Journal of Economic Literature
Graduate School of Business
Stanford University
Stanford CA 94305-5015

July 4, 2000

Dr. Andreas Ortmann
Center for Adaptive Behavior and Cognition
Max Planck Institute for Human Development
Lentzeallee 94
141954 Berlin
Germany

Dear Dr. Ortmann:

I enclose two referee reports on your paper "Biases and Heuristics in Psychology and Economics," submitted to the *Journal of Economic Literature.*

Referee A says that it is premature to cover this approach in the JEL since it has not yet had enough applications to economics. This referee suggests that perhaps in three to five years, after the approach you advocate has been tried out in economics applications, there might be room for a *JEL* article on it. My reading of the proposal leads me to the same conclusion. I personally am attracted to the point of view you present, but I think the topic will only be ready *for JEL* coverage when there is more to say in Section 3 of your outline: that is, more than just saving that the heuristics-and-biases program doesn't work, there would need to be some positive indication of what the alternative view has to offer economics. For example, after the "empirically better testable" explanations you talk of at the end of your proposal have been published in journals and subject to scrutiny from the profession, then would be the time for a *JEL* article.

The other referee also notes the lack of a positive account of reasoning.

Both referees offer a host of suggestions that I think you will find useful.

The paper deserves to be published somewhere, but the *JEL* is not the right place for it. You might send the full paper to *Economics and Philosophy* or some such journal.

Thank you for sending your work to the *JEL.*

Yours sincerely,
John McMillan

Referee 1:

JEL proposal, "Biases and heuristics in psychology and economics," Ortmann and Hertwig

The authors propose to describe an approach which emerged as a kind of criticism of the heuristics-and-biases (HB) approach to decision making which was fruitful in the 1970s-80s and is now finding application in economics. The critical approach, sometimes called ecological rationality or "fast and frugal", reacts to die following perceived problems in the HB paradigm (their p 4+): (1) Some effects can be shrunk by asking questions differently (chiefly, the relative frequency vs. single-event manipulation; also accurate calibration on how many of N answers will be right vs. single-answer narrow confidence intervals); (2) since normative principles often (though not always) produce point estimates, it is "easy" to produce large-sample directional violations; (3) rationality norms are misapplied; and (4) the heuristics arc not sharply defined.

I do not think this field is ready to summarize its results or arguments for economists quite yet. The proposal is essentially a warning to economists to "stay away from HB" and presumably move toward ecologically rational models. But not enough is known about the latter to be useful for economists at this point. In addition, you may know that virtually all the criticisms your camp has made of HB (and other elements of behavioral economics) have been made from inside economics as people have tried to argue for useful economic implications of these phenomena. You don't need to explain to economists how to criticize HB!

Finally, I think there is not so much fundamental disagreement between these two camps as you portray. A heuristic can be seen as falling short of optimal (measured by bias) or as doing nearly optimally given its simplicity (e.g., in operations research and computer science this is the normal meaning). Classic half-full, half-empty. KT were very clear early on (and later, in their "statistical intuitions" chapter in the 1982 blue-green KslovicT book) that the biases are just a way of guessing what the heuristics are, like the Muller-Lyer illusion tells you about principles of gestalt and contrast in perception. It **is** true that a large industry of people in professional schools, particularly, grabbed the biases part and emphasized that at the expense of understanding more about heuristics, which is a mistake. But there were parallel ideas in the air about heuristics as half-full (nearly-optimal), e.g., Dawes-Corrigan, Meehl, at the same time. (Incidentally, these early ideas are so far as I can tell rarely given enough credence or credit by the ecological rationality crew – e.g., linear models with specific weight structures are **exactly** the same in some cases as "take the best" models – which leaves a bad taste in the mouth of people like myself who are actually quite sympathetic but think you guys are overselling. If you were to say, "in the 1970s many people were interested in heuristics as half-full not half-empty, but the profession got caught up in the study of biases instead ... and we are here to rekindle interest in the former" that would work much better than pretending it all began with Gigerenzer, though he certainly deserves much credit too.)

