How Bad Is Incoherence?

Hal R. Arkes  
Ohio State University

Gerd Gigerenzer and Ralph Hertwig  
Max Planck Institute for Human Development, Berlin, Germany

Rationality is often defined in terms of coherence, assuming that a single syntactical rule such as consistency, transitivity, or Bayes’ rule suffices to evaluate behavior. Many normative claims made in psychological research follow this assumption. We argue that coherence-based norms are of limited value for evaluating behavior as rational. Specifically, we maintain that (a) there is little evidence that coherence violations are costly, or if they are, that people would fail to learn to avoid them; (b) adaptive rules of behavior can in fact imply incoherence; (c) computational intractability and conflicting goals can make coherence unattainable; and (d) coherence plays a key role in situations where it is instrumental in achieving functional goals. These observations lead to the conclusion that coherence cannot be a universal benchmark of rationality. We argue that smart choices need to be defined in terms of ecological rationality, which requires an analysis of the environmental structure and its match with cognitive strategies.

Keywords: coherence, correspondence, ecological rationality, intransitivity, propositional reasoning

A foolish consistency is the hobgoblin of little minds, adored by little statesmen and philosophers and divines. With consistency, a great soul has simply nothing to do.

—Ralph Waldo Emerson, from “Self Reliance,” 1841

[T]here is general agreement that rational choices should satisfy some elementary requirements of consistency and coherence.

—Amos Tversky and Daniel Kahneman, 1981

In psychological and economic research, rationality has often been equated with coherence. By coherence, we refer to a variety of rules for internal consistency, typically defined by propositional logic and probability theory. Coherence rules have been used as descriptive or normative models of behavior. In this article, we are concerned with the interpretation of these purely syntactical rules as normative for human behavior.

Table 1 provides a sample from this variety of coherence rules. In his influential research on reasoning and information search, Wason (1966, 1968) used two rules of propositional logic to define rational behavior: If $A \rightarrow B$, not-$B$, then not-$A$ (modus tollens), and if $A \rightarrow B$, $A$, then $B$ (modus ponens). Wason reported that most people’s judgments were influenced by the content of the $A$’s and $B$’s and thus predictably violated these rules of inference. He interpreted this phenomenon as irrational behavior, and moreover, attributed people’s rejection of his normative view to their “incorrigible conviction that they are right when they are, in fact, wrong” (Wason, 1981, p. 356). Similarly, Tversky and Kahneman (1983) used another syntactical rule to define rational behavior, set inclusion: $p(A) \geq p(A \setminus B)$, that is, the probability of an event $A$ cannot be smaller than that of events $A$ and $B$. For selected contents in the $A$’s and $B$’s, most people violated this rule. Tversky and Kahneman (1983) called this violation the con-
junction fallacy and from its “stubborn persistence” concluded “we suspect that incoherence is more than skin deep” (pp. 300, 313). Piaget and Inhelder (1975) had used the same logical rule as a standard of rationality, but reported that most children from age eight onward reason in accordance with class inclusion (see Hertwig and Gigerenzer, 1999, for a resolution of these seemingly contradicting findings). Whatever the descriptive result reported, in all of these cases the rationality of behavior is evaluated solely on the basis of a syntactical rule.

Table 1 provides further examples of coherence rules being equated with rationality and their violations being interpreted as cognitive errors. We discuss some of these in more detail shortly. In this article, we are not concerned with the question of how often people violate coherence, or under what conditions, but rather with the more fundamental question whether coherence rules actually provide sufficient criteria for rationality.

Those who emphasize that people violate one of these rules typically lay the blame on the human mind rather than on the syntactical norm. In their groundbreaking Science article, Tversky and Kahneman (1974) listed six cognitive “biases” that they attributed to the use of the representativeness heuristic; all of them—for instance, insensitivity to sample size, misconceptions of regression, and insensitivity to prior probabilities—represent violations of coherence rules. Most important for the present analysis, these violations have been suggested to have detrimental material consequences (e.g., Thaler & Sunstein, 2008; Yates, 1990).

In contrast, a small group of psychologists have advocated correspondence rather than coherence as the benchmark for rational behavior. Correspondence concerns purely syntactical relations between behaviors. Correspondence, in contrast, concerns relations between behavior and the environment, using measures such as how healthy, rich, successful in school, happy in marriage, or accurate in judgments people are, and ultimately how successful an organism is at survival and reproduction (Cosmides, 1989; Cosmides & Tooby, 1992; K. R. Hammond,
We will refer to these norms as *ecological*, that is, *functional* norms. Coherence and ecological norms are not the same. Believing that there is a 99% probability of HIV not being the cause of AIDS and therefore that antiretroviral drugs are not useful for patients, and believing there is a 1% probability that the opposite is true represents a fully coherent set of beliefs. Albeit coherent, this set of beliefs, judged by the available evidence, is both inaccurate and severely dysfunctional. In South Africa, a similar belief held by the Mbeki government is estimated to have cost the lives of some 330,000 people because a timely treatment program was not implemented (Chigwedere, Seage, Gruskin, Lee, & Essex, 2008). By contrast, believing that there is a 99.9% probability that HIV is the cause of AIDS and a 1% probability that this not true is incoherent but ecologically more reasonable in the sense that this set of beliefs corresponds more closely to what is known and would have led to the life-saving implementation of the antiretroviral treatment program.

Nevertheless, as Table 1 illustrates, one dominant view in the literature is that decisions should be evaluated by their internal quality, that is, coherence, rather than by their external consequences, that is, ecological rationality (Todd, Gigerenzer & the ABC Research Group, 2012). In the words of J. S. Hammond, Keeney, and Raiffa (1999, p. 110): “Although many people judge the quality of their own and others’ decisions by the quality of the consequences—how things turn out—this is an erroneous view.” Instead, J. S. Hammond et al. assert that rational decisions are ones that are coherent.¹

In this article, we argue that coherence rules do not comprise a universal bedrock principle of rationality. Specifically, we advance four arguments:

1. **Lack of evidence that incoherence is costly.** There is little evidence that coherence violations incur material costs, or if they do, that people would fail to learn to avoid them.

