ogy needs a shift in perspective to "a more balanced, full-range social psychology." The purpose of their review is to stimulate such a shift. K&F eloquently argue that there is much research demonstrating a large number of different behavioral and cognitive biases in social cognition. This is true; however, there is also a large body of research meeting their criterion, that is, the need to study a range of behavior and cognitive performance (some of these are presented above). In my opinion, therefore, a "perspective shift" is already in progress. Research and theories have been published, and are continuing to be published, that address normal social cognition and behavior without proposing that we reason or behave in error-prone ways. That said, K&F's article provides a timely reminder that we should seek to understand behavior as a whole and not simply focus on the ostensibly abnormal or unusual.

Authors' Response

Social psychology: A field in search of a center

Joachim I. Krueger[a] and David C. Funder[b]

[a]*Department of Psychology, Brown University, Providence, RI 02912;* [b]*Department of Psychology, University of California, Riverside, CA 92506.*
joachim_krueger@brown.edu **funder@citrus.ucr.edu**

Abstract: Many commentators agree with our view that the problem-oriented approach to social psychology has not fulfilled its promise, and they suggest new research directions that may contribute to the maturation of the field. Others suggest that social psychology is not as focused on negative phenomena as we claim, or that a negative focus does indeed lay the most efficient path toward a general understanding of social cognition and behavior. In this response, we organize the comments thematically, discuss them in light of our original exposition, and reiterate that we seek not a disproportionately positive social psychology but a balanced field that addresses the range of human performance.

In the target article, we argued that modern social psychology is characterized by an abiding preoccupation with troublesome behavior and flawed cognition. We traced this state of affairs to an underlying value orientation that accords primacy to negative phenomena and to the rigid way in which these phenomena tend to be cast in experimental design and statistical analysis. In conjunction, these properties of social–psychological research have impeded the development of theories with explanatory power and the ability to generate novel and nontrivial hypotheses. Our suggestions for a re-orientation were not radical. Instead, we sought to highlight several existing trends in both theorizing and methodology that could benefit the field if pursued more vigorously. Many of the commentators echo our concerns about the history and the current status of the field; they constructively elaborate on many of the proposed remedies, and they suggest new ones. Others defend the traditional view, arguing that social psychology should continue to focus on misbehavior and flawed judgment. We are indebted to all commentators for their carefully reasoned contributions. In this response, we highlight what we perceive to be recurring themes, and we delineate how the commentaries have shaped our thinking. As could be expected, we give more detailed consideration to commentaries that challenge important components of our original argument.

The relevant themes can be organized to parallel the organization of the target article. First, there is the question of *diagnosis*. Because we stressed the importance of studying the accuracy of social perception, it is only fair to ask whether our assessment of the state of the field is itself accurate. Second, there is the question of *methodology*. Our claim that the routine applications of null hypothesis significance testing contribute to the negative outlook turned out to be controversial; comments concerning moderator variables raised pertinent issues; and our proposal that research be oriented to examine the entire range of performance, rather than just the negative end, was in some cases misunderstood. Third, there are issues of *theory* and the kind of research most likely to help theory develop, which lie at the heart of the search for a balanced paradigm.

R1. An accurate diagnosis?

R1.1. Selectivity

There is no consensus among the commentators on whether social psychology is predominantly negative. Although many agree with our assessment that it is (**Hertwig & Wallin**, **Jussim**, **Kihlstrom**, **Ortmann & Ostatnicky**, **Schwarz**), others object (**Darley & Todorov**, **Gregg & Sedikides**, **Regan & Gilovich**, **Petty**, **Vitouch**). Still others feel that there is a negative orientation, but that this is as it should be (**Epley, Van Boven & Caruso [Epley et al.]**, **Friedrich**, **Klar & Levi**, **Shackelford & Vallacher**, **Stolarz-Fantino & Fantino**), or even, that this orientation is insufficiently negative (**Maratsos**). How then is one to arrive at a reasonably accurate negativity score? Database searches for relevant keywords such as *accuracy* or *bias*, of the kind provided by us or by **Kruger & Savitsky**, are only suggestive because they sample across diverse psychological subdisciplines and do not fully capture the impact of individual publications.

Our case for the overall negative orientation of social psychology traced its roots to an enduring ideological commitment that began with the idea that social groups are more likely to corrupt individuals rather than allow them to flourish (e.g., Allport 1924; Le Bon 1895). Although some later work (especially in the Lewinian tradition) examined effective leadership and heightened group performance, these topics faded from view as the cognitive revolution renewed interest in the psychology of stereotyping and prejudice. We also noted some of the rhetoric employed in the literature, which has included the characterization of human judgment as "ludicrous," "indefensible," and "self-defeating." **Regan & Gilovich** claim that in context these particular terms were justified. Besides questioning whether describing human behavior with a term like "ludicrous" is appropriate in *any* scientific context, we would note that these three terms were drawn from a longer list of examples of negative rhetoric. To quote another prominent example, none of the commentators claimed that the comment "How could people be so wrong?" (Ross & Nisbett 1991, p. 139) was either justified or quoted out of context. It would be hard to deny – and we are not certain whether Regan & Gilovich intend to deny – that overall the rhetoric of the heuristics and biases literature has been both remarkably negative and effectively attention-getting.

