Naturalistic Heuristics for Decision Making

Niklas Keller
Max Planck Institute for Human Development
Edward T. Cokely
Michigan Technological University
Konstantinos V. Katsikopoulos
Odette Wegwarth
Max Planck Institute for Human Development

ABSTRACT: Over the last 20 years, both naturalistic decision making and fast and frugal heuristics programs have radically broken with mainstream decision science, moving beyond the confines of artificial tasks and safe academic laboratories. We document commonalities of these programs and discuss ways in which a synthesis could contribute to a more relevant, precise, predictive, and effective decision science. We begin by reviewing the common roots and philosophies of the two programs, such as their respect for the capable decision maker and their acknowledgment of the importance of task ecology. We then identify four specific areas of synergetic potential, including ecological rationality and metacognition. Our review culminates in a case study of naturalistic heuristics based on a particular class of fast and frugal heuristics. These fast and frugal trees provide examples of effective, well-specified decision-making algorithms applied in a naturalistic domain: emergency medical diagnosis. By leveraging the strengths of each program, we point out some of the ways in which more sustainable progress can be fostered on issues that matter the most—for example, decisions that save and transform lives.

Introduction

OVER THE LAST 2 DECADES, TWO RESEARCH PROGRAMS HAVE MADE MAJOR CONTRIBUTIONS to the understanding of decision making in the world. These programs have radically broken with mainstream traditions that are limited to "small worlds" in the artificial confines of simple and safe academic laboratories. The first program, naturalistic decision making (NDM), is the focus of this special issue reviewing 20 years of decision-making research in the most taxing environments, including firefighting (Klein, Calderwood, & Clinton-Cirocco, 1986), critical care units (Crandall & Getchell-Reiter, 1993), and combat information centers (Kaempf, Wolf, & Miller, 1993). The second, called the fast and frugal heuristics (FFH) program, has documented the many ways in which simple decision algorithms can make people smarter, safer, and more efficient in medical, legal, engineering, and other applied domains (Galesic, Garcia-Retamero, & Gigerenzer, 2009;

ADDRESS CORRESPONDENCE TO: Niklas Keller, Max Planck Institute for Human Development, Centre for Adaptive Behaviour and Cognition, Lentzeallee 94, 14195 Berlin, Germany, nkeller@mpib-berlin.mpg.de.

Journal of Cognitive Engineering and Decision Making, Volume 4, Number 3, Fall 2010, pp. 256–274.
DOI 10.1518/155534310X12844000801168.

Gigerenzer & Engel, 2006; Gigerenzer, Todd, & the ABC Research Group, 1999; Katsikopoulos, 2009; Wegwarth, Gaissmaier, & Gigerenzer, 2009).

In what follows, we suggest that after decades of largely independent development, the two programs can learn a great deal from each other. We begin by providing an introduction to each program and a brief overview of their origins and commonalities. We next discuss differences with special emphasis on those that harbor synergetic potential (i.e., aspects that can be used to advance joint goals). Finally, we present a detailed case study of *naturalistic heuristics:* fast and frugal decision trees.

Fast and Frugal Heuristics

The FFH program investigates "simple heuristics that make us smart" (Gigerenzer et al., 1999). The program places emphasis on identifying the repertoire of heuristics (i.e., simple decision algorithms or rules of thumb) that people have at their disposal for adaptive decision making in this fundamentally risky and uncertain world. We define a heuristic as a computational algorithm that can be implemented in an information-processing system (human, animal, or machine). The approach takes opposition to the popular notion that heuristic decision processes are a dangerous cause of systematic error that must be stamped out. Instead, it embraces the analytically provable fact that no single decision process can always guarantee a good solution, given real-world constraints on computation and complexity. Thus, the investigation of superior decision making hinges on understanding the fit among task goals, contexts, and decision processes.

Just as there is no single effective decision-making strategy for all problems, there is no single normative benchmark for all environments. Hence, FFH analyses specify a normative benchmark relative to the structure and constraints of task ecology. This benchmark reflects the *ecological rationality* of a decision mechanism—that is, the way in which a particular decision-making heuristic can exploit the fit between the structure inherent in an environment (e.g., redundancies among cues) and the mind's evolved and developed cognitive capacities (Gigerenzer & Brighton, 2009; Gigerenzer & Selten, 2001).

In investigating ecological rationality, the FFH program has prioritized both the precision and predictive accuracy of its models (Marewski, Schooler, & Gigerenzer, 2010). The precise definition of a decision-making strategy (e.g., formalization as a computer program) fosters transparency and facilitates cumulative theory building and testing. The focus on precision is also a result of the desire to rectify the bad image of heuristics: Within the heuristics and biases program (Tversky & Kahneman, 1974), regions of good and bad performance were never defined, which led to misunderstandings as well as vague and untestable theory (Gigerenzer, 1996).

To avoid such obstacles, FFH specifies heuristics at the level of their constituent processes: (a) rules for information search, (b) rules for stopping search, and (c) decision rules. We refer to these as the building blocks of a heuristic. To date, the program has identified a number of ecologically rational (i.e., well suited to specific ecologies) and psychologically plausible decision-making processes that can be used to engineer decision aids or design better decision environments (Cokely, Schooler, & Gigerenzer, 2010; Todd & Gigerenzer, 2007).

One extensively studied heuristic is called *take-the-best* (Gigerenzer & Goldstein, 1996). When asked which of two objects has a higher value on a criterion (e.g., "Which of two stock portfolios will give me higher returns 1 year from now?") and the exact values are not immediately available, people are assumed to go through a number of cues (e.g., "Do I recognize the company names in the portfolio?") that might serve as indicators to the criterion. If people are well attuned to the information structure in their environment, they will search through these cues in the order of their validity (i.e., a measure of cue goodness). This is take-the-best's search rule. Take-the-best searches through this cue order until it finds the first cue that discriminates ("I recognize the companies in Portfolio A but not in Portfolio B."). This is take-the-best's stopping rule. Put simply, take-the-best chooses the object that was discriminated ("Choose Portfolio A"). This is take-the-best's decision rule.

The assumption that humans are sensitive to environmental frequencies is a well-studied phenomenon (e.g., Hasher & Zacks, 1979). Reference to this capacity in conjunction with the fact that only one cue is deliberated upon at a time makes take-the-best psychologically plausible. The effectiveness of the heuristic relies on its fit to the environment (e.g., accurate ordering of cues, sufficient cue validity), thus defining its regions of good and bad performance—its ecological rationality.

