work, but they fail to appreciate, I think, how high the stakes are in accepting their view of the world. They write enthusiastically about Uexküll, perhaps without fully appreciating that his species-typical subjectivity is different from their own claims about individually idiosyncratic subjectivity. Laboratory settings are created to simplify the complex, to establish precise control of conditions, and to keep out the “extraneous.” The presumption here is that successful laboratory settings will reveal pieces of what might be a complex process clearly and unambiguously in a way that would never be possible outside the laboratory. The analytic tool of the laboratory experiment has probably done more to enable scientific progress than any other aspect of scientific activity (see Horton, 1967, for an argument that the experiment is what most distinguishes scientific thought from folk–traditional thought). The presumption that justifies the experimental approach is that if one can take complex things apart in the laboratory, it is a small step (though one rarely taken in practice) to put them back together outside the laboratory. But if my interpretation of Felin et al. is correct, the task of putting things back together is neither small nor unproblematic. As Bennis, Medin, and Bartels (2010) put it, the world of the laboratory is a “closed system,” whereas the world we actually live in is an “open system.” What works in a closed system may not work in an open one, and what seems like an error in a closed system may be the best people can do in an open one. Consider again, for example, the effects of frames and mental accounts in decision making (Kahneman & Tversky, 1984; Thaler, 1999; Tversky & Kahneman, 1981). Framing and mental-accounting effects are often regarded as cognitive shortcomings, as “mistakes.” But rational decision making might be essentially impossible without such frames and accounts (B. Schwartz, 1986). Decision frames come into their own in open systems. If Fodor’s discussion of what central systems are and how they operate is roughly correct, then we will never understand them by creating an environment, like the laboratory, that distorts their fundamental nature.

The possibilities I raise here could apply quite broadly in psychology. As Gergen (1973) pointed out many years ago, in an article aptly titled “Social Psychology as History,” many of the phenomena that psychologists are most interested in understanding might be largely the province of Quinean and isotropic central systems.

In this time of “fake news” and “alternative truth,” I do not want to be understood as suggesting, nihilistically, that we can never really know anything. We can know plenty, and science has found out a great deal about the things it studies, if not so much about itself. A key reason for the progress that science makes, I believe, is not that science trains its practitioners to see the world as it is. No, I suspect that scientists are just as prone to effects of framing, context, and aspirations as anyone else. What makes science capable of real progress, I think, is that it is public. The community of scientists, in public conversation, corrects the “biases” and “illusions” that each of them has as individuals. Public science makes progress by means of what has been called “the wisdom of crowds” (Surowiecki, 2004)—in this case, highly trained crowds. Like proverbial blind men feeling around the parts of an elephant, scientists, like the rest of us, if not blind, are at least a little myopic. They rely on their colleagues to save them from embarrassment or worse.

In his thoughtful book, Time’s Arrow, Time’s Cycle, Stephen J. Gould (1987) distinguished between processes in nature that are repeatable (“time’s cycle”) and processes that are historical (“time’s arrow”). Gould regarded the theory of evolution as the paradigm case of a science that is essentially historical. As an enthusiastic contributor to geology and evolutionary science, Gould was hardly suggesting that because it was historical, evolutionary theory could not be scientific. What he was suggesting, however, was that to capture evolutionary processes, we needed a different model of science than the one handed down by physics. We needed explanation, not prediction. Exactly the same might be true when it comes to understanding the operation of the central system. And this is what I take to be the broadest implication of the Felin et al. argument. Where does that leave psychology? Not like physics, perhaps, but the science of psychology could do much worse than ending up as a science with the explanatory power of the theory of evolution.

---

Perceiving rationality correctly

Keith Stanovich

No important conclusions about rational thought depend on issues of perceptual theory at the level dealt with in Felin, Koenderink, and Krueger’s (2017) essay. It is true that several important theorists in the heuristics-and-biases literature have used analogies with perception to facilitate the understanding of cognitive biases. The perceptual examples used by Kahneman and others were used to highlight certain cognitive biases, but the implications of the heuristics-and-biases work for the study of rationality in no way depend on any theory of the visual illusions that were used. The arguments about human decision making that have formed the heart of the Great Rationality Debate (GRD) in cognitive science (Cohen, 1981;
Stanovich, 1999; Stein, 1996; Tetlock & Mellers, 2002) stand or fall on their own, independent of developments at this extremely abstract level of perceptual theory.

