Epistemic Closure in Folk Epistemology*

James R. Beebe and Jake Monaghan (University at Buffalo)

Forthcoming in Joshua Knobe, Tania Lombrozo, and Shaun Nichols (eds.),

Oxford Studies in Epistemology

We report the results of four empirical studies designed to investigate the extent to which an epistemic closure principle for knowledge is reflected in folk epistemology. Previous work by Turri (2015a) suggested that our shared epistemic practices may only include a source-relative closure principle—one that applies to perceptual beliefs but not to inferential beliefs. We argue that the results of our studies provide reason for thinking that individuals are making a performance error when their knowledge attributions and denials conflict with the closure principle. When we used research materials that overcome what we think are difficulties with Turri's original materials, we found that participants did not reject closure. Furthermore, when we presented Turri's original materials to non-philosophers with expertise in deductive reasoning (viz., professional mathematicians), they endorsed closure for both perceptual and inferential beliefs. Our results suggest that an unrestricted closure principle—one that applies to all beliefs, regardless of their source—provides a better model of folk patterns of knowledge attribution than a source-relative closure principle.

* This paper has benefited greatly from helpful comments and suggestions from John Turri, Wesley Buckwalter, two anonymous reviewers from *Oxford Studies in Experimental Philosophy*, an anonymous reviewer for the Second Annual Minds Online Conference, and audiences at the 2015 Experimental Philosophy Group UK conference, the 2016 Southern Society for Philosophy and Psychology conference, and University College Dublin.

Keywords: epistemic closure, folk epistemology, experimental philosophy, knowledge, expertise

1. Introduction

Epistemic closure principles have figured prominently in contemporary epistemology, particularly in discussions of radical skepticism. The most commonly discussed skeptical challenges in recent decades employ skeptical hypotheses that depict situations that are subjectively indistinguishable from what we take our normal circumstances to be but in which we fail to have knowledge. Where ‘*p*’ is some proposition we ordinarily take ourselves to know and ‘*SK*’ is a skeptical hypothesis, these challenges are typically represented as arguments of the following form[1]:

(1.1) If I know that *p*, then I know that not-*SK*.

(1.2) I do not know that not-*SK*.

(1.3) Therefore, I do not know that *p*.

Premise (1.2) is usually supported by considerations that purport to show that one’s current evidence somehow fails to favor *p* over *SK*. Premise (1.1) is usually taken to be an abbreviated substitution instance of the following closure principle for knowledge:

EC1. If *S* knows that *p*, and *S* knows that *p* entails *q*, then *S* knows that *q*.

Fred Dretske (1970) and Robert Nozick (1981) (in)famously denied that any closure principle for knowledge was true in their efforts to respond to radical skepticism and to construct a satisfactory theory of knowledge. However, it is well known that the tide of epistemological opinion has run strongly against Dretske and Nozick, with most philosophers agreeing that

[1] Cf. the presentations of skepticism in Brueckner (1985; 1994), DeRose (1999), Pritchard (2002a; 2002b; 2005), and Greco (2008).

closure is “an extremely intuitive idea” (Hawthorne 2004, 33) and that it captures the seemingly incontrovertible idea that “deduction is a way of extending one’s knowledge” (Williamson 2000, 117). Indeed, many would go further and agree with Richard Feldman’s (1995, 487) oft-quoted claim that the denial of closure was “one of the least plausible ideas to come down the philosophical pike in recent years.”

As John Turri (2015a, 3) notes, “closure’s proponents rely *heavily* on the claim that epistemic closure is a defining feature of ordinary thought and talk about knowledge—that is, of folk epistemology.” When philosophers claim a particular idea is extremely intuitive, it is common for experimental philosophers to wonder how many people actually agree with it and to look for ways to find out.[2] Accordingly, Turri (2015a) conducted a series of studies to investigate whether epistemic closure is in fact deeply embedded in folk epistemology. His results suggest that individuals may endorse closure only for perceptually based beliefs but not for inferential beliefs.

In a between-subjects design, Turri asked participants to read one of two versions of the following vignette:

> *Car Theft*. When Mr. Maxwell arrives at work in the morning, he always parks in one of two spots: C8 or D8. Half the time he parks in C8, and half the time he parks in D8. Today Maxwell parked in C8.
>
> It’s lunchtime at work. Maxwell and his assistant are up in the archives room searching for a particular document. Maxwell says, “I might have left the document in my car.” The

[2] Turri (2015, 3) writes, “This rather lavish gloss on ordinary practice and what ‘we’ find intuitively compelling or conversationally ‘abominable’ has thus far gone unchallenged, even by closure’s opponents (Nozick 1981, 205-6; Dretske 2013, 32). Thus it is widely assumed that closure’s opponents bear the revisionist’s burden of either explaining away contrary intuitions or showing that their view delivers otherwise unattainable benefits.”

> assistant asks, “Mr. Maxwell, is your car parked in space C8? It’s not unheard of for cars to be stolen.”

The perception version of Car Theft ends with the following words:

> Maxwell looks carefully out the window and then responds, “No, my car has not been stolen. It is parked in C8.”

The inference version ends with:

> Maxwell thinks carefully for a moment and then responds, “No, my car has not been stolen. It is parked in C8.”

Car Theft is based upon Jonathan Vogel’s (1990) well known discussion of car theft type cases that did more than any other article at the time to convince epistemologists that the rejection of closure was a mistake.

Turri then asked participants to select all of the statements below that are true in the Car Theft vignette[3]:

> (2.1) Maxwell knows that his car is parked in C8.
>
> (2.2) Maxwell knows that his car has not been stolen.
>
> (2.3) Maxwell is in the archives room.
>
> (2.4) Maxwell is in his assistant’s office.

Statements (2.3) and (2.4) served as comprehension checks. (2.1) and (2.2) tested for closure because Turri intended Maxwell to be depicted as (i) considering his belief that his car is parked in C8 and the basis upon which this belief rests, (ii) reaffirming his belief that the car is indeed parked there, (iii) appreciating that his car’s not being stolen follows from it’s being parked in

[3] In Turri’s first study, he presented participants with both (2.3) and (2.4) and either (2.1) or (2.2). In his second study, Turri presented each participant with all four statements. In a third study, he presented participants with conjunctions that combined (2.1) and (2.2), the negation of only one of these statements, or the negations of both. The overall pattern of data obtained in these different ways remained roughly the same.

C8, and (iv) affirming that it has not been stolen on this basis. If participants endorse closure, Turri thinks they should select (2.2) as often as they select (2.1). If, however, they reject closure, they should select (2.2) less often.

In Turri's fifth study, he gave participants a second vignette that had the same structure as Car Theft:

> *Computer*. When Mrs. Palmer arrives at work in the morning, she always starts her computer and puts it to work analyzing data sets. Today she put her computer to work analyzing demographic data and then walked into her assistant's office across the hall.
>
> Palmer says to her assistant, "It might take the computer a while to analyze the demographic data." The assistant asks, "Mrs. Palmer, is your computer analyzing the data? It's not unheard of for computer hackers to attack the system and crash a computer."

The perception and inference versions of Computer ended with one of the following statements:

> Mrs. Palmer looks carefully through the window into her office and then responds, "No, the hackers have not crashed it. The computer is analyzing data."
>
> Mrs. Palmer thinks carefully for a moment about the matter and then responds, "No, the hackers have not crashed it. The computer is analyzing data."

Turri then asked participants to select all of the statements below that are true in the vignette:

> (3.1) Palmer knows that the computer is analyzing data.
>
> (3.2) Palmer knows that hackers have not crashed the computer.
>
> (3.3) Palmer is in the archives room.
>
> (3.4) Palmer is in her assistant's office.

As above, (3.3) and (3.4) served as comprehension checks, while (3.1) and (3.2) tested for participants' endorsement of epistemic closure for knowledge.

Turri found that participants thought Maxwell both knew that his car was parked in C8 and knew that it had not been stolen in the perception condition. In the inference condition, participants thought that Maxwell knew that his car was parked in C8, but they were significantly less likely to think he knew that it had not been stolen. A similar pattern was found with Computer. Participants were as likely to think Palmer knew that her computer was analyzing data as they were to think that she knew that hackers had not crashed the computer in the perception condition. In the inference condition, however, they were less likely to think that she knew that the hackers had not crashed the computer than they were to think that she knew that the computer was analyzing data. Turri (2015a, 14) concludes:

> Overall, then, ordinary practice does not endorse an unqualified version of the epistemic closure principle. But it's arguably consistent with our results that ordinary practice is committed to a *source-relative* version of the epistemic closure principle. In particular, it's consistent with our results that a *perceptual epistemic closure principle* is a defining feature of folk epistemology. This is ironic given that the most famous attempted counterexamples feature perceptually based beliefs (e. g. Dretske's zebra case).

Turri's findings thus seem to challenge the received view in contemporary epistemology that an unrestricted epistemic closure principle is a central and non-negotiable feature of our shared epistemic practices.

