The Handbook of Rationality

Edited by: Markus Knauff, Wolfgang Spohn

Citation:

The Handbook of Rationality
Edited by: Markus Knauff, Wolfgang Spohn
DOI: 10.7551/mitpress/11252.001.0001
ISBN (electronic): 9780262366175
Publisher: The MIT Press
Published: 2021

Funding for the open access edition was provided by the MIT Libraries Open Monograph Fund.

The MIT Press

1.3 Evolution of Rationality

Gerhard Schurz

Summary

Rationality is a normative concept that has two dimensions: theoretical and practical rationality. We suggest evaluating theoretical rationality in terms of its cognitive success and practical rationality in terms of its promoting well-being and social cooperation. Based on this understanding, the evolution of rationality is investigated from both a philosophical and a psychological point of view. Human cognition has evolved in three dimensions: biological, cultural, and individual. Many researchers believe that human cognition originates from a sociolinguistic coevolution. A second characteristic is the coevolution of human technology and humans' ability to reason causally. Based on evolutionary considerations, it is argued that human cognition is based on a division of labor between specialized cognitive modules and general-purpose nonspecific mechanisms of learning and reasoning.

1. Factual and Normative Aspects of Rationality

Is it rational to increase human power by means of science and technology? Is it rational to believe in God because this increases one's happiness? The answer to these questions cannot solely be based on empirical facts and logical reasoning, because the notion of rationality used in these questions has a *normative* component. Since David Hume (1739/2004), we know that there is no logically valid inference from facts to norms.[1] The notion of rationality that one assumes relies at least partially on a value-loaded decision.

Thus, before we start this chapter, we have to explicate our assumed concept of rationality. In philosophy, one usually distinguishes between *theoretical* rationality (i.e., the rationality of our picture of the world) and *practical* rationality (i.e., the rationality of our norms of action). The aim of theoretical rationality is to acquire predictively and causally or explanatorily relevant information about the world. More generally formulated, the goal of theoretical rationality is *cognitive success*. Schurz and Hertwig (2019) define the cognitive success of a method as the product of its validity (percentage of correct inferences among all inferences rendered) times its applicability (percentage of all problems to which the method is applicable among all problems of a given type). When speaking in what follows about the evolution of theoretical rationality, we will focus on the dimension of cognitive success. There are other, *noncognitive* aspects of human belief systems that are the target of evolutionary selection. One such aspect played a significant role in the evolution of religion and has been called the *generalized placebo effect*: the mere *belief* in a benevolent God (or system of Gods) equipped with supernatural powers, who protects one and rewards godly behavior (either earthly or afterworldly), has strongly positive effects on one's psychological well-being, entirely *independently* of the cognitive implausibility of this belief.[2] Thus, although we focus our chapter on cognitive aspects, it should be clear that also certain noncognitive aspects play a role in the evolution of human belief systems.

The goal of practical rationality is usually seen in guiding our actions in the interest of human beings. What this means depends on one's accepted value judgments. As a minimal core meaning, we assume that acting in a practically rational way implies acting (a) in the egoistic interest of the acting person and (b) promoting cooperation between the interaction partners (possibly extended to nonhuman beings). Our discussion will focus on these two aspects of the evolution of practical rationality.

The remainder of this article is structured as follows: after laying down the foundations of generalized evolution theory, including biological, cultural, and individual evolution (section 2), we discuss the interrelation between the evolutions of social cognition, language, and mental representation (section 3); the evolution of causal reasoning (section 4); the relation between modularity and universality in human cognition (section 5); and the evolution of induction and deduction (section 6), and we conclude with a synopsis of the evolutionary architecture of human cognition (section 7).

2. Cognitive Evolution in Three Dimensions

Since our notion of rationality coincides with that of successful cognition, including social cognition, the focus of our chapter will be on the evolution of cognition. Human cognition is the product of *evolution in three dimensions*. The first and second dimension are:

(1) *Biological evolution* based on the biological inheritance of genetic information, including some recently explored epigenetic effects[3]

(2) *Cultural evolution*, where "culture" is intended in the broad sense of the generation-wise tradition of *acquired* information

According to a more recent research paradigm, also the evolution of culture can be understood as a generalized Darwinian process, which is not based on the inheritance of genes through (sexual) reproduction but rather on the inheritance of cultural traits based on *social learning*.[4] The latter notion means learning from other individuals, as opposed to individual learning from one's own trial-and-error experience. Knowledge acquired through social learning (i.e., not inherited genetically) is passed on from generation to generation and is successively improved through ongoing iterations of variation–selection (or trial-and-error) cycles.[5] This is most impressively exemplified in the evolution of technology from the tools of the Stone Age to contemporary automated industry and electronic intelligence (cf. Basalla, 1988), but it applies equally to other domains in the evolution of culture, such as science or law.

Biological- as well as cultural-level evolution is based on the three Darwinian modules of *reproduction*, *variation*, and *selection* (Dennett, 1995; Schurz, 2011; Sober, 1993). In contrast, the third dimension of evolution is only partially Darwinian:

(3) *Individual learning*, including inductive association learning and behavioral trial-and-error learning

The latter kind of learning, also called "operant conditioning," consists of Darwinian processes of variation and selection, insofar as spontaneously exhibited patterns of behavior are increased or decreased in their frequency, depending on whether the feedback from the environment was positive or negative for the individual. The only difference to a "full" evolutionary process is that individual trial-and-error learning is confined to the life span of an individual and cannot transcend it (except when it becomes part of cultural evolution). Therefore, Schurz (2011, section 11.1.2) characterizes individual trial-and-error learning by means of Campbell's notion of "retention" instead of reproduction (Campbell, 1960).

In association learning, two stimuli frequently observed to be co-occurring are associated with each other in one's mind and projected as expectation in regard to future observations. This is nothing but the inference of *induction* as described by David Hume. A subcase of association learning is *classical conditioning*, as exemplified by Pavlov's dog, which, due to a repeated ringing of a bell shortly prior to the giving of food, soon expects food already at hearing the bell ringing and starts to salivate. Classical conditioning is a biologically widespread mechanism demonstrated almost everywhere in the animal world, even in worms (Bitterman, 2000; Delius, Jitsumori, & Siemann, 2000). Moreover, classical conditioning has a known neurological basis, the so-called *Hebb rule* (Rojas, 1996, pp. 21, 258–259). However, classical conditioning does not follow the Darwinian modules of spontaneous variation and selection but is a form of directly environment-driven (or "Lamarckian") learning, since association hypotheses are directly "imprinted" upon the individual by the observed environment via inductive generalization.