We further noted that in this literature only negative and not positive effects become reified. The list of biases, errors, mistakes, illusions, and fallacies presented in Table 1 is a sobering illustration. If this sample were unduly biased, it ought to be possible to draw up an alternate list of positive effects. No commentator took the opportunity to do so, and we suspect that this is because it would be futile to try. With rationality (or accuracy) defined as a *point-hypothesis*, there are many ways to detect departures, but none to confirm the null hypothesis.

Finally, we argued that the heuristics-and-biases paradigm, as spearheaded by Kahneman and Tversky, was congenial to the pre-existing negative value orientation in social psychology, and its emergence coincided with the cognitive revolution. The adaptation of the heuristics-and-biases paradigm fueled the search for irrationalities, and the cognitive revolution introduced a variety of new places to look (e.g., automatic responses). The new hybrid paradigm achieved inordinate impact and public recognition as it swept over the field. As shown by the title of their review, Tetlock and Mellers (2002; cited by **Klar & Levi**) credited Kahneman and Tversky with stirring up "the great rationality debate." The work they reviewed was mainly concerned with exposing the limitations of Expected Utility Theory. In this, Kahneman and Tversky had great success, in part because they offered Prospect Theory as an alternative. The assumptions of Prospect Theory enjoy empirical support because they are grounded in well-established principles of the psychophysics of sensation and perception.

The derivative work within social psychology had no such general target, because social psychology lacks any overarching paradigm (such as Expected Utility Theory) that assumes behavior and cognition to be rational. As a result, Kahneman and Tversky's work only lent further ammunition to social-psychological research that was already premised on the pre-theoretical idea that people are socially and cognitively inept. Pinpoint hypotheses of rational thought popped up adventitiously only to be knocked down by the data. But these demonstrations of norm violations could not be regarded as anomalies calling for theory revision because there was no theory to be revised.

Still, many investigators, including some of the present commentators, endorse the idea that the study of anomalies yields the greatest theoretical benefits because errors open windows to the mind (**Epley et al.**, **Friedrich**, **Klar & Levi**, **Kruger & Savitsky**, **Shackelford & Vallacher**, but see **Gigerenzer**). This, indeed, is the key objection to our analysis. Our skepticism about this idea should now be clear. Experimentally demonstrated anomalies are informative only if there is a well-articulated theory mandating ethical behavior or rational thinking, and if the outcomes of experimental tests are not foregone conclusions. Neither condition is typically satisfied in social psychological work. Commonsense expectations of proper behavior and thought often pass for theory, and the sensors of data analysis are tuned to significant departures from narrowly defined norms.

It is astounding with what facility the idea that anomalies are most informative was imported to a field that had no general theory against which specific data could be judged to be anomalous. Although several commentators endorsed this idea in almost exactly the same words, no one presented a compelling case for *why* or *how* research on anomalies yields deeper insights than other research. One would think that this debate offered a splendid opportunity to convince skeptics of the theoretical value of errors such as "The Big Three" (the false consensus effect, the fundamental attribution error, and self-enhancement). If these phenomena open windows to the mind, what do we see when we look through them? That this opportunity for clarification should have been foregone is perhaps the most troubling outcome of the present exchange. Until further arguments are presented, we are inclined to think that the interest in errors and biases is indeed the kind of infatuation that we diagnosed it to be. We cannot resist quoting from Lessing's *Nathan the Wise* on the matter (Lessing 1779/1923, p. 139).

> NATHAN (to his maid Daya):
>
> And yet though it might sound but natural,
> An every-day and ordinary thing [. . .]
> Would it be less of a miracle?
> To me the greatest miracle is this,
> That many a veritable miracle
> By use and wont grows stale and commonplace.
> But for this universal miracle,
> A thinking man had ne'er confined the name
> To those reputed miracles alone
> Which startle children, ay, and older fools,
> Ever agape for what is strange and new,
> And out of nature's course.
>
> DAYA:
>
> Have you a mind
> With subtle instances like this to daze
> Her poor o'erheated brain?

Fiedler claims that "for a scientific contribution to be accepted as original, it has to deviate from established laws," and points to the work of Copernicus, Einstein, and Kahneman as prominent exemplars. According to this argument, demonstrations of errors are both original and theoretically progressive. We believe that they are neither. By definition, deviations from prior beliefs may qualify as original when first observed, but their originality should wear off with repetition (and Bayesian belief revision). This did not happen in social psychology, apparently because of an enduring commitment to the pre-theoretical idea that only the hypothesis of rationality needs refutation. Moreover, Kahneman and Tversky's work was original and impactful in relation to Expected Utility Theory, but the derivative work on social heuristics and biases was accepted precisely because it was *not* original. Instead, it was seen as a great fit with a great idea (a pre-theoretical, poorly justified one at that). The field slumped into a Trotskyite dystopia, in which the conduct of normal science (Kuhn 1970) was mistaken for a permanent revolution.

