TRUTH-SENSITIVITY AND
FOLK EPISTEMOLOGY

Mikkel Gerken
University of Southern Denmark

Abstract. Several studies have found a robust effect of truth on epistemic evaluation of belief, decision, action and assertion. Thus, truth has a significant effect on normative participant evaluations. Some theorists take this truth effect to motivate factive epistemic norms of belief, action, assertion etc. In contrast, I argue that the truth effect is best understood as an epistemic instance of the familiar and ubiquitous phenomenon of outcome bias. I support this diagnosis from three interrelating perspectives: (1) by epistemological theorizing, (2) by considerations from cognitive psychology and (3) by methodological reflections on the relationship between folk epistemology and epistemological theorizing.

1: INTRODUCTION. A series of studies have unearthed a robust effect of truth on epistemic evaluation of belief, decision, action and assertion (Turri 2013, 2015a–c, 2016a–b, 2017a–b; Turri and Buckwalter 2017). But while the studies provide compelling evidence for a truth effect on our folk epistemological assessments, they raise difficult epistemological and methodological questions. It is natural to take the evidence of a truth effect as evidence for factive epistemic norms – notably knowledge norms – of action, assertion and belief. Thus, the findings may appear to empirically settle prominent epistemological debates. In contrast, I will set forth a
novel empirical account of the truth effects. Specifically, I will argue that they are reasonably understood as manifesting an epistemic instance of a general outcome bias. If this diagnosis is correct, non-factive epistemic norms are not threatened by the findings. Apart from the debate over epistemic norms, we want clarity about whether robust patterns of folk epistemological judgments are biased or not. Thus, my main central aim is simply to motivate a novel and independently significant diagnosis of the truth effects on epistemic assessment.

After considering the evidence for a truth effect (Section 2), I motivate my diagnosis from several mutually reinforcing perspectives. First, I argue on epistemological grounds (Section 3). Second, I argue on empirical grounds (Section 4–5). Third, I argue on methodological grounds (Sections 6 and 7).

2: THE EVIDENCE FOR A TRUTH EFFECT ON EPISTEMIC ASSESSMENT.

The evidence for a truth effect on folk epistemological assessment of belief, action, assertion and decision comes from an impressive range of experiments that cannot be responsibly addressed in a single article (Turri 2013, 2015a–c, 2016a, b, 2017a, b; Turri and Buckwalter 2017). In consequence, I will primarily consider the truth effect on the epistemic assessment of belief (focusing Turri 2016a) and briefly relate the discussion to the findings concerning assertion and action.

2.1. Truth effects on belief. In order to provide evidence for a pervasive and robust truth effect, Turri measured participant evaluations of six kinds of normative vocabulary—namely, ‘evidence,’ ‘justification,’ ‘should,’ ‘rationality,’ ‘reasonableness,’ and ‘responsibility’ (Turri 2016a).

In a between-subject study design, participants considered cases divided in pairs that were alike in every regard with the exception of the truth-value of the protagonist’s belief. Thus, any significant difference in epistemic assessment would provide evidence for a truth effect. For example, participants considered cases such as the following:

(Watch)
Maria is a watch collector who owns over ten thousand watches. She cannot keep track of all her watches by memory alone, so her accountant maintains a detailed inventory of them. Maria knows that the inventory isn’t perfect, but it is extremely accurate. Someone asks Maria, “Do you own a 1990 Rolex Submariner?” Maria consults the inventory and it says that she does have one. In fact, she [does/does not] have one.

This article is protected by copyright. All rights reserved
Each participant responded to exactly one of the following evaluation types:

1. What does Maria’s evidence justify her in believing?
2. What is Maria justified in believing?
3. What should Maria believe?
4. What is it rational for Maria to believe?
5. What is the responsible thing for Maria to believe?
6. What is it reasonable for Maria to believe?

Participants responded on a 6-point Likert scale anchored with “I definitely have one” (3) and “I definitely do not have one” (-3). Turri found a strong effect of truth across evaluation types (M = 1.95 vs. M = 0.39). See Turri 2016a: 353 for details).\(^1\) A similar study design concerning the subject’s belief that a banging sound was a gunshot produced comparable results, and this indicates that the finding is not particular to the details of the vignette (Turri 2016a: 353). An earlier study yielded similar results insofar as the medians on a 7-point Likert scale anchored with -3 and 3 were M = 2.03 in true conditions vs. M = 0.01 in false conditions (Turri 2015c: 4013). In these studies, measures of effect magnitude revealed that “The size of truth’s effect was extremely large.” (Turri 2016a: 354; 2015c: 4013). Furthermore, Turri found a strong truth effect in a case of massive deception (a brain-in-a-vat case), and in this case, he even found a truth effect in a within-subject study design (Turri 2016a: 354ff).

Linear regression analysis revealed that belief evaluation was predicted by truth, although an even stronger predictor was knowledge scores (Turri 2015c: 4021. See also Turri and Buckwalter 2017). A further experiment indicated that “Even controlling for the influence of evaluations of evidence and knowledge judgments, the proposition’s truth value affected the evaluation of beliefs and decisions” (Turri 2015: 4023). However, the effect of truth on epistemic evaluations was found to be mediated by judgments about knowledge: “while truth has a large effect on these evaluations, knowledge judgments mediate its influence.” (Turri 2015c: 4024)

In sum, several studies indicate a robust significant effect of truth on epistemic evaluation of belief in several vignettes and across a range of types of evaluation, and some evidence suggests that a knowledge norm is best suited to explain this (Turri 2015: 4025).

2.2. Further truth effects—decision, action and assertion. In other work, the truth effect has been found for decision and actionability (Turri 2015c; Turri and Buckwalter 2017). In a study

---

\(^1\) I will only report statistically significant findings and refer to the original papers for the statistical analysis.

This article is protected by copyright. All rights reserved
design similar to the one discussed above, participants were asked to assess whether the protagonist should make a decision based on a true or false proposition on a 6-point Likert scale anchored with “definitely should” (3) and “definitely should not” (-3)” (Turri 2015c: 4015). As in the case of belief, the mean response was significantly higher in true than in false conditions (M = 2.50 vs. M = −0.53. See Turri 2015c: 4014). Linear regression analysis revealed the truth value to be the strongest predictor of evaluation of decision and action (Turri 2015c). On the basis of these findings, Turri argues that “Changing just the truth value of the source’s information on the present occasion radically changes what people say the agent should decide.” (Turri 2015c: 4013).