Naturalistic Decision Making

NDM is concerned with decisions outside the laboratory, investigating the most taxing and complex environments. Early NDM research drew heavily on work in expertise and skill acquisition (e.g., de Groot, 1946/1978), and this continues to be a primary focus of the approach. For example, the intuitive situation-action matching or recognition processes (Lipshitz, 1994) that come about as a result of expertise form the core of one flagship theory: recognition-primed decision making (Klein, 1993). NDM is interested in the macrocognitive processes that lead expert decision makers to recognize effective courses of actions, rather than merely investigate response (outcome) deviations from prespecified norms.

Moreover, NDM research stresses that decision accuracy (e.g., hit rate) is not all that is important in a decision (Hoffman & Militello, 2008). Being able to reach consensus among team members or to hedge against worst-case scenarios may negatively impact a decision strategy's accuracy in one particular case. However, such aspects are nonetheless desirable in many complex situations. A decision takes place in a dynamic context in which decision makers interact with a rapidly changing environment. Because of this, many NDM theories describe parallel processes.

Given complex actor-environment interactions, as expert knowledge is highly domain specific, and stressing that decision makers are sensitive to semantic as well as syntactic content, NDM develops *context-bound informal models* (Lipshitz, Klein, Orasanu, & Salas, 2001a). Critically, NDM's primary goal is applicability: improving decision making where it matters. Embracing *empirically based prescription*

(Lipshitz et al., 2001a), it investigates best practices from experts in the field and develops training and decision aids based on field research. Indeed, in this regard, NDM is a world leader (Klein, 1998; Lipshitz et al., 2001a).

An illustrative example of an NDM theory is recognition/metacognition (R/M). In R/M, "recognitional" activation of expectations and associated responses are supervised by metacognitive processes of "critiquing" and "correcting." A higher level process called "quick test" considers factors such as available time, the costs of an error, and the degree of uncertainty or novelty in a situation to judge whether correcting and critiquing are likely to be effective and/or feasible. Note that although all these processes occur in parallel, they are not assumed to be simultaneously available to consciousness. R/M processes are highly dynamic: Problems exposed by critiquing lead to corrective steps, which involve modification and elaboration of situations and plans, which can lead to detection of further problems, and so forth. R/M has been developed to describe the cognitive processes of expert decision makers when they need to go beyond their recognitional capacities to solve a problem. This was done by empirically based description of the decision makers in their natural decision context (e.g., ship-based tactical defense; Cohen, Freeman, & Wolf, 1996).

Origins and Commonalities

With the work of von Neumann and Morgenstern (1947) and Savage (1954/1972), subjective expected utility (SEU) became the dominant normative framework in the domain of decision making under risk after World War II. In the same year that Savage's SEU was published, Meehl (1954) published his influential critique of clinicians' judgments in comparison with multiple regression-based "actuarial" methods. Combining these two lines of thought, Howard (1966, 1968) developed decision analysis to provide prescriptive guidelines for decision makers in order to adhere to the norms of SEU.

However, in the 1970s data emerged indicating that decision makers seemed largely incapable of behaving according to either normative concepts or prescriptive aids. Accordingly, the influential heuristics and biases program (Tversky & Kahneman, 1974) documented an extensive map of deficiencies in "rational" thinking. This map, in turn, might provide directions for decision analysis. Together, these programs offered the hope that scientists could identify, catalogue, and repair the limited and flawed nature of human decision making. However, things were not as straightforward as that. People and organizational teams were often resistant to "debiasing" attempts (Fischhoff, 1982). As Dawes, Faust, and Meehl (1989) lamented 35 years after Meehl's (1954) initial call for greater reliance on decision support tools, many of the interventions had little impact and were not effectively adopted by professional decision makers.

In part, both NDM and FFH programs grew out of frustration with these failed interventions and distorted visions of people's capabilities. Both programs objected to the view of human decision making as fundamentally irrational and flawed, and both argued that the efficacy of intuitive aspects of decision making had been badly distorted by separating the laboratory from the natural ecology. Whereas NDM responded to the growing need to improve decision making in fast-paced, dynamic, high-stress, and high-stakes environments, the FFH program began to build precise models of simple heuristic processes, mapping out their environmental constraints and ecological rationality. Despite these different paths, however, the two approaches retain core commonalities.

The Capable Decision Maker

Human decision making is not essentially flawed and limited. Instead, FFH and NDM have documented how human decision making can be highly effective. At the time of their emergence 20 years ago, there was a huge gap in the literatures on judgment and decision making (JDM) and expertise (Shanteau, 1992). Spearheaded by the heuristics and biases program, research in JDM painted an almost uniformly abysmal picture of human decision making, both for novices and experts. Quite to the contrary, much of the expertise literature in cognitive science marveled at experts' superiority over novices in nearly every aspect of domain specific cognitive functioning, particularly when expertise was verified rather than assumed (Anderson, 1982; Chi, Feltovich, & Glaser, 1981; Ericsson, Prietula, & Cokely, 2007; Klein & Weitzenfeld, 1982; Lesgold, Feltovich, & Glaser, 1980). Having been used in the study of experts from the very beginning, NDM has adopted the expert literature's admiration of the efficacy of skilled human decision making (Lipshitz et al., 2001a; Zsambok, 1997).

The FFH approach extends the notion of the capable decision maker, arguing that laypeople can also be well attuned to the environments with which they interact (Gigerenzer, 1996; Gigerenzer, Hoffrage, & Kleinbölting, 1991). The assumption of a capable decision maker was key to the success of the theory of probabilistic mental models (PMMs)—a theory of overconfidence (Gigerenzer et al., 1991). At the time, overconfidence had been extensively studied following a heuristics and biases approach and was considered to be extremely robust and resistant to debiasing (Fischhoff, 1982). A typical way to study overconfidence was to ask participants a series of general knowledge questions, subsequently ask what percentage they thought they got right, and then compare that percentage with the percentage of correct answers.

The PMM theory is similar to the take-the-best heuristic: If direct knowledge is not available, people construct a reference class over which probabilistic cues are defined and searched in order of validity. People's overconfidence depends on how representative the sampled stimuli are of the constructed reference class. Gigerenzer et al. (1991) showed that in their drive to catalogue human deficiency, many previous studies on overconfidence had sampled unrepresentatively difficult or misleading questions. When representative questions were sampled, the overconfidence effect disappeared. When unrepresentatively easy questions were sampled, the effect could even be reversed, inducing people to display underconfidence.