R1.2. Omissions

We did point to a few exceptions to the predominant value orientation in social psychology, and we thank the commentators who discussed research directions that we had neglected, such as dual-process theories (**Petty**, **Slater**, **Stanovich**), social neuropsychology (**Wood**), social learning (**Stolarz-Fantino & Fantino**), and strategic interaction (**Hodges**, **Kameda & Hastie**, **Kenrick & Maner**, **Kihlstrom**).

Two specific omissions deserve further comment. The first and more general omission concerns cooperative behavior. From the perspective of orthodox game theory, cooperative behavior in non-zero-sum games such as the prisoner's dilemma appears anomalous and irrational. If rationality is defined in terms of self-interest, rational players are expected to defect, which leads to poor outcomes for everyone. Yet, many players cooperate (Komorita & Parks 1995; Sally 1995). **Darley & Todorov** imply that this finding can be counted as a positive contribution of social-psychological research, but we wonder how. Experimental games were first devised and studied by mathematicians and economists (e.g., Luce & Raiffa 1957). Subsequently, investigators in a variety of social-science disciplines observed levels of cooperation that were unexpectedly high from the game-theoretic perspective. The question is whether social psychology has succeeded where game theory failed. We do not think that it has (yet). It is true that some factors have been identified that increase levels of cooperation. Some of these require additional information about the other player. For example, people are more likely to cooperate with an attractive opponent (Mulford et al. 1998) or with someone belonging to a social ingroup (Kiyonari et al. 2000). The former effect seems irrational when observed in the anonymity of a one-shot game. The latter effect can be attributed to generalized expectations of reciprocity. Other effects fall near the boundary of social and personality psychology, such as the finding that people with a prosocial attitude (or disposition) are more likely to cooperate (van Lange & Semin-Goossens 1998). This effect too may be understood as a consequence of expectations of reciprocity (Acevedo & Krueger 2004; Krueger & Acevedo, in press). We think that generalized expectations of reciprocity will be a significant element in any social psychological account of cooperation in dilemmas (Baker & Rachlin 2001; Krueger 2003b; Krueger & Acevedo 2002). Perhaps ironically, this approach owes much of its potential to previous research on social projection, and thus the "false-consensus effect" (Acevedo & Krueger 2004; Krueger & Acevedo, in press; Orbell & Dawes 1991).

Dunning claims that another significant omission is our failure to consider further evidence relevant to Kruger and Dunning's (1999) study of the "unskilled and unaware" syndrome. We accorded this study a prominent place in the target article because it exemplifies many of the troublesome features of the heuristics-and-biases tradition. The focus of the research was unabashedly negative. Participants who performed poorly on a test (any test!) were charged with the dual failure of being unskilled and unaware of it, and the findings were heavily publicized as revealing deep failures of social cognition (see also Dunning et al. 2003). To elaborate on our foregoing discussion, we note four points.

First, regression to the mean guarantees that errors by low scorers will tend be overestimates, whereas errors by high scorers will tend to be underestimates (Krueger & Mueller 2002), but regression works in the other direction as well (Campbell & Kenny 1999). When actual test scores are regressed on the estimates, high estimates are associated with overestimation, whereas low estimates are associated with underestimation. The unskilled-and-unaware pattern is thus equivalent to over- and under-confidence. The latter pattern is well known, raising the question of whether the unskilled-and-unaware pattern is a new discovery.

Second, Kruger and Dunning did not demonstrate mediator effects of individual differences in metacognition. The mediator hypothesis assumes that individual differences in metacognitive skill (i.e., insight into one's own performance) are correlated with test scores. The negative correlation between test scores and estimation errors should then be attenuated when differences in metacognitive skill are controlled. A study testing this hypothesis was unsuccessful (Krueger & Mueller 2002). This may not be surprising because the presumed metacognitive skill involves the ability to discriminate individual test items on which one did well, from those on which one did poorly. Such a skill is a matter of sensitivity, whereas the overall estimate of one's own performance is a matter of bias (or threshold of optimism). These two facets of judgment are conceptually distinct (Swets et al. 2000), and there is no particular empirical reason to believe them to be related (**Lambert, Payne & Jacoby** [**Lambert et al.**]).

