And here, there seems to be nothing in the hypothesis about being hypnotized, drugged, etc. However, from my current point of view, there are some things that I could not believe unless I were crazy or subject to some kind of manipulation. That I am the Messiah is one of them. So, from my current point of view, the more hypothesis that I shall believe that I am the Messiah implies that I will be crazy, drugged, etc., and this does invalidate the conditional belief. To see this, just imagine the counterexample as described being offered to someone who now thinks it a genuine possibility that she is the Messiah (though she doesn't currently think she is).

This is illustrated by the famous case of Ulysses and the Sirens, and by the case of the Nineteenth Century Russian discussed by Derek Parfit in Reasons and Persons (Oxford: Clarendon Press, 1984), p. 327-8.


--

Michael A. Bishop

The Autonomy of Social Epistemology

Social epistemology is autonomous. When applied to the same evidential situations, the principles of social rationality and the principles of individual rationality sometimes recommend inconsistent beliefs. If we stipulate that reasoning rationally from justified beliefs to a true belief is normally sufficient for knowledge, the autonomy thesis implies that some knowledge is essentially social. When the principles of social and individual rationality are applied to justified evidence and recommended inconsistent beliefs and the belief endorsed by social rationality is true, then that true belief would be an instance of social knowledge but not individual knowledge.

This paper is organized into six sections. Section 1 describes a widely held conception of individual epistemology and the individualistic notion of rationality that drops out of it. A principle of rationality that has helped to shape judgments made within the contexts of social institutions is introduced in section 2. Sections 3-5 argue that this principle of social rationality is practically inconsistent with individual rationality. The final section concludes with a brief discussion of some objections and implications.

1. Epistemological Individualism

Epistemological individualism consists of two widely accepted theses. (1) Methodological individualism holds that knowledge of the causal and epistemic norms is somehow available to individuals and we can discover that knowledge via an appropriate process of self-exploration (Goldman and Pust 1994, Bostock 1987, Boudreau 1998). (2) Metaphysical individualism holds that only individual cognizers can have knowledge (for a holist, see Schmidt 1994). Methodological individualism is presupposed in the method of contemporary analytic epistemology, where successful theories are supposed to capture our epistemic intuitions about well-developed cases. Epistemic intuitions are usually taken to be non-discursive, though perhaps considered, judgments about the epistemic properties of some cognitive item (Cohen 1981, Boutilier 1987, Bonjour 1998, Pust 2000). Paradigm examples of epistemic intuitions are our judgments that subjects in Gethers cases do not have knowledge. It is useful to think about methodological individualism in terms of a performance-competence distinction. We tacitly possess, or in some way have access to, a reasoning competence: a set of principles about how we ought to reason. These principles define individual rationality. As a result, ordinary human reasoning... cannot be held to be faultily programmed: it sets its own standards (Cohen 1981, 317). How, then, are we to explain the fact that people sometimes reason poorly? Reasoning involves bringing to bear not only our reasoning competence but also our background beliefs and various ancillary cognitive systems, such as memory, attention and perception. Performance errors occur when we reason poorly as a result of a failure of our background beliefs or of our ancillary systems (e.g., see Cohen 1981, 323-326; Goldman and Pust 1994, 182-183). 2 The individualist view under consideration takes all reasoning errors to be performance errors. There can be no systematic flaws in our reasoning competence, since that competence defines individual rationality. 2

Metaphysical individualism, which holds that only individual persons can have knowledge, naturally coheres with a consequentialist conception of social epistemology: the epistemic quality of a social practice is a function solely of its epistemic consequences for individuals (Solomon 2000). On this view, social epistemology is a purely applied field. As with many applied fields, it can be excruciatingly...
difficult. Invoking novel and elaborate conceptual and mathematical machinery. But theoretically, it is completely derivative on the epistemology of individual knowers. This is many challenging and important views about the nature of social epistemology. But it would hardly be interesting to show that social epistemology is autonomous if I began by assuming (say) that knowledge is socially constructed or essentially distributed across social groups. If I can show that social epistemology is autonomous given these individualist assumptions, then it is autonomous, period.

2. The Flat Maximum Principle

For over 50 years, psychologists have developed rules for making judgments about matters of practical importance. These Statistical Prediction Rules (SPR) outperform human experts in making some important judgments. For example, SPs are more reliable than forensic psychologists in predicting the likelihood of violence and expert criminologists in predicting criminal recidivism (Loft and Zinik 1988; Carroll et al., 1988). They outperform clinicians in recently diagnosing progressive brain dysfunctions, and predicting the presence, location and cause of brain damage (Sell and Filisio, 1988; Melting, 1983). Paul Meehan’s evaluation of this literature is apt: “There is no controversy in social science which shows such a large body of qualitatively diverse studies coming out so uniformly in the same direction on this one. When you are pushing (scores of) investigations, predicting everything from the outcomes of football games to the diagnosis of liver disease

and when you can hardly come up with a half dozen studies showing even a weak tendency in favor of the clinician, it is time to draw a practical conclusion” (1980, 373–3).