Ecology Is Crucial

A major criticism of the traditional approach to human decision making, championed by both NDM and the FFH programs, is the disregard for task ecology in laboratory studies. In most of the decision sciences, concern for selecting representative stimulus samples has lagged far behind the concern for selecting either representative participant samples (e.g., Einhorn & Hogarth, 1981) or tractable laboratory problems (Klein, 1998). NDM has tackled these issues head-on, using and developing methods of cognitive field research (Hoffman & Militello, 2008) to observe cognition where it actually takes place in the world and where there is no question as to the representativeness of the environmental stimulus. The FFH program took a more theory-driven approach. It attempted to bring representative stimulus samples reflecting the underlying structure of a particular task ecology into the laboratory (e.g., the statistical structure of environmental cues), where environments and decision algorithms could be subjected to extensive, controlled analyses. Both methods have their virtues, and both have proven effective in confronting and revealing serious limitations of other nonecological methodologies.

Differences and Synergy

No single approach can cover every facet of the decision sciences. Even the most innovative methods and approaches have their limits. Whether or not criticisms are merited, major concerns should be acknowledged, examined, and addressed. We believe that such efforts can be greatly helped by utilizing the synergetic potential between the study of the FFH program and NDM. We argue that at least four areas hold such synergetic potential:

1. using ecological rationality as a normative benchmark for NDM;
2. emphasizing that dynamic processes (e.g., situation awareness, macrocognition) are essential for an understanding of heuristic selection;
3. using FFH's theoretical and methodological tools to specify and test NDM models with greater precision; and
4. using NDM methodology and theory to provide orientation when performing field studies concerning FFH.

Ecological Rationality as a Normative Benchmark for NDM

Ecological rationality is the normative concept of rationality used by the FFH program (Gigerenzer & Brighton, 2009; Gigerenzer & Selten, 2001; Gigerenzer et al., 1999). It refers to the fit between evolved (and developed) cognitive capacities, the heuristics that build on them, and the structure of the environment in which these heuristics are applied. Perhaps one of the most exciting and unexpected findings in the last 100 years of research in the decision sciences is the *less-is-more* effect—a crucial piece of evidence favoring ecological rationality vis-à-vis classical rationality as a normative principle. By analyzing the fit between cognition and environment, research has shown how, where, and why simple heuristics can provide not merely good enough but superior decision performance, outperforming even the most sophisticated optimization processes (Gigerenzer & Brighton, 2009; Gigenerzer et al., 1999). This can be well illustrated with the recognition heuristic (Goldstein & Gigerenzer, 2002).

The recognition heuristic simply states that if one object is recognized and the other is not, infer that the recognized object has a higher value on the criterion. In the classic study of the heuristic, two student populations, German and American, were asked which of two cities was larger: San Diego or San Antonio. Of the American students, 66% gave the correct answer (San Diego). Of the German students, almost all of whom had not heard of San Antonio, 100% gave the correct answer. As absence of information is not unstructured, the German students could make use of the fact that they did not recognize one of the cities to infer that it is smaller. The American students' greater knowledge (most had heard of both cities and knew additional information about them) actually interfered with their accuracy.

Although the city task may seem artificial, it provides a simple and powerful existence proof of the less-is-more effect. The recognition heuristic has successfully been extended to multialternative inference, used to predict political elections (Marewski, Gaissmaier, Schooler, Goldstein, & Gigerenzer, 2010) and the outcome of sports events (Serwe & Frings, 2006), and shown to beat far more complex models in stock portfolio selection (Gigerenzer & Brighton, 2009).

All previous interpretations of rationality operated under the assumption of an effort-accuracy trade-off: More computation, time, and information were always assumed to be better; simplifying heuristics could only provide approximate, second-best solutions. The less-is-more effect proves that improved decision accuracy is possible under conditions of less information, computation, or time (Gigerenzer, 2008). Note that "less is more" does not mean "the less, the better." It refers to the fact that there is often a point beyond which more knowledge or computation will not result in increased accuracy and may even result in decreased accuracy. Again, decision performance is about the fit between cognition and environment. Identifying when and under what conditions less-is-more effects occur is an essential part of research conducted within the FFH program.

In real-life decision making one often has to infer future states of the world from sparse, noisy information samples. Within the FFH program, "robustness" refers to small losses in performance (e.g., accuracy) when an inference has to be made for previously unencountered tasks. Many heuristics have been shown to be robust and outperform more complex models traditionally deemed as "optimal" decision-making benchmarks (multiple regression, classification and regression trees, naive Bayes, etc.), particularly when generalizing from limited amounts of data. These more sensitive algorithms (e.g., given their higher number of parameters) may overfit the available sample and fit noise rather than real patterns in the data.

Recently, insights from artificial intelligence have helped to specify under which conditions less-is-more effects can be expected to occur and their relation to overfitting. The prediction error of an algorithm, after having attempted to learn an underlying (true) environmental function from data samples, has three sources: bias, variance, and irreducible noise. Even a completely unbiased algorithm can still suffer from high variance error, and even a more biased algorithm can increase overall predictive accuracy by reducing the variance component of error (Gigerenzer & Brighton, 2009). Depending on the true underlying environmental function and the number of samples available, an organism can be better off with an induction process that is biased but avoids excessive variance.

The NDM community's decision to reject the notion of normativity altogether and opt for expert decisions as the benchmark has received criticism, as previous work has shown that experts can err in many situations (e.g., LeBoeuf & Shafir, 2001). A counterargument is that traditional decision-making research had been almost exclusively based on college freshmen. NDM has also responded by (a) pointing out that NDM studies do not take expert opinion at face value (Lipshitz, Klein, Orasanu, & Salas, 2001b); (b) specifying the necessary conditions for expertise to occur in greater detail (Kahneman & Klein, 2009—note that this is conceptually similar to specifying regions of good and bad performance of heuristics within the FFH program); and (c) pointing to its substantial track record of real improvements in decisions (e.g. Klein, 2008). The absence of any normative criteria for decision making in and of itself, however, remains a theoretical concern.

Investigating the ecological rationality of NDM models requires further specification of the conditions under which they are applied as well as the cognitive capacities upon which they build. But it would also allow NDM researchers to predict their models' performance across varied environments and test them against other, "rival" decision-making models or classical "optimal" benchmarks (Gigerenzer & Brighton, 2009).

