Why so negative? Evidence aggregation and armchair philosophy

Brian Talbot

Forthcoming in *Synthese*. The final publication is available at Springer via

http://dx.doi.org/10.1007/s11229-014-0509-z

0. Introduction

"Armchair philosophy," as I'll use the term, refers to a method typified in part by the use of philosophers' intuitions as evidence. A number of empirically based criticisms of this aspect of armchair philosophy have been advanced recently. Critics suggest either that the use of intuitions be largely abandoned, or that when intuitions are used that they be gathered through careful experiments or large-scale studies. The debate between armchair philosophers and their empirical critics would benefit from greater clarity and precision in our understanding of what it takes for intuition-based approaches to philosophy to make sense. This paper discusses some probability-based methods for determining what we can and cannot learn from intuitions in various conditions. Application of these shows that armchair philosophy makes sense in a broad range of situations, including situations that we are quite plausibly in at the moment. These methods also tell us something about what empirical findings would have to show to undermine appeals to intuitions.[1]

The paper starts by pointing out that armchair philosophers typically have more intuitions than just their own to draw upon. An understanding of when armchair philosophy is appropriate

[1] This introduction owes a lot to an anonymous referee who articulated what this paper intended to do much better than I could on my own.

must therefore involve an understanding of what can be learned from multiple intuitions on the same topic. At the same time, we are rarely certain about either the reliability of or the causal relationships between the multiple intuitions we have access to. So understanding what armchair philosophers can learn from multiple intuitions requires allowing for uncertainty about these factors. In light of this, I discuss some formal methods for determining what can be learned from multiple intuitions which allow for such uncertainty; full derivations of these are in the appendices at the end of the paper. I then apply these to explore the boundary conditions within which armchair philosophy makes sense. I argue that armchair philosophers and their empirical critics should agree that, at the moment, it is reasonable to think that we are within these boundary conditions. Finally, I draw some conclusions about what sort of evidence one would have to give to show that we are not.

1. Armchair philosophy and evidence aggregation

There is a stereotype about armchair philosophy that seems in the background of much discussion of philosophical methodology, which is that armchair philosophy is a mostly solitary endeavor. The stereotypical armchair philosopher consults her own intuitions on a question, writes a paper on the basis of these, sends the paper off, and moves on to the next question. But that's not how philosophy is ordinarily done. Armchair philosophy is highly social: philosophers talk their ideas over with their colleagues and students, present drafts of their papers at conferences, get responses from referees, and so forth. In the course of this process, we get a great deal of feedback about any intuition pumps we deploy. Fairly typical armchair philosophy will involve learning about the intuitions (about a single proposition or intuition pump) of anywhere from a handful to a couple of dozen people. In evaluating whether armchair uses of intuitions make sense as a practice, we need to ask whether it makes sense to base philosophical claims on the intuitions of *groups* of philosophers, albeit small groups.

The armchair use of intuitions that I am interested in here is using intuitions that *p* or that *~p* to learn about the truth of *p*. To discuss this, I will talk about how agreement or disagreement in intuitions change the *epistemic probability* of a given proposition for an agent. This is because it's easiest to frame discussion of evidence aggregation probabilistically. By "epistemic probability," I mean the probability that would be rational for an agent to assign to a proposition in light of that agent's evidence.[2] Sometimes (to avoid unpleasant repetition) I will use the term "credence" or "confidence" to refer to this number, and when I do I'm referring to the credence an agent *should* have in a proposition, not the one they do have. I take it to be epistemically desirable for an agent to believe propositions that have a very high epistemic probability for him or her, and I think that a proposition having a sufficiently high epistemic probability is often a sign that believing it is justified. In what follows, I will focus on epistemic probabilities above .9, but this is only for illustrative purposes.

Now that I've clarified what I have in mind when I talk about armchair philosophy, I should also briefly clarify the other views this paper will discuss. I will also talk about *empirical criticism* of armchair philosophy, which is the project of using empirical data to undermine the use of intuitions in philosophy. And I'll say something about *positive experimental philosophy*, a sort of philosophy that uses intuitions about *p* to provide evidence that *p* is true, but gathers these intuitions via experiments or large-scale surveys. There are other views on the use of intuitions in philosophy, but it's the debate between armchair philosophy, its experimental critics, and positive experimental philosophers that I will focus on. I will consider what empirical critics would have to show to show that the sort and amount of evidence that armchair philosophy can

2 Thank you to Tyler Hildebrand for helping me clarify this idea.

provide us with is insufficient. This will also help us see what sorts of findings might motivate a switch to the methods that positive experimental philosophers advocate.

If one makes a few assumptions – that the gathered intuitions are on average reliable, and that they are independent of each other – then when the majority of people whose intuitions are consulted on a question agree, the epistemic probability of the proposition they agree about will often be very high. Given these assumptions, achieving epistemic probabilities above .9 typically does not require either very large groups of intuitors or very large majorities. This is in some sense an old point, as it is just an application of the Condorcet Jury Theorem to intuitions (de Condorcet 1785).[3] More recently, writers have made similar points about eyewitness testimony, typically supported with Bayesian arguments (e.g. Schum & Martin 1982; Holder 1998). If intuitions were independent and on-average reliable, or if it were reasonable to think that they were, then armchair methods would be vindicated.

However, it is not reasonable to think that intuitions are independent in the relevant sense. It is at least possible that philosophers can causally influence each other's intuitions to some extent. Modeling this possibility means treating intuitions as not independent, as we'll see later in this paper. Even if we thought there was no possibility of causal influence between intuitions, independence would still be extremely implausible. It can be proved that, if intuitions were independent in the relevant sense, then the epistemic probability of a proposition would not depend on the size of the group who found it intuitive, or of the group that did not, but rather just on the difference in size between these two groups (see Appendix A for the proof).[4] If that were

[3] See Goldman (2010) for an instance of a philosopher discussing the idea that we might apply the Theorem to intuitions.

[4] Throughout this paper I only discuss those who have a certain intuition and those who have a contradictory one. Sometimes, however, we find that one has no intuition on a topic at all. If so, what should this tell us? A fairly

true, then finding out that 5,010 people intuit that *p* and 5,000 that ~*p* would give one just as good of evidence that *p* as finding that 11 people intuit that *p* and only 1 that ~*p*. This flies in the face of common sense. In the former case, where the proportion intuiting that *p* is lower than in the latter, the level of disagreement we find should suggest that these intuitions are not all that reliable.[5] This is a form of what we might call *informational non-independence*: knowing the content of some intuitions helps us to determine how good of information other intuitions on that topic are, which helps us to predict their content. Thus, to investigate when armchair philosophy makes sense, we need to go beyond the approach to evidence aggregation embodied by Condorcet's Jury Theorem.

Work has been done previously extending de Condorcet's Theorem to cases of non-independence. However, this work focuses on situations in which we know how causally dependent one piece of evidence is on another (e.g. Ladha 1992). Unfortunately, it is relatively rare that we know just how influenced one philosopher's intuitions have been by those of other philosophers. We need a way of calculating epistemic probabilities that allows for uncertainty about causal influence, and a way that allows for informational non-independence. I discuss some methods in the next section. Following that, I'll explore what the application of these tells us about armchair philosophy.

---

common response seems to be to treat this as no evidence at all. However, I have also seen philosophers respond to an absence of intuitions about a proposition as a sign that the proposition is implausible – in other words, treating the absence of intuitions that *h* as akin to intuitions that ~*h*. Either of these two approaches is already easily modeled in the approaches I discuss above, but future research might suggest that some other alternative is more appropriate.

5 My thanks to Michael Tooley for this example.

2. How to aggregate evidence with three types of dependence

Informational non-independence and different forms of causal influence must each be approached mathematically in somewhat different ways. Before I discuss these differences, I want to say something about what each approach has in common. We are considering how one should rationally respond to discovering that a certain number of intuitors intuit that *p* and a certain number intuit that ~*p*. This will be represented by how the epistemic probability of *p* changes for one in light of this discovery. The change in the epistemic probability of *p* in light of all these intuitions depends in part on how reliable one should see these intuitions as being. Since I'll be talking a lot about reliability, I must say something about the notion I'm using. Note that this is a technical definition, not an attempt to analyze any ordinary concept. I am using the term "reliability" to refer to the probability that an intuition will be the intuition that *p*, given the truth of *p*, and given the lack of causal influences from other intuitions of the sort I'll discuss in the next sections.[6]

Two questions might arise upon seeing this definition: why talk about the reliabilities of intuitions, rather than of intuitors or of a faculty for producing intuitions, and why this notion and not another? First, even if one prefers to think of reliability as a property of things that produce intuitions, in the end (when aggregating intuitions) one's evidence consists of intuitions themselves, not intuitors or a faculty. To calculate epistemic probabilities conditional on our

---

[6] One might think that what makes an intuition the intuition it is is the proposition that is intuited. That is, the intuition that *p* couldn't have been the intuition that ~*p*. If that is your view, then the notion of reliability I use here doesn't make sense. It can be replaced with something like: reliability is the probability that a certain event or moment which contains an intuition will contain the intuition that *p*, given that *p* is true. This doesn't affect any of the arguments I make in this paper.

evidence, we cannot avoid considering how likely it is that a given intuition is true, which is the sort of thing I have in mind when I talk about intuitions being reliable. So, why use my notion of reliability and not the most obvious alternative: the probability that *p* is true given the intuition that *p*? This latter notion makes it more difficult to assign intuitions an on-average reliability (which makes certain calculations much harder). To see why, consider a case where Fred has two intuitions, one that some fairly plausible claim is true, and one that some extremely implausible claim is true.[7] It would be quite odd for Fred to now assign both propositions the same probability. This is because the probability assigned to a proposition, given an intuition about that proposition, should take into consideration the prior probability that the proposition is true. Because of this, if "reliability" were used to mean the probability that *p* given the intuition that *p*, the reliabilities of intuitions should change from proposition intuited to proposition intuited, depending on the prior probabilities of these propositions, making it quite difficult to assign average reliabilities to intuitions. The notion of reliability I'm using need not change in that way.

On each approach to dependence I'll discuss, the epistemic probability of a proposition is determined (in part) by its prior probability, the number of intuitions for and against it that one has observed, and the reliability one should assign to these intuitions. How does non-independence factor in here?

Informally, intuitions are independent when they don't cause each other, nor have some mutual cause other than the truth, and when the content of one doesn't give us information about

[7] Fred can have intuitions that are implausible to him because intuitions need not be believed. He might have very good, intuition-independent, reasons to think that the intuited proposition is false. Or he might have many other intuitions that suggest that the intuition is false. E.g. Fred might intuit that there are more integers than even integers despite being aware of proofs that this is false.

the content of the other, except by giving us information about what the truth is (see Van Cleve 2011 for a discussion, and for pointers to further discussion on this topic). Put in terms of probability, they are independent when the probability of one intuition's having the content it does, conditional on the truth of that content, is not affected by information about others' intuitions on that topic. A formal definition of independence is:

A's intuition that p is independent of B's intuition iff

$$\Pr(\text{A intuits that } p \mid p \ \& \text{ B has her intuition}) = \Pr(\text{A intuits that } p \mid p)$$

$$\text{and } \Pr(\text{A intuits that } {\sim}p \mid {\sim}p \ \& \text{ B has her intuition}) = \Pr(\text{A intuits that } {\sim}p \mid {\sim}p)$$ [8]

Independence requires that the two sides of the above equation be equal, and so non-independence simply requires that they be unequal. Since there are infinite ways in which they can be unequal, non-independence can take an infinite variety of forms. I'll discuss three mathematical models of non-independence in this paper. I've picked these to model informational non-independence and two forms of causal influence. These should allow us to draw conclusions about a range of types of non-independence.

