Both General Domain Knowledge and Situation Assessment Are Needed To Better Understand How SJTs Work

Jinyan Fan and Melissa Stuhlman
Auburn University

Lijun Chen
Zhejiang University

Qingxiong Weng
University of Science and Technology of China

Although Lievens and Motowidlo (2016) made a strong case for reconceptualizing situational judgment tests (SJTs) as measures of general domain knowledge, we disagree with their view that the judgment or assessment of the situation itself is not important. We contend that situation assessment is an integral yet ignored factor in SJTs and that both general domain knowledge and situation assessment are needed to better understand how SJTs work.

Lievens and Motowidlo's unenthusiastic view of situation assessment in SJTs was based on (a) recent findings that test takers can successfully solve many SJT items even with the situation stems being removed (e.g., Krumm et al., 2015) and (b) the observation that situation assessment is not measured in traditional SJTs (Rockstuhl, Ang, Ng, Lievens, & Van Dyne, 2015). We believe that both arguments can be refuted. Regarding the first argument, a careful reading of Krumm et al.'s (2015) article suggests that their conclusions might have been overstated. For instance, although participants in the "decapitated" condition were not given the situation stems, they were nevertheless informed of the specific domain for each item (e.g., "teamwork"), and response options still contained rich contextual information, because they were developed based on specific situations. As such, the "decapitated" SJT items in Krumm et al.'s studies were probably less context-independent than assumed. Further, when comparing item-level scores between the traditional and "decapitated" SJT conditions, Krumm et al. used the Bonferroni correction to control for inflated family-wise alpha. Given

Jinyan Fan and Melissa Stuhlman, Department of Psychology, Auburn University; Lijun Chen, College of Public Affairs, Zhejiang University, Hangzhou, Zhejiang, People's Republic of China; Qingxiong Weng, School of Management, University of Science and Technology of China, Heifei, Anhui, People's Republic of China.

Correspondence concerning this article should be addressed to Jinyan Fan, Department of Psychology, Auburn University, 225 Thach Hall, Auburn, AL 36849. E-mail: jinyan.fan@auburn.edu

https://doi.org/10.1017/iop.2015.114 Published online by Cambridge University Press

the large number of *t* tests performed (35 and 30 in Studies 1 and 2, respectively), the Bonferroni correction was too conservative, unfairly favoring the null hypothesis. Were a less conservative correction method used, a much higher percentage of items would have shown significant group differences, which would have weakened their conclusions. In short, we are of the opinion that Krumm et al.'s studies did not convincingly prove that situation is not important in SJTs.

Lievens and Motowidlo's reasoning that situation assessment is not important in SJTs because it is not measured in traditional SJTs does not hold water, either. Ironically, Lievens and Motowidlo discussed an interesting study by Rockstuhl et al. (2015), who measured participants' assessment of the situation in a video-based intercultural SJT and found that the accuracy of situation assessment significantly predicted task performance and organizational citizenship behaviors, even after controlling for response effectiveness ratings. Rockstuhl et al.'s findings thus raise a provocative possibility that situation assessment is an important yet ignored factor in SJT research. In the remainder of this commentary, we explicate this idea by first briefly reviewing research on situation assessment in other selection procedures and then discussing how situation assessment and general domain knowledge can work together to further our understanding of SJTs.

Kleinmann et al. (2011) wrote that test takers in a selection procedure face two tasks: (a) figuring out what is being assessed and then (b) using this information to guide their behaviors to achieve good test performance. These two tasks roughly correspond to a cognitive component (understanding social situations) and a behavioral component (acting on that understanding) in social effectiveness (Jansen et al., 2013).

Different test takers may perceive the same selection situation differently, with varying degrees of accuracy, which gives rise to a construct referred to as the "ability to identify criteria" (ATIC; Kleinmann et al., 2011). ATIC is typically operationalized as the degree of similarity between test takers' perceptions of what is being assessed and the actual performance criteria in a selection procedure consensually determined by subject matter experts (SMEs; Kleinmann et al., 2011). Take Melchers et al.'s (2009) interview study as an example. After the completion of the interview component, participants were given a questionnaire containing all of the previously asked interview questions and were instructed to write down what they assumed each question was trying to assess. These assumptions were coded by trained assistants to determine whether each corresponded to one of the three consensually determined target dimensions and to what degree on a scale ranging from 1 = *fits somewhat* to 4 = *fits completely*. The assumption with the highest fit for the targeted dimension was used

https://doi.org/10.1017/iop.2015.114 Published online by Cambridge University Press

to compute the overall ATIC score by averaging these ratings across all interview questions.