In conclusion, and in anticipation of the following, human cognition turns out to be based on the interaction of a large diversity of cognitive processes. Some of them are genetically determined, others are culturally acquired, and still others are individually learned. Even if a cognitive trait has a genetic basis, it does *not* usually determine a person's mind (thinking this would be so constitutes a frequent misunderstanding of genetic accounts). Because of the enormous plasticity of the human mind (cf. Churchland, 1979), there is presumably not even a single innate cognitive mechanism that cannot be overruled or corrected by other cognitive processes and in particular by culturally acquired abilities. For example, humans experience the Mueller–Lyer illusion as illustrated in figure 1.3.1: they see the left line as longer than the right line, although their length is equal. But the human mind is plastic enough to recognize this error, to correct it by reflective reasoning, and to recognize its cause, which lies in mechanisms of three-dimensional vision: the arrows and inverted arrows produce the three-dimensional effect that the right stroke appears to be *closer* to the observer than the left stroke and is thus perceived as being shorter (Rock, 1984, pp. 163, 167).

As another example, consider Immanuel Kant's famous thesis that perceived space is necessarily Euclidean and has three dimensions, because this is dictated by the innate (a priori) structure of human visual cognition. However,

Figure 1.3.1
Mueller–Lyer illusion.

mathematically educated humans can overcome even this innate "barrier" and conceptualize non-Euclidean or more-than-three-dimensional spaces. Further examples will be given below.

In the next four sections, we start with the most basic components of human cognition and proceed to increasingly higher forms.

3. Social Cognition and the Evolution of Language and Mental Representation

The most striking cognitive difference between humans and all other animals is the human faculty of language. Rudimentary abilities of symbolic signaling systems are already present in nonhuman animals, but while great apes can be trained to learn at most a few hundred words, humans can, in principle, produce a potentially infinite number of grammatical sentences from several hundreds of thousands of words and abstract syntax. According to a well-known theory advanced by philosophers (Gehlen, 1956/1977; Topitsch, 1979), anthropologists (Boyer, 1994; Wenegrat, 1990), and biologists (Tomasello, 1999), the hominids' language faculties have developed (presumably over 1–2 millions of years) in close interrelation with their *social* cognition, their abilities to communicate and to organize joint intention and cooperative action. One also speaks of a *sociolinguistic coevolution* (Heyes, 2012).

Early hominid groups could even hunt animals as large as mammoths; a high amount of social coordination and division of labor is required for tasks of this sort. Although many aspects of human language appear to be cognitively universal, the initial selection pressure for the evolution of language was cultural. According to Thompson, Kirby, and Smith (2016), a strong genetic inheritance of grammar as postulated by Chomsky (1968) is not needed to explain the apparent universality of generative grammar in humans. By means of computer simulations and mathematical proof, Thompson et al. (2016) demonstrate that a *weak genetic bias* in favor of a certain trait is sufficient to explain its universality in human cultures, since the genetic bias is amplified by learning processes. Jablonka and Lamb (2005) and Jablonka, Ginsburg, and Dor (2012) explain language evolution as a *gene–cultural coevolution*: various aspects of language initially invented culturally were later genetically assimilated (i.e., became partially "inborn") because they created genetic selection pressures for cognitive abilities specialized for learning language; this mechanism is also known as the "Baldwin effect" (cf. Dennett, 1995, pp. 77ff.).

Communication—both about one's intentions and about the environment—is an obvious presupposition of successful cooperation. The coordination of individual actions that is required for successful cooperative behavior is systematically analyzed in *evolutionary game theory*. Weibull (1995, pp. 28–31) classifies symmetric two-person games into three kinds: (1) Hawk–Dove games (possessing a stable mixed equilibrium); (2) coordination games (possessing several stable pure equilibria), including the famous stag-hunt game as a subcase; and (3) prisoner's dilemma games (possessing no nontrivial equilibria). For the stag-hunt game, it has been shown (Skyrms, 2004) that under many but not all circumstances, successful cooperative behavior can emerge but is *not* evolutionary stable: if too many individuals prefer hunting small prey ("hare") instead of large prey ("stag"), because the success of the former does not depend on one's partners' cooperative behavior, then the evolution of cooperation breaks down.

Even worse are the evolutionary chances of cooperation in the prisoner's dilemma game, in which a cooperative action of *helping the other* is only optimal if it occurs reciprocally. Here cooperative actions are doomed to get exploited by "egoists" (i.e., noncooperators who receive help but don't return help). Thus, in the prisoner's dilemma game, cooperative action has to be stabilized by additional mechanisms of reciprocation.[6] Axelrod (1984/2006) proposed the strategy *tit-for-tat* (TFT), which initially acts cooperatively and then retaliates (i.e., punishes defection with defection and rewards cooperation with cooperation). Axelrod suggested that TFT is universally optimal and thus evolutionary stable, but later studies showed that there is no generally stable cooperation strategy in the iterated prisoner's dilemma, since for every refined cooperation strategy, one can invent a refined deception strategy that exploits it (Lorberbaum, 1994). However, there are strategies by means of which the evolution of cooperation can at least be imperfectly stabilized, such as (1) group selection with ongoing regrouping (Sober & Wilson, 1998), (2) communication and correlated pairing (Skyrms, 2010), (3) institutionalized sanctioning mechanisms (Gürerk, Irlenbusch, & Rockenbach, 2006), (4) cognitive mechanisms of cheater detection (Cosmides & Tooby, 1992), and (5) religious stabilization of altruism (Wilson, 2002).

Historically, the strategy of reciprocation has first been rationally reflected in the earliest human juridical document, the Babylonian *Codex Hammurabi* (1800 B.C.). But the evolution of cooperation begins much earlier. Many hunter–gatherer tribes were quite egalitarian, especially in places where food supply was instable and food sharing important. In their comparatively small groups, egoistic behavior of group members was much more easily controllable than in large agricultural societies (Sober & Wilson, 1998, p. 178; Wilson, 2002, p. 21). Thus, in all known hunter–gatherer societies, mechanisms of social reputation ("This guy is a good one") have developed and are still effective in peer groups of contemporary societies. In contrast, in the larger communities of premodern states formed after the agrarian revolution around 10,000 years B.C., reputation mechanisms alone were not enough to prevent cooperation-undermining exploitation behavior, whence in all known agrarian proto-states, governmental institutions that sanctioned rule breakers and collected taxes (among other activities) have emerged.

As explained, social cooperation requires the evolution of language and communication. But more is required for cooperation: the ability to represent and understand the minds, intentions, and beliefs of the other persons with whom one cooperates. Already at a very young age, human children develop the competence to understand other persons. At about nine months, babies acquire the ability of *joint attention* (i.e., they recognize the other's gaze direction and can direct their own gaze to the object spotted by the other). Tomasello (1999, pp. 61–70) speaks of the "nine-month revolution." A few months later, human start to *point* toward objects of their interest, and at the same time, the learning of language begins. Later, children acquire the ability of recognizing not only the others' intentions but also their *beliefs* and of distinguishing these from their own beliefs. This capability of children has been called their ability of *mindreading* (Nichols & Stich, 2003), or their possession of a *theory of mind* (Leslie, 1987). It was first demonstrated in an experiment called the *false-belief task* (Wimmer & Perner, 1983). The result of this experiment was based on verbal instructions and supported the conclusion that the ability to understand (false) beliefs of others emerges in children at the age of about four years. However, more recent experiments using a violation-of-expectation method demonstrated that already 15-month-old infants seem to understand when a person has a false belief (Onishi & Baillargeon, 2005).

