aps ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Current Directions in Psychological Science
2015, Vol. 24(5) 379–385

DOI: 10.1177/0963721415591878
cdps.sagepub.com
SAGE

The Two Settings of Kind and Wicked Learning Environments

Robin M. Hogarth[1], Tomás Lejarraga[2], and Emre Soyer[3]
[1]Department of Economics and Business, Universitat Pompeu Fabra; [2]Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany; and [3]Faculty of Business, Ozyegin University

Abstract
Inference involves two settings: In the first, information is acquired (learning); in the second, it is applied (predictions or choices). *Kind* learning environments involve close matches between the informational elements in the two settings and are a necessary condition for accurate inferences. *Wicked* learning environments involve mismatches. This conceptual framework facilitates identifying sources of inferential errors and can be used, among other things, to suggest how to target corrective procedures. For example, structuring learning environments to be kind improves probabilistic judgments. Potentially, it could also enable economic agents to exhibit maximizing behavior.

Keywords
kind and wicked learning environments, inference, judgment biases, decision making

Medical author Lewis Thomas recounted a story, dating from the early 20th century, of a physician in a New York hospital who acquired a reputation for accurately diagnosing typhoid fever in its early stages. The physician believed that the appearance of the tongue was highly diagnostic. Hence, his clinical technique included palpating patients' tongues before making his pessimistic forecasts. Unfortunately, he was invariably correct. But, as Thomas stated, "He was a more effective carrier, using only his hands, than Typhoid Mary" (Thomas, 1983, p. 22).

Hogarth (2001) used this example in introducing the concept of *wicked* learning environments. He described these as situations in which feedback in the form of outcomes of actions or observations is poor, misleading, or even missing. In contrast, in *kind* learning environments, feedback links outcomes directly to the appropriate actions or judgments and is both accurate and plentiful. In determining when people's intuitions are likely to be accurate, this framework emphasizes the importance of the conditions under which learning has taken place. Kind learning environments are a necessary condition for accurate intuitive judgments, whereas intuitions acquired in wicked environments are likely to be mistaken.

Our goal in this article is to elaborate on this distinction and to provide a more complete classification of types of learning environments. In doing so, we adopt the view that humans can be modeled as naive intuitive statisticians (Fiedler & Juslin, 2006) whose judgments mainly reflect the information available to them (see also the "what you see is all there is" metaphor of Kahneman, 2011). Thus, careful attention to the characteristics of learning environments is important for identifying sources of judgmental biases as well as suggesting corrective procedures.

The Two-Settings Framework

We conceptualize inference through the lens of probabilistic prediction. One observes a sample, calculates a statistic, and then estimates that statistic in the population or a different sample (as when, e.g., one estimates a mean). The theoretical justification relies on a simple assumption: Samples are randomly drawn from the same underlying population.

This formulation has been critical in judgment and decision-making research. It provides normative benchmarks (Tversky & Kahneman, 1974) and suggests descriptive models (Gigerenzer, 1991). However, we contend that

Corresponding Author:
Robin Hogarth, Department of Economics and Business, Universitat Pompeu Fabra, Ramon Trias Fargas, 25-27, 08005 Barcelona, Spain
E-mail: robin.hogarth@upf.edu

it is ill-suited for considering the psychological issues underlying decision making because, instead of *one* underlying population, people have to deal with *two* populations, or as we shall say, two *settings*.

In the first setting, people learn about a situation (e.g., how two variables covary). In the second, they take an action or make a prediction using the knowledge acquired in the first. One setting is characterized by learning and the other by choice or prediction. To illustrate, imagine you are a personnel manager who uses a test to select job candidates. This test has been accurate in the past (learning). Thus, for current decisions (predictions), the test can be expected to be accurate when the features of the two settings (past and present) match. For example, are the present candidates similar to those in the past? We emphasize that in this process, the manager must also have in mind (implicitly or explicitly) a *reference class* of relevant instances—that is, the specific group to which the inference refers. Different reference classes can imply different inferences.