The authors themselves, on page 14, say that "our arguments about perception may seem abstract and perhaps far removed from practical concerns about the study of rationality." I couldn’t agree more. These arguments about perception are indeed abstract. They are indeed far removed from practical concerns about the study of rationality. This far-fetched link between the literature on rational thinking and the literature on abstract level of perceptual theory seems to be employed here only to provide a seemingly new rationale for the authors to launch a largely redundant critique of the heuristics-and-biases literature.

It is a redundant critique because many of these criticisms have arisen and been dealt with throughout the last three decades of work on heuristics-and-biases tasks and the critiques of them. The Felin, Koenderink, and Krueger essay is backward-looking in that it revives old debates that have been resolved for some time now. The answers to virtually all of these criticisms are contained in the GRD synthesis that has been used in the field for over a decade.

That synthesis derives from works well into their second decade now, including, in chronological order: Stanovich (1999, 2004), Stanovich and West (2000), Kahneman and Frederick (2002), and Samuels and Stich (2004). The synthesis relies on contemporary dual-process theory (Evans & Stanovich, 2013; Kahneman, 2011). It also relies on two decades worth of work on individual differences in rational thought (Stanovich, West, & Toplak, 2016).

The synthesis follows from interpreting the responses primed by Type 1 and Type 2 processing as reflecting conflicts between two different types of optimization—fitness maximization at the subpersonal genetic level, and utility maximization at the personal level. The synthesis acknowledges a point that the critics of the heuristics-and-biases literature have stressed: that evolutionary psychologists have often shown that the adaptive response on a particular task is the modal response on the task—the one that most subjects give. However, that data pattern must be reconciled with another finding often obtained: that lower cognitive ability is often associated with the response deemed adaptive on an evolutionary analysis (Stanovich, 1999; Stanovich & West, 2000; Stanovich et al., 2016). The synthesis of the GRD referred to above argues that the evolutionary interpretations do not impeach the position of heuristics-and-biases researchers that the alternative response given by the minority of (more cognitively able) subjects is rational at the level of the individual. Subjects of higher analytic intelligence are simply more prone to override Type 1 processing in order to produce responses that are epistemically and instrumentally rational.

A point repeatedly made from within the GRD consensus position is that both Type 1 and Type 2 processing lead to rational responses most of the time. Thus, most of the time the outputs from the two systems are in sync, and there is no conflict. The controversy that spawned the GRD from the beginning was the invention of heuristics-and-biases tasks that primed two different responses, one from each of the systems. The assumption behind the current GRD synthesis is that the statistical distributions of the types of goals being pursued by Type 1 and Type 2 processing are different. The greater evolutionary age of some of the mechanisms underlying Type 1 processing accounts for why such processing more closely tracks ancient evolutionary goals (i.e., the genes’ goals) than does Type 2 processing, which instantiates a more flexible goal hierarchy that is oriented toward maximizing overall goal satisfaction at the level of the whole organism. Because Type 2 processing is more attuned to the person’s needs as a coherent organism than is Type 1 processing, in the minority of cases in which the outputs of the two systems conflict, people will be better off if they can accomplish a system override of the Type-1-triggered output (Stanovich, 2004). The response triggered by System 2 is the better statistical bet in such situations, and that is why it correlates with cognitive ability.