There is a controversy about the nature of the cognitive mechanisms underlying human "mindreading" (cf. Nichols & Stich, 2003). While the theory-of-mind approach assumes a specialized cognitive module responsible for this faculty, the *simulation approach* argues for the view that the same cognitive mechanism by which people represent their own intentions and beliefs is used for simulations of the (different) intentions and beliefs of other persons. An argument in favor of the simulation view is evolutionary simplicity: one cognitive system is evolutionary less costly and thus more probable than two separate systems (Schulz, 2011).[7]

At about six years, children acquire the *conventional moral stage*, according to the findings of Piaget (1932) and Kohlberg (1984). This stage goes beyond the mere understanding of the nature of reciprocity, which children already possess at the age of three or four (Cosmides & Tooby, 2015, p. 653): it consists in the interpersonal generalization of reciprocal rules (if I have to follow the rule, you have to follow it, too), which establishes a basic sense of normative justice together with the ability to detect rule-breakers.

On a deeper level, the sociocognitive coevolution was not confined to communication and mindreading but went hand in hand with the evolution of higher-level systems of *mental representation* that intermediated between stimulus and response. A recent analysis of their evolutionary advantage has been given by Schulz (2018), thereby integrating and complementing related accounts of Millikan (2002), Papineau (2003), and Sterelny (2012). Schulz argues that connecting behavioral responses not directly with perceived stimuli, but rather with mental representations of object-configurations (which are in a separate step related to perceived stimuli), is cognitively much more efficient, because behaviorally relevant object-configurations (e.g., a bear in a cave) are not always indicated by the same perceptual cue but by different cues, dependent on the given circumstances. Note that in a similar way, the cognitive efficiency of the common-cause abduction of unobservable entities has been demonstrated in the philosophy of science (Schurz, 2016).

4. Evolution of Causal Reasoning

Besides the sociolinguistic, there was a second coevolution, the *sociotechnical* coevolution (Heyes, 2012; Jablonka et al., 2012). Essential to it was the ability of hominids not only to predict but also to *manipulate* the objects in their environment, from fire usage, stone tools, weaving techniques, and advanced weapons to the invention of agriculture. Crucial for these developments was humans' capability to reason about the *causes* of (pleasant or unpleasant) events. Causal reasoning goes beyond mere inductive correlational reasoning: it reflects the causal structure behind correlations. For example, if

A is observed to be correlated with B, then according to widely accepted causal principles, (1) A may be the cause of B, or (2) B may be the cause of A, or (3) A and B may be effects of a common cause (which follows from the so-called *causal Markov* condition; cf. Pearl, 2009; Schurz & Gebharter, 2016; Spirtes, Glymour, & Scheines, 2000). Only in case (1) is it possible to change B by manipulating A. Thus, causal knowledge is crucial for humans' ability to manipulate and control their environment.

The earlier conjecture that causal reasoning is an exclusively human ability turned out to be untenable. Chimpanzees possess many systematic reasoning abilities, ranging from object classification and spatial movement to intelligent tool usage (Tomasello, 1999, pp. 16ff.). Also, other higher mammals have rudimentary causal reasoning abilities (Dunbar, 2000). However, humans' capabilities of causal modeling are by far superior compared to their nonhuman relatives. Herrmann, Call, Hernández-Lloreda, Hare, and Tomasello (2007) found that in simple causal tasks, great apes are about equivalent to three-year-old toddlers.[8] However, the big steps in causal and analytic reasoning in child development are achieved at ages between five and seven years (Brainerd, 1978). The ability to mentally slip into the role or "mind" of someone else is the basis of pretend play in childhood (Leslie, 1987) and a crucial precondition for the child's ability of *counterfactual* reasoning (Buchsbaum, Bridgers, Weisberg, & Gopnik, 2012; Rafetseder, Schwitalla, & Perner, 2013), which, according to philosophical analysis, constitutes an essential component of causal reasoning (Lewis, 1973; Woodward, 2003).

According to the anthropological theory mentioned at the beginning of section 3, humans' *causal* reasoning evolved in close relation to their *intentional* reasoning (i.e., their interpretation of the intentions of living beings as the causes of their behaviors). This hypothesis fits well with the cultural history of ideas. From the Stone Age to premodern civilizations, the major causal explanation mechanisms of *Homo sapiens* were twofold: some natural processes are causally explained by mechanical contact forces (e.g., throwing a stone or carrying a load), but natural processes such as the weather or the movement of celestial bodies, which emerge apparently unforced without an obvious mechanical cause, were explained by their *animation* according to the intentional model (Gehlen, 1956/1977). More specifically, these natural processes were explained in terms of the same social-intentional model that explained the behavior of fellow men—a process that led to the polytheistic conception of natural Gods (Topitsch, 1979; Wenegrat, 1990).

As rationality progressed in the cultural history of ideas, such explanations were rejected because of their lack of empirical support and replaced by explanations in terms of physical causes. However, the epistemic justification of the assumption of causes governing the behavior of nonliving physical objects turned out to be a highly difficult problem in the history of philosophy. The idealist philosopher Leibniz attempted to reduce causal to generalized intentional notions. In contrast, the empiricist philosopher David Hume argued that observational evidence can give us only knowledge of correlations, not of causation. In contemporary philosophy, there are promising suggestions for solving Hume's challenge (Schurz & Gebharter, 2016; Spirtes et al., 2000), but the controversy is still going on.

The fact that human reasoning operates with two distinct models of causality, physical versus intentional causality, is confirmed by several psychological studies. Already at the age of six months, babies distinguish between nonliving and living objects, to which they spontaneously apply these different paradigms of causality: while nonliving objects can be put in motion only by the transmission of a force by physical contact, living objects are also able to move spontaneously without any external force, solely due to their intentions or internal causes. This finding was established by means of experiments in which babies were presented two sequences of events (cf. Leslie, 1982; Spelke, Phillips, & Woodward, 1995): in the "normal" sequence, a nonliving object (e.g., a ball) moves toward a second one, hits it, and pushes it away. In the "abnormal" sequence, the first object likewise moves toward the second one but already comes to a halt a distance away from it, and the second object moves away without touch "as if by magic." While the babies hardly paid attention to the first event, their attention (measured by the duration of looking) lingers very long on the second event. In a second series, the same experiment was conducted with people instead of objects: a boy walks toward a girl; in the first case, he touches her and the girl walks away, and in the second case, he remains standing in front of the girl and the girl walks away after that. In this case, no significant difference was found in the duration of the babies' gaze: both sequences are equally "normal" for the babies.

