Experimental Evidence for the Existence of an External World

Eric Schwitzgebel and Alan T. Moore
Department of Philosophy
University of California at Riverside
Riverside, CA 92521
USA

Corresponding author:
Eric Schwitzgebel, eschwitz at domain ucr.edu.

Experimental Evidence for the Existence of an External World

Abstract:

In this essay I attempt to refute radical solipsism by means of a series of empirical experiments. In the first experiment, I exhibit unreliable judgment about the primeness or divisibility of four-digit numbers, in contrast to a seeming Excel program. In the second experiment, I exhibit an imperfect memory for arbitrary-seeming three-digit number and letter combinations, in contrast to my seeming collaborator with seemingly hidden notes. In the third experiment, I seem to suffer repeated defeats at chess. In all three experiments, the most straightforward interpretation of the experiential evidence is that something exists in the universe that is superior in the relevant respects – theoretical reasoning (about primes), memorial retention (for digits and letters), or practical reasoning (at chess) – to my own solipsistically-conceived self.

Experimental Evidence for the Existence of an External World

It occurs to me to wonder whether the external world exists – that is, whether anything exists other than my own stream of conscious experience. *Radical solipsism,* I'll say, is the view that my conscious mind is the only thing in the universe; there are no material objects, no other minds, not even a hidden unconscious side of myself. On radical solipsism, *this* [here I gesture inwardly at my sensory, emotional, and cognitive experiences] is all there is, nothing more. I find myself, now, contemplating radical solipsism. I want some sort of evidence or proof of its falsity.

You might think – if you exist – that my desire for proof is foolish. You might think it plain that I could never *show* radical solipsism to be false, that I can only *assume* that it's false, that any attempt to prove solipsism wrong would inevitably turn in a circle. You might think, with (my seeming memory of) Wittgenstein (1950-1951/1974), that the existence of an external world is an unchallengeable framework assumption necessary for any inquiry to make sense, that it's a kind of philosophical disease to want to rationally refute solipsism, that I might as well hope to establish the validity of logic using no logical assumptions. Now I'll grant that I might be philosophically sick. There's something admittedly crazy about solipsistic doubt. But it's not entirely clear to me, at least not yet, that I can't find my cure from within the disease, by giving my sick mind exactly the proof it wants.

At first blush, the historical evidence – or what I think of as the historical evidence – looks unpromising. The two most famous attempts to cure solipsism from within come from Descartes, in his *Meditations* (1641/1984)*,* and Kant, in his "Refutation of Idealism" (from his first *Critique*: 1781/1787/1929). Neither appears to work. Descartes's proof of the external

world requires accepting, as an intermediate step, the very dubious claim that the thought of a

perfect God could only arise from a being as perfect as God.[1] Kant's proof turns on the assertion

that I cannot be "conscious of my own existence as determined in time" or conscious of change

in my representations unless I perceive some permanent things that actually exist outside of me.

He offers no clear argument for this assertion. Why couldn't a sense of representational change

and of my determination in time arise innately or from temporally overlapping experiences or

from hallucinatory experiences *as of* seeing things that exist outside of me?[2] Most philosophers

today, it seems, regard as hopeless all such attempts to prove radical solipsism false using only

general logic and solipsism-compatible premises about one's own conscious experience.

So we might, with Hume (1740/1978), yield to the skeptic and grant her the argument,

then turn our minds aside for a while, play some backgammon, and go on living and

philosophizing just as before, only avoiding the question of radical solipsism. Or we might, with

Moore (1939), refute the solipsist by proving that the external world exists by means of

solipsism-*incompatible* premises that beg the question: "Here is a hand, here is another, therefore

there are external things; what, you want stronger proof than *that?*" Or we might, with

Wittgenstein, try to undercut the very desire for proof. However, none of these responses seems

to me to be preferable to actually delivering a non-question-begging proof if one is discoverable.

They are all *fallback maneuvers*. Another type of fallback maneuver, I think, can be found in

---

[1] See discussion in Broughton 2002; Nolan and Nelson 2006. Broughton explicitly notes
that few readers of the *Meditations* have found this aspect of Descartes's argument plausible
(2002: 154).

[2] Kant interpretation is always fraught. However, even sympathetic commentators seem
to agree that Kant's argument is problematic, at least if construed as a self-contained argument
with the positive conclusion that the external world exists. For charitably amended
reconstructions in light of subsequent Kantian texts – excessively charitable, we're inclined to
think, and still unsuccessful as proofs – see Guyer 1987 and Dicker 2008. See also Stroud 1984
and 1994 for a more general critique of Kantian "transcendental" refutations of skepticism.

those recent versions of contextualism and reliabilism that concede to the radical solipsist that we cannot know that the external world exists, once the question of its existence has been raised in a philosophical context, while insisting that we can still nonetheless have ordinary knowledge of the mundane facts of practical life.[3]

The historical landscape has been dominated by the two broad approaches described above. The first approach aims high, hoping to establish with apodictic or deductive or "transcendental" certainty, in a non-question-begging way, that the external world really does exist. The second approach abandons hope of non-question-begging proof, seeking in one way or another to make us comfortable with its absence. But there is a third approach, historically less influential, that has not yet been adequately explored. Its most famous advocate is Bertrand Russell.

Russell writes:

> In one sense it must be admitted that we can never *prove* the existence of things
> other than ourselves and our experiences.... There is no logical impossibility in
> the supposition that the whole of life is a dream, in which we ourselves create all
> the objects that come before us. But although this is not logically impossible,
> there is no reason whatsoever to suppose that it is true; and it is, in fact, a less
> simple hypothesis, viewed as a means of accounting for the facts of our own life,
> than the common-sense hypothesis that there really are objects independent of us,
> whose action on us causes our sensations (1912: 22-23, emphasis in original).

---

[3] See, e.g., Williams 1991; DeRose 1995; Lewis 1996; Dretske 2003. Note too that responses to skepticism that assume I believe that the external world exists (such as Sosa's 2000 "safety" response) do not apply to the present case, since I have suspended belief.

Russell also states that certain experiences are "utterly inexplicable" from the solipsistic point of view and that the belief in objects independent of us "tends to simplify and systematize our account of our experiences". For these reasons, he says, the evidence of our experience speaks against solipsism, at least as a "working hypothesis" (1912: 23-24; 1914: 103-104). Russell aims lower than do Descartes and Kant; and partly as a result his goal seems more plausibly attainable. Yet Russell also promises something that the Hume, Wittgenstein, Moore, et al. do not: a non-question-begging positive argument against solipsism. It's a kind of middle path between certainty and surrender or refusal.