2. **Adaptive incoherence.** Application of adaptive cognitive strategies can imply violations of coherence norms.

3. **Intractable coherence.** Coherence is often not attainable because of computational limits or conflicting goals.

4. **Ecologically rational coherence.** In specific environments, coherence can be instrumental in reaching functional goals such as fairness and survival.

We begin with a systematic review of the evidence for the widespread assumption in the literature that violations of coherence will incur material costs.

### Are Coherence Violations Costly?

One reason why judgment and decision making researchers have been primarily interested in coherence is the presumption that violations of coherence incur material costs. A person who violates norms based on logic or probability would seem to be highly vulnerable to exploitation and would also seem to be at a huge disadvantage in interacting successfully with the environment. Is there research that supports this presumption? To answer this question, we conducted several systematic Web of Knowledge searches for the major coherence rules reported in the literature, and, in addition, we conducted a survey among researchers. We begin with the most prominent argument, the “money pump.”

### Is There Evidence for a Money Pump?

Suppose a person prefers option A to B, B to C, and C to A, thus violating transitivity (see Table 1). If a person is willing to pay money to continually substitute a more preferred option

---

¹ We emphasize that neither coherence nor correspondence represents a unitary construct. Shepard (2001), for instance, defined correspondence in his mirror theory by arguing that through natural selection the evolved mind veridically mirrors significant objects and events of the external world. In his lens model, Brunswik (1952) defined correspondence by the degree to which perception and cognition accurately infer the probabilistic structure of the world. Simon (1990) later defined correspondence in terms of a pair of scissors, with cognition and the environment as the two blades. All these definitions are not identical: For Shepard, the mind reflects the world; for Brunswik, the mind infers the world; and for Simon, the mind must fit closely to the environment—although, in contrast to Shepard, Simon saw the two as complementary rather than as mirror images (Todd & Gigerenzer, 2001).
for a less preferred one, the intransitive relation between the three options will cause that unfortunate person to become a money pump (Condorcet, 1785).

Is there evidence that people actually turn into money pumps? We conducted a literature search, which found 107 articles. In reviewing them we could not find a single demonstration of a person becoming a money pump, that is, being continually exploited due to intransitive choices between three or more options. This result is consistent with arguments that the existence of money pumps is highly implausible (e.g., Schick, 1986; Levi, 2002) and that money pump arguments are “logical bogeymen” (Lopes, 1996, p. 187).

Is There Evidence That Preference Reversals Are Costly?

According to Tversky, Sattath, and Slovic (1988), “normatively equivalent procedures for assessing preferences should give rise to the same preference order” (p. 371). Studies indicate that this rule, called procedural invariance (see Table 1), is sometimes violated. Such a violation can lead to preference reversals (e.g., Lenert & Treadwell, 1999; Shafir, Simonson, & Tversky, 1993; for a critique see Sher & McKenzie, 2014), and a preference reversal could potentially be costly. For example, suppose a person is asked which of two lotteries, A and B, she would choose. The expected value of both lotteries is about equal, but A has a high probability of winning while B yields a high dollar value if she wins. She chooses A. Then she is asked to price each of the lotteries, that is, to state the lowest price at which she would be willing to sell each lottery if she owned it. She states a $3 value for A but a $5 value for ticket B. This discrepancy between choosing and pricing is a violation of procedural invariance, called preference reversal. Such a reversal could be easily exploited. Assume the person starts with $2 in her pocket and pays $5 for ticket B, thus incurring a $3 deficit. The market would then trade the “preferred” lottery ticket A to the person in exchange for B. The person would gladly accept, because of her preference for A. Now the market would buy ticket A for $3, the price the person stated for ticket A. The person would now have no tickets and no money, thus becoming worse off than before the series of exchanges began. This arbitrage procedure illustrates the manner in which a preference reversal could potentially be costly. Should this arbitrage procedure continue indefinitely, the victim would turn into a money pump.

We conducted another systematic Web of Science literature search using the key words procedural invariance and preference reversal. The search produced 1,036 articles, 32 from the first keyword search and 1,004 from the second, with surprisingly few articles appearing in both searches. We reviewed the titles and abstracts of each and found that only four of these studies empirically tested whether people committed a preference reversal and then incurred a cost due to their violation of procedural invariance (Berg, Dickhaut, & O’Brien, 1985; Cherry, Crocker, & Shogren, 2003; Cherry & Shogren, 2007; Chu & Chu, 1990).

The study by Cherry et al. (2003) showed that an initial preference reversal rate of 33% declined by half after three rounds and to below 5% after the ninth round of arbitrage. Berg et al. (1985) similarly showed that arbitrage reduced the average financial cost of each preference reversal by 39%, although it did not reduce the already low frequency of preference reversals. Chu and Chu (1990) reported that after a maximum of two arbitrage transactions, the number of preference reversals (summed across groups) dropped from 39 before arbitrage was introduced to 5 afterward. In a group that had no limit on the number of arbitrage transactions, this number dropped from 55 to only 1 after three transactions. Cherry and Shogren (2007) found that after three rounds of arbitrage, preference reversals disappeared in all but 2 of their 54 participants. Thus, results from these four studies indicate that participants quickly learn to avoid or at least strongly reduce preference reversals.

