

- Jansen, A., Melchers, K. G., Lievens, F., Kleinmann, M., Brändli, M., Fradfel, L., & König, C. J. (2013). Situation assessment as an ignored factor in the behavioral consistency paradigm underlying the validity of personnel selection procedures. *Journal of Applied Psychology, 98*, 326–341.
- Kleinmann, M., Ingold, P. V., Lievens, F., Jansen, A., Melchers, K. G., & König, C. J. (2011). A different look at why selection procedures work: The role of candidates' ability to identify criteria. *Organizational Psychology Review, 1*, 128–146.
- Krumm, S., Lievens, F., Hüffmeier, J., Lipnevich, A. A., Bendels, H., & Hertel, G. (2015). How “situational” is judgment in situational judgment tests? *Journal of Applied Psychology, 100*, 399–416.
- 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.
- Melchers, K. G., Klehe, U., Richter, G. M., Kleinmann, M., König, C. J., & Lievens, F. (2009). “I know what you want to know”: The impact of interviewees' ability to identify criteria on interview performance and construct-related validity. *Human Performance, 22*, 355–374.
- Motowidlo, S. J., & Beier, M. E. (2010). Differentiating specific job knowledge from implicit trait policies in procedural knowledge measured by a situational judgment test. *Journal of Applied Psychology, 95*, 321–333.
- Motowidlo, S. J., Hooper, A. C., & Jackson, H. L. (2006a). Implicit policies about relations between personality traits and behavioral effectiveness in situational judgment items. *Journal of Applied Psychology, 91*, 749–761.
- Motowidlo, S. J., Hooper, A. C., & Jackson, H. L. (2006b). A theoretical basis for situational judgment tests. In J. A. Weekley & R. E. Ployhart (Eds.), *Situational judgment tests: Theory, measurement, and application* (pp. 57–82). Mahwah, NJ: Erlbaum.
- Rockstuhl, T., Ang, S., Ng, K. Y., Lievens, F., & Van Dyne, L. (2015). Putting judging situations into situational judgment tests: Evidence from intercultural multimedia SJTs. *Journal of Applied Psychology, 100*, 464–480.

The “Hot Mess” of Situational Judgment Test Construct Validity and Other Issues

Michael A. McDaniel, Sheila K. List, and Sven Kepes
Virginia Commonwealth University

The construct validity of situational judgment tests (SJTs) is a “hot mess.” The suggestions of Lievens and Motowidlo (2016) concerning a strategy to make the constructs assessed by an SJT more “clear and explicit” (p. 5)

Michael A. McDaniel, Sheila K. List, and Sven Kepes, Department of Management, Virginia Commonwealth University.

Correspondence concerning this article should be addressed to Michael A. McDaniel, Virginia Commonwealth University, 301 West Main Street, P.O. Box 844000, Richmond, VA 23284. E-mail: mamcdani@vcu.edu

are worthy of serious consideration. In this commentary, we highlight two challenges that will likely need to be addressed before one can develop SJTs with clear and explicit constructs. We also offer critiques of four positions presented by Lievens and Motowidlo that are not well supported by evidence.

Challenges to Establishing SJT Construct Clarity

The two main challenges likely to complicate the effort of developing SJTs with clearly defined constructs are that (a) SJT items are typically heterogeneous at the item level and (b) SJT scales will typically not show discriminant validity.

SJT Items Are Typically Heterogeneous at the Item Level

As illustrated by McDaniel and Whetzel (2005), SJT items are heterogeneous at the item level in that they have correlations with constructs that are not related to each other. For example, an item may have meaningfully large correlations with both general cognitive ability and the personality trait of agreeableness. This makes it difficult to obtain an interpretable factor structure that could be used to determine the constructs measured. Indeed, very few interpretable factor analyses have been reported (for an exception, see Legree, Heffner, Psotka, Martin, & Medsker, 2003). In brief, evidence supporting the construct validity of SJTs is unlikely to come through exploratory or confirmatory factor analyses. Instead, alternative strategies may need to be used to establish the construct validity of an SJT.