Rather than assuming that both situations (e.g., past and present) are random samples from the same underlying population, we posit two distinct settings. We refer to the first as *L* (for *l*earning) and the second as *T* (for *t*arget) and ask how these match. On the left-hand side of Figure 1, we consider six ways in which the elements of information in *L* and *T* do or do not match, and these, in turn, allow us to define different task structures for kind and wicked learning environments. Clearly, kindness or wickedness can vary in degree. However, our intention here is limited to classification.

The right-hand side of Figure 1 illustrates the cases on the left using the job-selection scenario. Each scatter plot shows the data experienced by the manager in learning about the relation between test scores and job performance from past applicants (*L*). Subsequently, this information is used to predict the performance of new candidates (*T*).

Cases A and B represent kind learning environments. In A, there is a perfect match between the elements of *L* and *T*. In the example on the right, the correlation between *X* and *Y* is 1.0. Performance can be predicted perfectly from the test.

Case B reflects that the presence of random error means that matches are at best approximate. The relation between *X* and *Y* on the right is represented by an ellipse as opposed to a straight line. Technically, such mismatches imply an intersection between *L* and *T* (as does E, explained below; however, B differs from E in that in the former, the mismatch is entirely due to random factors).

Cases C through F represent wicked learning environments. In C, *L* is a subset of *T*. There are elements in *T* that cannot be inferred from *L*. Examples include the survivorship bias, in which data have been systematically restricted by events or actions (Einhorn & Hogarth, 1978). In the example on the right, performance data are not available for people scoring low on the test ($X < 10$) because they were not selected for the job.

In D, *T* is a subset of *L*. This can occur when the person is unaware that there has been a change in the composition of the reference class between learning and prediction. For example, imagine that the applicant pool changes because the local university has lowered its admission standards, such that there are no highly qualified candidates among the graduates applying for the job. However, the personnel manager does not realize this.

In E, the elements of *L* and *T* intersect because of systematic factors, and the ability to predict in *T* is limited. This case captures self-fulfilling prophecies or so-called treatment effects. In terms of the selection model, those chosen ($X > 10$) receive special "treatment" that systematically biases job performance positively (e.g., they have excellent mentors). The personnel manager is exposed to a biased learning sample.

Case E also captures the conditions of both C *and* D, where, in our example, an employer does not observe performance measures for candidates with low test scores ($X < 10$) and the learning sample is biased by the change in the applicant pool, of which the manager is unaware.

Finally, we note Case F, in which *T* and *L* have no elements in common. In this case, the variable used to predict performance is not related to it (e.g., physical appearance).

Features of Wickedness

A wicked learning environment can emerge as a result of actions taken by the person making the inferences (as in self-fulfilling prophecies, Case E) as well as the characteristics of the environment. For example, a Case C situation could arise if someone were asked to make predictions beyond the range of data observed in the past (Feiler, Tong, & Larrick, 2012). Here, the mismatch is not triggered by the individual's actions.

Although discrete in our classification scheme, kindness and wickedness can vary in degree as on a continuum. For instance, Case A is kinder than B, which is kinder than E or F. But what happens when mismatches are due to random factors? In B, for example, noise attenuates predictive ability. In fact, with much noise, predictive ability could be inherently lower in Case B than in some wicked environments, such as Case C. However, our framework clearly indicates that whereas the underlying cause of mismatch is random in the former, it is systematic in the latter.

We envisage learning as involving the sequential accumulation of information, such that the size and variability