Consider a common SPR, a linear model for making a judgment about a target property, $P$, on the basis of predictor cues $c_1, c_2, \ldots, c_n$,

$$P = w_{c_1} + w_{c_2} + \ldots + w_{c_n}$$

where $c_i$ is the value for the $i$th predictor cue, and $w_i$ is the weight assigned to the $i$th cue. Raw values are not plugged into the variables. They are first normalized (they are z-scored) and then positively correlated with the target property. For example, a linear model that predicts the quality of the vintages for a Red Bordeaux wine might reflect the fact that Bordeaux grapes thrive best in hot dry summers (Auberle, et al., 1992). So, if the average rainfall in a given summer is one standard deviation above (below) the mean, the value for the rainfall cue might be $-1$ (or 1).

Given predictor cues that are normalized and positively correlated with the target property, we can distinguish three different kinds of linear model according to how their weights are chosen (Dawes, 1982): "1. A proper model, weights are chosen so as to best fit the data. Imagine a graph in which the predictor cues are plotted against the target property. The proper linear model draws a straight line that best fits those points.

2. In a unit model, all the weights are equal to 1. Each cue has an equal "say" in the final prediction. This is an improper linear model because its weights are not chosen so as to best fit the data.

3. In a random model, the weights are assigned randomly (but between 0 and 1)."

A proper model will have many correlate improper models. Many of these will differ from the proper model only in terms of their weights; but some will drop those cues used by the proper model that are weakly correlated with the target. Most people are not particularly surprised that these models outperform human experts. Such models make judgments that best fit the available evidence — something experts cannot be expected to do on the fly. Even so, the reaction of witnessing experts to successful proper linear models has been "somewhat between violent and hysterical" (Passol, 1990). Most people are surprised, however, at the obvious accuracy of improper linear models. In a classic essay, Daves and Corrigan (1974) found that for problems of social prediction, random and unit models were about as reliable as proper models, and therefore more reliable than human experts. Daves reports that this result "when published engendered two responses. First, many people didn't believe them — until they tested out random and unit models on their own data sets. Then, other people showed that the results were trivial. I concur with those proclaiming the results trivial, but not realizing the triviality at the time. I luckily produced a "citation classic"..." (1988, 209, n. 17). The result is trivial because of a finding in statistics called the flat maximum principle (Efron and Hoggart, 1975).

The flat maximum principle (FMP) says that for certain sorts of problem, the weights (between 0 and 1) assigned to the predictor cues of a linear model do not affect the model's long-term accuracy as long as:

1. The judgment problem is difficult. What is being predicted is messy — the data points are spread out enough so that no linear model will be especially reliable.

2. The evidential cues are reasonably predictive. The evidential cues are somewhat redundant. As a result, adding extra cues to a successful model might add very little (if any) accuracy.

Problems of social judgment tend to share these features. So for such problems, the FMP implies that improper models will be about as reliable as proper models.

It is easy to underestimate the counterintuitive nature of the FMP. To fully understand its oddness, consider figure 1, which compares a Proper Model (PM) and an Improper Model (IM) applied to a binary (yes/no) problem. The FMP implies that PM and IM will be about equally reliable overall. But it doesn't follow that they will always agree. In fact, when proper and improper models are applied to the same set of problems, it is typical for one to be more reliable than the other.

This is not inconsistent with the FMP for two reasons. First, two models might be equally reliable in the long run but not equally reliable on a short run of problems. Second, and less obviously, the FMP does not imply that a particular proper linear model will be exactly as reliable in the long run as all of its correlate improper models. The proper model and its correlate improper models will typically have a range of different long-term reliability scores. What the FMP implies is that we should not expect proper models to be significantly more reliable on average than their correlate improper models.
with social rationality are not the result of performance errors (i.e., ignorance, false beliefs, or poor reasoning). Many readers will object because they believe it is irrational for us not to accept and reason in accordance with the FWP. I agree (see section 6, as well as Bishop and Toot (2005). But this is not inconsistent with the purely descriptive claim I'm arguing for here. My ultimate view is that reasoning rationally about selective defense cases will require a significant revision in the way we reasoning competence works (see also Bishop, 2000). To make this case, I will document people's resistance to using improper models and then argue that the best explanation for this resistance is that the FWP is inconsistent with the principles that make up our reasoning competence.

In selective defense studies, subjects are given a successful SPR and allowed to override its judgment. The consistent result of these studies is that subjects are still outperformed by the SPR, even when they're told that it has been shown to be more reliable than experts (Sawyer, 1995, Leli and Filkow 1984). The Goldberg Rule is a successful unit-weight model that predicts whether a psychiatric patient is neurotic or psychotic on the basis of a MMPI (Minnesota Multiphasic Personality Inventory) profile. It diagnoses a patient as neurotic if $P_1 > P_3 > P_5 > P_4 > P_2$, where $P_1$ is a validity scale and the rest are personality scales of the MMPI. Otherwise it diagnoses the patient as psychotic. In particular, interesting selective defense study, one group of subjects was given the Goldberg Rule and a second group of subjects was given only the numerical values of the formula for each profile and the optimum cutting score. Both groups were told that 'this formula would achieve approximately 70% accuracy and that it would be more accurate for extreme values than for values close to the cutting score' (Goldberg, 1968, 403). The results are represented in figure 2. Those who were not given the formula but only its results (A) "increased their accuracy to a bit below 70% correct"—still not as accurate as the formula by all itself. For those who were given the formula but had to take it themselves (B), there was a small increase in accuracy. But this improvement "gradually wore away over time" until it disappeared altogether (Goldberg, 1968, 493). Group B ended up reasoning about as accurately as if they had never been given the SPR. Let's distinguish two aspects of this result.

