The Bootstrapping Problem

Jonathan Weisberg
University of Toronto

Abstract. Bootstrapping is a suspicious form of reasoning that verifies a source's reliability by checking it against itself. Theories that endorse such reasoning face the bootstrapping problem. This article considers which theories face the problem, and surveys potential solutions. The initial focus is on theories like reliabilism and dogmatism, which allow one to gain knowledge from a source without knowing that it is reliable. But the discussion quickly turns to a more general version of the problem that does not depend on this allowance. Five potential solutions to the general problem are evaluated, and some implications for the literature on peer disagreement are considered.

Suppose a witness to a crime reports that the perpetrator's shirt was green, and we believe her. It would be illegitimate to then corroborate her credibility as follows: "the offender's shirt was green, just as the witness said. Since she spoke the truth, we have evidence that she is a reliable witness." After all, our only basis for believing that the offender's shirt was green is that the witness said so. If we verify her reliability by relying on her testimony in this way, we pull ourselves up by our bootstraps. *The bootstrapping problem* arises when an epistemological theory licenses such reasoning, saying it is good when it seems clearly bad.

Bootstrapping started out as a problem for reliabilism, but it's been spreading. Versions of the objection have now been brought against a range of foundationalist theories, and even against various views on peer disagreement. It has even been argued that bootstrapping is a paradox, a problem for everyone that arises from widely held and intuitively alluring assumptions. So, who actually faces the bootstrapping problem? What solutions have been proposed and what are their virtues and vices? This article surveys the literature with an eye to answering these questions.

We begin in §1 by considering how the original bootstrapping problem, posed as an objection to reliabilism, generalizes to other theories. In §2 we turn to proposed diagnoses of the generalized problem, examining various attempts to isolate what

I am grateful to Igor Douven, Adam Elga, Jennifer Nagel, Sherrilyn Roush, Mike Titelbaum, Brian Weatherson, and an anonymous referee for their generous help.

is wrong with bootstrappish reasoning. Finally, in §3 we consider a bootstrapping objection from the literature on peer disagreement.

1. SCOPE OF THE PROBLEM

The hubbub began when Fumerton (1995) and Vogel (2000) raised bootstrapping as an objection to *reliabilism*, the view that knowledge is reliably formed true belief.[1] Vogel illustrates the problem with a case like the following:

> *Abigail.* The gas gauge in Abigail's car is reliable, though she has no evidence about its reliability. On one occasion the gauge reads 'full', leading her to believe that the tank is full, which it is. She also notes that the gauge reads 'full', and concludes by deduction that the gauge read correctly on this occasion. She then repeats this procedure many times on other occasions, eventually coming to believe that the gauge is reliable, since it read correctly on each occasion.[2]

Abigail relies solely on reliable methods: perception, trust in her gauge, deduction, and induction are all reliable processes. So if reliabilism is true, each of her beliefs constitutes knowledge, including her conclusion that the gauge is reliable. But this seems absurd: Abigail has bootstrapped her way into knowledge of the gauge's reliability.

Cohen (2002) generalizes the objection to all *basic knowledge theories*, i.e. theories that endorse:

Basic Knowledge: Some sources of knowledge do not require that the subject know the source is reliable in order to gain knowledge from it.

Basic Knowledge is motivated by the regress problem. If one must always know that one's source is reliable in order to gain knowledge by it, an infinite regress threatens. To know that one has hands, one must establish the reliability of perception, presumably via some other source, but then that source's reliability must also be established, and so on. Basic Knowledge promises to stop the regress. If one can gain knowledge from (say) perception without knowing that perception is reliable, one can know that one has hands without facing infinite regress.

Reliabilism is one example of a basic knowledge theory. According to reliabilism, Abigail's gauge can provide her with knowledge about the level in the tank without

[1] Reliabilism comes in many forms. This is a very simple characterization of *process reliabilism*, sweeping many details under the rug. See (Vogel, 2000) for more detail on reliabilism in the context of bootstrapping; for a thorough survey and history, see (Goldman, 2008).

[2] Subjects will be given names in alphabetical order. Abigail corresponds to Vogel's "Roxanne".

her knowing that the gauge is reliable, because the gauge is in fact reliable. Another basic knowledge theory of recent interest is *dogmatism*, which says that perception can grant knowledge even absent knowledge that perception is reliable. If it appears to you that P, you can[3] know that P even if you don't know[4] that perception is reliable. For example, if a red table looks red to you under ordinary circumstances, and you believe it is red based on its appearance, then you know it is red, even if you don't know that your vision is reliable. (Pollock, 1974; Pryor, 2000)

How does dogmatism run afoul of bootstrapping? The following case is adapted from Cohen (2002):

> *Bernard.* Bernard is shopping for a table. His vision is reliable, though he has no evidence bearing on its reliability.[5] He sees a red table and forms the belief that it is red. He also notes that the table appears red, and concludes by deduction that it is as it appears. He repeats this procedure many times with other tables and colours. Eventually, he concludes by induction that his colour vision is reliable.

According to dogmatism, Bernard knows the first table is red, the second is blue, and so on. In each case he believes P because it appears to him that P, and dogmatism says he knows P in such cases (provided his belief is true, there are no defeaters, etc.). Since introspection and deduction presumably grant knowledge too, Bernard also knows how each table appears to him, and thus that each table is as it appears. Finally, since induction presumably grants knowledge as well, Bernard knows that his vision is reliable.

