Le, H., Oh, I.-S., Robbins, S. B., Ilies, R., Holland, E., & Westrick, P. (2011). Too much of a good thing: Curvilinear relationships between personality traits and job performance. *Journal of Applied Psychology, 96*(1), 113–133. doi:10.1037/a0021016

Lievens, F., & Motowidlo, S. J. (2016). Situational judgment tests: From measures of situational judgment to measures of general domain knowledge. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 9*, 3–22.

Motowidlo, S. J., Hooper, A. C., & Jackson, H. L. (2006). A theoretical basis for situational judgment tests. In J. A. Weekley & R. E. Ployhart (Eds.), *Situational judgment tests: Theory, measurement, application* (pp. 57–81). Mahwah, NJ: Erlbaum.

Unintended Consequences: Narrowing SJT Usage and Losing Credibility With Applicants

Amy E. Crook
Belmont University

Lievens and Motowidlo (2016) present a convincing case for why situational judgment tests (SJTs) should be developed specifically to measure general domain knowledge, but I have two concerns regarding the use of SJTs in psychological research and practical settings if the reconceptualization of SJTs offered by the authors is adopted to the exclusion of other current approaches. The first concerns abandoning SJTs as a helpful job analysis tool if we encourage intentional conflation of trait expression with job effectiveness in the development of SJT items. Second, diluting the situational component of SJTs may reduce their acceptance as selection tools by job applicants and practitioners in organizations.

Issues With SJTs Developed To Perfectly Correlate Effectiveness With Trait Expression

Although it has been widely accepted that SJTs are a method to measure various types of job knowledge, the characterization of SJTs as atheoretical if not designed specifically to test knowledge of the effectiveness of a specific trait (implicit trait policies) downplays their utility when intentionally used in this manner to analyze a job. Some SJTs have been developed by first focusing on wide coverage of job behaviors and then carefully examining conceptually relevant traits in the behaviors described in the SJT items. Indeed, as many SJTs are developed from critical incidents, SJTs have been used as a job analysis tool to actually better understand how trait expression relates to job

Amy E. Crook, Jack C. Massey College of Business, Belmont University.
Correspondence concerning this article should be addressed to Amy E. Crook, Jack C. Massey College of Business, Belmont University, 1900 Belmont Boulevard, Nashville, TN 37212. E-mail: amy.crook@belmont.edu

https://doi.org/10.1017/iop.2015.118 Published online by Cambridge University Press

effectiveness (Crook, 2014; Crook & Beier, 2015). In other words, SJTs can be used as a way to empirically approach job analysis and determine which implicit trait policies are important for performance, guided by theoretical underpinnings and not just a scattershot of dozens of traits.

If test developers decide a priori which traits should be effective and then simply design the response options to be high on the trait and effective or low on the trait and ineffective, as the authors advocate, one must discontinue using SJTs as a job analysis tool. How do we appropriately determine which traits should be effective and ineffective on the job in such cases? Presumably, we can use concurrent validity studies to measure relationships between personality and job performance, but previous studies have suggested that SJTs measure knowledge of trait effectiveness, which is more proximal to job performance than a person's standing on the trait itself (Crook et al., 2011). Such relationships were observed by using a single-response format where the behavioral responses to the situations were ones that had been observed directly on the job. Test creators did not generate response options designed to be high or low on traits; the behaviors, reported by subject matter experts (e.g., supervisors), occurred on the job and were noticed as particularly effective or ineffective. Ratings of trait expressions within the behaviors were done separately by independent raters blind to the investigation, thus giving more credibility to the relationships discovered between trait expressions and job effectiveness.