Thus, already babies distinguish spontaneously between contact causality for nonliving and intentional causality for living objects. In other words, intentional causality and contact causality are two distinct causal models. This also fits with the human history of ideas, insofar as contact causality for inanimate objects not only is well anchored in intuitive human physics but has also dominated the history of philosophy from Aristotle to Descartes. In contrast, objects that moved without apparent contact causes (e.g., heavenly bodies) were consistently explained using the intentional model of causality as explained above.

5. Modularity and Universality in Human Cognition

A further debate concerns the generality versus specificity of human mechanisms of cognition. According to the school of evolutionary psychology (Barkow, Cosmides, & Tooby, 1992), the human mind consists of various cognitive "modules." These are cognitive programs or processes that were evolutionarily selected for specific tasks (Carruthers & Chamberlain, 2000); they are genetically anchored and little mutable.[9] Examples of such area- and purpose-specific modules have already been mentioned and include (1) the module of shared attention (developed in the child at 9 months), (2) the modules of intentionality and contact causality (at 6 months), (3) the modules of language acquisition (from 18 months on), (4) mental spatial models (at the sensorimotorical level with 0–2 years; Brainerd, 1978), (5) the theory-of-mind module (at 4–5 years), and (6) the development of conventional moral values (at 6 years), which are associated with (7) the module of cheater detection (see below). Related is the research program of *adaptive rationality* according to which humans' cognitive mechanisms are adapted to the structure of local environments, being tailored to the specific tasks for which they provide highly efficient solutions (Gigerenzer, Todd, & ABC Research Group, 1999; Todd, Gigerenzer, & ABC Research Group, 2012). A second claim of proponents of this paradigm is that simple heuristics are frequently more successful than computationally costly general reasoning mechanisms, following the slogan "Less can be more."

The *modularist* view has also been called the "Swiss-Army-knife model" (Tooby & Cosmides, 1992) or the "adaptive toolbox model" (Gigerenzer & Selten, 2001). The opposite view is defended by *generalist* models of cognition. Until the 1960s, classical logic was the dominant generalistic normative standard of theories of rational reasoning (cf. Inhelder & Piaget, 1958). When psychologists discovered empirically that, in many domains, human reasoning did not accord with the principles of logic (e.g., Wason, 1966), these findings were interpreted as signs of human irrationality (cf. Evans, 2002). It was suggested that psychologists should adopt an alternative general normative system, such as Bayesian probability and decision theory (e.g., Oaksford & Chater, 1991). However, human reasoning has been observed to deviate from the norms of probability, too (Barbey & Sloman, 2007; Kahneman 2011; Kahneman & Tversky, 1972). These findings led many psychologists toward a general skepticism in regard to the "normative" status of generalist reasoning accounts.

Contemporary defenders of generalistic reasoning accounts are more modest. They agree that the early models of reasoning in terms of logic or probability theory were too simplistic or unrealistic from a cognitive viewpoint. Moreover, they do not deny that human cognition hosts a multitude of specialized cognitive abilities or modules. However, it is implausible that the human brain contains a module for each specific purpose or area of application; 200,000 years of evolution are insufficient to allow so many modules (Tomasello, 1999, pp. 204–205). Thus, contemporary generalists argue that besides specialized modules, cognitive evolution has also brought about a set of purpose- and area-unspecific cognition mechanisms. This view has been defended, among others, by Over (2003, p. 122), Almor (2003, p. 104), and, more recently, in a volume on the "new thinking on the evolution of cognition" (Heyes & Frith, 2012). The authors in this book argue that human cognition is largely purpose-general; it does not resemble a Swiss army knife but rather a human *hand*, whose fingers can serve an extremely large variety of purposes (Heyes, 2012, p. 2092).

Recently, the existence of generalistic "improvisational intelligence" has also been acknowledged by some evolutionary psychologists (Barrett, Cosmides, & Tooby, 2007). For semantic clarification, note that the notion of a "cognitive module" in the mainstream literature is always related to a *specific type* of adaptation problem, purpose, and environmental condition. Thus, one should not also call general cognitive abilities "modules," as this would stretch this notion over all bounds and make it empty. Moreover, it has been argued that purpose-specificity and domain-specificity are two distinct *dimensions* of specificity/generality (Duchaine, Cosmides, & Tooby, 2001). However, it seems that the two—although clearly nonidentical—are definitely not independent: the number and diversity of *purposes* to which a cognitive ability is applicable and the number and diversity of *conditions* under which it is applicable are strongly correlated.

6. Induction and Deduction

The most important example of a *general* mechanism of learning and cognition is *induction*, which in the form of conditioning or association learning is common in virtually all animals (recall section 2). In the view of earlier behavioral psychologists such as Thorndike (1911), the merit of learning by conditioning consists precisely in the fact that this is *not* an area-specific but a universally applicable learning process.

General cognitive mechanisms are particularly important for sustainable evolutionary success under environmental conditions that are changing in unforeseeable ways—a situation that occurred frequently in hominid evolution and even more frequently in human cultural evolution. This argument has been put forward by Schurz and Thorn (2016) as an important complementation of the adaptive rationality account (see also Thorn & Schurz, 2019). To be cognitively successful under changing environments, one needs strategies for learning which cognitive methods perform best in which environment, or temporal phase of the environment (a point acknowledged within the adaptive rationality program; cf. Todd et al., 2012, p. 15). Obviously, meta-cognitive selection strategies have to be sufficiently general; otherwise, they could not serve the purpose of selecting the optimal method in a wide range of environments.

In view of the cognitive universality of inductive reasoning, it is remarkable that the epistemic justification of induction turned out to be one of the most difficult problems in philosophy. According to David Hume's famous challenge (1748/2006, chapters 4, 6), the reliability of inductive reasoning from past observations to future (unobserved) cases can in no way be demonstrated by rational argument, neither strictly nor even probabilistically. For this reasoning rests on the assumption that nature is uniform, that is, that the future resembles the past, but that this is so cannot be demonstrated by observation or by logic or probability calculus. Nor can it be demonstrated by induction, by pointing to the fact that so far, the method of induction was successful, because the "inductive justification of induction" is circular and thus epistemically worthless (for details, see Schurz, 2019, chapters 2–4). An alternative account of justifying induction, going back to Reichenbach (1935, §80), is based on the idea that induction is the *best that we can do* in order to achieve successful predictions. Schurz (2019) defends this idea in terms of the *optimality of meta-induction*. Given a type of problem (a prediction task or decision problem), the meta-inductivist observes the success records of all methods that are *accessible* to her (i.e., methods whose success rates can be observed). Using these records, the meta-inductivist constructs an optimal combination of these methods, in the form of a weighted average that is periodically updated in a success-dependent way. Based on discoveries in computational learning theory (Cesa-Bianchi & Lugosi, 2006), Schurz (2008, 2019) demonstrates that meta-induction achieves an optimal success rate in all possible worlds among all competing methods (of prediction or decision) that are accessible to the given agent. Schurz (2008, 2019) proposes meta-induction as a new solution to Hume's problem of the justification of induction mentioned above. Since meta-induction is primarily a strategy of social learning—that is, learning from the performance of others (although it can also be applied to the results of individual learning)—its optimality constitutes an essential foundation for the social propagation of information in cultural evolution (Schurz, 2012, 2019, section 10.2).

