

Why Situational Judgment Is a Missing Component in the Theory of SJTs

Klaus G. Melchers
Universität Ulm

Martin Kleinmann
Universität Zürich

In their focal article, Lievens and Motowidlo (2016) consider procedural knowledge about effective actions in work situations as the key component of their theory of situational judgment tests (SJTs). In our commentary we want to suggest that situational judgment should nevertheless not be neglected in such a theory.

Specifically, Lievens and Motowidlo stress the importance of individuals' knowledge—be it of effective behavior in particular jobs or in the form of general knowledge about the pros and cons of expressing traits in situations like the ones described in an SJT—but they mention the role of individuals' perceptions of these situations only very briefly. This seems surprising, given that they describe research by Rockstuhl, Ang, Ng, Lievens, and Van Dyne (2015) that looked at the issue of situational judgment that is usually not explicitly measured in SJTs. In that research, substantial correlations between situational judgment (measured as test takers' "thoughts, feelings, and ideas of the people in the situation," Lievens & Motowidlo, p. 11) and judgments of response effectiveness were found. Based on these results and on related research on selection interviews and assessment centers (ACs), we want to argue that situational judgment plays an important role for test takers' performance in SJTs and probably also for the criterion-related validity of these SJTs.

Why SJTs Capture Situational Judgment Even Though It Is Not Explicitly Measured

We agree that relevant knowledge is important when test takers in a usual SJT have to make judgments about the effectiveness of different response options—or about their potential behavioral intentions. However, we doubt that knowledge about the general effectiveness of certain behaviors or of ex-

Klaus G. Melchers, Institut für Psychologie und Pädagogik, Universität Ulm, Ulm, Germany; Martin Kleinmann, Psychologisches Institut, Universität Zürich, Zürich, Switzerland.

Correspondence concerning this article should be addressed to Klaus G. Melchers, Institut für Psychologie und Pädagogik, Universität Ulm, Abteilung Arbeits- & Organisationspsychologie, Albert-Einstein-Allee 41, D-89069 Ulm, Germany. E-mail: klaus.melchers@uni-ulm.de

pressing certain traits is sufficient to answer an SJT effectively because different individuals might interpret the situations that are described in the different items very differently, even when they have similar work experience. This is where test takers' situational judgment—or more generally their interpretation of the situations—becomes relevant.

For test takers, the situations that are described in SJT items are often similar to the situations that interviewees face in questions in a selection interview or that participants face in exercises in an AC or in other kinds of work simulations. Especially, when these questions or exercises deal with interpersonal situations, then there is often more than an obvious one and only way to handle them—or said differently, many of these situations are not “strong situations” that are interpreted in the same uniform manner and that allow only a single, uniform behavioral response (Mischel, 1973), but they are often ambiguous so that different individuals might interpret them differently. Furthermore, based on their respective perceptions of a situation, different individuals might choose very different responses to deal with this situation in addition to different responses that are related to differences in individuals' procedural knowledge.

In line with this, measures of situational judgment (or of situation assessment or of individuals' ability to identify evaluative criteria, ATIC) have been found to be related to performance in selection interviews (e.g., Melchers et al., 2009), ACs (e.g., Jansen et al., 2013), and other selection simulations (e.g., Oostrom, Melchers, Ingold, & Kleinmann, in press). In the present context, situational interviews are particularly relevant, because they are similar to SJTs in many regards: They describe hypothetical work-related situations and ask interviewees what they would do in these situations (Maurer, Sue-Chan, & Latham, 1999). Furthermore, similar to SJTs, interviews and ACs have been challenged by evidence concerning their construct-related validity. This research raised concerns about whether it is indeed possible to measure the targeted performance constructs that they were designed to measure (e.g., Lance, 2008; Macan, 2009) even though this has been a less severe issue in the interview domain because interviews are often not developed in a construct-oriented way. Thus, similar to the available evidence that situation assessment correlates with interview or AC performance (cf. Kleinmann et al., 2011), we would consider the findings by Rockstuhl et al. (2015) as support for the suggestion that SJT scores that are based on effectiveness judgments nevertheless capture systematic variance that is related to situational judgment.

Why Situational Judgment Is Relevant for the Criterion-Related Validity of SJTs

The ability to correctly interpret a social situation is relevant not only when applicants have to take an SJT, an interview, or an AC but also later on the job

when they face ambiguous social situations. In line with this, Rockstuhl et al. (2015) found that situational judgment was a valid predictor of task performance. Similarly, previous research also found that situation assessment in interviews and ACs was a valid predictor of task performance and that the criterion-related validity of these selection procedures was higher when the variance related to situation assessment was not partialled out (e.g., Ingold, Kleinmann, König, Melchers, & Van Iddekinge, 2015; Jansen et al., 2013; Oostrom et al., in press). Although this research used a different approach to measure situation assessment by focusing on candidates' ability to correctly interpret performance criteria, it nevertheless stressed the importance of situation assessment even in samples that were relatively homogeneous in their work and educational background.