Similar effects have been found for epistemic assessments of assertion in a variety of studies (Turri 2013; 2015a. For an overview, see Turri 2017). For example, in a variety of (Watch), Turri found that “over 80% of participants in the False condition responded that Maria should not make the statement compared to only 3% in the True condition” (Turri 2013: 283). Several studies focus on knowledge and provide evidence that knowledge is a better predictor than truth of proper belief, action and assertion (Turri 2015a, c; Turri, Friedman & Keefner 2017, Turri and Buckwalter 2017). For example, Turri concludes a survey by taking a wider set of evidence to support a knowledge norm of assertion: “In summary, convergent evidence from animal communication studies, developmental findings on human children, observations of patterns in everyday discourse, and experiments with human adults all strongly support the conclusion that knowledge is a central norm of the social practice of assertion” (Turri 2017: 5).

I should note two things in this regard: First, given the factivity of knowledge, knowledge norms are a subset of factive norms. Any argument against factive norms will a fortiori be an argument against knowledge norms.2 So, I will consider the truth effect in its own right. Second, rather than addressing all the relevant evidence, my restricted aim is to examine whether the specific findings of a truth effect on epistemic evaluation of belief, action and assertion support factive epistemic norms (including knowledge norms) thereof. As noted, my primary focus will be on the case of belief. However, I note the findings of similar effects for decision, action and assertion because they bolster the assumption that the truth effect on epistemic evaluation is pervasive and robust.

2.3. Initial objections and responses. Initial responses have questioned whether such studies really indicate a truth effect on epistemic evaluation by suggesting that they do not reflect

---

epistemic requirements but rather requirements concerning what is moral, prudential or polite to assert (Kvanvig 2011; Pagin 2015). One thing to note is that this objection does not appear to be fully general since it concerns assertability. Moreover, Turri has augmented the findings of a truth effect by studies designed to rule out such pragmatic confounders.

To do so, Turri had participants consider simple scenarios and recorded their judgments about assertability, truth value, morality, rationality, etiquette, and legality (Turri 2017b). After controlling for the influence of these further factors, truth remained the strongest predictor of assertability attributions. This finding runs contrary to the assumption that the truth effect only reflects non-epistemic concerns having to do with conversational pragmatics. In contrast, the finding is what should be expected if the truth effects reflect folk epistemological concerns. So, the study does not only compromise confounding explanations, it also lends positive support to the postulation of a truth effect on epistemic evaluation.

Overall, then, Turri’s findings provide strong evidence of a robust and general effect of truth on folk epistemological assessments. Thus, I will, in what follows, accept the existence of a truth effect as an important aspect of our folk epistemology.4

2.4. The epistemological significance of the findings. My criticism differs from responses that question whether the sketched body of empirical work really indicates a truth effect on genuinely epistemic assessment. In contrast, I accept that the findings provide impressive evidence for a robust truth effect on folk epistemical assessment. What I will challenge is rather the step from the assumption that there is a robust truth effect on folk epistemic assessment to any strong epistemological conclusion.

In one place, Turri concludes that “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.” (Turri 2015c: 4026). In Turri 2016a, a central conclusion is drawn in terms of ‘truth-sensitivity’ insofar as Turri’s title suggests that truth-insensitive theories are radical. However, in the present context, it is critical to be careful about the precise meaning of ‘truth-sensitive.’ So, I will initiate the discussion of the epistemological debates that the findings bear on by considering the notion of truth-sensitivity.

---

3 See also Turri 2013, Experiments 2 and 3.
4 One can always quarrel with details. For example, the subjects in some of Turri’s cases only possessed rather weak warrant. It would be interesting to see whether the effect of truth would be diminished in probes where the subjects are extremely well warranted. For further criticism, see (Brössel and Reuter forthcoming)
3: TRUTH-SENSITIVITY IN EPISTEMOLOGY.

In this section, I will consider the notion of truth-sensitivity and its role in epistemology. In particular, I will argue that there is a non-factive notion of truth-sensitivity that is central to externalist epistemology.

3. Simple and subtle truth-sensitivity. Whereas most epistemologists hold that epistemic rationality is constitutively sensitive to truth, disputes arise over the nature of truth-sensitivity. Fixing some terminology may help clarify the dialectics of such disputes. I will use the term ‘warrant’ to denote the genus of epistemic rationality. ‘Entitlement’ will denote its externalist species and ‘justification’ will denote its internalist species. The idea of truth-sensitivity calls for yet further terminology. Turri targets “truth-insensitive theories of core epistemological properties such as justification and rationality” according which belief is “…insensitive to truth in a very specific way. Suppose that two possible agents believe the same proposition based on the same evidence. Either both are justified or neither is; either both have good evidence for holding the belief or neither does. This does not change if, on this particular occasion, it turns out that only one of the two agents has a true belief.” (Turri 2016a: 348).

Since the dichotomy between truth-sensitivity and truth-insensitivity may be too simple, it is worth drawing a trichotomous distinction. Some theorists require that a belief must be true to be rational in any genuinely epistemic sense. A prominent brand of such a factivist approach is the knowledge-first approach (Williamson 2000; Sutton 2007; Turri 2016a). Let us call such theories simple truth-sensitive theories due to the simple requirement that a token belief must be true to be warranted.

Other epistemologists only require a looser truth-connection, and these include both epistemic internalists and epistemic externalists. For example, Goldman does not require that a warranted token belief be true but only that it be generated by a reliable (truth-sensitive) type of cognitive process (Goldman 1979). Generally, many epistemologists do not identify the truth-sensitivity requirement with simple truth (factivity) but with a more subtle albeit still constitutive truth-connection. So, the many varieties of such views are aptly named subtle truth-sensitive theories.

Finally, we may recognize views according to which truth is not constitutively associated with warrant. Turri cites Chisholm as holding the view that “there is no connection between

---

5 The terminology derives from Burge 2003. See also Pedersen and Graham forthcoming; Gerken 2013b, forthcoming.
what an agent should believe and truth” (Chisholm 1989: 76 cited in Turri 2016a: 350).⁶ Theories according to which there is no constitutive connection between epistemic rationality and truth are appropriately labeled truth-insensitive theories.

This addition to Turri’s terminology helps ensure that we do not take arguments against truth-insensitive views as arguments against subtly truth-sensitive views or for simple truth-sensitive views. I do not suggest that Turri commits this fallacy. But since the dichotomous distinction may encourage it, the more fine-grained terminology will serve us better.

In fact, yet more fine-grained distinctions may be called for since the category of subtly truth-sensitive theories harbors both epistemic internalists and epistemic externalists (Gerken 2013b, 2017b, 2018). While I will not tire the reader with further terminology, it is important to emphasize that many epistemically externalist views are subtly rather than simply truth-sensitive (Dretske 1971; Goldman 1979, 1986; Burge 2003; Sosa 2007; Graham 2012). Although simple truth-sensitive theorists (including Turri) are well aware of this, they often target epistemic internalists. For example, in setting up the dialectic, Turri cites epistemic internalists such as (Cohen 1984; Chisholm 1989; BonJour & Sosa 2003; Conee and Feldman 2004 and Fumerton 2006) but none of the mentioned externalists. So, it is worth highlighting the epistemically externalist tradition of subtle truth-sensitivity to avoid a false dilemma between epistemic internalist versions of subtle truth-sensitivity and simple truth-sensitivity.⁷

Thus, my aim is only to argue that Turri’s important findings do not favor simple truth-sensitive theories over epistemically externalist subtle truth-sensitive theories. So, I will briefly specify the view that I will defend.