Abigail and Bernard have something striking in common: each trusts a source s/he does not know to be reliable, Abigail her gauge, Bernard his vision. If we count the beliefs they base on these sources as knowledge, bootstrapping looks unstoppable. All the subject must do is note what the source says on each occasion, deduce that it was correct on each occasion, and conclude by induction that it is generally reliable. Each of the later steps looks innocent, so the first step looks to be the culprit.

But in (Weisberg, 2010) I argue that bootstrapping remains a problem even if we reject Basic Knowledge. Let's call a source *super-reliable* if it is even more reliable than it needs to be in order to be a source of knowledge. Now consider the following variation on Abigail's case:

---

[3]Why only "can" know? Dogmatism says that perception defeasibly justifies, but more is required in order to know. Your belief must be true and there must be no defeaters. There may be other requirements too, like degettierization, tracking, or safety.

[4]On some formulations, "even if you don't possess justification for believing." (Pryor, 2000, 519)

[5]Also, circumstances are such that prima facie justification is enough for a true belief to count as knowledge (see fn. 3): there are no gettierizing conditions present, his belief tracks the truth, etc.

> *Charlie.* "Charlie knows that the gauge in his car is reliable, and it is in fact super-reliable. On one occasion the gauge reads 'full', leading him to believe that the tank is full, which it is. He notes that on this occasion the tank reads 'full' and is full. He then repeats this procedure many times on other occasions, coming to believe that the gauge is not only reliable, but super-reliable." (2010, 528)

Unlike Abigail, Charlie starts out knowing that his gauge is reliable. So we don't need to rely on Basic Knowledge to get a problematic result here, since common sense says that Charlie knows his tank is full. Yet bootstrapping is still a threat if we just change the conclusion the subject draws. Instead of concluding that his gauge is reliable (which he already knows), Charlie concludes that his gauge is *super*-reliable.[6]

2. Diagnoses of the Bootstrapping Problem

The bootstrapping problem emerges from the preceding discussion a paradox, rather than an objection to this or that epistemological view. Rejecting controversial theses like reliabilism, dogmatism, and Basic Knowledge does not eliminate the problem; seemingly common-sense and innocuous assumptions lead to the absurd conclusion that Charlie knows his gauge to be super-reliable. How are we to resolve the paradox? Here we'll survey five proposals.

2.1. **No Rule Circularity (Vogel).**

Vogel (2008) suggests that our bootstrapping agents are guilty of rule-circular reasoning. That is, they violate the following principle:

> **No Rule Circularity:** A belief that an epistemic rule *R* is reliable cannot be justified by the application of *R*. That is, neither the conclusion itself nor any belief which supports the conclusion may be justified in virtue of the application of *R*.

Abigail, for example, applies the rule *Trust the gauge* in the reasoning she uses to conclude that her gauge is reliable.

While No Rule Circularity has strong prima facie appeal, White points out that there are cases where rule-circular reasoning is good: "Doing well in a memory game can suggest that I have a good memory, even though I can't help but use my memory

[6] *Objection*. To support the strong hypothesis of super-reliability, Charlie must know that the gauge read correctly in many, many instances. But given that he only knows that the gauge is reliable, he cannot know this—the conjunction of so many instances is too improbable to be justified. *Reply*. Even a single reading of the gauge generates a bootstrapping problem. If Charlie reads and trusts the gauge only once, he can conclude that the gauge read correctly on this occasion, which is (slight) evidence for super-reliability. (Weisberg, 2010, 530–2)

to evaluate my performance." (2006, 530) For this diagnosis to work, we need some way of distinguishing between good and bad rule-circular reasoning.[7]

2.2. **Tracking (Roush).** Roush (2005) suggests that our bootstrapping agents fail to track the truth. According to Nozick (1981), for a subject's belief in *P* to count as knowledge it must be that (roughly) she would not have believed *P* had it been false. Because of the way Abigail is disposed to reason, she would believe that her gauge is reliable even if it were not. She takes the gauge at its word, so she would think it was establishing a long and perfect track-record even if it were not reliable. So the reason Abigail does not know her gauge is reliable is that she does not track the truth.

One challenge for the truth-tracking diagnosis is to explain not only why Abigail fails to know that her gauge is reliable, but also why she is irrational or unjustified in her belief that it does. Abigail's failure appears to be internal, at least in part. The tracking theorist might bring Abigail's external failure to track the truth indoors by saying one should not believe something when one knows the belief will not constitute knowledge (Williamson, 2000, 255–6). This is an intuitively appealing proposal, though McGlynn (forthcoming) argues that it is subject to counterexample: it is reasonable to believe a lottery ticket will lose even if one realizes that belief does not constitute knowledge (or track the truth).