Induction and meta-induction are presumably the evolutionarily most fundamental cognitive mechanisms (insofar as all forms of learning from feedback are based on induction). They are certainly more fundamental than causal reasoning (see below) and even more fundamental than logical reasoning and deduction. On the other hand, from a philosophical viewpoint, logical reasoning is more fundamental than induction, since induction presupposes some amount of logic and not vice versa.

Psychologists have found out that human reasoning's fit with the rules of classical propositional logic is weak: while some rules (e.g., modus ponens) are already mastered by children at the age of six (Brainerd, 1978), other logical rules such as modus tollens are not even understood by the majority of adults. The latter fact is exemplified in Wason's famous card selection task (Evans, 1982, chapter 9; Wason, 1966). In this experiment (which is also discussed in chapter 2.3 by Johnson-Laird and chapter 4.5 by Chater & Oaksford, both in this handbook), test persons (TPs) are given four cards from a pack of cards and told to test the following rule: *If there is an A on the front side, then there is a 1 on the back side*. They see four cards lying on a table, two face-up, showing an A and a B, respectively, and two face-down, showing a 1 and a 2. They were asked, Which of these four cards do you have to turn over in order to check whether the rule actually applies to these cards? According to the laws of classical propositional logic, one should turn over the first (A) card and the fourth (2) card: turning over the first card corresponds to the logically valid modus ponens (MP) inference (from "If A, then 1" and "A" infer "1"). Turning the fourth card over corresponds to the equally valid modus tollens (MT) inference (from "If A, then 1" and "not-1" follows "not-A"). In contrast, turning over the second and third cards corresponds respectively to the two invalid inferences NA ("negating the antecedent") and AC ("affirming the consequent"). The empirical performance results, however, are MP 100%, NA 5%, AC 10%, and MT 5%. Thus, normal adults only master the logical rule of MP but fail to master the rule of MT and apply equally or even more often the invalid rules NA and AC.

To defenders of the logico-generalist paradigm of cognition, these results came like a shock. Equally disturbing was a further result first brought to light by Griggs and Cox (1982), which demonstrated that people apply the rule of MT perfectly in the social context of detecting rule breakers. Here, the TPs had to test the following consumption regulation in a youth club: *Those who drink alcohol have to be at least 16 years old*. TPs were confronted with four adolescents; of two of them (Berta, Klaus), they only knew the beverage but not the age (14 y., 18 y.), and of the other two (Lisa, Martin), they only knew the age but not the beverage (cola, beer). They were asked, *Whom do you have to check to determine whether he/she has broken the rule?* with the result that now all TPs checked Berta and Martin (i.e., they mastered this instantiation of MP and MT perfectly).

As an explanation of this finding, Cosmides and Tooby (1992) proposed the thesis that humans have a *specialized* cognitive module for the detection of social rule breakers (*cheater detection*), which emerged through selection in the course of social evolution. More recent experiments (Cosmides, Barrett, & Tooby, 2010) show that these special inference abilities cannot be explained by assuming a more specific but still general reasoning ability, such as reasoning with *deontic rules* (as argued by Cheng & Holyoak, 1985) or with *utilities* (as proposed by Manktelow & Over, 1991); rather, they are restricted to situations in which the action (whose permission is governed by the rule) is beneficial to the actor, so that the rule violation is indeed an act of cheating.

The fact that humans possess *content-specific* reasoning modules does not exclude that their conditional reasoning is also governed by some purpose-general rules. In this line, psychologists and philosophers have developed the following evolutionary explanation for the deviations of human conditional (if–then) reasoning from the norms of deductive logic: the if–then relations of our natural environment are almost never *strictly* true but rather admit of exceptions (i.e., they express relations of *high conditional probability*). The hypothesis of ordinary “If A then B” statements as high-conditional-probability statements (“B is highly probable given A”) was proposed by Adams (1975), evolutionarily substantiated by Schurz (2001), and has been experimentally confirmed (Evans, Handley, & Over, 2003; more on this in chapter 4.6 by Oberauer & Pessach, this handbook). It turns out that humans’ intuitive if–then reasoning fits well with the rules of probability logic (Pfeifer & Kleiter, 2010; Schurz & Thorn, 2012). The MT rule, however, is *not* probabilistically valid. This may explain why subjects rarely apply the rule MT in Wason’s experiment. In contrast, in the second experiment, the more specific cognitive module of cheater detection becomes activated. In this way, both the specific module of cheater detection and the general reasoning mechanisms for uncertain conditionals have a plausible evolutionary explanation, *without* coming into conflict with each other.

The interaction of specialized cognitive modules and general cognition mechanisms can also be excellently studied in the abovementioned example of the cognitive model of contact causality. Although already toddlers have the intuition that inanimate objects can be put in motion only through contact or pushing, people have known *magnets* for centuries. After a short time of adaptation to the seemingly “magical” actions at a distance of magnets, children just as adults become used to this without problems. Apparently, the innate but, in the case of magnets, inapplicable model of contact forces is *overwritten* and thus corrected by inductive learning. We learn from this consideration that humans’ innate model of causality does not determine human reasoning by “a priori necessity” (as believed by the philosopher Immanuel Kant) but just acts as one innate disposition among many others and can be *corrected* by other cognitive mechanisms.

Through examples like these, it becomes apparent how efficiently special and general cognitive mechanisms can work together. Over (2003, pp. 124–125) and Schurz and Thorn (2016) therefore propose a *dualistic* theory of cognition, according to which human cognition consists of area-specific modules *and* general cognitive processes. A further example splendidly illustrating the division of labor between general and purpose-specific reasoning processes is the selection of optimal methods from the “adaptive toolbox” by the general strategy of meta-induction, as explained above.