2.1 Informational non-independence

The pattern of intuitions on a topic that we observe might tell us something about how reliable intuitions are on that topic. As we discussed above, having a very large percentage of a group agreeing on a topic suggests to some extent that the individuals are either fairly reliable or fairly anti-reliable, whereas a 50/50 split suggests more that they are unreliable rather than reliable or anti-reliable. This is a form of non-independence, since learning about how reliable some intuitions are gives us information about how likely similar intuitions are to have certain contents. For example, more reliable intuitions are more likely to have content that we deem as

[8] My thanks to an anonymous referee for helping me to properly articulate this formulation.

independently likely to be true. To determine what groups of intuitions tell us about the probability of their content, we need to be able to appropriately take into consideration what these intuitions tell us about their own reliability.

So, how do we assign reliabilities to intuitions we have gathered, keeping in mind that the intuitions we have gathered are data about how reliable they are? We start with the idea that different intuitions can (in principle) have different levels of reliability. Since reliabilities are probabilities, these can take any value between 0 and 1. If we think about the whole population of all intuitions, or the whole population of intuitions relevantly similar to the ones we are observing (e.g. the intuitions of philosophers, or the intuitions of philosophers about ethics), we are probably warranted in thinking that certain levels of reliability are more common or more likely than others. For example, if we are warranted in thinking intuitions are highly reliable, then we will see high reliability values as very likely, although we should probably think that *some* intuitions are very unreliable, and so it is possible that some randomly selected intuition has a very low reliability value. So, for any intuition we confront, when we have little information about its reliability in particular, we can assign each possible reliability value it could have a probability according to how often we think that reliability occurs in the relevant overall population of intuitions.

When presented with a group of intuitions about *p*, we can use the degree to which they agree or disagree about *p* to narrow down our views of how reliable they each are likely to be. Let's call this the *narrowing down* approach to non-independence. To implement this, we consider the group of intuitions we have as a series of intuitions (the order does not matter). As we consider each intuition in the series, we also consider each possible reliability value it could have. For each of these values, we ask, "If this were the average reliability value of these intuitions, how likely would the observed distribution of intuitions be to occur?" This allows us to update our views on how likely it is that this reliability is the average reliability of this set of intuitions.

How do we then determine the probability that p? We know that, for the range of possible reliability values $r_1 \ldots r_n$, each intuition either has reliability value r_1 or r_2 or r_3, and so forth.[9] For each intuition, we can determine what it would tell us about how likely p is were r_1 the average reliability value, or were r_2, or were r_3, and so forth. We weigh what each intuition would tell us about the likelihood of p, given each specific reliability value, by the likelihood of that particular reliability value obtaining (which we obtained from narrowing down our prior views). This can be seen as calculating an expected probability that p. I derive this approach from the probability calculus and give more details on just how it works in Appendix B.

Employing this narrowing down approach requires having some warranted views about the distribution of reliability values in the relevant overall population of intuitions. The relevant population might be, for example, the population of intuitions of philosophers, or intuitions of philosophers who work in a particular area, or the intuitions of philosophers on a certain subject; I won't take a position on this issue. In my discussion below, I assume reliabilities are normally distributed and consider a range of possible averages and standard deviations.

Olsson and Shogenji (2004) discuss a similar notion of dependence, which they attribute to C.I. Lewis (1946). Their discussion differs in two significant ways from that in this paper. First, they do not factor in the possibility of anti-reliability: that the probability of a source of evidence giving the correct answer might be less than chance. I will allow for that in this paper.

---

[9] Since reliability values are probabilities, they can take any real number between 0 and 1 as their value. So I am really describing a method for closely approximating the probability that p; see Appendix B.

Second, they assume that levels of reliability are evenly distributed – that every level of reliability has an equal chance of being the actual level of reliability of some source or piece of evidence. My approach can be used given any distribution that can be mathematically modeled, although in this paper I assume a normal distribution. These two differences have significant consequences. As Van Cleve (2011) points out, "[Olsson & Shogenji's approach] means we should assign a probability of ¾ to the testimony of a witness of unknown reliability on a true-false question [where true and false have equal prior probabilities]." (Van Cleve 2011, 362) To briefly explain with some slight over-simplifications: on their model, the witness can be either fully reliable or only reliable as chance, which are equally likely (we get similar results if we allow any possible reliability between full and chance). If the witness is entirely reliable, there is a 100% chance they are correct, and if they are only as reliable as chance they have a 50% chance of getting the right answer. Since these are equally likely, the probability of getting the right answer is 75%. To see why this is implausible as a general rule, consider three cases: 1) Esther asks a randomly chosen person if the number of stars is odd or even; 2) Esther asks her poker opponent if Esther should call or not; 3) Esther asks Tim if Tim's favorite singer is a man or woman. In each case, assume the *prior* probability of each possible answer is .5. In none of these should Esther have a 75% confidence in the answer that is given. In 1 and 2, she should see the given answer as less likely to be true, and in 3 more likely. The problem with 1 and 3 is that they assume that 50% reliability is just as likely as 100% reliability. But a person is much more likely to be 50% reliability than 100% reliable when guessing about the number of stars, and is much more likely to be 100% reliable than 50% reliable in stating their preferences. The problem with 2 is that our opponent is likely *anti*-reliable, and this approach does not model that.

2.2 Causal non-independence

Another type of non-independence is causal dependence. Intuitions might influence each other: for example, hearing that I intuit that *p* might make my advisees more likely to have this intuition. Two intuitions on a topic can also be both influenced by some common cause that is somewhat unrelated to the truth of the matter they are both about. For example, intuitions might be influenced by media portrayals of the topic they are about. These sorts of causal dependences can vary in ways that require distinct mathematical models. Rather than discussing every possibility, I'll instead consider two types of causal dependence. These will serve as a template for modeling other types, and will allow us to extrapolate or interpolate to a wide range of plausible sorts of causal dependence.

Symmetric dependence occurs when each intuition on a topic has some chance of determining the content of each other intuition on the topic. For *asymmetric dependence*, only one subset of the intuitions gathered – in this paper, those in the majority– has a chance of determining the content of individual intuitions, and the chance that this occurs is independent of the size of the group.[10] One could also define a form of asymmetric dependence in which non-majority intuitions can be the product of influence. This will be less interesting for our purposes, as such dependence does not pose much of a threat to armchair philosophy. This is because it would make divergence in intuitions less worrisome, as it would reduce the significance of the

---

[10] Note that causal dependence is only worrisome, for our purposes, when it holds between two intuitions in the set of intuitions one is using as evidence about a given proposition. If you gather my intuition that *p*, and my intuition that *p* is dependent on Fred's, this doesn't matter if you don't also gather Fred's intuition. This is because dependence is bad (for our purposes) when it reduces the information that an intuition gives us. This only occurs when intuitions in the set gathered are dependent on each other, as some will be redundant given the other; dependence on non-members of the set gathered doesn't create this redundancy.

disagreeing minority intuitions.

For both the symmetric and asymmetric notion of dependence, the credence that a set of intuitions licenses in a proposition depends on how dependent intuitions are on each other. I will call this the "degree of dependence." The *degree of symmetric dependence* is the probability that any given intuition in a group will have its content determined by any other given intuition in that group, independent of the truth or falsity of that content. Note that each intuition has this probability of influencing each other intuition in the group. Formally:

> intuition A is symmetrically dependent to degree *d* on intuitions $\{i_0\ldots i_n\}$ iff for each of $\{i_0\ldots i_n\}$, there is a *d* chance that the content of A is determined by it

The *degree of asymmetric dependence* is the probability that any given intuition in a group will have the same content as the majority of intuitions in that group, independent of the truth or falsity of that content. Formally,

> intuition A is asymmetrically dependent to degree *d* on intuitions $\{i_0\ldots i_n\}$ iff there is a *d* chance that the content of A is determined by the content shared by the majority of $\{i_0\ldots i_n\}$

If we know the degree of dependence of intuitions on some topic, and their reliability, it is relatively easy to calculate epistemic probabilities from sets of intuitions, but what do we do if we don't know the degree of dependence? As with the narrowing-down approach, we have to use the set of intuitions gathered about *p* as evidence both about *p* and about the degree of symmetric/asymmetric dependence of these intuitions. We start with our view about the distribution of degrees of dependence in the relevant population of intuitions. We then consider the observed intuitions as a series. For each intuition in the group, we consider each possible degree of dependence it could have and ask, "How likely would the observed distribution of intuitions so far be given this degree of dependence?" This updates our prior views on the likelihood of each degree of dependence obtaining. We then determine what each intuition would tell us about the probability of *p* given each possible degree of dependence, and we weigh this by how likely each degree of dependence is. This gives us an expected value for the probability of *p*. Appendix C gives more details on these calculations.

To do these calculations, we need to have views about the prior probability of each possible degree of dependence. As for levels of reliability, I assume that degrees of dependence are normally distributed in the relevant population of intuitions. In my calculations below, I'll consider a range of possible averages and standard deviations.

The above approaches to modeling (possible) non-independence give us tools that we can use to figure out how confident we should be in specific propositions in specific circumstances. These will allow us to bring more rigor and precision to the discussion of empirical findings and philosophical methodology, which is one of the goals of this paper. The appendices discuss how to implement these tools in actual practice. Note that if we had reason to think that intuitions were causally dependent in a way that does not map onto either symmetric or asymmetric dependences, the methods I discuss can be easily adapted to model this.

In the next section, I will consider what these approaches to non-independence tell us about when armchair philosophy would and would not be appropriate.

3. When does the use of intuitions make sense?

The practice of armchair philosophy is characterized by the use of intuitions as evidence. The circumstances in which one might attempt to learn something from intuitions can vary in a number of ways. The variables include: the number of intuitions one has access to, the prior probability of the proposition the intuitions are about, the perceived average reliability of intuitions in general, the likely distribution of reliability values around that average, the type of dependence relevant, and the likely distribution of dependence values in the population of intuitions. I want to discuss what armchair philosophers can learn from intuitions, but it's infeasible to consider what we can learn in all the possible scenarios we could be in. Instead, we'll condense a great deal of information by making some simplifying distinctions.

Since we want to know when the practice of armchair philosophy is appropriate, let's first decide what "appropriate" means. A minimum requirement is that it must be possible to achieve a high enough epistemic probability in a proposition by doing armchair philosophy. This raises two further questions: what is a "high enough" epistemic probability, and what counts as "doing armchair philosophy?" I'll use .9 as the probability threshold that must be crossed. While this is arbitrary, it seems worth aiming for and will give us a good sense of the landscape. It also turns out that the results I discuss below won't change in very interesting ways were we to use a somewhat higher or lower threshold. To characterize armchair philosophy, we need to differentiate it from something like positive experimental philosophy, which also uses intuitions from multiple sources as evidence. The difference is that armchair philosophers draw just on the intuitions that they have relatively easy access to – those of their colleagues, or of the audiences at their talks. I'll consider what armchair philosophers can learn about *p* from 20 intuitions about *p* (their own included). The number is again arbitrary but I think it fits well enough with ordinary practice. So, to investigate the conditions in which armchair philosophy makes sense, I am going to look for conditions in which we can get rational credences above .9 in propositions by considering the intuitions of 20 intuitors.

It might turn out that we can get a high credence in a proposition by talking to 20 people as long as they all agree. That *that's* possible in some circumstances doesn't count very much in favor of armchair philosophy. What would be more interesting would be if one could achieve a credence over .9 in some proposition by talking to twenty or fewer other people even if many of them *dis*agreed. The best-case scenario for armchair philosophy would be if one could achieve a .9 credence that *p* when just a bare majority of intuitions favoring *p* – when 11 out of 20, or 55%, agree that *p*. But armchair philosophy would be a worthwhile practice even if this were not possible. After all, most of us aren't going to deploy intuition pumps when we anticipate that close to half of our audience is going to disagree with us on them. Armchair philosophy, as actually practiced, makes sense if we can get a high credence in a proposition whenever a decently sized majority intuitively agrees about the proposition. I'll use 75% as my cut off, or 15 out of 20 people.[11] So, I'll say *armchair philosophy makes sense* (or "is appropriate") in conditions where achieving a .9 credence that *p* requires 15 or fewer intuitors out of 20 consulted to agree about *p*.