We agree with the received view that individuals' are making a mistake when they endorse (2.1) and (3.1) but deny (2.2) and (3.2). However, we think the error lies not in denying that Maxwell knows that his car has not been stolen or that Palmer knows that hackers have not crashed the computer when these beliefs are based upon inference. Rather, we think the error lies in agreeing that Maxwell knows that his car is parked in C8 or that Palmer knows that the computer is analyzing data in Turri's inference conditions. We contend that the studies we report below provide support for our claim that participants in the inference conditions who agree with (2.1) and (3.1) are making a performance error. We use these data to argue that an unrestricted closure principle provides the best model of folk knowledge attributions.

There are a few other things about Turri's method of examining closure that we think call for further investigation. One is that Turri believes that both the perception and inference versions of Car Theft and Computer depict the protagonists as believing one proposition, seeing that this proposition entails another, and (allegedly) believing the second on the basis of inferring it from the first. However, in the perception version of Car Theft, Maxwell "looks carefully out the window" and sees that his car parked in C8. Like Turri, most philosophers will probably think that he only sees that it is parked in C8 but does not see that it has been stolen. Thus, if he is to believe that his car has not been stolen, it will be necessary for him to draw an inference. However, the metaphysics and epistemology of perception are quite complex, and it is not clear to us that ordinary participants will think that an inference must be made in order for Maxwell to arrive at the belief that his car has not been stolen. In our view, it seems quite likely that many ordinary individuals will think that, in addition to seeing that his car is parked in C8, Maxwell will also see—directly and without the aid of inference (or at least without any inference in addition to those that may already be involved in perception)—that it has not been stolen. It seems quite intuitive to think that the fact that Maxwell is laying eyes on his car directly confirms his belief that his car has not been stolen and unintuitive to think that seeing his car can confirm his belief only if this sensory experience is coupled with the belief that seeing his car entails that it has not been stolen. But if Turri’s participants understand Maxwell as seeing directly that his car has not been stolen, then the perception version of Car Theft will not test individuals’ endorsement of closure at all. Similar remarks apply to the Computer case, in which Palmer “looks carefully through the window into her office” and sees that the computer is analyzing the data and is perhaps also viewed as seeing that the computer has not crashed. Consequently, we may not be able to conclude (as Turri does) that ordinary individuals endorse closure for perceptually based beliefs but not for inferential ones. The lesson may simply be that when individuals are presented with clear tests for closure (e.g., in the inference conditions), they fail to endorse it.

The inference versions of Car Theft and Computer seem better poised to test for folk endorsement of closure because the protagonists in these cases are not in a position to directly perceive that the car has not been stolen or that the computer has not been crashed. These beliefs must be arrived at by inference. Of course, the protagonists are not in a position to directly verify that the car is parked in C8 or that the computer is analyzing the data either. But the important thing is that the beliefs about car theft and computer hacking can only be known on the basis of inference.

It is somewhat odd that Turri emphasizes the inferential component in Maxwell’s belief that his car is parked in C8 and Palmer’s belief that the computer is analyzing data without acknowledging that their beliefs are based upon perception and memory as well. Maxwell remembers where he parked his car when he arrived at work in the morning. Palmer remembers that when she arrived at work in the morning she started her computer and put it to work analyzing data sets. Furthermore, their memorial beliefs were ultimately based upon perception themselves. On the basis of what they remember, together with background beliefs about how likely car thefts and hacker attacks are, each makes an inference about whether the car is still parked in C8 or the computer is still analyzing data. Since the ultimate basis of the beliefs in question are the protagonists' (perceptually based) memories, it may be that participants reject closure in these cases because of the role of memory or background beliefs play in their justification rather than the role of inference.

A second issue with Turri's studies that we think calls for more careful scrutiny is that in each condition he asked participants to select which of four statements—(2.1) through (2.4) or (3.1) through (3.4)—are true. We agree that if participants endorse closure, they should indicate that both (2.1) and (2.2) are true in the Car Theft case and (3.1) and (3.2) are true in Computer. However, the task demands imposed by Turri's experimental set-up are such that participants might stop selecting true items not because they have fully thought through the relevant issue and determined that no more knowledge attributions should be made but rather that they have failed to reflect fully upon the issue. Consider the fact that someone might know that *p* is true and know that *p* entails *q* but not get around to forming the belief that *q* is true. In like fashion, participants might believe that Maxwell knows that his car is parked in C8 and that he knows that his car's being parked entails that it has not been stolen. But a lack of sufficient motivation might keep them from thinking about, forming beliefs about, or bothering to select an answer about whether Maxwell knows that his car has not been stolen. Laziness seems far more likely to keep participants from attributing knowledge to Maxwell that he knows that his car has not been stolen than it does to keep them from attributing knowledge to him that his car is in C8. Turri, however, interprets failures to select (2.2) or (3.2) as true as meaning that participants do not endorse closure when the failure may have merely resulted from lack of motivation. We think that it would be more ideal if the tasks were structured so that each participant is required to respond to statements (2.2) and (3.2) and is not given the option to simply skip facing them directly.

A third issue with Turri's studies concerns the samples he used in his studies. Proponents of the 'expertise defense' (e.g., Sosa 2007, Williamson 2011) complain that when experimental philosophers use undergraduates, MTurk workers, or other adults who have not received formal philosophical training in their studies, the data they obtain are of rather poor quality and consequently fail to tell us anything interesting about how we should view the cases presented in the research materials. The complaint alleges that it is only the reflective intuitions of the suitably trained that would warrant the often surprisingly conclusions some experimental philosophers seek to support. In some concluding remarks, Turri (2015a, 15) considers the possibility that epistemic closure is "too subtle or complicated for ordinary people to understand, which is why they exhibit patterns that violate it," but he notes that "ordinary people are remarkably sensitive to the complex factors that undermine knowledge in Gettier cases (Turri, Buckwalter, & Blouw 2014), which are at least as subtle and complicated as the epistemic closure principle."

While we do not endorse the expertise defense as a general response to surprising results in experimental philosophy, we wondered whether non-philosophers who had more formal training in deductive reasoning than the average person might be better able to appreciate whatever complexities are involved in deductive closure. In particular, we wondered how professional mathematicians—who are like the folk in their general lack of familiarity with debates among philosophers concerning closure and skepticism but unlike the folk in regard to their expertise in deductive reasoning—might respond to Turri's cases. If non-philosophers fail to indicate that Maxwell knows both that his car is parked in C8 and that his car has not been stolen or fail to indicate that Palmer knows both that the computer is analyzing data and that it has not been crashed, it might be due to their lack of sufficient skill in handling entailment relations. In other words, their responses might reflect a performance error rather than the fact that closure is not an important part of folk epistemology. However, if professional mathematicians rejected closure at the same rates as other non-philosophers, we think we would be entitled to greater confidence in believing that this pattern of responses was evidence that an unrestricted closure principle is not in fact part of our core folk epistemic practices.

We performed a series of empirical studies designed to address each of the issues described above. We first replicated Turri's basic results (Study 1) using Car Theft. In each of our studies, we required participants to respond to both the positive and the negative knowledge claims—i.e., to both (2.1) and (2.2) or both (3.1) and (3.2). We thereby addressed the second issue raised above concerning the possibility that a lack of sufficient motivation might interact with Turri's experimental set-up in a way that kept participants from directly reflecting upon whether Maxwell knows that his car has not been stolen.

Above we noted that many participants might think that Maxwell can both see that his car is parked in C8 and see that it has not been stolen and thus that Car Theft may not really test for participants' endorsement of closure at all. In order to provide participants with a perceptual case that involved inferring something from a perceptual belief that could not itself be directly verified by perception, in Study 2 we gave participants a modified version of Car Theft and asked them whether Maxwell knows that his car is parked in C8 and whether he knows that it has not been destroyed and replaced by an exact replica. We believe this study presents a more significant test of Turri's source-relative closure hypothesis than his original Car Theft vignette because Maxwell is no longer able (if he ever was) to have direct perceptual access to the second proposition in question. We found that participants in Study 2 overwhelmingly endorsed closure in both the perception and the inference conditions. Due to some concerns that Turri (2016) and Wesley Buckwalter (2016) raised about the fanciful nature of the vignette used in Study 2, we ran additional study (Study 3) to address these concerns and found further support for our case.

In our fourth study, we compared the knowledge attributions and denials of professional mathematicians with university undergraduates and, in accord with our hypothesis, found that mathematicians endorsed closure for both perceptual and inference beliefs. Also in accord with our expectations, they agreed with both positive and negative knowledge claims at much lower levels in the inference conditions than in the perception conditions. These findings support our contention that the non-philosophers in Turri's studies are making a performance error when they attribute and deny knowledge in ways that conflict with the closure principle.

We maintain that an unrestricted closure principle that applies to all beliefs, regardless of source, provides a better model of the patterns of folk knowledge attributions that have been observed and thus that closure does seem to be a core part of folk epistemology after all. In the course of providing evidence for this conclusion, we also show that skeptical pressure from relevant alternatives that are difficult to rule out can significantly decrease knowledge attributions—an effect that experimental philosophers have not always been able to observe.