1. Resistance. We detect from SPRs even when we know they are more reliable than we are in the long run.

2. Suboptimality. Experts with SPRs are less reliable in the long run than SPRs. It is useful to keep these phenomena distinct. While some resistance is obviously necessary for suboptimality, it is not sufficient. We might resist and improve upon the SPRs reliability, which is what experts certainly thought they were doing. This is different because it would have to appeal to our ignorance and poor reasoning. In particular, subjects are ignorant of the APR, the SPR results, and they perhaps don't fully believe that the SPR is more reliable than they are, and when it comes to complex issues, we don't reason as reliably as we think we do. Although R.D. Toot and I have defended an explanation along these lines (Bishop and Toot, 2005, 37-53), there are two prima facie problems with it:

a. It doesn't account for the difference between groups A and B in resistance and suboptimality. Are we supposed to believe that the members of group A know more about the FWP, SPRs or the Goldberg Rule than the members of group B, or that group A consists of better reasoners than group B? Surely not.

b. It doesn't account for the delay in group B's resistance and suboptimality reasoning.

While ignorance and poor reasoning are part of the explanation for the selective defense results, they can't be the full story.

A more complete explanation begins with the assumption that people are quite good, but not perfect, at tracking episodic support (i.e., when the evidence supports one belief better than another). In figure 3, Expert has only the IV, not the PM. Since the PM is constructed so as to be best viewed as evidence, we'll suppose that it always delivers the belief that is best supported by the evidence. (This assumption is usually, but not always true, see section 4.)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>55%</td>
<td>15%</td>
<td>15%</td>
<td>Correct</td>
</tr>
<tr>
<td>Proper model (PM)</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>70%</td>
</tr>
<tr>
<td>Improve model (IM)</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>70%</td>
</tr>
<tr>
<td>Expert Performance</td>
<td>T</td>
<td>F</td>
<td>F</td>
<td>70%</td>
</tr>
</tbody>
</table>

Figure 3: Expert tracks episodic support 95% of the time.

Let's assume that Expert tracks episodic support 95% of the time. As a result, she detects from IV 5% of the time when it agrees with PM (columns 1 and 2) and 95% of the time when it disagrees with PM (columns 3 and 4). With this pattern of detection, Expert loses 2% accuracy (19.75% loss in column 1, .75% gain column 2, and no net change in columns 3 and 4 combined). But Expert comes to the belief that is best supported by the evidence (i.e., agrees with the PM) 95% of the time (as opposed to only 70% for the IV). So Expert reasons less reliably than the IV even though she is much more likely to adopt the belief that is best supported by the evidence.

An essential part of the correct explanation of the selective defense results must appeal to the fact that we track episodic support better than improper models. We typically detect from a SPR the basis of a second-order belief to the effect that the SPR recommends a belief that is not best supported by the evidence. If figure 3 represents the reasoner's situation in a selective defense study, note two points about these secondorder beliefs:

1. They are usually true (e.g., in figure 3, they're true 95% of the time).
2. They are typically supported by defeasible (and often valid) arguments.

Given (1) and (2) and almost any reasonable theory of justification, when we detect from a SPR, we are justified in believing that the SPR recommends a belief that is not best supported by the evidence. When this second-order belief is true (in columns 3 and 4) we can draw an even stronger conclusion. We know, sometimes on the basis of a deductively sound argument, that the SPR-supported belief is not best supported by the evidence. No wonder we defect.

But why do we reason suboptimally? The answer is that high sensitivity brings false alarms. We are like the overly sensitive fire alarm that reliably detects fires but also smokes. By being so sensitive to the evidence, we usually defect when the evidence doesn't support the belief recommended by the SPR (columns 3 and 4), but we also occasionally defect when the evidence does support the belief recommended by the SPR (columns 1 and 2). It is reasonable to suppose that ignorance and poor reasoning play a significant role in cases of inappropriate detection. The problem is that detecting when we should (columns 3 and 4) does not change our reliability whereas detecting when we should (columns 1 and 2) does not change our reliability.

This explanation has the resources to account for the pattern of detection in figure 2. Group B detects more often than group A simply because members of B more often come to the relevant conclusion (i.e., that the evidence doesn't support the SPR recommended belief). And this is because group B's use of the SPRs puts them in a better position to know when the SPR is ignoring or otherwise misweighing evidence. Consider a schematic example. Group B members (like group A members) know which MMPI scores the SPR ignores (or misweights); so they are in a better position to detect cases in which those ignored (or misweighted) scores indicate a diagnosis that contradicts the SPRs diagnosis. When this happens, members of group B detect because they recognize that abiding by the SPR requires them to adopt a belief that is not supported by the evidence.
4. Everybody's Objection

In my experience, almost everybody responds to these studies as follows: “I'll accept the selective detection results and your explanation for them. But nothing here rules out our reasoning competence. Subjects reason suboptimally in the selective detection studies because they fall victim to a performance error. They fail to give sufficient weight to an important piece of evidence, namely, that the QR is in the long run more reliable than they are. If they were to weigh this evidence properly, they wouldn't reason less reliably than the QR.” In thinking clearly about Everybody’s Objection, it's important to keep in mind that the issue is not whether subjects in the selective detection studies are somehow irrational. It's on record arguing that they are. The issue is whether their reasoning in the selective detection studies is consistent with the principles that constitute their reasoning competence or whether it is the result of a performance error (ignorance or a false belief or a failure of an ancillary cognitive system).