This is arbitrary in a number of ways, but these sorts of simplifications are necessary for us to have a fruitful conversation. Of course one can rationally do armchair philosophy in some conditions where one cannot achieve a .9 credence with only 75% agreement; one might reasonably expect (or find) almost universal agreement on some topics, and this might warrant over a .9 credence. (On the other hand, in certain conditions one would not be able to cross the .9 threshold even with full agreement) In talking about when armchair philosophy does or does not make sense, I'm really just giving a general lay of the land.

If we are in circumstances in which armchair philosophy makes sense, as I'm using the

[11] One thing to note here is that, the credence we can achieve through 75% agreement among twenty intuitors will not be the same as the credence we can achieve through 75% agreement among a different number of intuitors. This is because, the more intuitors we consult, the more we learn about the reliability or dependence values of intuitions, which will affect what we learn about the propositions intuited.

term, it will be rational to use armchair methods before considering positive experimental philosophy. Doing experiments or studies consumes a great deal of time and effort. If we can cross the .9 epistemic probability threshold without this time and effort, it doesn't make sense to use more laborious methods. Even if we desire a higher epistemic probability, we should still look to armchair methods first when in these circumstances, as one or two additional intuitions will usually be all it takes to get from .9 to very close to certainty. What should we do in circumstances where armchair philosophy does generally make sense, but where the distribution of intuitions we actually gather by armchair methods does not push us over .9 with regards to either either p or $\sim p$? I don't have the space to address this in any detail; however, if we have achieved something close to an even split among intuitors, and we are in conditions that generally favor armchair philosophy, I would suggest that it will typically be more efficient to work on developing new intuition pumps than to switch to experimental methods.

3.1 Findings

To explore the conditions in which armchair philosophy does and does not make sense, I considered a large range of situations involving either informational dependence (modeled with the narrowing down approach) or causal dependence using either the symmetric or asymmetric dependence models. For causal dependence, I considered situations where the degrees of dependence were known, and situations in which they were unknown. When applying the narrowing down approach or considering causal dependence with unknown degrees of dependence, I looked at a range of possible distributions of reliabilities or degrees of dependence, modeled by different averages and standard deviations. I looked at how causal dependence interacted with different average reliabilities of intuitions. And I considered how causal and information dependence interacted with the prior probability of the proposition in question.

My findings are illustrated in the accompanying graphs (Figures 1-13). Each bar on each graph tells one the *minimum* percentage of 20 intuitors who must agree that *p* for the epistemic probability that *p* to be greater than .9. The bars on each graph are clustered according to average reliabilities of intuitions, and shaded to correspond to different prior probabilities that *p*. In picking which graphs to include in this paper, I tried to illustrate the boundary conditions at which armchair philosophy mostly will make sense, but beyond which it mostly will not. The information presented is not very fine grained, but it allows us to draw some general conclusions. I'll summarize here the findings illustrated in the tables.

Perhaps unsurprisingly, reliability is very significant in determining whether or not armchair philosophy makes sense. When typical intuitions are just minimally reliable – when the average reliability of the relevant population of intuitions is around .6 – relatively small amounts of potential causal dependence are enough to pose a threat to armchair uses of intuitions. When typical intuitions are highly reliable, even fairly high amounts of known causal dependence, or high likelihoods of causal dependence, are consistent with armchair methods. In between these extremes of reliability, we see a wide range of conditions in which armchair philosophy makes sense. For the rest of this section, I will focus on what happens in these middle conditions, conditions in which the average reliability of the relevant population of intuitions should be seen as .7 or .8.

Informational non-independence, as modeled by the narrowing-down approach, is largely not an issue for armchair philosophy. As noted above, I assumed that reliability values are normally distributed. At moderately wide distributions of reliability values (sd = .2), armchair philosophy makes sense unless both the average reliability of all intuitions and the prior probability of *p* (the proposition one has gathered intuitions on) are low (.6 and .1 or .2 respectively). When the distribution of reliability values is much wider (sd = .5), armchair philosophy largely makes sense at the middle average reliabilities. When the distribution of reliability values gets *very* wide (sd = .7), armchair philosophy starts to break down at .7 reliability, but still makes sense at higher average reliability levels.

What about when we know what degree of causal dependence obtains for the relevant set of intuitions? Here, we can give rough "breaking points," where armchair methods start to make less sense, and much beyond which they won't make sense even at high reliability values. This breaking point for known degrees of asymmetric dependence is about .3, and for known degree of symmetric dependence is around .07. (In the next section I illustrate what these levels of dependence would look like were they to manifest)

Similarly, there are breaking points for *un*known degrees of symmetric and asymmetric dependence. Degrees of causal dependence were assumed to be normally distributed, and so average values and standard deviations can vary. Because of this, there really is not one breaking point, but rather many. But we can illustrate something quite interesting by noting the following: for both symmetric and asymmetric dependence, there is a breaking point around an average degree of dependence of .3, with a standard deviation of .3. This gives us a rough sense of how sustainable armchair practices are given various data about the possibility of causal dependence.

These breaking point values for possible symmetric and asymmetric dependence are worth discussing. For asymmetric dependence, the breaking point when we suspect causal dependence but don't know its precise degree is roughly the same degree of dependence that, if we *knew* it obtained, would start to significantly undermine armchair uses of philosophy. This seems intuitive – if the use of intuitions doesn't make sense when we know they are much more than .3 dependent, then it probably won't make sense when we know that most of the time they are more than .3 dependent. However, if we think this way, the results for *symmetric* dependence look surprising. If we knew that certain intuitions were symmetrically dependent on one another to a degree above .07, armchair use of those intuitions wouldn't make much sense. However, if we knew that, *on average*, intuitions were symmetrically dependent to a much higher degree, even to an average degree of .1 or .2, armchair uses of intuitions *would* make sense (note that .2 was tested but is not illustrated in the accompanying figures). How can that be? If the average intuition is symmetrically dependent to a degree of .1 or .2, then most intuitions are significantly beyond the .07 breaking point – why would armchair philosophy still make sense in such conditions?

This will be less puzzling with a bit of reflection. Recall that, when we have an unknown degree of dependence, to calculate how likely *p* is given a group of intuitions for and against *p*, we consider all the possible degrees of dependence that these intuitions could have, and get a weighted average of what each intuition would tell us about *p* given each possible degree of dependence. The weighting is provided by the likelihood of each degree of dependence obtaining. Even if the distribution of degrees of dependence is centered around a high value, lower degrees of dependence will often be fairly likely as well. Intuitions provide good information at these lower degrees of symmetric dependence. So these contribute to the expected amount of information we get from each intuition. Further, we tend not to get *dis*information at high degrees of symmetric dependence, as long as intuitions are typically reliable – intuitions in these conditions will tend to simply reiterate the truth, since the intuitions that influence them are likely to be correct. So, the expected amount of information we get from a set of intuitions can still be fairly high, even when the average intuition in the whole population is too influenced by other intuitions to be good information.

Why, then, don't we see this pattern with asymmetric dependence? Why is the average degree of asymmetric dependence that makes armchair philosophy problematic right around the known degree of asymmetric dependence that causes problems? Asymmetric dependence is the chance of an intuition being influenced by the majority view, and so intuitions in the minority cannot have been influenced. So, as degrees of asymmetric dependence rise, the information one gets from the majority intuitions goes down but the information one gets from those in the minority does not. This also explains an odd pattern one sees in some of the data (e.g. Figures 11 and 12): in some conditions where we have unknown degrees of asymmetric dependence, we see that it is harder to achieve a .9 confidence that *p* when intuitions are .9 reliable than when they are .8 reliable. Why does increasing reliability decrease the warrant we get from agreement in intuitions? This is because the minority intuitions get more benefit, in some sense, from the increase in reliability than do the majority. The majority intuitions (but not those in the minority) might be the product of causal dependence, and this makes it harder to trust what they say as minority intuitions become more reliable.

While space limits the figures I can present, I'll mention that I also tested a third kind of causal dependence. This is intermediate between asymmetric and symmetric dependence. On this notion, only intuitions in the majority can causally affect other intuitions (so this is like asymmetric dependence in that way); however, each of these intuitions has a chance of influencing each other intuition (this is similar to symmetric dependence). When this level of dependence is unknown, the "breaking point" beyond which armchair philosophy largely does not make sense is around an average degree of dependence of .15, with a standard deviation of

.3.

The data suggests that we should often believe what groups of intuitors intuit. It's worth noting one caveat. In some cases where we know that one particular intuitor (or reasoner) is quite significantly more reliable than the average member of some group of intuitors, we cannot reach a high level of confidence by trusting the group over that individual; this is true especially when we are dealing with small groups of intuitors. On the other hand, it will be relatively rare that we have good reason to think that one particular intuitor is much more reliable than the average member of a group of intuitors, at least when we are dealing with groups of philosophers.

We have seen that armchair philosophy makes sense in a wide range of scenarios. It will only be dubious when intuitions are not very reliable, or given certain expected levels of causal dependence between intuitions. What does this tell us about armchair philosophy at the current time? Should we think we are in a scenario conducive to armchair philosophy? If we should currently think so, what would empirical critics how to show to undermine the practices of armchair philosophy? I address this in the next two sections. First, I will talk about what we should currently believe about causal dependence, and then go on to talk about reliability.

4. Armchair philosophy now, and empirical criticism in the future

In this section I will suggest that, given what we currently know, it is implausible that the intuitions armchair philosophers use are dependent to a high enough degree to make armchair philosophy useless. We could discover that such high degrees of dependence *do* obtain, but until we have convincing evidence for this, it will be reasonable to assume that they do not. I'll also discuss what that evidence should look like, given the results I've discussed.

Symmetric dependence is a very strong notion of dependence: when intuitions are potentially symmetrically dependent on one another, each intuition in a group has a chance of *completely determining* the content of each other intuition in the group. This kind of strong causal connection between intuitions is the sort of thing that might sometimes occur in small groups of people who talk to each other often, but it's *prima facie* implausible that it occurs so much that degrees of symmetric dependence often get above very low levels. To see how implausible this is, let's consider precisely what even low levels of symmetric dependence mean. For degree of symmetric dependence d, the probability that a given intuition out of a set of n intuitions was not influenced by any other is $(1\text{-}d)^{n\text{-}1}$. So a degree of symmetric dependence of .05 means that, for a set of intuitions gathered from a group of 20 people, each intuition has a roughly 62% chance of being wholly determined by some other intuition. This chance goes up to 75% given a .07 degree of symmetric dependence, and to almost 90% for a .1 degree of symmetric dependence. Think about how many topics we rarely talk about with many of our colleagues; further, consider the fact that we often try to use quite novel intuitions pumps. Intuitions little discussed, or intuitions on novel topics, should be less susceptible to influence of this sort. So degrees of symmetric dependence of even .05 should be quite surprising. If that's right, then it would be surprising if the average degree of symmetric dependence in the relevant population of intuitions was as high as .1, or the standard deviation of degrees of symmetric dependence over .1, as either would make make degrees of symmetric dependence of .05 and above common. If we look at figures 8 through 10, we see that this puts us in conditions good enough for armchair philosophy, as long as intuitions as decently reliable. It seems safe to assume, prior to empirical findings on this topic, that degrees of symmetric dependence are distributed in a way that would largely be compatible with armchair philosophy.

It's plausible that asymmetric dependence is much more widespread than symmetric dependence. But is it plausible that degrees of asymmetric dependence worrisome for armchair philosophy are widespread? A .3 degree of asymmetric dependence between intuitions is roughly the line below which armchair philosophy makes sense, and above which it largely won't. A .3 degree of asymmetric dependence would mean that any intuition in a group has a 30% chance of being *wholly determined* by the majority view. Again, given that many of our intuitions are never discussed, and that we often use novel thought experiments, we shouldn't expect such high degrees of asymmetric dependence to be so common, and it would take significant empirical findings to justify the claim that they were.

Symmetric and asymmetric dependence model just one form of causal dependence, and I have not talked about any other forms at length. However, I think the lessons we learn about them should generalize to any plausible form of causal dependence. What we'll see if we considered other models of causal dependence is that degrees of dependence incompatible with armchair philosophy will be *prima facie* implausibly high.[12] This is because these will have to be degrees of dependence that make it very likely that our intuitions are determined by the intuitions of others. The distributions of degrees of such dependences that are reasonable to currently accept should, as we see with symmetric and asymmetric dependence, be compatible with armchair philosophy.