Dynamic Processes Are Essential for an Understanding of Heuristic Selection

Metacognitive processes (e.g., cognitive regulation of performance via monitoring and control processes; Flavell, 1979) deeply influence both the development and the expression of problem solving and expertise. For example, metacognition is a crucial aspect of deliberate practice, which is widely accepted as a prerequisite for the development of superior and reproducible expert performance (Ericsson, 2006; Ericsson, Krampe, Teschromer, 1993). It is also the hallmark of expert performers' ability to "know what they don't know"—that is, to have an acute awareness of the limits of their knowledge and its applicability within or removed from a particular context (Kahneman & Klein, 2009). The NDM community has extensively studied sensemaking (Klein, Moon, & Hoffman, 2006a, 2006b), situation awareness (Endsley, 1995a, 1995b; Zsambok & Klein, 1997), and other metacognitive processes such as critical thinking (Cohen & Freeman, 1997; Cohen et al., 1996) and team decision making (Cannon-Bowers, Salas, & Converse, 1993; Salas, Cooke, & Rosen, 2008).

Understanding metacognition is highly relevant for the study of heuristics. For example, in a test of the recognition heuristic, Oppenheimer (2003) presented people with pairs of cities of which one was familiar but, because of reasons unrelated to their size (e.g., Chernobyl), another was unfamiliar but could be assumed to be larger (it had a Chinese name). The participants did not apply the heuristic. Thus, even if the recognition heuristic builds on a relatively simple process, its effective use involves an awareness of how dynamic mind-environment interactions change the validity of cues or strategies.

Indeed, such metacognitive processes are necessary for successful heuristic application and performance. Schooler and Hertwig (2005) have shown how fluency (e.g., the speed of memory retrieval in an ACT-R framework) can provide an accurate basis for judgment. Recognition and memory-retrieval dynamics reflect a crucial aspect of the cognitive capacities of humans. Knowledge and dynamic interaction with such capacities provides a basis for the evaluation of the validity of particular cues. Similarly, other fast and frugal research has examined how memory and decision strategies shape judgment and alter cognitive abilities (Cokely & Kelley, 2009; Cokely, Kelley, & Gilchrist, 2006; Gaissmaier, Schooler, & Rieskamp, 2006).

Any multiple-strategy approach brings with it a *strategy selection problem:* how to decide which decision strategy to use. This problem is currently a major focus of investigation within the FFH program (e.g., Gaissmaier et al., 2006; Rieskamp & Otto, 2006). The extensive experience NDM researchers have with metacognitive processes offers the opportunity to better inform aspects of strategy selection in the natural world that are essential to a science of heuristics.

FFH's Theoretical and Methodological Tools Specify and Test NDM Models With Greater Precision

The heuristics and biases program has been criticized for having poorly defined and unspecified heuristics (e.g. Einhorn & Hogarth, 1981; Gigerenzer, 1998). This underspecification enabled the same heuristics to explain (but not predict) opposing, inconsistent phenomena (e.g., the representativeness heuristic; see Ayton & Fisher, 2004). The use of vague decision biases that explain everything and nothing was also a concern raised by Klein (1989, 1998) in response to the official evaluations of the *Vincennes* incident of 1988 (Fogarty, 1988). To avoid the major issues associated with vague theory, the FFH program emphasizes precise, computational models of heuristics that can be submitted to extensive testing and analysis (see Gigerenzer & Brighton, 2009; Marewski, Schooler, & Gigerenzer, 2010).

One standard of the fast and frugal heuristic program is specification of the building blocks of a heuristic (i.e., information search, stopping, decision). Although there have been efforts to formalize some NDM theories as computational models (e.g., Warwick, McIlwaine, Hutton, & McDermott, 2001), there is little if any work attempting to clarify the major NDM theories in terms of their building blocks or their ecological niches. Consider R/M: The "quick test" within the model surely is a heuristic process that could be fleshed out usefully. For example, the order in which the important factors (time, stakes, novelty) are "ticked off" and how knowledge, goals, and environmental factors (e.g., missing information) influence this order could be specified. Also, the devil's advocate technique or the STEP (construct a story, test, evaluate, and plan) procedure, which was developed from the R/M model (Cohen et al., 1996), at times run the risk of reverting to old, implausible, and—in light of the findings we noted previously on the less-is-more effect and robustness—possibly ineffective ideals of full information search and integration (Todd & Gigerenzer, 2001).

Greater specification allows more detailed reference to evolved cognitive capacities and environmental constraints, and therefore it is crucial if one wishes to map out the ecological rationality of decision processes. Furthermore, it opens up a new set of methodological tools: competitive testing of multiple postulated models under varying environmental conditions or the possibility of making specific predictions (e.g., regarding information search times). The use of such methods may help further address concerns that NDM studies suffer from low experimental control (Jungermann, 2001; Roelofsma, 2001) or produce only data that favor NDM models (e.g., Caverni, 2001; Jungermann, 2001; but see Lipshitz et al., 2001b, for a discussion).

NDM Methodology and Theory Provide Orientation When Venturing Into the Field to Study FFH

After disappointing progress on the development of decision aids and training methods based on formal standards, NDM researchers made major contributions by moving from theory-based laboratory evaluations to careful and precise fieldwork. In the process, they developed highly effective methodologies for the explication of cognitive processes, such as the critical decision method (Crandall, Klein, & Hoffman, 2006; Klein, Calderwood, & McGregor, 1989), the knowledge audit (Klein & Militello, 2005), applied cognitive task analysis (Militello & Hutton, 1998), and goal-directed task analysis (Endsley, Bolte, & Jones, 2003), which cover complementary aspects of proficient decision makers' domains of work and decision strategies.

Identifying a useful and ecologically rational heuristic for a particular decision initially requires a significant amount of information (Gigerenzer et al., 1999). Particularly in dynamic environments, it will be at the very least uneconomical, and often impossible, to model all relevant decision-making processes and their interactions at the level of FFH. The cognitive field research (CFR) methods developed and employed by NDM could greatly help here by, for example, identifying vital cues or "leverage points" (Klein, 1998; Klein & Wolf, 1998). What information is accessible to the decision maker at critical decision nodes? What are the motivations and goals? What alternative courses of action exist to achieve these goals?