Third, **Dunning** asserts that we ignored Studies 3 and 4 in the Kruger and Dunning (1999) article, which he claims established the validity of the mediator hypothesis. Study 3 showed that low scorers do poorly when evaluating the test answers of others. This was construed as an indication of low metacognitive skill, but it is hardly a parsimonious account. The finding may reflect the low scorers' limited ability to solve the test problems. If it cannot be shown that metacognition involves skills that are conceptually distinct from those needed to do well on the test, there is no need to invoke them as variables mediating the correlation between performance and estimation errors. Study 4 showed that low scorers' overestimation errors disappeared when these participants were trained to do well. Kruger and Dunning seemed to realize that they could not train participants to acquire better metacognitive skills. Instead, they manipulated the original predictor variable (i.e., actual performance), which reconfounds test-taking and metacognitive skill even while it restricts the range of the outcome variable.

Fourth, one might ask precisely what normative standard is violated by the "unskilled and unaware" effect. One possibility is that people should predict their own performance perfectly. A standard of perfection is unrealistic, however, if only because both estimated and actual performance scores are inevitably affected by random error components. Interestingly, if the correlation between the two were close to perfect, there would be no need to score psychometric tests. We could simply ask people how they think they did. In an alternative ideal universe, individual differences in test performance could be diminished (as was attempted in their Study 4, when low scorers were trained to score high). In the limiting case, measurement would again become superfluous. More realistically, statistical regression would still assert itself. As Galton put it, "the deviations from the average – upwards towards genius, and downwards towards stupidity – must follow the law that governs deviations from all true averages" (Galton 1892, p. 28).

R1.3. Interpretation

Some commentators, especially **Darley & Todorov** and **Regan & Gilovich**, suggest that we portrayed major social psychological work as more negative than it really is. Particularly with regard to the classic behavioral research, these commentators emphasize how the findings can be construed as showing how normal people try to cope with difficult situations. This perspective is indeed more compassionate than many secondary treatments. Thus, we only partially disagree with these commentators. As we noted in the article, the original researchers went to great lengths to map out the conditions under which social pressures sway people. Nevertheless, the nature of these pressures was typically directed toward non-normative, or even offensive, behavior. Countervailing situational forces (e.g., the liberating effects of accomplices in the Asch paradigm) played the role of reducing the prevalence of negative behavior. In a related vein, **Klar & Levi** maintain that error research implies a compassionate stance because "it is impossible to recognize how remarkable an achievement occasional accuracy is, without first appreciating to what extent human judgment is prone to error." We disagree with this assessment. Accurate judgment is difficult because the context is difficult; the necessary information is often missing, ambiguous, misleading, complex, and confusing. Rather than being prone to error, human judgment is adapted for accuracy under difficult conditions. When information is missing, false, or ambiguous, the accuracy of social judgments suffers without implying the operation of faulty mental processes.

Darley & Todorov claim that work on bystander nonintervention was not meant to imply normative or ethical failure. As we see it, the social-psychological experiments were designed to see if *simple* social situational variables could produce the same kind of nonbehavior seen among the bystanders in the murder of Kitty Genovese. The inaction of these bystanders was a brute fact, and the question was whether it could be attributed to some particular characteristic of them, such as the exceptional callousness of New Yorkers. The studies by Darley and Latané (1968) suggested that this was not so because under certain conditions, prosocial acts could be inhibited among normal, randomly selected people, not from New York. This made the implications more universal and therefore even more jarring. Now, Darley & Todorov suggest that the passive bystanders in the experimental emergency paradigm "had not decided *not* to respond" and remained "in a state of indecision and conflict concerning whether to respond or not." This is perhaps a more sympathetic construal than sheer callousness, but not by a wide margin.

With regard to the Good-Samaritan study (Darley & Batson 1973), **Darley & Todorov** suggest that an unspecified proportion of participants were "in conflict between stopping to help the victim and continuing on [their] way to help the experimenter." From this perspective, all of the participants' behavior was both altruistic and unresponsive, depending on whether the victim or the experimenter was seen as the beneficiary. The "help-as-a-zero-sum-game" hypothesis is only credible when the predicaments of the two parties in need are equivalent. This was hardly the case in the Good-Samaritan study, and the predicaments were even more strikingly different in the Epileptic-Seizure study (Darley & Latané 1968). Moreover, the hypothesis would vitiate the relevance of the bystander studies to the Kitty Genovese murder. In that gruesome situation, who would benefit from a person not calling the police? Darley & Todorov seem to recognize that the two targets of help are out of balance when they suggest that "people can train themselves to resist these forces." Such a decision to improve one's behavior presupposes a valid judgment of who needs help more urgently.

Fiedler raises a question concerning the interpretation of social-cognitive research that parallels **Darley & Todorov**'s argument regarding the classic behavioral work. He suggests that our argument boils down to a one-sided, negativistic interpretation of the findings when, instead, many presumably negative effects can easily be recast as positive. It seems to us, however, that when biases such as overconfidence or false consensus take on a more positive look, it is not a matter of a changed interpretation, but a change in the normative model or the context of the task (see **Gigerenzer** for examples, or Fiedler's [2000] ecological approach to social perception).

