Croatian Journal of Philosophy
Vol. VII, No. 20, 2007

Book Reviews

Michael A. Bishop and J. D. Trout, *Epistemology and the Psychology of Human Judgment* (Oxford: Oxford University Press, 2005), 224 pp.

The book brings fresh insights into contemporary epistemology and, more importantly, it radicalizes and sharpens recent naturalistic approaches, which keeps it in line with some similar approaches to epistemology (like, for instance, Nozick's in *The Nature of Rationality*, ch. III, Henderson's in "Epistemic Competence and Contextualist Epistemology", *The Journal of Philosophy* 91 (1994), 627–49, and R. Cummins's in his criticism of reflective equilibrium). It shares two common traits with them: *ameliorative* character of the proposed alternative approach to standard analytic epistemology, and extensive criticism of this classical analytic epistemology. Ameliorativism is a twofold thesis: there is bad news, people are endowed with relatively unreliable, and therefore poor, reasoning strategies; and the good news is that such strategies can (and should) be repaired or even substituted with more reliable ones. Classical or standard analytic epistemology is usually conceived as a thesis claiming that our reasoning processes are to be justified, and the justification procedure is crucially taken to have a form of internal, reflective acceptance of beliefs as an output of these processes. The amelioративists take this characteristic as a target of their criticism.

Bishop and Trout's book radicalizes the views from amelioративist tradition in both points, positive proposals and criticism. Chapters containing positive proposals are 2, 3, 4, 5, and 9, while those critically oriented are 7 and 8. I am going to overview briefly the structure of the book and its main theses and arguments, and then scrutinize and challenge some points I find questionable or implausible.

After the Introduction follows the chapter "Laying cards on the table" where the authors clearly and concisely set forth their views on epistemology and its place in human life. The authors hold that epistemology should be a serious business having an important and practical task. Unfortunately, epistemology in its standard, classical appearance does not satisfy this lofty ideal. The very first sentence of the chapter unambiguously stresses it: "It is time for epistemology to take its rightful place alongside ethics as a discipline that offers practical, real-world recommendation for living" (p. 6). The proclaimed aim of this book is to offer an alternative to standard epistemology attaining the tasks it fails to do.

The subsequent chapter displays a kind of "starting point" for such renewed epistemology that should be "normatively reason-guiding and genuinely capable of benefiting the world" (p. 7). The starting point that the authors rely upon is the view supported by a number of psychological experiments that have been performed since the early sixties. The book is mostly relying on the experiment type discussed in detail by Hastie and Dawies[1] (very frequently quoted in Bishop and Trout's book). The experiments address "the question of whether trained experts' predictions were better than statistically weighted averages of the relevant predictors".[2] Interestingly, the experiments confirm the thesis that some statistical rules (the paradigmatic case is Godberg's rule) significantly outperformed predictions made by experts. The authors find the experiments enormously important for epistemology and introduce the topic in the chapter "The amazing success of statistical prediction rules" (Chapter 2). They call the branch of psychology which tackles this kind of finding "Ameliorative Psychology". To get a richer insight into the story, let me paraphrase the authors' description of the Goldberg's rule.

The Goldberg Rule is the most documented success of Ameliorative Psychology. It predicts whether a psychiatric patient is neurotic or psychotic on the basis of a famous clinical profile, so-called "MMPI profile". Lewis Goldberg (1965) found that the rule expressed by the following formula outperformed 29 clinical judges:

$$x = (L + Pa + Sc) - (Hy + Pt)$$

where L is a validity scale and Pa, Sc, Hy and Pt are clinical scales of the MMPI.

If $x < 45$, diagnose a patient as neurotic. If $x \geq 45$, diagnose a patient as psychotic. When tested on a set of 861 patients, the Goldberg Rule had a 70% hit rate; clinicians' hit rates varied from a low of 55% to a high of 67%. (13 of the 29 clinical judges in the above study were experienced PhDs, while the other 16 were PhD students. The PhDs were no more accurate than the students.)

The Goldberg Rule is an instance of the *Improper Linear Models* (PLM), which, with *Proper Linear Models* and *Bootstrapping Models*, forms a family of superior reasoning strategies called *Statistical Prediction Rules* (SPR). Reasoning strategies are "rules for making judgment on the basis of certain cues" (71). There are four elements characterizing those rules:

a) the cues used to make the prediction;
b) the formula for combining the cues to make the prediction;
c) the target of prediction (what the prediction is about);
d) the range of objects (states, properties, processes,…) defined by detectable cues, about which the rule makes judgment that are thought to be reliable.