So far I've just discussed the current state of the art. What about the future of armchair philosophy? We could certainly learn new information about causal dependence that undermined armchair philosophy. What would that have to look like? Let's say we learned that

[12] For example, in the previous section I briefly mention the breaking point for an intermediate form of dependence. This was .15, which is quite high, since this is like symmetric dependence in that each intuition in the majority has a chance of determining the content of each other intuition.

the intuitions of participants in one experiment were *symmetrically* dependent on one another, or that we learned even more: that the intuitions of all people about a certain thought experiment, or a certain type of thought experiment, or even about one domain of philosophy, were symmetrically dependent on one another to a specific degree. This would suggest either that intuitions of other people on the same case, or on other cases or topics will be symmetrically dependent as well. However, it only gives us limited information about these other people, cases, or topics. The data would suggest that the average degree of dependence in the population of intuitions is higher, or that the distribution of is wider, than we might have expected. But it doesn't allow us to assign anything like a precise degree of dependence to any set of intuitions that hasn't been directly studied. We've just seen that, in the absence of a more or less known degree of symmetric dependence value for a group of intuitions, even a fairly strong possibility of symmetric dependence still allows for armchair practices to make sense. If the average degree of symmetric dependence in the population of intuitions is below .3, armchair philosophy mostly makes sense. A .3 degree of symmetric dependence is extremely high: it means that, in a group of just *10* intuitors, an intuition has a 98% chance of being wholly caused by the other intuitions in the group. It's going to be very hard to find inductive evidence that shows that the average intuition is so strongly symmetrically dependent. This means that it is difficult to use data on symmetric dependence to cast doubt on the use of intuitions of a group of people on a given topic without very specific research on that particular group and topic.

It will also be difficult to cast doubt on armchair philosophy via inductive evidence of asymmetric dependence, although not quite as hard. One would have to show that the average intuition (on a certain topic) has a .3 degree of asymmetric dependence. This will be tough without specific research on that topic, but since .3 is not such a startling degree of asymmetric dependence, less evidence would be required than to show high average degrees of symmetric dependence.[13]

Armchair uses of intuitions do not tell us much if we know that intuitions are overly causally dependent on one another, or if we know that they are extremely likely to be so dependent. I've argued that, at the time being, armchair philosophers don't have good reasons to worry about causal dependence, and that it will be somewhat difficult for empirical critics to show that intuitions are overly causal dependent. However, one of the other lessons of my findings is that the appropriateness of the use of intuitions will often depend on intuitions being more than minimally reliable. The next section will address what we should think about reliability at the present moment.

5. Reliability

Empirical critics of intuitions, positive experimental philosophers, and armchair philosophers should share the view that, prior to substantial empirical research, the use of intuitions in philosophy is reasonable, which implies treating intuitions as reliable. I'll start with a fairly simple argument for this, and then explore two objections.

Armchair philosophy involves philosophers relying at least on their own intuitions to justify beliefs. One's own intuitions could not justify beliefs if one did not see them as more than minimally reliable, so it seems that armchair philosophers are committed, at least *a priori*, to seeing their own intuitions as fairly highly reliable. It would be unreasonable for any philosopher to see their own intuitions as significantly more reliable than average without

[13] One should also note that inductive generalization from empirical research on intuitions is already a bit risky. The intuitions that get studied and the data that gets reported are not randomly selected, but instead will tend to be selected for their surprising nature.

substantial empirical data to back this up.[14] So armchair philosophy brings with it a commitment to the claim that philosophers are, on average, fairly reliable intuitors.

Positive experimental philosophy experimentally elicits intuitions and then uses them as evidence that what was intuited is true, which requires assuming that intuitions are at least minimally reliable. Positive experimental philosophers aren't obviously committed to the claim that *philosophers*' intuitions are reliable, but they have to think that some people's intuitions are, and in the absence of data showing that philosophers and lay-people are differently reliable intuitors, we should treat them more or less the same. There is little reason, *a priori*, to assign intuitions reliabilities low enough that armchair philosophy is not justified but positive experimental philosophy is; *a priori* arguments will tend to be for conclusions at the "almost entirely reliable" or "no more reliable than chance" extremes. So both positive experimental philosophers and armchair philosophers should probably assign intuitions *a priori* average reliabilities above .6.

Finally, empirical critics of armchair philosophy look to empirical research to undermine the use of intuitions in philosophy. The most plausible interpretation of this is that empirical critics believe that undermining the use of intuitions requires empirical research.[15] This means

---

[14] Michael Huemer has argued that it is reasonable to treat one's intuitions as more reliable than those of others in some situations (Huemer 2011). However, on his view this is only reasonable when one lacks evidence about the reliability of others; once one realizes that humans, or humans of certain types, generally share cognitive capacities, one has such evidence.

[15] As philosophers are notoriously skeptical of *a posteriori* arguments, doing empirical criticism of armchair philosophy if *a priori* arguments against intuitions were available would be making one's life unnecessary difficult. So charity suggests that empirical critics of armchair philosophy either do not have *a priori* arguments against the use of intuitions, or see these as insufficiently strong.

that the empirical critic should be willing to assign intuitions an *a priori* average reliability above .6. Thus, those pursuing these three projects should agree that the absence of data showing unreliability licenses seeing intuitions as reliable enough to justify armchair philosophy.[16]

One might worry, however, that it is not reasonable to see intuitions as reliable enough to justify armchair philosophy in light of findings on "group intelligence."[17] Empirical research on group intelligence has demonstrated that, in the long run, groups perform very well on tasks with the following characteristics: a) the tasks are performed with repeated trials, and b) the correct responses to the tasks can be verified without relying on the group. The use of intuitions in philosophy doesn't have these two characteristics; in fact, this use is often predicated on the idea that there is no wholly intuition-independent way to discover the relevant truths. However, (a) and (b) are not necessary for groups to actually perform well on tasks: all that is needed is sufficient reliability and low enough causal dependence. What (a) and (b) are necessary for is *proving* that a group performs well in the long run on a task. Demonstrations of long run group performance on a task require knowledge of what correct performance looks like, and long-run data. This raises the question of whether it is necessary to be able to show that armchair uses of intuitions have been or will be successful in order to justify these uses of intuitions. I do not have the space in this paper to address this at length. However, it seems implausible to generalize this standard and say that, for any source of evidence *e*, before we can trust *e* we must be able to know when *e* gives correct results without employing *e*. Many standard views in

---

[16] Another aspect of my approaches to aggregating intuitions is that I treat all of the intuitions elicited as equally reliable. Of course, they probably are not. However, we typically won't know just how reliable each of our individual colleagues are. So we should calculate using the expected reliability of each of our intuitions, and the expected value of a normally distributed variable just is the mean value of the variable (Ross, 2007).

[17] This concern was raised by two anonymous reviewers.

epistemology would deny this generalization, claiming that it leads to skepticism. Further, intuitions are often claimed to be a particular source of evidence that can be trusted without having to be verified (see, e.g. BonJour 1998; Huemer 2001; Sosa 1998). From looking at philosophical practice, it seems that a great many philosophers must see intuitions this way, and, as I have argued, all parties to the debate that I am addressing should agree this is *a priori* reasonable.

There is a second objection to the above argument that I should address. Understanding the response to this objection is also important for understanding what empirical findings must do to show that intuitions are less reliable than we currently think. The objection is based in ideas articulated by Michael Huemer (2008). He has argued that intuitions can be good evidence in philosophy even if they are less than 50% reliable. If that's right, then one might use intuitions without believing that they have the sort of on-average reliability that I've been discussing. Huemer argues for this in part by using an analogy: imagine that several eyewitnesses to an accident are asked what the license plate of one of the cars involved is. We know that none of them are more reliable than 50%. Most of them give widely varying answers, but two agree that the license number was X7841A. In that case, we have good evidence that this was in fact the correct license number, even though none of the witnesses is more reliable than 50%, because the chances that two would agree on a false answer is miniscule. So we can have good evidence even if the source of that evidence is not more than 50% reliable. And so perhaps armchair and positive experimental philosophers can agree that intuitions are good evidence without agreeing that they are more than 50% reliable.

However, the analogy doesn't show us much about intuitions. The reason the agreement in Huemer's case provides good evidence is that there is an enormous range of possible answers that can be given to the question, "What did the license plate of that car say?" The odds that any two witnesses would agree through chance alone are astronomically low.[18] For the coherence of bits of putative evidence to be informative, the bits of evidence must be more reliable than chance (ignoring the possibility that we know that they are less reliable than chance). In the license plate case, more reliable than chance could still be much less than 50% reliable. On the other hand, if there are only two possible answers to a question, witnesses must be more than 50% reliable for their coherence to be informative (again ignoring known anti-reliable witnesses). Were the witnesses less than 50% reliable, they would be more likely to agree on an incorrect answer than on the correct answer. To have an intuition is, roughly, to have the feeling that a certain proposition is true, or perhaps that it is false; there are only two possible ways our intuitions can go.[19] And so "witnesses" telling us about their intuitions cannot give us good evidence if they are less than 50% reliable. Since the three positions in the debate that I'm discussing agree that, *a priori*, its reasonable to see intuitions as good evidence, they should agree that, *a priori*, it is reasonable to see intuitors as more than 50% reliable.

One might protest, however, that experimental philosophers often do ask people questions with more than one answer. A typical experiment involves giving participants a proposition and asking them to rate their agreement or disagreement with that proposition on a 5 or 7 point Likert scale. Intuitors would only need to be more than 15 or 20% reliable for intuitions about the correct level of agreement to be evidence (depending on the size of the Likert

[18] Huemer acknowledges this in a footnote, but doesn't discuss its ramifications for intuitions in philosophy.

[19] This might not be entirely true: we might ask someone, "What ethical theory is intuitively true?" and there are more than two possible answers here. But I don't think that this is a model for typical appeals to intuition, which are about the truth or falsity of single propositions. When we present thought experiments to people, we rarely ask, "What is your intuition here?" without specifying what question the intuition is about.

scale).[20] So perhaps positive experimental philosophers need not assume that intuitions are over 50% reliable.

This point about Likert scales is largely irrelevant to positive experimental philosophy, however. Positive experimental philosophy uses intuitions that *p* as evidence that *p*. Agreement on a specific point on a Likert scale is evidence for propositions such as, "People *somewhat* agree with [some proposition]," versus "People *strongly* agree with [some proposition]." If these were the sorts of propositions that positive experimental philosophy was interested in – that is, if each point on a Likert scale represented a distinct proposition whose truth intuitions were supposed to be evidence for – then it would be the case that people would not need to be on average reliable for intuitions to be useful evidence. However, no one doing positive or armchair philosophy is really having a debate about the truth of these sorts of propositions. Rather, what they want to study are intuitions about the proposition intuitors are agreeing or disagreeing with. To be good evidence for the sorts of propositions philosophers tend to care about, intuitions must fall on the correct side of this mid-point more often than not: they must be on average reliable to be evidence for their content.

The above point has a consequence worth noting. For experimental results to tell us about factors that affect the reliability of intuitions, they must show us that these factors affect the *content* of intuitions, and not just their strength. That is, it's not enough to show that an experimental manipulation brought about a significant difference in average reported Likert scores; rather, the manipulation should cause a change in *content* – a move from one side of the mid-point of the scale to the other.

Once we see that the sides of the debate I'm addressing agree that it is reasonable, *a*

[20] My thanks to an anonymous referee for bringing up this point.

priori, to see intuitions as reliable, we then have to wonder just what views of the distribution of likely reliability values of intuitions are reasonable. As noted above, I think *a priori* reasonable views will put average reliability fairly high – it's hard to imagine an *a priori* argument that intuitions are just 60% reliable. However, I can't make more of an argument about that here, as it depends on arguing for very substantive epistemic claims about the source and nature of justification. I will settle for having shown that armchair philosophy works for a very wide range of levels of reliability, and merely suggest that whatever reliability we assign prior to substantial empirical research will quite likely be on high side of the range I've been discussing.