CFR methods also provide important information on aspects of decision making that the FFH program has so far investigated comparatively less (e.g., metacognitive functions and group decision making). The bird's-eye view of CFR methods eases the appreciation of dynamic interactions between choices made and the environment. Currently, many of the scenarios tested and decision rules postulated for FFH are akin to one-shot situations. They do not necessarily acknowledge a decision maker actively influencing his or her environment or using his or her experience to shape it toward a state that allows for implementation of a preferred action.

At this level of analysis, the classical distinctions between the different postulated types of mechanisms for strategy selection—decision/metacognition, learning, and context—become porous. Multiple mechanisms of strategy selection can therefore constrain the application of decision-making processes within a single meaningful event for the decision maker. CFR methods could reveal how conflicts and interdependencies among goals, available information, and alternative courses of action shape the heuristics used or constrain the selection of particular heuristics over others.

What is the promise of a science that synthesizes both programs' methodologies and theoretical strengths to assess critical decisions? In the next section, we present a family of fast and frugal decision mechanisms that we believe harbors major opportunities for integration: fast and frugal trees. These algorithms have already been applied to a number of fields, including biology, educational training, engineering, medicine, and legal decision making (for an introduction to the construction of fast and frugal trees, see Martignon, Katsikopoulos, & Woike, in press).

Naturalistic Heuristics Via Fast and Frugal Trees: A Case Study

Doctors are regularly faced with a difficult decision: Does a patient with intense chest pain have a high or low risk of heart failure? In one Michigan hospital, the doctors sent about 90% of patients with this symptom to the coronary care unit. This was a "defensive" strategy that was believed could only benefit patients while simultaneously protecting doctors against the risk of lawsuits (Gigerenzer & Engel, 2006). However, the resultant overcrowding of the coronary care unit led to excessive costs, decreased the quality of care provided, and introduced risks (e.g., secondary infections) for patients who should not have been there.

To solve this problem, physicians were trained to use the Heart Disease Predictive Instrument (HDPI), which consists of a chart with more than 50 probabilities and a pocket calculator with a logistic regression program. Unsurprisingly, physicians were unhappy using this instrument, and it was never successfully adopted. It was regarded as too opaque and appeared to make clinical judgment dependant on a pocket calculator. Subsequently, Green and Mehr (1997) constructed a fast and frugal tree as decision aid, with much better results. For one, the doctors found it very easy to internalize, freeing them of the need to use physical decision aids. In addition, however, the resultant accuracy in classification was equal to the optimal use of the HDPI.

Given the complexities of this situation, at the core of this decision lays a categorization task: Is the patient at high risk or low risk? A categorization tree can be graphically represented by the root node, the tree's first level, and subsequent levels with one cue processed at each level (see Figure 1). There are two types of nodes: nodes in which a question is asked about the value of the object on a cue, and exit nodes, at which point the object is categorized and the process stops. A categorization tree is fast and frugal if and only if it has at least one exit at each

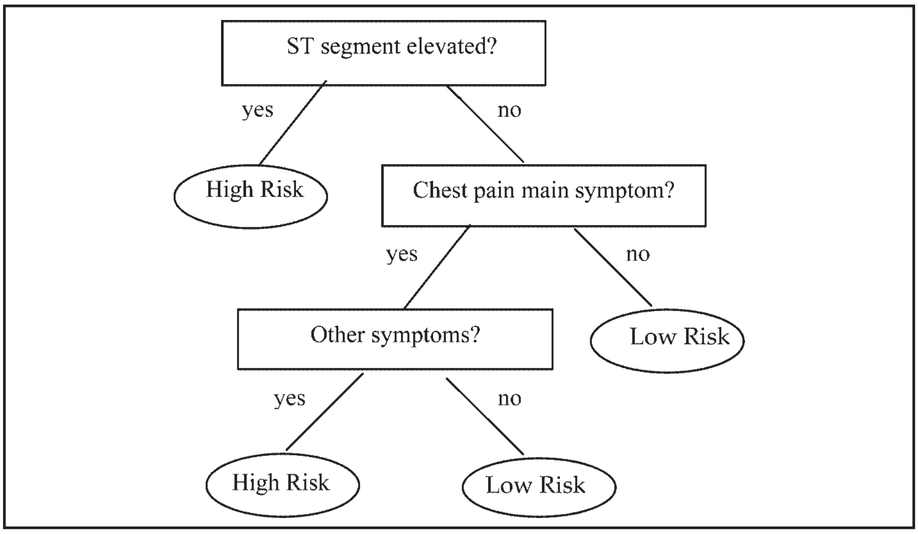

Figure 1. A fast frugal tree for categorizing patients as having a high or low risk of ischemic heart disease. (For background information, see Green & Mehr, 1997.) The ST segment is an electrocardiographic measure (the portion between the QRS segment and the beginning of the T wave) indicative of ischemic heart disease.

level (Martignon, Katsikopoulos, & Woike, 2008). If a second question had been asked for all patients with an elevated ST segment, the tree in Figure 1 would not have been fast and frugal.

The labels "fast" and "frugal" have precise meanings: The frugality of a tree for a set of objects is the mean number of cues it uses for making a categorization, across these objects. The speed of a tree, also for a set of objects, is the mean number of basic operations—arithmetic and logical—used for making a categorization. By these definitions, changing question nodes into exits makes a tree faster and more frugal.

The Robustness of Fast and Frugal Trees

Figure 2 compares the accuracy of two types of fast and frugal trees (called Zig and Max; for details see Martignon et al., 2008) with two classic benchmarks from statistics and artificial intelligence: logistic regression (LR) and the classification and regression trees (CARTs) of Breiman, Friedman, Stone, and Olshen (1984). Martignon et al. (in press) used 11 medical categorization problems from the Machine Learning Repository at the University of California, Irvine. The accuracy of the four models was assessed in four cases each. In the first case, the parameters of the models were estimated based on all data—that is, the models were *fitted* to the whole data set. In the other three cases, the parameters of the models were estimated based on a subset of the data (90%, 50%, or 15% of the whole data set), and the same parameters were used to assess how well the models *predicted* the remaining data.