While my job at this preliminary stage is not really to argue with you about these things, it is important to make clear my views and it might help you to get a sense of how an economist who takes these ideas seriously (though shares some reservations) will react. I think point (1) is not fundamental to whether HB should be used to do economics. There seems to be an odd implicit criticism that the HB approach used the "wrong" or "unnatural" format. What is the implication for economics? Well if economic decision makers answer questions like, "For 100 businesses just like yours, how many would fail?" then they are likely to be less overconfident, say, than if they think about their own idiosyncratic odds of success. But such managers often very actively resist the "outside view" (Kahneman and Lovallo's term) in which their business is one of a class, rather than special. So I just don't see why the evidence on sensitivity to representation means economists shouldn't pay attention to these phenomena.

Point (2) is not fundamental by any means. First, the claim that any difference can be inflated to conventional significance by jacking up N is just wrong. There are Bayesian corrections which require pseudo-t statistics to rise with N, and informal norms of proof like requiring very large p-values when samples are large, which combat precisely this problem. (As for your en passant point, this boundary problem is well-recognized now in experimental economics and in most areas, or at least many, where a boundary prediction is made and violated in the direction of something like bounded rationality, the same behavior can be reproduced in a game or choice with an interior solution. For example, for ultimatums you get a boundary prediction in standard games but get similar results in alternating-offer sequential bargaining which has interior predictions, see Ochs and Roth 1989 AER. For public goods games there are specific designs with interior self-interest solutions, see Lcdyard 1995 Handbook Expl Econ. Also in many cases results are specifically compared against rational predictions with structural models of error to precisely deal with this problem. Be careful to do your scholarship **very** diligently in this area. And do you really want to end up asserting that these phenomena, like fairness theories used to explain dictator allocations and ultimatum rejections, don't exist? Of course they do! The evidence and everyday intuition is just overwhelming. If you just want to scold exp'l economists for not being sensitive enough about the boundary problem, you should know this has been on people's minds for years and some studies have worked around it.)

Point (3) is controversial and will likely remain so. I've never been convinced by critiques like Birnbaum's or Gigerenzer's claims about what "practicing statisticians" would do. The fact is that inferring that P-> Q implies Q-> P, for example, is a logic error. Ceteris paribus, knowing that such logic errors are common in abstract problems, for example, can't help but have some economic content. Perhaps this error is a byproduct of an adapted mind which miscodes the question that is being asked. That interpretation just substitutes "error in classification" for "logic error". A better example may be ultimatum games. It is fashionable among gamc theorists to, curiously, deny the possibility that responder subjects thoughtfully sacrifice a dollar or two to punish a complete stranger who they think has behaved greedily (even though usually in economics a person's tastes are beyond dispute – they can spend money on a cheesy action movie in which greed is punished, but can't spend money to punish greed themselves!) The alternative argument is that people are used to playing repeated games with people they know and cannot turn off their repeated-game habit, or cannot distinguish one-shot and repeated games. Thus, they are playing "ecologically rationally" in one-shot games. But there remains a sense in which they are making a mistake – they are treating one kind of game as if it is like another, which it is not. It is ridiculous to say that the "wrong norm of game-theoretic equilibrium" is being used to judge the one-shot behavior. It is the right norm because a creature who is ecologically rational **and** can distinguish one-shot and repeated games will do even better than one who can't. You can argue that such a creature would **not** be well-adapted because they'd have to big a brain or would be using scarce brainpower for distinguishing rare events (one-shot games), but such an argument quickly becomes tautological. And in any case, it is useful to have a category of "super rational" or "idealized rationality" to describe a kind of rational behavior), which may not be evolutionary feasible or adapted but it is still a good benchmark. (After all, if some individual differences are better than others, it is good to have a high standard for selection; and we may also want to design organizations to be super-rational if possible.)

Point (4) is well-taken. However, note that incorporating some of these ideas into economics is likely to sharpen the definitions since economists hate words and love math. For example, the following papers offer reasonably precise theories of heuristic updating: Rabin-Schrag QJE 1997; Rabin unpub'd; Barberis, Shleifer and Vishnv J Fin Ec 1998?; Daniel et at J Finance 1998. So while have been slow to be precise about the mechanisms or their implications, it may be that economists will do this cleanup (and show that it is possible).