6. Conclusion

One goal of this paper has been to show how to more rigorously and precisely think about the use of intuitions in philosophy and what empirical research tells us about this use. The mathematical tools I have discussed are intended to be independent of controversial assumptions about the epistemology of intuitions, and I hope that all parties in this debate benefit from them. These tools can help us to interpret experimental research about the reliability and independence of intuitions, can shape the direction of future empirical research, and can help armchair philosophers see just how they much weight they should assign various claims based on the results one has obtained from talking to colleagues together with the experimental data out there.

The second goal of the paper has been to apply these tools to the question of whether philosophers should abandon the armchair. We see that armchair uses of intuitions can give us a high rational confidence in propositions as long as it is rational to see intuitions as more than minimally reliable and as not highly causally dependent on one another.[21] It seems that armchair

[21] Some might think that philosophers should not employ methodology that is not actually or necessarily truth conducive, even if such methodology is rational in light of what one reasonably believes; this concern was raised

philosophers and their empirical critics should agree that, *a priori*, it is reasonable to see intuitions as more than minimally reliable. Further, it seems quite implausible, *a priori*, that the intuitions we use in philosophy are causally dependent enough on one another to cause problems. So, at the present moment, it is responsible to use intuitions as evidence, even if one does not gather them experimentally or from large-scale studies. In fact, it seems typically more reasonable to gather intuitions from the armchair than via experiments or surveys. However, I have not shown that armchair philosophy is actually truth conducive, nor have I shown how we could prove that it is. While I have tried to show it is reasonable to *see* armchair uses of intuitions as truth conducive in a range of plausibly actual situations, we could discover that we are not in such a situation. Only substantial empirical data can undermine armchair uses of intuitions, and this data should meet certain constraints.

The most effective way to undermine armchair philosophy is to attack the reliability of intuitions. This is because it will be hard to give strong enough evidence that worrisome levels of dependence obtain for any particular armchair use of intuitions through inductive generalization, and thus dependence-based criticism will tend to require empirical findings directly about the particular intuitions being criticized (this is especially true for symmetric dependence, but is an issue for asymmetric types of dependence as well). However, there is a challenge to undermining armchair philosophy via data about reliability. The notion of reliability that I have used, and that supports armchair philosophy, has to do with how likely one

by one anonymous reviewer. This concern seems to reflect debates between Bayesians and classical statisticians about the use of necessarily truth conducive methodology versus the "confirmation conducive" methodology of Bayesian epistemology. I have hoped to set this sort of debate aside and focus on what makes our beliefs rational, an approach widely accepted in philosophy. My thanks to Hanti Lin for help in understanding this worry.

is to have an intuition with the correct content, and has nothing do with the strength of the intuition. When looking at data to determine what it tells us about the reliability of intuitions, we should focus our attention on changes in content of intuitions, and not just changes in Likert scores.

Acknowledgements

I owe a great deal of thanks to Julia Staffel, Kenny Easwaran, Dom Bailey, Chris Heathwood, Eric Chwang, Tyler Hildebrand, Adam Keeney, Michael Tooley, Hanti Lin, and Christian Lee for helping me to sort out my arguments and ideas. Thank you also to Jonah Miller for his suggestions on some of the math. This paper also greatly benefitted from anonymous journal referees, whose feedback on previous versions pushed me to really work out the material in this final version.

References

Alexander, J., Mallon, R., Weinberg, J.M. (2010). Accentuate the negative. The Review of *Philosophy and Psychology*, 1(2), 297-314.

BonJour, L. (1998) *In Defense of Pure Reason*. Cambridge: Cambridge University Press.

de Condorcet (1785). Essai sur l'application de l'analyse a la probabilite des decisions rendues a la pluraite des voix. http://gallica.bnf.fr/ark:/12148/bpt6k417181/f4.image. Accessed 18 June 2012.

Goldman, A. (2010). Philosophical naturalism and intuitional methodology. *Proceedings and Addresses of the American Philosophical Association*, 84(2), 115-150.

Holder, R.D. (1998). Hume on miracles: Bayesian interpretation, multiple testimony, and the existence of God. *British Journal for the Philosophy of Science*, 49(1), 49-65.

Huemer, M. (2011). Epistemological Egoism and Agent-Centered Norms. In Dougherty, T. (Ed.) *Evidentialism and its Discontents* (17-33), Oxford: Oxford University Press.

Huemer, M. (2008). Revisionary intuitionism. *Social Philosophy and Policy*, 25(1), 368-392.

Huemer, M. (2001) *Skepticism and the Veil of Perception*. Lanham: Rowman & Littlefield.

L'Ecuyer, P. (2012) SSJ : A Java library for a stochastic simulation. http://www.iro.umontreal.ca/~simardr/ssj/indexe.html. Accessed 20 June 2012.

Ladha, K.K. (1992). The Condorcet Jury Theorem, free speech, and correlated votes. *American Journal of Political Science*, 36(3), 617 – 634.

Ross, S.M. (2007). §2.4 Expectation of a random variable. *Introduction to probability models* (9th ed.). Burlington: Academic Press.

Schum, D.A. & Martin, A.W. (1982). Formal and empirical research on cascaded inference in jurisprudence. *Law & Society Review*, 17(1), 105-152.

Sosa, E. (1998). Minimal intuitions. In DePaul, M., Ramsey, W. (Eds) *Rethinking Intuitions*. Lanham: Rowman & Littlefied.

Van Cleve, J. (2011). Can coherence generate warrant *ex nihilo*? Probability and the logic of concurring witnesses. *Philosophy and Phenomenological Research*, 82(2), 337-380.

Appendix A: Aggregating independent intuitions

In this section I show that, if intuitions are independent, the absolute sizes of the groups intuiting some proposition and its negation are irrelevant to the epistemic probability one should assign to the proposition. What matters instead is the difference in size between these two groups.

If we have some proposition *P*, and some evidence *E* which consists of intuitions for and against P, we can calculate the credence we should assign to P via Bayes' theorem:

$$1.\quad \Pr(P|E) = \frac{\Pr(P)\Pr(E|P)}{\Pr(P)\Pr(E|P)+\Pr(\sim P)\Pr(E|\sim P)}$$

Let's define some useful variables:

f is the number of people who intuit that *P*; by stipulation, *P* will always be the proposition found intuitive by the majority.

a is the number of people who find ~*P* intuitive. I will assume no one is agnostic.

r is the on-average reliability of the relevant body of intuitions.

To calculate Pr(*P*|*E*) via Bayes' theorem, we need to calculate Pr(*E*|*P*) and Pr(*E*|~*P*). Pr(*E*|*P*) is the probability that *f* people find *P* intuitive and *a* people find it counter-intuitive, given that *P* is true. Given that *P* is true, *f* people can intuit that *P* only if they get things right, and *a* can intuit that ~*P* only if they get things wrong. For each intuitor, the probability of getting things right is *r* and the probability of getting things wrong is 1 – *r*. Because we are assuming independence, the probability that *f* people get things right and *a* get things wrong is $r^f(1\text{-}r)^a$, times the number of combinations of people that can result in this split of intuitions. Note that this last factor will appear in Pr(*E*|~*P*), and so can be cancelled out of the equation.

To find Pr(*E*|~*P*), we instead multiply $(1\text{-}r)^f$ times r^a, again times the number of possible combinations (which cancels out).

Putting this all together, we get:

$$2.\quad \Pr(P|E) = \frac{\Pr(P)r^f(1-r)^a}{\Pr(P)r^f(1-r)^a+\Pr(\sim P)r^a(1-r)^f}$$

We can factor r^a and $(1\text{-}r)^a$ out of both the numerator and denominator (since we are assuming *f* > *a*), and get the following:

$$3.\quad \Pr(P|E) = \frac{\Pr(P)r^{f-a}}{\Pr(P)r^{f-a}+\Pr(\sim P)(1-r)^{f-a}}$$

Here we see that only the difference between the number intuiting that *P* and that ~*P*, but not the size of either group, matters.

Appendix B: Implementing the narrowing down approach

This section will discuss how to implement the narrowing down approach, assuming that there is no possible causal dependence between intuitions.

For any proposition *P* and intuitions $I_1 \ldots I_n$, which are intuitions that *P* or intuitions that ~*P*, by Bayes' theorem:

$$4.\quad \Pr(P|I_1 \ldots I_n) = \frac{\Pr(P)\Pr(I_1 \ldots I_n|P)}{\Pr(P)\Pr(I_1 \ldots I_n|P) + \Pr(\sim P)\Pr(I_1 \ldots I_n|\sim P)}$$

Pr(*P*) and Pr(~*P*) are the prior probabilities of *P* and ~*P*, and I'll assume that these are known (and rational). So what needs to be calculated are $\Pr(I_1 \ldots I_n|P)$ and $\Pr(I_1 \ldots I_n|\sim P)$.

Let's start with $\Pr(I_1 \ldots I_n|P)$. Since $I_1 \ldots I_n$ is a conjunction of intuitions, $\Pr(I_1 \ldots I_n|P)$ is the product of the probability of each intuition in the set given *P*; because these intuitions are not independent, the probability of each intuition in the set must be calculated given all intuitions before it (by the multiplication rule). Formally,

$$5.\quad \Pr(I_1 \ldots I_n|P) = \prod_{i=1}^{n} \Pr(I_i|P\&[I_1 \ldots I_{i-1}])$$

$I_1 \ldots I_{i-1}$ is in brackets because, when i=1, no term should be substituted in that place.

How do we calculate $\Pr(I_i|P\&I_1 \ldots I_{i-1})$ for each intuition in the set? We don't know precisely what reliability value each intuition in the group has. However, we do know all the possible reliability values each *could* have. If the possible reliability values are $r_1 \ldots r_m$, the probability of any given intuition occurring, given P, is the probability of it occurring and r_1 obtaining, or it occurring and r_2 obtaining, or it occurring and r_3 obtaining, and so forth. Formally,

$$6.\quad \Pr\left(I_i \,\middle|\, P\&[I_1 \ldots I_{i-1}]\right) \approx \sum_{j=1}^{j=m} \Pr(r_j|P\&[I_1 \ldots I_{i-1}])\Pr(I_i|P \,\&\, r_j\&[I_1 \ldots I_{i-1}])$$

It should be noted that reliability can take any value from 0 to 1, so there are uncountable possible reliability values and they cannot actually be listed as a series $r_1 \dots r_m$. To be really precise, equation 6 should be an integral, rather than a summation. However, since in practice we want a numeric value for $\Pr(I_i|P\&I_1\dots I_{i\text{-}1})$, we must use some method for approximating the value of this integral. Treating r as a discrete variable and summing the results for each possible value (as in equation 6) is one such method. To generate the values discussed in the main text, I treated r as a discrete variable that could take one of 10,000 values between 0 and 1 (this required a computer to do the math for me). This should give us a very close approximation of the value of $\Pr(I_i|P\&I_1\dots I_{i\text{-}1})$.

We now have two terms to calculate: $\Pr(I_i|P\&r_j\&I_1\dots I_{i\text{-}1})$ and $\Pr(r_j|P\&I_1\dots I_{i\text{-}1})$. We can simplify $\Pr(I_i|P\&r_j\&I_1\dots I_{i\text{-}1})$. We know the content of the intuition I_i, and are given a reliability value and the truth of the relevant P. With these givens, $I_1\dots I_{i\text{-}1}$ make no difference to the probability of I_i. So

$$7. \quad \Pr\big(I_i \,\big|\, P \,\&\, r_j \& [I_1 \dots I_{i-1}]\big) = \Pr\big(I_i \,\big|\, P \,\&\, r_j\big)$$

We'll discuss how to calculate this at the end of the section (following equation 19).