A more pressing concern is that the tracking requirement faces well-known counterexamples, and handling those counterexamples may undermine the tracking theorist's ability to handle bootstrapping. Consider a counterexample due to Sosa (1999): I throw a bag of trash down the chute in my apartment building and assume it reaches the bottom. In fact it does, but I would have no idea if it hadn't; it would have gotten stuck on the way down and I would have noticed no difference. My belief that the bag made it to the bottom does not track the truth but it still constitutes knowledge, especially if bags getting stuck is a highly unusual occurrence (suppose it's never actually happened). Roush handles examples like this one by saying that I can deduce that the bag probably made it to the bottom from something I do track, namely that bags usually make it to the bottom (if they didn't, I'd have heard). From my knowledge that bags usually make it to the bottom, and that I've deposited my bag, I can deduce that it probably made it to the bottom.[8] So I know the bag probably

[7] I would add that No Rule Circularity is also too weak: some cases of bootstrapping do not exhibit rule-circularity (Weisberg, 2010, 530).

[8] Actually, there must be more to the deduction than this. Deductive reasoning is monotonic while this inference is defeasible. Were my neighbour to inform me that her bag got stuck in the chute just previously, it would no longer be probable that my bag made it all the way down.

made it to the bottom.[9] This solution requires that we make an exception to the tracking requirement, allowing that beliefs deduced from other knowledge count as knowledge themselves even when they don't track the truth. But Roush argues extensively that allowing this exception yields a superior account of knowledge.

The problem is that allowing this exception hampers tracking theory's ability to handle bootstrapping. Just as I can deduce that the bag probably made it to the bottom, Abigail can deduce that her gauge is probably reliable. Presumably she knows that gauges with long and perfect track-records are usually reliable; we can stipulate that she tracks this fact. And, according to the tracking theorist, she knows that the gauge has a long, perfect track-record; perception and the gauge are reliable by hypothesis, so her beliefs about what the gauge says and how the tank is track the truth. So, though she does not track the truth about the gauge's reliability, she can deduce its probable reliability from things she does track, which seems bad enough. The challenge for the tracking theorist is to find a way to say that I know my trash (probably) made it to the bottom of the chute without saying that Abigail knows her gauge is (probably) reliable.[10]

2.3. **Restrict Closure (Cohen).** Cohen (2002) suggests that bootstrapping agents violate a restriction on combining different kinds of knowledge in reasoning. His proposal begins with Sosa's (1991; 1997) distinction between animal knowledge and reflective knowledge. Animal knowledge merely requires that one's belief track reality, while reflective knowledge further requires "awareness of how one knows, in a way that precludes the unreliability of one's faculties." (Sosa, 1997, 427) Cohen's idea is that our bootstrapping agents illegitimately combine the two kinds of knowledge in their reasoning. Bernard, for example, combines his animal knowledge that the table is red with his reflective knowledge that it appears red to him[11] to deduce that his vision was accurate on this occasion. Cohen's proposal is that such combinations are not kosher. Combining animal and reflective knowledge in reasoning does not

[9] That's not the same as knowing that it did make it to the bottom, but Roush (2005, 66) argues that it's close enough. Even granting that it is though, there's a worry that tracking theory is collapsing into a form of internalism. Sensitivity to the truth has been replaced by sensitivity to its probability, which may just amount to being responsive to one's evidence.

[10] In a new paper, Roush (2012) takes a different view of deduction's knowledge-generating powers, but a proper assessment of this new view would take us too far afield.

[11] There is a worry here: this knowledge may not be reflective in the relevant sense. Bernard's knowledge that the table appears red to him may be reflective in the sense that he is reflecting on his own state of mind. But this knowledge isn't obviously reflective in the sense of being accompanied by an "awareness of how one knows, in a way that precludes the unreliability of one's faculties." If introspection tracks mental states in the way vision tracks visible properties, Bernard's knowledge of how things appear to him may just be animal knowledge.

produce knowledge. So Bernard does not know that his vision worked correctly on a given occasion, and hence does not know that his colour vision is reliable in general.

Why think that animal knowledge and reflective knowledge cannot be combined in reasoning? As Cohen puts it, animal knowledge is "a relatively minimal cognitive achievement" (2002, 326), so it is plausible that its role in reasoning differs from that of reflective knowledge. Also, animal knowledge does not respect closure principles like:

Single-Premise Closure: If *S* knows that *P*, competently deduces *Q* from that knowledge, and thereby comes to believe *Q*, while retaining knowledge of *P* throughout, then *S* knows that *Q*. (Williamson, 2000)

One's belief that a table is red may track reality even though the entailed belief that the table is not a trickily lit white table does not. (Nozick, 1981) So it is not outlandish to suppose that animal and reflective knowledge play quite different roles in reasoning.

One problem for Cohen's proposal is that some combinations of animal and reflective knowledge do seem kosher. If a subject notices a constant conjunction of sunset with hunger, it seems fine for her to conclude that she generally gets hungry at sunset, even though her knowledge that the sun is setting is animal knowledge.[12] So for this proposal to work, we need a more nuanced story about when animal and reflective knowledge may be legitimately combined.

Another concern is that the proposal does not address Charlie's case. Charlie's knowledge that his tank is full looks like reflective knowledge, since he knows that his gauge is reliable. So restricting animal knowledge's role in reasoning will not prevent him from bootstrapping.[13]

2.4. **No Risk, No Gain (Titelbaum, Douven & Kelp).** A fourth proposed diagnosis: perhaps our bootstrapping agents are guilty of carrying out no-lose investigations. That is, their methods for determining whether a source is reliable cannot condemn the source, they can only yield positive verdicts. Charlie, for example, cannot gather anything but positive data by reading his gauge and trusting what it says. According to this diagnosis, such methods cannot justify a belief.