With that said, I don't think your paper will make a good JEL paper at this time:

First, the HB ideas are **not** at the heart of the sweeping change in economics taking psychology seriously. There are actually few papers using availability, rep'ness, or anchoring heuristics directly. More influential are theories of hyperbolic discounting, nonexpected utility (prospect theory), social utility functions, and so forth. So you may warn economists to stay away from HB, but they are not embracing the impact of that work in any case. (An exception may be overconfidence, but this is simply a pattern rather than a mechanism and **has** proved modellable and useful in explaining many patterns.) If you think your 4 criticisms of HB can be extended to all these domains then you could write a good JEL paper but I don't see how since generally they don't extend, and if you think so you have **a lot** of work to do.

Second, I don't see that the ecological rationality approach offers a better alternative way of getting psychology in economics than HB at the present time. (I.e., even if you can sell the argument that HB should be ignored, what's the sales pitch for ecological rationality.) Quick example (maybe an unfair one): I read w/ care the finance chapter in the ABC Group book. It is really dreadful. First, there are very high standards of proof in detecting anomalies in financial markets (careful control for survivor bias, comparison to control groups for risk, long samples with typically split-sample control for artifacts, etc.) which that chapter didn't meet. Second, it cannot be an equilibrium for more familiar stocks to always go up (when does it stop)? Third, the available evidence suggests the opposite tendency overall – small firms tend to outperform large, and similarly with glamour stocks (which do badly). In any case, my point is this: That chapter is as close as the ecol-rationality's have gotten to something which is familiar and important for economics, and did a very poor job. If there are better applications, that would make a great JEL paper but it seems likely to come 5-10 yrs down the road.

On an even more pessimistic note: A shortcoming of the ecol-rationality approach may be that, almost by definition, the sorts of heuristics you want to discover and argue for will be idiosyncratic to specialized types of adaptive problems. Economists love, and perhaps can't live without, very general principles which are deliberately **not** domain-specific. The two seem fundamentally at odds. Now you could well argue – I mostly agree – that if the brain is a collection of specialized modules, **that** sort of brain should be the foundation for microeconomics. But do to so would require a lot of detail and parsimonious specification for how to do economics with ecologically rational agents. That is a really hard enterprise. I think if you began doing it you would end up with an appreciation for how powerful some of the tools of behavioral economics are (e.g., prospect theory), **because** they are domain-general.

Specific example: On p 7 you write (line -5) that the high failure rate of small businesses [large business too, by the way] or day traders may be due to overconfidence, "but one can surely think of more precise, and empirically better testable, explanations." It's not that simple. There are lots of competing explanations, but they are not always precise (e.g, it may involve some unobservable artifact like utility for running your own business-- is that precise!?). Furthermore, the best work in this area (e.g. Odean) tries **very** hard to test competing explanations; if you can't rule them out then economists are always skeptical that it's overconfidence or something else behavioral. You should be more respectful of how hard behavioral economists have worked at trying to figure out what is due to behavioral limits and what is due to rational calculation (learning, etc.)

Referee 2:

Review of "Biases and Heuristics in psychology and Economics"

This proposal attempts to summarize various criticisms of the heuristics and biases tradition in judgment and decision making. It focuses on three major biases: the conjunction fallacy, the overconfidence bias, and the base rate fallacy. It offers criticisms such as: 1) biases are reduced when people reason about frequencies rather than probabilities, 2) the widespread asymmetric null-hypothesis testing of nominative models has a built in bias because normative predictions are point-specific, 3) there are alternative norms of rationality that may be based more on process than structure, 4) the heuristics-and-biases tradition does not specify precise processes nor makc precise predictions.

A critical review of this tradition could be valuable to the economics literature in light of the increasing attention given to "behavioral" economics. I'm not sure, however, that the current proposed format could make a very persuasive case. Here are the major problems.

The paper jumps back and forth too much between types of biases and types of criticisms. According to the proposal, the paper is going to present "three major biases – the conjunction fallacy, the overconfidence bias, and the base rate fallacy" (p. 3). But some of the criticisms draw on criticisms directed at other biases (e.g., illusion of control, p. 4, the wason selection task, p.6). Other criticisms intermingle biases and classes of criticisms. On pages 4 and 5, the argument is presented that frequency formats reduce both overconfidence and base rate neglect. The paragraph ends with Koehler's conclusion that "in some studies judgments were found appropriately sensitive to base rates, in others too little, and yet others too much." This doesn't follow from the preceding review of research on frequency judgments. There isn't a well-developed line of argument here.