How do we calculate $\Pr(r_j|P\&I_1\dots I_{i\text{-}1})$? By Bayes' theorem:

$$8. \quad Pr(\, r_j \mid P \,\&\, I_1 \dots I_{i-1}) \;=\; \frac{Pr(r_j)\; Pr(P\&I_1 \dots I_{i-1} \mid r_j)}{Pr(P\&I_1 \dots I_{i-1})}$$

$\Pr(r_j)$ is the prior probability of r_j, and we'll assume it is known and rational. (In my calculations in the main text, I assumed that the prior probabilities of reliabilities are normally distributed, and used a programming library by L'Ecuyer (2012) to calculate these priors) By the multiplication rule:

$$9. \quad Pr(P\&I_1 \dots I_{i-1} \mid r_j) = \Pr\,(P|r_j) \Pr\,(I_1 \dots I_{i-1}|P\&r_j)$$

Being given r_j doesn't by itself affect $\Pr(P)$, so $\Pr(P|r_j) = \Pr(P)$. Substituting this into equation 9, and equation 9 into equation 8, we get:

$$10.\ Pr(\,r_j \mid P \,\&\, I_1 \ldots I_{i-1}) \;=\; \frac{Pr(r_j)\,\Pr\,(P)\Pr\,(I_1 \ldots I_{i-1}|P\&r_j)}{Pr(P\&I_1 \ldots I_{i-1})}$$

For a moment, I want to focus on the denominator here. By the multiplication rule:

$$11.\ Pr(P\&I_1 \ldots I_{i-1}) = \Pr(P)\,\Pr\,(I_1 \ldots I_{i-1}|P)$$

We can substitute this into equation 10, and cancel the Pr(P) in the numerator and denominator, getting:

$$12.\ Pr(\,r_j \mid P \,\&\, I_1 \ldots I_{i-1}) \;=\; \frac{Pr(r_j)\,\Pr\,(I_1 \ldots I_{i-1}|P\&r_j)}{\Pr\,(I_1 \ldots I_{i-1}|P)}$$

Some of these terms will cancel out with other terms from elsewhere in our equation. So, rather than simplifying or expanding equation 12 further, let's see how what we've done fits into what we started out trying to do.

Remember, everything we have done so far has been to calculate Pr($I_1 \ldots I_n|P$), which is a term in the numerator of Bayes' theorem from equation 4. As we saw in equation 5, Pr($I_1 \ldots I_n|P$) is equal to the product of the probability of each intuition in $I_1 \ldots I_n$, given P (since these intuitions are dependent on each other, each must be calculated given all the prior intuitions). If we have n intuitions, we can also think of think of this as the product of the probabilities of intuitions 1 through n-1, times the probability of intuition n:

$$13.\ \Pr(I_1 \ldots I_n|P) = \Pr(I_1 \ldots I_{n-1}|P)\,\Pr\,(I_n|P\&\,I_1 \ldots I_{n-1})$$

This will allow us to cancel some terms, as we'll see in a moment. For now, let's simplify equation 13. Equation 6 shows us how to calculate the probability of a single intuition, and we can substitute it for Pr($I_n|P\&I_1 \ldots I_{n-1}$); equation 7 shows us how to simplify one of the terms in equation 6, giving us:

$$14.\ \Pr(I_1 \ldots I_n|P) \approx \Pr(I_1 \ldots I_{n-1}|P)\sum_{j=1}^{j=m}\left[\Pr\,(r_j|P\&I_1 \ldots I_{n-1})\Pr\,(I_n|P\,\&\,r_j)\right]$$

Equation 12 tells us how to expand Pr($r_j|P\&I_1 \ldots I_{n-1}$); substituting that into equation 14, we get:

$$15.\ \Pr(I_1 \ldots I_n|P) \approx \Pr(I_1 \ldots I_{n-1}|P) \sum_{j=1}^{j=m} \left[\frac{Pr(r_j)\ \Pr\ (I_1 \ldots I_{n-1}|P\&r_j)\Pr\ (I_n|P\ \&\ r_j)}{Pr(I_1 \ldots I_{n-1}|P)}\right]$$

Let's focus on the summation in equation 15. It sums the value of the given fraction for all possible *r* values. However, since the probability term in the denominator does not take an *r* value as given, it does not change as the *r* values considered vary. So every value in this summation has $\Pr(I_1 \ldots I_{n-1}|P)$ as a common denominator. Thus, rather than considering this as the summation of values of a fraction, we can consider it as a fraction with a summation in its numerator:

$$16.\ \sum_{j=1}^{j=m} \frac{Pr(r_j)\ \Pr\ (I_1 \ldots I_{n-1}|P\&r_j)\Pr\ (I_n|P\ \&\ r_j)}{Pr(I_1 \ldots I_{n-1}|P)} = \frac{\sum_{j=1}^{j=m}[Pr(r_j)\ \Pr(I_1 \ldots I_{n-1}\,|\,P\&r_j)\Pr\ (I_n|P\ \&\ r_j)]}{Pr(I_1 \ldots I_{n-1}|P)}$$

Substituting into equation 15, we get:

$$17.\ \Pr(I_1 \ldots I_n|P) \approx \Pr(I_1 \ldots I_{n-1}|P) \frac{\sum_{j=1}^{j=m}[Pr(r_j)\ \Pr(I_1 \ldots I_{n-1}\,|\,P\&r_j)\Pr\ (I_n|P\ \&\ r_j)]}{Pr(I_1 \ldots I_{n-1}|P)}$$

The denominator of the fraction cancels with the previous term in the equation, leaving us with:

$$18.\ \Pr(I_1 \ldots I_n|P) \approx \sum_{j=1}^{j=m}\left[Pr(r_j)\ \Pr\ (I_1 \ldots I_{n-1}|P\&r_j)\Pr\ (I_n|P\ \&\ r_j)\right]$$

This can be slightly simplified: the probability of $I_1 \ldots I_{n-1}$ (given *P* and some *r*) times the probability of I_n (given *P* and the same *r*) is just the probability of $I_1 \ldots I_n$ (given *P* and that *r*):

$$19.\ \Pr(I_1 \ldots I_n|P) \approx \sum_{j=1}^{j=m}\left[Pr(r_j)\ \Pr\ (I_1 \ldots I_n|P\&r_j)\right]$$

To calculate this, we need to know how to calculate $\Pr(I_1 \ldots I_n)$ given the truth of *P* and some reliability value. This is now relatively easy, because we can now act as if the intuitions are independent of each other. When implementing the narrowing down approach by itself, we are assuming that there is no causal dependence between intuitions. The only dependence relation we are concerned with is informational dependence – we learn something about how reliable intuitions are by observing what pattern of intuitions we find on this topic. If we are *given* a reliability value, however, then this sort of dependence is no longer a factor. Thus, given *P* and some reliability value r_j, we can calculate $\Pr(I_1 \ldots I_{i-1})$ by simply multiplying the probability of each intuition in the set together. The probability of each intuition depends on its content. Given *P*, intuitions that *P* are correct, and so have a probability of occurring equal to the given reliability value r_j. Intuitions that ~*P* have a probability of $(1\text{-}r_j)$.

So, if there are *f* intuitions for *P* and *a* intuitions that ~*P*:[22]

$$20.\ \Pr\left(I_1 \ldots I_{i-1} \,\middle|\, P\&r_j\right) = r_j^f (1-r_j)^a$$

Substituting into equation 19:

$$21.\ \Pr(I_1 \ldots I_n|P) \approx \sum_{j=1}^{j=m}\left[Pr(r_j)\, r_j^f (1-r_j)^a\right]$$

If we look back at equation 4, we see that we now only have to calculate the denominator of Bayes' theorem. The term we need to calculate is $\Pr(I_1 \ldots I_n|{\sim}P)$. Almost everything I have just said applies to calculating this term; the only difference is how we calculate the probability of each intuition given some *r* value and ~*P*. For all the intuitions that *P*, the probability of them occurring given some *r* value and ~*P* is 1-*r*; for intuitions that ~*P*, their probability given some *r* and ~*P* is *r*. So

$$22.\ \Pr(I_1 \ldots I_n|{\sim}P) \approx \sum_{j=1}^{j=m}\left[Pr(r_j)\,(1-r_j)^f r_j^a\right]$$

Plugging all of this in to Equation 4, we get:

$$23.\ \Pr(P|I_1 \ldots I_n) \approx \frac{\Pr(P)\sum_{j=1}^{j=m}\left[Pr(r_j)\, r_j^f (1-r_j)^a\right]}{\Pr(P)\sum_{j=1}^{j=m}\left[Pr(r_j)\, r_j^f (1-r_j)^a\right] + \Pr({\sim}P)\sum_{j=1}^{j=m}\left[Pr(r_j)\,(1-r_j)^f r_j^a\right]}$$

Appendix C: Causal dependence

In this section, we'll discuss how to deal with possible causal dependence, assuming we

[22] Please note: the following equation should also consider the number of combinations of ways to get *f* and *a* intuitions that *P* and ~*P*. However, this will cancel out in later equations, so I do not consider it here.

have a given reliability value for our set of intuitions.

The calculations for the symmetric and asymmetric dependence approaches are derived almost exactly as above. Instead of narrowing down a reliability value, we consider a range of degrees of dependence; we'll use d to stand for these. This doesn't make a big difference to the derivations: for most of the discussion in Appendix B, the actual meaning of the r term – that it represents a reliability value – didn't make a difference to the algebra being done. We can simply swap d for r for most of the derivation without issue, as long as we then add a reliability value as a given.

Where things diverge from the discussion in Appendix B substantially is in the final steps in the derivation: equation 20 through 23. This is where we finally calculate the probability of an intuition, or of a series of intuitions. Modifying equation 19 so that it is relevant to causal dependence, we get:

$$24.\ \Pr(I_1 \ldots I_n | P\&r) \approx \sum_{j=1}^{j=m} \left[Pr(d_j) \Pr\left(I_1 \ldots I_n | P\&r\&d_j\right) \right]$$

Note that r is a given on the left side of the equation. However, r need not be a given in Pr(d), as Pr($d|r$) = Pr(d).

The calculation of Pr($I_1 \ldots I_n | P\&r\&d$) varies depending on what notion of dependence we use. Consider first symmetric dependence. For intuitions that P, the probability that they obtain is equal to the probability that: one intuits correctly and is not influenced by previous intuitions or that one has been influenced by previous intuitions that P but not by previous intuitions that ~P. The probability of not being influenced by n intuitions is $(1\text{-}d)^n$, and the probability of being influenced by at least one out of n intuitions is $(1\text{-}(1\text{-}d)^n)$. Thus, the probability of intuiting that P, given r, P, d, and previous intuitions, is

$$25.\ \Pr\left(intuiting\ that\ P \,\middle|\, P\&d\&r\&previous\ intuitions\right) = r(1-d)^{all\ prior\ intuitions} + (1-(1-$$

$$d)^{prior\ intuitions\ that\ P})(1 - d)^{prior\ intuitions\ that\ {\sim}P}$$

The probability of intuiting that ~*P*, given *r*, *P*, *d*, and previous intuitions is:

$$26.\ \Pr\left(intuiting\ that\ {\sim}P \,\middle|\, P\&d\&r\&previous\ intuitions\right) = (1 - r)(1 - d)^{all\ prior\ intuitions} + ((1 - d)^{prior\ intuitions\ that\ P})(1 - (1 - d)^{prior\ intuitions\ that\ {\sim}P})$$

The order we consider intuitions in doesn't matter, so to simplify our calculations, we can consider all intuitions that *P* first, and then all intuitions that ~*P*. For intuitions that *P*, this allows us to ignore the possibility of influence by intuitions that ~*P*.

To calculate $\Pr(I_1 \ldots I_n | P\&r\&d_j)$ let *f* be the number of intuitions that *P* and *a* the number that ~*P*:

$$27.\ \Pr\left(I_1 \ldots I_n \,\middle|\, P\&r\&d_j\right) = \prod_{i=1}^{i=f}\left[r\left(1 - d_j\right)^{i-1} + (1 - (1 - d)^{i-1})\right]\prod_{i=1}^{i=a}\left[(1 - r)\left(1 - d_j\right)^{f+i-1} + (1 - d)^{f}(1 - (1 - d)^{i-1})\right]$$

This can be substituted into equation 24.

The calculation for $\Pr(I_1 \ldots I_n | {\sim}P\&r)$ can be derived from what I've just said, with minor changes.

