

# Scoring Situational Judgment Tests: Does the Middle Matter?

By Jeff Weekley, Ph.D., Kenexa

In the typical situational judgment test (SJT), an applicant is presented with a series of situations he or she could be expected to encounter on the job (these situations are usually determined via a job analysis). In most cases, each situation has four or five possible responses and the applicant is asked to indicate how he or she would handle that situation from the choices available (see the following example). While situational judgment tests have been around for many years, they have recently begun to garner much attention in the research literature (e.g., Chan & Schmitt, 1997; Clevenger, Pereira, Wiechmann, Schmitt, & Harvey, 2001; McDaniel, Morgeson, Finnegan, Campion, & Braverman, 2001; Olson-Buchanan, Drasgow, Moberg, Mead, Keenan, & Donovan, 1998; Smith & McDaniel, 1998; Weekley & Jones, 1997; 1999; Weekley & Ployhart, 2002). This increased popularity of SJTs is undoubtedly due to research demonstrating a number of positive features. First, research indicates that SJTs can validly predict performance and do so incrementally over the typical “benchmark” test of mental ability. Second, SJTs produce smaller subgroup differences (less potential for adverse impact) and more favorable applicant reactions than do tests of general mental ability.

Your top performing sales person is having personal problems. She is widely respected by the other employees. Recently, her performance has slipped badly and her attendance has become irregular. What are you most likely to do?

- Give her some time to work it out.
- Reduce her hours and have a newer person help with some of her customers.
- Counsel her about the importance of continued good performance.
- Ask the Employee to come up with a performance improvement plan.
- Warn the Employee that you will take disciplinary action if the problems continue.

Although SJTs are increasingly popular for practical reasons, there remain many unanswered questions. In the current research, we sought to compare alternative means of scoring SJTs. Typically, researchers have looked for the single best answer to the situation or for the best and worst alternatives. In either case, the response options not chosen are treated as unscored distracters. As yet unexamined is the issue of whether including all response options in the scoring of an SJT has measurement benefits. Rank-ordering of response options (e.g., from “best” to “worst”) is one means of including all options in the scoring. Doing so would seem to offer an advantage in measurement precision over the “pick best” or “pick best/pick worst” protocols that currently prevails. Our research was designed to test the question of whether there is anything to be gained (in terms of validity and reliability) by including all of the response options in the scoring of an SJT. Specifically, we compared three alternatives to scoring an SJT where the “correct” answers were determined by a panel of subject matter experts: identifying the best option, identifying the best and worst options, and ranking all options from best to worst.

Four SJTs were developed for use as potential predictors in a professional services firm, measuring sales-related judgment, administration-related judgment, leadership-related judgment and client service-related judgment. Each of the items had five possible response options. Data were gathered from 212 employees of the firm. For each of the situational questions, participants rank-ordered the five response options from 1=best to 5=worst. Because participants rank-ordered the five options for all items, it was possible to compare the three scoring formats using the same data.

Across the four SJTs, our results showed there to be little difference between the “pick the best” and the “pick best/pick worst” scoring options in terms of validity. Rank ordered scoring, however, showed a significant improvement in performance prediction in two of the four SJTs. Internal consistency reliabilities for all three approaches

were very low and none offered any reliability advantages over the others. In short, there does appear to be an advantage in terms of validity (but not reliability) to capturing respondent's evaluations to the full range of options presented. The greater validity of the rank-ordering method over the more typical "pick the best/pick the worst" format suggests that the middle of the range of answers do matter. This difference is not trivial from a practical perspective, as the rank order scoring method for the leadership SJT showed validity approximately 1.5 times greater than the most commonly used method (pick the best/worst), and 1.78 times greater than the pick the best scoring method. This would result in a sizeable improvement in overall utility.

In conclusion, situational judgment tests have many characteristics that make them attractive predictor measures. Research which compares and contrasts various scoring methods should continue and replication of the current results needed. A validity increase on the order of 1.5 to 1.8 times greater than existing scoring methods has real practical utility. Only by understanding more thoroughly the issues involved in optimally scoring SJTs can the field hope to advance SJTs from a measurement method to a means of measuring theoretically relevant constructs. ■

## About the Author

### Jeff Weekley, Ph.D.

Jeff Weekley, Ph.D., serves as a Kenexa executive consultant. Before joining Kenexa, Dr. Weekley held senior human resource management positions with Zale Corporation, Southland Corporation and Greyhound Lines. Weekley has designed, validated and implemented numerous large-scale employee selection systems for retail store managers and associates, customer service employees, health care providers, hospitality guest-contact employees, drivers and mechanics. He has also created many organizational development programs including succession planning processes, performance management systems, leadership training and internal customer satisfaction surveys.

Dr. Weekley has authored numerous articles for the Journal of Applied Psychology, Personnel Psychology, Academy of Management Journal, Human Performance and Journal of Management. He is a member of the American Psychological Association, the Society for Industrial and Organizational Psychology and the Academy of Management. Weekley holds a Doctorate degree in organizational behavior from the University of Texas at Dallas, a Master of Science degree in industrial and organizational psychology and Bachelor of Science degree in psychology from Texas A&M University.

[www.kenexa.com](http://www.kenexa.com)  
866.391.9557