Unfortunately, there are two major shortcomings in Russell's argument. One is Russell's emphasis on simplicity. The most natural way to develop the external world hypothesis, it seems, involves committing to the real existence of billions of people, many more billions of artifacts, and naturally-occurring entities vastly more numerous even than that, manifesting in highly complex and unpredictable patterns. On the face of it, it's odd to say that such a picture of the world is simpler than radical solipsism.[4]

The second shortcoming is the uncompelling, gestural nature of Russell's supporting examples. *What* is it, exactly, that is "utterly inexplicable" for the solipsist? The utterly inexplicable thing, Russell says, is a cat's seeming hunger, after an interval during which the cat was not experienced:

> if [the cat] does not exist when I am not seeing it, it seems odd that appetite should grow during non-existence as fast as during existence. And if the cat consists only of sense-data, it cannot be *hungry,* since no hunger but my own can be a sense-datum to me. Thus the behaviour of the sense-data which represent the

---

[4] Though simplicity is a complex business. See Sober 1975; Vogel 1990; Zellner et al., eds., 2001; Peacocke 2004.

cat to me, though it seems quite natural when regarded as an expression of hunger, becomes utterly inexplicable when regarded as mere movements and changes of patches of colour, which are as incapable of hunger as a triangle is of playing football (1912: 23, emphasis in original).

To this example, Russell appends a second one: that when people seem to speak "it is very difficult to suppose that what we hear is not the expression of a thought" (24).

But are such experiences really *so* utterly inexplicable for the solipsist? Consider hallucinations or dreams, which arguably can involve apparent hungry cats and apparent human voices, without, behind them, real feline hunger or real human minds independent of my own. Russell considers this objection, but he seems satisfied to give only a cursory, single-sentence response: "But dreams are more or less suggested by what we call waking life and are capable of being more or less accounted for on scientific principles if we assume that there really is a physical world" (24). The inadequacy of this response is revealed by the fact that the solipsist might be able to say something quite similar: If I assume *solipsism* then I can account for the appearances of the cat and the interlocutor as imperfect projections of myself upon my imagined world, grounded in what I know about myself though introspection. Such an explanation is sketchy, to be sure, but so also is the current scientific explanation of dreaming and hallucination. At best, Russell's argument is problematically underdeveloped.

Although I am dissatisfied with Russell's particular argument, as I sit here (or seem to) with my solipsistic doubts, I still feel the attraction of that general approach. The core Russellian idea I want to preserve is this: Although *in principle* (contra Descartes and Kant) all the patterns in my experience are compatible with the non-existence of anything behind that experience, I still can evaluate two competing hypotheses about the origins of those experiences: the

solipsistic hypothesis according to which all there is in the universe are those experiences themselves, and the external world hypothesis which holds that there is something more. I can consider these hypotheses not by the standards of apodictic or transcendental proof, but rather as something like scientific hypotheses, with epistemic favor accruing to the one that appears to do the better job accounting for the experiential evidence at hand. Although Russell's own remarks are too cursory, that doesn't speak against the general project. Indeed, it invites the hope that a more patient effort might still bear fruit.

More recently, Laurence BonJour (1985, 2003), Jonathan Vogel (1990, 2005, 2008), Christopher Peacocke (2004), and others have argued against radical skepticism on explanationist grounds, thus pushing Russell's project a bit farther.[5] Vogel, for example, asserts that the skeptic will be forced between leaving her experiences unexplained and deploying ad-hoc explanations: "niceties aside, the fact that something is spherical explains why it behaves like a sphere (in its interactions with us and with other things). If something that is *not* a sphere behaves like one, this will call for a more extended explanation" (Vogel 1990: 663-664). This seems exactly right. But despite these authors' good start, their remarks remain scientifically sketchy – and the discussions tend to focus on deceiver skepticism rather than radical solipsism. What I really want is a rigorous, scientific experimental test in which I, from within the position of solipsistic doubt, give the solipsistic hypothesis what I judge to be its most plausible and appealing form, make specific predictions from it, see whether those specific predictions are borne out, and if the predictions are not borne out consider whether there are plausible post-hoc explanations. As far as I know, this sort of thing has never been properly done.

---

[5] For reviews, see Beebe 2009, McCain 2011. See also Reichenbach 1938/2006 and Sober 2011, discussed in Schwitzgebel 2013, and McCain 2012. Chalmers 2003/2010 and 2012 provides an interesting structural argument against treating a global deceiver scenario as a skeptical scenario, but he does not engage with radical solipsism.

Thus, in place of Russell's vaguely scientific appeal, I will try an *actual empirical test* of the two competing hypotheses, in formal scientific style. I will generate specific empirical predictions from radical solipsism and from non-skeptical external-world realism, pitting the two hypotheses against each other in what I hope is a non-question-begging way, and then I will quantitatively analyze the empirical results and consider how well the competing views can accommodate them.

In other words, I aim to do some solipsistic science. There is no contradiction in this. Skepticism about the external world is one thing, skepticism about induction and scientific reasoning quite another. I aim to see whether, from assumptions and procedures that even a radical solipsist can accept, I can generate experimental evidence for the existence of an external world – an exercise in what I might think of as the experimental philosophy of cosmology.

Let me emphasize: I don't hope to prove something from nothing. The skeptic's position is unassailable if her opponent must prove all the premises of any potential argument. I aim to refute not all of skepticism, but rather only radical solipsism. I aim only to move from solipsism-compatible premises to an anti-solipsistic conclusion. Accordingly, for purposes of this project, I don't plan to entertain any more than the usual doubt about (solipsism-compatible versions of) induction or deduction or short-term memory or introspective self-knowledge.

More specifically, I will allow myself to assume the following:

- introspective knowledge of sensory experience and of other happenings in the stream of experience;

- memories of past experience from the time of the beginning of the series of experiments but not before;

- concepts and categories arrived at I-know-not-how and shorn of any presupposition of grounding in a really-existing external world;

- the general tools of reason and scientific evaluation to the extent those rules don't build in any assumptions about affairs beyond my stream of experience.

Leaning only on these resources, I will try to establish, to a reasonable standard of scientific confidence, the existence of an external world.