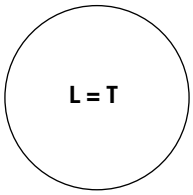
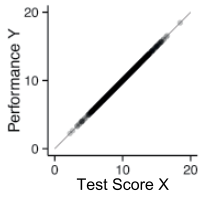
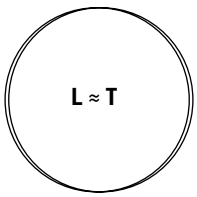
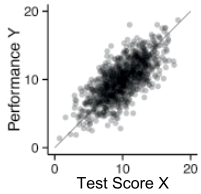
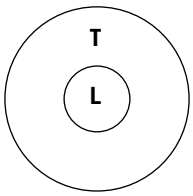
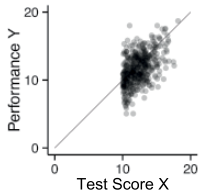
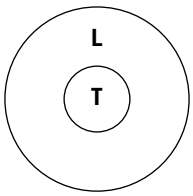
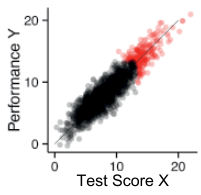
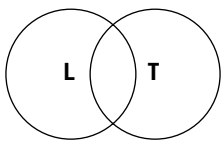
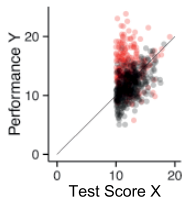
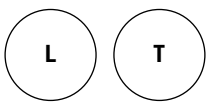
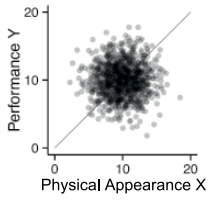
		Venn Diagrams Corresponding to Kind and Wicked Environments	Examples for Job Selection	
Kind Environments	A. The elements of *L* match those of *T*.			Judgments are based on test scores that predict performance perfectly.
	B. The elements of *L* and *T* are approximately the same.			Judgments are based on test scores that are imperfectly related to performance.
Wicked Environments	C. Elements of *L* have been "filtered" out. There is missing data.			Performance can only be observed for high test scores (X >10) because candidates with low sores are not selected.
	D. The elements of *T* are a subset of *L*.			The composition of the applicant pool has changed and candidates with high scores no longer apply (red dots).
	E. There is an intersection of elements of *L* and *T*.			Selected candidates (X >10) receive special treatment that inflates performance (red dots).
	F. *L* and *T* have nothing in common.			Judgments of performance are based on an unrelated variable, physical appearance.

Fig. 1. The two-settings framework. On the left, we show six ways in which the elements of information in the learning setting (*L*) and the target setting (*T*) do or do not match. On the right, we show an example scenario involving job selection.

of samples also play important roles—particularly when samples are small. Often, however, mismatches involve both systematic and random factors, and observing larger samples might not help.

Our framework deals only with the elements of information in *L* and *T*. It does not explain, for example, the reasons why individuals consider extraneous information (e.g., as in priming) or how information is aggregated in making inferences. These issues are important because many errors can be attributed to attention paid to extraneous information (Kahneman, 2011) and/or inappropriate aggregation rules (e.g., using additive aggregation

when it should be multiplicative; Larrick & Soll, 2008). However, by distinguishing information matches between *L* and *T*, we can better isolate the underlying sources of judgmental errors that are due to task features as opposed to idiosyncratic psychological processes.

Our framework reveals that some biases identified by specific labels in the literature can have multiple causes. Consider, for example, illusory correlation (Fiedler, 2000b). On the one hand, this can be induced by experiencing filtered observations. That is, the individual's experience in *L* is biased because part of a bivariate distribution is obscured from observation. On the other hand, the phenomenon investigated by Chapman and Chapman (1969) is about the role of prior beliefs on perceived correlations. (See also Denrell & Le Mens, 2011.)

Our perspective also speaks to the predictive accuracy of some heuristic decision processes that typically ignore information and involve simple decision rules (Gigerenzer & Gaissmaier, 2011). Successful heuristics exploit two key features of the environment: how information is aggregated and redundancy (Hogarth & Karelaia, 2007). As such, they operate in the intersection of *L* and *T*. For example, when people employ the recognition heuristic to select one of two alternatives (Goldstein & Gigerenzer, 2002), they base their judgments on information available in memory that happens to be correlated with what they are trying to predict.

Matching as a Default

People often use a default strategy that projects a match from *L* to *T* (Kahneman & Tversky, 1973). There could be several reasons for this. First, inferences often need only to suggest a direction as opposed to providing precise answers (Hogarth, 1981).