A number of authors have considered and/or defended this general approach. White (2006) discusses (but does not endorse) the following principle:

[12] *Objection*. The agent's knowledge that she is hungry is also animal knowledge, not reflective knowledge. *Reply*. Then Bernard's knowledge that the table appears red to him should count as animal knowledge too (see fn. 11).

[13] Cohen (2010, 146–8) responds that Charlie's reasoning is flawed for other, more mundane reasons, but we don't have the space to evaluate that response here.

Disconfirmability: If you know that a test cannot disconfirm a hypothesis, no result of the test can confirm the hypothesis either.[14]

Douven & Kelp (forthcoming) endorse the following principle from Glymour:

No Risk, No Gain: "To test a hypothesis we must do something that could result in presumptive evidence against the hypothesis." (Glymour, 1980, 115)

And Titelbaum (2010, 123) endorses a ban on "no-lose investigations", a notion he formally defines as follows:

No-Lose Investigation (NLI): A *no-lose investigation* is one where:

(1) P is not justified for the agent at t_1,
(2) at t_1 the agent knows that $\neg P$ will not be justified for her at t_2, and
(3) at t_1 the agent knows that if P is true, P will be justified for her at t_2.

Despite the differences in detail, these three formulations share an appealing core idea: without any epistemic risk, there is no epistemic gain. Let's call this general approach to bootstrapping *the NRNG diagnosis*.

In addition to its inherent plausibility, the NRNG diagnosis seems to draw support from probability theory. The following is an elementary theorem of the probability calculus:

NRNG Theorem: Let $\{O_i\}$ be a partition. If $p(H|O_i) \not< p(H)$ for every O_i, then $p(H|O_i) \not> p(H)$ for every O_i.

Think of H as a hypothesis and the O_i's as the possible outcomes of a test. Then the theorem tells us that if no possible outcome of the test can disconfirm the hypothesis, none can confirm it either.[15]

But when we apply the theorem to cases, the support for the NRNG diagnosis dissipates. In Charlie's case, H is the proposition *The gauge is super-reliable*. But what are the O_i? Charlie is testing H by reading the gauge and then taking it at its word, so it seems the O_i are propositions of the form *The gauge says that* $X \wedge X$. Now it's true that none of these possible outcomes disconfirms H, so it *seems* the theorem tells us that none of them can confirm H either. The trouble is that these O_i do not form a partition; not only is it possible for the gauge to say X when $\neg X$, but Charlie

[14] White draws the idea from (Pryor, manuscript), an unpublished, early version of (Pryor, 2004).

[15] This gloss makes some substantive assumptions, most notably that confirmation can be understood in terms of positive probabilistic relevance. For background on confirmation and positive relevance, see (Fitelson, 1999, 2001).

attributes positive probability to such possibilities. He knows the gauge is reliable, but he does not think it is perfect. So, presumably, he thinks there is some small chance that the gauge will read 'full' when the tank is not full.

The trouble with using the NRNG Theorem to underwrite the NRNG diagnosis is, it's simply false that there are no possible outcomes that decrease the probability of H. What is true is that, if such an outcome obtains, Charlie will not realize that it does. While every outcome is one where the things Charlie comes to believe do not disconfirm the hypothesis, those things do not form a partition. So closer examination suggests that the NRNG Theorem does not actually underwrite the NRNG diagnosis.[16]

Now that does not mean the NRNG diagnosis is mistaken. But it does cast doubt on its prima facie appeal. Perhaps the plausibility of the principles proposed by Douven & Kelp and by Titelbaum derives from our intuitive recognition that the NRNG Theorem is true. If so, the fact that the theorem does not apply to bootstrapping cases suggests that the NRNG diagnosis does not apply either.

That doubt is amplified by a point raised by Vogel: "the process by which I know I am conscious when I am [...] could not return a verdict other than that I am conscious." (2000, 615) Vogel's case is like Charlie's in an important respect: there is a possible outcome that disconfirms the hypothesis, it's just that the truth of the outcome will not be believed should it obtain.[17]

Further concerns about the NRNG diagnosis arise from variations on bootstrapping cases like the following:

> *Denise.* Denise is just like Charlie except that, every time she reads the gauge, she flips a coin to decide how to proceed. If it comes up heads, she will take the gauge at its word. If it comes up tails, she will use a dipstick to corroborate. As it happens, the coin lands heads over and over again, so Denise never resorts to the dipstick.

---

[16] *Objection.* We simply chose the wrong O_i in our analysis. The O_i are instead the propositions of the form *The gauge says X*. These do plausibly form a partition, and none of them could disconfirm H, so the theorem really does tell us that none could confirm H either. *Reply.* Now we are simply noting that the gauge's reading does not, *by itself*, confirm H. What we wanted to know was why the gauge's reading *together with knowledge of the tank's state* does not confirm H.