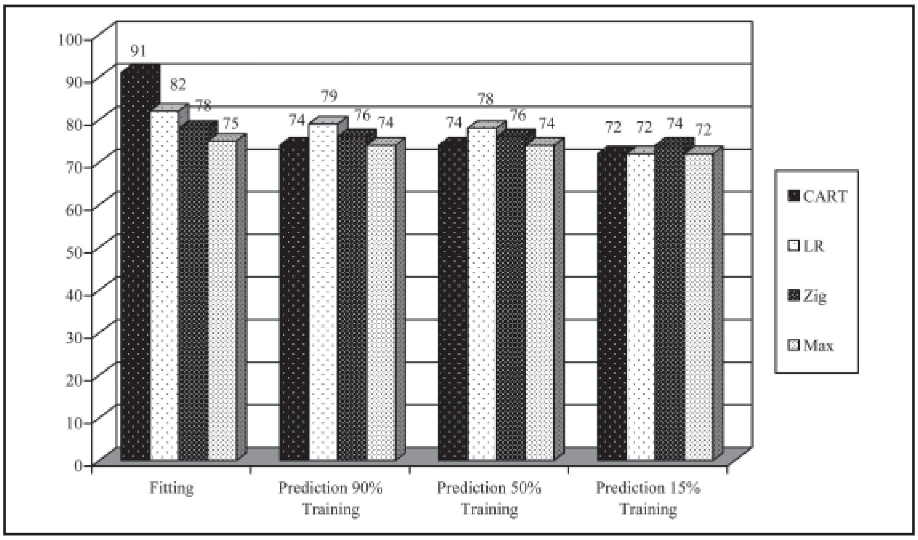

Figure 2. Average performance of four categorization models (classification and regression trees, or CART; logistic regression, or LR; and two types of fast and frugal trees, called Zig and Max; see Martignon et al., 2008 for details), across 11 medical problems, in fitting and three cases of prediction (training set was 90%, 50%, or 15% of the whole data set).

As is often the case, the more complex models (LR and CART) did much better in fitting. If it is to be used in the real world, however, a good categorization model needs to predict unknown cases, not explain the past by hindsight. Here, the simple Zig trees match or come close to the accuracy of CART and logistic regression, whereas Max lags a few percentage points behind. Zig even outperforms CART and LR when there is little available information (15% of the whole data set).

The Synergetic Potential of Fast and Frugal Trees

Fast and frugal trees have been readily accepted by decision makers in the field, have been shown themselves to be more robust (predict better) than more complex models using more information, and are transparent and can be easily conveyed. The heuristic's ecological rationality has been spelled out precisely. The analysis defines conditions regarding good and bad performance as well as the heuristic building blocks of searching, stopping, and decision making in detail. Fast and frugal trees are also psychologically plausible, assuming prediction and limited information search rather than fitting and optimization. That it beats some of the classical "optimal" benchmarks in prediction could be ascertained only through the new methodological tools that were made available as a result of greater specification. Last, the transparency and psychological satisfaction of fast and frugal trees is the primary reason they have been so successfully adopted by decision makers in the field.

However, in bringing decisions from the wild into the laboratory, it is important to ascertain at every stage whether or not crucial aspects of the decision and its environment have been lost in the process. Some of these aspects are nigh impossible to recreate in the lab, such as the level of motivation that comes about as a result of a threat to one's own or other people's lives. Others, such as the frequent occurrence of distracting factors or issues of vigilance, are often overlooked in laboratory research or studied in unstructured isolation. A thorough Cognitive Field Research analysis can provide a map of such factors, ensuring they are not neglected in the final analysis.

Regarding the case study discussed previously, the critical decision method may reveal that it is much more common for doctors to be able to detect the "chest pain" cue prior to the "ST segment" cue in a natural setting. It may also give hints as to how the tree could be adapted to changing circumstances on the fly or at which points such trees could be usefully incorporated into the larger decision-making process. A goal-directed task analysis may identify that a deeper problem lies somewhere else entirely (e.g., that the goal of working without fear from lawsuit interferes with the goal of providing appropriate care). The synergetic potential between NDM and fast and frugal trees arises from the possibility that researchers of the microcognitive and researchers of the macrocognitive, with similar attitudes, may study cognition in coherence.

Conclusion

In fundamental and important ways, NDM and FFH are already interconnected. The two approaches to decision making share common roots, philosophies, and values. Critically, however, the two programs substantively diverge in their methods, strengths, and limitations. As such, we have argued that there is considerable yet-to-be-realized synergetic potential offering opportunities for those who might endeavor toward a tighter integration of methods and approaches. Some 20 years after its first conference, NDM as a scientific endeavor is performing a difficult, perhaps even dangerous, balancing act. NDM has the admirable goal of making its results relevant for decision making in this fundamentally complex and uncertain world; accordingly, its practitioners have opted to study decision making in highly specific environments. However, this need not come at the cost of theoretical specification. To be clear, NDM has many, many achievements to show for its efforts. Nonetheless, as has been documented here and elsewhere (Gigerenzer, 1996, 1998; Simon, 1990), the loss of theoretical precision and predictive accuracy can potentially undermine what is actually dearest to this program, its applicability.

As the NDM and FFH programs prepare to welcome a new generation of researchers, we feel it is crucial to acknowledge the necessity of creating a science that is precise, sustainable, and makes a difference. A call for integration is a call to reflect upon and distill one's core values—to bring them to the fore. We hope that the present effort contributes to the sustainability of the ecological decision sciences, and we offer an open invitation to other decision researchers, in NDM and beyond, willing to join us in this integrative endeavor.

Acknowledgment

We thank Markus Feufel and Frank Detje for helpful comments regarding earlier drafts of this paper.

References

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89,* 369–406.

Ayton, P., & Fisher, I. (2004). The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness? *Memory and Cognition, 32,* 1369–1378.

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees.* London: Chapman and Hall.

Cannon-Bowers, J. A., Salas, E., & Converse, S. (1993). Shared mental models in expert team decision making. In N. J. Castellan (Ed.), *Individual and group decision making: Current issues* (pp. 221–246). Hillsdale, NJ: Erlbaum.

Caverni, J. P. (2001). What is really naturalistic about naturalistic decision-making research? *Journal of Behavioral Decision Making, 14,* 357–358.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5,* 121–152.

Cohen, M. S., & Freeman, J. T. (1997). Understanding and enhancing critical thinking in recognition-based decision making. In R. Flin & L. Martin (Eds.), *Decision making under stress: Emerging themes and applications* (pp. 161–169). Aldershot, UK: Ashgate.

Cohen, M. S., Freeman, J. T., & Wolf, S. (1996). Meta-recognition in time-stressed decision making: Recognizing, critiquing, and correcting. *Human Factors, 38,* 206–219.

Cokely, E. T., & Kelley, C. M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgment and Decision Making, 4,* 20–33.