I will not assume that I or anything else existed yesterday. Solipsistic doubt invites doubts about the past: From what I'm experiencing *now*, and going forward, can I establish an external world? From a radical skeptical break in which I cut away the past and all things external, can I build things back up? Furthermore, if my memory of days before today does in fact serve me correctly, then one of things it tells me is that it itself is highly selective and unsystematic, and that the free recall of what is salient and readily available is not usually good scientific method. Though I can pull together memories of events that seem best explained by the hypothesis that there is an external world, it is more scientifically rigorous to restart at square one.

If solipsism implies that I have complete control over my stream of experience, I can easily refute it. I could, for example, take in my hands a deck of cards (or at least seem to myself to do so) and resolutely *will* that I draw a queen of clubs. Then I might note the failure of the world to comply with my will. In fact, I have now attempted exactly this, with an apparent ten of diamonds as my result. But unfortunately for the prospects of such an easy proof, solipsism has no such voluntaristic implications and thus admits of no such anti-voluntaristic refutation.[6]

---

[6] Certain remarks of Locke (1689/1975, IV.xi), Berkeley (1710/1965, §29 ff.), and Fichte (1797/2000, §2) suggest a view of solipsism on which the solipsist has full voluntary control of

To think through this last issue more clearly, I close my eyes – or rather, not to assume the existence of my body, I do something that seems to me to be a closing of the eyes. What I visually experience is an unpredictable and uncontrollable array of colors against a dark gray background, the *Eigenlicht*.[7] This uncompliant *Eigenlicht* is entirely compatible with radical solipsism as long as I conceptualize the patterns it contains as nothing but patterns *in,* or randomnesses *in,* my stream of sensory experience, patterns governed by their own internal coherences rather than by anything further that stands behind them. The unpredictability and uncontrollability of these visual patterns no more compels me to accept the existence of non-experiential entities than irresolvable randomness and unexplained laws in the material world, as I conceptualize it, would compel me to accept the existence of immaterial entities behind the material ones.

What sorts of tests, then, might put radical solipsism at risk? Interpreting this question as straightforwardly as possible, I see three types of potential evidence that would be difficult to accommodate on a solipsistic view: evidence of the existence of something with theoretical reasoning powers that exceed my own, evidence of the existence of something that can retain its properties over a period during which those properties are lost to my sensory experience and memory, and evidence of something with practical reasoning powers that exceed my own.

I could ensure the impossibility of refuting solipsism by insulating it in advance against making any predictions that conflict with the predictions of the external world hypothesis. But although this might be an interesting form of solipsism to consider aprioristically, empirically-insulated solipsism is defective hypothesis by ordinary scientific standards. Either it avoids

all experience. But such *compliant-world solipsism* is a flimsy cousin of the more robust version of solipsism I have in view. See also Descartes 1641/1984, Meditations 3 and 6.

   [7] See Schwitzgebel 2011.

making predictions that conflict with the external world hypothesis by predicting nothing – and thus is unfalsifiable and less confirmed by observations compatible with it than are theories that make risky predictions – or it avoids the conflict by being jury-rigged post-hoc to precisely match the predictions of the competitor hypothesis without generating any new organic predictions of its own. Better, I think, to treat the solipsistic hypothesis as one that accommodates ordinary sensory patterns, like *Eigenlicht* patterns, but that generates riskily different predictions about how those patterns will evolve in cases of the sort soon to be described.

I have characterized solipsism as the view that only my stream of conscious experience exists. Another, wider solipsism might allow also the existence of an unconscious part of my mind but no world beyond that. I will consider only narrow solipsism. It is the more natural target, both historically[8] and epistemically. The epistemic asymmetry between consciousness and everything else is what is striking.

I will now describe three experiments, all conducted in one uninterrupted episode on a single day. To the extent possible, the remaining text, apart from the final concluding section, reflects real thoughts on the day of experimentation, with a few subsequent modifications for clarity. To fit all of these thoughts into the time-span of a single day, I drafted a version of the material below in the present tense using dummy results based on pilot experiments. I entered into the experiment with the intention of genuinely thinking the thoughts below with real data as the final results came in.

---

[8] E.g., Quine: "solipsism, according to which there is nobody else in the world, not indeed any world but the pageants of one's own sense data" (1966/1976: 250); Thornton: "Solipsism is therefore more properly regarded as the doctrine that existence is everything that I experience… and is necessarily construed by me as part of the content of *my* consciousness" (2004: introduction)

Experiment 1: The Prime Number Experiment.

Method. I have prepared for this experiment by programming Microsoft Excel to calculate whether a four-digit number is prime, displaying "prime" next to the number if it is prime and "nonprime" if it is not. Then I programmed Microsoft Excel to generate arbitrary numbers between 1000 and 4000, excluding numbers divisible by 2, 3, and 5. Or rather, I should say, without assuming the existence of a computer or Microsoft Excel, this is what I now seem to remember having done. Version A of this experiment will proceed in four stages, if all goes according to plan. First, I will generate a fresh set of 20 new qualifying four-digit numbers. Second, I will take my best guess at which of those 20 numbers are prime, allotting myself approximately two seconds for each guess. Third, I will paste this set of numbers into my seeming prime number calculator function, noting which ones are marked as prime by the seeming machine. Finally, by laborious manual calculation, I will determine which among those twenty numbers actually are prime. Version B will proceed the same way, except using Roman numerals as the initial basis for my guesses.

My hypothesis is this: If nothing exists in the world apart from my stream of conscious experience, then the swift, seemingly Excel-generated answers should not be statistically more accurate than my own best guesses. For if they were more accurate, that would suggest the existence of something more capable of swift, accurate prime-number detection than is my own solipsistically-conceived conscious self.

Results. I have just now conducted the experiment as described. Then main results are displayed in Figure 1. For Version A, my best guesses yielded an estimate of 11 primes. In most cases, this felt like simple hunchy guessing, though 3913 did pop out as nonprime. The apparent

Excel calculation also yielded an output of 11 seemingly machine-calculated primes. Manual calculation confirmed the seeming machine results in 19 out of the 20 cases.[9] In contrast, manual calculation confirmed my best-guess judgments in only 11 out of the 20 cases. The difference in accuracy between 19/20 and 11/20 is statistically significant by Fisher's exact test (manually calculated), with a two-tailed p value of < .02. For Version B, again both my best guesses and the apparent Excel outputs yielded 11 estimated primes, and again manual calculation confirmed the apparent Excel outputs in 19 of the 20 cases,[10] while manual calculation confirmed my best guesses in 13 of the 20 cases. The difference in accuracy between 19/30 and 13/20 is marginally significant by Fisher's exact test (manually calculated), with a two-tailed p value of approximately .05.[11]

---

[9] Subsequent examination of my notes suggests a long-division error on my part and thus that the apparent Excel output was correct after all.