Another interpretative issue arose in the wake of the "Merlot metaphor." When it dawned on us that in the situationist paradigm it is paradoxical to see the fundamental attribution error (FAE) as a disposition of research participants, we expected vigorous objections. There were none. Indeed, it is remarkable that of 35 commentaries, *not one* defended the common interpretation of this putatively "fundamental" phenomenon. We suggest this is because the FAE falls apart under close scrutiny, and continues to be famous only because it seldom receives any.

Instead of defending the FAE, some commentators (**Fiedler**, **Kruger & Savitsky**) try to turn the paradox around. After saying that we "point out, correctly, that FAE researchers themselves commit the FAE," Fiedler suggests that *we* "commit the FAE by blaming researchers rather than the scientific situation." What we argued was this: If one considers statistically significant effects to be products of the experimental design, one should credit the experimental situation as the FAE's cause. The attributional zero-sum logic then implies that FAE-making dispositions cannot be the cause. Ergo, attributing the FAE to such a disposition is an instance of the same. Alternatively, one could relinquish the claim that statistically significant effects stem from the power of experimental situations. In the case of the FAE, this would mean that a situation that was designed to eliminate dispositional inferences failed to do so. This too is an inconvenient conclusion for a situationist paradigm. In the Jones-and-Harris design this imbalance means, distressingly, that the researchers had greater success changing behavior (i.e., getting compliance for the writing of counterattitudinal essays) than changing judgments (i.e., from the dispositional to the situational). Our point is that one cannot have it both ways: professing the power of situations, and blaming the dispositions of research participants when that power falls flat.

Fiedler's suggestion that our pointing up this paradox is itself an example of the FAE can mean one of two things. It might mean that all dispositional inferences are paradoxical and thus invalid. We see no support for this extreme view, either in attribution theory or elsewhere. Or it means that the FAE paradox can only be stated at the cost of superimposing a second layer of the same paradox. This view would then, of course, also apply to Fiedler's own suggestion that we too have a disposition to err on this matter. The prospect of an infinite regress here is not an appealing one.

Perhaps, as in the analysis of behavior more generally, it is not fruitful to try to separate dispositional from situational factors when making sense of the conclusions drawn by individuals or groups of investigators. Any particular study simultaneously reflects the beliefs and preferences of individual researchers and the ideology prevalent in the field. We may then agree with **Fiedler** that the problem lies, in part, "in ideological constraints imposed on research" (also **Haslam, Postmes & Jetten [Haslam et al.]**). Ideologies are represented in individual minds, often with a great deal of consensus. They are both personal and cultural. Ross's

(1977) statistical criterion of detecting dispositions notwithstanding, to define the personal only as that which is unique (e.g., Karniol 2003), is too limiting (Krueger 2003a). By Ross's definition, claims about the FAE could be attributed to individual researchers (or research participants) only if these researchers (or participants) held a minority view. If that were the case, the FAE would hardly be regarded as "fundamental."

R2. An "aye" for methods

A central theme of our target article is that the conventional methodology of social psychology can produce misleading results and has helped to create a literature that exaggerates misbehavior and flawed judgment. We offered suggestions for increased methodological sensibility, and several commentators took these ideas further. **Borkenau & Mauer** and **Gosling** elaborate on the need to study individual differences in social judgment, and to take advantage of the Brunswikian framework (**Hammond**). Other commentators propose improved experimental designs (**Ortmann & Ostatnicky**) or model-fitting techniques (**Jussim**). In addition, we would urge greater awareness of the role of random error in judgment and the concomitant regression artifacts (**Gigerenzer**). We are encouraged by the increasing popularity of methods that decompose judgment into theoretically meaningful components (**Lambert et al.**; Swets et al. 2000). The specific appeal of these methods is that they guard against the temptation to treat any significant sign of bias as an indication of inaccuracy (see also Hastie & Rasinski 1987; Wright & Drinkwater 1997). For a recent compendium of further methodological advances (e.g., simulations, connectionist modeling, meta-analysis), we recommend the handbook published by Reis and Judd (2000).

Three specific methodological issues arise from our article and the commentaries. The first concerns the pitfalls of null hypothesis significance testing (NHST) and our suggestions for the refinement of statistical inference. The second issue concerns the role of moderator variables in research on misbehavior and flawed judgment. The third issue is our suggestion that research be reoriented to study the whole range of human performance, not just the negative end.