[17] Titelbaum responds that Vogel's example does not satisfy his definition of an NLI since it is synchronic; Titelbaum's formal definition of an NLI is diachronic, requiring the existence of distinct times t_1 and t_2. A natural response is to modify Vogel's example so that P is *I will be conscious tomorrow at noon* instead of *I am conscious now*. Titelbaum (personal correspondence) replies to this variation by citing a caveat to his proposal designed to handle cases of "diminished capacity." (Titelbaum, 2010, 128, fn. 16) The idea is that you know now that you will not be justified in believing $\neg P$ at noon tomorrow only because, if $\neg P$ is true, you will be unconscious and so not believe anything at all. I am not sure what to make of this reply.

Denise does not satisfy part (2) of Titelbaum's definition of an NLI: she does not know ahead of time that she will not be justified in rejecting the hypothesis of super-reliability. For all she knows, the coin will come up tails at some point, and resorting to the dipstick will show that the gauge misread on that occasion. Yet surely Denise is no more justified in concluding that her gauge is super-reliable than Charlie is.

The formulations discussed by White and by Douven & Kelp may do better here. In one sense, Denise does do something that could result in presumptive evidence against the hypothesis, viz. flipping the coin. But in another sense, what she does could not have resulted in counter-evidence, since what she actually does is to trust the gauge every time. So what these formulations say about Denise's case depends on how we type her actions: if the coin flip is "something she does" for Glymour/Douven & Kelp, or part of the "test" for White, then the NRNG diagnosis does not work. Otherwise, it does work. Clearly the latter result is what we want, but a principled way of typing actions in order to get it is what's needed.

2.5. **No Feedback (Weisberg).** Our fifth diagnosis focuses directly on the probabilistic aspects of bootstrapping cases. Begin by observing that probabilistic support is intransitive: sometimes A supports B and B supports C, yet A does not support C.[18] For example, Canadian citizens are likely to live in North America, and people living in North America are likely to be U. S. citizens, but Canadian citizens are not likely to be U. S. citizens. In fact, probabilistic support also violates *cumulative transitivity*: sometimes A supports B and $A \wedge B$ supports C, yet A does not support C. For example, if someone scores high on a test, they are probably competent with the material. And if they scored high and are competent, they probably did not cheat. But their high score does not support the hypothesis that they did not cheat. We would expect them to score high whether they cheated or came by their score honestly; a high score cannot discriminate between these two possibilities.

The current diagnosis suggests that bootstrappish reasoning like Charlie's fallaciously presupposes that probabilistic support does obey cumulative transitivity. Charlie begins with observations about the gauge's readings (A), and draws conclusions about the states of the tank (B). He then conjoins his observations about the readings with his conclusions about the states of the tank ($A \wedge B$) to infer that the gauge is super-reliable (C). But the observations he began with (A) do not increase the probability that the tank is super-reliable (C). So, the suggestion goes, Charlie is forbidden from combining his conclusions about the states of the tank with his observations about the gauge's readings in assessing the gauge's reliability. To conjoin

[18] See (Douven, 2011) for an excellent study of the intransitivity of probabilistic support.

his conclusions about the states of the tank with the readings that supported them would be to pretend that probabilistic support obeys cumulative transitivity when it doesn't.

Generalizing this idea, I propose in (Weisberg, 2010) a restriction on epistemic "feedback": when feeding the conclusions of one's ampliative reasoning back into the pool of premises from which they were drawn, one may only infer conclusions from the augmented premise-set that were already supported by the original, *un*augmented premise-set. Having inferred B from A, one may subsequently infer from $A \wedge B$ only what was already supported by A alone. Otherwise we get compounded amplification, distorting the ampliative potential of our premises and resulting in epistemic "feedback".

The following defeater for inductive reasoning formalizes that idea:

No Feedback: If (i) B_1–B_n are inferred from A_1–A_m, and (ii) C is inferred from B_1–B_n (and possibly some of A_1–A_m) by an argument whose justificatory power depends on making C at least x probable, and (iii) A_1–A_m do not make C at least x probable without the help of B_1–B_n, then the argument for C is defeated. (Weisberg, 2010, 533–4)

The intuitive idea is that the premises at the root of an inferential chain must probabilistically support the conclusion at the end of the chain if the chain is to provide justification. So, for example, the chain *The Times reports that unemployment is down* → *Unemployment is down* → *The economy is improving* is good because the fact that the *Times* reports a drop in unemployment supports the conclusion that the economy is improving. But the inference chain *The Times reports that unemployment is down* → *Unemployment is down* → *The Times' report is not a fabrication* is not good, because the fact that the *Times* reports a drop in unemployment does not probabilistically support the conclusion that that report is not a fabrication.

No Feedback gets the right results in the cases we've considered so far. In the cases of Abigail, Charlie, and Denise, the readings of the gauge do not bear on the gauge's reliability. And in Bernard's case, the fact that a table looks a certain colour is irrelevant to the reliability of his colour vision.

Still, No Feedback may be problematic. According to No Feedback, what one may infer from a piece of knowledge depends on where that knowledge comes from. Is this true for deductive inferences as well as inductive ones? If so, No Feedback may conflict with closure principles like SPC. When one infers B from A, whether one may then deduce C from B depends on wether A makes C probable. And, sometimes, A won't make C probable; there are cases where A makes B very probable, B entails

C, and yet *A* does not make *C* probable. For example, I parked my car on Avenue A an hour ago, so it is probably parked there now, which entails that it has not been stolen and moved (Vogel, 1990). If No Feedback restricts deductive reasoning as well as inductive reasoning, then it will prevent me from deducing that my car has not been stolen, which appears to conflict with SPC. On the other hand, if No Feedback does not apply to deductive reasoning, then our diagnosis of bootstrapping threatens to be ad hoc.