The authors suggest prosociality and conscientiousness as two theoretically relevant traits to effectiveness in many jobs. If we determine these underlying antecedents are relevant for job performance through a concurrent validation study, however, and then build an SJT that purposefully (one might even state "artificially") conflates the trait with effectiveness by writing response options such that the correlation of trait expression and effectiveness approaches 1, what exactly are we measuring? Within the authors' proposed framework, it is the respondent's understanding of how effective expressing that trait is for the job, but when trait expression and effectiveness covary completely, does it become a situationally based personality test that is presumably more resistant to faking than self-report measures of the direct trait? Although there certainly could be value in such measures, I do wonder how much incremental validity this trait-engineered SJT would have above and beyond self-report measures of the personality trait in a selection context where the applicant is already considering how to best represent oneself. Though previous studies of SJTs do not find this to be an issue, trait expressions were not intentionally bimodally associated with effectiveness. I imagine that the perfect correlation between the trait and effectiveness might make the relationship too obvious to respondents. Would the SJT, with the contextualized behavioral expression of the trait, be more related to

https://doi.org/10.1017/iop.2015.118 Published online by Cambridge University Press

cognitive ability than the straightforward self-report measure of the trait? Although studies in the past have shown that SJTs have lower adverse impact than other traditional cognitive selection tests (see Lievens, Peeters, & Schollaert, 2008, for an overview), this newly designed method for building trait-focused SJT response options may reveal new unintended consequences in this area.

Reducing Situational Context and Applicant Reactions

With regard to reducing the situational specificity of items and moving toward more generic SJTs, we do have substantial empirical support that general domain knowledge can accurately predict performance across jobs, even without situational content or situational stems as shown in the studies by Lievens. There remain some practical concerns around generating standardized judgment tests that are essentially situationally neutral and can be applied across jobs, however. Such tests may be more susceptible to coaching or faking, as acknowledged by the authors. Another concern not yet discussed is the decrease in face validity for such decontextualized tests and, hence, potentially lower favorability in applicant reactions. Prior studies have shown that SJTs are appreciated at face value in selection settings due to their clear relationship to the job at hand (Salgado, Viswesvaran, & Ones, 2001; Whetzel & McDaniel, 2009). Removing the job-specific situations may reduce favorable applicant reactions and the ease with which these measures will be embraced by managers for selection purposes. For example, I can imagine the confusion and potential distaste experienced by applicants for a banking job if they take a generic prosocial SJT that asks them what they would do in work situations as a lawyer, doctor, or human factors engineer (as per the example SJT described in the focal article; Motowidlo, Ghosh, Mendoza, Buchanen, & Lerma, 2015). Some evidence suggests that affective applicant reactions can influence test scores on SJTs (McCarthy et al., 2013), so these concerns are relevant to both practical affective and empirical issues.

Furthermore, the discussion of situational importance left unclear whether the authors favor removing context completely or using contexts that are similar across all jobs to make the SJT more targeted at general domain knowledge. If we generate more generic SJTs for general domain knowledge areas, it will be interesting to examine whether no context is actually preferable to "wrong-job" contexts, in terms of applicant reactions, administrator reactions, and criterion-related validity. Similarly, I think a name change is in order if we do indeed find that the situation is truly unimportant in specifically designed tests for general domain knowledge areas: prosocial judgment test, for example, would seem more apt.

https://doi.org/10.1017/iop.2015.118 Published online by Cambridge University Press

Concluding Thoughts

Lievens and Motowidlo have provided persuasive arguments for the maturation of the SJT field to move toward more standardized, generic formats purposefully designed to align certain trait expressions with effectiveness. Overall, I agree that it is time to solidify our efforts in SJTs by having some commonly accepted tests that tap into general domain knowledge areas relevant to many jobs (e.g., the prosocial judgment test for customer service roles), and creating such tests in a theoretically driven framework regarding trait expression and behavioral effectiveness can be a sound strategy. It would be disappointing, though, if the field generally took this call for consolidation to mean that we must reduce the diverse ways in which we use SJTs, as job analysis tools, for training content or assessments (Crook, 2012), or for other purposes. To be clear, the authors did not specifically advise eliminating these other uses for SJTs; these issues are just not directly addressed and could end up being unintended consequences of the systematic SJT development guidelines put forth.