To put the idea into a broader frame, the authors bring up a theoretically interesting claim that ameliorative psychology should serve as an empirical basis for plausible epistemology (their idiom is *Strategic reliabilism*, or SR for short) and supply it with new reliable reasoning models. This idea

[1] Reid Hastie and Robin M. Dawes, *Rational Choice in an Uncertain World* (London: Sage Publications, 2001).

[2] Ibid., p. 55.

is elaborated in the chapter "Extracting Epistemic Lessons from Ameliorative Psychology". Strategic Reliabilism, the vision of epistemology that the authors promote, gives "a systematic voice and the theoretical foundation to the long-standing success of SPRs while at a same time avoiding the most serious objections to traditional *process reliabilism*" (p. 71).

Bishop and Trout elaborate further traits of their theoretical views on *Strategic reliabilism* in three chapters under a common header (with the following subtitles: Robust Reliability, The Cost and Benefit of Excellent Judgment, Epistemic Significance). One of the distinguishing features of strategic reliabilism is that, in contrast with standard epistemology which aims at *justified* beliefs, it looks for *excellent* beliefs or judgments. To possess distinctive excellence, strategies must be *reliable*, having a high hit score, and be *tractable* enough so that a naïve thinker can make use of them. Using economic parlance, excellent strategies are detectable by cost-benefit analysis. What is required is not absolute but relative quality of the reasoning strategy. "According to SR, the relative quality of a reasoning strategy is the function of its expected costs and benefits, as well as those of its competitors. It is an empirical question which of the two reasoning strategies has a better cost-benefit ratio for a particular reasoner" (p. 21).

The parts of the book containing criticism of standard analytic epistemology (SAE) come up in chapters 7 and 8. Chapter 7 offers core criticism of standard analytic epistemology while chapter 8 criticizes the so called *rejecting-the-norms* arguments, which a number of authors have employed against the *heuristics and biases* approach. Chapter "Positive Advice" concludes the book.

Let me articulate the central idea of the book: the essential task of the renewed epistemology (in the form of *strategic reliabilism*) is to improve and repair the reasoning strategies people utilize in different domains. According to this, epistemology is a practical discipline that, having beneficial potential, can and should advise and thus normatively guide real lives of average people. That sounds simple and clear. Beyond simplicity and clearness, it looks convincing and elegant. Still, I am not quite persuaded that SR is so virtuous.

To begin examining the proposed doctrine, two issues seem to me insufficiently transparent. The first issue concerns the clear determination of rules and strategies falling into the SPR category. The reasoning strategies described and offered as typical examples are those applied in making diagnoses of psychiatric patients, hiring new employees or making predictions about someone's creditworthiness. In short, all of them are related to predictive tasks. Those strategies, structurally akin to Goldberg's rule, are indeed well documented and experimentally proved to be superior over expert's reasoning. But, one would expect from those who are going to ground new epistemology that its building blocks, set of predictive strategies, extend to a much broader base. This is probably the reason why the authors, in their optimistic enthusiasm, count among SPRs many other rules and models having little or nothing to do with the predictive function of SPRs. Among the others, they mention "neural networks, naïve Bayes classifiers, Markov Chain Monte Carlo algorithms, decision tree models and support vector machines". Some of them cannot be counted as predictive strategies. Some of

the mentioned models certainly are not designed to improve human thinking. It would be very interesting to know how neural networks, for example, can possibly be used to improve thinking.

The other neuralgic point is the relation between the criticized theory, SAE, and the theory supposed to avoid its failures, SR. If SR has to replace the standard epistemology, it is supposed to be a sum total of epistemology. SAE itself covers very different topics and fields concerning people's knowledge, reasoning processes and judging procedures, perceptual and other cognitive abilities, their structure and, possibly, their origin, not to mention the emotions and testimony. On the other hand, the authors seem to be unconcerned with those topics. But the question is, does it mean that all other epistemological interests, fields and domains that standard epistemology tackles, should be neglected and even jettisoned? The authors are silent about that, but their criticism of SAE is aimed to show SR's superiority over all other forms of standard epistemology. If RS is superior over SAE in the sense that it avoids a great part of problems that SAE is unable to handle, it is obvious that RS should substitute SAE. The substitution claim heavily depends on the superiority claim. In the rest of the review I will scrutinize the authors' attack on standard analytic epistemology and challenge their claim of superiority of SR over SAE.