What I am calling the GRD synthesis reconciles most of the debates between the heuristics-and-biases researchers and their critics. The GRD synthesis has been around for quite some time now and has been reiterated in the literature many times. This is why it is surprising to see some of the same old shopworn issues coming up again in this essay. The authors keep reiterating the point that heuristics (Type 1 processing) are useful most of the time, often give the normative response, and that they are adaptively efficient (“many of the seeming biases have heuristic value and lead to better judgments and outcomes,” p. 14; “the vast amount of decision making that humans get right receives little attention,” p. 16; “apparent biases might be seen as rational and adaptive heuristics,” p. 16). But, as I noted previously, dual-process theorists have been at pains to state that Type 1 processing is efficacious most of the time and that reliance on Type 1 processing does not always lead to error. Evans and I pointed out that the equation of Type 1 processes with all bad thinking and Type 2 processes with correct responding is the most persistent fallacy in the history of dual-process theory (now reaching its 40th anniversary; Posner & Snyder, 1975; Shiffrin & Schneider, 1977; Wason & Evans, 1975). Likewise, the early originators of the heuristics-and-biases research tradition consistently reiterated that Type 1 processing modes often lead to normative responding and efficient task performance (Kahneman, 2000, 2011). The GRD synthesis long ago gave up this fallacy, so it is surprising to see it reiterated so often here, or used to create a straw man in statements like “the human susceptibility to priming and sensitivity to salient cues is not prima facie evidence of irrationality” (p. 16). Of course System 1 priming is not prima facie evidence of irrationality! No dual-process theorist has ever made this claim. All of the early dual-process theorists (e.g., Posner and Shiffrin; see
above) assumed that priming in the human brain was efficacious, as have all subsequent theorists.

Other critiques in this essay likewise seem to take us backward to old issues long resolved. The end of the essay reads like a Panglossian litany. In the GRD literature, a Panglossian is the type of theorist who tries to close every gap between the descriptive and the normative that is revealed by empirical research (Stanovich, 1999, 2004; Stein, 1996). Such a theorist has many options. First, instances of reasoning might depart from normative standards due to performance errors (temporary lapses of attention and other sporadic information-processing mishaps). Second, computational limitations may prevent the normative response. Third, in interpreting performance, we might be applying the wrong normative model to the task. Alternatively, we may be applying the correct normative model to the problem as set, but the subject might have construed the problem differently and be providing the normatively appropriate answer to a different problem.

All of these (random performance errors, computational limitations, incorrect norm application, and alternative problem construal) are alternative explanations that avoid ascribing subpar rationality to a response—and they have all been extensively discussed in the literature. But numerous theorists have warned that it is all too easy to use the alternative interpretations in an unprincipled, cherry-picked way that makes Panglossianism unfalsifiable. Rips (1994) warned that “a determined skeptic can usually explain away any instance of what seems at first to be a logical mistake” (p. 393). Kahneman (1981) argued that Panglossians seem to recognize only two categories of errors, “pardonable errors by subjects and unpardonable ones by psychologists” (p. 340). Referring to the four classes of alternative explanation discussed above—random performance errors, computational limitations, alternative problem construal, and incorrect norm application—Kahneman noted that Panglossians have “a handy kit of defenses that may be used if subjects are accused of errors: temporary insanity, a difficult childhood, entrapment, or judicial mistakes—one of them will surely work, and will restore the presumption of rationality” (p. 340).

In short, the toolkit of the Panglossian is too large and too prone to be applied in an unprincipled manner. For years, theorists have pointed to the need for principled constraints on the alternative explanations of normative/descriptive discrepancies. Our own work on individual differences (Stanovich et al., 2016; Toplak, West, & Stanovich, 2011; West, Toplak, & Stanovich, 2008) was originally motivated by the need to provide such principled constraints (Stanovich & West, 1998, 2000). Yet the critiques in the last three pages of this essay simply proceed as if these debates had not occurred and already generated a research literature—almost as if we were back in the time of Cohen (1981), at the root of the GRD. As if this were a new insight, we are repeatedly warned about alternative construals:

Furthermore, this alternative theory needs to recognize that many of the simplistic tests of rationality omit important contextual information and also do not recognize that even simple stimuli, cues, and primes can be interpreted in many different ways. (p. 14)

There is a large variety of stimuli that could be pointed to (and proven) but missed by human subjects in the lab or in the wild. But these types of findings can be interpreted in a number of different ways. (p. 15)