Cokely, E. T., Kelley, C. M., & Gilchrist, A. L. (2006). Sources of individual differences in working memory: Contributions of strategy to capacity. *Psychonomic Bulletin and Review, 13,* 991–997.

Cokely, E. T., Schooler, L. J., & Gigerenzer, G. (2010). Information use for decision making. In M. N. Maack & M. J. Bates (Eds.), *Encyclopedia of library and information sciences* (3rd ed., pp. 2727–2734). New York: Taylor & Francis.

Crandall, B., & Getchell-Reiter, K. (1993). Critical decision method: A technique for eliciting concrete assessment indicators from the "intuition" of NICU nurses. *Advances in Nursing Sciences, 16,* 42–51.

Crandall, B., Klein, G., & Hoffman, R. R. (2006). *Working minds: A practitioner's guide to cognitive task analysis.* Cambridge, MA: MIT Press.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science, 243,* 1668–1674.

de Groot, A. D. (1978). *Thought and choice in chess.* The Hague: Mouton. (Original work published 1946)

Einhorn, H. J., & Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgment and choice. *Annual Review of Psychology, 32,* 53–88.

Endsley, M. R. (1995a). Measurement of situation awareness in dynamic systems. *Human Factors, 37,* 65–84.

Endsley, M. R. (1995b). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37,* 32–64.

Endsley, M. R., Bolte, B., & Jones, D. G. (2003). *Designing for situation awareness: An approach to human-centered design.* London: Taylor & Francis.

Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 39–68). Cambridge, UK: Cambridge University Press.

Ericsson, K. A., Krampe, R. T., & Teschromer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100,* 363–406.

Ericsson, K. A., Prietula, M. J., & Cokely, E. T. (2007). The making of an expert. *Harvard Business Review, 85,* 114–121.

Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgement under uncertainty: Heuristic and biases* (pp. 422–444). Cambridge, UK: Cambridge University Press.

Flavell, J. H. (1979). Macrocognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, *34*, 906–911.

Fogarty, W. M. (1988). *Investigation report: Formal investigation into the circumstances surrounding the downing of Iran Air Flight 655 on 3 July 1988.* Arlington, VA: U.S. Department of Defense.

Gaissmaier, W., Schooler, L. J., & Rieskamp, J. (2006). Simple predictions fueled by capacity limitations: When are they successful? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 966–982.

Galesic, M., Garcia-Retamero, R., & Gigerenzer, G. (2009). Using icon arrays to communicate medical risks: Overcoming low numeracy. *Health Psychology, 28,* 210–216.

Gigerenzer, G. (1996). On narrow norms and vague heuristics: Reply. *Psychological Review, 103,* 592–596.

Gigerenzer, G. (1998). Surrogates for theories. *Theory and Psychology, 8,* 195–204.

Gigerenzer, G. (2008). *Rationality for mortals.* New York: Oxford University Press.

Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science, 1,* 107–143.

Gigerenzer, G., & Engel, C. (Eds.). (2006). *Heuristics and the law.* Cambridge, MA: MIT Press.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review, 103,* 650–699.

Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models—A Brunswikian theory of confidence. *Psychological Review, 98,* 506–528.

Gigerenzer, G., & Selten, R. (2001). *Bounded rationality—The adaptive toolbox.* Cambridge, MA: MIT Press.

Gigerenzer, G., Todd, P. M., & the ABC Research Group. (1999). *Simple heuristics that make us smart.* New York: Oxford University Press.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review, 109,* 75–90.

Green, L., & Mehr, D. R. (1997). What alters physicians' decisions to admit to the coronary care unit? *Journal of Family Practice, 45,* 219–226.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108,* 356–388.

Hoffman, R. R., & Militello, L. G. (2008). *Perspectives on cognitive task analysis: Historical origins and modern communities of practice.* Boca Raton, FL: CRC Press/Taylor & Francis.

Howard, R. A. (1966). Decision analysis: Applied decision theory. In *Proceedings of the Fourth International Conference on Operational Research* (pp. 51–71). New York: Wiley.

Howard, R. A. (1968). Foundations of decision analysis. *IEEE Transactions on Systems Science and Cybernetics, 4,* 211–219.

Jungermann, H. (2001). One mode of decision making (research)—No less, no more. *Journal of Behavioral Decision Making, 14,* 367–369.

Kahneman, D., & Klein, G. A. (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist, 64,* 515–526.

Kaempf, G. L., Wolf, S., & Miller, T. E. (1993). Decision making in the AEGIS combat information center. In *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting* (pp. 1107–1111). Santa Monica, CA: Human Factors and Ergonomics Society.

Katsikopoulos, K. V. (2009). Coherence and correspondence in engineering design: Informing conversation and connecting with judgment and decision-making research. *Judgment and Decision Making, 4,* 147–153.

Klein, G. (1989, May). Do decision biases explain too much? *Human Factors Society Bulletin, 23*(5), 1–3.

Klein, G. (1993). A recognition primed decision (RPD) model of rapid decision making. In G. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), *Decision making in action: Models and methods* (pp. 205–218). Cambridge, MA: MIT Press.

Klein, G. (1998). *Sources of power—How people make decisions*. Cambridge, MA: MIT Press.

Klein, G. (2008). Naturalistic decision making. *Human Factors, 50,* 456–460.

Klein, G., Calderwood, R., & Clinton-Cirocco, A. (1986). Rapid decision making on the fire-ground. In *Proceedings of the Human Factors and Ergonomics Society 30th Annual Meeting* (pp. 576–580). Santa Monica, CA: Human Factors and Ergonomics Society.

Klein, G., Calderwood, R., & MacGregor, D. (1989). Critical decision method for eliciting knowledge. *IEEE Transactions on Systems, Man, and Cybernetics, 19,* 462–472.

Klein, G., & Militello, L. (2005). The knowledge audit as a method for cognitive task analysis. In H. Montgomery, R. Lipshitz, & B. Brehmer (Eds.), *How professionals make decisions* (pp. 335–342). Mahwah, NJ: Erlbaum.

Klein, G., Moon, B., & Hoffman, R. R. (2006a). Making sense of sensemaking: 1. Alternative perspectives. *IEEE Intelligent Systems, 21,* 70–73.

Klein, G., Moon, B., & Hoffman, R. R. (2006b). Making sense of sensemaking: 2. A macrocognitive model. *IEEE Intelligent Systems, 21,* 88–92.