For asymmetric dependence, one needs to make changes to equations 25 and 26. On this approach, an intuition can only be influenced by the majority view. The probability that an intuition was or was not influenced is simply based on the degree of asymmetric dependence *d* and is independent of the size of the majority. Given *P*, for intuitions that agree with the majority, their probability given some degree of asymmetric dependence is the chance that they were correct and were not influenced, or that they were influenced by the majority. For intuitions that disagree with the majority (given *P*), their probability is the probability that they were *in*correct and were not influenced.

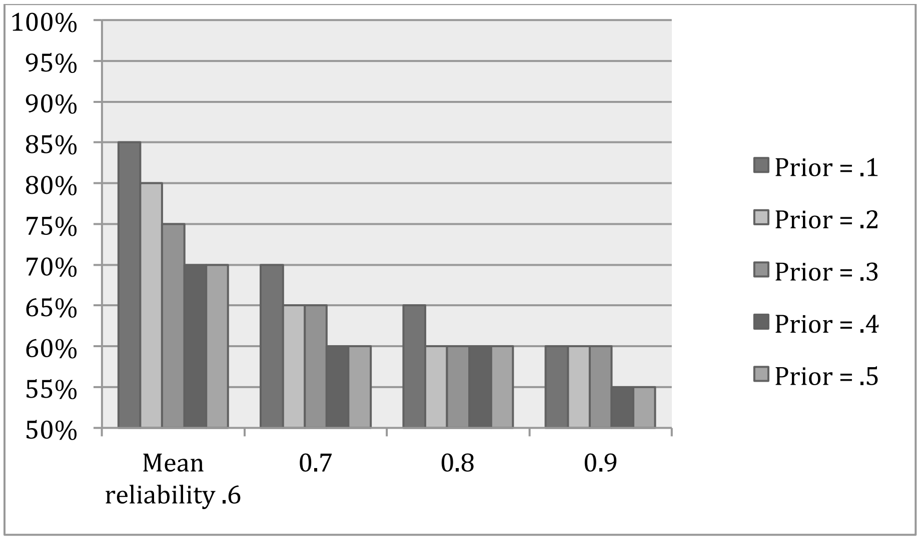

Fig. 1 What percent of 20 intuitors must agree that *p* for the probability that *p* to be greater than .9? Calculated using the narrowing down approach, with a standard deviation of reliabilities of .2 and the listed prior probabilities that *p*

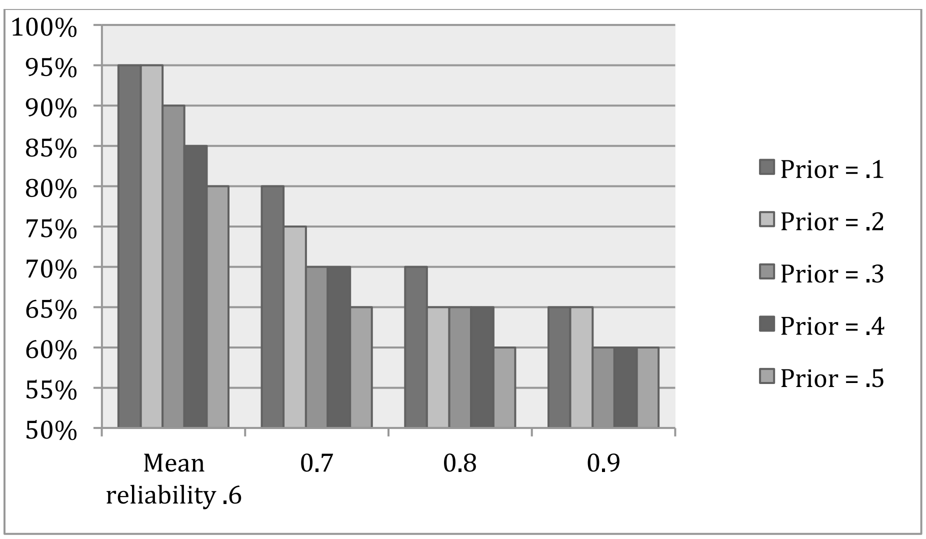

Fig. 2 What percent of 20 intuitors must agree that *p* for the probability that *p* to be greater than .9? Calculated using the narrowing down approach, with a standard deviation of reliabilities of .5

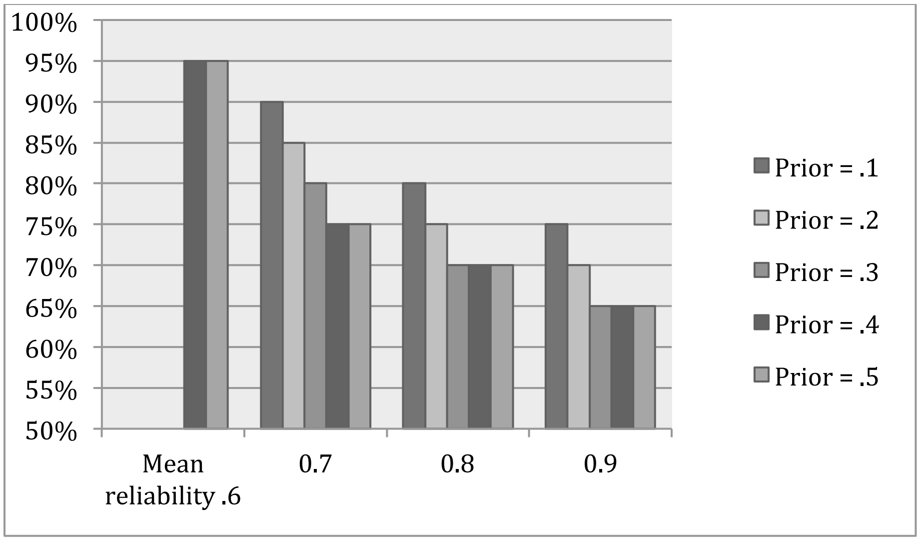

Fig. 3 What percent of 20 intuitors must agree that *p* for the probability that *p* to be greater than .9? Calculated using the narrowing down approach, with a standard deviation of reliabilities of .7

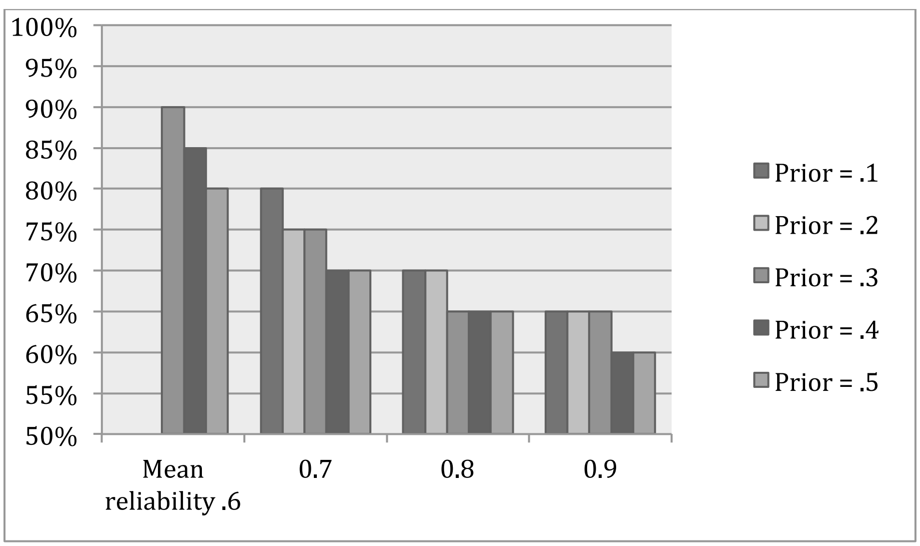

Fig. 4 What percent of 20 intuitors must agree that *p* for the probability that *p* to be greater than .9? Calculated assuming a known degree of symmetric dependence of .05

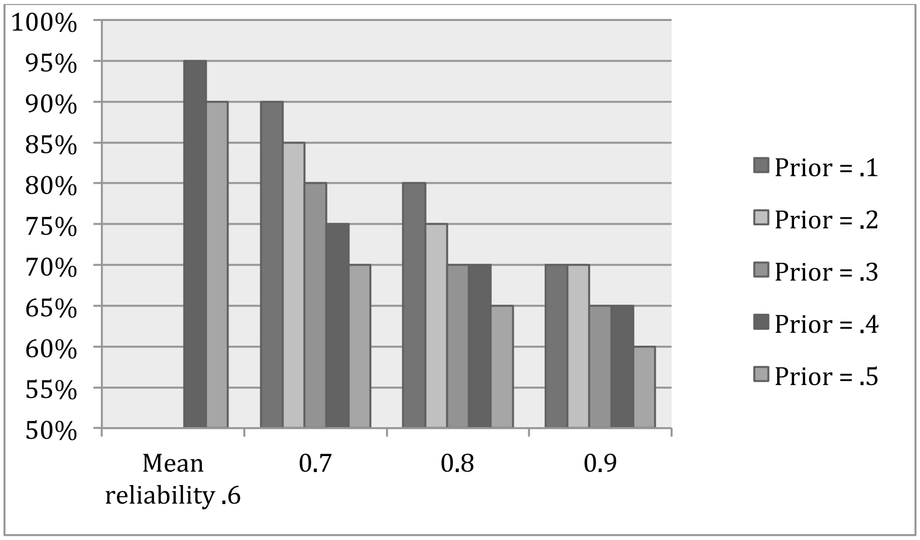

Fig. 5 What percent of 20 intuitors must agree that p for the probability that *p* to be greater than .9? Calculated assuming a known symmetric dependence of .07

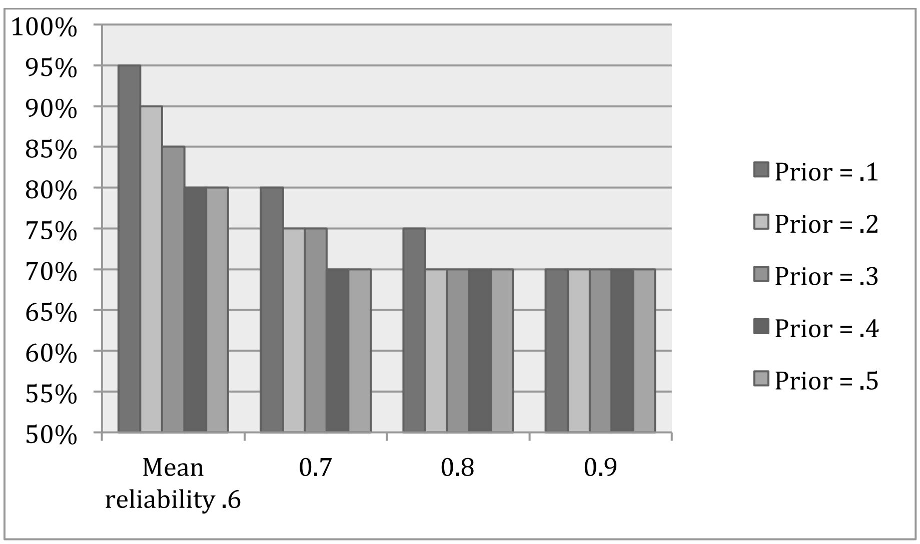

Fig. 6 What percent of 20 intuitors must agree that *p* for the probability that *p* to be greater than .9? Calculated assuming a known asymmetric dependence of .2