As if this were a new insight, we are repeatedly warned about alternative norms:

granting scientists themselves an all-seeing position—against which human decision making is measured. The conventional and even ritualistic use of this null hypothesis has endowed it a normative force. Yet, repeated rejections of this null hypothesis are of limited interest or concern when the normative status of the theory is itself questionable. (p. 14)

Visual illusions reveal that multiple responses, or ways of seeing, are equally rational and plausible. (p. 16)

we argue that even simple stimuli are characterized by indeterminacy and ambiguity. Perception is multistable, as almost any percept or physical stimulus—even something as simple as color or luminance (Koenderink, 2010)—is prone to carry some irreducible ambiguity and is susceptible to multiple different interpretations. (p. 16)

Unmentioned are the constraints on the alternative construals and alternative norms that have been empirically investigated in the years since Cohen (1981). Also unmentioned is the fact that embarrasses many of these Panglossian critiques: Most subjects in heuristics-and-biases experiments retrospectively endorse the Bayesian and subjective expected utility norms that they violate. That is, after responding—usually after failing to override the response that comes naturally (Kahneman, 2003)—subjects choose the correct norm that they were led to violate (Kahneman & Tversky, 1982; Shafir, 1993, 1998; Shafir & Tversky, 1995; Thaler, 1987). When shown the multiple norms that Felin et al. stress repeatedly, subjects are more likely to endorse the Bayesian norm than alternatives (Stanovich & West, 1999). In introducing the collection of Amos Tversky’s writings, Shafir (2003) stressed this very point: “The research showed that people’s judgments often violate basic normative principles. At the same time, it showed that they exhibit sensitivity to these principles’ normative appeal” (p. x). For example, Koehler and James (2009) found that nonnormative “probability matchers rate an alternative strategy (maximizing) as superior when it is described to them” (p. 123).

In short, when presented with a rational-choice axiom that they have just violated in a choice situation, most subjects will actually endorse the axiom. If people nevertheless make irrational choices despite consciously endorsing rational principles, this suggests that the ultimate cause of the irrational
choices might reside in Type 1 processing and the miserly tendency not to override it with Type 2 processing.

Consider framing effects and preference reversals, two of the most researched ways of demonstrating deviations from instrumental rationality (Kahneman, 2011; Lichtenstein & Slovic, 2006). In such problems, subjects often agree in postexperimental interviews that the two versions are identical and that they should not be affected by the wording. In short, preference reversals or framing effects do not represent alternative contextualizations that subjects want to have. Instead, such alternative construals represent mental contamination (Wilson & Brekke, 1994) that the subjects would choose to avoid.

The issue of postexperimental endorsement is just one way of employing the understanding/acceptance assumption in the GRD—that more reflective and engaged reasoners are more likely to affirm the appropriate normative model for a particular situation (Slovic & Tversky, 1974; Stanovich & West, 1999). Subjects actively reflecting on the norms are more likely to indicate the norms they want to follow. Likewise, individuals with cognitive/personality characteristics more conducive to deeper understanding are more accepting of the appropriate normative principles for a particular problem. That is the result of the individual-differences work I mentioned above.

The authors keep reiterating that the extant literature emphasizes bias too much. In fact, there is no way to tell whether there has been too much or too little emphasis on bias. To know that, someone would have to know the exact distribution of benign and hostile environments a person must operate in and the exact costs and benefits of defaulting to Type 1 processing in every single environment (talk about omniscience!). The point (extensively discussed by Kahneman, 2011) is that an attribute-substituting System 1 and a lazy System 2 can combine to yield rational behavior in benign environments but yield seriously suboptimal behavior in hostile environments. A benign environment is an environment that contains useful cues that, via practice, have been well represented in System 1. Additionally, for an environment to be classified as benign, it must not contain other individuals who will adjust their behavior to exploit those relying only on System 1 heuristics. In contrast, a hostile environment for Type 1 processing is one in which none of the available cues are usable by System 1 (causing the substitution of an attribute only weakly correlated with the true target). Another way that an environment can turn hostile is if other agents discern the simple cues that are triggering the cognitive miser’s System 1—and the other agents start to arrange the cues for their own advantage (e.g., in advertisements or the deliberate design of supermarket floor space to maximize revenue).