[10] Subsequent examination of my notes reveals an error in my manual conversion from Roman to Arabic numerals.

[11] Subsequent analysis with a statistical program yields p = .044. For Version A, subsequent analysis yields p = .008.

Figure 1: Accuracy of prime number estimates, as judged by manual calculation, for my best guesses before calculating, compared to what would seem to be the output of an Excel spreadsheet programmed to detect primes. Error bars are manually-calculated 95% confidence intervals using the normal approximation and a ceiling of 100%.
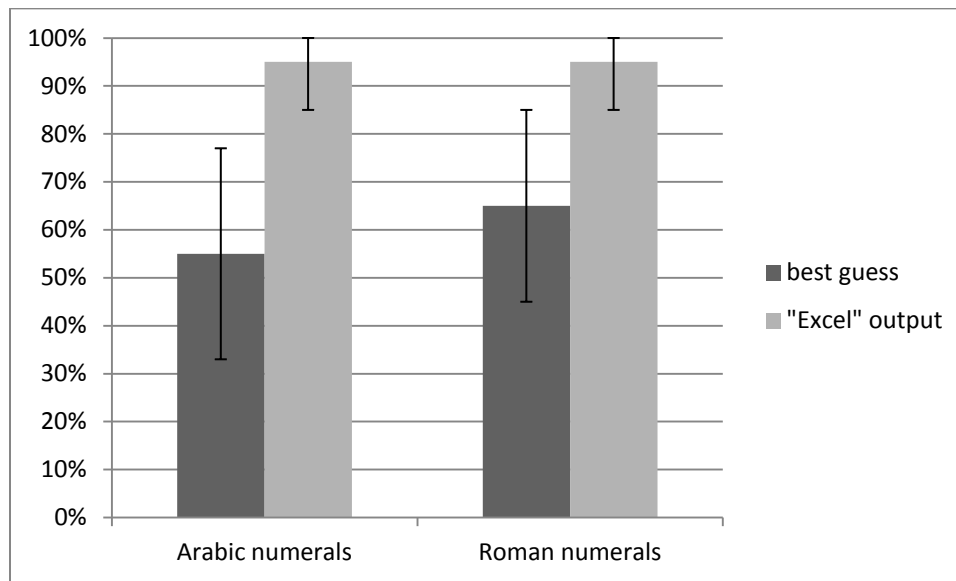
Discussion. I believe the most natural interpretation of these results is that something

exists in the external world that has calculation capacities exceeding my own. Although I was

able to hand-confirm most of the answers, I could not do so swiftly, nor recognize them as

correct when they arose. The best explanation seems to be that, in that instant that I seemed to

drag down the Excel function, something outside my solipsistically-conceived stream of

experience performed calculations of which I was not conscious.

As mentioned before, I am setting aside radically skeptical concerns about memory

within the span of the experiment (what if the world was created two seconds ago?) and

introspection[12] (what if I delusionally misjudged all my intentions and/or sensory experiences?).

My aim, as I have emphasized, is not to employ radically skeptical standards generally, but

rather to employ the normal standards of science insofar as they can be employed by someone

open-minded about radical solipsism. I also set aside concerns about whether this seeming

computer really *calculates* rather than only being designed to give outputs interpretable by users

as the results of calculation. Either way, the results suggest the existence of someone or

something with prime-number-detection capacities that exceed those present in my

solipsistically-conceived stream of experience. Even if that thing is only my own unconscious

mind, bent on tricking my conscious mind into misinterpreting my experimental results, radical

solipsism as I've defined it would still be false, since radical solipsism denies the existence of

anything outside of my stream of experience.

As I reflected earlier, I believe that solipsism can readily allow that the stream of

experience contains patterns within it, as long as those patterns are not caused by anything

---

[12] From within the perspective of this project I intend that my judgments be not perceptual judgments about the outside world but rather introspective judgments about my experiences of that seeming outside world. See Schwitzgebel 2012 for more discussion of the relationship between perception and introspection.

behind that experience. The anti-solipsistic interpretation of these results thus turns crucially on the question of whether the outcome of this experiment might plausibly be only a manifestation of such solipsistic patterns of experience. So, how plausible would such a pattern be, really, given solipsistic assumptions? What should I expect patterns of experience to look like if solipsism is true?

These are hard questions to answer. And yet I don't want to be *too* hard on myself. I'm looking only for scientific plausibility, not absolute certainty.

Examining my own experience, one typical kind of pattern is this: When I do something that feels like shifting my eyes to the left, my experienced visual field seems to shift to the right – a fairly simple law of experience, a simple way in which two experiences might be directly related with no compelling explanatory need of a non-experiential intermediary. Likewise, when I seem to see a spherical thing and then seem to reach out to touch it, I seem also to have tactile experience of something spherical. This more complex pattern is not fully expressible by me, but still it seems a fairly straightforward set of relationships among experiences. It's tempting to think that there must be a genuine mind-independent physical sphere that unifies and structures those cross-modal experiences. But if I am to be genuinely open-minded about solipsism, I must admit that the existence of a radically new ontological type is a heavy cost to pay to explain this relationship among my experiences. They might be related to each other directly by an admittedly somewhat complex set of intrinsically experiential laws (as in Mill 1867). Similarly when I close my eyes. There are regularities – the visual field changes radically in roughly the way I expect, though it also gains some highly unpredictable elements. Solipsism can also allow the existence of unrecognized patterns of relationship among my experiences. For example, afterimages of bright seeming-objects might be in perfect complementary colors even if I don't

realize that fact. There might be discernible but as-yet-undiscerned regularities governing

temporal evolutions in the flight of colors I experience when my eyes are closed. All of this

seems plausibly explicable by solipsistic laws that relate experiences directly to one to another.

The core question is: Could laws of broadly this sort suffice to explain my present results?

To explain the results of Experiment 1 solipsistically via such unmediated laws of

experience, something like the following would have to be true. There would have to be an

unmediated relationship between, on the one hand, (a.) the visual experience, for example, of the

numeral "2837" in apparent black Calibri 11 font in an apparent Excel spreadsheet, accompanied

by the inner speech experience of saying to myself, with understanding, "twenty-eight thirty-

seven" in English, and, on the other hand (b.) the visual experience of suddenly seeing "prime"

in the matched column to its right if the number is prime or "nonprime" if the number is

nonprime. At first blush such an unmediated relationship seems to me at least a little strange.