During the last decade, research on ATIC has been conducted within several selection procedures such as structured interviews, assessment centers, and integrity tests. Two robust findings about ATIC directly relevant to the current discussion are as follows (for comprehensive reviews, see Jansen et al., 2013; Kleinmann et al., 2011). First, ATIC scores positively relate to selection test scores, meaning that a high ATIC benefits performance in the selection procedure. Second, ATIC scores demonstrate incremental validity over selection test scores in predicting job performance. This finding suggests that ATIC captures important predictive information independent of test performance.

As pointed out by Lievens and Motowidlo, situation assessment (or ATIC) has never been examined in the context of SJTs, with the lone exception of Rockstuhl et al.'s (2015) study. SJT scholars instead have focused on a form of general domain knowledge, called "implicit trait policy" (ITP; Motowidlo, Hooper, & Jackson, 2006a, 2006b). ITP refers to the implicit beliefs about the utility of expressing certain traits in test and/or job situations. ITP in SJTs is operationalized as the strength of association between response options' trait expression levels and effectiveness ratings (Motowidlo et al., 2006a, 2006b). Motowidlo and colleagues propose that individuals may have a variety of ITPs—for instance, ITP for agreeableness, ITP for extroversion, ITP for prosociality, and so on. Two major findings have emerged from empirical research on ITP (e.g., Motowidlo & Beier, 2010; Motowidlo et al., 2006a). First, ITP scores predict trait-related behaviors and job performance. Second, there are positive correlations between traits and ITPs for the corresponding traits; for instance, trait agreeableness scores positively predict ITP scores for agreeableness.

Although existent research on SJTs has established the existence and usefulness of ITPs, the inclusion of ATIC in SJTs may further our understanding of how SJTs work. We submit that ATIC and ITP fit nicely into the two tasks test takers face in a selection procedure mentioned earlier, with ATIC tapping into the task of figuring out what is being assessed and ITP taping into the task of using this information to guide their responses in SJTs.

We posit that ATIC and ITP should complement each other. For instance, a high ATIC should help test takers figure out what specific trait is being assessed in an SJT item, which will lead to the activation of the most appropriate (or *correct*) ITP from a large number of potential ITPs to aid with effectiveness judgment in the SJT item. For example, imagine that two candidates have the same latent level of general domain knowledge, for instance, ITP for agreeableness, but candidate A has a high ATIC and candidate B

https://doi.org/10.1017/iop.2015.114 Published online by Cambridge University Press

has a low ATIC. When an SJT item assesses agreeableness, candidate A correctly identifies this criteria, activates his/her ITP for agreeableness, and then engages in effectiveness judgment accordingly, resulting in a high observed ITP score. In contrast, candidate B mistakenly identifies conscientiousness as the criteria, and he/she then judges the effectiveness of response options according to conscientiousness expression rather than agreeableness expression, resulting in a low observed ITP score. This noise in observed ITP scores introduced by a low ATIC should attenuate the criterion-related validity of ITP scores. Another way to conceptualize the above moderation effect is to recognize that although ATIC scores in general should be positively related to job performance (Kleinmann et al., 2011), not knowing how to act on correctly identified criteria in SJT items, that is, having a low ITP, should weaken the ATIC–job performance link.

Jansen, Lievens, and Kleinmann (2011) provided preliminary evidence that may support the above moderation argument. These authors found in the context of an assessment center that participants who were highly conscientious or agreeable (and therefore most likely had high ITPs for each of these traits) were better able to express these traits in behaviors but only when they correctly perceived that these traits were being demanded by the situation (i.e., having a high ATIC score). Future research is needed to expressly test the ATIC $\times$ ITP interaction effect in SJTs. Further, it would be interesting to examine whether removing the situation stems in SJTs would lower the criterion-related validity of SJT scores and whether such an effect can be attributed to lowered ATIC through the decapitation of situation stems. If the above hypotheses are confirmed by empirical data, one practical implication would be that organizations interested in using SJTs to train and develop employees' general domain knowledge should also simultaneously include a training component that is focused on improving employees' ability to correctly interpret various situational cues at the workplace.

In closing, we conclude that it is premature to dismiss the situation assessment or judgment component of SJTs. To the contrary, we advocate that explicit incorporation of situation assessment into SJT research shall not only help us gain a fuller understanding of the mechanisms of SJTs but also facilitate the integration of two emerging bodies of research on ATIC and ITP, which have been treated separately for the most part thus far.

References

Jansen, A., Lievens, F., & Kleinmann, M. (2011). Do individual differences in perceiving situational demands moderate the relationship between personality and assessment center dimension ratings? *Human Performance, 24*, 231–250.

https://doi.org/10.1017/iop.2015.114 Published online by Cambridge University Press

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

https://doi.org/10.1017/iop.2015.114 Published online by Cambridge University Press