

correct responses, but also produces more erroneous responses than the other student does, simply by virtue of responding to more items. Which student knows most about the studied material? To answer that question, one cannot look only at errors, or only at accurate responses. Rather, accurate responses and errors must be *jointly* considered to separate the contribution of knowledge from that of bias, willingness to guess, and, perhaps, a preference for some particular alternative (e.g., alternative “C” on a multiple choice) when guessing. Similarly, we have combined correct responses and errors to estimate the contributions of accurate perception and bias. One “bias” that has been of interest to us is a bias toward racial stereotypes.

In one representative paradigm (Payne 2001; see also Lambert et al. 2003), an image either of a gun or of a small hand-tool is presented, and participants are asked to correctly identify the object by pressing the corresponding key (marked GUN or TOOL). In this paradigm – and unlike many social perception experiments – there is an objective criterion for accuracy. Just before the target object appears on the screen, participants are randomly “primed” with a picture of either a Black or a White face. We find that participants are biased to make stereotype-consistent errors. For example, they are more likely to mistake a wrench for a gun when primed with a Black face, as opposed to a White face.

An important element of this paradigm is that it allows use of Jacoby’s (1991) process-dissociation procedure to estimate the relative roles of cognitive control and automaticity in driving behavior. Cognitive control in this paradigm corresponds to participants’ ability to respond to the veridical properties of the target, ignoring information from the nonpredictive racial cues. The other parameter, accessibility bias, is relevant to how participants respond in the *absence* of cognitive control. It is here that automatic reactions come into play, determining whether the *gun* response is likely to be chosen differentially, dependent upon racial cues, when participants are unable to fully control their responses (Payne et al., in press).

In this, as well as other, paradigms it is virtually meaningless to ask whether people are accurate or not. Indeed, we have found that overall accuracy rates can be varied greatly by simply changing the parameters of the task (e.g., giving participants less time to respond). Of greater importance, analyzing the *pattern* of accuracy and errors permits us to address process-level questions. For example, Lambert et al. (2003) used process dissociation in order to shed light on a decades-long debate as to why people show greater reliance on well-learned responses in public settings, otherwise known as a social facilitation effect (Zajonc 1965). Rather than being the result of strengthening of bias (Hull 1943), such effects were caused by a loss of cognitive control.

Whereas the absolute levels of accuracy may change from person to person or from context to context, the basic processes are likely to remain the same, varying only in magnitude. K&F advocate a Bayesian approach as a way of accounting for both accuracy and bias. However, this is a rather descriptive approach, much like the null hypothesis testing it is meant to replace. Process dissociation is one kind of model aimed at quantifying the mental processes at work behind observable outcomes. Other process models, such as signal detection theory, multinomial models, and connectionist models, have recently made entries into the social psychological literature as well. The advantage of process models is that once the basic processes are understood, one can predict and interpret both accuracy and bias naturally from the same underlying framework.

Conclusion. K&F’s article is important and timely, and we are largely, although not entirely, in agreement with their main points. K&F charge that social psychologists have devised clever paradigms that paint people as inappropriately foolish. Should social psychologists endeavor to “balance the score” by devising clever paradigms to show higher levels of absolute accuracy? We are not sure that this represents a productive line of inquiry. Social psychologists should not have to choose between emphasizing accuracy *or* errors. The important question is not whether humans

should be portrayed as noble or foolish. Instead, we might do better to focus on models of the processes driving human judgment, and let the portraits emerge as they will.

People actually are about as bad as social psychologists say, or worse

Michael P. Maratsos

Institute of Child Development, University of Minnesota, Minneapolis, MN 55414. Marat001@umn.edu

Abstract: Experimental studies are not representative of how badly people function. We study people under relatively innocuous conditions, where their self-interests are very low. In the real world, where people’s self-interests are much higher, people are much worse a good deal of the time (some illustrations are cited). This is often “adaptive” for the perpetrators, but that doesn’t make it “good” behavior. That people function so badly in our experiments, where self-interest is relatively minimal, is what is really terrifying.