The initial problem might be that SAE, as a target of criticism is, also according to Bishop and Trout, an umbrella term for quite different approaches. Some of them are very close to SR. Nozick's approach, for example, seems to be reliabilistic in the way similar to Bishop and Trout's. Henderson, on the other hand, advocates a version of ameliorativism quite close to the one our authors have in mind. But I am going to neglect it and discuss Bishop and Trout's criticism in treating all different approaches as if they belong to the same front. Nonetheless, one has to admit that all these different approaches have in common a tendency toward some form of justification.

Assailing SAE, the authors condemn it for a number of incorrigible mistakes, pathologies, or, as I will name them, "deadly sins". Some of those sins are more deadly than others. To start with less deadly ones, the sin of *descriptivism* should be mentioned. It might sound a bit surprising, for, as a matter of fact, descriptivism is a characteristic frequently used by critics of naturalistic epistemology who argue that descriptive naturalism can hardly bridge the *is-ought* gap and so attain a genuinely normative character as a necessary mark for epistemological theory. In their maneuver[3] Bishop and Trout turn the critical blade toward the critics claiming that SAE theories are descriptive in a more malicious way than the naturalistic theories.

So, SR and SAE are both descriptive, but differ in the consequences that descriptivism has on bridging the *is-ought* gap and on objects of description. In the first case, those of SR, the descriptive character of the theory facilitates bridging the *is-ought* gap (people simply *ought* to employ strategies that are more reliable and have a higher hit rate), while in the latter case it hinders the normative status of the theory ("SAE will never provide effective normative guidance, because … its goals and methods are beyond

[3] The position is at length exposed in Bishop and Trout's article "The Pathologies of Standard Analytic Epistemology", to appear in *Nous*.

repair" (p. 22). Regarding the object of description, as a descriptive basis SR has reasoning strategies of average people. Although its descriptivism is not focused on revealing the causal structure of the reasoning processes, it "tends to focus on the assessment of reasoning strategies in terms of their reliability or their relative reliability". On the other hand, SAE's object of description is something quite different. If SAE gives us descriptive knowledge, what is this knowledge about? The answer leads us to a more transgressive sin of SAE, that of *Parochialism*. Talking about the method of SAE, Bishop and Trout say that "… this method is geared to give us descriptive knowledge about the epistemic judgment of a relatively small group of idiosyncratic people". The SAE theories, hold Bishop and Trout, describe only epistemologists' judgments, which are not shared by a vast majority of laymen, the non-philosophical population. This severe charge requires a more detailed elucidation. So, where does this parochialism comes from? It is a consequence of another deadly sin of SAE, that of *conservativism*, described as *stasis requirement:* "Philosophers accept or reject an epistemological theory on the basis of whether it accords with their considered judgments" (p. 105). "Rejecting theories solely because they do violence to our considered judgments," they continue "is a *shockingly conservative* [italics N. S.] principle of theory choice. This may only become clear if we compare it to methods in other fields of inquiry. The special theory of relativity does extreme violence to our considered judgments about simultaneity. But that is hardly a reason to reject it."[4]

Finally, the reason why SAE's method is so *shockingly conservative*, why it prescribes acceptance of only those beliefs or strategies that are in accordance with our considered judgments, lies in the fundamental, original sin, that of *justificationism*. The goal of justification is common to all SAE theories, while, on the other hand, SR, supported by AP, assesses reasoning strategies aiming to distinguish those with greatest *epistemic excellence*. Let us see what is so sinful in the justification process. Justification as internal, reflective acceptance of beliefs is, according to Bishop and Trout, clearly expressed by L. BonJour's formulation: "[O]ne's cognitive endeavors are epistemically justified only if and to the extent that they are aimed at [truth], which means very roughly that one accepts all and only those beliefs which one has good reason to think are true. To accept a belief in the absence of such a reason, however appealing or even mandatory such acceptance might be from some other standpoint, is to neglect the pursuit of truth" (**BonJour 1985???**, p. 8). Taking into consideration a good reason (that p is true) is, according to this remarkably clear formulation, the crux of *internal* justification.

We are now in a position to disclose the salient points of the criticized theory and to briefly summarize the criticism. I am focusing here on the pure structural frame of SAE, neglecting its values consequences. Here is the structure: to justify some new belief, strategy or theory, the cognizer has to have good reasons to reflectively, internally accept it as true. What provides such good reasons is a kind of reflective coherentism in which one accepts those beliefs that accord with his considered judgments and rejects

[4] I cannot resist the temptation to add that the reason to accept the theory certainly is not its *predictive power*.

those that violate them. Such a procedure, as the accusation goes, would never allow correction, improvement or reparation of a reasoning process, even if it is proved to be inadequate or poor.