In line with the argument that situational judgment plays an important role, we echo Lievens and Motowidlo's suggestion that it might be valuable to collect situational judgments in addition to judgments concerning the effectiveness of different response options. However, we are skeptical about whether SJTs should always "be designed to determine whether people know that actions that express high levels of the relevant compound trait in a particular job are effective and that actions that express low and polar opposite levels of the compound trait are ineffective" (Lievens & Motowidlo, p. 16).

The reason for our skepticism is that this suggestion seems to be based on the assumption that actions related to a given trait are always effective. However, in many jobs that involve a broad range of social interactions—like, for example, in leadership positions—different actions that are related to different traits might be necessary to be successful across different situations: In some circumstances, it might be more important to be sympathetic or to show cooperative behavior, whereas other situations might require more assertive behavior. Thus, good situational judgment is necessary to correctly interpret these situations and to decide about the appropriate actions that need to be taken. In line with this, evidence from the AC domain suggests that good situational judgment is reflected in the ability to also discern targeted performance requirements across exercises that put very different demands on assesseees (Speer, Christiansen, Melchers, König, & Kleinmann, 2014).

Why Does Situational Judgment Matter Even Without a Description of the Situation?

Lievens and Motowidlo also state that the results of a series of studies by Krumm et al. (2015) question the assumption that SJTs tap into situational judgment. In these studies, test takers had to choose the most effective response option on the basis of these response options alone because the usual item stems were deleted. Results from verbal protocol analyses were taken

as support for the notion that reliance on general domain knowledge was related to better performance in the SJT.

At this point, we would still argue that test takers try to understand the situation to which the response options are related, with the only difference that an understanding of this situation is (at least somewhat) more difficult in comparison with when an item stem with a description of the situation is included. Furthermore, the evidence provided by Krumm et al. (2015) does not rule out this suggestion. Even though Krumm et al. (2015) write that “use of general knowledge” was related to a higher likelihood of choosing the correct response option (p. 410), many of the examples in their article suggest to us that this knowledge was related to the correct interpretation of the situation in the first step. Thus, test takers, for example, first had to discern that the situation *required* effective team performance before they could make a judgment about the most effective behavior to ensure effective team performance in the second step.

How Shall Situational Judgment Be Taken Into Account in the Development of SJTs?

One way, concerning the development of more construct valid SJTs, is to target only a single construct in an SJT. In line with this, the overview by Christian, Edwards, and Bradley (2010) showed that many SJTs focus on only one construct—or at least on only a single construct category or domain. A potential drawback with this approach is that the targeted kind of effective behavior might become too obvious across a set of SJT items that are always related to the same trait and requiring the same kind of action. In such a case, an SJT might become too transparent for test takers, and it might possibly be that this might ultimately lead to an impairment of the SJT’s criterion-related validity. Thus, if the degree to which SJTs require situational judgment becomes too limited, systematic variance that is related to performance in the SJT as well as on the job might be lost. In line with this, evidence from ACs has shown that criterion-related validity diminished once the targeted AC dimensions were made very transparent to participants (Ingold, Kleinmann, König, & Melchers, in press).

An approach that would alleviate the risk to make SJTs too transparent is to capture more than a single construct. Similarly, as suggested by Lievens and Motowidlo, if one targets compound traits, which consist of relatively different facets that are relevant for effective performance, then the risk of a harmful increase in transparency might be prevented.

Another option that might be used—at least when an SJT is not too transparent—would be to make use of situational judgments in addition to judgments of response effectiveness. On the basis of the SJT research by Rockstuhl et al. (2015) as well as on results from the interview and the AC

domains (e.g., Ingold et al., 2015; Jansen et al., 2013), one could try to additionally collect situational judgments and to use them for making selection decisions. However, in this case, more research would be necessary to determine the acceptability among recruiters as well as applicants of the different ways in which such judgments might be collected.