There are also experiments that tested the persistence of preference reversals in repeated nonarbitrage situations, such as in auctions. In

\footnote{We used each of the key words money pump and Dutch book in combination with each of the two stems intransitive and transitive. We checked the title and abstract of each article to identify those that test the claim that persons who violate transitivity are at risk of being continually exploited. When there was any indication that the study empirically investigated this risk, we read the complete article.}
these experiments a person who violated procedural invariance could not be money pumped, but it was possible to assess whether their violations diminished over repeated trials. This is a much weaker intervention than arbitrage, because the cost of procedural invariance violations is not apparent. For example, Cox and Grether (1996) asked subjects to consider the following two gambles taken from the landmark study of Lichtenstein and Slovic (1971):

- "P bet": 35/36 chance to win $4 and 1/36 chance to lose $1
- "$ bet": 11/36 chance to win $16 and 25/36 chance to lose $1.50

The former is called the P bet, because it has a high probability of winning. The latter is called the $ bet, because it has the higher dollar amount to be won. Lichtenstein and Slovic (1971) found that many people preferred to play the P bet but assigned a higher price to the $ bet, thereby exhibiting a preference reversal and thus a violation of procedural invariance. Cox and Grether (1996) assessed each person’s preference for each gamble using a nonmarket-pricing and a market-pricing procedure. If a person exhibited a higher price for one of the two gambles using one procedure but a higher price for the other gamble using the other procedure, procedural invariance would be violated. Cox and Grether found that with a second price sealed bid auction as the market pricing task, the frequency of such preference reversals dropped substantially over five contests of P bets versus $ bets. Ordóñez, Mellers, Chang, and Roberts (1995) found analogous results. Wedell and Böckenholt (1990), using only hypothetical gambles, found that if each gamble was played multiple times rather than once, the preference reversals greatly diminished. Braga, Humphrey, and Starmer (2009) found that market exposure eliminated the standard preference reversal (choice of the P bet but higher price for the $ bet) but increased the opposite preference reversal (choice of the $ bet but higher price for the P bet). The authors attributed this unusual finding to an aspect of their procedure (see p. 404 of their article). Note that no participant’s preference reversal was punished with any arbitrage procedure in any of these studies.

Many researchers have found that the use of real as opposed to hypothetical rewards significantly reduced the frequency of preference reversals even in nonarbitrage situations (Bohm & Lind, 1993; Bohm, 1994; Berg, Dickhaut, & Rietz, 2010, who provide a review of this issue). On the other hand, Loomes, Starmer, and Sugden (2010) found that over the course of six P bet/$ bet contests using real money, the frequency of procedural invariance violations did not diminish. Lichtenstein and Slovic (1973) and Grether and Plott (1979) concur that contests between P bets and $ bets using real money do result in violations of procedural invariance. Again, we point out that such studies provided no arbitrage opportunity.

With regard to violations of procedural invariance, let us acknowledge two important points. First, we do not question the existence of preference reversals in nonarbitrage situations, particularly when there is a single P bet versus $ bet contest. However such studies generally do not address the existence or magnitude of any cost of such violations of procedural invariance. Second, there are many different types of preference reversals. Two of the more prominent are differing preferences exhibited by (a) joint versus separate assessment of two items (e.g., Hsee, Loewenstein, Blount, & Bazerman, 1999) and (b) choice versus matching procedures (e.g., Slovic, 1995). To the best of our knowledge no studies using these procedures have either assessed the cost of such violations or tested their durability in the face of arbitrage.3

In conclusion, the overwhelming majority of 1,036 articles did not test whether violations of procedural invariance turn people into money pumps or otherwise impose any costs; the only four that did so found that arbitrage substantially reduced the number of preference reversals or their costs. In addition, even in many nonarbitrage market situations, persons who received financial feedback after each of multiple

---

3 One reviewer suggested that irrationalities occur in the financial markets as a result of violations of coherence norms and that an army of analysts continually seek arbitrage opportunities engendered by these violations. However, we question whether such irrationalities are necessarily coherence violations. For example, Hirshleifer and Shumway (2003) found that returns on stock exchanges varied with the weather. Such financial activity may be potentially costly to those who manifest such behavior, but it does not follow that this behavior violates one of the coherence norms (such as in Table 1).
trials significantly diminished their frequency of preference reversals.

Is There Evidence That Inconsistency Attributable to Framing Is Costly?

According to Tversky and Kahneman (1981), different representations of information that are logically equivalent should yield the same preference. Once again, our question is not whether logically equivalent frames are truly informationally equivalent (for a critique, see Sher & McKenzie, 2006; Frisch, 1993; Kühberger, 1998; Mandel, 2001), but whether there is evidence that violating this coherence rule, called descriptive invariance (see Table 1), is costly.

We performed a systematic Web of Science literature search using the search term descriptive invariance. Among the 11 hits we found no empirical demonstration of costs. We then consulted Kühberger’s (1998) meta-analysis of “framing” and examined every citation in the bibliography that appeared to be relevant. We also examined the 237 articles published after 1998 that referenced Kühberger (1998) and the reference list of Keren’s (2011) Perspectives on Framing. Although there are hundreds of studies on the Asian Disease Problem and similar demonstrations, we found only one group of studies that sometimes used indirect indicators of cost.4

These studies pertain to negotiating situations in which factors to be negotiated are described to the participants in terms of either potential gains or potential losses (e.g., Neale & Bazerman, 1985; Bottom & Studt, 1993). Depending on various contextual factors, positive framing usually leads to more settlements than does negative framing, but negative framing can also lead to more combined profits for the negotiating parties (e.g., Bottom, 1990). When these experiments are run in graduate classes of business administration (MBA), the number of successful settlements can influence the course grade. This outcome could be said to comprise an indirect measure of costs. Note that these studies did not include arbitrage or opportunities for learning.

In conclusion, a large number of studies have demonstrated that descriptive invariance is often violated. But apart from a few negotiation studies that make course credits dependent on violating or obeying descriptive invariance, there is little empirical evidence that violations of descriptive invariance incur costs.

Is There Evidence That Violations of Independence Are Costly?

Property Alpha (see Table 1) is one formulation of what is called the independence condition, or the independence of irrelevant alternatives. If a person prefers A over B, adding a third option C to the choice set should not reverse the preference of A over B. Violations of independence have been demonstrated in both humans (e.g., Huber, Payne, & Puto, 1982; Shafir et al., 1993) and other animals (Latty & Beekman, 2011). We performed a systematic Web of Science literature search using the search terms regularity axiom, Chernoff condition, and independence of irrelevant alternatives. None of the 200 publications identified by the search demonstrated actual costs of either single or repeated violations of regularity.

Did We Overlook Evidence for Costs of Coherence Violations?