3.2. Epistemically externalist subtle truth-sensitivity. The epistemically externalist brand of a subtly truth-insensitive account of warranted belief may be illustrated from a number of angles. I will start with a case of one-off perceptual illusion:

**WAX APPLE**

When Ivan arrives at his daughter’s day care, there are bowls of fruit on the tables. However, a red wax apple, a, has been misplaced in a fruit bowl that does not contain any real red apples. The wax apple is visually indiscriminable from a real red apple. The light that hits Ivan’s retina is photon-to-photon identical to the

---

6 I will not weigh in on the exegetical question concerning whether Chisholm holds a truth-insensitive or a subtly truth-sensitive view.
7 In (Gerken 2017b), I argue that the debate between (Dougherty and Rysiew 2013) and (Williamson 2013) suggests such a false dilemma. Furthermore, many externalists will reject Turri’s “based on the same evidence” formulation because they reject evidentialism.
light that would have hit his retina had there been a real red apple in the bowl. When a child asks for a red apple, Ivan spots a, and on this basis, he forms the belief that there is a red apple in the bowl.

The rest of the story involves surprise, tears and, in the end, good laughs. But the question before us is the epistemic status of Ivan’s belief that there is a red apple in the bowl at the time just after its formation. According to subtle truth-sensitive theories of the externalist kind, Ivan is as warranted (entitled and justified) in his belief as Ivana who spots a real red apple on the adjacent table and forms a similar belief.

The apple-belief in question is the only false fruit-related belief that Ivan entertains. So, the case exemplifies a one-off false belief explained in terms of an abnormal external circumstance. Of course, the case indicates that Ivan’s fruit-recognizing competence is fallible. But this is no reason to claim that he is unwarranted. His true belief that there is a banana in the bowl is equally fallible. As is Ivana’s true belief that there is a red apple in the bowl on the next table. Indeed, epistemic externalists think that accounting for warranted false beliefs is a desideratum because it helps explain that we rely on fallible cognitive competences even when we acquire true belief and knowledge. But fallible competences sometimes fail. When this is due to indiscernible highly local environmental abnormalities, warrant is typically preserved. Thus, it would be just as epistemically irrational of Ivan to suspend his false belief that there is a red apple in the bowl as it would be to suspend his true belief that there is a yellow banana in the bowl.

The account is epistemically externalist in the positive sense that the belief type is objectively truth-conducive in the epistemically normal environment even though this token belief is false due to an environmental abnormality. The account is also externalist in virtue of centrally invoking objective facts about general regularities in the environment Ivan is embedded in. A thing that looks precisely like a red apple is in fact a red apple in the vast majority of the cases in Ivan’s environment. This objective general fact about Ivan’s environment helps explain why his apple-belief is objectively, but subtly, truth-sensitive.8

Finally, the account is epistemically externalist in the negative sense that it is not required that Ivan be aware or have first-person reflective access to the epistemically relevant facts about the environment or his cognitive competences. It is enough that his cognitive competences are in fact subtly truth-sensitive relative to his normal environment (Burge 2003; Gerken 2013a, 2018, forthcoming). Truth is no less constitutive of epistemic warrant in an account cast in terms of

8 Given this explanatory role of the general external environment, the account should be integrated with externalism about mental states (Burge 2003; Graham 2012; Gerken 2013a–b, 2018, forthcoming).
truth-conduciveness than in one cast in terms of truth *simpliciter*. So, charges that subtle truth-sensitive accounts must be epistemically internalist are mistaken (Gerken 2017b, 2018, forthcoming).

I do not take this externalist account to align perfectly with our folk epistemology. Indeed, it may clash with even robust patterns of folk epistemological judgments (Gerken 2017a). I will argue that Turri’s empirical findings exemplify one such clash. But first cases of massive deception should be briefly addressed.

### 3.3. One-off deception vs. massive deception:

In characterizing truth-insensitive theories, Turri makes the following suggestion: “Epitomizing this line of thought are thought experiments about radically deceived “brains in vats.” It is widely and uncritically assumed that such a brain is equally justified as its normally embodied human “twin.”” (Turri 2016a: 348).

However, especially epistemically externalist proponents of a subtle truth-sensitivity often claim that the subject is warranted in one-off illusion cases but not in massive deception cases. I have devoted a full paper to this point (Gerken 2018). So, rather than repeat myself, I will just regurgitate the point that is most central for the present purpose: It is compatible with subtle truth-sensitivism to reject that a subject in a brain-in-vat (BIV) case is as warranted as one in an ordinary case of veridical perception.

Consequently, I restrict the discussion to the assumption that cognizers, such as Ivan in the one-off case *WAX APPLE*, may be as warranted in one-off illusions as their twins (e.g., Ivana) who are in veridical counterpart cases. Since factivist epistemologists target this view, it is one worth defending. So, I will focus on one-off cases while setting aside the harder cases of massive deception (but see Gerken 2013a, 2018).

### 3.4. Standards and norms:

Subtle truth-sensitivists do not reject that a belief is epistemically flawed in a fundamental way when it is false. They just claim that truth is distinct from the kind of truth-sensitivity that warrant requires. Even simple truth-sensitivists recognize this much insofar as they deny that all true beliefs are warranted. For example, simple truth-sensitivists do not regard a true belief formed on the basis of wishful thinking as warranted. Subtle truth-sensitivists insist that the flip-side of this insight must also be acknowledged: Just as some true beliefs are unwarranted, some subtly truth-sensitive beliefs are warranted.

Thus, epistemic norms governing epistemic warrant are constitutively associated with truth but they are distinct from truth itself. This point may be brought out by Thomson’s distinction between standards and norms (Thomson 2008 Ch. IV). Truth is the epistemic
standard of belief, and the norms that govern epistemic rationality are constitutively related to the standard of truth but in a more complex manner. For example, whether a belief meets the epistemic norms has to do with whether it is being truth-conducive in virtue of the exercise of a cognitive competence. Adopting Thomson’s terminology, a belief is epistemically correct when meeting the epistemic standard of truth and epistemically sound when meeting the subtle truth-sensitive norm.9

Now subtle truth-sensitivists may more precisely characterize wherein a false warranted belief—such as a one-off illusion—is and is not epistemically flawed. It is incorrect in virtue of failing to meet the standard of truth, but it is sound in virtue of meeting the more complex epistemic norm that pertains to truth-conduciveness in virtue of competence. Likewise the epistemic properties of a lucky guess may be characterized as correct but unsound.

