
Stubborn Reliance on Human Nature in Employee Selection: Statistical Decision Aids Are Evolutionarily Novel

STEPHEN M. COLARELLI AND MATTHEW THOMPSON
Central Michigan University

The limited use of decision aids has been a conundrum in our field for years, and Highhouse's (2008) article is one of several attempts to explain and solve the problem. Lawshe and Bolda (1958) argued that managers could not understand the metric of validity coefficients, so they developed a more user-friendly metric, the expectancy chart. Others suggested that managers were not aware of the economic value of valid predictors, and they worked out economic utility equations (e.g., Schmidt, Hunter, McKenzie, & Muldrow, 1979). Dawes (1971) complained that decision makers did not understand the value of mechanical methods for combining predictor scores, and he encouraged better communication about their benefits. Yet, none of this had much effect on persuading managers to use decision aids (Hastie & Dawes, 2001; Latham & Whyte, 1994). Highhouse offers another interpretation: employers' belief systems. People still believe (erroneously) in near-perfect prediction estimates and that predictive ability can be improved with experience. He recommends that we do a better job of communi-

cating to managers what the decision-making literature has found about uncertainty and error.

These explanations and prescriptions have two elements in common. First, they assume that managers are "rational" actors: If presented with convincing data, they can be persuaded to change their habits. Second, there is a communication problem. Industrial-organizational psychologists are not effectively communicating the benefits of selection technologies to managers; if we did a better job communicating, managers would see the light. Of course, people can, to some degree, be persuaded by scientific evidence, and clear communication can be effective. However, the problem goes deeper than this. Some beliefs are readily changed by new information, reflection, and analysis; others are more stubborn. Beliefs about assessing people and predicting their behavior fall into the latter category. Therefore, a *factual* understanding of the value of selection decision aids does not guarantee that these technologies will be incorporated into people's belief systems or that they will be readily learned and used. To make headway into changing the effects of stubborn belief systems, we need to understand both *what* the beliefs are and *why* people have them.

People's beliefs about, and preferred methods for, predicting behavior are more than products of current culture; rather, they stem from evolved psychological mechanisms.

Correspondence concerning this article should be addressed to Stephen M. Colarelli. E-mail: colar1sm@cmich.edu

Address: Department of Psychology, Central Michigan University, Mt. Pleasant, MI 48859

Stephen M. Colarelli and Matthew Thompson, Department of Psychology, Central Michigan University.

Thanks to Geeta D'Souza and Kate Unterborn for comments on an earlier version of this article.

These are *adaptations* (Williams, 1966): heritable features of the brain that evolved in ancestral environments to solve problems related to survival and reproduction. It is important to keep in mind two points when invoking psychological mechanisms as (partial) explanations for behavior in the modern world. First, psychological mechanisms do not deny learning. However, they do produce preferences that make it easier to learn behaviors triggered by psychological mechanisms (e.g., sex) than those that go against the grain of a mechanism (e.g., dieting). Second, our modern environment is frequently “mismatched” with adaptations that evolved in the Pleistocene. For example, our preferences for salty, sweet, and fatty foods were adaptive in our ancestral environment when salt, carbohydrates, and calorie- and protein-dense foods were difficult to get. However, these preferences are mismatched to our current environment—where junk food with unnaturally high concentrations of salt, carbohydrates, and fat is easily obtainable, resulting in overconsumption and health problems, such as obesity.

How Psychological Mechanisms Influence the Use of Decision Aids

Psychological decision aids have been around for only about a 100 years; they are evolutionarily novel. Although basing decisions on test scores and probability statements are often appropriate for personnel decisions in organizations, humans are predisposed to assess others with the psychological mechanisms that have been used for millennia, which are frequently mismatched to modern organizations.

For most of our evolutionary history, we lived in small bands of about 150 people; we tended to know one another personally, their habits, personalities, and reputations (Megarry, 1995). Information about other people was gathered either through face-to-face interaction or by word of mouth (Moore, 1996). *People who made faulty judgments about others under these conditions would be less likely to survive and reproduce than those who made more accu-*

rate judgments. Therefore, humans developed (through natural selection) mental mechanisms that predisposed them to seek out and process information about people that is transmitted during face-to-face interaction (Moore, 1996; Sugiyama 2001). People still learn most of what they know about their social world through face-to-face interaction, beginning with parent–child interaction to the rich milieu of adult gossip (Dunbar, 1996). Indeed, much of our ability to perceive traits, intentions, and emotions of others is reasonably accurate, universal across cultures, and appears to be based on evolved mechanisms (Ekman, 1993).

Even those of us steeped in probability theory, psychometrics, and the scientific selection literature have difficulty overcoming the pull of these mechanisms. Most I–O faculty still use interviews and letters of recommendation when assessing job candidates (York & Cranny, 1989), and we suspect that most of us select our romantic partners through face-to-face interaction.¹

It is also important to consider the structure of the adaptive problems that our mental circuitry evolved to solve. Some of the most pressing problems in the Pleistocene probably involved assessing potential mates, allies, enemies, and partners for social exchange (Buss, 2008). It is less likely that making predictions about abstract, pan-utilitarian notions of job performance was an adaptive problem faced by our ancestors. Everyday prediction of behavior still remains driven by fundamental social motives—protecting oneself from harm, seeking romantic relationships, striving for status, and coalition building (Kenrick, Neuberg, & Cialdini, 2006).