The Meliorist (see Stanovich, 1999, 2004) supporters of the heuristics-and-biases approach see that approach as ideally suited to studying cognition in the modern world. The beguiling (but wrong) intuitive response in heuristics-and-biases tasks is viewed as a strength and not a weakness. It is a design feature, not a bug. Why? Because the modern world is, in many ways, becoming hostile for individuals who rely solely on Type 1 processing. The Panglossian theorists have shown us that many reasoning errors might have an evolutionary or adaptive basis. But the Meliorist perspective on this is that the modern world is increasingly changing so as to render those responses less than instrumentally rational for an individual. Einhorn and Hogarth (1981) long ago made the telling point that “in a rapidly changing world it is unclear what the relevant natural ecology will be. Thus, although the laboratory may be an unfamiliar environment, lack of ability to perform well in unfamiliar situations takes on added importance” (p. 82).

Critics of the abstract content of most laboratory tasks and standardized tests have been misguided on this very point. Evolutionary psychologists have singularly failed to understand the implications of Einhorn and Hogarth’s warning. They regularly bemoan the “abstract” problems and tasks in the heuristics-and-biases literature and imply that since these tasks are not like “real life,” we need not worry that people do poorly on them. The issue is that, ironically, the argument that the laboratory tasks and tests are not like “real life” is becoming less and less true. “Life,” in fact, is becoming more like the tests! Try arguing with your health insurer about a disallowed medical procedure, for example. The social context, the idiosyncrasies of individual experience, the personal narrative—the “natural” aspects of Type 1 processing—all are abstracted away as the representatives of modernist technological-based services attempt to “apply the rules.” Unfortunately, the modern world tends to create situations in which the default values of evolutionarily adapted cognitive systems are not optimal. Modern technological societies continually spawn situations in which humans must decontextualize information—where they must deal abstractly and in a depersonalized manner with information rather than in the context-specific way of the Type 1 processing modules discussed by evolutionary psychologists. The abstract tasks studied by the heuristics-and-biases researchers often accurately capture this real-life conflict. Likewise, market economies contain agents who will exploit automatic Type 1 responding for profit (better buy that “extended warranty” on a $150 electronic device!). This again puts a premium on overriding Type 1 responses that will be exploited by others in a market economy. The commercial environment of a modern city is not a benign environment for a cognitive miser. To the extent that modern society increasingly requires the Type 1 computational biases to be overridden, then Type 2 overrides will be more essential to personal well-being.

Evolutionary psychologists have tended to minimize the importance of the requirements for decontextualizing and abstraction in modern life (the “unnaturalness” of the modern world that in fact matches the
“unnaturalness” of many laboratory tasks!). For example, Tooby and Cosmides (1992) use the example of how our color constancy mechanisms fail under modern sodium vapor lamps; the authors warn that “attempting to understand color constancy mechanisms under such unnatural illumination would have been a major impediment to progress” (p. 73)—a fair enough point. But what it misses is that if the modern world were structured such that making color judgments under sodium lights was critical to one’s well-being, then this would be troublesome for us because our evolutionary mechanisms have not naturally equipped us for this. In fact, humans in the modern world are in just this situation vis-à-vis the mechanisms needed for fully rational action in highly industrialized and bureaucratized societies.

Thus, the longstanding debate between the Panglossians and the Meliorists can be viewed as an issue of figure and ground reversal. It is possible to accept most of the conclusions of the work of Panglossian theorists but to draw completely different morals from them. For example, evolutionary psychologists want to celebrate the astonishing job that evolution did in adapting the human cognitive apparatus to the Pleistocene environment. Certainly they are right to do so. But at the same time, it is not inconsistent for a person to be horrified that a multimillion dollar advertising industry is in part predicated on creating stimuli that will trigger Type 1 processing heuristics that many of us will not have the disposition to override. To Meliorists, it is no great consolation that the heuristics so triggered were evolutionarily adaptive in their day.