2. Study 1

In our first study, we attempted to replicate Turri's (2015a) original results. We recruited 120 participants (average age = 35, 34% female, 81% Caucasian, 100% native English speakers) from Amazon's Mechanical Turk (www.mturk.com) and presented them with either the perception or the inference version of Car Theft.[4] Turri's first three studies used vignettes that ended with Maxwell stating "No, my car has not been stolen. It is parked in C8." However, in his fourth study, in an effort to make clear to participants that Maxwell believed that his car had not been stolen because he believed it was parked in C8, Turri has Maxwell respond with "My car is parked in C8, so it has not been stolen." We used the latter formulation in each of our studies.[5]

Participants were asked to indicate the extent to which they agreed or disagreed with statements (2.1) and (2.2), the order of which was counterbalanced. Participants were asked to select one of the following seven options as their answer: Completely Disagree, Mostly Disagree, Slightly Disagree, Neither Agree nor Disagree, Slightly Agree, Mostly Agree, and Completely Agree. For purposes of analysis, 'Completely Disagree' was coded as '1,' 'Mostly Disagree' as 2, and so on. Participant responses are summarized in Table 1.

	Perception	**Inference**
Positive Knowledge	6.45*** (1.27)	6.20*** (1.22)
Negative Knowledge	5.62*** (2.00)	3.75 (1.99)

Table 1. Mean levels of agreement (and standard deviations) for each knowledge claim in the perception and inference conditions of Study 1. In all tables, an '*,' '**,' or '***' indicates that the designated mean differs significantly from the neutral midpoint at the .05, .01, or .001 level.

[4] Mturk workers in all studies were required to have at least a 98% approval rating on at least 5,000 tasks and were paid $.35 for their participation. Workers were prevented from participating in more than one study in this paper.

[5] Turri's use of this alternative formulation did not lead to a significant difference in participants' responses, but we thought it was nonetheless preferable to use it in our own. In a fifth study that used Computer, Turri has Palmer state "No, the hackers have not crashed it. The computer is analyzing data." In Study 5 below, where we use Computer, we changed this to "The computer is analyzing the data, so hackers have not crashed it."

Mean levels of agreement differed significantly from the neutral midpoint in three out of four cases (with very large effect sizes).[6] A two-way mixed ANOVA revealed a significant main effect for source (perception vs. inference) that was medium in size and one for the valence (positive vs. negative) of the knowledge claim in question that was large in size.[7] In other words, participants attributed knowledge in significantly greater numbers in the perception condition than in the inference condition and expressed more agreement with the positive knowledge claim than with the negative knowledge claim. A significant and medium-sized interaction was observed between source and valence.[8] In other words, the lower agreement registered for the negative knowledge claim in the inference condition was the result of more than just the source and valence factors operating independently.[9][10]

In order to investigate possible differences in participants' endorsement or rejection of closure, we grouped together all those who selected the 'Completely Disagree,' 'Mostly Disagree,' or 'Slightly Disagree' into a Disagree (or knowledge-denying) category, all who selected 'Neither Agree Nor Disagree' into a Neutral category, and all who selected 'Slightly Agree,' 'Mostly Agree,' or 'Completely Agree' into an Agree (or knowledge-affirming) category. We then classified participants as rejecting closure if they either (i) agreed (with some

[6] Perception/positive knowledge: $t(59) = 14.97$, $p < .001$, $r = .89$. Perception/negative knowledge: $t(59) = 6.26$, $p < .001$, $r = .63$. Inference/positive knowledge: $t(59) = 13.99$, $p < .001$, $r = .88$. Inference/negative knowledge: $t(59) = -.97$, $p > .05$.

[7] Source: $F(1, 118) = 19.57$, $p < .001$, $r = .38$. Valence: $F(1, 118) = 77.38$, $p < .001$, $r = .63$.

[8] $F(1, 118) = 18.76$, $p < .001$, $r = .37$.

[9] This result provides further confirmation Turri's (2015b, 307) claim to have uncovered a new kind of cognitive bias, which he calls the 'source-content bias,' according to which "people evaluate inferential belief more harshly than perceptual belief," "people evaluate inferential belief more harshly when its content is negative (i.e., that something is *not* the case) than when it is positive (i.e., that something is the case)," and these two factors interact to produce a stronger effect in combination than would be predicted by the sum of their independent contributions.

[10] There was no significant main effect for gender ($F(1, 115) = .01$, $p > .05$), and there were no significant interaction effects for gender (gender x source: $F(1, 115) = 1.05$, $p > .05$; gender x valence: $F(1, 115) = .46$, $p > .05$). Age was positively correlated with negative knowledge attributions in the perception condition ($r = .34$, $p < .01$). A three-way (source x valence x order) mixed ANOVA showed no main effect for order, $F(1, 116) = .53$, $p > .05$.

degree of confidence) with a positive knowledge claim but failed to agree with the negative knowledge claim (either by disagreeing with the latter or remaining neutral about it) or (ii) disagreed (with some degree of confidence) with a negative knowledge claim but failed to disagree with the positive knowledge claim (either by agreeing with it or remaining neutral). Participants were classified as endorsing closure if they did anything else. Participants' responses are summarized in Figure 1.

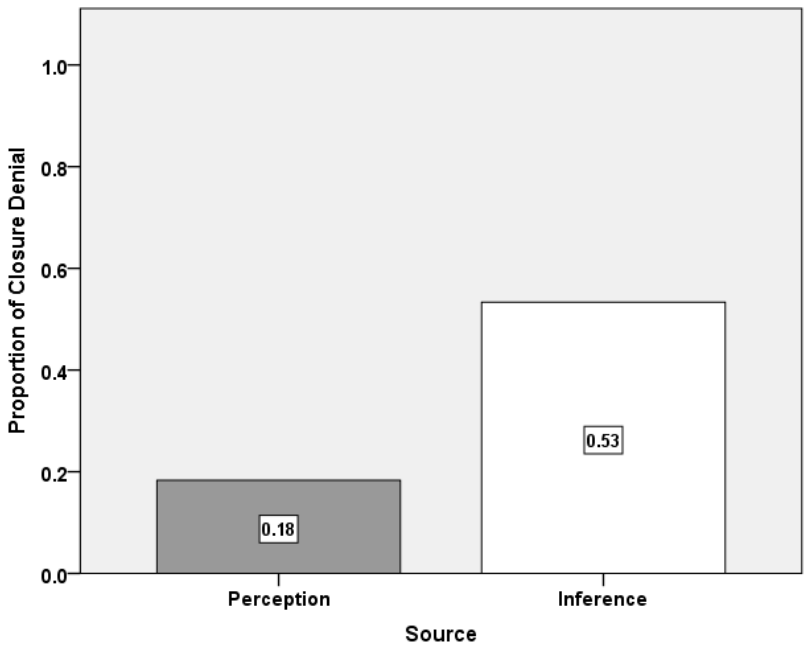


Figure 1. Proportions of participants who rejected closure in the perception and inference conditions of Study 1.

82% of participants in the perception condition attributed knowledge in accordance with the closure principle, while 53% of participants in the inference condition attributed knowledge in a way that was inconsistent with the principle. These differences are statistically significant, with a medium effect size.[11][12]

Study 1 thus replicates Turri's initial finding that participants treat positive and negative knowledge claims very differently in the two conditions of Car Theft. Our results are also consistent with the hypothesis that closure applies to perceptually based beliefs but not to inferential beliefs.

3. Study 2

As we noted above, we harbored doubts about whether Car Theft was a good case to use to test for individuals' endorsement of closure for perceptual beliefs. In the perception version of Car Theft, when Maxwell "looks carefully out the window" and sees that his car parked in C8, we think that many participants will think that Maxwell also sees—directly and without the aid of inference—that it has not been stolen. If this is how participants interpret the vignette, it will fail to serve as a test for closure.

Therefore, in Study 2, we constructed the following variant of Car Theft in which Maxwell comes to believe a proposition on the basis of perception, recognizes that this proposition entails a second proposition that he cannot directly verify by perception, and comes to believe this second proposition on the basis of the recognized entailment:

> *Replica*. Mr. Maxwell lives in a world where wizards like to practice their magical skills by destroying ordinary objects and replacing them with exact replicas. One problem with

[11] χ^2 (1, N = 120) = 15.98, $p < .001$, Cramér's V = .37.

[12] There were no gender differences in participants' rates of the endorsement or rejection of closure in either the perception or the inference conditions. Perception: χ^2 (1, N = 60) = .89, $p > .05$. Inference: χ^2 (1, N = 59) = .32, $p > .05$. There was a positive correlation between age and the endorsement of closure in the perception condition but not in the inference condition. Perception: $r = .29$, $p < .05$. Inference: $r = .10$, $p > .05$.

this is that while the wizards create exact replicas of the objects they destroy, they do not always create replicas of the internal contents of those objects. So, if they destroy your wallet and replace it with a replica, the replacement may not have copies of your money or credit cards in it.

When Mr. Maxwell arrives at work in the morning, he parks his car in his usual parking spot, space C8. At lunchtime, Maxwell and his assistant are up in the archives room searching for a particular document. Maxwell says, “I might have left the document in my car.” The assistant asks, “Mr. Maxwell, is your car still parked in space C8? It’s not unheard of for wizards to destroy ordinary objects and replace them with exact replicas.”

The perception and inference versions of Replica end with one of the following sentences:

Maxwell looks carefully out the window and then responds, “My car is parked in C8, so it has not been destroyed and replaced with an exact replica.”