7. The Evolutionary Architecture of Human Cognition

We have so far explained the cooperation of specialized (modular) and general mechanisms of human cognition. Most modular cognitive processes, but also many general processes such as inductive inference, work for the most part *unconsciously*. What, then, is the actual role of *consciousness* in human cognition? There is no consensus on this question. An extreme view is *epiphenomenalism*, which claims that consciousness only plays the role of a subsequently summarizing *reporter* on our unconscious mental processes but not the role of a causal trigger (Block, Flanagan, & Güzeldere, 1996,

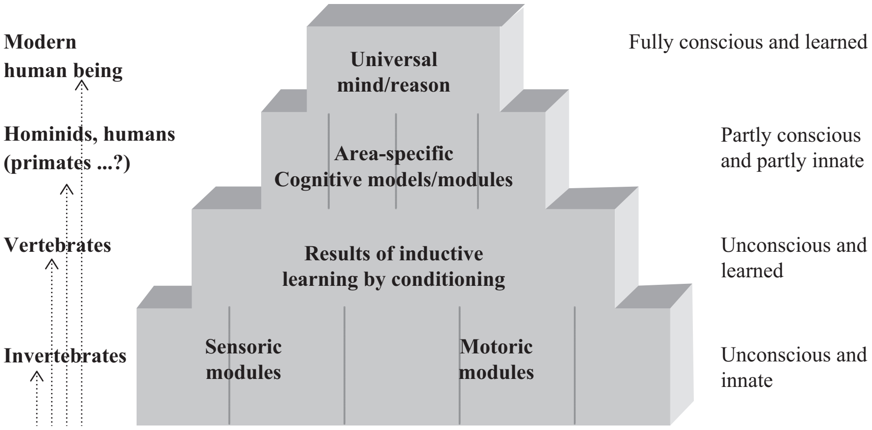

Figure 1.3.2
Evolution-based structure of human cognition.

chapter 19). This view appears to me to be exaggerated and even untenable. It is admittedly true that the conscious human mind reflects only a fraction of the unconscious processes of human cognition and often does play the role of a summarizing reporter. Moreover, conscious human reasoning processes are slower by decimal powers than the unconscious and modular cognitive processes. When we make practical decisions, for example, about whether we should quickly run across an intersection or not, then we gauge the involved probabilities and utilities intuitively within a fraction of a second, while if we calculated such an estimate in a precise decision-theoretic manner, it would take us hours.

And still, this much-slower conscious, logical mind has proved in the course of cultural evolution that—if given enough time to pursue its activities in a secure space—it is capable of much greater cognitive achievements than all cognitive heuristics put together. Its merits lie in its ability to comprehend its domain of application in a logically consistent and empirically controlled way, without any simplifications and distortions. Through the institutional establishment of an area of research and education sealed off from the pressure of practical demands, the conscious mind was able to develop our scientific-technological civilization. Conscious reasoning, from Euclid and Leonardo to Minkowski, was able to systematize the mechanisms of Euclidean geometry and perspectival projection so completely that all deceptions of our perception modules were resolvable. From Aristotle to Boole and Gödel, conscious reasoning was able to capture the laws of logical and probabilistic inference so completely that thereby all intuitive errors of reasoning could be recognized. Through abstract-mathematical thinking, human reasoning could penetrate areas exceeding everything that natural imagination and heuristics could ever give us: with Einstein (and others), it was able to intrude into the laws of the unimaginably large, with Bohr (and others) into the laws of the unimaginably small, with Darwin (and others) to transcend the border between the nonliving and the living, and with computer technology to transcend the border between nature and mind. The overwhelming success of conscious scientific reasoning extended humans' range of possible action and gave space to new technologies from synthetic chemistry and nuclear energy to genetic engineering and robotics.

In conclusion, we can say that even if humans' general reasoning capacities become efficient through their interaction with the specialized cognitive modules, it is still these general reasoning capacities that made scientific-technological progress possible and distinguish humans most strongly from their nonhuman ancestors. This leads to modeling human cognition as a layered structure that is illustrated by the architectonics displayed in figure 1.3.2.

Notes

1. For the proof of this thesis of Hume by means of modern logic, see Schurz (1997).

2. Religious beliefs in this generalized characterization are found in more or less all known cultures (including China), although there are important differences, for example, concerning polytheism versus monotheism or omnipotence versus power-restricted Gods or "ghosts" and so on (cf. Boyer, 1994; Schurz, 2011, section 17.5; Wenegrat, 1990). Another approach to the generalized placebo effect is the psychology of "positive thinking" (Taylor, 1989).

3. Cf. Jablonka and Lamb (2005). These authors even postulate four dimensions of evolution.

4. Pioneers of the modern theory of cultural evolution (to mention just a few) are Cavalli-Sforza and Feldman (1973), Dawkins (1976/1987), Boyd and Richerson (1985), and Donald

Campbell (1960); more recent developments are Mesoudi (2011), Schurz (2011), Dennett (2018), and Henrich (2016).

5. Dawkins (1976/1987) suggested the name “meme” for the cultural counterpart to genes, but this notion is controversial (cf. Aunger, 2002).

6. The notions of “egoistic” versus “cooperative” action are understood here in the evolutionary sense, not necessarily implying a corresponding psychological mechanism (cf. Sober & Wilson, 1998, p. 201).

7. Tomasello (1999) has argued that the ability to conceive their conspecifics as *intentional* beings discriminates humans from nonhuman great apes. Later this thesis turned out to be untenably strong: also great apes possess these abilities but to a much lesser extent (Call & Tomasello, 2008). Based on the observation of eye movements, Krupenye, Kano, Hirata, Call, and Tomasello (2016) found evidence that chimps can even anticipate actions based on false beliefs.

8. In contrast, human toddlers are better than chimpanzees in tasks of social cognition.

9. Note that a genetically determined mechanism need not be “single track” but can be “multitrack,” leading to different behaviors in different environments (like a hardwired computer program can react differently to different inputs). In such a case, the *phenotypic manifestation* of the hardwired mechanism is mutable but not the mechanism itself.

References

Adams, E. W. (1975). *The logic of conditionals*. Dordrecht, Netherlands: Reidel.

Almor, A. (2003). Specialized behaviour without specialized modules. In D. E. Over (Ed.), *Evolution and the psychology of thinking* (pp. 101–120). Hove, England: Psychology Press.

Aunger, R. (Ed.). (2002). *Darwinizing culture: The status of memetics as a science*. Oxford, England: Oxford University Press.

Axelrod, R. (2006). *The evolution of cooperation* (Rev. ed.). New York, NY: Basic Books. (Original work published 1984)

Barbey, A. K., & Sloman, S. A. (2007). Base rate respect: From ecological rationality to dual processes. *Behavioral and Brain Sciences*, *30*, 241–297.

Barkow, J. H., Cosmides, L., & Tooby, J. (Eds.). (1992). *The adapted mind: Evolutionary psychology and the generation of culture*. New York, NY: Oxford University Press.

Barrett, H. C., Cosmides, L., & Tooby, J. (2007). The hominid entry into the cognitive niche. In S. W. Gangestad & J. A. Simpson (Eds.), *The evolution of mind: Fundamental questions and controversies* (pp. 241–248). New York, NY: Guilford Press.