Face-to-face methods are more attuned to individual, self-interest than actuarial

1. Every year one of us asks graduate students in his personnel psychology class—most of whom do not need convincing that statistical decision aids are a good thing—if they would forego the ritual of face-to-face courtship and choose a mate based on a test score, if a paper-and-pencil test were developed that could accurately assess a potential mate and predict the quality of a marriage. Inevitably, almost all students say they still want to interact with a potential mate before making a marriage decision.

methods—further suggesting that they are evolved adaptations. Because people are prone to behave in ways that are in their self-interest, they usually prefer selection methods, such as the interview, that allow them to gather information relevant to their own interests. An interview provides a wealth of information (e.g., sex, age, values, coalitional potential). A paper-and-pencil test provides only one piece of information: a test score. Because individual interests rarely mesh fully with organizational goals (Colarelli, 2003), traditional methods often win out.

Is the Situation as Bad as Highhouse Suggests?

Nevertheless, Highhouse has painted an overly dismal picture about the use of modern methods and the predictive accuracy of traditional methods. In some situations, people are *more* likely to use decision aids. For over 50 years, universities have included well-validated tests in their procedures for selecting undergraduates and graduate students. Modern decision aids have been a mainstay in the U.S. military's selection protocols since World War I, and the U.S. government uses validated tests in its selection of civil servants (Wigdor & Garner, 1982). However, context is critical. People are likely to use decision aids when mechanisms for face-to-face assessment are not triggered or when face-to-face methods convey little personal advantage. This would typically be the case in large organizations where people doing the hiring may not interact with applicants or where they are unlikely to work directly with people they hire (Colarelli, 1996).

In certain contexts, folk methods can assess traits and predict behavior reasonably well. Humans have been assessing one another for thousands of years, so it is unlikely that people were incompetent in such an important area of life until modern decision aids were invented. Methods that have withstood the test of time probably did so because they were useful in some ways (Colarelli, 2003). People's intuitive

ability to assess some personality traits, likeability, and social competence can be quite accurate (Funder, 1999). If one believes that a large portion of the variance in teaching ratings reflects likeability and social competence, a study by Ambady and Rosenthal (1993) provides strong evidence of the accuracy of intuitive judgments. They found that students' ratings of college professors based on a 10-second video clip were almost the same as ratings made by students who were enrolled in professors' classes for a full semester.

The difference between traditional and modern assessment methods may have to do with criteria and time frame. Traditional face-to-face methods are more attuned to assessing qualities and predicting criteria associated with reproduction and survival: mate value (fertility, kindness, and commitment), relative status (dominance and ambition), and coalitional behavior (likelihood of cooperation). Modern decision aids tend to be associated with predicting criteria related to performance in modern organizations. The research evidence is still equivocal about when people prefer traditional face-to-face or actuarial methods to assess all traits. Topor, Colarelli, and Han (2006) found that human resource practitioners gave higher hirability ratings to hypothetical applicants when intelligence was assessed by a valid paper-and-pencil test than when it was assessed by an interview or an assessment center. However, they gave higher hirability ratings when Conscientiousness and Agreeableness were assessed by interviews and assessment centers, respectively, than by paper-and-pencil tests. Lievens, Highhouse, and DeCorte (2005) found somewhat different results. Retail store supervisors gave higher hirability ratings to hypothetical applicants when intelligence and Extraversion were assessed by a paper-and-pencil test than an interview; however, they found no difference when Conscientiousness was assessed by either method. It is hard to know why the two studies produced discrepant results—possibly because of methodological and sampling differences. Nevertheless, more of this type of research

is needed to tease out how people respond to traits assessed by different methods and why they do so (cf. Tett & Christiansen, 2008).

Ecological Validity and Probability Estimates

Highhouse may also be overstating the case that people are averse to making probability judgments. Many animals perform intuitive probability calculations when foraging for food (Real, 1991); it is also likely that mental mechanisms for inductive probability judgments evolved in humans. However, when studying human inductive reasoning, it is important to consider the *form* that information relevant to solving adaptive problems regularly took in ancestral environments. Cosmides and Tooby (1996) argued that for most of our evolutionary history, people made probability estimates based on *frequencies of events*. As a result, when people are asked to make a single probability estimate, they do poorly. However, when they are asked to make the same estimate using frequencies as input and output, they do remarkably well. They examined this experimentally by varying the presentation of the Casscells, Schoenberger, and Graboys (1978) medical diagnosis problem (Cosmides & Tooby, 1996). This problem is famous in the decision-making literature because highly educated people, with training in statistics, neglected the base rate. In the original study, only 18% of the Harvard Medical School students and staff gave the correct answer. When Cosmides and Tooby reframed the problem in terms of frequencies (rather than asking for single event probability), 76–92% of the subjects gave the correct answer (depending on the particular version of how the problem was presented).