Maxwell thinks carefully for a moment and then responds, “My car is parked in C8, so it has not been destroyed and replaced with an exact replica.”

Like Car Theft, Replica features Maxwell wondering where a particular document is and thinking that he might have left it in his car. His assistant is again portrayed as raising a certain kind of error possibility—this time that it has been destroyed and replaced by an exact replica rather than that it has been stolen—after which Maxwell confidently affirms that he knows his car’s whereabouts. However, it is much less plausible to think that in the perception condition of Replica Maxwell can directly perceive that the error possibility in question does not obtain. We think this places the perception and inference conditions on a more equal footing in Replica than they are in Car Theft.

After reading one of the versions of Replica, participants were asked the following comprehension question:

> Do the wizards always create replicas of the internal contents of the objects they destroy and recreate?

Participants were directed to select either ‘Yes’ or ‘No.’ Those who failed to answer the question correctly had their responses to the positive and negative knowledge claims excluded from the statistical analysis. Participants were then asked to indicate the extent to which they agreed or disagreed with both (2.1) and the following claim, the order of which was counterbalanced:

> (2.5) Maxwell knows that his car has not been destroyed and replaced with an exact replica.

Participants were given the same answer choices as in Study 1, and their responses were coded in the same fashion.

Using a between-subjects design, 120 participants (average age = 37, 45% female, 71% Caucasian, 95% native English speakers) were recruited from Amazon’s Mechanical Turk and were presented with either the perception or the inference condition of Replica. 11 participants failed to answer the comprehension question correctly. Responses from the remaining participants to the positive and negative knowledge claims are summarized in Table 2.

	Perception	**Inference**
Positive Knowledge	4.28 (2.27)	4.58 (2.16)
Negative Knowledge	4.08 (2.48)	4.50 (2.30)

Table 2. Mean levels of agreement (and standard deviations) for each knowledge claim in the perception and inference conditions of Study 2.

In sharp contrast to the results of Study 1 and Turri's original studies, none of the mean levels of agreement differed significantly from the neutral midpoint[13], and neither the source of belief nor the valence of the knowledge claim had any appreciable impact upon participants' responses.[14] In other words, on the whole participants were not inclined to think that Maxwell knows either that his car is parked in C8 or that it has not been destroyed and replaced with an exact replica. They were also not significantly inclined to deny these knowledge claims.[15]

As in Study 1, we grouped 'Slightly Agree,' 'Mostly Agree,' or 'Completely Agree' responses into a broad knowledge-affirming category and 'Completely Disagree,' 'Mostly Disagree,' or 'Slightly Disagree' into a knowledge-denying category, with 'Neither Agree Nor Disagree' responses falling into a Neutral category. When we looked at how often participants endorsed or rejected closure, we failed to observe the patterns that the source-relative closure hypothesis would lead us to expect (cf. Figure 2).

[13] Perception/positive knowledge: $t(52) = .91$, $p > .05$. Perception/negative knowledge: $t(52) = .22$, $p > .05$. Inference/positive knowledge: $t(54) = 2.00$, $p = .051$. Inference/negative knowledge: $t(55) = 1.62$, $p > .05$.

[14] Source: $F(1, 106) = 1.46$, $p > .05$. Valence: $F(1, 106) = .19$, $p > .05$. Source x valence: $F(1, 106) = .04$, $p > .05$.

[15] There were no gender effects. A three-way (source x valence x gender) mixed ANOVA showed no main effect for gender, $F(1, 104) = 1.62$, $p > .05$. A three-way (source x valence x order) mixed ANOVA showed no main effect for order, $F(1, 104) = .72$, $p > .05$. And age did not correlate with agreement to positive or negative knowledge claims in either the perception or inference conditions. Perception/positive knowledge: $r = .15$, $p > .05$. Perception/negative knowledge: $r = -.06$, $p > .05$. Inference/positive knowledge: $r = -.04$, $p > .05$. Inference/negative knowledge: $r = .14$, $p > .05$.

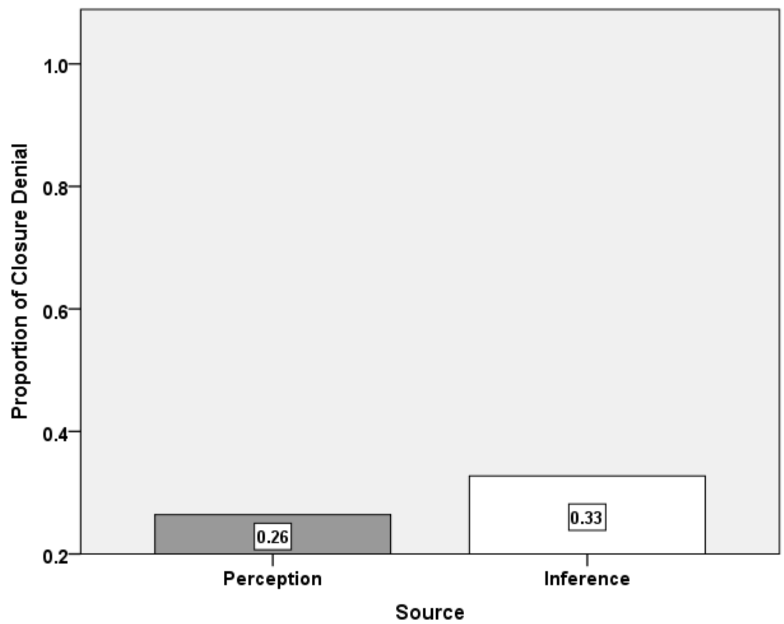


Figure 2. Proportions of participants who rejected closure in the perception and inference conditions of Study 2.

Roughly three-fourths of participants in the perception condition and two-thirds in the inference condition attribute knowledge in ways that were consistent with closure. The differences between these distributions failed to be significant.[16][17] When we compared the perception data from Study 1 to the perception data from Study 2 and the inference data from Study 1 to the inference data from Study 2 using two chi squared tests of independence, we found that the frequency of closure denial in Study 2 was not significantly different than in Study 1 in the perception condition but was significantly lower in the inference condition, with a small effect size.[18]

[16] χ^2 (1, N = 108) = .52, p > .05.

[17] There were no gender differences in participants' acceptance or rejection of closure and no correlations between age and closure. Gender/perception: χ^2 (1, N = 53) = .68, p > .05. Gender/inference: χ^2 (1, N = 55) = .08, p > .05. Age/perception: r = -.08, p > .05. Age/inference: r = .16, p > .05.

[18] Perception: χ^2 (1, N = 113) = 1.07, p > .05. Inference: χ^2 (1, N = 115) = 4.96, p < .05, Cramér's V = .21.

It is crucial to note that the positive knowledge claim used in Study 2 is identical to the positive knowledge claims used in previous studies, and Maxwell's evidence for this claim is identical as well. The only things that have changed are the relevant alternative that is incompatible with Maxwell knowing where his car is parked, Maxwell's inability in the perception condition to have direct perceptual access to the falsity of this alternative, and how difficult it appears to be for Maxwell to rule out this alternative with his present evidence. Participants seemed to realize that Maxwell could not rule out the possibility of a wizarding deception and thus was not in a position to know that (2.5) was true. And in accordance with closure, they appeared to reason that Maxwell was therefore not in a position to know that (2.1) was true either. The source-relative closure hypothesis seems unable to explain why most of our participants attributed knowledge in accordance with the closure principle in both the perception and inference conditions in Study 2 and why the wizarding relevant alternative led them to attribute knowledge at lower rates than in previous studies.

As we argued above, we believe that the perception condition of Turri's original Car Theft vignette does not serve as a good test case for closure-based knowledge. When Maxwell can look out the window and see that his car is parked in C8, participants may not think there is a need for him to infer that it has not been stolen. They may view him as being able to see this directly. We believe that Replica serves as a better test for the endorsement of closure for perceptual beliefs because Maxwell is unable to directly perceive whether or not the relevant alternative in question obtained.

However, in their comments on a previous draft of this paper presented at a recent conference, Turri (2016) and Buckwalter (2016) raise some concerns about Study 2. One of Buckwalter's concerns is the following:

> In any event, knowledge was not initially attributed in these cases, so an optimal test of closure cannot be performed. How can one evaluate what follows from a knowledge attribution in folk psychology if no attribution is present in the first place?

Focusing on the fact that participants were not on the whole inclined to attribute positive knowledge in Replica, Buckwalter expresses doubts about Study 2's ability to test whether participants' think that knowledge is closed under known entailment.

However, we do not believe that the mean participant response to the positive knowledge claim is the right place to look to determine whether we have constructed an adequate test for closure. For one thing, individuals' acceptance of a conditional should be tested in two different ways. One should look not only at whether those who endorsed the antecedent also endorsed the consequent. One should also look at whether those who rejected the consequent also rejected the antecedent. The data represented in Figure 2 take into account both of these possibilities. Furthermore, the data behind Figure 2 focus on the level of individual responses rather than on the broad level of group means. The analysis behind this figure looks at whether each individual who attributed positive knowledge also attributed negative knowledge and whether each individual who denied negative knowledge also denied positive knowledge. This analysis does not attempt to compute anything about closure denial on the basis of mean responses. On the basis of this individual-level analysis, we can see that three-fourths of participants in the perception condition and two-thirds in the inference condition attribute knowledge in ways that were consistent with closure. Thus, we do not think that Study 2 fails to provide an adequate test for closure for the reason in question.

