledge and action. *Journal of Philosophy*, *105*(10), 571.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: Guilford Press.

Hill, C., & Schechter, J. (2007). Hawthorne's lottery puzzle and the nature of belief. *Philosophical Issues*, *17*, 102–122.

Huemer, M. (2007). Moore's paradox and the norm of belief. In S. Nuccetelli & G. Seay (Eds.), *Themes from G.E. Moore: New essays in epistemology and ethics*. Oxford: Oxford University Press.

Huemer, M. (2011). The puzzle of metacoherence. *Philosophy and Phenomenological Research*, *82*(1), 1–21.

Kvanvig, J. (2009). Assertion, knowledge, and lotteries. In D. Pritchard & P. Greenough (Eds.), *Williamson on knowledge* (pp. 140–160). Oxford: Oxford University Press.

Lackey, J. (2007). Norms of assertion. *Nous*, *41*(4), 594–626.

Lehrer, K., & Cohen, S. (1983). Justification, truth, and coherence. *Synthese*, *55*(2), 191–207.

Levin, J. (2008). Assertion, practical reason, and pragmatic theories of knowledge. *Philosophy and Phenomenological Research*, *76*(2), 359–384.

Lieberman, M. D. (2013). *Social: Why our brains are wired to connect*. New York: Crown Publishers.

Littlejohn, C. (2013). The Russellian retreat. *Proceedings of the Aristotelian Society*, *113*(3pt3), 293–320. doi:10.1111/j.1467-9264.2013.00356.x.

Locke, J. (1690/1975). In P. H. Nidditch (Ed.), *An essay concerning human understanding*. Oxford: Clarendon Press.

MacIver, A. M. (1938). Some questions about 'know' and 'think'. *Analysis*, *5*(3), 43–50.

McGlynn, A. (2013). Believing things unknown. *Nous*, *47*(2), 385–407.

Milgram, S. (1974). *Obedience to authority: An experimental view*. New York: Harper Perennial.

Montminy, M. (2013). Why assertion and practical reasoning must be governed by the same epistemic norm. *Pacific Philosophical Quarterly*, *94*(1), 57–68.

Moore, G. E. (1959). *Philosophical papers*. New York: Collier Books.

Noveck, I. A., & Sperber, D. (2004). *Experimental pragmatics*. Basingstoke: Palgrave Macmillan.

Nozick, R. (1981). *Philosophical explanations*. Cambridge, MA: Harvard University Press.

Pritchard, D. (2014). Epistemic luck, safety, and assertion. In C. Littlejohn & J. Turri (Eds.), *Epistemic norms: New essays on action, belief and assertion*. Oxford: Oxford University Press.

Reed, B. (2013). Practical matters do not affect whether you know. In M. Steup, J. Turri, & E. Sosa (Eds.), *Contemporary debates in epistemology* (2nd ed., pp. 95–106). Malden, MA: Wiley.

Reid, T. (1764/1997). In D. R. Brookes (Ed.), *An inquiry into the human mind on the principles of common sense*. University Park, PA: Pennsylvania State University Press.

Robinson, P. H. (1982). Criminal law defenses: A systematic analysis. *Columbia Law Review*, *82*(2), 199–291.

Robinson, P. H., & Darley, J. M. (1998). Testing competing theories of justification. *North Carolina Law Review*, *76*, 1095–1145.

Ross, L., & Nisbett, R. E. (2011). *The person and the situation: Perspectives of social psychology*. London: Pinter & Martin.

Schwitzgebel, E. (2008). The unreliability of naive introspection. *Philosophical Review*, *117*(2), 245–273. doi:10.1215/00318108-2007-037.

Searle, J. R. (1979). *Expression and meaning*. Cambridge: Cambridge University Press.

Sellars, W. (1963). *Science, perception and reality*. Atascadero, CA: Ridgeview Publishing Company.

Smith, M. (2012). Some thoughts on the JK-rule. *Nous*, *46*(4), 791–802.

Smithies, D. (2012). The normative role of knowledge. *Nous*, *46*(2), 265–288. doi:10.1111/j.1468-0068.2010.00787.x.

Stanley, J. (2005). *Knowledge and Practical Interests*. Oxford: Oxford University Press.

Stroud, B. (1984). *The significance of philosophical skepticism*. Oxford: Clarendon Press.

Sutton, J. (2007). *Without justification*. Cambridge: MIT Press.

Turri, J. (2010). Prompting challenges. *Analysis*, *70*(3), 456–462. doi:10.1093/analys/anq027.

Turri, J. (2011). The express knowledge account of assertion. *Australasian Journal of Philosophy*, *89*(1), 37–45. doi:10.1080/00048401003660333.

Turri, J. (2013). The test of truth: An experimental investigation of the norm of assertion. *Cognition*, *129*(2), 279–291. doi:10.1016/j.cognition.2013.06.012.

Turri, J. (2014a). Knowledge and suberogatory assertion. *Philosophical Studies*, *167*(3), 557–567. doi:10.1007/s11098-013-0112-z.

Turri, J. (2014b). Selfless assertions: Some empirical evidence. *Synthese*. doi:10.1007/s11229-014-0621-0.

Turri, J. (2015). Knowledge and the norm of assertion: A simple test. *Synthese*, *192*(2), 385–392. doi:10.1007/s11229-014-0573-4.

Turri, J. (in press a). Sustaining rules: A model and application. In J. A. Carter, E. C. Gordon, & B. Jarvis (Eds.), *Knowledge first*. Oxford: Oxford University Press.

Turri, J. (in press b). The radicalism of truth-insensitive epistemology: Truth's profound effect on the evaluation of belief. *Philosophy and Phenomenological Research*.

Turri, J. (under review a). *Knowledge and assertion in Gettier cases*. Waterloo: University of Waterloo.

Turri, J. (under review b). *Vision, knowledge, and assertion*. Waterloo: University of Waterloo.

Turri, J. (under review c). *Knowledge and the norm of assertion: An essay in philosophical science*. Waterloo: University of Waterloo.

Turri, J., & Blouw, P. (2015). Excuse validation: A study in rule-breaking. *Philosophical Studies*, *172*(3), 615–634. doi:10.1007/s11098-014-0322-z.

Turri, J., Friedman, O., & Keefner, A. (under review). *Knowledge central: a central role for knowledge attributions in social evaluations*. Waterloo: University of Waterloo.

Unger, P. (1975). *Ignorance: A case for skepticism*. Oxford: Oxford University Press.

Vogel, J. (1990). Are there counterexamples to the closure principle? In M. D. Roth & G. Ross (Eds.), *Philosophical studies series* (Vol. 48, pp. 13–27). Berlin: Springer. doi:10.1007/978-94-009-1942-6_2.

Weiner, M. (2005). Must we know what we say? *Philosophical Review*, *114*(2), 227–251.

Williamson, T. (2000). *Knowledge and its limits*. Oxford: Oxford University Press.

Wright, S. (2014). The dual-aspect norms of belief and assertion. In C. Littlejohn & J. Turri (Eds.), *Epistemic norms: New essays on action, belief and assertion* (pp. 239–258). Oxford: Oxford University Press.