Cues, minds, and equilibria: Responses and extensions

Teppo Felin, Jan Koenderink, Joachim Krueger

We are thrilled to receive thoughtful commentaries on our article from prominent scholars in psychology, cognitive science, decision science, and biology. The commentaries range from highly critical to broadly supportive. We welcome the opportunity to respond to these comments and to highlight extensions and implications of our all-seeing eye argument as it applies to rationality, perception, and cognition.

Space considerations unfortunately prohibit us from addressing the commentaries point by point. Thus, we will largely focus on the commentaries that raise fundamental concerns and counterexamples to our argument (Chater & Oaksford, 2017; Funder, 2017; Gigerenzer, 2017; Nordli, Todd, & Gigerenzer, 2017; Stanovich, 2017). However, we also make passing reference to the commentaries more favorable to our original argument (Noble, 2017; Schwartz, 2017). Our response is organized around three fundamental issues—namely, (1) the problem of cues; (2) what is the question?; and (3) equilibria, $500 bills, and the axioms of rationality.

The problem of cues

Several commentaries argue that we either take our rationality and all-seeing eye argument too far or that the argument does not apply to their particular conception of rationality, cognition, or judgment. Funder (2017), for example, points to the situation-construal literature and argues that situational analysis in fact allows scientists to objectively study perception—thus challenging our all-seeing eye argument. Gigerenzer and colleagues (Gigerenzer, 2017; Nordli et al., 2017) argue that the all-seeing eye assumption does not pertain to the literature on heuristics and ecological rationality. We respectfully disagree with these views.

A straightforward way to illustrate how the all-seeing eye plagues both the situation construal and the heuristics or ecological rationality literatures is to point to the theoretical primacy that is placed on “cues.” Cue-focused approaches necessarily imply an all-seeing eye. The focus on cues within these two theoretical traditions—though other cognitive and psychological theories could also be used as examples—makes them theories of perception, as well. As we discussed in our original article, most theories of rationality and cognition tend to feature implicit meta-theories about perception and observation, and thus the underlying perceptual assumptions deserve careful scrutiny. We first discuss the situation-construal literature, in conjunction with Funder’s commentary, and then discuss Gigerenzer and colleagues’ commentaries and the particular emphases they place on cues, heuristics, and ecological rationality.

Funder and colleagues (e.g., Funder, 2016) build their model of situation construal on a general “model of situation perception” (see Fig. 1 and Table 2 of Rauthmann et al., 2014, pp. 679, 686). The perceptual focus is evident in the focus on cues: The word “cue” is mentioned 92 times in a recent article introducing a taxonomy of situations (Rauthmann et al., 2014). What, then, are cues? For Funder and colleagues “situation cues are physical or objective elements that comprise the environment. They can be objectively measured and quantified” (Rauthmann et al., 2014, p. 680; cf. Funder, 2016). Cues represent the “composition of the situation” (Rauthmann, Sherman, & Funder, 2015, p. 364)—the ecology or environment (cf. Pervin, 1978)—
rationality (Gigerenzer, 2017; Nordli et al., 2017), functional and equilibrium analysis (Chater & Oaksford, 2017), and the literature on rationality and the psychology of reasoning (Stanovich, 2017). Beyond these literatures, the all-seeing eye is also the central assumption among many in philosophy (e.g., Block, 2015; Burge, 2010), vision science (e.g., Geisler, 2011; Ma, 2012), computer science (e.g., Gershman et al., 2015), and economics (e.g., Frydman & Phelps, 2001; Muth, 1961; Thaler, 2016). Perceptual assumptions tend to be deeply hidden within most theories, and of course deeply embedded in the very nature of empirical observation and science itself. Thus, we hope that this debate and set of commentaries will open up further discussion and dialogue, which in turn will allow for productive theoretical and empirical investigations to further our understanding of rationality, mind, and cognition across the sciences.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References