Turri (2016) and Buckwalter (2016) both raise an additional concern that we think poses an interesting challenge to Study 2. Turri (2016) contends that "the wizardry scenario was conspicuously extravagant and unrealistic," suggesting that this might have "incited ambivalence, thereby masking the underlying effect." Similarly, Buckwalter (2016) writes, "Stemming from the fanciful nature of the stimuli, there could be many reasons why attribution in magical worlds occurred at chance rates incidental to closure." In an effort to examine the possible effects that the unrealistic Harry Potter-esque features of Replica might have had on participant responses in Study 2, we ran a follow-up study that was free of any non-Muggle elements.

4. Study 3

We have argued that the best way to test for the endorsement or denial of closure in cases of perceptually based belief is to have the agent in question form a belief in a proposition on the basis of perception, recognize that this proposition entails a second proposition, and yet be unable to tell directly by perception alone whether the second proposition is true. The Replica story satisfies this requirement, but Car Theft and Computer do not. For Study 3, we constructed two additional vignettes that satisfy this same requirement but lack the allegedly "extravagant" or "fanciful" characteristics of Replica.

The first story is based upon a case that has been widely discussed in the mainstream epistemological literature on Gettier cases and skepticism and that has been studied by some experimental philosophers (e.g., Nagel et al. 2013, Alexander et al. 2014):

> *Table*. James and Madeline are in a furniture store, shopping for a new table. James looks at a bright red table under normal lighting conditions and forms the belief that the table is red. After walking to the other side of the store and looking at several other pieces of furniture, James and Madeline begin talking about the red table again. Madeline says, "A

white table under red lighting conditions would look exactly the same to you, and you have not checked whether the lighting is normal, or whether there might be a red spotlight shining on the table."

The perception version of Table ends with the following:

James walks back to the table, looks at it carefully, and then responds, "The table is red, so it is not a white table that has been illuminated to look red."

The inference version ends as follows:

James does not walk back to the table, but he thinks carefully for a moment and then responds, "The table is red, so it is not a white table that has been illuminated to look red."

Participants in both conditions were asked to indicate the extent to which they agreed or disagreed with the following two statements, the order of which was counterbalanced:

(4.1) James knows that the table is red.

(4.2) James knows that the table is not a white one that has been illuminated to look red.

Participant responses were recorded on a seven-point in the same fashion as in previous studies.

The second vignette used in Study 3 was derived from one used by John Waterman, Chad Gonnerman, Karen Yan, and Joshua Alexander (forthcoming):

Jaguar. Christine and Aleah go to the zoo. As they walk around, they pause in front of an exhibit marked "Brazilian Jaguar Enclosure." Christine reads about jaguars from the sign, and looks out and sees a jaguar sleeping on the branch of a tree in the enclosure. After walking to the other side of the zoo and looking at several other exhibits, Christine and Aleah begin talking about the jaguar again. Aleah says, "African leopards look very much like Brazilian jaguars, and the signs in the zoo have recently been replaced by an inexperienced crew of workers. If a worker had accidentally switched the signs on their exhibits, you wouldn't be able to tell the difference between a jaguar and a leopard."

The perception and inference versions end with one the following two sentences:

Christine walks back to the jaguar exhibit, looks carefully at the animal sleeping in the tree, and then responds, "The animal is a jaguar, so it is not a leopard."

Christine does not walk back to the jaguar exhibit, but she thinks carefully for a moment and then responds, "The animal is a jaguar, so it is not a leopard."

Participants in both conditions were asked to indicate the extent to which they agreed or disagreed with the following statements, the order of which was counterbalanced:

(5.1) Christine knows that the animal in the pen is a jaguar.

(5.2) Christine knows that the animal in the pen is not a leopard.

100 workers (average age = 37, 47% female, 75% Caucasian, 97% native English speakers) from Amazon's Mechanical Turk were presented with either the perception version of Table and the inference version of Jaguar or the inference version of Table and the perception version of Jaguar. The order of presentation of the vignettes was counterbalanced. Participant responses to Table and Jaguar are summarized in Table 3.

	Table		**Jaguar**	
	Perception	**Inference**	**Perception**	**Inference**
Positive Knowledge	5.54*** (1.63)	5.10*** (1.84)	4.62* (1.73)	4.39 (1.99)
Negative Knowledge	5.14*** (1.89)	5.02*** (1.88)	4.36 (1.79)	3.94 (1.96)

Table 3. Mean levels of agreement (and standard deviations) for each knowledge claim in the perception and inference conditions of Study 3.

Mean levels of agreement with (4.1) and (4.2) fell significantly above the neutral midpoint in all cases, with large effect sizes.[19] With only one exception, mean levels of agreement with (5.1) and (5.2) did not differ significantly from the midpoint.[20] A three-way (vignette x source x valence) mixed ANOVA revealed a small main effect for vignette.[21] In other words, the difference between the typical level of agreement with knowledge attributions in response to Table and the typical level of agreement with knowledge attributions in Jaguar was significant.

Importantly, and in contrast to Turri's studies and Study 1 (but not Study 2), the source of the agents' beliefs in Table and Jaguar made no difference to participants' level of agreement with the four knowledge attributing statements.[22] In other words, participants were neither more nor less inclined to attribute knowledge in Table and Jaguar when the belief was perceptually based than when it was inferential. There was also no interaction between vignette and source.[23]

There was a small, significant main effect for valence.[24] In other words, participants were somewhat more inclined to agree with positive knowledge claims than negative knowledge claims. There were no significant interactions between valence and vignette, valence and source, or valence, vignette, and source.[25][26]

---

[19] Table/perception/positive knowledge: $t(49) = 6.68$, $p < .001$, $r = .69$. Table/perception/negative knowledge: $t(49) = 4.28$, $p < .001$, $r = .52$. Table/inference/positive knowledge: $t(49) = 4.22$, $p < .001$, $r = .52$. Table/inference/negative knowledge: $t(49) = 3.84$, $p < .001$, $r = .48$.

[20] Jaguar/perception/positive knowledge: $t(49) = 2.54$, $p < .05$, $r = .34$. Jaguar/perception/negative knowledge: $t(49) = 1.42$, $p > .05$. Jaguar/inference/positive knowledge: $t(48) = 1.37$, $p > .05$. Jaguar/inference/negative knowledge: $t(49) = -.22$, $p > .05$.

[21] $F(1, 195) = 12.39$, $p < .01$, $r = .24$.

[22] $F(1, 195) = 1.46$, $p > .05$.

[23] $F(1, 195) = .01$, $p > .05$.

[24] $F(1, 195) = 11.45$, $p < .01$, $r = .24$.

[25] Valence x vignette: $F(1, 195) = .37$, $p > .05$. Valence x source: $F(1, 195) = .19$, $p > .05$. Valence x vignette x source: $F(1, 195) = 2.00$, $p > .05$.

[26] There was no significant main effect for gender, and there were no significant interaction effects for gender. Main effect: $F(1, 195) = 2.29$, $p > .05$. Gender x vignette: $F(1, 195) = 1.35$, $p > .05$. Gender x source: $F(1, 195) = .00$, $p > .05$. Gender x valence: $F(1, 195) = .00$, $p > .05$. Age did not correlate with either positive or negative

In regard to the endorsement or denial of closure, the results of Study 3 are quite striking (cf. Figure 3).

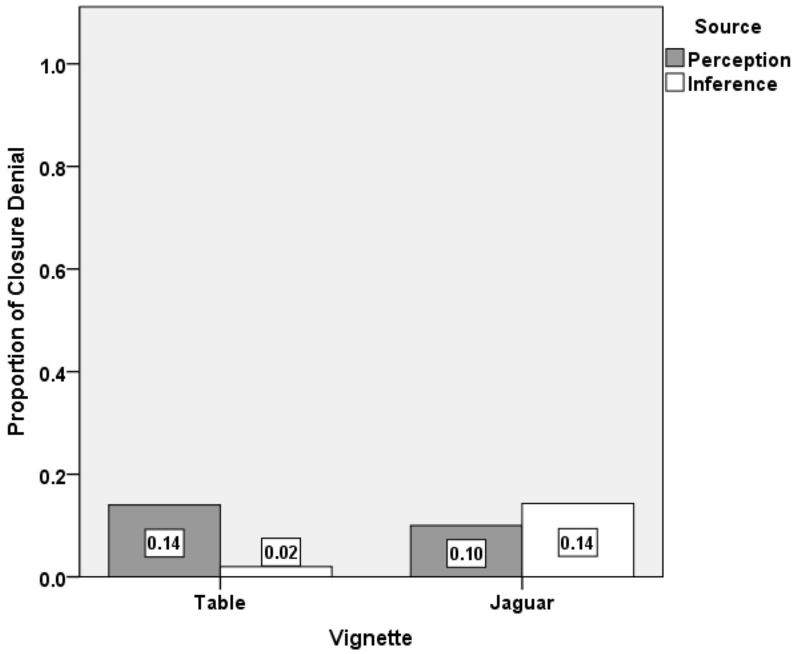


Figure 3. Proportions of participants who rejected closure in the perception and inference conditions of Table and Jaguar in Study 3.