Klein, G., & Weitzenfeld, J. (1982). The use of analogs in comparability analysis. *Applied Ergonomics, 13,* 99–104.

Klein, G., & Wolf, S. (1998). The role of leverage points in option generation. *IEEE Transactions on Systems, Man, and Cybernetics Part C—Applications and Reviews, 28,* 157–160.

LeBoeuf, R. A., & Shafir, E. (2001). Problems and methods in naturalistic decision-making research. *Journal of Behavioral Decision Making, 14,* 373–375.

Lesgold, A. M., Feltovich, P., & Glaser, R. (1980). Radiological expertise. *Bulletin of the Psychonomic Society, 16,* 169.

Lipshitz, R. (1994). Decision making in three modes. *Journal for the Theory of Social Behavior, 24,* 47–66.

Lipshitz, R., Klein, G., Orasanu, J., & Salas, E. (2001a). Focus article: Taking stock of NDM. *Journal of Behavioral Decision Making, 14,* 331–352.

Lipshitz, R., Klein, G., Orasanu, J., & Salas, E. (2001b). Rejoinder: A welcome dialogue—And the need to continue. *Journal of Behavioral Decision Making, 14,* 385–389.

Marewski, J. N., Gaissmaier, W., Schooler, L. J., Goldstein, D. G., & Gigerenzer, G. (2010). From recognition to decisions: Extending and testing recognition-based models for multi-alternative inference. *Psychonomic Bulletin and Review, 17,* 287–309.

Marewski, J. N., Schooler, L. J., & Gigerenzer, G. (2010). Five principles for studying people's use of heuristics. *Acta Psychologica Sinica, 42,* 72–87.

Martignon, L., Katsikopoulos, K. V., & Woike, J. K. (2008). Categorization with limited resources: A family of simple heuristics. *Journal of Mathematical Psychology, 52,* 352–361.

Martignon, L., Katsikopoulos, K. V., & Woike, J. K. (in press). Fast and frugal trees for categorization. In P. M. Todd, G. Gigerenzer, & the ABC Research Group (Eds.), *Ecological rationality: Intelligence in the world*. Oxford, UK: Oxford University Press.

Meehl, P. E. (1954). *Clinical vs. statistical prediction: A theoretical analysis and review of the evidence*. Minneapolis: University of Minnesota Press.
Militello, L. G., & Hutton, R. J. B. (1998). Applied cognitive task analysis (ACTA): A practitioner's toolkit for understanding cognitive task demands. *Ergonomics, 41,* 1618–1641.
Oppenheimer, D. M. (2003). Not so fast! (and not so frugal!): Rethinking the recognition heuristic. *Cognition, 90,* B1–B9.
Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General, 135,* 207–236.
Roelofsma, P. H. M. P. (2001). Evaluating ten years of NDM: Welcome back in the lab! *Journal of Behavioral Decision Making, 14,* 377–379.
Salas, E., Cooke, N. J., & Rosen, M. A. (2008). On teams, teamwork, and team performance: Discoveries and developments. *Human Factors, 50,* 540–547.
Savage, L. J. (1952/1972). *The foundations of statistics* (1st ed./2nd rev. ed.). New York: Dover Publications.
Schooler, L. J., & Hertwig, R. (2005). How forgetting aids heuristic inference. *Psychological Review, 112,* 610–628.
Serwe, S., & Frings, C. (2006). Who will win Wimbledon? The recognition heuristic in predicting sports events. *Journal of Behavioral Decision Making, 19,* 321–332.
Shanteau, J. (1992). Competence in experts—The role of task characteristics. *Organizational Behavior and Human Decision Processes, 53,* 252–266.
Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology, 14,* 1–19.
Todd, P. M., & Gigerenzer, G. (2001). Putting NDM into the adaptive toolbox. *Journal of Behavioral Decision Making, 14,* 381–383.
Todd, P. M., & Gigerenzer, G. (2007). Environments that make us smart: Ecological rationality. *Current Directions in Psychological Science, 16,* 167–171.
Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty—Heuristics and biases. *Science, 185,* 1124–1131.
von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior* (2nd ed.). Princeton, NJ: Princeton University Press.
Warwick, W., McIlwaine, S., Hutton, R., & McDermott, P. (2001). Developing computational models of recognition-primed decision making. In *Proceedings of the Tenth Conference on Computer Generated Force* (pp. 232–331), Norfolk, VA.
Wegwarth, O., Gaissmaier, W., & Gigerenzer, G. (2009). Smart strategies for doctors and doctors in-training: Heuristics in medicine. *Medical Education, 43,* 721–728.
Zsambok, C. E. (1997). Naturalistic decision making: Where are we now? In C. E. Zsambok & G. A. Klein (Eds.), *Naturalistic decision making* (pp. 3–16). Mahwah, NJ: Erlbaum.
Zsambok, C. E., & Klein, G. A. (Eds.). (1997). *Naturalistic decision making.* Mahwah, NJ: Erlbaum.

Niklas Keller is pursuing his PhD in psychology at the University of Bamberg. He is a doctoral contractor with the Human Factors Division at the Industrieanlagen-Betriebsgesellschaft mbH (IABG), Munich, and a visiting predoctoral research fellow with the Center for Adaptive Behavior and Cognition at the Max Planck Institute for Human Development, Berlin. His research focuses on applying the theory of bounded rationality to decision making in high-stakes, high-risk domains, such as the military or police.

Edward T. Cokely is an assistant professor of psychology and director of the Decision Science and Decision Engineering Laboratory (DeSciDE) at the Michigan Technological University. He is also an associated scientist with the Center for Adaptive Behavior and Cognition at the Max Planck Institute for Human Development. His research focuses on cognitive process

investigations of individual differences in human performance and decision making, with emphasis on risk and ethics in medical, financial, legal, and engineering domains.

Konstantinos V. Katsikopoulos works on the theory of bounded rationality. He has contributed to the development and testing of heuristics that people (laypersons and experts) should or do use to make decisions with limited time, information, and computation. He received a PhD in industrial engineering and operations research from the University of Massachusetts and is a senior research scientist (W2 associate professor) at the Max Planck Institute for Human Development in Berlin, Germany.

Odette Wegwarth is a research scientist with the Harding Center for Risk Literacy at the Max Planck Institute for Human Development. While working on her PhD she gained wide-ranging expertise of different models of humans' decision making. Her current research focused on what statistical formats foster and limit doctors' and patients' risk understanding and how this understanding would shape these groups' decision making.