It is easy to badly asst: Everybody's Objection, confident that your reasoning competence would direct you to weigh the QR as reliably as the QR. But in constructing her epistemological theories, the individual does not rely primarily on these sorts of abstract judgments but on our epistemic intuitions about particular, well-defined cases (like the Getter examples). Everybody’s Objection needs to be tested against some concrete, albeit hypothetical, cases — preferably cases involving subjects about which you have some expertise. If you are a philosopher with some experience on hiring committees, you probably consider yourself a reasonably good judge of philosophical promise. So let’s suppose there are two SRPs for hiring philosophers that have been shown to be more reliable in the long run than experts are, you, etc. One is a proper model (PM) that considers all the evidence you normally consider in making hiring decisions. For the sake of convenience, let's suppose this involves four lines of evidence: letters of recommendation, Ph.D. program, writing sample, and interview. The second SRP is an improper model (IM) that weights three lines of evidence equally (so it is a unit model) and ignores the least predictive line of evidence — let’s say it ignores the interview evidence.

We'll consider four hypothetical selective detection cases to test your epistemic intuitions. In every case, we'll assume that you only have the IM available, and that you and the IM agree about who the two finalists are — Theresa (T) and Fabiola (F). Since you know the IM's results and the QR predicts, we'll suppose you have well-grounded confidence that figure 4 represents your situation reasonably accurately.

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>55%</td>
<td>55%</td>
<td>55%</td>
<td>55%</td>
</tr>
<tr>
<td>Perfect model (PM) F F F F 75%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improper model (IM) F F F 75%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4: The hiring cases

So in these four cases, you will have two epistemological considerations pulling against each other. On the one hand, considering only the evidence you have about the candidates, you will have a justified belief that the evidence supports rejecting the QR. On the other hand, the QR recommends a higher frequency of true beliefs in the long run than experts like you. How strong is the frequency consideration? In figure 4, you are inclined to reject 25% of the time (by column: 2.75% + 7.5% + 14.25% + 14.25%), and you will defect to the true belief 15% of the time and to the false belief 17% of the time. Not a huge difference! If you defect, you can expect to defect to the true belief almost 47% of the time. 10 Keep in mind, this is a long run frequency; it says nothing about whether you're right to defect in any particular case.

Case #1: You have applied the IW to the dossiers of T and F. This has involved grading the applicants’ letters, writing samples, and Ph.D. programs, normalizing those grades (according to simple functions), and then adding those transformed grades together. F has a slightly higher total grade than T, and so the IM recommends hiring F. You are fairly sure, however, that the PM would recommend hiring T, and that the entirety of the evidence supports hiring T. The reason is that the IM ignores interview information, T gave a terrific interview. However, just after F’s interview began, F verbally abused and then physically assaulted the interviewers.

If you defended Everybody’s Objection because you were confident that you would never defect from a successful QR, you were kidding yourself. The fact that one line of evidence (e.g., interview) is not strongly correlated with job performance is a fact about long run frequencies. On any particular occasion, that evidence might be highly relevant. But this will be insufficient enough that it won’t make a significant difference to the QR’s long term reliability.

Case #2: You have applied the IM to the dossiers of T and F. F has a higher total grade than T, and so the IM recommends hiring F. You are fairly sure that the PM would also recommend hiring F. However, you have extra evidence about F that is not usually considered in hiring. At the national conference, you observed F (but no other candidates) after hours. F was engaged in shocking and illegal behavior — walking down the street with a weapon and a bottle of whiskey, yelling, “I'm going to bag me some liberals tonight!”

The first two cases show that SRPs, including proper models, ignore relevant evidence. The ignored evidence is relevant in the sense that one could in principle improve on the SRPs long term reliability by taking it into account. Of course, cases #1 and #2 are extreme and unrealistic, but then so are many important examples that inform contemporary epistemology. More importantly, the staunchest defenders of SRPs recognize that SRPs ignore relevant evidence and that detection is sometimes justified. 11 A psychologist who developed a SRP for predicting recidivism told me of a striking example of legitimate detection: one should always correct his model if it doesn't predict a zero chance of recidivism for dead people. 12 At this point, Everybody’s Objection becomes: OK, my reasoning competence would direct me to sometimes defect from a successful QR. But I wouldn't insist as much as subjects do in selective detection tasks, because I would properly weigh the fact that the QR is more reliable in the long run than I am.” The problem, of course, is that instances of legitimate detection don't announce themselves with fanfare and confetti.

Case #3: You have applied the IM and F to the dossiers of T and F. Comparing the dossiers, T and F went to the same university (for the program quality is the same), T's writing sample is inferior to F's, but T's letters are superior to F's. But the difference in the quality of the writing samples is appreciably greater than the difference in the quality of the letters. Since the IM weighs the three pieces of evidence equally, it ranks F higher than T. But given your experience, you justifiably believe that letters are stronger evidence of future job performance than the writing sample; if one had to make a decision based on only one of these, one would choose the letters. Had the model given somewhat greater weight to the letters than to the writing sample, as some proper model would, you were convinced that it would have recommended hiring T.