Participants overwhelmingly attributed or denied knowledge in accordance with the closure principle for knowledge in both the perception and the inference conditions of Table and Jaguar. Furthermore, the degree to which their knowledge judgments accorded with closure was higher than in previous studies. A chi squared test of independence that combines data from Table and

knowledge claims in either the perception or the inference conditions. There was, however, a significant order effect, with participants being (surprisingly) slightly more likely to agree with both knowledge claims when the negative knowledge claim was presented first. Four-way (vignette x source x valence x order) mixed ANOVA: $F(1, 191) = 9.79$, $p < .01$, $r = .22$.

Jaguar reveals no significant difference between rates of closure denial in the perception and inference conditions.[27]

Recall that Turri (2016) and Buckwalter (2016) were concerned that the magical features of Replica might render it unsuitable as a tool to test for ordinary epistemic judgments. However, when we constructed cases with the same overall structure as Replica but without these magical elements, our results did not change. Turri and Buckwalter also expressed doubts about what could be concluded from our Replica data because participants were not significantly inclined to agree with the positive knowledge claim in question. However, participants who were presented with Table were strongly inclined to agree with the positive knowledge claim in both the perception and the inference conditions. This means that our Table data are the sort of data that Turri and Buckwalter think should be used to test for closure. Yet we did not find that participants' knowledge attributions and denials conflicted with closure.

For reasons described above, we contend that Car Theft and Computer do not represent good tests of closure for perceptual beliefs. When we constructed three vignettes that are capable of serving as better tests for closure in these conditions, we found that participants do not deny closure for perceptual beliefs. This result, of course, is one that Turri (2015a) already claims to have established. Nonetheless, we think that the result is on a stronger footing now that better tools have been used to study it.

Significantly, we failed to find the same kind of violations of closure for inferential beliefs that Turri obtained and that we observed in Study 1 using Turri's materials. The source-relative closure hypothesis suggested but not fully endorsed by Turri (2015a) should not be misinterpreted as the thesis that individuals never attribute or deny knowledge in ways that are

[27] χ^2 (1, N = 199) = .85, $p > .05$.

consistent with the closure principle. Because the closure principle is supposed to apply to all pairs of propositions, the rejection of closure simply means that it does not always hold. Nevertheless, an important question is why individuals would deny closure in the inference conditions of Car Theft and Computer. A limitation of Studies 2 and 3 is that they cannot directly speak to this question. Therefore, we designed an additional study that explored the possibility that participants might be making a performance error when they deny closure.

5. Study 4

In Section 1, we noted that Turri (2015a, 15) considers the possibility that epistemic closure is "too subtle or complicated for ordinary people to understand, which is why they exhibit patterns that violate it." Although Turri does not think this possibility underlies the actual pattern of results he obtained, we suggested that non-philosophers might failed to attribute knowledge in accordance with the closure principle either because they reject closure or because they lacked sufficient skill in handling whatever subtlety or complexity is involved in reasoning in accordance with the closure principle. We hypothesized that if closure were a central feature of folk epistemology, non-philosophers who had formal training in deductive reasoning would thus attribute knowledge in ways that were more consistent with the closure principle than the average undergraduate or MTurk worker who lacked the relevant training.

Consequently, for Study 4, we recruited two groups of participants. One group was composed of 254 undergraduates from a large public institution in the northeastern United States (average age = 20, 63% female, 61% Caucasian, 81% native English speakers). The second set of participants consisted of 208 professional mathematicians from large and regional public institutions across the United States (average age = 48, 14% female, 73% Caucasian, 73% native

English speakers). The mathematicians were recruited via personal email invitation. Their emails were obtained from the webpages of their respective departments.[28] In addition to being trained in formal methods of reasoning, we also presumed these mathematicians would be more intelligent than the average person and thus would commit fewer performance errors in their application of their concept of knowledge than ordinary participants.

Each participant in Study 4 was given two vignettes to read and evaluate. Participants received either the perception version of Car Theft and the inference version of Computer or the inference version of Car Theft and the perception version of Computer. The order with which participants responded to these vignettes was counterbalanced. Each participant who received a Car Theft vignette was then asked to indicate the extent to which they agreed or disagreed with (2.1) and (2.2), and each participant who received some version of Computer was asked whether they agreed or disagreed with (3.2) and (3.3). The order of these statements was also counterbalanced. Participant responses were recorded in the same fashion as in the studies above and are summarized in Table 4.

		Car Theft (Perception)	**Car Theft (Inference)**	**Computer (Perception)**	**Computer (Inference)**
Undergraduates	**Positive Knowledge**	6.16*** (1.32)	5.70*** (1.83)	5.74*** (1.42)	5.26*** (1.86)
	Negative Knowledge	5.19*** (1.89)	3.52* (2.12)	4.90*** (1.93)	3.92 (1.99)
Mathematicians	**Positive Knowledge**	6.14*** (1.34)	4.13 (2.47)	5.14*** (2.02)	3.69 (2.37)
	Negative Knowledge	5.44*** (1.89)	2.46*** (1.97)	4.65** (2.14)	2.93*** (2.09)

Table 4. Mean levels of agreement (and standard deviations) for each participant group and each knowledge claim in the perception and inference conditions of Study 4.

[28] 58% of the mathematicians claimed some area of discrete mathematics as their primary area of specialization, with 14% in analysis and 18% in applied mathematics.

The mean levels of agreement differed from the neutral midpoint in thirteen of the sixteen cases.[29] Ten of these thirteen were above the midpoint, and three were below. A four-way mixed ANOVA revealed a significant main effect for each of the following independent variables: group (undergraduates vs. mathematicians), vignette (Car Theft vs. Computer), source (perception vs. inference), and valence (positive vs. negative).[30] In other words, (i) there were significant differences in how undergraduates and mathematicians responded to the research materials (with mathematicians reporting lower levels of agreement with all knowledge claims), (ii) participants were less inclined to attribute knowledge in Computer than in Car Theft, (iii) participants attributed knowledge less often in the inference conditions than in the perception conditions, and (iv) participants expressed lower levels of agreement with negative knowledge claims than with positive knowledge claims.

The following two-way interactions were also significant: group x source, group x valence, vignette x source, vignette x valence, and valence x source.[31] In other words, the two

---

[29] Undergraduates/Car Theft/perception/positive knowledge: $t(127) = 18.54$, $p < .001$, $r = .85$ Undergraduates/Car Theft/perception/negative knowledge: $t(95) = 7.45$, $p < .001$, $r = .61$. Undergraduates/Car Theft/inference/positive knowledge: $t(125) = 10.42$, $p < .001$, $r = .68$. Undergraduates/Car Theft/inference/negative knowledge: $t(125) = -2.53$, $p < .05$, $r = .22$. Undergraduates/Computer/perception/positive knowledge: $t(125) = 13.73$, $p < .001$, $r = .78$. Undergraduates/Computer/perception/negative knowledge: $t(124) = 5.19$, $p < .001$, $r = .42$. Undergraduates/Computer/inference/positive knowledge: $t(127) = 7.66$, $p < .001$, $r = .56$. Undergraduates/Computer/inference/negative knowledge: $t(127) = -.45$, $p > .05$. Mathematicians/Car Theft/perception/positive knowledge: $t(96) = 15.78$, $p < .001$, $r = .85$. Mathematicians/Car Theft/perception/negative knowledge: $t(95) = 7.45$, $p < .001$, $r = .61$. Mathematicians/Car Theft/inference/positive knowledge: $t(109) = .54$, $p > .05$. Mathematicians/Car Theft/inference/negative knowledge: $t(109) = -8.20$, $p < .001$, $r = .62$. Mathematicians/Computer/perception/positive knowledge: $t(110) = 5.96$, $p < .001$, $r = .49$. Mathematicians/Computer/perception/negative knowledge: $t(110) = 3.20$, $p < .01$, $r = .29$. Mathematicians/Computer/inference/positive knowledge: $t(96) = -1.29$, $p > .05$. Mathematicians/Computer/inference/negative knowledge: $t(95) = -5.02$, $p < .001$, $r = .46$.

[30] Group: $F(1, 912) = 44.66$, $p < .001$, $r = .22$. Vignette: $F(1, 912) = 8.31$, $p < .01$, $r = .10$. Source: $F(1, 912) = 180.97$, $p < .001$, $r = .41$. Valence: $F(1, 912) = 280.34$, $p < .001$, $r = .48$.