R2.1. The NHST bugaboo

Our article pointed out several of the increasingly often-recognized pitfalls of null hypothesis significance testing, leading some commentators to respond that NHST "has not caused the current crisis" (**Fiedler**, see also **Chow**, **Goodie**, **Gregg & Sedikides**). But we made no such claim. Instead, we pointed to the consequences of NHST in the context of narrow, unforgiving, point-specific norms of good behavior or rationality. The target article was not the place for a full or decontextualized discussion of NHST. Chow presents some general arguments, and for a full review and discussion of his ideas, we recommend an earlier exchange in this journal (Chow 1998). Gregg & Sedikides downplay the significance of NHST by arguing that the acid test for inductive inferences is the replicability of the findings. We agree, but not totally. As shown elsewhere, the p value obtained from significance testing is a valid cue towards replicability. Its validity increases inasmuch a replication study closely resembles the original work (Greenwald et al. 1996; Krueger 2001). NHST proscribes the use of prior probabilities of the hypotheses, whereas Bayesian analyses make explicit use of them. Thus, we disagree with Fiedler, who suggests that the Bayesian perspective offers no qualitative improvement because different statistical indices (p values) can be converted into each other.

Our main recommendation for reform was to ask researchers to begin to think within a Bayesian framework (see also Rorer 1991). In doing so, we attempted to dispel the view that Bayesian thinking and NHST are incompatible. Often, the two frameworks are viewed as a choice between objective and subjective data analysis. When framed this way, who would not prefer the objective approach? We strove to show that Bayesian calculations of inverse probabilities (i.e., the probabilities of certain hypotheses given the empirical data) can be grafted on the better-known procedures employed within NHST. This idea has been expressed by several writers within social psychology (Krueger 2001; Trafimow 2003) and elsewhere (Cohen 1994; Hagen 1997). Most notably, the American Psychological Association's task force on statistical inference came to endorse this view (belatedly) in its rejoinder to commentaries on its original set of recommendations (Task Force on Statistical Inference 2000).

R2.2. Salvation in moderation?

The search for moderator variables has a long tradition in social psychology (Greenwald et al. 1986). As we noted, the classic work on conformity, obedience, and bystander behavior was designed to identify important situational constraints on the basic effects. **Kruger & Savitsky** extend this approach to the study of cognitive-perceptual biases, suggesting that "so-called contradictory biases typically lead to the investigation of moderating variables." This may be so in some cases, but we suggest that much more common is their safe mutual isolation in independent literatures employing different jargon. For example, we have yet to see research reconciling the "hot hand" with the "gambler's fallacy," or the overuse of stereotypes with the underuse of base rates, and in general the errors in Table 1 are treated as general and not moderated phenomena. Even when moderation is pursued, the results may be less than wholly illuminating.

For example, numerous studies of the false consensus effect (FCE) have explored ways the effect could be made larger or smaller, yet failed to produce a coherent view of how social projection comes to operate in the first place. Instead, the FCE became overdetermined as more and more causal factors were "ruled in" rather than ruled out (Krueger 1998b). The impasse was broken not by the discovery of further moderator variables, but by the introduction of a new normative model with a fresh view of how projection can produce predictive accuracy (Dawes 1989; Hoch 1987). Significantly, no reliable moderator variable was identified that would predict the occurrence of the opposite of the FCE (i.e., the "false uniqueness effect," Krueger 2000b). A similar critique applies to research on the fundamental attribution error (FAE). The more factors are identified that increase people's propensity to make dispositional attributions, the less we know about how and when these attributions might be correct, or what dispositional inferences are good for.

Self-enhancement has also seen some moderator research. When conceptualized as a motive, for example, situational variables can be found that trigger or inhibit it (Shepperd et al. 1996). Kruger (1999) identified task difficulty as a moderator variable. His analysis showed that people self-enhance when tasks are easy, but that they self-diminish when tasks are difficult. This finding suggests that self-enhancement is not an immutable, built-in bias, but can be made to disappear or even reverse depending on the selection of tasks.

The task difficulty effect on self-enhancement can be examined further from a regression perspective. With regard to the "unskilled-and-unaware" hypothesis (Kruger & Dunning 1999), we began with the empirically based expectation that actual performance (A) is positively correlated with estimated own performance (E). It follows that A is negatively correlated with estimation errors (E-A). This analysis holds regardless of whether A refers to individual differences in ability or to differences in task difficulty. As tasks get easier, the probability of success and the probability of underestimation errors increase. At the same time, people self-enhance *more*. How can that be? Kruger (1999) suggested that, taking the subjective experience of task strain as a cue, people infer low and high performance, respectively, from difficult and easy tasks. A second reasonable assumption is that people will project their own experience on the average other person (O) (Kelley & Jacoby 1996). The correlation between these O ratings and actual task difficulty (A), because it is mediated by E, will then be lower than the correlation between E and A. It follows that self-enhancement, when construed as the "better-than-average" effect, or E-O, will be positively correlated with the probability of success (A) (see Asendorpf & Ostendorf 1998 for derivations). In short, opposite patterns can be obtained depending on whether self-enhancement is construed as an overestimation of reality or as a favorable social comparison. What is more, neither pattern reflects a genuine moderator effect, which requires two predictor variables, whose cross products contribute to the prediction of the outcome (Aiken & West 1991).