The overall thrust of Krueger & Funder’s (K&F’s) article is really “are people as bad, morally and cognitively, as social psychologists say or imply?” They want to say no, the present literature is unbalanced. I agree with many of K&F’s analyses of the extant social psychological data; their calls for greater balance and completeness seem well justified. But in some major ways, they are wrong. First, the experiments are unrepresentative, in a way not considered by K&F: In these experiments, very little self-interest is actually at stake for the subjects; in the real world, much more is typically at stake. Consider the subjects in the famous Asch or Milgram experiments (cf. Asch 1956; Milgram 1963; 1974). They won’t have to continue to live with the other people in the experiment afterwards. They won’t receive promotions, demotions, or firings from them; they won’t be accused of heresy or treason or witchcraft by them; they aren’t friends they could lose; they won’t be cast out to starve. What is so shocking about the Asch and Milgram experiments is that there was so much conformity and cruelty, *given how little the subjects had at stake*.

In real life, people have real self-interest and real passions at stake. The results are quite often horrible. I will only cite a few historical and current examples of the multitude available. None of these concern terrible behavior in wars or massacres, or the Holocaust, which might be (wrongly) written off as “exceptions.”

My first example is polygamy: As soon as there were surplus agricultural resources, men in most societies took up hoarding women for themselves, perhaps two or three or four, or more (e.g., harems) if they could. This women-hoarding is “adaptive” for the favored men, but is hard on other men, who then lack mates; it often has made more miserable lives for the women. It is ordinary unkindness.

Also ordinary is the horrible behavior that has been used to control women. Take, for example, the practice of footbinding in China which consisted of crushing, for years, the feet of young girls to keep them small, and unable to sustain walking. X-rays of the feet are horrifying. The practice started with Chinese emperors who wanted to control their harems, but soon spread to prosperous men with multiple wives; it continued to spread throughout society as a requirement for upwardly mobile marriage. By the early twentieth century, in some large areas, *half* the girls were footbound. Everyone accepted the results as “attractive,” and mothers argued it was also *healthy* for the girls (it isn’t). Another example is the practice of clitorrectomy. In modern Africa, millions of girls are clitorrectomized to control them sexually; their mothers claim that it is healthy (it isn’t). And, of course, killing unfaithful wives has been commonly accepted everywhere.

Slavery lasted for centuries in ancient Greece, with very few moral objections from those who benefited; the Church did not declare slavery immoral. Conditions were often horrible; in the

Laurion silver mines, which supported Athenian democracy, male slaves typically died in 2 to 3 years. People's supporting cognitions were appropriately flawed. Aristotle argued that slavery is justified because slaves come from losers in wars, and losing wars shows that the losers are inferior in merit. Aside from whether "merit" should mean merit in warfare, and whether this "merit" should spread to a man's wives, women, and descendants, consider the following: Is every man on a losing side actually inferior in manly "merit" to every man on the winning side? Of course not. By this logic, the great Trojan Hector "deserved" to be a slave to any warrior on the winning Greek side. Aristotle benefited from slavery, and this corroded his reasoning, making it, I believe, "contemptible."

The examples proliferate. Doctors, through the centuries, were one of the best-educated classes, but, as Montaigne wrote, they did not use formal operational thought. For example, for more than two thousand years, doctors followed the practice of bleeding people, which killed many and cured none; during these centuries, no doctors (an educated class) tested whether bled people actually recovered better than non-bled people; no one proposed it, either, apparently. Self-interest (a doctor has to have something to do) impaired cognition, as it always does.

Until unions formed, employers always paid employees as little as possible, just enough to get workers, and to have surviving children as laborers (the "iron law of wages"). When the English government passed laws against children working for a shift longer than ten hours, manufacturers employing child labor invented the "split shift" (e.g., dinner ends one shift; a new one begins). These (usually evangelical Christian) manufacturers generally thought God wanted them to prosper this way. In much of Asia today, if someone is raped, or steps on a land mine, or is a permanent social leper (untouchable), you don't have to pity them, or help them. They did something in a former life to deserve this (Karma); religious cognition obviates the burden of sympathy. On Wall Street, according to Scott Paltrow in "Heard on the Street," scandals occur and will continue to occur because (1) there's no money in playing straight with small investors (commission regulations); (2) there's money in helping big guys; (3) you're almost never caught; (4) big executives nearly always negotiate no punishment for themselves as part of the settlement with the government (e.g., Sandy Weill, Citibank); and (5) small investors are viewed as contemptible suckers who deserve it (Scott Paltrow, "Heard on the Street," *Wall Street Journal*).