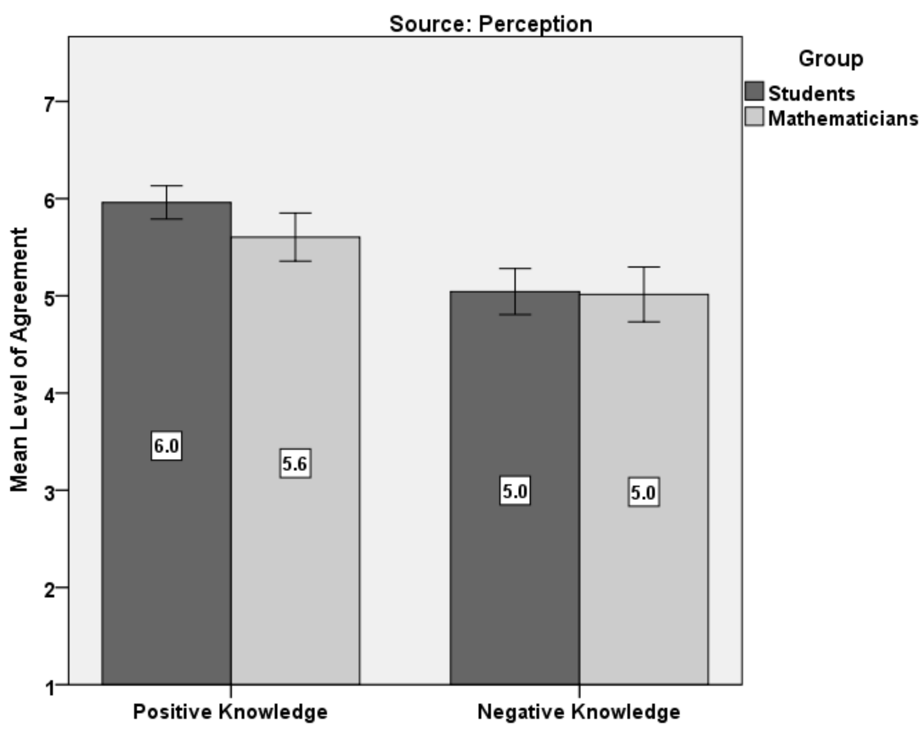
[31] Group x source: $F(1, 912) = 27.40$, $p < .001$, $r = .17$. Group x valence: $F(1, 912) = 10.70$, $p < .01$, $r = .11$. Vignette x source: $F(1, 912) = 7.79$, $p < .01$, $r = .09$. Vignette x valence: $F(1, 912) = 15.26$, $p < .001$, $r = .13$. Valence x source: $F(1, 912) = 29.25$, $p < .001$, $r = .18$. There was also a significant three-way interaction between vignette, source, and valence: $F(1, 912) = 7.27$, $p < .01$, $r = .09$.

groups of participants responded differently to information regarding the source of the protagonist's belief and to the valence of each knowledge claim. Furthermore, how participants were affected by the valence of the knowledge claims determined in part by which vignette they read and the source of the protagonist's belief.[32]

The most important finding at this stage of analysis concerns the very different ways that undergraduates and mathematicians responded in the two inference conditions. In the inference versions of Car Theft and Computer, undergraduates strongly agreed with the positive knowledge claims that Maxwell knows that his car is parked in C8 and that Palmer knows that the computer is analyzing data. By contrast, mathematicians' responses to these knowledge claims did not differ from chance. In response to the negative knowledge claims that Maxwell knows his car has not been stolen and that Palmer knows that hackers have not crashed the computer, undergraduates' responses did not differ from chance, while the responses of mathematicians fell significantly below chance. In short, mathematicians were much more skeptical about the protagonist's ability to obtain knowledge of the propositions in question simply by thinking carefully about the matter for a moment. We think this is the correct perspective to take on the inference-based knowledge claims, given the lower reliability of thinking carefully for a moment, as compared to looking carefully out the window directly at the objects in question.

[32] Analyzing our two participant groups separately, among the mathematicians there was no main effect for gender and no significant interaction between gender and source, but there was a small but significant interaction between gender and valence. Gender: $F(1, 403) = .72$, $p > .05$. Gender x source: $F(1, 403) = .07$, $p > .05$. Gender x valence: $F(1, 403) = 4.58$, $p < .05$, $r = .11$. For undergraduates, there was no significant main effect for gender (although it approached significance) and no interaction effect between gender and source, but there was again a significant interaction between gender and valence. Gender: $F(1, 503) = 3.81$, $p = .052$, $r = .09$. Gender x source: $F(1, 503) = .08$, $p > .05$. Gender x valence: $F(1, 503) = 11.21$, $p < .01$, $r = .15$. Both significant interaction effects manifested themselves with females indicating levels of agreement with positive and negative knowledge claims that were further apart than the levels of agreement males indicated for positive and negative knowledge claims. There were no significant correlation between age and either positive or negative knowledge ratings among either group of participants in either the perception or the inference conditions.

Nevertheless, mathematicians continued to indicate greater levels of agreement with positive knowledge claims than with negative knowledge claims (cf. Figures 4 and 5, which combine data from the Car Theft and Computer vignettes). As predicted, however, the difference between the mean levels of agreement expressed toward positive and negative knowledge claims was smaller for mathematicians than for undergraduates. In the perception conditions, the difference between undergraduates' mean level of agreement with positive and negative knowledge claims was 1.0, whereas for mathematicians it was 0.6. In the inference conditions, the difference between undergraduates' ratings of positive and negative knowledge claims was 1.8, while for mathematicians it was 1.2.[33]



[33] Our findings provide a modest degree of support for Turri's (2015b) source-content bias hypothesis, inasmuch as there was a significant interaction between source and valence in the mathematicians' responses to the positive and negative knowledge claims ($F(1, 411) = 10.22$, $p < .01$, $r = .16$.). Although the effect size was smaller than for other participant groups, the fact remains that mathematicians indicated lower levels of agreement with negative inferential beliefs, as compared to other kinds of beliefs.

Figure 4. Mean levels of agreement with positive and negative knowledge claims by undergraduates and mathematicians in the perception conditions of Study 4. Error bars represent 95% confidence intervals in all figures.

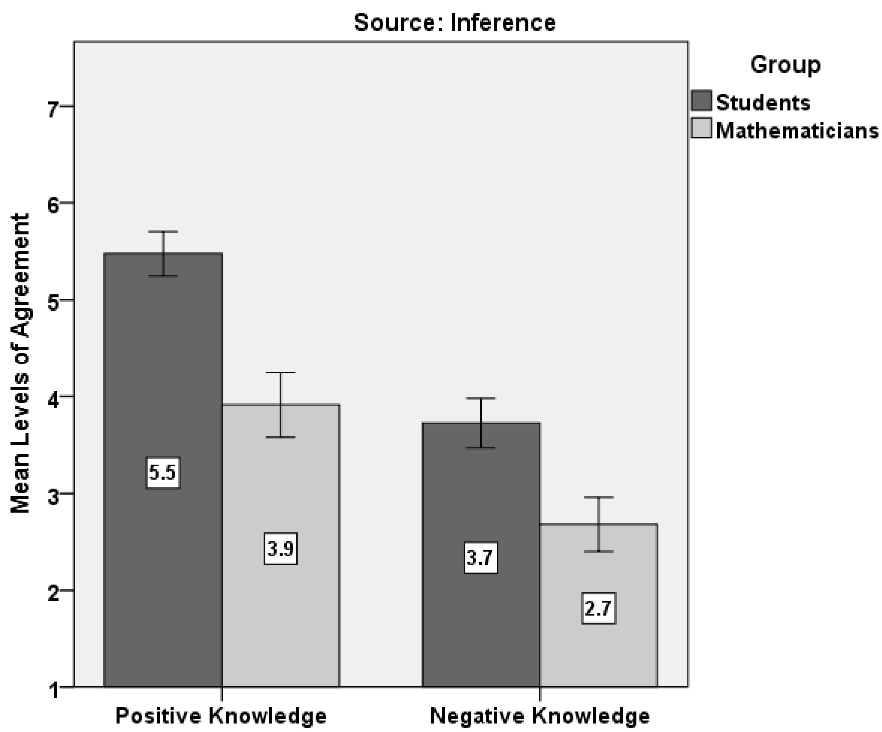


Figure 5. Mean levels of agreement with positive and negative knowledge claims by undergraduates and mathematicians in the inference conditions of Study 4.

Combining data from both the Car Theft and the Computer vignettes and assessing participants' endorsement or rejection of closure leads to the results depicted in Figure 6.

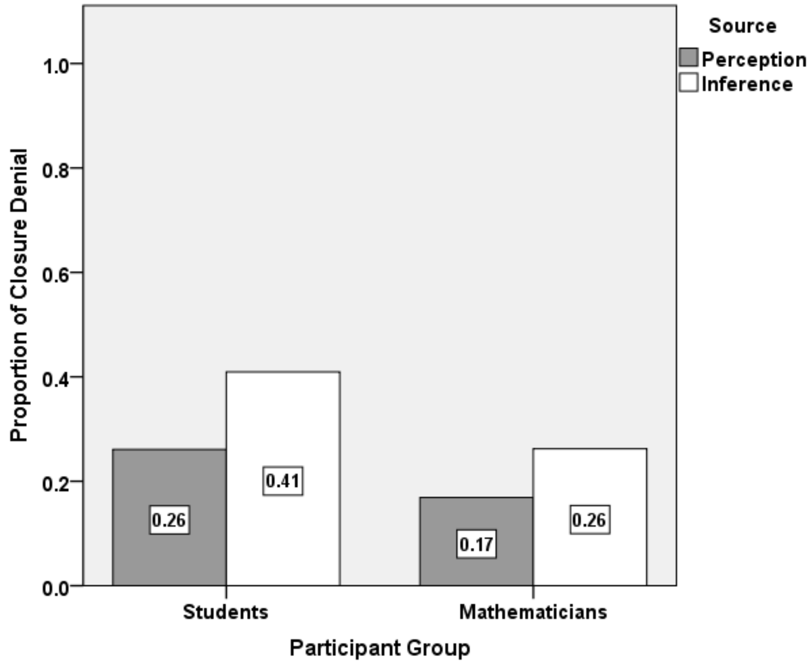


Figure 6. Proportions of undergraduates and mathematicians who rejected closure in the perception conditions of Study 4.