Second, assume that a person knows that elements are missing from *L*. What should be done? Much depends on what is known to be missing (Elwin, 2013). However, from a normative perspective, it is unclear how to correct for missing observations (Case C) and unrepresentative learning sets (Case D).

Third, default matching strategies are cognitively simple. Adjusting defaults requires meta-cognitive ability that people may not possess (Fiedler & Kutzner, in press).[1]

Relationships to Other Frameworks

Other scholars have used differences between two settings to explain bias. In their work on affective forecasting, Gilbert and Wilson (2007) contrasted people's images of future outcomes with what actually happens. For example, when buying a convertible, a person may imagine the joys of driving in beautiful weather but fail to consider other scenarios involving bad weather.

The importance of matches between two settings is acknowledged in the literature on transfer of learning (Barnett & Ceci, 2002). Interestingly, Thorndike's (1903) influential theory was framed in terms of "identical elements" and the match between these elements in the settings where learning is acquired and applied. However, his concern was with learning facts or skills (e.g., does learning to play tennis transfer to other racquet sports?).

Other social scientists show concern about the matches between two settings when exploiting data sets. For example, statisticians and machine-learning experts know that results obtained in samples do not necessarily generalize and have developed techniques for testing out-of-sample inferences.

We have not explicitly considered dynamic or nonstationary environments. At one level, such environments are wicked (a likely Case E). However, if the nature of the dynamic change can be inferred from the first setting, these environments can be kind. Consider, for example, learning seasonal cycles from experience.

Implications

Our framework has descriptive and prescriptive implications. In the context of examining inferential judgments in a particular task, it first draws our attention to whether this is kind or wicked. If kind, we have the necessary conditions for accurate inference. Therefore, any errors must be attributed to the person (e.g., inappropriate information aggregation). If wicked, we can identify how error results from task features, although these can also be affected by human actions. In short, our framework facilitates pinpointing the sources of errors (task structure and/or person). Table 1 lists some phenomena in the literature viewed from this perspective. For example, consider the "hot stove" effect, the fourth entry. Here, a person's experience of past outcomes (learning) determines what she selects currently (target), but then the outcome of this biases her subsequent learning.

There have been many attempts to correct judgmental biases (Soll, Milkman, & Payne, in press). Since kind environments are a necessary condition for accurate judgments, our framework suggests deliberately creating kind environments. Indeed, this reasoning motivated our work on *simulated experience*, in which we engineered kind environments by letting people experience sequential outcomes of probabilistic processes (Hogarth & Soyer, 2011) and investigated their ability to make appropriate probabilistic statements. Facing problems that are typically answered erroneously, participants' judgments in these kind environments were quite accurate. Moreover, the participants were confident in their responses.

Table 1. Four Illustrative Phenomena Viewed From the Two-Settings Framework

Phenomenon	Description	Reference	Examples	Further comments
Survivorship bias	The environment eliminates failures, so people tend to consider only the survivors of a process while ignoring the cases that did not survive.	Einhorn and Hogarth (1978)	1. Because failed entrepreneurs disappear, the probability of success in a new venture may be overestimated. 2. Mutual fund companies drop poorly performing mutual funds from their portfolios. This results in the overestimation of past returns based on the surviving ones (Elton, Gruber, & Blake, 1996).	In evaluating a process, *L* excludes past failures, thereby biasing extrapolation to *T*.
Censorship bias	Observations from a population are not observable beyond a specific "censorship" point. They are either ignored or treated as having values at the censorship point.	Feiler, Tong, and Larrick (2012)	Managers tend to observe when an employee falls short in a task, but they are unlikely to observe how much more employees are capable of doing on occasions in which they complete the work assigned to them (Feiler et al., 2012).	In evaluating performance, *L* excludes information needed to make an accurate assessment in *T*.
Selection bias	Decision makers focus on a specific subset of the observations to make inferences about the population.	Denrell (2005b); Koehler and Mercer (2009)	1. Journalists study successful businesses (excluding unsuccessful ones) to discover what makes a business successful (Denrell, 2005b). 2. Investors judge the future performance of a mutual fund by considering only data concerning the more successful related funds as opposed to all related funds (Koehler & Mercer, 2009).	In evaluating a process, cases are excluded from *L* that should be considered in *T*.
The "hot stove" effect	Decision makers avoid options that led to unfavorable outcomes in the past. Therefore, negative experiences tend to remain uncorrected.	Denrell and March (2001)	Managers may step back from implementing a new process because it led to immediate negative effects, without giving the new process a second chance (Denrell & March, 2001).	Early negative experiences make people stop search in *L*, such that it excludes elements that are relevant in *T*.