Our systematic literature searches suggest that there is little empirical evidence that violations of coherence norms are costly, or if they are, that they survive arbitrage and learning. To ascertain whether we had overlooked any evidence, we conducted a survey among experts. On October 26, 2011 we posted a request on the electronic mailing list of the Society for Judgment and Decision Making. One of us asked the approximately 1,000 researchers on the list: “I am seeking examples of instances in which violations of coherence make people less rich, less healthy, or have less accurate beliefs. . . .” A total of 10 studies

---

4 The term framing has been used to describe many phenomena that do not strictly pertain to the original usage of the term by Kahneman and Tversky (1984). For example, “framing studies” comparing graphical versus numerical presentation of data were not included in our search. We included only studies in which the independent variable comprised manipulation of the data presentation in either the loss or gain frame.
were proposed. Of these 10, some were new but only one warranted further scrutiny.\(^5\)

In that study, Bonini, Tentori, and Osherson (2004) showed that participants allocated monetary bets that violated the conjunction rule. Although their betting scenario was hypothetical, Tversky and Kahneman (1983, p. 303) found a similar result in a betting scenario involving real payoffs. These two studies come as close as any study we know of to demonstrating costly effects of coherence violations. However, neither of these studies tested whether these violations would survive arbitrage. Thus, it seems fair to conclude that neither our systematic literature searches nor the researchers in the field could identify substantial evidence indicating that violations of coherence are costly.

We emphasize that lack of evidence of costs is not the same as evidence for lack of costs. The fact that empirical studies of the costs of coherence violations are so rare could mean several things. First, many researchers appear to assume that it is self-evident that violations of coherence norms are costly, thus requiring no empirical proof. Second, researchers may have investigated how costly the violations are and found no evidence, thus finding it difficult to publish this null result (the “file-drawer problem”; Rosenthal, 1979).

Are Coherent Beliefs More Valid?

Why is there no evidence for the widespread belief that violations of coherence are costly? One explanation is that researchers took this belief for granted and did not even look for evidence. A more radical explanation, however, is that, in many domains, there actually exists no correlation between incoherence and costs. Rather, the correlation might be between invalid beliefs and costs, while there is none between incoherent beliefs and their validity. We are aware of only one study that compared coherence and validity of knowledge about a consequential decision, whether or not to take a PSA (prostate specific antigen) test. Berg, Biele, and Gigerenzer (in press) measured the consistency and the validity of beliefs about prostate cancer screening, as well as the actual history of having taken a PSA test, among 125 academics, almost all economists. Consistency was measured asking the participants for their estimates of base rates, sensitivities, specificities, and posterior probabilities, and then determining whether the set of estimates was mutually consistent in terms of Bayes’ Rule. Validity of belief was measured by comparing the same estimates with the scientific evidence. The correlation between consistency and accuracy was precisely zero. For instance, the most consistent economist had the most invalid beliefs. Having valid beliefs and thereby understanding the pros and cons of PSA testing is essential for making good decisions. At least in this study, coherence proved to be uncorrelated with valid knowledge, and thus did not foster good health decisions.

Can Violations of Coherence Be Beneficial?

Oddly, despite the scarcity of evidence, the reigning assumption in the experimental literature is that violations of coherence rules are costly. Rarely has the opposite possibility been considered: Is there experimental evidence that violations of coherence have material benefits in specific situations? We are aware of only one study that investigated this possibility. Berg, Eckel, and Johnson (2011) tested more than 800 participants on 72 problems, measuring consistency of preferences over time in three different ways. Together with expected utility maximization and Bayesian probability updating, consistency in time is, according to Arrow (1982), the most important manifestation of rationality. Berg et al. reported that those who violated time-consistency and expected utility theory earned higher average payoffs than those who did not, even after controlling for risk attitude, demographic variables such as household income, and success in school.

\(^5\) Lichtenstein and Slovic’s classic article (1973) was mentioned by multiple researchers. However, this study did not demonstrate that the amount of money earned was related to the number of preference reversals. The other suggested studies used only hypothetical outcomes (Azar, 2011), demonstrated judgment errors that were not coherence violations (DiBonaventura & Chapman, 2008), showed that respondents’ low numeracy impaired their decision-making performance (Reyna, Nelson, Han, & Dieckmann, 2009), or showed that persons who performed poorly on a battery of judgment/decision-making problems also manifested certain maladaptive risk-taking behaviors (Parker & Fischhoff, 2005). In the latter study the authors did not relate coherence violations to diminished outcomes in that specific domain.
Adaptive Rules Can Imply Incoherence

In contrast to the assumption that some form of coherence—logical or probabilistic—is a defining feature of rationality, successful behavior can imply coherence violations. We refer to this as adaptive incoherence. This includes adaptive behavior that violates norms based on (a) transitivity, (b) noncontradiction, (c) consistency, and (d) truth-table logic. Again, we point out that these coherence rules are content-blind; that is, they are impervious to the biological, social, or psychological goals of organisms. Adaptive rules of behavior implement these goals. Adaptive rules of behavior do not generally lead to violations of coherence but, depending on the structure of the environment, can result in adaptive incoherence.

Fitness Maximization Can Imply Intransitivity of Choice

Transitivity is an axiom on which virtually every theory of rational choice is based:

If $A \succeq B$, and $B \succeq C$, then $A \succeq C$,

where $A$, $B$, $C$ are elements of the choice set $S$, and $A \succeq B$ denotes that $A$ is weakly preferred to $B$.

In their article “Violations of Transitivity Under Fitness Maximization,” Houston, McNamara, and Steer (2007) considered the case of an animal facing the dangers of starvation and predation during the winter. The animal will perish if its energy reserves drop to zero, while the maximal level of energy reserves is 40. One unit of energy expenditure is needed to survive every time unit. The choice set consists of three patches to forage, each with a probability $p$ of getting 2 energy units, and a probability $q$ of death from a predator:

$A: p = 50, q = 0,$

$B: p = 55, q = .00001,$

$C: p = .75, q = .00001.$

Houston et al. (2007) showed that for certain ranges of energy reserve levels the ordering of choices is intransitive. For instance, if the reserve levels are anywhere between 15 and 38, the optimal choice in terms of fitness is to choose $B > A$, $C > B$, but $A > C$—a violation of transitivity.