You have two powerful arguments for detection:

1) The quality of the letters is more strongly correlated with job performance than the quality of the writing sample; so T's superior writing sample is less diagnostic of job performance than T's superior letters. The evidence best supports the belief that T has stronger credentials.

2) The improper model comes to the opposite conclusion because it does not weigh these lines of evidence appropriately. If it did, it would have come to the conclusion that T has stronger credentials than F.

Of course, you know that when you defect, your long term accuracy is slightly worse than the QR's (about 47% to 53%). But you also know there are legitimate cases for detection. When you defect, you know that T has stronger credentials than F (by 89% of the time). 13 So on just about anybody's view of justification, it would seem that you are justified in believing that T has stronger credentials than F (And if T really does
Michael A. Bishop

have stronger credentials, you know this.) In this epidemic situation, what do your epistemic intuitions—your reflective, non-discursive judgments—tell you about what you ought to believe? Let's consider one final case.

Case #4: You have applied the IM to the classes of T and F. The IM ranks F somewhat higher than T. But IM ignores interview evidence. As it happens, T had a much better interview than F. In fact, T sparkled, while F appeared almost incompetent, sometimes unable to answer questions coherently. You have no objection to the IM recommendation given only the evidence the IM considered. But you are certain that if the IM were amended so as to give even a slight degree of weight to the interview evidence, the large difference in the quality of their interviews would lead the model to recommend hiring T rather than F.

You have two powerful arguments for defecting from the IM:

(a) While interviews are only weakly correlated with job performance, in this case, T and F are sufficiently closely matched on other measures that the difference in their interviews Gray in the scale in T's favor. The evidence best supports the belief that T has the stronger credentials.

(b) The improper model comes to the opposite conclusion because it completely ignores interview evidence. If it weighed interview evidence appropriately, it would have come to the conclusion that T has stronger credentials than F.

Once again, you know that when you defect (your long term accuracy is slightly worse than the SIRS). But you also know that there are legitimate cases for defection. In this case, you are justified in accepting T as the stronger credentials than F. What do your epistemic intuitions—your reflective, non-discursive judgments—tell you about what you ought to believe?

Those who are inclined to defect in cases #3 or #4 might object that I have rigged the game by describing cases in which one has strong evidence in favor of defecting. If you are inclined to raise this objection, I have to wonder: What were you expecting? That expert clinicians with years of experience would defect from SIRSs arbitrarily or when they had clearly insufficient reason? Such assumptions are unacceptable.

In order to defend the final test in such cases, you would have to argue that the interview evidence is relevant and that it would level the playing field between T and F.

You might object that ignoring evidence you knew was relevant, the next time your department hires you would be less inclined to ignore evidence you were justified in believing was relevant.

Consider again the resistance finding represented in figure 6. Group B doesn't resist much at first; it's only later, after they had experience with the SIR that they really resist ignoring relevant evidence. So being able to ignore irrelevant evidence once or even on a few occasions, isn't enough. You must remain determined to ignore evidence you know is relevant, even after you've been Poverty by doing so. And the longer you keep it up, the more you can expect to be burned by it. If you are inclined to lose patience with making errors as the result of ignoring evidence you know is relevant, then you are inclined to exhibit a classic pattern of resistance.

5. Explaining Excessive Defection

It is time to confront the consideration that I think ultimately drives Everybody's Objection: "Surely there is some mistake that we are making when we defect excessively from successful SIRS. After all, this defecting leads us to reason less reliably than the SIRS. Once we put our finger on this mistake, it will be obvious that in a performance error and not a result of the proper operation of our reasoning competence." This consideration, however, is not a new one. We will continue to drive Everybody's Objection unless I can explain, even tentatively, how the selective decision results might be the result of the normal operation of our reasoning competence, rather than the result of ignorance or false beliefs about the normal. The origin of the epistemic critique is that our reasoning competence takes non-defection errors to be epistemically worse than defection errors. Our reasoning competence operates in such a way that a prepared to make more total errors in order to make fewer non-defection errors. It is important to keep in mind that the outputs of our reasoning competence are epistemic intuitions reflective but non-discursive judgmental errors. So our decision to accept this trade-off (more errors for fewer non-defection errors) is no more the result of conscious deliberation than are our Gettier judgments.

Why might our reasoning competence take non-defection errors to be worse than defection errors? Here are two mutually supporting hypotheses. First, psychologists have shown that regret plays an important role in our judgment and decision making. We know we will feel regret if we make an error. We take this anticipated regret into account when making decisions. Second, people feel accurate perceptions are more acceptable than inaccurate perceptions. Consequently, if we think that we will bring greater regret if we turn out badly, we are more likely to make accurate perceptions. Perhaps our epistemic intuitions in the selective decision studies are driven in part by the fact that we anticipate greater regret from non-defection errors than from defection errors. Consider the regret you would feel if you made a bad hire and your mistake was "We didn't pay attention to what we knew the evidence was saying" as opposed to "We defected from a formula that we knew was about 2% more reliable than we are in the long run." It is not implausible to suppose we anticipate considerably greater regret with the former mistake than the latter. If so, perhaps our reasoning competence reflects this fact when guiding how we reason.