3.5: Not about the blame. The philosophical rationale for subtle truth-conduciveness does not centrally involve appeal to blamelessness, excusability or responsibility. Epistemic warrant and epistemic norms concern objective truth-conduciveness (in virtue of competence) and not merely blamelessness or excusability. At least epistemically externalist brands of subtle truth-sensitivity may hold that one may be blameless or excused for beliefs that are not truth-conducive but one is not thereby warranted (Gerken 2013a, 2017a).

Of course, ascriptions of blame and epistemic warrant sometimes go hand in hand. After all, the present view is compatible with the idea that blamelessness is necessary for—and hence not irrelevant for—warrant. But the distinction between blamelessness and warrant is recognizable by systematic reflection on cases. In one set of cases, S is blameless for failing to meet the epistemic norm. In such cases, S is not warranted. In another set of cases, S is blameless for failing to meet the epistemic standard of truth. In such cases, S may remain warranted. These case types are different, but it requires systematic reflection to recognize that this is so. Consequently, blamelessness and warrant may often be run together in our folk epistemology. So, this is another area in which the theoretical view and the patterns of folk epistemological assessment may diverge.10

3.6: Epistemically externalist subtle truth-sensitivity in summary. The brisk tour of an epistemic externalist version of subtle truth-sensitivity is only a rough sketch and not meant as an

---

9 One need not adopt every aspect of Thomson’s 2008 account. One might also draw a further distinction between epistemic norms and epistemic guidelines as I do in (Gerken 2017a, Ch. 6) where this framework is developed.

10 Turri has also found important empirical data concerning excusability. But since the interpretation about this data set raises related but further issues of its own, it would be irresponsible to also take on this here.
argument that will convince simple truth-sensitivists. Rather, it is meant to highlight key aspects of the subtle truth-sensitive view that I will defend and as a sketch of the basic rationales for it.

4: THE EMPIRICAL CASE FOR OUTCOME BIAS. In this section, I will briefly present some of the empirical literature on outcome bias – roughly, the effect of knowledge of an outcome on the evaluation of the decision or procedure leading to it (Baron and Hershey 1988; Mazzocco et al. 2004). I will also briskly note the closely related phenomenon of hindsight bias (Fishhoff 1975, Hawkins and Hastie 1990; Guilbault et al. 2004) and, even more briskly, the phenomenon of fundamental error attribution (Ross 1977; Stalder 2009).

4.1. Outcome bias introduced: Recognition of outcome bias has been traced back to at least Arnould (1662/1964: 285). However, since the 1970s, empirical work on the phenomenon has accelerated. In 1974, Brown et al. noted that “Since good decisions can lead to bad outcomes (and vice versa), decision makers cannot fallibly be graded by their results (Brown, Karr, Peterson 1974: 4). Similarly, Edwards remarked that “A good decision cannot guarantee a good outcome. All real decisions are made under uncertainty. A good decision is therefore a bet, and evaluating it as good or not must depend on the stakes and odds, not on the outcome” (Edwards 1984: 7).

These considerations set the stage for a seminal paper on outcome bias (Baron and Hershey 1988). In a series of studies, they found evidence that outcome has an effect on participants’ assessments of “…the quality of thinking of the decisions, the competence of the decision maker, or their willingness to let the decision maker decide on their behalf” (Baron and Hershey 1988: 569).

Baron and Hershey’s participants evaluated the decision to operate on a 55-year-old man with a heart condition where there was an 8% risk of him dying from the operation. Participants in both the positive and negative outcome conditions evaluated the decision to operate on a 7-point Likert scale (anchored with 3: Clearly correct, the opposite decision would be inexcusable and -3: Incorrect and inexcusable). In this case and a range of variations, Baron and Hershey found that “Overall there was an outcome bias. Cases in which the outcome was success… …were rated higher than matched cases in which the outcome was failure” (Baron and Hershey 1988: 572).

The finding is replicated with many variations of the stimuli. Consequently, outcome bias is widely regarded as a central bias in human cognition. For example, it was included in
Stanovich and West’s study that sought to correlate cognitive ability with what they regard as “…some of the most classic and well-studied biases in the heuristics and biases literature” (Stanovich and West 2008: 672). Apart from outcome bias, they included biases such as the conjunction effect, framing effects, anchoring effects, base-rate neglect, myside bias, sunk-cost effect etc. (Stanovich and West 2008; Stanovich 2011).

In Stanovich and West’s replication, participants in the negative outcome condition “evaluated a medical decision that was designed to be objectively better than the first: 2% chance of death rather than 8%; 10-year increase in life expectancy versus 5-year increase, etc. However, it had an unfortunate negative outcome—death of the patient” (Stanovich and West 2008: 673). Participants in both conditions evaluated the decision to operate on a 7-point Likert scale anchored by ‘1: Incorrect, a very bad decision’ and ‘7: Clearly correct, an excellent decision.’

Stanovich and West found that both high and low cognitive ability groups “displayed the expected outcome bias—the decision with the positive outcome was rated as a better decision than the decision with the negative outcome, despite the fact that the latter was objectively better.” (Stanovich and West 2008: 675). However, “the low-SAT group displayed more outcome bias.” (Stanovich and West 2008: 675). Such correlations may be taken as indirect evidence that the outcome effect is indeed a bias (Stanovich 2011).

Outcome bias may be found with many manipulations of outcome. For example, Mackie and colleagues held fixed the description of a teacher facing a test as well as her actual performance (a success rate of 69% or 71%). However, half the participants were informed that a 75% success rate was required to pass whereas the other half were informed that passing only required a 65% success rate (Mackie et al. 2001: 77). Participants supportive of mandatory testing of teachers evaluated the teacher better on properties such as intelligence and competence in the passing conditions than in the failing conditions (Mackie et al. 2001: 79).

More generally, evidence for outcome bias has been found in fields as diverse as the medical sciences (Berlin 2004), management studies (Marshall and Mowen 1993; Sezer et al. 2016), finance and business (Fisher and Selling 1993; Tan and Lipe 1997), legal studies (Clarkson et al. 2002), social cognition (Allison et al. 1996; Mackie et al. 2001; Mazzocco et al. 2004) etc. Thus, Stanovich and West appear to be on solid ground when they regard it as among the most well-studied biases (Stanovich and West 2008).

11 Interestingly, the participants opposed to mandatory teacher testing evaluated the teacher in a manner independent of the outcome. Mackie et al. take this as evidence that motivation is central to outcome bias. But the point here is merely that outcome bias may occur for many outcome manipulations.
4.2. Hindsight bias and the curse of knowledge. Outcome bias belongs to a family of biases sometimes subsumed under the heading ‘curse of knowledge’ (Stanovich 2011: 149; Roese and Vohs 2012). These are phenomena that occur when knowledge of an outcome impacts the assessment of the objective properties of a decision or procedure even though the study is designed to render the outcome irrelevant to the assessment.