On radical solipsism, I'm inclined to think, it should be some feature of the visual experience

that drives the appearance of "prime" or "nonprime" on the seeming-screen; but what feature

could that be? The visual appearance of the numeral "1867" (prime) doesn't seem to share

anything particularly more in common with the visual appearance of "3023" (also prime) than

with "1387" (nonprime) – and still less with "MMCCCXXXIII". What seems to unify the

pattern of results is instead a semantic feature, primeness of the represented number, something

that is neither on the sensory surface nor, seemingly, present anywhere else in my conscious

mind at the beginning of the experiment; and such a property seems an unlikely candidate to

drive unmediated regularities in a solipsistic universe. How would it get traction? "Display

'prime' if and only if the squiggles happen to represent a prime number in some recognizable

numeral system" seems too complex or arbitrary to be a fundamental law, but also too

disconnected from the particular complexities on the surface of my experience at the moment to be driven directly by those surface complexities. Solipsism more naturally predicts, I think, what I did in fact predict on its behalf: that just as I might expect in a dream, the assortment of "prime" and "nonprime" should be unrelated to actual primeness except insofar as I have guessed correctly.

Given my acknowledged bias against solipsism, though, it would be imprudent for me to leap too swiftly to the conclusion that solipsism is false from this evidence alone. Maybe this is exactly the sort of law of experience that I should expect on a solipsistic worldview, charitably enough understood? Maybe solipsism can be developed to accommodate laws that directly relate unrecognized semantic properties of numeral experiences to semantic properties of English orthography experiences, or something. One awkward result does not a decisive refutation make. So I have some further experiments in mind.

Experiment 2: Two Memory Tests.

Method. I am currently having visual experience of an apparent person. I am inclined to think of this apparent person as my graduate student collaborator, Alan. I have arranged for this seeming "Alan" to test my memory. In the first test, he will orally present to me an arbitrary-seeming series of 20 three-digit numbers. He will present this list to me twice. I will first attempt to freely recall items from that list, and then I will attempt to recognize those items from among a 40-item list, half of which consists of new three-number combinations. The second test will be the same procedure with 20 sets of three letters each, and with a two-minute enforced delay to further impair my performance. In both cases, I expect that seeming-Alan will tell me that my memory has been less than perfect. He will then, if all goes according to plan, tell me

that he is visually re-presenting the original lists. If solipsism is true, nothing should exist to anchor the advertised visual "re-presentation" to the earlier orally-presented lists, apart from my own memory. In those cases, then, where my memory has failed, the supposed "re-presentation" should not contain a greater number of the originally presented elements than would be generated by chance. I shouldn't be able to step in the same stream twice except insofar as I can remember that stream (though maybe I could have the illusory feeling of having done so). The contents of experience should not have a fixity that exceeds that of my memory, because there is nothing beyond my own experience that can do the fixing. At least, this seems to me the most straightforward prediction from within the solipsistic worldview.

The final move in the experiment will be to test whether the re-presented list does indeed match the original list despite the gap in my memory. The method of test will be this: Seeming-Alan will state the procedure by which he generated the seemingly arbitrary lists. By (seeming) prior arrangement, he will have used a simple arithmetic procedure for the numbers and a simple semantic procedure for the letters (such as the decimal expansion of a four-digit fraction and a simple transformation of a familiar text). I should then be able to test whether the later-presented full lists of 20 three-element items are indeed consistent with generation by the claimed procedures. If so, this will suggest that the original lists were also generated by those same procedures. It will do so, if all goes well, because (as I will later estimate) there's only a very small chance that two arbitrary lists of 20 three-element items would have several items in common – the several items I will presumably remember across the temporal gap – unless they were generated by the same procedure. This would then allow me to infer that the entire "re-presented" list does indeed match the entire original list despite my failure to recall some items across the interval – contra the no-same-stream prediction of the solipsistic hypothesis.

Results. The main results are displayed in Figure 2. According to seeming-Alan, in the number test, I correctly recalled 7 of the 20 three-digit items (with no false recall) and I accurately recognized 14 of the items. In the letter test, I correctly recalled 8 of the 20 three-letter items (with no false recall) and accurately recognized 18. The generating patterns, he claims, were the decimal expansion of 1/2012, excluding the initial zeroes, and the most famous lines of Martin Luther King, Jr.'s, "I Have a Dream" speech, skipping every other letter and excluding one repeated item. In both cases I manually confirmed that the "re-presented" lists conformed to the purported generation procedure. Manual application of the two-tailed Fisher's exact test shows my recollections to match significantly less well to the "re-presented" lists than do the manually confirmed results (both p's < .001). At a p < .05 confidence level, the recognition results are statistically significant for the three-digit items but not for the three-letter items.

Figure 2: Number correct out of 20 as judged by comparison with the lists "re-presented" by seeming-Alan, for my recall guesses, my recognition-test guesses, and my manual confirmation of the purported generating procedure. Error bars for "recalled" and "recognized" are manually calculated 95% confidence intervals using the normal approximation. Error bars for the ceiling results use the "rule of three".
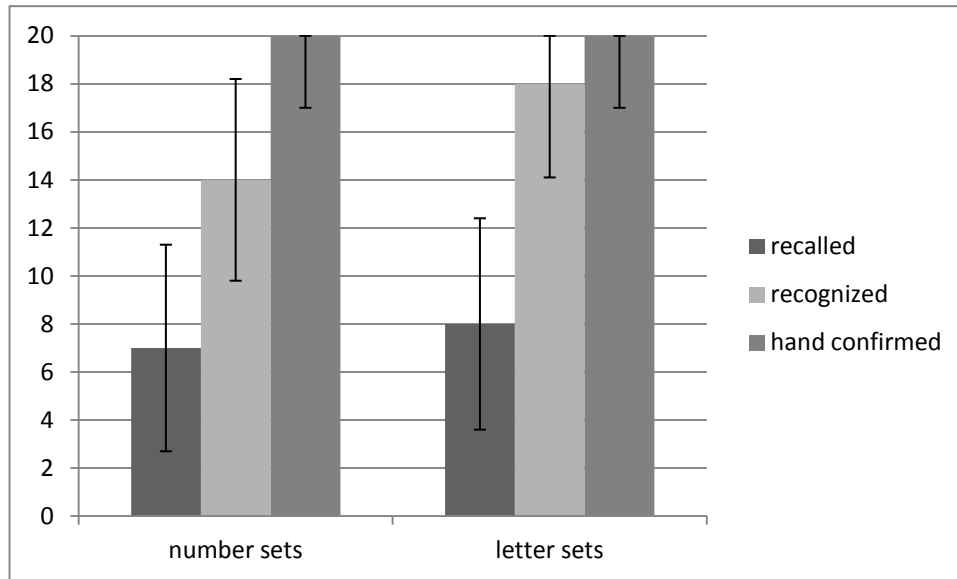
I found myself trying hard in both memory tasks. Since I am inclined to believe that the external world does exist and thus that some other people might read what I now appear to be writing, I was motivated not to come across as sieve-brained. This created a substantial incentive to answer correctly, which I did not feel as strongly in Experiment 1.