The second hypothesis is suggested by a parole board commissioner who explains why he refuses to use SIRSs. "If you had the ability to look at a paper and make a judgment as you know a conscientious decision, but you chose to look at a piece of paper instead and made the wrong decision. Then you have a very good chance of making a greater number of generally consequential errors as a result of ignoring the SIRS. Perhaps we find non-defection errors epistemically worse than defect errors because we feel more epistemically responsible for non-defection errors than for defect errors. We feel greater responsibility when we adopt a belief we

The Autonomy of Social Episteme

July 2003

Michael A. Bishop

Michael A. Bishop

have stronger credentials, you know this.) In this epidemic situation, what do your epistemic intuitions—your reflective, non-discursive judgments—tell you about what you ought to believe? Let's consider one final case.

Case #4: You have applied the IM to the classes of T and F. The IM ranks F somewhat higher than T. But IM ignores interview evidence. As it happens, T had a much better interview than F. In fact, T sparkled, while F appeared almost incompetent, sometimes unable to answer questions coherently. You have no objection to the IM recommendation given only the evidence the IM considered. But you are certain that if the IM were amended so as to give even a slight degree of weight to the interview evidence, the large difference in the quality of their interviews would lead the model to recommend hiring T rather than F.

You have two powerful arguments for defecting from the IM:

(a) While interviews are only weakly correlated with job performance, in this case, T and F are sufficiently closely matched on other measures that the difference in their interviews Gray in the scale in T's favor. The evidence best supports the belief that T has the stronger credentials.

(b) The improper model comes to the opposite conclusion because it completely ignores interview evidence. If it weighed interview evidence appropriately, it would have come to the conclusion that T has stronger credentials than F.

Once again, you know that when you defect (your long term accuracy is slightly worse than the SIRS). But you also know that there are legitimate cases for defection. In this case, you are justified in accepting T as the stronger credentials than F. What do your epistemic intuitions—your reflective, non-discursive judgments—tell you about what you ought to believe?

Those who are inclined to defect in cases #3 or #4 might object that I have rigged the game by describing cases in which one has strong evidence in favor of defecting. If you are inclined to raise this objection, I have to wonder: What were you expecting? That expert clinicians with years of experience would defect from SIRSs arbitrarily or when they had clearly insufficient reason? Such assumptions are unacceptable.

In order to defend the final test in such cases, you would have to argue that the interview evidence is relevant and that it would level the playing field between T and F.

You might object that ignoring evidence you knew was relevant, the next time your department hires you would be less inclined to ignore evidence you were justified in believing was relevant.

Consider again the resistance finding represented in figure 6. Group B doesn't resist much at first; it's only later, after they had experience with the SIR that they really resist ignoring relevant evidence. So being able to ignore irrelevant evidence once or even on a few occasions, isn't enough. You must remain determined to ignore evidence you know is relevant, even after you've been Poverty by doing so. And the longer you keep it up, the more you can expect to be burned by it. If you are inclined to lose patience with making errors as the result of ignoring evidence you know is relevant, then you are inclined to exhibit a classic pattern of resistance.

5. Explaining Excessive Defection

It is time to confront the consideration that I think ultimately drives Everybody's Objection: "Surely there is some mistake that we are making when we defect excessively from successful SIRS. After all, this defecting leads us to reason less reliably than the SIRS. Once we put our finger on this mistake, it will be obvious that in a performance error and not a result of the proper operation of our reasoning competence." This consideration, however, is not a new one. We will continue to drive Everybody's Objection unless I can explain, even tentatively, how the selective decision results might be the result of the normal operation of our reasoning competence, rather than the result of ignorance or false beliefs about the normal. The origin of the epistemic critique is that our reasoning competence takes non-defection errors to be epistemically worse than defection errors. Our reasoning competence operates in such a way that a prepared to make more total errors in order to make fewer non-defection errors. It is important to keep in mind that the outputs of our reasoning competence are epistemic intuitions reflective but non-discursive judgmental errors. So our decision to accept this trade-off (more errors for fewer non-defection errors) is no more the result of conscious deliberation than are our Gettier judgments.

Why might our reasoning competence take non-defection errors to be worse than defection errors? Here are two mutually supporting hypotheses. First, psychologists have shown that regret plays an important role in our judgment and decision making. We know we will feel regret if we make an error. We take this anticipated regret into account when making decisions. Second, people feel accurate perceptions are more acceptable than inaccurate perceptions. Consequently, if we think that we will bring greater regret if we turn out badly, we are more likely to make accurate perceptions. Perhaps our epistemic intuitions in the selective decision studies are driven in part by the fact that we anticipate greater regret from non-defection errors than from defection errors. Consider the regret you would feel if you made a bad hire and your mistake was "We didn't pay attention to what we knew the evidence was saying" as opposed to "We defected from a formula that we knew was about 2% more reliable than we are in the long run." It is not implausible to suppose we anticipate considerably greater regret with the former mistake than the latter. If so, perhaps our reasoning competence reflects this fact when guiding how we reason.