Whereas outcome bias concerns the evaluation of an individual or procedure responsible for the outcome, classic hindsight bias studies showed that knowledge of the outcome causes participants to overestimate its probability (Fischhoff 1975). This may lead people to think that they knew the outcome before it occurred—a “knew-it-all-along” effect (Fischhoff and Beyth 1975; Bradfield and Wells 2005; Blank et al. 2008).

A seminal study by Fischhoff made use of the methodology which has become the paradigm in the study of hindsight bias: “… the Before group read a brief (150 word) description of a historical or clinical event for which four possible outcomes were provided. The After groups read identical passages to which a final sentence presenting one of the possible outcomes as the “true” outcome had been added… … Subjects in all groups were asked to (a) estimate the likelihood of occurrence of each of the four possible outcomes, and (b) evaluate the relevance of each datum in the event description.” (Fischhoff 1975: 289). Fischhoff found was that the likelihood of the outcome was estimated to be higher when it was said to obtain.

The finding has been replicated with wide variations in the stimuli. A 2012 discussion estimated hindsight bias to “…having been featured in more than 800 scholarly papers” across multiple disciplines (Roese and Vohs 2012: 411). Hence, hindsight bias is widely regarded as ubiquitous (see Hawkins and Hastie 1990 for a survey and Guilbault et al. 2004 for a meta-analysis).

Whereas there is little skepticism about the ubiquity or robustness of hindsight bias, it is discussed how to taxonomize varieties of the phenomenon (Bradfield and Wells 2005; Blank et al. 2008). Likewise, the exact relation between hindsight bias and outcome bias is debated (Baron and Hershey 2008; Clarkson et al. 2001). But cognitive scientists widely accept that hindsight bias and outcome bias share the common denominator that knowledge of an outcome that is irrelevant for some assessment nevertheless affects it. So, according to a broad consensus in cognitive science, there is strong empirical evidence for the overarching “curse of knowledge” phenomenon according to which knowledge of the outcome may unduly influence our intuitive assessments.  

I conjecture that the same effect will be found for belief. If so, ‘curse of belief’ might be more accurate.
4.3: Fundamental attribution error. A final psychological phenomenon that I want to call attention to (although that's all space permits for) is fundamental attribution error. Roughly, this is the tendency to underestimate situational factors and overestimate features of individuals in our explanations (Ross 1977; Gilbert and Malone 1995).

Central evidence for fundamental attribution error comes via the questioner–contestant paradigm (Block and Funder 1986; Stalder 2009) in which participants observe two individuals. One of them—the questioner constructs 10 trivia questions and poses them to the other individual—the contestant—who can answer only 3 of the 10 questions correctly. Participants then comparatively rank the two individuals’ general knowledge. The key finding is that participants tend to rank the questioner as more knowledgeable than the contestant. Because this response disregards the situational disadvantage of the contestant, it indicates the fundamental attribution error: Participants make their judgment on the basis of perceived properties of the individual when an appropriate basis would be the relevant aspects of the situation—i.e., that the questioner got to construct the questions (Block and Funder 1986).

Since these studies concern evaluations of individuals qua knowers, they are highly relevant to the study of folk epistemology. More specifically, the fundamental attribution error is potentially relevant for understanding one-off cases of illusion because they indicate that laypersons tend to underestimate situational factors in epistemic assessments.

4.4: The bounded rationality of outcome heuristics. Outcome bias is postulated on the basis of considerations that explain biases generally. The explanation starts with the assumption that human cognizers are capacity-limited and, therefore, relying on cognitive heuristics that are associated with biases (Evans 2010; Kahnemann 2011; Stanovich 2011). Biased responses may occur when a generally useful heuristic is overgeneralized—i.e., deployed in conditions in which it is inaccurate.13 Such a heuristic account plausibly applies to outcome effects, and this augments the assumption that they manifest a bias. Specifically, the truth effects are explained as the product of a cognitively cost-effective heuristic: Roughly, if the outcome is bad, the procedure is bad. Such an outcome heuristic is reasonably accurate in the subject’s normal environment, although it is systematically inaccurate in specifiable conditions. Thus, Baron and Hershey “…expected to find outcome bias because the generally useful heuristic according to their outcome may be overgeneralize to situations in which it is inappropriate” (Baron and Hershey 1988: 571).

---

13 I’ve argued elsewhere that we often use ascriptions of knowledge as heuristic proxies for other and more complex epistemic assessments (Gerken 2015b, 2017a–b).
It is important to this rationale for regarding outcome effects as biased that it is often reasonable to use the outcome heuristic. If you have limited information about some procedure and gain information that it succeeds on a particular occasion, then it is \textit{ceteris paribus} boundedly reasonable to let that information partly influence your assessment of the procedure. This point is generally recognized in empirical work on outcome and hindsight bias (e.g., Baron and Hershey 1988: 569; Mackie et al. 2001). Consequently, there is, for a given task that elicits an outcome effect, two interpretations that the empirical researchers consider.

\textbf{Cue Response:}

The participants reasonably use the outcome information as a cue because it is relevant information for the task.

\textbf{Bias Response:}

The participants use the outcome information as a cue although it is not relevant information for the task.

To control for the \textit{Cue Response}, researchers on outcome bias and related phenomena design their studies to make it clear to the participants that they have received all the relevant information prior to receiving information about the outcome. For example, the participants are typically told the exact probability of success of the procedure they are asked to assess.\footnote{It is, of course, hard to control for every confounder. Perhaps participants evaluate the token process in abstraction from its type or perhaps they assume that bad outcomes are not properly based on the specified procedure. Thanks to Dirk Kindermann and Julien Durant for these suggestions. Importantly, however, if such suggestions explain the findings, they provide no threat to subtle truth-sensitivity.} This renders the outcome irrelevant information. So, when participants nevertheless rely on an outcome heuristic, there are good grounds for postulating a bias response. Baron and Hershey make the point as follows: “While it is not in general unreasonable to use outcome as information bearing on the assessment of the basis of the outcome, this information may be \textit{overused}.” (Baron and Hershey 1988: 570). Mackie et al. also note that it is understandable that outcome information may be used even when it is irrelevant because it has “… the heuristic property of allowing inferences to be drawn even with a minimum of cognitive investment” (Mackie et al. 2001: 72).

In conclusion, outcome bias is postulated on the basis of well-established considerations about the use of reasonably accurate but biased heuristics. Moreover, studies on outcome effects go some way to control for the \textit{Cue Response}. Explaining the outcome effects in terms of an underlying heuristic augments the assumption that they manifest a bias. Such an account does
not claim that a biased response is altogether irrational but only that it exemplifies a (boundedly rational) systematic mistake. However, this is enough to motivate that the response pattern should not straightforwardly inform our theoretical views of the relevant issues. This, in turn, is what I seek to argue here.