Optimal rules of foraging are state-dependent. For example, if the reserve levels are greater than 38, then the animal’s choices among $A$, $B$, and $C$ will be transitive, not intransitive. However, the energy reserve of the animal making the choice among three options is not part of the definition of transitivity, which is intended to be a domain-general criterion of rationality. Our point is that coherence cannot be a universal criterion for rational behavior: animals whose reserves are between 15 and 38 and who honor rationality will be at a nutritional disadvantage compared to those who do not. Being dead strikes us as too great a cost to honor transitivity. Houston et al. (2007, p. 367) conjectured that state-dependent intransitive choice patterns may occur under many common conditions in the animal kingdom.

Houston et al.’s (2007) analysis concerns risk, not uncertainty. That is, even when all probabilities are known for sure, coherence as defined by the transitivity axiom is not sufficient for defining adaptive behavior. Similar analyses have been conducted by Gilpin (1975), Laird and Schamp (2006), and discussed by Kacelnik, Schuck-Paim, and Pompilio (2006) and Wimsatt (2006). The issue is not coherence per se, but solely that on the basis of the transitivity axiom alone—without taking into account the state of the organism, its goals and the environment—we cannot distinguish between rational and irrational behavior.

Adaptive Social Rules Can Imply Inconsistency of Choice

One basic form of “contraction consistency” is known as Property Alpha (see Table 1), or “independence of irrelevant alternatives” (Sen, 1993):

$A(\Omega)$ and $A \in \Psi \subseteq \Omega \Rightarrow A(\Psi),$

where $\Omega$ and $\Psi$ denote two (nonempty) sets of alternatives and $A(\Omega)$ denotes that alternative $A$ is chosen from the set $\Omega$. Property Alpha says that if $A$ is chosen from $\Omega$, and if $A$ belongs to a subset $\Psi$ of $\Omega$, then $A$ must be chosen from $\Psi$ as well. The following two choices are logically
inconsistent in the sense that they violate Property Alpha:

1. A is chosen over B given the options \{A, B\}, and
2. B is chosen over A given the options \{A, B, C\}.

The question is again whether consistency of choice, and by extension, its rationality, can be decided upon independent of context, without referring to something external to choice, such as a person’s goals, values, and environment. Consider the following problem adapted from Sen (1993):

At a dinner party, a fruit basket is passed around. When it reaches Mr. Polite, one apple is left in the basket. Mr. Polite decides to behave decently and picks nothing (A) rather than the apple (B). If the basket had contained another apple (C), he could reasonably have chosen B over A without violating standards of good behavior.

Choosing A over B from the choice set \{A, B\} and choosing B over A from the choice set \{A, B, C\} violates Property Alpha, even though there is nothing irrational about Mr. Polite doing justice to his name. In fact, to avoid violating consistency, he would have to suspend his good manners. Sen (2002) concluded that the idea of internal consistency of choice “is essentially confused, and there is no way of determining whether a choice function is consistent or not without referring to something external to choice behavior (such as objectives, values, or norms)” (pp. 121, 122). We emphasize that the term environment needs to be added to this list—after all, the consistency of behavior is not determined by its politeness or impoliteness alone, but also by the available choice set and others’ behavior. Violating Property Alpha can be functional in social situations.

Sen’s (1993) dinner party example suggests another important observation that has received little attention in the treatment of consistency, namely that behavioral inconsistencies do not imply inconsistent strategies. Since the economist Paul Samuelson pioneered the revealed preference approach in the 1930s (e.g., Samuelson, 1938), treatments of consistency have focused on observable choice rather than on strategies producing choice. Inconsistency in behavior, however, does not necessarily mean that the person uses an inconsistent strategy (Hertwig & Gigerenzer, 2011). Instead, it can be a direct consequence of consistently using a strategy, such as “If there is only one food item left, take it only when everyone else has had at least one.” Depending on the environment—his position in the choice sequence, the number of items left, and the possibility of replenishment—Mr. Polite, who consistently relies on this heuristic, can behave in a seemingly inconsistent way. This argument has a strong implication: If one analyzes observable choice only, without a theory of personal values or strategies underlying choice, it is impossible to determine the rationality of the choice. This implication applies to all presumed cognitive fallacies in Table 1.

Adaptive Social Rules Can Imply Violations of Propositional Logic

Adaptive behavior has goals other than logical consistency. One is to exchange resources (e.g., information, food, support) with other people. According to Trivers’ (1971, 2002) theory of reciprocal altruism, humans possess both altruistic and cheating tendencies. Therefore, agents in a social contract need to search for information revealing possible cheating (Cosmides, 1989). To do so efficiently, the agent’s perspective is essential: People need to find out whether they are being cheated, not whether they are doing the cheating. Logic, in contrast, has no perspective; it is content-blind.

Consider the day-off problem, similar to the four-card Wason selection task (see Table 1). It describes a social contract between an employer and an employee (Gigerenzer & Hug, 1992):

If an employee works on the weekend, then that person gets a day off during the week.

In this task there were four cards (representing files of previous employees), which specified on one side whether an employee worked on the weekend and on the other whether that person got a day off. The front side of the four cards read “worked on the weekend,” “did get a day off,” “did not work on the weekend,” and “did not get a day off.” In Gigerenzer and Hug’s study, one group of participants was cued into the role of an employee and asked to check only those cards that could reveal whether the rule was violated. The majority (75%) selected “worked on the weekend” and “did not get a day off.” This combination of cards is consistent
with both truth table logic and the functional goal of detecting a cheater. But when another group of participants was cued into the role of an employer, the large majority selected “did not work on the weekend” and “did get a day off.” In the employee’s condition, by contrast, only 2% of the participants had selected this combination of cards. Cued into the employer’s perspective, people’s reasoning was inconsistent with the truth table but again consistent with the goal of not being cheated. Search for information thus obeyed a “Machiavellian logic” (avoid being cheated) rather than propositional logic (avoid being logically incoherent). Other social rules showed the same strong effects (Gigerenzer & Hug, 1992).