SJT Scales Will Typically Not Show Discriminant Validity

Consistent with the typical finding of uninterpretable factors, it is improbable that SJT scales will be able to show discriminant validity. As an example, if one designs SJT scales to measure the Big Five personality traits, the resulting five SJT scales will tend to have much larger magnitude correlations with each other than would be desirable. Furthermore, these correlations will tend to be larger in magnitude than are those found in scales using personality items.

Position Critiques

In addition to the two challenges noted above, we also offer four critiques of Lievens and Motowidlo's positions. We suggest that (a) the inclusion of situational scenarios may help to reduce ambiguity in response options, (b) an alternate view of knowledge acquisition could account for overlap between job-related knowledge and the general knowledge domain, (c) compound traits may lead to unnecessary construct proliferation, and (d) there is no

documented empirical evidence that single-item SJTs are more time and cost effective to develop.

Situational Scenarios May Help To Reduce Ambiguity in Response Options

In regard to the authors' assertion that recent evidence from Krumm et al. (2015) showed that situational scenarios are not necessary for high performance on SJTs, we argue that situational scenarios can help reduce ambiguities in item responses. McDaniel, Psotka, Legree, Yost, and Weekley (2011) have argued that SJT items vary in ambiguity such that the respondent may need to make specific assumptions in order to respond. Such items are associated with near zero validity. Consider this single-response SJT item: “You complete the work assigned to you by two different supervisors, both of whom consider their work to have priority, in the order it was assigned.” Depending on unstated aspects of the situation, doing work in the order in which it was assigned could be either an effective or an ineffective behavior. An SJT that incorporates a scenario can provide needed context to reduce the ambiguity of the response and improve the validity of the item. Thus, Lievens and Motowidlo's suggestion that situational scenarios are not necessary is unlikely to generalize across many SJTs and situations. More research is needed to determine the conditions under which situational scenarios are not required or necessary.

An Alternative View of Knowledge Acquisition

The authors of the focal article also claim that the findings from Krumm et al. (2015) support the reconceptualization of SJTs as measuring general domain knowledge. However, there can be considerable overlap between job specific knowledge and general domain knowledge. For instance, the claim that “job-specific knowledge can be learned only through exposure to that job or jobs like it” (Lievens & Motowidlo, p. 8) is not necessarily correct. Job specific knowledge can be obtained through formal education and training. For example, a doctoral student can learn about item response theory (a job specific knowledge for a test developer) in graduate school, retain this knowledge, and apply it when employed as test developer. Furthermore, the authors' assertion that “general domain knowledge is . . . not acquired from specific job experiences” (Lievens & Motowidlo, p. 8) is also likely to be inaccurate. Consider an adolescent, employed in a fast food restaurant, who arrives to work late. Upon arrival, the supervisor counsels the adolescent regarding the inappropriateness of being late. Due to this situation, the adolescent has obtained general knowledge about the value of being on time for work. Therefore, the authors' claim may not be correct. Any (re-) conceptualization of SJTs should account for the strong likelihood of considerable overlap in the acquisition of job specific and general domain knowledge.

Compound Traits Such as Prosocial Action May Contribute to Construct Proliferation

The authors suggest that their approach can be used to measure compound traits. Although SJTs can be developed to measure compound traits, we caution against this. Compound traits likely contribute to the construct proliferation that plagues the industrial–organizational psychology and management literature like an ever-expanding clump of fungus devouring our discipline (Le, Schmidt, Harter, & Lauver, 2010; Schwab, 1980). In addition, designing an SJT to measure a compound trait ultimately runs contrary to the authors’ stated goal, namely, to develop SJTs with “clear and explicit constructs” (Lievens & Motowidlo, p. 5), because compound traits are inherently multidimensional. Therefore, we encourage the reevaluation of the suggestion that SJTs should be developed to assess compound traits.