I think No Feedback does apply to deductive reasoning as well as inductive reasoning. In (Weisberg, 2010, 542), I borrow an idea from (Nagel, 2011) to argue that this does not conflict with SPC, properly understood. No Feedback does forbid me from deducing that my car has not been stolen and moved, but SPC does not say that this deduction grants knowledge. For I would not be deducing that my car has not been stolen *from my knowledge* that it is parked on Avenue A. If I actively judge that my car is parked on Avenue A in the course of deducing that it has not been stolen, that judgment does not constitute knowledge (though I may nevertheless still know that it is on Avenue A). As long as I am actively considering whether my car has been stolen and moved, I cannot justifiedly judge that my car is still on Avenue A. I must first rule out the possibility of theft on some rational grounds; I must cite further evidence that my car has not been stolen, or give some reason why this possibility is too irrelevant or remote to merit consideration in the circumstances, or something similar. Until I do, no active judgment that my car on Avenue A is justified. Hence any attempt to deduce from there that my car has not been stolen and moved is an attempt to deduce from non-knowledge. So it is true that whatever I deduce from my knowledge is knowledge too; it's just that I cannot deduce this particular proposition from my knowledge under the circumstances.

3. Bootstrapping & Peer Disagreement

We've surveyed the literature emerging from the bootstrapping objection to reliabilism. But bootstrapping objections have started to show up in the apparently unconnected literature on peer disagreement. In this section we'll examine one such argument.

Suppose you form an opinion based on the available evidence and arguments, but you learn that an equally informed and capable reasoner—a *peer*—formed a different opinion in response to the same evidence and arguments. What should you think? Should you steadfastly stick to your view, or should you be concessive, perhaps suspending judgment or adjusting your credence to bring it closer to your peer's?

Elga (2007) defends the *equal weight view*, which says that you should think your peer's opinion equally likely to be correct as your own. If you do anything else, he argues, you will be in a position to bootstrap into a deflated estimate of your peer's reliability:

> *Emma & Fahid.* Emma believes Fahid is her epistemic peer, but they have run into a disagreement. She responds by sticking to her opinion, concluding that Fahid must be wrong since she is right. She then takes this as evidence that Fahid is less reliable than she is; here is an instance where she has gotten things right and Fahid has not. After several more such instances, she becomes confident that Fahid is actually not as reliable as she is.[19]

Elga concludes that any divergence from the equal weight view is wrong because it leads to bootstrapping.

What should we make of this argument? It might show that the equal weight view is true. But Emma's bootstrappish reasoning might instead be flawed for some other reason completely independent of her response to Fahid's disagreement. Maybe her reasoning is rule-circular, fails to track the truth, constitutes a no-lose investigation, or is a case of epistemic feedback. Let's consider each option in turn.

Is Emma's reasoning rule-circular? If there is a rule her bootstrapping is being used to support, it looks to be something like, *Stick to your guns in cases of peer disagreement*. And she does seem to apply this rule in the reasoning she uses to support it; she must trust her own reasoning over Fahid's in order to conclude in a given instance that she is more likely to be right than he is. So Emma's reasoning may well be rule-circular.[20] Still, we've seen that rule-circular reasoning can be legitimate (§2.1), so perhaps it's best not to settle for this diagnosis.

Does Emma's belief fail to track the truth? It seems so. Even if Fahid were in fact her epistemic peer, sticking to her guns in the face of each disagreement would lead her to conclude that Fahid is wrong each time, and thus that he is less reliable than she is. Still, as in Sosa's garbage chute example (§2.2), Emma might be able to deduce that Fahid is probably her epistemic inferior from facts she does track. If she is in fact right in each case of disagreement, while Fahid makes some silly

[19] I've described the case so that Emma favours her own view conclusively, but this is not essential. She can favour her own view a little bit and still bootstrap. If she gives only slightly more weight to her view than to Fahid's, she has some slight evidence that Fahid is less reliable than she is. For here is an instance where she is (slightly) more likely to have been right. The more such instances accrue, the more evidence she gathers that Fahid makes mistakes more frequently than she does.

[20] "May" because I suspect her reasoning could be reconstructed in a way that does not rely on the rule in question.

mistake, her beliefs about the points of disagreement may track the truth, so that she can deduce and know that Fahid is wrong. From a long track-record of such instances, and her knowledge that track-records of failure are generally indicative of unreliability, she could then deduce that Fahid is probably less reliable than she is. So, what tracking theory says about Emma's case depends on whether the tracking theorist makes an exception for deduction (and on external facts about Emma's reasoning and opinions).

Is Emma guilty of carrying out a no-lose investigation? It seems not: by consulting Fahid, Emma does something that could provide presumptive evidence against the hypothesis that he is less reliable. For they might have found themselves in agreement, providing evidence that they are equally reliable. In terms of Titelbaum's formal definition of an NLI, the problem is that condition (2) of the definition fails for Emma. P is *Fahid is less reliable than I am*, but Emma does not know ahead of time that she will not be justified in believing $\neg P$. If Fahid agrees with her in each instance, she will be justified in continuing to believe that he is equally reliable, hence justified in believing it is not the case that he is less reliable.