In sum, we cannot determine the rationality of behavior by logic alone, without referring to something external such as the organism’s goals. In the context of social contracts, the variation of an agent’s perspective via social roles suggests that the goal people pursue is cheater detection, not following the modus tollens (Cosmides & Tooby, 1992; Gigerenzer, 2000; see also Cheng & Holyoak, 1985). Our point is not new. Wilhelm Wundt (1912/1973), the father of experimental psychology, concluded that logical norms have little to do with thought processes, and that attempts to apply them in order to learn about psychological processes have been “absolutely fruitless” (p. 149). Some psychologists have slowly come around to acknowledging Wundt’s insight (Evans, 2002; Oaksford & Chater, 1991). For instance, Lance Rips, who had previously argued that deductive logic plays a central role in cognitive architecture (Rips, 1994), declared that he would no longer defend this expansionist “Theory of Everything” anymore (Rips, 2001, p. 215).

On the Impossibility of Always Being Coherent

We will now address another limit to coherence as a universal norm of behavior. Even if there were only one coherence rule (see Table 1) and even if coherence served the goals of an organism, there would nonetheless be (numerous) situations in which it is impossible for the organism to be coherent. Specifically, there are situations in which (a) consistency is computationally intractable and (b) people have multiple conflicting goals.

Consistency Can Be Computationally Intractable

Several cognitive consistency theories (Arkes & Garske, 1977, Chapter 9) have been proposed, the most prominent of which is cognitive dissonance theory. Festinger (1957) pointed out that he replaced the term inconsistency with dissonance, a term with a less logical connotation. The theory has two basic hypotheses (p. 3):

1. The existence of dissonance, being psychologically uncomfortable, will motivate the person to try to reduce the dissonance and achieve consonance.
2. When dissonance is present, the person will not only try to reduce it but actively avoid situations and information that would likely increase it.

Festinger explicitly treated dissonance like other basic motivating factors, such as hunger and frustration, which need to be resolved (p. 3). By cognition, he meant “any knowledge, opinion, or belief about the environment, about oneself, or about one’s behavior” (p. 3). He discussed how dissonance is resolved, and where it persists. Yet a fundamental problem arises for his and similar consistency theories (Heider, 1946, 1958; Newcomb, 1956; Osgood & Tannenbaum, 1955).

The problem is how to determine which beliefs are inconsistent. Festinger defined dissonance and consonance as “relations which exist between pairs of ‘elements’” (p. 9). Two elements “x and y are dissonant if not-x follows from y” (p. 13). This formulation is identical with the principle of noncontradiction discussed earlier. In reality, the cognitive elements (beliefs) may not always be defined precisely enough for their logical status to be decided. Also, many cognitive elements pertain to ambivalent evaluations that may be somewhat consistent and somewhat inconsistent with other elements. Dissonance researcher Aronson (1968, p. 8) expressed his exasperation with the definition of dissonance by stating that if one wanted to ascertain whether two elements were dissonant, one had to ask Festinger. But even if semantic ambiguity can be resolved in order to
ascertain whether two elements are consistent, a serious computational problem will emerge.

Consider a person with \( n \) elements or beliefs. In order to check the pairwise consistency of \( n \) elements, \( n(n-1)/2 \) pairs need to be compared. According to evolutionary anthropologists, modern humans lived in groups of approximately 150 members, and these groups depended on extensive personal knowledge based on face-to-face interaction for their stability and coherence [emphasis added] through time (Dunbar, 1993, p. 691). But is checking consistency computationally possible? Assume that a person has a relationship with each of these 150 people and holds 20 beliefs about each person, meaning that the person holds a total of 3,000 beliefs. Consequently, it would be necessary to check 4,498,500 pairs for their consistency. If checking a single pair takes as little as one second, that would necessitate spending over six years (8 hours per workday) checking for the consistency of those beliefs. Admittedly, reducing the consistency checks to only those pairs that are relevant for the specific purpose may appear to contain this computational explosion. Yet this strategy would first require a check on each candidate’s beliefs to ascertain whether they might be relevant. Even with only 20 relevant beliefs about two other adults, a person would still have to check 1,770 pairs of elements. However, researchers who champion cognitive consistency typically do not limit the scope of beliefs that are being checked for consistency. In addition, any boundaries to the scope of checking would require a model of the stopping rule people use to decide when to stop checking. None of the consistency theorists listed above have addressed this limitation to their theory, although philosophers (e.g., Cherniak, 1986) and artificial intelligence researchers (e.g., van Rooij & Wareham, 2012) have noted this fundamental problem.

Because the number of pairs increases exponentially with the number of elements, checking for consistency quickly becomes a computationally intractable problem when \( n \) increases. The same problem emerges for theories about networks of beliefs, such as probabilistic inference in Bayesian belief networks. Checking for consistency within these networks is NP-hard (Cooper, 1990), as are approximations (Dagum & Luby, 1993). For mere mortals, holding fully consistent beliefs—if one has more than a few to begin with—becomes quickly computationally intractable.

**Multiple Goals Can Make Coherence Impossible**

Rules of coherence—such as transitivity, consistency, and noncontradiction—are typically defined while assuming a single dimension of preference or goal. Our argument is this: If a person has more than one goal, or if the environment imposes conflicting goals (see below), it is not invariably possible to be simultaneously consistent on all. This can force people to violate noncontradiction:

\[
\text{If } p_i, \text{ then not } \neg p_i, \text{ for all } i \in G,
\]

where \( G \) is the set of all relevant goals for a given decision. This argument is orthogonal to the previous one, because it can hold for as small a number of beliefs as two. The general point is that the attempt to be consistent with respect to one goal or dimension (e.g., physicians trying to do what is in the best interest for one’s patients) can leave a person no choice but to be inconsistent with respect to a second one (e.g., trying to avoid lawsuits by practicing defensive medicine).