The question of moderator variables has larger, paradigmatic importance in social psychology. Over the last decade, the study of culture and cultural differences has received increased attention. This is an important and potentially fruitful development. In the context of the present discussion, however, we note that a good deal of research continues to rely on ostensibly established phenomena from the heuristics-and-biases paradigm. Culture is thereby conceived as a moderator variable that tells us whether certain errors or biases are more or less prevalent in one culture or another. Nisbett and his colleagues, for example, found that the effect sizes of many of the standard phenomena rise or fall depending on whether studies are conducted with participants from individualist or collectivist cultures (Nisbett 2003; Nisbett et al. 2001). East Asians show a stronger hindsight bias but a smaller fundamental attribution error than do Americans. Although these findings are intriguing, their interpretation will likely spawn some controversy. Nisbett and colleagues take a relativistic stand with regard to cognitive norms, noting that most are of Western, and specifically Aristotelian, origin. Yet, these norms are used to evaluate the performance of East Asians, whose Taoist or Confucian framework has little use for such concepts of logic or rationality. Thus, it remains to be seen whether cross-cultural work can complement the heuristics-and-biases paradigm as hoped.

R2.3. Taking the Bad with the Good

A few commentators read our paper as prescribing research that "congratulat[es] whatever positivity is out there already" (**Dunning**; also see **Epley et al.**, **Regan & Gilovich**) and even as seeking to establish that "everything's super" (**Kruger & Savitsky**). This is a fundamental misreading of our intention, and while we appreciate that other commentators did recognize that we do not advocate "an imbalanced focus on the 'sunny side' of social behavior" (**Figueredo, Landau & Sefcek [Figueredo et al.]**), some further clarification might be in order. To reiterate, we argued in section 4.1.1 that

> a one-sided research emphasis on positive behavior, perhaps complete with null hypotheses where *bad* behavior represents the null to be disconfirmed, might eventually generate problems parallel to those besetting the one-sided emphasis on negative behavior. We recommend that the *range* of behavior be studied, rather than showing that behavior is bad – or good – more often than people would expect.

In this vein, we agree with **Gregg & Sedikides** that heuristics and biases do sometimes lead to harmful outcomes, but problems arise when these biases are assumed to cause harm in *all* contexts and when correcting them across the board would cause more harm than good. By the same token, some heuristics might be mental *appendixes*, that like the gastrointestinal kind cause occasional harm and no good, but we will move closer to identifying which these are only when we stop assuming (or acting as if) they *all* are. Fine-tuning the analysis of the implications of heuristics is exactly what we would support.

We believe that the suggestion that our target article fails to describe "in more detail and precision what such a [balanced] psychology would look like, even by example" (**Dunning**) seems to overlook section 4, which addresses this topic and occupies almost a third of the article's length. Specific and detailed examples are included in sections 4.1.2 and 4.3.3.1. The central intention of the Realistic Accuracy Model (section 4.3.3.2) is to "point to four specific stages where efforts to improve accuracy might productively be directed." Efforts to improve accuracy, of course, presume that not everything is "super."

A related misunderstanding is evidenced by **Kruger & Savitsky**, who identify what they see as an inconsistency in the prior work of one of us. They note that "whereas Taylor and Brown [1988] emphasized the positive implications of judgmental errors, Funder and colleagues emphasized the negative implications." The Taylor and Brown thesis is based on research showing that people who have positive views of themselves and their prospects generally have good outcomes. Funder's misgivings stem from the characterization of this effect as being due to illusions, when it seems reasonable to expect that well-functioning people will develop positive and *accurate* self-views. The possibility that "adaptive illusions" might not be illusions at all is supported by findings that overly positive "narcissists" have negative outcomes (e.g., Paulhus 1998; Robins & Beer 2001). In a broader perspective, our point of view does not seek to defend inaccuracy; it does in many cases lead us to doubt claims that inaccuracy (a.k.a., "errors," "illusions") has been found.

We take extra space to reiterate this point because we suspect this is precisely where misunderstanding of our intention is most likely to occur. Our critique of the negativity of behavioral and cognitive social psychology is not meant to lead to a new and positive social psychology, but rather to a *balanced* field of research where both across and within studies the entire range of performance would receive attention. This is why, for example, we endorse **Gosling**'s call for more research in real-world contexts – not necessarily because accuracy is more likely in such contexts, but because, unlike in some experimental situations, accuracy is at least *possible.* We believe it is axiomatic – if still not universally appreciated – that demonstrations of the circumstances that promote error will be most informative when they can be contrasted with the circumstances that promote accuracy, and this cannot occur unless both are studied.