Implications for Personnel Selection

We are not arguing that we throw the baby out with the bath water. However, working with (rather than against) people's evolved nature will improve the acceptance and possibly the effectiveness of our interven-

tions. People tend to consider self-interest over abstract organizational interests. Given this human predilection and the inevitable conflicts of interest within organizations, we need to learn more about how the operation of selection systems interacts with those interests (e.g., Brown, 1979). It is also important to better understand preferences for different types of selection methods and how they interact with self-interest and organizational context. For example, given people's preferences and motives for face-to-face information, we need to be realistic about the unstructured interview. In circumstances where people making hiring decisions will be working with new hires, interview decisions are more likely to reflect personal than organizational interests. People may be more willing to use paper-and-pencil tests when they are used in conjunction with an interview or when those doing the hiring are unlikely to work with new hires. Some traditional methods work reasonably well (work sample tests and job tryouts). In the contexts where they work well, we should probably leave well enough alone. Finally, we should continue to conduct research on which formats allow probabilistic information to be easily understood.

References

- Ambady, N., & Rosenthal, R. (1993). Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality & Social Psychology, 64*, 431–441.
- Brown, S. H. (1979). Validity distortions associated with a test in use. *Journal of Applied Psychology, 64*, 460–462.
- Buss, D. M. (2008). *Evolutionary psychology* (3rd. ed.). Boston: Pearson.
- Casscells, W., Schoenberger, A., & Graboys, T. B. (1978). Interpretation by physicians of clinical laboratory results. *New England Journal of Medicine, 299*, 999–1001.
- Colarelli, S. M. (1996). Establishment and job context influences on the use of hiring practices. *Applied Psychology: An International Review, 45*, 153–176.
- Colarelli, S. M. (2003). *No best way: An evolutionary perspective on human resource management*. Greenwich, CT: Praeger.
- Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all?: Rethinking some conclusions of the literature on judgment under uncertainty. *Cognition, 58*, 1–73.

- Dawes, R. M. (1971). A case study of graduate admissions: Applications of three principles of human decision making. *American Psychologist*, *26*, 180–188.
- Dunbar, R. (1996). *Grooming, gossip, and the evolution of language*. Cambridge, MA: Harvard University Press.
- Ekman, P. (1993). Facial expression and emotion. *American Psychologist*, *48*, 384–392.
- Funder, D. C. (1999). *Personality Judgment: A realistic approach to person perception*. San Diego, CA: Academic Press.
- Hastie, R., & Dawes, R. M. (2001). *Rational choice in an uncertain world*. Thousand Oaks, CA: Sage.
- Highhouse, S. (2008). Stubborn reliance on intuition and subjectivity in employee selection. *Industrial and Organizational Psychology; Perspectives on Science Practice*, *1*, 333–342.
- Kenrick, D. T., Neuberg, S., & Cialdini, R. (2006). *Social psychology* (3rd ed.). Boston: Allyn & Bacon.
- Latham, G. P., & Whyte, G. (1994). The futility of utility analysis. *Personnel Psychology*, *47*, 31–46.
- Lawshe, C. H., & Bolda, R. A. (1958). Expectancy charts. I. Their use and empirical development. *Personnel Psychology*, *11*, 353–365.
- Lievens, F., Highhouse, S., & DeCorte, W. (2005). The importance of traits and abilities in supervisors' hirability decisions as a function of method of assessment. *Journal of Occupational and Organizational Psychology*, *78*, 453–470.
- Megarry, T. (1995). *Society in prehistory: The origins of human culture*. New York: New York University Press.
- Moore, R. F. (1996). Caring for identified versus statistical lives: An evolutionary view of medical distributive justice. *Ethology and Sociobiology*, *17*, 379–401.
- Real, L. A. (1991). Animal choice behavior and the evolution of cognitive architecture. *Science*, *253*, 980–986.
- Schmidt, F. L., Hunter, J. E., McKenzie, R. C., & Muldrow, T. W. (1979). Impact of valid selection procedures on work-force productivity. *Journal of Applied Psychology*, *64*, 609–626.
- Sugiyama, M. S. (2001). Food, foragers, and folklore: The role of narrative in human subsistence. *Evolution & Human Behavior*, *22*, 221–240.
- Tett, R., & Christiansen, N. D. (2008). Personality assessment in organizations. In G. Boyle, G. Matthews, & D. Saklofske (Eds.), *Personality measurement and assessment*. pp. 720–742. Thousand Oaks, CA: Sage Publications.
- Topor, D., Colarelli, S. M., & Han, K. (2006). The influence of personnel selection constructs and methods on evaluations of job applicants. *Journal of Business and Psychology*, *21*, 361–376.
- Wigdor, A. K. & Garner, W. R. (Eds.). (1982). *Ability testing: Uses consequences, and controversies (Parts 1 & 2)*. Washington, DC: National Academy Press.
- Williams, G. C. (1966). *Adaptation and natural selection*. Princeton, NJ: Princeton University Press.
- York, K. M., & Cranny, C. J. (1989). Job hunting and applicant recruiting: Expectations in the interview process and suggestions for improvement. *Industrial-Organizational Psychologist*, *26*, 19–23.