Kohlberg (1981), the stage-theorist of the development of moral reasoning, devised the Heinz dilemma to demonstrate such a conundrum. A woman was near death from cancer. A drug might save her, but the pharmacist charged 10 times what the drug cost him to produce. Heinz, the sick woman’s husband, could not raise enough money and asked the pharmacist to sell it at a lower price or let him pay later. The pharmacist said no. Heinz, a consistently law-abiding citizen, grew desperate and broke into the store to steal the drug. Kohlberg’s famous dilemma represents a situation in which a person cannot simultaneously conform to both goals: to always abide the law and to always look after her or his partner’s welfare.

Kohlberg’s story is hypothetical, but dilemmas between conflicting goals appear to be the rule rather than the exception. Merenstein (2004) was a physician in Virginia who practiced evidence-based medicine. Because there is no evidence that early detection of prostate can-
The Ecological Rationality of Coherence

Our treatment of coherence up to this point could be misconstrued to imply that coherence is of no value as a benchmark of adaptive behavior. This is not our view. What we do oppose is the widespread assumption that coherence is a universal, domain-general criterion of rationality (see Table 1). There are many environments—and we have described some—in which violations of coherence are a consequence of adaptive strategies. However there are also environments in which cognitive faculties for reasoning in accordance with norms of coherence are crucial in the service of an organism’s goals. For illustration, we turn to two recurrent classes of tasks in which adhering to coherence may be particularly important in serving correspondence goals. The first is an “old” evolutionary task; the second is germane to a cultural institution, the law.

Transitivity Enables Inferring Status Hierarchies

Consider the ecological rationality—as opposed to the logical rationality—of transitivity. The ability to reason in accordance with transitivity can enable organisms to exploit known relationships in order to deduce unknown ones: For example, using the known relationship $A > B$ and $B > C$ it is possible to infer—assuming transitivity—that $A > C$. Transitive inference represents an adaptive cognitive ability in species in which status hierarchy is a component of social organizations, a ubiquitous component across a wide variety of species including fish, birds, and primates (Chiao, 2010). Grosenick, Clement, and Fernald (2007) showed that using transitive inference, male fish (A. burtoni) can successfully infer a conspecific’s social rank within a hierarchy by watching pairwise fights between rivals. Because male fish engage in regular aggressive bouts that determine their access to territory and resources, winning these bouts is crucial to male reproductive fitness. Hence, the “ability to infer the relative strength of rivals before engaging them in potentially costly fights should be highly adaptive” (p. 429).

Grosenick et al.’s (2007) investigation offers two insights: First, the ability to draw transitive inferences is a domain-specific adaptation. It
helps an organism to determine whether it should engage in a fight prior to the fight; once the fight has begun, the ability to display unpredictable behavior becomes important. Generalizing Sperber’s (1996) distinction between “proper” (the context in which a strategy evolved) and “actual” domains (the context in which it is now used) to our topic, we can say that at least some coherence rules such as transitivity have proper domains. That is, cognitive abilities and decision-making strategies have likely evolved to enable organisms to perform better in real-world environments. Thus, evolutionary adaptiveness is domain-specific. Modern theories of coherence rationality divorce these rules and associated abilities from their proper evolutionary domain and treat coherence as a domain-general norm of rationality (in Sperber’s terminology, the “actual domain”).

Furthermore, the ability to make transitive inferences is not important in and of itself. Rather, it operates in the service of functional goals such as avoiding the risk of getting injured in contests, losing contests, and ultimately losing access to territory and resources. A male fish, oblivious to transitivity, may observe other fish fighting and then attack the alpha male. Such an inability to make transitive inferences may result in serious injury or death. Hence, there is evolutionary pressure to infer the hierarchy so as not to jeopardize vital correspondence goals.

**Consistent Framing Fosters Equal Protection of the Law**

For any legal system that honors equal protection of the law—the right of all persons to have the same access to the law and courts and to be treated equally, both in procedures and in the substance of the law—framing of evidence is crucial. Framing is defined as different representations of logically equivalent information, and one speaks of a “framing effect” if people react differently to these frames. Consider two such frames, conditional probabilities and natural frequencies, and the presentation of evidence in the courtroom. If one expert witness testifies in terms of natural frequencies in one trial, and a second expert witness testifies in terms of conditional probabilities in another trial, equal procedures are not followed. As a consequence, different verdicts are likely to result, because conditional probabilities trigger the overestimation of the probability that a defendant is the source of a DNA trace, and by implication, the probability of guilt. In contrast, natural frequencies strongly reduce the risk of this systematic confusion and, in addition, foster transparency (Gigerenzer, 2002). Thus, consistent use of the same transparent frame increases understanding of evidence by jurors and judges and the chance that defendants be treated equally (Koehler, 2001).

Consider a defendant accused of rape, where the only available evidence against him is that his DNA matches the traces found on the victim. The relevant information about the accuracy of the DNA analysis can be framed either in terms of conditional probabilities (such as the sensitivity and specificity of the DNA test) or in natural frequencies (such as the number of people in every 100,000 who are not the source of the DNA trace but nevertheless show a match). Both frames are logically equivalent for calculating the Bayesian posterior probability of the defendant being the source of the DNA trace.

In a study, 27 judges, professors of law, and other professional jurists, and 127 advanced law students evaluated two realistic criminal court rape case files (Lindsey, Hertwig, & Gigerenzer, 2003). When presented with conditional probabilities, only 13% of professionals and 1% of students were able to correctly estimate the probability that the defendant was the source of the DNA traces on the victim. With natural frequencies, 68% of the professional and 44% of the students arrived at the correct estimates. Equally important, when the information was framed in conditional probabilities, 45% of the professionals (55% of the law students) found the defendant guilty, compared with 32% of the professionals (33% of the law students) when the information was framed in natural frequencies. Similarly large effects of natural frequencies on the ability to make Bayesian inferences have been reported in medical diagnosis (Hofrage, Lindsey, Hertwig, & Gigerenzer, 2000).

Our point here is not whether framing effects violate descriptive invariance (see Table 1) and thus generally count as errors (Tversky & Kahneman, 1981) or that they can be defended as normative (McKenzie, 2004; McKenzie & Nelson, 2003; Sher & McKenzie, 2006). Rather, the legal case illustrates a situation where consistency, here, the consistent use of the same frame
of evidence, operates in the service of functional goal: the equal protection of the law. Note that we are not arguing for consistency in the law as an end in itself; there exist a number of reasons why a certain degree of inconsistency in the law should be tolerated and in fact, may be even be functional (Engel, 2006).