The second hypothesis is suggested by a parole board commissioner who explains why he refuses to use SIRSs. "If you had the ability to look at a paper and make a judgment as you know a conscientious decision, but you chose to look at a piece of paper instead and made the wrong decision. Then you have a very good chance of making a greater number of generally consequential errors as a result of ignoring the SIRS. Perhaps we find non-defection errors epistemically worse than defect errors because we feel more epistemically responsible for non-defection errors than for defect errors. We feel greater responsibility when we adopt a belief we
to be an error, an error we could have prevented but didn’t, then when we adopt a belief about η that we were justified in thinking was best supported by the evidence we had about η. Many epistemologists have argued that our notion of epistemic justification is tied to the notion of epistemic responsibility (e.g., Kornblith 1983, Plantinga 1993). If so, then we would expect our epistemic intuitions, if they track justification reasonably well, to incline us away from non-detection errors.

My suggestion is that our suboptimal reasoning in the selective detection studies is the result of ignorance or false beliefs or poor reasoning. It is the result of the proper operation of our reasoning competence, which directs us to accept more total errors in order to make fewer non-detection errors.

6. Objections and Implications

Social institutions have no trouble abiding by the FMP. A financial institution might make thousands of credit decisions a day. A hospital might make thousands of medical diagnoses a year. From the perspective of these institutions, error is inevitable. What matters in making these thousands of judgments is minimizing significant errors. So if institutional decision-makers are faced with using either a SP that minimizes significant error or human experts who make considerably more significant errors, it’s a no-brainer. That is, the fact that the experts would make fewer non-detection errors cuts no ice.

I believe that our reasoning competence can be returned to better handle selective detection cases. In this spirit, the individualist can (and, I believe, should) accept the FMP as a principle of individual rationality. This will be particularly easy if she accepts the method of wide reflective equilibrium (Goodman 1965, Daniloff 1979). But however the individualist goes about satisfying the revision of our notion of individual rationality in the light of the FMP, isn’t there something disingenuous about the individualist absorbing the FMP while denying the autonomy of social epistemology? After all, the individualist has to revise their tacitly held principles of individual rationality because they are mistaken; they are practically inconsistent with a correct principle of rationality, the FMP. So when the principles of our reasoning competence are right, they’re right; and when they’re wrong, the individualist revises them without guilt or penalty. Nice racket.

An individualistic theory of rationality that simply articulates and clarifies the principles of our reasoning competence will end up being incorrect. Such a theory must be informed and improved by principles of social rationality — principles governing epistemically well-ordered social systems. It’s perfectly reasonable for the individualist to turn around and absorb elements of social epistemology. But it would be delusional for him to think that social epistemology does not have its own, distinctive theory — that it simply borrows the principles of individual epistemology and applies them in cases in which social factors are crucial.

While relativists are likely to be sympathetic to the argument presented here, they might not take it as a case for autonomy. They might argue that the autonomy thesis is false because social epistemology does not have a unique set of basic normative principles; one normative theory, relativism, is all we need for both individual and social epistemology. But this objection fails to appreciate the logic of the case for autonomy. The case depends essentially on the fact that relativism is inconsistent with the principles that constitute our reasoning competence. So relativism cannot be a particularly good theory for social epistemology theories that capture the contemporary judgments of our reasoning competence. Of course, the relativist is free to join the other individualists in insisting that we should revise our reasoning competence and our epistemic intuitions in light of what we have learned about the principles that govern epistemically well-ordered social systems. But then she is in the same predicament as the individualist in the last two paragraphs: correcting our principles of individual rationality with the principles of social rationality. If we take relativism to be a promising theory of social epistemology rather than a doomed theory of individual rationality, then relativists have always embraced the autonomy thesis, whether they’ve known it or not.

Some might be disappointed that the conception of social epistemology I have employed in this paper is a rather conservative, relativist one. For the disappointed, let me emphasize two points. First, my goal has been to argue that even on very conventional, individualism-friendly conceptions of rationality, social epistemology is autonomous. Nothing in this paper commits anyone to these assumptions. And second, the autonomy thesis I have defended isn’t especially modest. Let me briefly (and without much argument) mention three ways in which the autonomy thesis is boldier than it might first appear.

1. The autonomy of social epistemology does not mean that it has no connections to other disciplines. It is likely to depend crucially on findings in the social sciences — psychology, economics, history and sociology. But armed with pragmatic and semantic notions (e.g., useful truth), social epistemology is unlikely to need any complicated results from traditional, individual epistemology. Good social epistemologists can (and often do) ignore traditional, individual epistemology as it has been practiced since Plato.

2. The case for autonomy raises worries about the prospects of methodological individualism, which is practiced by most contemporary epistemologists. If our reasoning competence can conflict with correct epistemological principles, then it seems perverse to insist upon methodological theories that capture the contemporary judgments of our reasoning competence. This worry is magnified by evidence showing that people’s reasoning competence seems to vary by culture and socioeconomic class (Weinberg, Nichols & Stich 2001).

3. Social epistemology is autonomous because there is some (in)ferential knowledge that social theories of social epistemology are not. This knowledge is not being applied (perhaps badly applied) to the individual. Epistemology has always been social.