Discussion. I believe the most natural interpretation of these results is that something existed in the external world that retained the originally presented information across my gap of memory.

Alternative interpretations are possible as always. The question is whether I should find those alternative interpretations scientifically attractive, considering the truth or falsity of solipsism as an open question. Can solipsism accommodate the results naturally, without the strained, ad-hoc excuses that are the tell-tale sign of a theory in trouble?

Might there have been a direct, unmediated connection between the original auditory experience of "IAE", "DEM", etc., and the later visual experience of those same letter arrangements? If I am willing to countenance backward causation, then a neat explanation offers itself: I might have solipsistically concocted the generating patterns at the moment seeming-Alan seemed to be informing me of them, and then these generating patterns might have backwardly caused my initial auditory experiences and guesses. However, temporally backward causation seems a desperate move if I'm aiming to apply normal scientific standards as I conceptualize them. A somewhat less radical possibility is temporally gappy cross-modal causation from auditory experience at the beginning of the experiment to visual "re-presentation" later; but this requires, in addition to the somewhat uncomfortable provision of temporally gappy causation, a further seeming implausibility absent from the backward causation case – the seeming implausibility, similar to that in Experiment 1, of an unmediated law of experience that operates

upon semantic contents of those experiences that are unrecognized as such by me at the time of the law's operation. In this case the relevant semantic contents would be nothing as elegant as primeness but rather the decimal expansion of one divided by the current calendar year, excluding the initial zeroes, and the English orthography of the words of a familiar-seeming speech, skipping alternate letters and excluding one repeated triplet.

The thought occurs to me that some of the laws of external-world psychology, as I conceive of it, are also weird and semantical. For example, an advertisement might trigger a tangentially associated memory. But the crucial difference is this: In the case of external-world psychology, the semantic associations, even if not conscious, I assume to be grounded in mundane facts about neural firing patterns and the like. A bare solipsistic tendency to create and then recreate, unbeknownst to myself, the same partial orthography of a familiar speech, while meanwhile being unable to produce that partial orthography when I consciously try to do so – well, that's not impossible perhaps, but neither does it seem as natural a development of solipsism as does the view that the stability of experience should not exceed the stability of memory.

My argument would be defeated if I could have easily found some simple scheme, post-hoc, that could generate twenty items including exactly those seven recalled numbers and eight recalled letter sets. My anti-solipsistic interpretation requires that there be only one plausible generating scheme for each set; otherwise there is no reason to think the unrecalled items would be the same on the initially presented list as on the subsequently presented list. So, then, what

are the odds of a post-hoc fit of seven or more items from each set? Fortunately, very low –

about one in a million, given some plausible assumptions and the mathematics of combination.[13]

Perhaps, then, the best defense for solipsism, if it's not to collapse into a general radical

skepticism about short-term memory or induction or arithmetic, is temporally gappy cross-modal

forward causation grounded in unrecognized weird semantic features of the relevant experiences.

I'm inclined to think this is a somewhat awkward position for the solipsistic hypothesis. But

maybe I'm still being too unsympathetic to solipsism? Maybe I should have expected that

scientific laws would look somewhat weird in a solipsistic world, and rather unlike the scientific

laws I think of as characteristic of the natural sciences and naturalistic psychology? So I have

planned for myself one final experiment of a rather different sort.

Experiment 3: Defeat at Chess.

Method. Seeming-Alan tells me that he is good at chess. I believe that I stink at chess.

Thus, I have arranged to play 20 games of speed chess against seeming-Alan, with a limit of

approximately five seconds per move. If solipsism is true, nothing in the universe should exist

that has chess-playing practical reasoning capacities that exceed my own, and so I should not

experience defeat at rates above statistical chance when directing all of my conscious efforts on

---

[13] Consider the odds of hitting one specific familiar sentence or one specific number-generating sequence. Given 1000 possible three-digit numbers, equally likely, or 1000 equally likely letter sets, the odds of exactly seven matches among the 20 items generated by that specific sentence or number-generating scheme are approximately $(1/10^3)^7$ times 20-choose-7 possible arrangements (about 10^6), i.e., approximately one in 10^{15}. The odds of 8 or more matches add only negligibly to this probability. Even if we assume a billion possible simple generating schemes along roughly the lines I recall having suggested to seeming-Alan, the odds of a chance match of at least seven items out of 20 to any one of those billion generating schemes are about one in a million (i.e., $10^9/10^{15} = 1/10^6$).

behalf of one color. Figure 3 displays the procedure, as presented to me by a seeming camera

held by a seeming Gerardo Sanchez.

Figure 3: The procedure of Experiment 3.

Results. Seeming-Alan defeated me in 17 games of 20, with one stalemate. 17/19 is statistically higher than 50% with a p-value < .001 (manually calculated).

Discussion. It occurs to me that I might have hoped to lose, so as to generate results confirming my preferred hypothesis that the external world exists. Against this concern, I reassure myself with the following thoughts. If it was an *unconscious* desire to lose, then that seems to imply that something exists in the world besides my stream of conscious experience, namely, an unconscious part of my mind, and thus radical solipsism as I have defined it is false. If it was, instead, a *conscious* desire to lose, I should have been able to detect it, barring an unusual degree of introspective skepticism. What I detected instead was a desire to win as many games as I could manage, given my background assumption that if Alan actually exists I would be hard pressed to win even one or two games. Playing with what I experienced as determination and energy, I found myself forcefully and repeatedly struck by the impression that the universe contained a practical intelligence superior to my own and bent on defeating me. The most natural scientific explanation of the pattern in my experience is that that impression was correct.