Research Strategy: Logical Versus Ecological Rationality

Coherence and correspondence not only epitomize two distinct benchmarks for rational behavior but also two different research strategies. We refer to these as the logical-rationality and the ecological-rationality research strategies. The first is typically based on a simple three-step protocol. First, take some coherence rule from logic, probability theory, statistics, or decision theory (Table 1 provides examples such as modus tollens, class inclusion, or descriptive invariance) that is claimed to be “accepted” as a necessary condition for rational decision making. Second, wrap the rule into a thematic garb (e.g., the Wason task, the Linda problem, the Asian disease problem) and investigate the extent to which people’s behavior conforms to the rule. Third, if behavior and rule conform, interpret behavior as rational, otherwise as irrational (see the list of presumed fallacies in Table 1).

In contrast, the ecological-rationality research strategy does not equate rationality with following a coherence rule. Rather, it measures rationality in terms of success in the world, such as making competitive decisions and accurate predictions. In this research strategy, the content, context, and environment are considered irrelevant for the normative evaluation of behavior that solely rests on a coherence rule.

3. Identify the structural properties of the environment in which the strategies should operate.
4. Determine the “ecological rationality” of the strategies in the environment, that is, a set of environmental conditions that specify when a given strategy is likely to reach the goal better than a competing strategy.

Thus, the ecological-rationality program replaces the a priori imposition of norms in the logical-rationality program by an empirical investigation of conditions for the actual success of strategies.

Let us illustrate the ecological-rationality strategy by the common task of choosing between two alternatives, such as investment plans, based on a number of predictors. Assume the goal is to predict which alternative will perform better. As to Step 2, we consider two classes of strategies: lexicographic and linear rules. We choose lexicographic rules such as take-the-best and elimination-by-aspect, given the abundant experimental evidence that people’s inferences can be modeled by these rules (Gigerenzer, Hertwig, & Pachur, 2011).

We choose linear rules, such as linear regression, given that these are widely believed to embody rationality because they satisfy coherence rules such as transitivity. Lexicographic rules, in contrast, ignore part of the predictors (without computing an “optimal stopping rule”), ignore the dependencies between predictors, and can lead to systematic intransitivities.

33

Keeney and Raiffa (1993) warn that such a lexicographic heuristic is “more widely adopted in practice that it deserves to be,” is “naively simple,” and “will rarely pass a test of ‘reasonableness’” (p. 78).

We can only briefly describe some of the results here. Research on Step 3 has identified three structural properties: noncompensatoriness (Martignon & Hoffrage, 2002), dominance, and cumulative dominance (Baucells, Carrasco, & Hogarth, 2006). Research on Step 4 has shown that if one or more of these properties hold, a lexicographic rule will make exactly the same predictions as any linear rule. Moreover, if in addition, parameters have to be estimated from limited samples, then a lexicographic heuristic is likely to make more
accurate predictions than a linear rule (Gigerenzer & Brighton, 2009).

The study of the ecological rationality of lexicographic rules provides a new explanation as to why individuals have often been reported to rely on these rules. The explanation no longer is predicated on internal reasons such as cognitive or motivational limitations, but on the match between the rule and the environment. For instance, a study of 51 environments showed that one or more of the three conditions mentioned above holds in 90% to 97% of all comparisons (Şimşek, 2013). In these situations, it is utterly reasonable to rely on a simple heuristic, which is likely to be more accurate than the linear rule—and faster and more frugal, too. The danger of imposing coherence rules as normative, independent of an analysis of their ecological rationality, is that we will continue to misclassify such intelligent behaviors as incoherent and irrational.

The study of ecological rationality is related to the program of rational analysis (e.g., Anderson, 1990; Anderson & Schooler, 1991; Chater & Oaksford, 1999), which typically relies on Bayesian optimization rather than studying the decision process itself.

What does all this mean for research on human rationality? In our view, two implications follow:

No a Priori Imposition of Content-Free Norms

Researchers should not impose a coherence rule as an a priori norm of rationality on behavior and cognition. Instead, it is essential to analyze the specific structures of the environment, and how they interact with the cognitive machinery and the goals of the organism. That is the study of ecological rationality (Todd et al., 2012). As a consequence, claims of irrationality based on coherence rules (see Table 1) should be revisited; some of this work has already been completed (Gigerenzer, Fiedler & Olsson, 2012; Pleskac & Hertwig, 2014).

Analyze Ecological Rationality

Studying ecological rationality does not mean that biases in human cognition are unimportant or impossible to demonstrate, as suggested by Cohen (1981). On the contrary, biases can in fact enable better decision making, as the analysis of lexicographic strategies has demonstrated. Similarly, perception works well not despite but because of biases, such as the perceptual bias for rising tones (Neuhoff, 1998). Furthermore, a bias can be a necessary byproduct of an adaptive process (Arkes, 1991), as illustrated by the hindsight bias (Hertwig, Fanselow, & Hoffrage, 2003).

In this article, we have not dealt with the bias-variance dilemma, which provides the statistical foundation for understanding the beneficial nature of bias (Gigerenzer & Brighton, 2009). What we tried to show is that there is little evidence that incoherence hurts; that adaptive behavior can imply incoherence; that coherence may not be tractable; and that when coherence has instrumental value, it is not a goal by itself but serves different, functional goals. To paraphrase Ralph Waldo Emerson’s aperçu in the epigram, with coherence, a smart choice has surprisingly little to do.

References

Azar, O. H. (2011). Do consumers make too much effort to save on cheap items and too little to save on expensive items? Experimental results and im-


Berg, N., Biele, G., & Gigerenzer, G. (in press). Consistent Bayesians are no more accurate than non-Bayesians: Economists surveyed about PSA. *Review of Behavioral Economics*


Cosmides, L., & Tooby, J. (1992). Cognitive adap-


Received January 9, 2015
Revision received August 24, 2015
Accepted August 24, 2015