Mathematicians were much less likely to reject closure in either the perception or inference conditions. Two chi squared tests of independence conducted separately on the perception and inference data revealed that the differences between undergraduates and mathematicians in both conditions were significant.[34] Mathematicians did reject closure at slightly higher rates in the inference conditions than in the perception conditions, but note that mathematicians' rejection of closure in the inference conditions was roughly equivalent to undergraduates' rejection of closure in the perception conditions of this and other studies. On the

[34] Perception: χ^2 (1, N = 460) = 5.60, p < .05, Cramér's V = .11. Inference: χ^2 (1, N = 460) = 10.95, p < .01, Cramér's V = .15.

whole, then, mathematicians strongly endorsed closure for both perceptual and inferential beliefs.[35]

According to the source-relative closure hypothesis articulated by Turri (2015a), our shared epistemic practices include a closure principle that applies to perceptually based beliefs but not to inferential beliefs. The mathematicians that participated in Study 4 are folk-like in that they are not actively involved in philosophical debates about closure and related matters such as skepticism, but they are unlike ordinary folk in regard to their expertise in deductive reasoning. If the source-relative closure hypothesis were true, there should not be any reason why their formal training should lead them to depart from the folk epistemological dictum that closure only applies to perceptually based beliefs. Our findings do not accord with this prediction. In fact, they provide significant support for the claim that an unrestricted closure principle figures importantly in folk epistemology.

6. Conclusion

The three most important features of the studies we report in this contribution are the following. First, we constructed three vignettes in Studies 2 and 3 that provide better tests for the folk endorsement of closure for perceptual beliefs than Turri's (2015a) original Car Theft and Computer vignettes. We found that individuals' knowledge attributions in the perception conditions of these studies strongly conform to closure. Because of concerns we had about the way Turri designed his perception conditions, we had doubts about whether his studies showed

[35] There were no significant correlations between age and endorsement of closure among undergraduates or mathematicians. Undergraduates: $r = .06$, $p > .05$. Mathematicians: $r = -.05$, $p > .05$. Among undergraduate participants, males were significantly more likely to endorse closure in the inference conditions than females, but there were no significant differences in the perception conditions. Perception: $r = .09$, $p > .05$. Inference: $r = .16$, $p < .05$. Among mathematicians, there were no significant correlations between gender and closure, although the correlation in the inference conditions approached significance. Perception: $r = .04$, $p > .05$. Inference: $r = .13$, $p = .07$.

that individuals endorsed closure for perceptual beliefs. We believe that the conclusion that they do so is now on a stronger footing in light of these new studies. We would also like to echo Turri’s (2015a, 14) claim that “[t]his is ironic given that the most famous attempted counterexamples [to closure] feature perceptually based beliefs (e.g. Dretske’s zebra case).”

Secondly, in the inference conditions of Studies 2 and 3 our participants endorsed closure for beliefs that we thought would be strongly analogous to the inferential beliefs in Car Theft and Computer, where Turri (2015a) and we observed violations of closure. The perception conditions of Replica, Table, and Jaguar were supposed to represent improvements over earlier materials, but we did not intend the inference conditions to be significantly different. The only factor that we think might be responsible for this difference is the fact that the relevant alternatives in Replica, Table, and Jaguar (viz., Maxwell’s car being destroyed and replaced with an exact replica by wizards, a white table being illuminated to look red, and a leopard looking like a jaguar to Christine) seem to be more difficult for the agents in question to rule out on the basis of the evidence available to them than the ones in Car Theft and Computer (viz., a car being stolen, hackers crashing a computer). We hypothesis that relevant alternatives that are more difficult to rule out make an agent’s epistemic shortcomings or inabilities more salient than alternatives that are less difficult to rule out. This increased salience then leads participants to decrease knowledge attributions all around. However, it should be admitted that this hypothesis requires further investigation before we can place much confidence in it.

Thirdly, we found that the knowledge attributions and denials of non-philosophers who are folk-like with respect to debates about closure and skepticism but who are experts in deductive reasoning strongly adhered to the closure principle for knowledge. The mathematicians in Study 4 endorsed closure for inferential beliefs at rates similar to those observed with MTurk workers and undergraduate populations for perceptual beliefs. We believe that the epistemic behavior of these model folk provide reason for thinking that other individuals are making a performance error when their knowledge attributions and denials conflict with the closure principle. We conclude that the best explanation of these findings is that an epistemic closure principle for knowledge appears to be a central feature of folk epistemology after all.

References

Alexander, Joshua, Chad Gonnerman, and John Waterman. 2014. “Salience and Epistemic Egocentrism: An Empirical Study.” In James R. Beebe (ed.) *Advances in Experimental Epistemology*. Bloomsbury, pp. 97-118.

Brueckner, Anthony. 1985. “Skepticism and Epistemic Closure.” *Philosophical Topics* 13: 89-117.

Brueckner, Anthony. 1994. “The Structure of the Skeptical Argument.” *Philosophy and Phenomenological Research* 54:827-35.

Buckwalter, Wesley. 2016. Invited comments on “Epistemic Closure in Folk Epistemology” by James R. Beebe and Jake Monaghan at the *Second Annual Minds Online Conference*. Sept. 29, 2016. Available at: http://mindsonline.philosophyofbrains.com/2016/2016-3/epistemic-closure-in-folk-epistemology/#comment-370.

DeRose, Keith. 1999. “Introduction: Responding to Skepticism.” In K. DeRose & T. A. Warfield (eds.), *Skepticism: A Contemporary Reader*. Oxford: Oxford University Press, pp. 1-24.

Dretske, Fred. 1970. “Epistemic Operators.” *Journal of Philosophy* 67: 1007-23.

Feldman, Richard. 1995. “In Defence of Closure.” *Philosophical Quarterly* 45: 487-94.

Greco, John. 2008. “Skepticism About the External World.” In J. Greco (ed.), *The Oxford Handbook of Skepticism*. Oxford: Oxford University Press, pp. 108-28.

Hawthorne, John. 2004. *Knowledge and Lotteries*. Oxford: Oxford University Press.

Nagel, Jennifer, Valarie San Juan, and Raymond Mar. 2013. “Lay Denial of Knowledge for Justified True Beliefs.” *Cognition* 129: 652-61.

Nozick, Robert. 1981. *Philosophical Explanations*. Cambridge, MA: Harvard University Press.

Pritchard, Duncan. 2002a. “Contemporary Skepticism.” In J. Fieser & B. Dowden (eds.), *The Internet Encyclopedia of Philosophy*, http://www.iep.utm.edu/skepcont/.

Pritchard, Duncan. 2002b. “Recent Work on Radical Skepticism.” *American Philosophical Quarterly* 39: 215-57.

Pritchard, Duncan. 2005. “The Structure of Sceptical Arguments.” *Philosophical Quarterly* 55: 37-52.

Sosa, Ernest. 2007. “Experimental Philosophy and Philosophical Intuition.” *Philosophical Studies* 132: 99-107.

Turri, John. 2015a. “An Open and Shut Case: Epistemic Closure in the Manifest Image.” *Philosophers' Imprint* 15: 1-18.

Turri, John. 2015b. “Skeptical Appeal: The Source-Content Bias.” *Cognitive Science* 39: 307-24.

Turri, John. 2016. Invited comments on “Epistemic Closure in Folk Epistemology” by James R. Beebe and Jake Monaghan at the *Second Annual Minds Online Conference*. Sept. 29, 2016. Available at: http://mindsonline.philosophyofbrains.com/2016/2016-3/epistemic-closure-in-folk-epistemology/#comment-371.

Turri, John., Wesley Buckwalter, and Peter Blouw. 2015. “Knowledge and Luck.” *Psychonomic Bulletin and Review* 22: 378-390.

Vogel, Jonathan. 1990. "Are There Counterexamples to the Closure Principle?" In Michael D. Roth & Glenn Ross (eds.), *Doubting: Contemporary Perspectives on Skepticism*. Dordrecht: Reidel, pp. 13-27.

Waterman, John, Chad Gonnerman, Karen Yan, and Joshua Alexander. (forthcoming). "Knowledge, Certainty, and Skepticism: A Cross-Cultural Study." In Masaharu Mizumoto, Stephen Stich, Eric McCready, and Jason Stanley (eds.), *Epistemology for the Rest of the World*.

Williamson, Timothy. 2000. *Knowledge and Its Limits*. Oxford: Oxford University Press.

Williamson, Timothy. 2011. "Philosophical Expertise and the Burden of Proof." *Metaphilosophy* 42: 215-229.