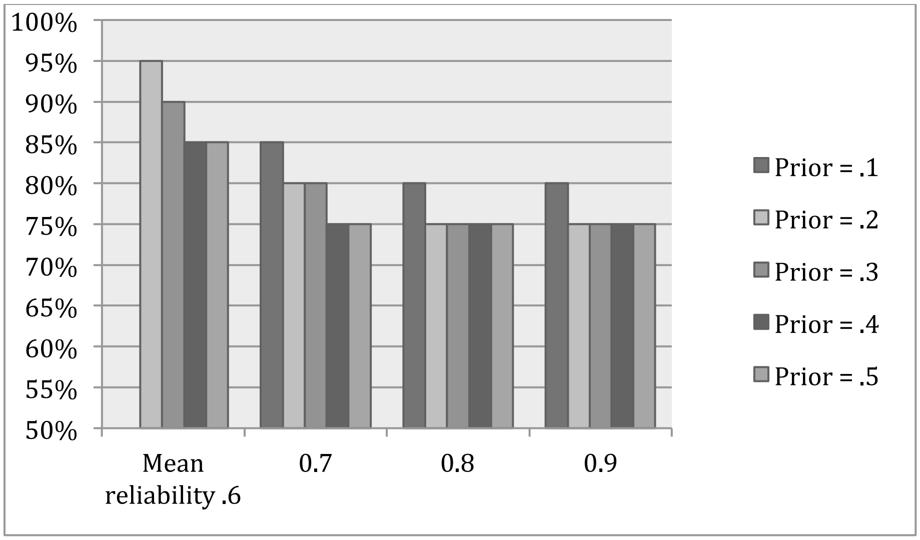

Fig. 7 What percent of 20 intuitors must agree that *p* for the probability that *p* to be greater than .9? Calculated assuming a known asymmetric dependence of .3

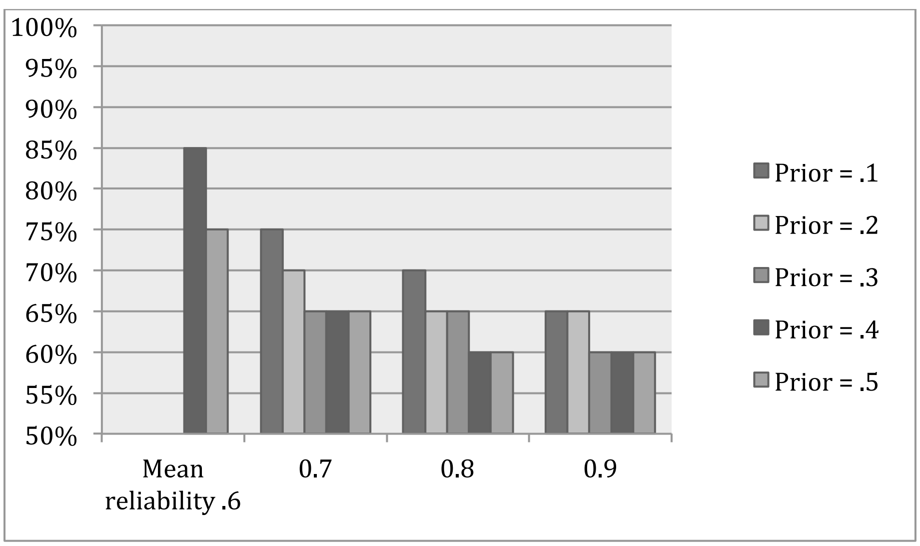

Fig. 8 What percent of 20 intuitors must agree that *p* for the probability that *p* to be greater than .9? Calculated using an unknown degree of symmetric dependence, with a population average degree of dependence of 0 and a standard deviation of .3

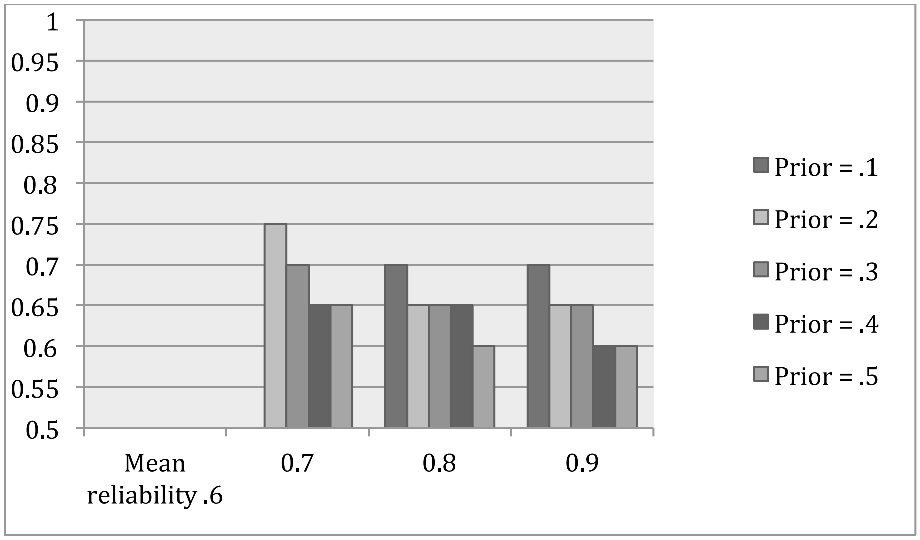

Fig. 9 What percent of 20 intuitors must agree that *p* for the probability that *p* to be greater than .9? Calculated using an unknown degree of symmetric dependence, with a population average degree of dependence of .1 and a standard deviation of .3

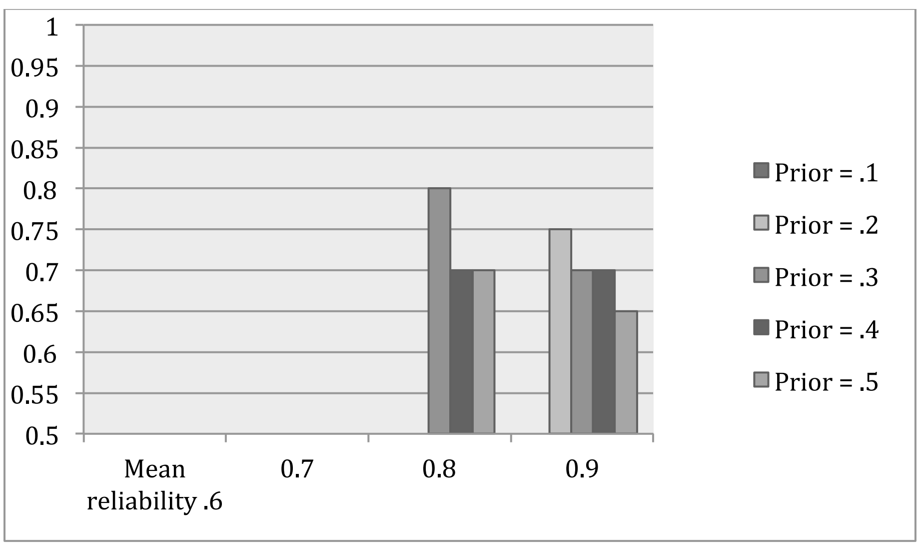

Fig. 10 What percent of 20 intuitors must agree that *p* for the probability that *p* to be greater than .9? Calculated using an unknown degree of symmetric dependence, with a population average degree of dependence of .3 and a standard deviation of .3

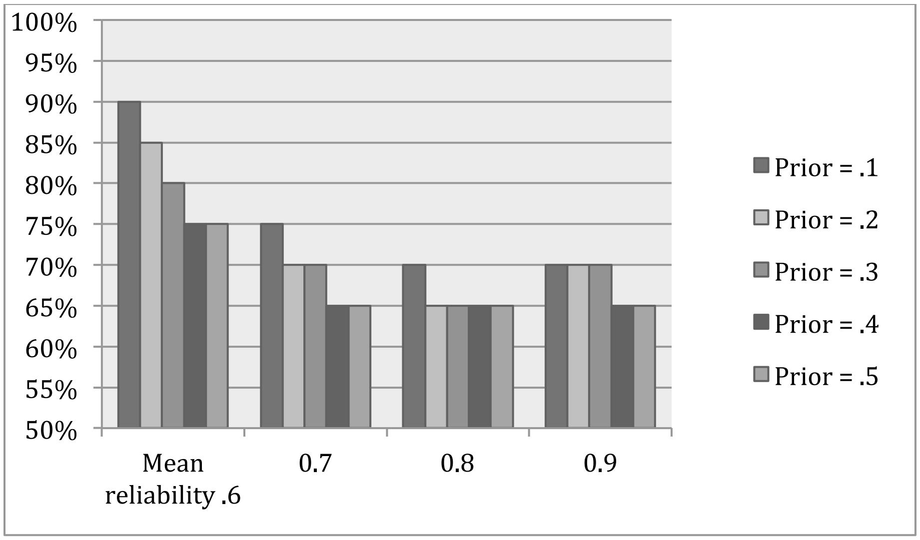

Fig. 11 What percent of 20 intuitors must agree that *p* for the probability that *p* to be greater than .9? Calculated using an unknown degree of asymmetric dependence, with a population average degree of dependence of 0 and a standard deviation of .5

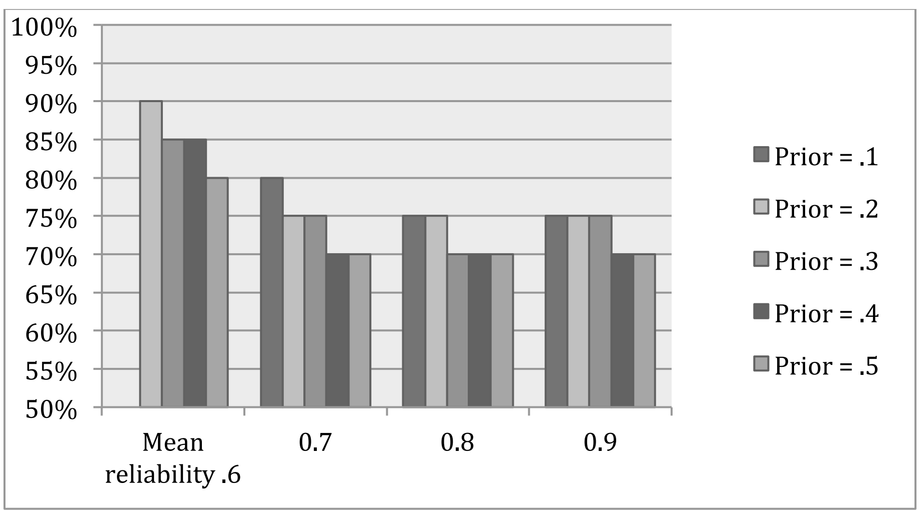

Fig. 12 What percent of 20 intuitors must agree that *p* for the probability that *p* to be greater than .9? Calculated using an unknown degree of asymmetric dependence, with a population average degree of dependence of .2 and a standard deviation of .3

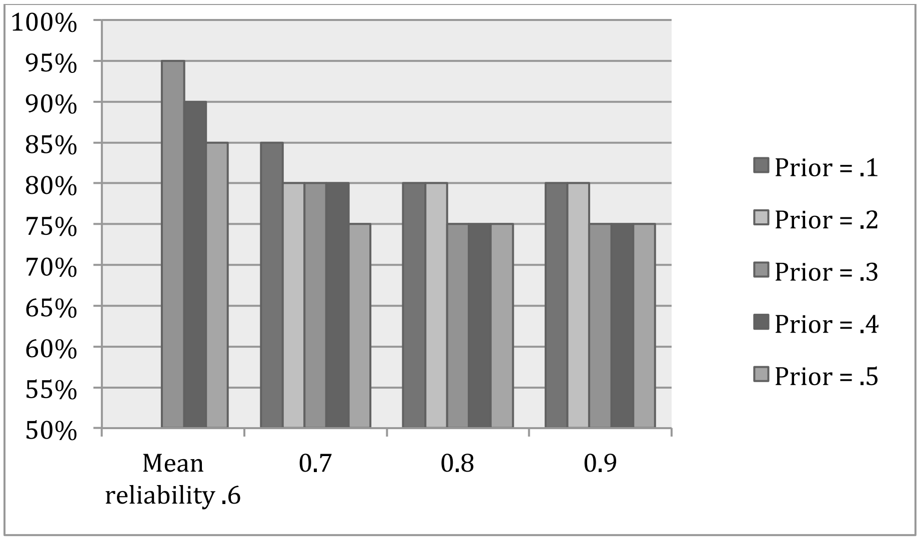

Fig. 13 What percent of 20 intuitors must agree that *p* for the probability that *p* to be greater than .9? Calculated using an unknown degree of asymmetric dependence, with a population average degree of dependence of .3 and a standard deviation of .3