In a similar vein, many of the criticisms are offered at a superficial level. Why should human beings be better at reasoning about frequencies than probabilities? What is an independent basis to choose other normative principles besides laws of logic, probability, and expected utility theories? What is the alternative to using point-predictions from nominative models-if they make them, shouldn't they be held to them?

Finally, the criticisms, even if accepted, don't present a strong case *for* human rationality, but simply raise doubts about some methods and some claims in the heuristics and biases tradition. Thus, it reads more like an apologia than a positive account of rationality. Normative theories do make point predictions – why shouldn't they be used as the null-hypothesis? What's the alternative? And, in explaining away some of the effects, claims arc offered such as "the fact that many of the alleged cognitive biases can be made to disappear, or at least reduced, and in some cases inverted, prompts questions about experimental design and implementation (p. 5)." Inverting a bias does not seem like something to celebrate! It demonstrates that people do not conform to a Normative principle, but that the process that causes the bias is more complicated than the current hypothesis, which then spurs the search for a new process. (As a concrete example, the "regression-to-the-mean" explanations offered by Erev, Wallsten, & Budescu and by Soll for overconfidence successfully predicts both over- and underconfidence – see cites below.) A similar apology is offered for people's failure to use Bayes' rule. Birnbaum is paraphrased as saying, "in terms of a signal detection model, for instance, a witness may try to minimize some error function" and then different error functions are proposed. These criticisms are unsatisfactory because they are unorganized and ad hoc. In the end, there is no positive account of "limited but adaptive reasoning" offered in the place of the heuristics and biases tradition.

Recommendations

I think the paper would be much improved if it focused an only one type of bias – a very celebrated and important one – and thoroughly critiqued it. One advantage of this is practical. Given that the current paper would like to address three major biases and "a half dozen other cognitive illusions", there will have to be a lot of citations, basic review, and explication to cover these basic topics and get the average, uninformed economics reader up to speed. But the other advantage is rhetorical: The current arguments jump between biases and criticisms, without pinning down the problems with any one bias. So, for example, the paper could focus on a prominent and much celebrated bias, such as "overconfidence," and critique it thoroughly.

Given the controversy that swirls around the heuristics-and-biases tradition, the authors also need to be balanced and inclusive. Here are some suggestions, starting with overconfidence.

There are several lines of research on overconfidence that are neglected in the current review. First, there is a new tradition (Erev, Wallsten, & Budescu in Psychological Review, 1994; Jack Soll in OBHDP, 1995; see also Juslin, Psychological Review, 2000) focussing on a regression to the mean problem. They offer a statistical explanation (along the lines of "biased samples") for its existence, but don't deny it exists! Second, there is also research on how frequency formats do not eliminate overconfidence (Klayman, et al, 1999, OBHDP) or introduce new types of errors in confidence judgments, including underconfidence (Buehler & Griffin, 1999, Cog Psych). (And, once again, I don't see why inverting an error should be a cause to celebrate. For the frequency claim to be persuasive, it needs to lead to better calibration – i.e., confidence should match accuracy.)

Other comments on "balance" and "thoroughness":

To an outsider to this debate, it seems that both Gigcrenzer's work and Kaluieman and Tversky's early work could be characterized as falling in the Simon tradition of bounded rationality. But the rhetoric of this work has focussed on exaggerating these differences. Perhaps it comes down to the fact that Gigerenzer and colleagues are trying, to "redefine what constitutes rationality by taking into account constraints on resources such as time, knowledge, and cognitive processing" (p. 2). This quote admits that both groups share a common perspective on limitations. It shifts the focus to definition. But a paper focussing on definition – and justifying the rationality of limitations – would be a quite different paper than the one proposed.

Kahneman and Tversky in their original Psych Review article also showed that the conjunction fallacy was reduced using a frequency format.

There are other perspectives besides the Kahneman and Tversky (people are inherently flawed) and Gigerenzer (people are adoptively rational). The first is a tradition of corrigibility. Richard Nisbett and his colleagues have shown that people's reasoning can be improved. This tradition starts with the presumption of flaws, but shows that they are not inherent. A second tradition is that people have a repertoire of strategies, and tailor their strategies to the environment. John Payne and his colleagues have argued that people are "adaptive" in this sense.