In addition, I think we should consider the creation of more generic, off-the-shelf SJTs as a potential boon for developing *more valuable* situationally specific SJTs for jobs. For example, although it is generally effective to respond to requests for volunteering with expressions of conscientiousness, what are the specific situations in a particular job that call for unconscientious responses instead? (I am considering cases where prioritization of one's time is critical in the situation, and conscientiousness on a lesser priority may be sacrificed to ensure adequate thoroughness on a higher priority.) Under what circumstances does acting prosocially toward customers no longer benefit the organization but actually create problems and become ineffective? As such, I can see a future where practitioners use one or two general domain knowledge SJTs (potentially devoid of any situational information; context-independent SJTs) appropriate for a position in tandem with a job-specific knowledge SJT (context-dependent SJTs) that is incredibly specific to the key situations in the job that require "abnormal" responding. This wraparound approach could lead to better incremental validity and offset some of the previous concerns expressed.

References

Crook, A. E. (2012). *Training interpersonal skills for interviews: The value of behavioral models and the role of personality* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Accession No. 3534247)

Crook, A. E. (2014, April). Comparing single-response and multiple-response SJTs. In M. C. Reeder & J. Golubovich (Chairs), *Situational judgment test design and measurement informed by psychological theory*. Symposium conducted at the 30th Annual Conference of the Society for Industrial and Organizational Psychology, Philadelphia, PA.

https://doi.org/10.1017/iop.2015.118 Published online by Cambridge University Press

Crook, A. E., & Beier, M. E. (2015). *Using SJTs to understand relations between prosocial knowledge, coping with stress, and OCBs*. Unpublished manuscript.

Crook, A. E., Beier, M. E., Cox, C. B., Kell, H. J., Hanks, A. R., & Motowidlo, S. J. (2011). Measuring relationships between personality, knowledge, and performance using single-response situational judgment tests. *International Journal of Selection and Assessment, 19*, 363–373.

Lievens, F., & Motowidlo, S. J. (2016). Situational judgment tests: From measures of situational judgment to measures of general domain knowledge. *Industrial & Organizational Psychology: Perspectives and Practice, 9*, 3–22.

Lievens, F., Peeters, H., & Schollaert, E. (2008). Situational judgment tests: A review of recent research. *Personnel Review, 37*, 426–441.

McCarthy, J. M., Van Iddekinge, C. H., Lievens, F., Kung, M., Sinar, E. F., & Campion, M. A. (2013). Do candidate reactions relate to job performance or affect criterion-related validity? A multistudy investigation of relations among reactions, selection test scores, and job performance. *Journal of Applied Psychology, 98*, 701–719.

Motowidlo, S. J., Ghosh, K., Mendoza, A. M., Buchanen, A. E., & Lerma, M. N. (2015). *A context-independent situational judgment test to measure prosocial implicit trait policy*. Unpublished manuscript.

Salgado, J. F., Viswesvaran, C., & Ones, D. S. (2001). Predictors used for personnel selection: An overview of constructs, methods, and techniques. In N. R. Anderson, D. S. Ones, H. K. Sinangil, & C. Viswesvaran (Eds.), *Handbook of industrial, work, & organizational psychology* (pp. 165–199). London, UK: Sage.

Whetzel, D. L., & McDaniel, M. A. (2009). Situational judgment tests: An overview of current research. *Human Resource Management Review, 19*, 188–202.

Scoring SJTs for Traits and Situational Effectiveness

Robert J. Harvey
Virginia Tech

Lievens and Motowidlo (2016) addressed three of the most important unanswered questions regarding situational judgment tests (SJTs): (a) Should we view them as tests that can assess relatively generic constructs that predict performance across settings, (b) what constructs can they assess, and (c) how should they be scored? They suggested fundamentally changing the SJT development process by targeting the specific constructs we measure, using scoring systems that address both the targeted traits and their situational effectiveness, examining construct validity, and evaluating the criterion-related validity of SJT traits (Lievens & Motowidlo, pp. 11–12).

Robert J. Harvey, Department of Psychology, Virginia Tech.
Correspondence concerning this article should be addressed to Robert J. Harvey, Department of Psychology, Virginia Tech, 121 Williams Hall, Blacksburg, VA 24061. E-mail: harveyrj@vt.edu

https://doi.org/10.1017/iop.2015.118 Published online by Cambridge University Press