R3. Theory

Unlike some other social sciences, social psychology does not have a master theory. Instead, there are numerous theories of small or intermediate range (**Brase**, **Hammond**). There are unifying themes, but these are rather pre-theoretical. We think that there are two reasons why these pre-theoretical commitments have been so influential. One is Asch's observation that "there is an inescapable moral dimension to human existence" (cited by **Hodges**). The other is that a great part of the phenomena relevant to the field are also the kinds of phenomena that people can observe in their own lives and form opinions about. The confluence of the moral dimension and common sense provided an epistemological framework that has given the field a moralistic aspect. At the risk of exaggerating this point, one might say that the move from the classic behavioral work to social-cognitive work on heuristics and biases was a move from the "crisis of conscience" metaphor to a psychology of "dumb and dumber" (**Kihlstrom**).

The present exchange points to a variety of theoretical developments that can go a long way to detach social psychology from its pre-theoretical premises, and thus to restore balance. The Realistic Accuracy Model (RAM) was offered as one way to incorporate both error and accuracy within a single framework, and to extend the analysis of social judgment beyond the "utilization" (cognitive) stage to include interpersonal and contextual influences. **Kameda & Hastie** are correct to observe that RAM remains a basic schematic for research and theorizing, rather than a complete account. RAM claims that accurate judgment can occur only if relevant information is available to a judge, who then detects and correctly utilizes that information. Although this claim adds three steps to the usual analysis of accuracy, and has generated new research, the bare-bones Figure 2 of the target article does not explicitly address how multiple sources of information arrive in an interacting stream, while their meaning changes according to context. But to add these elements would require simply a straightforward extension of the basic model (and a much messier Fig. 2). A more significant omission is any representation of the judge's goals. The purpose of RAM is to explain the necessary steps towards accurate judgment, where accuracy is defined as a veridical correspondence between the distal properties of an object and the understanding of that object achieved by an observer. Whether accuracy itself is a worthy goal, for example, according to a cost-benefit analysis, is a separate (and interesting) issue that at present lies outside of the model.

In forecasting future theoretical developments, a conservative estimate is that some progress will be made by importing theoretical advances from neighboring fields. One example is the increasing popularity and generativity of evolutionary theory. The relevance of this theory was first pursued in the area of mate attraction and selection. As **Kenrick & Maner** show, its influence is now spreading. Similar developments are under way in the area where social psychology overlaps with the psychology of judgment and decision making (**Gigerenzer**, **Hertwig & Wallin**, **Ortmann & Ostatnicky**) and with behavioral economics (Colman 2003; Hertwig & Ortmann 2001). The latter development appears to be particularly promising in focusing attention on the dynamics of strategic interpersonal behavior, as noted by **Kameda & Hastie**, **Hodges**, and **Kihlstrom.**

We do not expect any of the various new approaches to take over the field, but rather, to offer considerable integrative momentum. Perhaps, the coexistence of these approaches will put too much strain on the field, leading researchers and students to feel they must commit themselves to one metatheory or another. We do not think, however, that this is a grave risk. At least one popular introductory textbook was written by three researchers representing strikingly different theoretical orientations (Kenrick et al. 2005). These authors have set an example of how the field can be jointly illuminated by the evolutionary, cognitive, and cultural-behavioral approaches.

In conclusion, we wish to reaffirm our hope that social psychology will benefit from reform. As we hope the target article made clear, we expect that certain epistemological core values can be retained, such as a realist approach to the subject matter, a vigilant dedication to the principles of theoretical coherence and parsimony, and a continued effort to demystify social realities.

References

Letters "a" and "r" appearing before authors' initials refer to target article and response respectively.

Abbey, A. (1982) Sex differences in attributions for friendly behavior: Do males misperceive females' friendliness? *Journal of Personality and Social Psychology* 42:830–38. [DTK]

Abelson, R. P. (1997) On the surprising longevity of flogged horses: Why there is a case for the significance test. *Psychological Science* 8:12–15. [APG]

Abelson, R. P., Leddo, J. & Gross, P. H. (1987) The strength of conjunctive explanations. *Personality and Social Psychology* 13:141–55. [SS-F]

Acevedo, M. & Krueger, J. I. (2004) Two egocentric sources of the decision to vote: The voter's illusion and the belief in personal relevance. *Political Psychology* 25(1):115–34. [arJIK]

Ackerman, P. L., Beier, M. E. & Bowen, K. R. (2002) What we really know about our abilities and our knowledge. *Personality and Individual Differences* 33:587–605. [aJIK]

Adolphs, R. (2003) Cognitive neuroscience of human social behaviour. *Nature Reviews Neuroscience* 4(3):165–78. [JNW]

Ahadi, S. & Diener, E. (1989) Multiple determinants and effect size. *Journal of Personality and Social Psychology* 56:398–406. [aJIK]

Aiken, L. S. & West, S. G. (1991) *Multiple regression: Testing and interpreting interactions.* Sage. [rJIK]

Allport, F. H. (1924) *Social psychology.* Houghton-Mifflin. [arJIK]

Allport, G. W. (1935) Attitudes. In: *Handbook of social psychology*, ed. C. Murchison, pp. 798–884. Clark University Press. [REP]