Note: *L* = learning setting; *T* = target setting.

We are not alone in suggesting simulation methodology. These methods have proven useful, for example, in financial decisions (Goldstein, Johnson, & Sharpe, 2008; Kaufmann, Weber, & Haisley, 2013) and understanding the implications of climate change (Sterman, 2011). Our framework can contribute to specifying when simulation methods are likely to be useful.

Although we highlight the advantages of making environments kind, we note that it may sometimes also pay to exploit wicked environments. In providing placebos, for example, the goal is that people should draw the wrong lesson from experience. Our framework can be used to conceptualize such interventions.

Recently, Erev and Roth (2014) examined deviations from economic rationality from the perspective of learning behavior. They argued that maximizing behavior is likely when the learning environment leads agents to "the best payoff for all agents on average, and most of the time" (p. 10818). Their contribution is important because, instead of postulating the use of conventional maximization models, Erev and Roth attributed successes and failures in maximizing behavior to what and how agents have learned and thus, implicitly, to whether they have been exposed to kind or wicked learning environments.

These ideas suggest that the concepts of kind and wicked learning environments can be useful in the design of economic incentive schemes. That is, instead of assuming that economic agents can calculate maximizing solutions, one should provide experiences that lead to appropriate responses—that is, in kind environments. Although the way to achieve this remains uncertain, posing the problem in these terms is a major step forward.

Recommended Reading

Denrell, J. (2005a). Why most people disapprove of me: Experience sampling in impression formation. *Psychological Review*, *112*, 951–978. An enlightening and comprehensive review that illustrates how biased judgment can result from the information people happen to sample.

Einhorn, H. J., & Hogarth, R. M. (1978). (See References). The origin of the ideas presented here—how environments can lead us to learn the wrong things.

Fiedler, K. (2000a). Beware of samples! A cognitive–ecological sampling approach to judgment biases. *Psychological Review*, *107*, 659–676. A clear statement of the need to understand how mind and task interact when judgments are formed.

Hogarth, R. M. (2001). (See References). Explains the development of the concepts of kind and wicked learning environments—see, in particular, Chapter 3.

Kahneman, D. (2011). (See References). An enchanting semi-autobiographical narrative that lays out many of the issues in the psychology of judgment and choice.

Acknowledgments

The authors are thankful for constructive comments from Tim Pleskac, Gaël Le Mens, Jordi Quoidbach, Klaus Fiedler, and anonymous referees as well as attendees at the Management breakfast seminar at Universitat Pompeu Fabra and members of the Center for Adaptive Rationality at the Max Planck Institute for Human Development.

Declaration of Conflicting Interests

The authors declared no conflicts of interest with respect to their authorship or the publication of this article.

Funding

Hogarth gratefully acknowledges the financial support of the Spanish Ministerio de Economía y Competitividad (Grant ECO2012-35545). Lejarraga acknowledges the financial support of the Swiss National Science Foundation (grant CRSII1_136227).

Note

1. For a case suggesting where this might be possible, see Koehler and Mercer (2009).

References

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin*, *128*, 612–637.

Chapman, L. J., & Chapman, J. R. (1969). Illusory correlations as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology*, *74*, 271–280.