Basalla, G. (1988). *The evolution of technology*. Cambridge, England: Cambridge University Press.

Bitterman, M. E. (2000). Cognitive evolution. In C. M. Heyes & L. Huber (Eds.), *The evolution of cognition* (pp. 61–80). Cambridge, MA: MIT Press.

Block, N., Flanagan, O., & Güzeldere, G. (Eds.). (1996). *The nature of consciousness*. Cambridge, MA: MIT Press.

Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. Chicago, IL: University of Chicago Press.

Boyer, P. (1994). *The naturalness of religious ideas*. Berkeley: University of California Press.

Brainerd, C. J. (1978). *Piaget’s theory of intelligence*. Englewood Cliffs, NJ: Prentice-Hall.

Buchsbaum, D., Bridgers, S., Weisberg, D. S., & Gopnik, A. (2012). The power of possibility: Causal learning, counterfactual reasoning, and pretend play. *Philosophical Transactions of the Royal Society B*, *367*(1599), 2202–2212.

Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences*, *12*(5), 187–192.

Campbell, D. T. (1960). Blind variation and selective retention in creative thought as in other knowledge processes. *Psychological Review*, *67*, 380–400.

Carruthers, P., & Chamberlain, A. (Eds.). (2000). *Evolution and the human mind*. Cambridge, England: Cambridge University Press.

Cavalli-Sforza, L. L., & Feldman, M. W. (1973). Models for cultural inheritance I: Group mean and within group variation. *Theoretical Population Biology*, *4*, 42–55.

Cesa-Bianchi, N., & Lugosi, G. (2006). *Prediction, learning, and games*. Cambridge, England: Cambridge University Press.

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, *17*, 391–416.

Chomsky, N. (1968). *Language and mind*. New York, NY: Harcourt Brace & World.

Churchland, P. M. (1979). *Scientific realism and the plasticity of the mind*. Cambridge, England: Cambridge University Press.

Cosmides, L., Barrett, H. C., & Tooby, J. (2010). Adaptive specializations, social exchange, and the evolution of human intelligence. *PNAS*, *107*(Suppl. 2), 9007–9014.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 163–228). New York, NY: Oxford University Press.

Cosmides, L., & Tooby, J. (2015). Adaptations for reasoning about social exchange. In D. M. Buss (Ed.), *The handbook of evolutionary psychology: Vol. 2. Integrations* (2nd ed., pp. 625–668). Hoboken, NJ: Wiley.

Dawkins, R. (1987). *The selfish gene* (2nd, expanded ed.). Oxford, England: Oxford University Press. (Original work published 1976)

Delius, J., Jitsumori, M., & Siemann, M. (2000). Stimulus equivalencies through discrimination research. In C. M. Heyes & L. Huber (Eds.), *The evolution of cognition* (pp. 103–122). Cambridge, MA: MIT Press.

Dennett, D. C. (1995). *Darwin's dangerous idea: Evolution and the meanings of life*. New York, NY: Simon & Schuster.

Dennett, D. C. (2018). *From bacteria to Bach and back: The evolution of minds*. London, England: Penguin Books.

Duchaine, B., Cosmides, L., & Tooby, J. (2001). Evolutionary psychology and the brain. *Current Opinion in Neurobiology*, *11*(2), 225–230.

Dunbar, R. M. (2000). Causal reasoning and mental rehearsal in primates. In C. M. Heyes & L. Huber (Eds.), *The evolution of cognition* (pp. 205–220). Cambridge, MA: MIT Press.

Evans, J. St. B. T. (1982). *The psychology of deductive reasoning*. London, England: Routledge & Kegan Paul.

Evans, J. St. B. T. (2002). Logic and human reasoning. *Psychological Bulletin*, *128*, 978–996.

Evans, J. St. B. T., Handley, S. J., & Over, D. E. (2003). Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(2), 321–335.

Gehlen, A. (1977). *Urmensch und Spätkultur* [Prehistoric man and late culture] (4th corr. ed.). Frankfurt/Main, Germany: Athenaion. (Original work published 1956)

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

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

Griggs, R. A., & Cox, J. R. (1982). The elusive thematic-materials effect in Wason's selection task. *British Journal of Psychology*, *73*, 407–420.

Gürerk, Ö., Irlenbusch, B., & Rockenbach, B. (2006). The competitive advantage of sanctioning institutions. *Science*, *312*, 108–111.

Henrich, J. (2016). *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. Princeton, NJ: Princeton University Press.

Herrmann, E., Call, J., Hernández-Lloreda, M. V., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *Science*, *317*, 1360–1366.

Heyes, C. M. (2012). Introduction. *Philosophical Transactions of the Royal Society B*, *367*(1599), 2091–2096.

Heyes, C. M., & Frith, U. (Eds.). (2012). New thinking: The evolution of human cognition. *Philosophical Transactions of the Royal Society B*, *367*(1599).

Hume, D. (2004). *A treatise of human nature: Book III. Of morals*. Mineola, NY: Dover. (Original work published 1739)

Hume, D. (2006). *An inquiry concerning human understanding* (S. Butler, Ed.). Fairford, England: Echo Library. (Original work published 1748)

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking: From childhood to adolescence*. New York, NY: Basic Books.

Jablonka, E., Ginsburg, S., & Dor, D. (2012). The co-evolution of language and emotions. *Philosophical Transactions of the Royal Society B*, *367*(1599), 2152–2159.

Jablonka, E., & Lamb, M. J. (2005). *Evolution in four dimensions: Genetic, epigenetic, behavioral, and symbolic variation in the history of life*. Cambridge, MA: MIT Press.

Kahneman, D. (2011). *Thinking, fast and slow*. London, England: Macmillan.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, *3*(3), 430–454.

Kohlberg, L. (1984). *The psychology of moral development*. New York, NY: Harper & Row.

Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science*, *354*(6308), 110–114.

Leslie, A. M. (1982). The perception of causality in infants. *Perception*, *11*, 173–186.

Leslie, A. M. (1987). Pretense and representation: The origins of "theory of mind." *Psychological Review*, *94*(4), 412–426.

Lewis, D. (1973). Causation. *Journal of Philosophy*, *70*, 556–567.

Lorberbaum, J. (1994). No strategy is evolutionarily stable in the repeated prisoner's dilemma. *Journal of Theoretical Biology*, *168*(2), 117–130.

Manktelow, K., & Over, D. (1991). Social roles and utilities in reasoning with deontic conditionals. *Cognition*, *39*, 85–105.

Mesoudi, A. (2011). *Cultural evolution: How Darwinian theory can explain human culture and synthesize the social sciences*. Chicago, IL: University of Chicago Press.

Millikan, R. G. (2002). *Varieties of meaning: The 2002 Jean Nicod Lectures*. Cambridge, MA: MIT Press.