Does it matter that, if the external world exists in something like the form I think it does, some chess-playing computer programs could have defeated me as handily as seeming-Alan did? I don't see why it should. Whether the strategy and intentionality is manifested directly by a human being or instead through the medium of a human-programmed chess-playing computer – or in any other way – as long as that strategy and intentionality exceeds the abilities of my own conscious mind, the solipsistic hypothesis faces a substantial explanatory challenge. It can try to address this challenge by appealing to intrinsic relationships among experiences of mine – relationships among seeming chess moves countered by other seeming chess moves whose

power I only recognize in retrospect – but the more elegant unifying explanation of the results would seem to be the existence of a hostile goal-directed intelligence.

Could I then maybe just abandon the pursuit of explanation? Could I just say it's a regularity unexplained, end of story? But why settle myself so quickly into defeat, when the existence of a hostile intelligence seems so readily available as an alternative explanation? Simply shrugging, when I am not forced to do so, runs contrary to the exploratory, scientific spirit of this exercise.

I might easily enough *dream* of being consistently defeated in chess. Maybe some dreamlike concoction of a seeming chess master could equally well explain my pattern of experience, without need of an external world? But dreams of this sort, as I seem to remember, differ from the present experiment in one crucial way: They are vague on the specific moves or unrealistic about those moves. In the same way, I might dream of proving Fermat's last theorem. Such cases involve dream-like gappiness or irrationality or delusive credulity – the type of gappiness or irrationality or delusive credulity that might make me see nothing remarkable in discovering I am my own father or in discovering a new whole number between 5 and 6. Genuine dream doubt might involve doubt about my basic rational capacities, but if so, such doubts outrun simply solipsistic doubts. And even if I am dreaming or in some other way concocting for myself an imaginary chess master, radical solipsism in the intended sense would still be defeated, if the dream explanation implies that there is some unconscious part of my mind that excels at chess and on which I draw, unwittingly, to select those startlingly clever moves. I can easily imagine a world in which I am regularly defeated at chess; but leaning on the resources only of my own chess-inept conscious mind, I can no more specifically imagine the truly brilliant play of a chess expert than I can specifically imagine a world in which twenty four-

digit numbers are all, in an instant, properly coded as "prime" or "non-prime". If I consistently experience genuinely clever and perceptive chess moves that repeatedly exploit flaws in my own conscious strategizing, moves that I experience as surprising but which I retrospectively appreciate, it seems hard to avoid the conclusion that something exists that exceeds my own conscious intelligence in at least this one area.

General Discussion.

When I examine my stream of experience casually, nothing in it seems to compel the rejection of solipsism. My experience contains seeming representations of outward objects; it follows patterns unknown to me and that resist my will. But those basic facts of experience are readily compatible with the truth of radical solipsism. Once I find myself with solipsistic doubts, "here is a hand" does not help me recover; but neither do ambitious proofs in the spirit of Descartes and Kant seem to succeed. I could try to reconcile myself to the impossibility of proof, but that feels like giving up.

Fortunately, the external world hypothesis and the solipsistic hypothesis do appear to make different empirical predictions under certain conditions, at least when interpreted straightforwardly. The external world hypothesis predicts that I will see evidence of theoretical reasoning capacities, memorial retention, and practical reasoning capacities exceeding those of my narrowly-conceived conscious self, while solipsism appears to predict the contrary. I can then scientifically test these predictions, avoiding begging the question by using only tools that are available to me from within a solipsistic perspective.

The results come out badly for solipsism. To escape my seemingly anti-solipsistic results requires either adopting other forms of radical skepticism in addition to solipsism (for example

about memory, even in the short duration of these studies) or adopting what strike me as increasingly ad hoc, strained, and convoluted accounts of the nature of the laws or regularities connecting one experience to the next.

Did I really need to do science to arrive at this conclusion, though? Maybe instead of running formal experiments could I have simply consulted long-term memory for evidence of my frustration by superior intelligences and the like? Surely so! And thus maybe also even before conducting these exercises I implicitly relied upon such evidence informally to support my knowledge that the external world exists. Indeed, it would be nice to grant this point, since then I can rightly say that I have known for a long time that the external world exists. But still, the present procedure has several advantages over attempts to remember past frustrations and failures. For one thing, it achieves its goal despite conceding more to the skeptic from the outset, for example, unbelief in yesterday. For another, it more rigorously excludes chance and confirmation bias in evidence selection. And for still another, it forces me starkly and explicitly to consider the best possible alternative solipsistic explanations I can devise to account for specific, concrete pieces of evidence – giving solipsism a chance, I hope a fair chance, to strut its stuff, if stuff it has.

Perhaps it's worth noting that the best experiments I could concoct all involved pitting my intelligence against another intelligence, or against a device created by another intelligence – a device or intelligence capable of generating semantic or strategic patterns that I could only subsequently appreciate. Whether this is an accidental feature of the experiments I happened to design or whether it reflects some deeper truth, I am unsure.[14]

---

[14] Maybe Fichte (1797/2000) and Hegel (1807/1977) wouldn't be surprised by this feature of my findings. See discussion in Beiser 2005 and Stern 2012. However, I share Stern's concerns about the force of these responses to solipsism as historically developed.

I conclude that the external world exists. To be clear, I don't at this point conclude anything about its metaphysical character. It is consistent with my experimental results that the external world be material or divine or an unconscious part of myself or an evil demon or an ectoplasmic computer, as long as it can host intelligence. I'll go out on a limb, though, and tentatively conclude that Alan exists,[15] a genuine independent-minded co-author, and I will do something that seems to me to be the act of circulating this essay for him and others to read.[16]

References

Beebe, James (2009) 'The abductivist reply to skepticism'. *Philosophy & Phenomenological Research,* 79, 605-636.

Beiser, Frederick (2005) *Hegel.* New York: Routledge.

Berkeley, George (1710/1965) *A treatise concerning the principles of human knowledge.* In C.M. Turbayne (ed.), *Principles, dialogues, and philosophical correspondence.* New York: Macmillan.

BonJour, Laurence (1985) *The structure of empirical knowledge.* Cambridge, MA: Harvard.

BonJour, Laurence (2003) 'A version of internalist foundationalism'. In L. BonJour and E. Sosa, (eds.), *Epistemic justification.* Malden, MA: Blackwell.

Broughton, Janet (2002) *Descartes's method of doubt.* Princeton: Princeton.