Denrell, J. (2005b, April). Selection bias and the perils of benchmarking. *Harvard Business Review*. Retrieved from https://hbr.org/2005/04/selection-bias-and-the-perils-of-benchmarking/ar/1

Denrell, J., & Le Mens, G. (2011). Seeking positive experiences can produce illusory correlations. *Cognition*, *119*, 313–324.

Denrell, J., & March, J. G. (2001). Adaptation as information restriction: The hot stove effect. *Organization Science*, *12*, 523–538.

Einhorn, H. J., & Hogarth, R. M. (1978). Confidence in judgment: Persistence of the illusion of validity. *Psychological Review*, *85*, 395–416.

Elton, E. J., Gruber, M. J., & Blake, C. R. (1996). Survivorship bias and mutual fund performance. *The Review of Financial Studies*, *9*, 1097–1120.

Elwin, E. (2013). Living and learning: Reproducing beliefs in selective experience. *Journal of Behavioral Decision Making*, *26*, 327–337.

Erev, I., & Roth, A. E. (2014). Maximization, learning and economic behavior. *Proceedings of the National Academy of Sciences, USA*, *111*(Suppl. 3), 10818–10825.

Feiler, D. C., Tong, J. D., & Larrick, R. P. (2012). Biased judgment in censored environments. *Management Science*, *59*, 573–591.

Fiedler, K. (2000b). Illusory correlations: A simple associative algorithm provides a convergent account of seemingly divergent paradigms. *Review of General Psychology*, *4*, 25–58.

Fiedler, K., & Juslin, P. (2006). Taking the interface between mind and environment seriously. In K. Fiedler & P. Juslin (Eds.), *Information sampling and adaptive cognition* (pp. 3–32). New York, NY: Cambridge University Press.

Fiedler, K., & Kutzner, F. (in press). Information sampling and reasoning biases: Implications for research in judgment and decision making. In G. Keren & G. Wu (Eds.), *Wiley Blackwell handbook of judgment and decision making*.

Gigerenzer, G. (1991). From tools to theories: A heuristic of discovery in cognitive psychology. *Psychological Review*, *98*, 254–267.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, *62*, 451–482.

Gilbert, D. T., & Wilson, T. D. (2007). Prospection: Experiencing the future. *Science*, *317*, 1351–1354.

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

Goldstein, D. G., Johnson, E. J., & Sharpe, W. F. (2008). Choosing outcomes versus choosing products: Consumer-focused retirement investment advice. *Journal of Consumer Research*, *35*, 440–456.

Hogarth, R. M. (1981). Beyond discrete biases: Functional and dysfunctional aspects of judgmental heuristics. *Psychological Bulletin*, *90*, 197–217.

Hogarth, R. M. (2001). *Educating intuition*. Chicago, IL: The University of Chicago Press.

Hogarth, R. M., & Karelaia, N. (2007). Heuristic and linear models of judgment: Matching rules and environments. *Psychological Review*, *114*, 733–758.

Hogarth, R. M., & Soyer, E. (2011). Sequentially simulated outcomes: Kind experience vs. non-transparent description. *Journal of Experimental Psychology: General*, *140*, 434–463.

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Strauss and Giroux.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*, 237–251.

Kaufmann, C., Weber, M., & Haisley, E. (2013). The role of experience sampling and graphical displays on one's investment risk appetite. *Management Science*, *59*, 323–340.

Koehler, J. J., & Mercer, M. (2009). Selection neglect in mutual fund advertisements. *Management Science*, *55*, 1107–1121.

Larrick, R. P., & Soll, J. B. (2008). The MPG illusion. *Science*, *320*, 1593–1594.

Soll, J. B., Milkman, K. L., & Payne, J. W. (in press). A user's guide to debiasing. In G. Keren & G. Wu (Eds.), *Wiley Blackwell handbook of judgment and decision making*.

Sterman, J. D. (2011). Communicating climate change risks in a skeptical world. *Climatic Change*, *108*, 811–826.

Thomas, L. (1983). *The youngest science: Notes of a medicine watcher*. New York, NY: Viking.

Thorndike, E. L. (1903). *Educational psychology*. New York, NY: Lemcke and Buechner.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131.