4.5: Still not about the blame. On occasion, the literature on outcome bias is concerned with ascription of excuse, blame and other ethical categories. An example is its application to legal matters (Baron and Hershey 1988; Gino, Moore and Bazerman 2008). However, outcome bias does not merely or even primarily concern responsibility, blame, excuse, exemption or such. Rather, studies on outcome bias are primarily about assessment of objective qualities of, for example, the objective effectiveness of procedures. Indeed, much of the literature follows Baron and Hershey’s explicit concern with the evaluation of “…the true quality of the decision” (Baron and Hershey 1988: 569).

Baron and Hershey are explicit that “…the term responsibility need not only refer to the quality of the decision” (Baron and Hershey 1988: 570). Since they are interested in the objective properties, they did not investigate responsibility judgments but more objective measures such as “quality of thinking” (Baron and Hershey 1988: 571). In subsequent studies, the decision resulting in a bad outcome was explicitly “designed to be objectively better” than the one resulting in a good outcome (Stanovich and West 2008: 673). Generally, the difference between more subjective notions such as responsibility and objective properties is recognized in research on outcome and hindsight bias.

Empirical studies of outcome bias have applications for blame ascription. However, outcome and hindsight studies are primarily about the accuracy of assessments of the objective basis of belief, action or decision procedures. Hence, the cognitive scientists are hardly mistaking the relevant objective qualities for mere responsibility, blamelessness or excusability. If they were simply concerned with such phenomena, they would not explicitly focus on the objective properties of the procedure as when, for example, Stanovich and West anchor their Likert scale with ‘Clearly correct, an excellent decision.’ (Stanovich and West 2008).

Management theorists are interested in outcome bias because they are interested in properly evaluating procedures with regard to their effectiveness. They are concerned that evaluations may mistakenly favor an ineffective procedure type, P1, over a more effective procedure type, P2, when an occurrence of P1 yields the desired outcome as a matter of luck whereas an occurrence of P2 fails to do so as a matter of bad luck. Often the focus is on the procedure rather than the individual. Lefgren et al. consider the assessment of strategies that
have the aim of maximizing winning in sports and assume that such assessments may be biased (Lefgren, Platt and Price 2015). Given their interests, study designs and explicit remarks, it is clear enough that cognitive scientists who research outcome bias are often concerned with objective properties such as reliability or probability rather than with excusability or blame.

5: THE TRUTH EFFECT AS EPISTEMIC OUTCOME BIAS.
Reflective readers will already have noted the connection between the surveyed empirical work on outcome bias and the truth effects on epistemic assessments. But let me, for the sake of explicitness, state how I see the matter:

*Epistemic Outcome Bias*

The thesis *Epistemic Outcome Bias* explains truth effects on epistemic assessment in a manner that is compatible, indeed coherent, with subtle truth-sensitivity. When postulating an explanation in terms of a cognitive bias, it is important to embed it in existing empirical work. That’s why I have spent some energy going over some psychological findings of outcome and hindsight bias. This important empirical background has been strikingly absent from the philosophical discussions of truth effects on epistemic assessment.

Note that *Epistemic Outcome Bias* does not explain away Turri’s data, and it is not *ad hoc*. Rather, it is an application of a well-established way of accounting for this type of data. To anyone familiar with the research on outcome bias, the truth effect on epistemic assessments bears an uncanny resemblance: Participants in outcome bias studies rate procedure A with a good outcome as better than procedure B with a bad outcome although they are equal in regard to the probability of a successful outcome. Likewise, Turri’s participants rate S who holds a true belief as more epistemically warranted (justified, rational, reasonable etc.) than S* who holds a false belief although they are equal in terms of truth-conduciveness. Given the vast amount of empirical evidence that outcome bias is ubiquitous and pervasive in human cognition, *we should expect a truth effect on epistemic assessment.*

This prediction is supported by the work on hindsight bias which also indicates how knowledge of outcome may bias assessment of objective properties. Likewise, the prediction is supported by the work on fundamental attribution error. Recall that here participants
underestimate the situational factors (e.g., the illusionary environment) in favor of overestimating the individuals properties (e.g., the fallibility of vision) in explanations of a bad outcome.

The diagnosis offered by *Epistemic Outcome Bias* does not suggest that Turri’s findings are insignificant. On the contrary, it is extremely important to provide evidence that a bias that is found in some domains extends to folk epistemology. So, as I see things, Turri’s findings may teach us something extremely important—namely, that we are prone to *epistemic* outcome bias.

5.1. The challenge to simple truth-sensitivists. If the truth effects are manifestations of outcome bias, they provide no evidence for simple truth-sensitivity. However, there are at least two responses available to simple truth-sensitivists:

Response 1: Respect the analogy and reinterpret the empirical studies on outcome bias by rejecting that they indicate an outcome bias.

Response 2: Reject the analogy by providing an empirical or philosophical argument that the truth effect is different from the structurally similar studies on outcome bias.

A more fine-grained taxonomy of response possibilities may be required since each response may be developed in importantly different ways. For example, Response 1 may involve questioning the empirical evidence for postulating a bias. But it could also consist in arguments that the intuitive responses are not biased in any illegitimate manner. I will structure the present discussion by considering Response 1 in the next section and Response 2 in this one.

While there are differences between Turri’s studies and the studies on outcome bias, these differences do not provide comfort for simple truth-sensitivists. A central difference is that cognitive psychologists seek to explicitly control for a *Cue Response*. That is, they seek to rule out that the outcome information is a good cue because it is relevant information for the task. For example, participants are often given an explicit numerical reliability measure of the relevant processes (Baron and Hershey 1988; Stanovich and West 2008).

A potential difference is that Turri’s study design does not explicitly control for the *Cue Response*. (This is fair enough since Turri did not pursue evidence for outcome bias). For example, participants in (Watch) are informed that “…inventory isn’t perfect, but it is extremely accurate.” (Turri 2016a). This is more open to interpretation than a firm numerical measure. Consequently, one might argue that Turri’s participants are reasonably using truth as an indicator of truth-conduciveness. If participants are confronted with limited information about the truth-
conduciveness of a given process (checking the inventory), information about the outcome on a particular occasion may reasonably be regarded as relevant information. When the outcome is a false belief, participants experience incongruence with the information that the process is reliable.

However, if Turri’s findings were merely indicative of a *Cue Response*, they would not provide evidence for simple truth-sensitivity. They would just indicate that the folk appreciate that truth value is a reasonable indicator of truth-conduciveness. Alternatively, they would indicate that participants took the outcome as evidence that bad outcomes are not properly based on the specified procedure. But it is consistent with subtle truth-sensitivity that we should assess violations of the basing-relation as inferior to cases of proper basing.