Nichols, S., & Stich, S. (2003). *Mindreading: An integrated account of pretence, self-awareness, and understanding other minds*. Oxford, England: Clarendon Press.

Oaksford, M., & Chater, N. (1991). Against logicist cognitive science. *Mind & Language*, *6*, 1–38.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*(5719), 355–358.

Over, D. E. (2003). From massive modularity to metarepresentation: The evolution of higher cognition. In D. E. Over (Ed.), *Evolution and the psychology of thinking* (pp. 121–144). Hove, England: Psychology Press.

Papineau, D. (2003). *The roots of reason: Philosophical essays on rationality, evolution, and probability*. Oxford, England: Clarendon Press.

Pearl, J. (2009). *Causality: Models, reasoning, and inference* (2nd ed.). Cambridge, England: Cambridge University Press. (Original work published 2000)

Pfeifer, N., & Kleiter, G. D. (2010). The conditional in mental probability logic. In M. Oaksford & N. Chater (Eds.), *Cognition and conditionals: Probability and logic in human thinking* (pp. 153–173). Oxford, England: Oxford University Press.

Piaget, J. (1932). *The moral judgement of the child*. London, England: Kegan Paul, Trench, Trubner.

Rafetseder, E., Schwitalla, M., & Perner, J. (2013). Counterfactual reasoning: From childhood to adulthood. *Journal of Experimental Child Psychology*, *114*(3), 389–404.

Reichenbach, H. (1938). *Experience and prediction: An analysis of the foundations and the structure of knowledge*. Chicago, IL: University of Chicago Press.

Rock, I. (1984). *Perception*. New York, NY:: Scientific American Books.

Rojas, R. (1996). *Neural networks: A systematic introduction*. Berlin, Germany: Springer.

Schulz, A. W. (2011). Simulation, simplicity, and selection: An evolutionary perspective on high-level mindreading. *Philosophical Studies*, *152*(2), 271–285.

Schulz, A. W. (2018). *Efficient cognition: The evolution of representational decision making*. Cambridge, MA: MIT Press.

Schurz, G. (1997). *The is–ought problem: An investigation in philosophical logic* (Studia Logica Library). Dordrecht, Netherlands: Kluwer.

Schurz, G. (2001). What is 'normal'? An evolution-theoretic foundation of normic laws and their relation to statistical normality. *Philosophy of Science*, *68*(4), 476–497.

Schurz, G. (2008). The meta-inductivist's winning strategy in the prediction game: A new approach to Hume's problem. *Philosophy of Science*, *75*(3), 278–305.

Schurz, G. (2011). *Evolution in Natur und Kultur: Eine Einführung in die verallgemeinerte Evolutionstheorie* [Evolution in nature and culture: An introduction to the generalized theory of evolution]. Heidelberg, Germany: Spektrum Akademischer Verlag.

Schurz, G. (2012). Meta-induction in epistemic networks and the social spread of knowledge. *Episteme*, *9*(2), 151–170.

Schurz, G. (2016). Common cause abduction: The formation of theoretical concepts and models in science. *Logic Journal of the IGPL*, *24*(4), 494–509.

Schurz, G. (2019). *Hume's problem solved: The optimality of meta-induction*. Cambridge, MA: MIT Press.

Schurz, G., & Gebharter, A. (2016). Causality as a theoretical concept. *Synthese*, *193*(4), 1071–1103.

Schurz, G., & Hertwig, R. (2019). Cognitive success: A consequentialist account of rationality in cognition. *Topics in Cognitive Science*, *11*(1), 7–36.

Schurz, G., & Thorn, P. D. (2012). Reward versus risk in uncertain inference: Theorems and simulations. *Review of Symbolic Logic*, *5*(4), 574–612.

Schurz, G., & Thorn, P. D. (2016). The revenge of ecological rationality: Strategy-selection by meta-induction within changing environments. *Minds and Machines*, *26*(1–2), 31–59.

Skyrms, B. (2004). *The stag hunt and the evolution of social structure*. Cambridge, England: Cambridge University Press.

Skyrms, B. (2010). *Signals: Evolution, learning, & information*. Oxford, England: Oxford University Press.

Sober, E. (1993). *Philosophy of biology*. Boulder, CO: Westview Press.

Sober, E., & Wilson, D. S. (1998). *Unto others: The evolution and psychology of unselfish behavior*. Cambridge, MA: Harvard University Press.

Spelke, E. S., Phillips, A., & Woodward, A. L. (1995). Infants' knowledge of object motion and human action. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 44–78). Oxford, England: Clarendon Press.

Spirtes, P., Glymour, C., & Scheines, R. (2000). *Causation, prediction, and search* (Lecture Notes in Statistics, Vol. 81). Cambridge, MA: MIT Press.

Sterelny, K. (2012). *The evolved apprentice: How evolution made humans unique*. Cambridge, MA: MIT Press.

Taylor, S. E. (1989). *Positive illusions: Creative self-deception and the healthy mind*. New York, NY: Basic Books.

Thompson, B., Kirby, S., & Smith, K. (2016). Culture shapes the evolution of cognition. *PNAS*, *113*(16), 4530–4535.

Thorn, P. D., & Schurz, G. (2019). Meta-inductive prediction based on attractivity weighting: Mathematical and empirical performance evaluation. *Journal of Mathematical Psychology*, *89*, 13–30.

Thorndike, E. L. (1911). *Animal intelligence: Experimental studies*. New York, NY: Macmillan.

Todd, P. M., Gigerenzer, G., & ABC Research Group (Eds.). (2012). *Ecological rationality: Intelligence in the world*. Oxford, England: Oxford University Press.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 19–136). New York, NY: Oxford University Press.

Topitsch, E. (1979). *Erkenntnis und Illusion* [Knowledge and illusion]. Hamburg, Germany: Hoffman & Campe.

Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology* (Vol. 1, pp. 106–136). London, England: Penguin.

Weibull, J. W. (1995). *Evolutionary game theory*. Cambridge, MA: MIT Press.

Wenegrat, B. (1990). *The divine archetype: The sociobiology and psychology of religion*. Lexington, MA: Lexington Books.

Wilson, D. S. (2002). *Darwin's cathedral: Evolution, religion, and the nature of society*. Chicago, IL: University of Chicago Press.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103–128.

Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford, England: Oxford University Press.

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Knauff, Markus, editor. | Spohn, Wolfgang, editor.
Title: The handbook of rationality / edited by Markus Knauff and Wolfgang Spohn.
Description: Cambridge : The MIT Press, 2021. | Includes bibliographical references and index.
Identifiers: LCCN 2020048455 | ISBN 9780262045070 (hardcover)
Subjects: LCSH: Reasoning (Psychology) | Reason. | Cognitive psychology. | Logic. | Philosophy of mind.
Classification: LCC BF442 .H36 2021 | DDC 153.4/3—dc23
LC record available at https://lccn.loc.gov/2020048455