Burge, Tyler (1999/2013) 'A warrant for belief in other minds'. In *Cognition Through Understanding*. Oxford: Oxford.

Chalmers, David J. (2003/2010) 'The Matrix as metaphysics'. In *The character of consciousness.* Oxford: Oxford.

Chalmers, David J. (2012) *Constructing the world.* Oxford: Oxford.

DeRose, Keith (1995) 'Solving the skeptical problem'. *Philosophical Review,* 104, 1-52.

Descartes, René (1641/1984) *Meditations on first philosophy.* In J. Cottingham, R. Stoothoff, and D. Murdoch (eds.), *The philosophical writings of Descartes, vol. II.* Cambridge: Cambridge.

Dicker, Georges (2008) 'Kant's refutation of idealism'. *Noûs,* 42, 80-108.

Dretske, Fred (2003) 'Skepticism: What perception teaches'. In S. Luper (ed.,) *The skeptics.* Aldershot, England: Ashgate.

Fichte, Johann G. (1797/2000) *Foundations of natural right.* Ed. F. Neuhouser. Trans. M.

    Bauer. Cambridge: Cambridge.

Guyer, Paul (1987) *Kant and the claims of knowledge.* Cambridge: Cambridge.

Hill, Christopher S. (1991) *Sensations.* Cambridge: Cambridge.

Hegel, G.W.F. (1807/1977) *Phenomenology of spirit.* Trans. A.V. Miller. Oxford: Oxford.

Hume, David (1740/1978) *A treatise of human nature,* ed. L.A. Selby-Bigge and P.H. Nidditch.

    Oxford: Oxford.

Kant, Immanuel (1781/1787/1929) *Critique of pure reason.* Trans. N.K. Smith. New York: St

    Martin's.

Lem, Stanislaw (1961/1970) *Solaris,* trans. J. Kilmartin and S. Cox. San Diego, CA: Harcourt.

Lewis, David (1996) 'Elusive knowledge'. *Australasian Journal of Philosophy,* 74, 549-567.

Locke, John (1689/1975) *Essay concerning human understanding,* ed. P.H. Nidditch. Oxford:

    Oxford.

McCain, Kevin (2011) *Inference to the best explanation and the external world.* PhD.

    Dissertation, University of Rochester.

McCain, Kevin (2012) 'The predictivist argument against skepticism'. *Analysis,* 72, 660-665.

Mill, John Stuart (1867) *An examination of Sir William Hamilton's philosophy.* London:

    Longmans, Green, Reader, and Dyer.

Moore, G.E. (1939) 'Proof of an external world'. *Proceedings of the British Academy,* 25, 273-

    300.

Nolan, Lawrence, and Alan Nelson (2006) 'Proofs for the existence of God'. In Stephen

    Gaukroger, (ed.), *The Blackwell Guide to Descartes' Meditations.* Oxford: Blackwell.

Pargetter, Robert (1984) 'The scientific inference to other minds'. *Australasian Journal of Philosophy,* 62, 158-163.

Peacocke, Christopher (2004) *The realm of reason.* Oxford: Oxford.

Price, H.H. (1938) 'Our evidence for the existence of other minds'. *Philosophy,* 13, 425-456.

Quine, W.V.O. (1966/1976). *Ways of paradox.* Cambridge, MA: Harvard.

Reichenbach, Hans (1938/2006) *Experience and prediction.* Notre Dame, IN: University of Notre Dame.

Russell, Bertrand (1912) *The problems of philosophy.* Oxford: Oxford.

Russell, Bertrand (1914) *Our knowledge of the external world.* London: George Allen & Unwin.

Schwitzgebel, Eric (2011) *Perplexities of consciousness.* Cambridge, MA: MIT.

Schwitzgebel, Eric (2012) 'Introspection, what?'. In D. Smithies and D. Stoljar (eds.), *Introspection and consciousness.* Oxford: Oxford

Schwitzgebel, Eric (2013) 'Hans Reichenbach's cubical world and Elliott Sober's beach', blog post at *The Splintered Mind,* March 19, 2013: http://schwitzsplinters.blogspot.com/2013/03/hans-reichenbachs-cubical-world-and.html (accessed Oct. 21, 2014).

Schwitzgebel, Eric (2014) 'Stanislaw Lem's proof that the external world exists', blog post at *The Splintered Mind,* January 21, 2014: http://schwitzsplinters.blogspot.com/2014/01/stanislaw-lems-proof-that-external.html (accessed Oct. 21, 2014).

Sober, Elliott (1975) *Simplicity.* Oxford: Oxford.

Sober, Elliott (2011) 'Reichenbach's cubical universe and the problem of the external world'. *Synthese,* 181, 3-21.

Sosa, Ernest (2000) 'Skepticism and contextualism'. *Philosophical Issues, 10,* 1-18.

Stern, Robert (2012) 'Is Hegel's master-slave dialectic a refutation of solipsism?' *British Journal for the History of Philosophy, 20,* 333-361.

Strawson, Peter (1959) *Individuals.* London: Methuen.

Stroud, Barry (1984) *The significance of philosophical scepticism.* Oxford: Oxford.

Stroud, Barry (1994) 'Kantian arguments, conceptual capacities and invulnerability'. In P. Parrini (ed.), *Kant and contemporary epistemology.* Dordrecht: Kluwer.

Thornton, Stephen P. (2004). Solipsism and the problem of other minds. *Internet Encyclopedia of Philosophy:* www.iep.utm.edu/solipsis (accessed August 25, 2014).

Vogel, Jonathan (1990) 'Cartesian skepticism and inference to the best explanation'. *Journal of Philosophy,* 87, 658-666.

Vogel, Jonathan (2005) 'The refutation of skepticism'. In M. Steup and E. Sosa (eds.), *Contemporary debates in epistemology.* Malden, MA: Blackwell.

Vogel, Jonathan (2008) 'Internalist responses to skepticism'. In J. Greco (ed.), *The Oxford handbook of skepticism.* Oxford: Oxford.

Williams, Michael (1991) *Unnatural doubts.* Cambridge, MA: Blackwell.

Wittgenstein, Ludwig (1950-1951/1974) *On certainty.* Ed. G.E.M. Anscombe and G.H. von Wright. Malden, MA: Blackwell.

Zellner, Arnold, Hugo A. Zeuzenkamp, and Michael McAleer, eds. (2001) *Simplicity, inference, and modeling.* Cambridge: Cambridge.