Moreover, even if some of Turri’s findings were best interpreted as manifesting a *Cue Response*, it would be very surprising if there were no truth effect on epistemic assessment in studies that were explicitly designed to control for the *Cue Response*. That said, I think that Turri’s studies as they stand are best understood as providing evidence for a *Bias Response* and, therefore, for *Epistemic Outcome Bias*. But it is important to note that on neither interpretation, do they provide evidence for simple truth-sensitivity over subtle truth-sensitivity.

5.2. Blame and excuse maneuvers: As noted, knowledge of an outcome may bias our attributions of responsibility, excusability and blame. But recognizing this fact does nothing to suggest that knowledge of an outcome does not also bias our assessment of objective epistemic properties. Many factors may excuse (or even exempt from evaluation) someone for reaching a bad outcome by relying on a bad procedure. But this is a different matter from someone reaching a bad outcome due to the reliance on a procedure that is the best available by objective parameters such as degree of reliability. It is not clear that a subject who deploys such a procedure stands in need of an excuse (Gerken 2011, Schechter 2017). Insofar as she stands in need of an excuse, it is only for not meeting the standard of true belief.15

Of course, it should be recognized that the fact that it is unreasonable to blame S more than S* does not entail that S and S* are equi-warranted. But it should also be recognized that this non-entailment does not entail, or even suggest, that S and S* are not-equi-warranted. Consequently, the fact that ascriptions of blame, excusability and responsibility are prone to outcome bias is compatible with more objective epistemic assessments of, for example, the quality of a decision or procedure being prone to outcome bias.

15 I find it is more accurate to say that bad luck *explains* the failure to meet the standard than that it excuses the agent for not meeting it. But ordinary language is too imprecise to rely on in this area.
Notably, Turri found truth effects on epistemic evaluations and responsibility attributions but not on blame ascriptions (Turri 2016a, Experiment 2). If this finding is robust, it may raise a challenge for views that the participants are in effect concerned with blameworthiness. The finding that would provide evidence for such claim would be a more pronounced truth effect on blame ascriptions than on assessments of objective epistemic properties. More generally, asymmetries between response patterns for epistemic assessments and ascriptions of blame challenge claims that participants mix them up.\textsuperscript{16}

As noted, cognitive scientists are explicitly designing studies to keep outcome bias on ascriptions of responsibility and excuse apart from outcome bias on assessment of objective properties of a decision or procedure. So, it is not unreasonable to take studies that find outcome effects on the latter at face value.

\textbf{6: METHODOLOGICAL CONSIDERATIONS.}

Before concluding I will briefly consider the methodology underlying the step from a truth effect on folk epistemological assessments to epistemological conclusions. In doing so, I will – albeit even more briefly – sketch the contours of a methodological alternative.

\textbf{6.1. Straightforward vs. equilibristic methodology.} In one place, Turri provides the following summary of what one set of findings of truth effects accomplish:

“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.” (Turri 2015c: 4025)

I agree with the first suggestion that the findings compromise arguments that simply appeal to alleged “parity intuitions” in folk epistemology.\textsuperscript{17} However, the second and third suggested accomplishments concern the step from a truth effect in our folk epistemology to factive epistemic norms. As noted, I take the \textit{Epistemic Outcome Bias} hypothesis to compromise such a

\textsuperscript{16} Of course, Turri’s finding suggests an asymmetry with cases in which the outcome has an effect on blame ascriptions (Gino, Moore and Bazerman 2008). It is hard on the basis of the available evidence to determine whether this is best explained by differences in probes or study design or by distinctively epistemological aspects of Turri’s study. Also relevant are Brössel and Reuter’s findings of differences between ‘should’-questions and ‘permissibility’-questions which they take to support non-factive epistemic norms (Brössel and Reuter forthcoming).

\textsuperscript{17} I have not myself argued from such claims about folk intuitions although I have given other arguments in favor of non-factive norms (Gerken 2011, 2012, 2014, 2015a, 2015c, 2017a–b, 2018, forthcoming).
step. Notably, Turri does not argue solely from the truth effects to factive norms but provides a wider variety of empirical considerations (Turri 2017a gives an overview). Moreover, in the case of assertion, he is primarily concerned with our normative practices and does not identify these with the epistemic norms that epistemologists are concerned with. In the case of epistemic norms of belief, decision and evidence, Turri’s goal is the more restricted one of identifying central tendencies in folk epistemology.\(^\text{18}\)

However, it is for substantive reasons well worth considering the step from the truth effects to factive epistemic norms. One proponent of a knowledge norm of assertion, DeRose, has sketched a methodology that could vindicate such a step. This methodology of the straightforward “takes very seriously the simple positive and negative claims speakers make utilizing the piece of language being studied, and puts a very high priority on making those natural and appropriate straightforward uses come out true, at least when that use is not based on some false belief the speaker has about some underlying matter of fact.” (DeRose 2009: 153. See also Stanley 2005: 5. For criticism, see Gerken 2017a, Chapter 3.3).

What underlies the present response is a deep methodological disagreement with the assumption that folk epistemological patterns may straightforwardly inform epistemological theory. I agree with Turri’s claim that the findings indicate a competence: “…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.” (Turri 2015c: 4011). But, crucially, this competence is a systematically fallible one.

Our folk epistemological competences are, like the rest of our cognitive resources, severely capacity-limited and, consequently, driven by heuristics. Folk epistemological heuristics are, like heuristics in general, associated with biases. Therefore, we should be cautious in letting folk epistemological judgments and practices straightforwardly inform our epistemological theory. In fact, epistemological theory may sometimes be invoked to interpret folk epistemological judgments and practices straightforwardly inform our epistemological theory. I call the methodology that consists in critical use of empirical work on folk epistemology as well as epistemological theorizing ‘equilibristic methodology’ (I elaborate in Gerken 2017a Chapter 3.3 by providing some broad methodological principles).

An equilibristic methodology requires that folk epistemological biases are only postulated in a manner that is principled and not \textit{ad hoc}. Consequently, a specific psychological explanation must be offered and embedded in independently motivated assumptions (Gerken 2017a). This is why I have argued for \textit{Epistemic Outcome Bias} by way of emphasizing existing empirical work on outcome bias.

\(^{18}\) Thanks to John Turri for clarifying correspondence.
6.2. Does postulating epistemic outcome bias amount to an error theory? May one object that *Epistemic Outcome Bias* amounts to an unacceptable error theory?

No. *Epistemic Outcome Bias* is not an error-theory in the sense that people are generally unreliable in their epistemic assessments since it is compatible with assuming that such assessments are largely reliable and true. Furthermore, *Epistemic Outcome Bias* is not a performance error theory. Rather, it postulates a systematic psychological bias that is explained in reference to signature cognitive limitations rather than haphazard performance errors (See Gerken 2017a on this distinction).