Undocumented Claims Concerning Time and Cost Efficiency of Single-Response SJT Items

Although we have nothing against single-response SJTs, the authors of the focal article assert, without support, that single-response “item development . . . is further simplified and made more efficient” (Lievens & Motowidlo, p. 17). We disagree. For example, with a Likert rating format (e.g., “rate each of the responses using the 1–7 scale of effectiveness”), each response associated with a scenario is a scorable item. Thus, for a scenario with five response options, one can obtain five scorable items with one scenario. With a single-response item, one needs one scenario for each scorable item. We suggest that our difference of opinion with the focal article authors would best be addressed empirically.

Conclusion

Taken together, we concur with Lievens and Motowidlo that the construct validity of SJTs could benefit greatly from additional research attention. In fact, we believe that it is currently a “hot mess” without much theoretical or empirical guidance. We also agree with several of the authors’ suggestions. However, we see some challenges in the strategies offered by Lievens and Motowidlo and disagree with several of their assertions.

References

- Krumm, S., Lievens, F., Hüffmeier, J., Lipnevich, A. A., Bendels, H., & Hertel, G. (2015). How “situational” is judgment in situational judgment tests? *Journal of Applied Psychology*, *100*, 399–416. doi:10.1037/a0037674
- Le, H., Schmidt, F. L., Harter, J. K., & Lauver, K. J. (2010). The problem of empirical redundancy of constructs in organizational research: An empirical investigation. *Organizational Behavior and Human Decision Processes*, *112*, 112–125. doi:10.1016/j.obhdp.2010.02.003

- Legree, P. J., Heffner, T. S., Psotka, J., Martin, D. E., & Medsker, G. J. (2003). Traffic crash involvement: Experiential driving knowledge and stressful contextual antecedents. *Journal of Applied Psychology*, 88, 15–26. doi:[10.1037/0021-9010.88.1.15](https://doi.org/10.1037/0021-9010.88.1.15)
- 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.
- McDaniel, M. A., Psotka, J., Legree, P. J., Yost, A. P., & Weekley, J. A. (2011). Toward an understanding of situational judgment item validity and group differences. *Journal of Applied Psychology*, 96, 327–336. doi:[10.1037/a0021983](https://doi.org/10.1037/a0021983)
- McDaniel, M. A., & Whetzel, D. L. (2005). Situational judgment test research: Informing the debate on practical intelligence theory. *Intelligence*, 33, 515–525. doi:[10.1016/j.intell.2005.02.001](https://doi.org/10.1016/j.intell.2005.02.001)
- Schwab, D. P. (1980). Construct validity in organizational behavior. In B. M. Staw & L. L. Cummings (Eds.), *Research in organizational behavior* (Vol. 2, pp. 3–43). Greenwich, CT: JAI Press.

It's Time To Examine the Nomological Net of Job Knowledge

W. Jackeline Torres and Margaret E. Beier
Rice University

Lievens and Motowidlo (2016) argue compellingly that situational judgment tests (SJTs) measure job-relevant general domain knowledge, conceptualized as implicit trait policies (ITPs). ITPs are defined as a person's knowledge about the utility of expressing certain traits. They develop through the feedback a person receives when acting in accordance with their trait profiles in different environments (work, life, leisure). Positive feedback reinforces the knowledge that behavior in accordance with one's own traits is appropriate, and negative feedback reinforces the knowledge that an approach that differs from one's trait tendencies may be more effective. As such, ITPs represent a person's knowledge about the effectiveness of behaviors across a variety of contexts.

Job knowledge has been recognized as an important determinant of job performance throughout the history of industrial and organizational (I-O) psychology, and because it is more proximal to the performance context, it generally accounts for more variance in performance than cognitive ability (Schmidt & Hunter, 1998). Indeed, more than 30 years ago, Hunter

W. Jackeline Torres and Margaret E. Beier, Department of Psychology, Rice University.

Correspondence concerning this article should be addressed to W. Jackeline Torres, Department of Psychology, MS-25, Rice University, 6100 Main Street, Houston, TX 77005. E-mail: jackietorres@rice.edu