Is she guilty of violating No Feedback then? This turns out to be a tricky question. The answer depends on what the true epistemic probabilities are in a case of disagreement, which is precisely what's at issue between proponents and opponents of the equal weight view. To see why, let's examine Emma's reasoning from both points of view; first from the perspective of the equal weight view, then from an opponent's perspective.

Let E be the evidence Emma and Fahid share at the outset, and H their point of disagreement. Emma's reasoning then looks like this:

1.	E	Assumption
2.	H	from (1)
3.	I believe H and Fahid believes $\neg H$.	Assumption
4.	My opinion about H is true and Fahid's is false.	from (2), (3)
5.	Fahid is less reliable than me.	from (4)

The premises at the root of her reasoning are the Assumption lines (1) and (3), i.e. the shared evidence E and the fact that she and Fahid disagree. Emma is guilty of feedback if lines (1) and (3) do not probabilistically support her conclusion on line (5). Do they?

According to the equal weight view, they do not. For according to the equal weight view, when she and Fahid disagree, Emma should think it just as likely that she is

right as that Fahid is right:

$$p(H|E \wedge \text{I believe } H \wedge \text{Fahid believes } \neg H) \quad = \quad 1/2.$$

So her root premises, (1) and (3), do not support her ultimate conclusion, and she is not justified in concluding (5) according to No Feedback. She can only get to (5) by feeding her conclusion in (2) back into her premise-set. So according to the equal weight view, Emma is guilty of feedback.

What about opponents of the equal weight view, will they think Emma is guilty of feedback? Not necessarily: they may say that Emma's root premises do probabilistically support her conclusion. To see this, consider one alternative to the equal weight view, what Elga calls the *right reasons view*. On this view, when Emma encounters Fahid's disagreement, she should stick to her guns if E really does support H. If Emma was right that E is a reason for believing H, she should continue to rely on that reason even in the face of Fahid's disagreement.[21] Presumably, a proponent of this view thinks that:

$$p(H|E \wedge \text{I believe } H \wedge \text{Fahid believes } \neg H) \quad = \quad p(H|E).$$

So if Emma was right to infer H from E in the first place, her root premises will probabilistically support her conclusion in (5). And in that case, she is not guilty of feedback.

This result will be agreeable to the proponent of the right reasons view though. In the case where Emma is right that E supports H, the right reasons advocate will want to say that Emma's reasoning is good. Emma is right to conclude that Fahid is less reliable than she is, since she really does have evidence that Fahid is less reliable: given E, H is probably true, and yet Fahid believes $\neg H$. (And in the case where Emma is wrong that E supports H, she is wrong to conclude that Fahid is less reliable, but the blame can be placed on her mistaken inference from line (1) to (2).)

What is the upshot? As far as the debate between the equal weight view and its competitors go, this analysis weakens Elga's argument for the equal weight view. We observed that the right reasons view cannot blame Emma's bootstrapping on a violation of No Feedback, but we also observed that right reasoners would not want to. From their point of view, Emma is blameless. Accusing her of "bootstrapping" is unfair, since she really does have good reasons for adjusting her estimate of Fahid's reliability.

[21] On this view, we blithely disregard the fact that, in a sense, Emma may not be able to tell what circumstance she is in: one where E really does support H vs. one where she merely mistakenly thinks it does.

As far as testing competing diagnoses of the bootstrapping problem goes, No Feedback scores a point. Both parties to the peer disagreement debate will say that No Feedback makes the right prediction in Emma's case. They will disagree about what the right prediction is, but they will also disagree about what prediction the diagnosis makes. If we are trying to adjudicate between various diagnoses of bootstrappish reasoning, as we were in §2, the discussion here supports No Feedback.

But there's much more to think about when it comes to bootstrapping and peer disagreement. We should consider other alternatives to the equal weight view besides the right reasons view, like Kelly's "total evidence view" (2010). In fact, Kelly provides his own bootstrapping arguments for his total evidence view, hence against the equal weight view. How do his arguments square with Elga's, and with the diagnoses we considered in §2? Even just sticking to the territory we did cover, there's more to consider. Emma's reasoning might be reconstructed in other ways importantly different from the five-step argument outlined above. An alternate reconstruction might get different predictions from No Feedback, and from our other diagnoses. With so much left to consider, the conclusions of this section are only tentative.

4. Conclusion

Through a series of cases we saw that the bootstrapping problem reaches well beyond reliabilism. After formulating the challenge for reliabilism in the case of Abigail, we used the case of Bernard to show how other Basic Knowledge theories, especially dogmatism, face the same problem. Then, on the basis of Charlie's case, we concluded that non-Basic Knowledge theories face the problem too.