There is at least one step in the argumentation that goes wrong. I mean the part of the argument saying that having a good reason is coherentisticaly determined. Thus formulated, it gives the impression that the only permitted way of reflective justification or having a good reason from the thinker's perspective is a matter of equilibrium that justifies beliefs in an entirely a priori way. It is not true because it is not in contrast with the fact that our best warrant or justification has to be sensitive to the success of our reasoning processes. To be justified in accepting p or to have a good reason for taking it as true (as a matter of fact, these are two sides of the same thing)[5] includes sensitivity to the success of our action, namely, the reliability of the process the particular belief proceeds from. The process remains internal and reflective even if *having a good reason that p is true* is based on empirical evidence.

SR, on the contrary, offers a justification-free epistemology and declares a thoroughgoing externalism. To refute justificationism, the proponent of SR has to show that a purely externalist theory can normatively guide the agent to choose the particular reasoning strategy *without having* a good internal reason for it. If there is any reason at all, it must be external. That means that having such a reason the agent must be able to choose a course of action without being internally justified. Here we have a dilemma. Either such an external reason may function as a blind compulsion, in which case an agent does not really choose anything, or she is actually in a position to choose, in which case she must do it internally. Since we know that average agents in everyday situations act as conscious decision-makers, the solution with blind compulsion in not an alternative.

Hence, my objection to SR is quite simple. Being normatively advised to choose that reasoning strategy that has a relatively higher hit rate, the agent decision procedure must be internal and reflexive. She must decide, relying on some decision procedure and weighting some reasons, which strategy is more reliable and more feasible. Therefore, the structural difference between SAE and SR turns out to be quite thin.

Bishop and Trout can answer, as they actually do, that there are two kinds of reasons and, therefore, two kinds of rationality assessing those reasons. One kind of rationality is appropriate for standard, individual epistemology, while the other fits SR that has "essentially" a social character. Since there are clear cases in which these two types of rationality, advising two behavioral patterns, are inconsistent, one of these patterns must be more or less irrational. A good way to prove the superiority of SR is to show that the behavior governed by principles of individual epistemology must be displayed as irrational. The frame for discussing this argumentation is not present in the book but is offered in Bishop's article "The Autonomy of Social Epistemology" (draft).

[5] I can have the warrant about strategies I imply, in some kind of a priori manner, but to be reflectively justified, to know about this warrant, I must be in some degree sensitive to the success or reliability of those strategies. This a posteriori or empirical element in justification process is not in contrast with justification-based epistemology.

To present the issue, let me start with a case where a cognizer has to choose between strategies *x* and *y*. Let *x* be rule-based on the linear statistical model (statistical prediction rule, SPR) and y be strategy based on individual, possibly expert's judgment. According to SR (social) standard of rationality, our cognizer is strongly recommended to choose x. Taking into consideration what is going on from the cognizer's perspective, however, the situation becomes pretty complex. To describe the differences between strategies x and y, I shall briefly picture the core of the problem and ignore some details.

The predictor is given the formula containing cues that SPR takes in consideration. She also knows that SPR obtains score of 70% correct predictions in, say, 1000 cases, while experts' predictions has a rate of 68% correct predictions. In addition, she also knows that there are some cues SPR does not consider. Let's imagine that our predictor has to make a decision in the particular situation where she has evidence that in this very situation a cue that SPR does not consider strongly recommends defection from SPR. It is obvious that in this situation our (individual) epistemic intuition tells us to defect *although we know that in the long run it is worse than non-defection.* But, it is also true that in *this* situation defection almost inevitably leads to correct prediction.

Bishop's diagnosis is that the defection is irrational because "our reasoning competence operates in such a way that it is prepared to make more total errors in order to make fewer non-defection errors", and "our reasoning competence takes non-defection errors to be epistemically worse than defection errors". It is not clear at all why defection in the particular situation, when predictor knows that such behavior will lead to correct prediction, should be wrong. All what we have are two equally strong intuitions, one advising non-defection and the other advising defection. None of them is decisively wrong or irrational. If that is true, Bishop and Trout have no firm ground to declare superiority of SR over SAE.

NENAD SMOKROVIĆ
Department of Philosophy,
University of Rijeka