References


Notes

1. For helpful comments on earlier versions of this paper, I want to thank Jennifer Tadlock, Barn Reed, J.D. Trout and the audience at the Second Annual EPISTEME Conference, particularly Dennis Whiting, David Chrisansen, Don Falls, Alkin Goldman, Darren Matthews, Gary S. Ratten, Harvey Siegel, and Jonathan Vogel. This research was supported by a National Science Foundation grant (SES#0334536).

2. Performance errors include errors that arise from the failure of ancillary cognitive systems (which is the standard notion of performance errors) taken from linguistics) and errors that are the result of ignorance of all false background beliefs.

3. The qualification 'systematic' recognizes that most methodological individualists will permit some conservative revision of the principles that constitute our reasoning competence in the service of coherence and clarification.

4. I will assume that we are very good trackers of epistemic support. While I worry that this is too optimistic, it is crucial for the methodological individualist. Why should she build epistemological theories that capture our epistemic intuitions if she thought we're likely at tracking important epistemic properties?

5. If I chose 70% reliability scores for the IM and PM because that is the reliability score of the Goldberg Rule. But the distribution of agreement between the IM and PM (55.1/15.1/15) is purely hypothetical. The claim made here is insensitive to this distribution. As long as the IM and PM have 70% reliability scores, a 95% tracker of PMs judgments will lose 2% reliability.

6. The delay occurs because members of group B have to undergo a process of coming to recognize that the SRF is improper and that it sometimes recommends beliefs that are not supported by the evidence. And then they have to be motivated to defect (more often than group A) in the face of the knowledge that the SRF is more reliable than they are. This process takes time.

7. Perhaps you are skeptical that such a model could exist. This is to be expected. Recall the wide- laders' reaction to the successful wine prediction SRF - "somewhere between violent and hysterical.,

8. If you track epistemic support 80% of the time, you’ll defect to the true belief a bit less than 40% of the time. If you track epistemic support much worse than this, one wonders: Why does the individualist seek epistemological theories that capture our epistemic intuitions if we track epistemic support so poorly?

9. The pattern of agreement between IM and PM makes a difference here. If we suppose they agree 90% of the time, so that the pattern of agreement is (55/25/5) rather than (55/15/15), then you defect 14% of the time and you defect to the true belief 6% of the time. So if you defect, you can expect to defect to the true belief almost 43% of the time. Even if they always agree.
Circles of Reason: Some Feminist Reflections
on Reason and Rationality

Rationality and reason are topics so fraught for feminists that any useful reflection on them requires some prior exploration of the difficulties they have caused. One of those difficulties for feminists and, I suspect, for others in the margins of modernity, is the rhetoric of reason — the ways reason is bandied about as a qualification differentially bestowed on different types of person. Rhetorically, it functions in different ways depending on whether it is being denied or affirmed. In this paper, I want to explore these rhetorics of reason as they are considered in the work of two feminist philosophers. I shall draw on their work for some suggestions about how to think about rationality, and begin to use those suggestions to develop a constructive account that withstands the rhetorical temptations.

Philosopher Michelle LeDoux, in her recently translated, The Sex of Knowing (LeDoux 2003), takes note of Jeneáquins Rousseau’s somewhat bifurcated views regarding rationality. In his Lettre à D’Alembert, text and the accompanying notes offer accounts quite at variance with each other. In the notes, LeDoux finds a complex and open view of reason, quoting Rousseau: “Human reason has no well-defined common measure and... it is unjust for any man to give his own as the rule to that of others, let us suppose good faith, without which all dispute is only cackle. Up to a certain point there are common principles, a common evidence, and in addition each has his own reason which determines him. Thus this sentiment does not lead to skepticism, but also since the general limits of reason are not fixed and no one can inspect another’s, here, with one stroke, the proud dogmatist is stopped.” (Rousseau 1984, quoted in LeDoux, p. 174).

There are some sticky points in these remarks and in others in the same vein in the notes. One wants to know the extent of the common principles, what it is for one’s own reason to determine one, what reasons could be offered for the principles he affirms. But LeDoux’s point is that Rousseau’s vision is an expansive one, of a reason that is open, that is not exhausted by a set of rules, that accommodates diversity, that incorporates dialogue and criticism. Surely a promising vision to explore and elaborate. But this wouldn’t be Rousseau without a catch, and LeDoux is on to him, drawing her reader’s attention to what he says in the text. Here he chastises the men’s club of Geneva as the venue of True Reason, largely because here the men are “exempted from having to lower their ideas to the range of women.” So the best exemplar of Reason is a closed society, best just because it is closed, restricted to men who are Genevan citizens. This reason flourishes in closed circles, closed in terms of membership and in terms of the principles that might constitute reason.

Apparently, there is a common measure to which some fail to measure up. LeDoux’s exhibition of Rousseau’s duplicity helps us see the historical baggage the concepts of reason and/or rationality carry — at one and the same time, desiderata of a kind of capacity, thought to be a human universal, and specification of rules for the proper exercise of that capacity. The specification of rules or status inevitably closes down what and who will be counted as rational. It disqualifies as possessors of the capacity those who don’t or who are presumed not to conform. Citizenship, male, of course, whether in Athens or Geneva, has often functioned as a criterion of performance according to the rules, enabling those within the circle to dismiss those outside, heedless of their protests. Twenty years ago, historian of philosophy Genevieve Lloyd published a stunning short book, The Man of Reason,