References

- Christian, M. S., Edwards, B. D., & Bradley, J. C. (2010). Situational judgment tests: Constructs assessed and a meta-analysis of their criterion-related validities. *Personnel Psychology*, 63, 83–117. doi:[10.1111/j.1744-6570.2009.01163.x](https://doi.org/10.1111/j.1744-6570.2009.01163.x)
- Ingold, P. V., Kleinmann, M., König, C. J., & Melchers, K. G. (in press). Transparency of assessment centers: Lower criterion-related validity but greater opportunity to perform? Advance online publication. *Personnel Psychology*. doi:[10.1111/peps.12105](https://doi.org/10.1111/peps.12105)
- Ingold, P. V., Kleinmann, M., König, C. J., Melchers, K. G., & Van Iddekinge, C. H. (2015). Why do situational interviews predict job performance? The role of interviewees' ability to identify criteria. *Journal of Business and Psychology*, 30, 387–398. doi:[10.1007/s10869-014-9368-3](https://doi.org/10.1007/s10869-014-9368-3)
- Jansen, A., Melchers, K. G., Lievens, F., Kleinmann, M., Brändli, M., Fraefel, 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. doi:[10.1037/a0031257](https://doi.org/10.1037/a0031257)
- Kleinmann, M., Ingold, P. V., Lievens, F., König, C. J., Melchers, K. G., & Jansen, A. (2011). A different look at why selection procedures work: The role of candidates' ability to identify criteria. *Organizational Psychology Review*, 1, 128–146. doi:[10.1177/2041386610387000](https://doi.org/10.1177/2041386610387000)
- 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](https://doi.org/10.1037/a0037674)
- Lance, C. E. (2008). Why assessment centers do not work the way they are supposed to. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 1, 84–97. doi:[10.1111/j.1754-9434.2007.00017.x](https://doi.org/10.1111/j.1754-9434.2007.00017.x)
- 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.
- Macan, T. (2009). The employment interview: A review of current studies and directions for future research. *Human Resource Management Review*, 19, 203–218. doi:[10.1016/j.hrmr.2009.03.006](https://doi.org/10.1016/j.hrmr.2009.03.006)
- Maurer, S. D., Sue-Chan, C., & Latham, G. P. (1999). The situational interview. In R. W. Eder & M. M. Harris (Eds.), *The employment interview handbook* (pp. 159–177). Thousand Oaks, CA: Sage.
- Melchers, K. G., Klehe, U.-C., 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. doi:[10.1080/08959280903120295](https://doi.org/10.1080/08959280903120295)
- Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. *Psychological Review*, 80, 252–283. doi:[10.1037/h0035002](https://doi.org/10.1037/h0035002)

- Oostrom, J. K., Melchers, K. G., Ingold, P. V., & Kleinmann, M. (in press). Why do situational interviews predict performance? Is it saying how you would behave or knowing how you should behave? Advance online publication. *Journal of Business and Psychology*. doi:[10.1007/s10869-015-9410-0](https://doi.org/10.1007/s10869-015-9410-0)
- 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. doi:[10.1037/a0038098](https://doi.org/10.1037/a0038098)
- Speer, A. B., Christiansen, N. D., Melchers, K. G., König, C. J., & Kleinmann, M. (2014). Establishing the cross-situational convergence of the ability to identify criteria: Consistency and prediction across similar and dissimilar assessment center exercises. *Human Performance, 27*, 44–60. doi:[10.1080/08959285.2013.854364](https://doi.org/10.1080/08959285.2013.854364)

Clearly Defined Constructs and Specific Situations Are the Currency of SJTs

Lijun Chen
Zhejiang University

Jinyan Fan, Lu Zheng, and Elissa Hack
Auburn University

Although we echo Lievens and Motowidlo's (2016) view that situational judgment test (SJT) research should subscribe to the construct-driven approach, we disagree with their argument on two counts. First, we question whether measuring general domain knowledge represents the only way to advance SJT research. Second, we question whether it is appropriate to downplay the importance of situations in SJTs. In this commentary, we first briefly review construct-driven SJT studies and then share our own experience in developing an SJT for integrity in China using the construct-driven approach. Based on the review and reflection, we come to two major conclusions: (a) construct-driven SJT research has progressed well so far without the reconceptualization of SJTs as measures of general domain knowledge, and (b) specific situations are an important feature of SJTs that should not yet be dismissed.

Lijun Chen, College of Public Affairs, Zhejiang University, Hangzhou, Zhejiang, People's Republic of China; Jinyan Fan, Lu Zheng, and Elissa Hack, Department of Psychology, Auburn University.

Correspondence concerning this article should be addressed to Lijun Chen, College of Public Affairs, Zhejiang University, 234 Mengmingwei Hall, 866 Yuhangtang Road, Hangzhou, Zhejiang, People's Republic of China 310058. E-mail: lijunchen@zju.edu.cn