Very few westerners ever trouble themselves seriously over the poverty-stricken conditions of the third-world people whose cheap labor helps support their lives.

These and many other everyday things are, of course, all "adaptive" for the perpetrators; but K&F think "adaptive" for the self and its favored associates means, somehow "generally good." This is K&F's second major mistake, one that evolutionary theorists do not make.

I don't rest my case (or refine it; unfortunately, there isn't enough space for that here). Self-interest makes people worse, and the real world is full of it, *much more* so than in our pallid experimental situations; that people commonly act or think so badly in these experimental situations, only adds to the terrible knowledge we have of ordinary people and human nature in the real world.

Proper experimental design and implementation are necessary conditions for a balanced social psychology

Andreas Ortmann and Michal Ostadnický

Center for Economic Research and Graduate Education, Politických veznu 7, 111 21 Prague 1, Czech Republic. Andreas.Ortmann@cerge-ei.cz

Michal.Ostadnický@cerge-ei.cz

<http://home.cerge-ei.cz/ortmann>

Abstract: We applaud the authors' basic message. We note that the negative research emphasis is not special solely to social psychology and judgment and decision-making. We argue that the proposed integration of null hypothesis significance testing (NHST) and Bayesian analysis is promising but will ultimately succeed only if more attention is paid to proper experimental design and implementation.

We do subscribe to the basic message of Krueger & Funder (K&F), that there is a negative research emphasis in social psychology and judgment and decision-making, and that this negative research emphasis hinders theory developments, such as programs that try to understand to what extent seemingly maladapted heuristics in laboratory settings may be quite reasonable in real-life settings (e.g., Gigerenzer et al. 1999).

K&F persuasively lay out the allure of such a negative research emphasis. Indeed, it is much more interesting (and, we submit, on average easier, faster, and less expensive) to generate violations of norms or conventions than to explain why they have arisen in the first place. Although we are as surprised as the authors that the persistent emphasis on norm violations has not yet decisively eliminated its allure, we do see evidence that, at least in psychology, the tide is turning (e.g., Gigerenzer 1991; 1996b; Gigerenzer et al., in press; Juslin et al. 2000; Koehler 1996). The target article strikes us as yet another good example of that encouraging trend.

Curiously, but maybe not surprisingly, although the unbalanced view of humans as cognitive misers seems slowly but surely on its way out in social psychology and judgment and decision-making, the heuristics-and-biases program, which seems mostly responsible for the unbalanced view, has during the past decade invaded economics with little resistance (e.g., Rabin 1998; see Friedman 1998 for an early and lone attempt to stem the tide), amidst outrageous claims. To wit, "mental illusions should be considered the rule rather than the exception" (Thaler 1991, p. 4). Sounds familiar?

It is easy to see why the widespread practice of taking the predictions of canonical decision and game theory as an explicit or implicit null hypothesis (e.g., the predictions of no giving in standard one-shot dictator, ultimatum, or various social dilemma games), has facilitated this development. Although the simplistic rational actor paradigm surely deserves to be questioned – and experimental evidence questioning it has generated some intriguing theory developments recently (e.g., Goeree & Holt 2001) – the rational actor paradigm is often questioned by perfunctory reference to the various "anomalies" that psychologists in the heuristics-and-biases tradition claim to have discovered. This negative research strategy nowadays often goes under the name of behavioral economics and finance.

Alleged errors of judgment and decision-making, such as the overconfidence bias or the false consensus effect (or any other choice anomaly of the list provided in Table 1 in the target article), are taken to be stable and systematically replicable phenomena.¹ Rabin (1998), whose article has become the symbolic reference for most self-anointed experts in the areas of behavioral economics and finance, is particularly explicit about it when he says, "I emphasize what psychologists and experimental economists have learned about people, rather than how they have learned about it" (Rabin 1998, p. 12).

Of course, there is no such thing as an empirical insight per se; each and every empirical result is a joint test of some (null) hypothesis about the behavior of people and of the way the test was designed and implemented. Think of the giving behavior in dicta-