We then looked at five proposed solutions to the problem. First, Vogel proposed a ban on rule-circular reasoning, but White observed that rule-circular reasoning is sometimes ok. Second, Roush suggested that bootstrapping fails to track the truth. I observed that this may not account for the internal irrationality of bootstrapping, and that handling familiar counterexamples to tracking theory may disable its ability to handle bootstrapping. Third, Cohen proposed restricting deductive reasoning, suggesting that animal and reflective knowledge should not be combined in deduction. But, I suggested, this restriction is too strict, and it does not handle Charlie's case anyway. Fourth, Titelbaum and Douven & Kelp suggested that bootstrapping is illegitimate because it is a kind of no-lose investigation. I observed that, while bootstrapping may be a no-lose investigation in some sense, it is not the kind of no-lose investigation ruled out by probability theory. Moreover, Vogel pointed out that some no-lose investigations are legitimate. Also, the case of Denise suggested that bootstrapping needn't be a no-lose investigation anyway. Fifth, I proposed a ban on

feedback, formalized in the No Feedback defeater for inductive reasoning. No Feedback handled all the cases we'd considered so far, but extending it to cover deductive reasoning risked collision with widely held closure principles for knowledge.

Finally, we turned to peer disagreement, focusing on Elga's bootstrapping argument for the equal weight view. We asked whether Elga's bootstrapping subject might go wrong, not by flouting the equal weight view, but in some other way. We observed that she might be guilty of rule-circular reasoning, but we also recalled that rule-circular reasoning is not always illegitimate. We observed that she fails to track the truth, but that she might nevertheless be able to recover her bootstrappish conclusion by deducing it from things she does track. We then observed that she is not guilty of conducting a no-lose investigation, though she might be guilty of violating No Feedback. Interestingly though, whether she does violate No Feedback depends on whether one agrees with the equal weight view. Thus, as an argument for the equal weight view, Elga's bootstrapping argument may beg the question.

References

Cohen, Stewart. 2002. Basic knowledge and the problem of easy knowledge. *Philosophy and Phenomenological Research* 65(2): 309–29.

Cohen, Stewart. 2010. Bootstrapping, defeasible reasoning, and *a priori* justification. *Philosophical Perspectives* 24: 141–59.

Douven, Igor. 2011. Further results on the intransitivity of evidential support. *The Review of Symbolic Logic* 4(4): 487–97.

Douven, Igor & Christoph Kelp. forthcoming. Proper bootstrapping. *Synthese* .

Elga, Adam. 2007. Reflection and disagreement. *Noûs* 41(3): 478–502.

Fitelson, Branden. 1999. The plurality of bayesian measures of confirmation and the problem of measure sensitivity. *Philosophy of Science* 66 (Proceedings): S362–S378.

Fitelson, Branden. 2001. *Studies in bayesian confirmation theory*: University of Wisconsin, Madison dissertation. `http://fitelson.org/thesis.pdf`.

Fumerton, Richard. 1995. *Metaepistemology and skepticism*. Rowman & Littlefield.

Glymour, Clark. 1980. *Theory and evidence*. Princeton University Press.

Goldman, Alvin I. 2008. Reliabilism. `http://plato.stanford.edu/entries/reliabilism/`.

Kelly, Thomas. 2010. Peer disagreement and higher order evidence. In Richard Feldman & Ted Warfield (eds.), *Disagreement*, 111–74. Oxford University Press.

McGlynn, Aidan. forthcoming. Believing things unknown. *Noûs* .

Nagel, Jennifer. 2011. The psychological basis of the harman-vogel paradox. *Philosophers' Imprint* 11(5): 1–28.

Nozick, Robert. 1981. *Philosophical explanations*. Harvard University Press.

Pollock, John L. 1974. *Knowledge and justification*. Princeton.

Pryor, James. 2000. The skeptic and the dogmatist. *Noûs* 34(4): 517–49.

Pryor, James. 2004. What's wrong with moore's argument. *Philosophical Issues* 14: 349–78.

Pryor, James. manuscript. Is moore's argument an example of transmission-failure? `http://www.jimpryor.net/research/papers/Moore2001.pdf`.

Roush, Sherrilyn. 2005. *Tracking truth: Knowledge, evidence, and science*. Oxford University Press.

Roush, Sherrilyn. 2012. Sensitivity and closure. In Kelly Becker & Tim Black (eds.), *The sensitivity principle in epistemology*, Cambridge University Press.

Sosa, Ernest. 1991. Knowledge and intellectual virtue. In *Knowledge in perspective*, 225–44. Cambridge University Press.

Sosa, Ernest. 1997. Reflective knowledge in the best circles. *Journal of Philosophy* 94(8): 410–30.

Sosa, Ernest. 1999. How must knowledge be modally related to what is known? *Philosophical Topics* 26(1/2): 273–84.

Titelbaum, Michael. 2010. Tell me you love me: Bootstrapping, externalism, and no-lose epistemology. *Philosophical Studies* 149(1): 119–34.

Vogel, Jonathan. 1990. Are there counterexamples to the closure principle? In M. Roth & G. Ross (eds.), *Doubting: Contemporary perspectives on skepticism*, 13–27. Kluwer.

Vogel, Jonathan. 2000. Reliabilism leveled. *Journal of Philosophy* 97(11): 602–23.

Vogel, Jonathan. 2008. Epistemic bootstrapping. *Journal of Philosophy* 105(9): 518–39.

Weisberg, Jonathan. 2010. Bootstrapping in general. *Philosophy and Phenomenological Research* 81(3): 525–48.

White, Roger. 2006. Problems for dogmatism. *Philosophical Studies* 131(3): 525–57.

Williamson, Timothy. 2000. *Knowledge and its limits*. Oxford University Press.