A central reason why cognitive scientists postulate outcome bias in general is that doing so does not amount to postulating wholesale irrationality among the folk exhibiting it. Rather, the folk’s use of an outcome heuristic may be seen as boundedly rational insofar as it is a cost-effective, and reasonably accurate, way of assessing decisions and procedures. Likewise, *Epistemic Outcome Bias* is not an error theory but rather a standard account of empirical findings in terms of a bias.

May the idea that the truth effects manifest a reasonable bias (or perhaps no bias at all) be invoked in a version of Response 1? One version of this response appeals to Gigerenzer’s work showing that heuristics can outperform reflective cognition on certain tasks (Gigerenzer 2008). Indeed, this type a response would be situated in a grand debate called ‘the rationality wars’ (Rysiew 2008; Sturm 2012). An extreme position in this debate is what Stanovich labels the “panglossian” view according to which there is “no difference between descriptive and normative models of performance because human performance is actually normative” (Stanovich 2011: 9). This view would underwrite a straightforward argument from the truth effects to simple truth-sensitivism. Crucially, however, there is no similar argument from the weaker claim that that the truth effects manifest a boundedly rational outcome bias. As noted, the fact that outcome effects can be explained in terms of a boundedly reasonable outcome heuristic is a reason to regard them as exemplifying a bias. But the responses that are biased in this manner should not inform epistemology in any straightforward manner.

The debate over the nature of biases is of a magnitude that calls for further discussion (see Gerken 2017a). For the present purpose, I have only highlighted that arguing that the truth effects are boundedly rational does not show that they are not biased or biased in a way that permits modelling our epistemology directly on them.
6.3. Is it revisionary to postulate epistemic outcome bias? Might simple truth-sensitivists argue that my central hypothesis, *Epistemic Outcome Bias*, uncritically follows cognitive scientists in a revisionary approach to folk epistemology? Given the similarity between the truth effects and outcome effects generally, this amounts to rejecting the general idea that outcome bias is a bias. This should be recognized as strategy for developing Response 1. Of course, we should not uncritically presuppose that the cognitive scientists are right in their presupposition that outcome bias is a bias. However, to avoid the charge of doing so, I started out with epistemological arguments that align with the parity assumption in cognitive science. Moreover, I noted that some confounders have been controlled for and that the postulation of a bias is reinforced by an account of the underlying heuristics. While none of these arguments are conclusive, they lend abductive weight to *Epistemic Outcome Bias* and the claim that we should not straightforwardly model our epistemology on such a bias – even if it is boundedly rational.

Crucially, Turri’s findings do not compromise the claim that the folk are generally prone to outcome bias. After all, this general claim predicts the findings. So, an argument is required before dismissing this long-standing claim which has guided a range of fruitful work in the cognitive sciences (Stanovich and West 2008). Whereas arguments that cognitive science is dramatically mistaken should be welcomed, they should also be required. Dismissing an assumption that is central to a long-standing empirical research program is not a thing to take lightly. It calls for careful criticism of the operative scientific framework and a viable alternative. The findings of truth effects on epistemic assessment do not by themselves provide either.

Although philosophical considerations can sometimes provide the basis for corrections and revisions of cognitive science, cognitive psychology does not in this case appear to rest on philosophical naïveté. As emphasized, its key assumptions align with influential philosophical assumptions such as the externalist ones sketched in Section 3.

So, in the absence of cogent empirical or philosophical arguments, it seems methodologically problematic to take the evidence of a truth effect as evidence for simple truth-sensitivity. Given that coherent philosophical and empirical considerations support *Epistemic Outcome Bias*, the hypothesis is not philosophically, empirically or methodologically revisionary.

6.4. Concluding methodological considerations. Since *Epistemic Outcome Bias* is an empirically based hypothesis with philosophical impact, it might be compromised by both empirical and philosophical arguments. However, such arguments must be given before it is dismissed. Unfortunately, the hypothesis has not yet been included in discussions of the candidate explanations of the truth effects on epistemic assessments.
In articulating *Epistemic Outcome Bias*, I have argued that it is not revisionary or otherwise methodologically suspect since it is embedded in existing philosophical and empirical work. While it would be premature to think of *Epistemic Outcome Bias* as established, it is a viable hypothesis that predicts findings of a truth effect in a manner consistent with subtle truth-sensitivism. So, it would be equally premature to declare simple truth-sensitivism as empirically established without considering *Epistemic Outcome Bias*. In an area as poorly understood as folk epistemology, it is crucial to proceed with caution. Doing so involves considering the philosophically reasonable hypotheses that are compatible with the empirical findings. Thus, I have sought to make the case that *Epistemic Outcome Bias* is among the philosophically and empirically reasonable hypotheses that deserve consideration.

7: CONCLUSION. So, how do the findings of truth effects, according to the present diagnosis, bear on the debates over epistemic norms? First, subtle truth-sensitivists should recognize that Turri succeeds in his negative goal of empirically compromising arguments that proceed from alleged folk epistemological “parity intuitions.” Secondly, but as importantly, the findings do not support simple truth-sensitivism over subtle truth-sensitivism if *Epistemic Outcome Bias* is the correct diagnosis. Both of these conclusions are important because they curb both empirically detached speculations as well as overly straightforward construction of philosophical theory from empirical findings. More generally, it is important to consider the truth effects in relation to, rather than in abstraction from, relevant existing work in cognitive psychology. So, whether further work corroborates or compromises *Epistemic Outcome Bias*, it is likely to improve our understanding of the nature of folk epistemology and its relation to epistemological theorizing.

Furthermore, if *Epistemic Outcome Bias* is correct, findings from the truth effect studies are very valuable in a further sense. Our epistemic assessments of each other are among the most consequential aspects of our folk epistemology. Whether we regard someone’s actions or assertions as meeting the relevant epistemic norms has a very significant bearing on how we will treat her as a member of the epistemic community. So, if our folk epistemological assessments are prone to a robust epistemic outcome bias, this is something that we must address in the same way that we address other biases with potentially damaging ramifications. Thus, Turri’s findings may be significant not because they should lead us to adopt a simple truth-sensitive epistemology
but because they should lead us to put our epistemology to use in addressing systematic shortcomings of our folk epistemology.19

LITERATURE


19 The paper grew out of *On Folk Epistemology* (cf. Section 6.1.c), since a responsible discussion called for a paper of its own. My main gratitude is to John Turri who read a full draft and found time for extended correspondence. The paper is much better for it, and I am extremely grateful for the help. The paper was presented at the Bled Conference in June 2017, at the ECAP9 at LMU München in August 2017 and at a workshop at the University of Copenhagen in November 2017. I am grateful to the participants for helpful discussion. Finally, I am grateful to a referee for a helpful organizational suggestion and to Julie Brummer for some stylistic ones.

This article is protected by copyright. All rights reserved
Press.


This article is protected by copyright. All rights reserved


This article is protected by copyright. All rights reserved

