Industrial and Organizational Psychology (2024), **17**, 149–153
doi:10.1017/iop.2024.8

CAMBRIDGE UNIVERSITY PRESS

FOCAL ARTICLE

Assessment centers: Reflections, developments, and empirical insights

Duncan J. R. Jackson[1], Michael D. Blair[2], and Pia V. Ingold[3]

[1]King's Business School, King's College London, London, UK, [2]U.S. Office of Personnel Management, Kansas City, MO, USA and [3]Department of Psychology, University of Copenhagen, Copenhagen, Denmark
Corresponding author: Duncan J. R. Jackson; Email: duncan.jackson@kcl.ac.uk

(Received 19 February 2024; accepted 19 February 2024; first published online 06 May 2024)

Keywords: Introduction

Assessment centers (ACs) are a popular evaluation approach often applied for the purposes of guiding employment selection and development decisions. AC participants are required to engage in a series of work simulation exercises (e.g., role plays, group discussions, and presentations), and their performance on those exercises is rated by trained assessors. It is these interactions between participants and work simulations that inspired early organizational interest in the AC approach in the late 1940s (Handyside & Duncan, 1954; Highhouse & Nolan, 2012) that has remained into the present.

ACs continue to hold appeal in contemporary organizations, which is likely due, in part, to the their interpersonal nature (Kleinmann & Ingold, 2019) and the rich source of job-relevant information they provide, particularly on job candidates and for employee development (Lievens, 2009). In the same manner, ACs continue to motivate the interests of researchers, as evidenced by the volume of empirical articles on ACs published over the last 10 years (e.g., Breil et al., 2023; Dimotakis et al., 2017; Heimann et al., 2022; Hickman et al., 2023; Hoffman et al., 2015; Ingold et al., 2016, 2018; Jackson et al., 2016, 2022; Jansen et al., 2013; Kuncel & Sackett, 2014; Lievens et al., 2015; Meriac et al., 2014; Monahan et al., 2013; Oliver et al., 2016; Putka & Hoffman, 2013; Sackett et al., 2017; Speer et al., 2014; Thornton et al., 2019; Wirz et al., 2020).

A link to the inaugural issue

Even though it only dates back roughly 15 years, it is a testament to research- and practice-based interest in ACs that the first issue of *Industrial-Organizational Psychology: Perspectives on Science and Practice* (*IOP*) included a focal article on ACs by Charles Lance along with responses to his article.[1] Lance (2008) contributed a critique of how ratings from ACs are scored. He concluded that ACs do not measure dimensions and that attempts to use ACs to generate dimension scores should be abandoned. Points for consideration, reactions, and diverging points of view raised by Lance's critique (e.g., Arthur et al., 2008; Howard, 2008; Rupp et al., 2008) have consistently elicited novel research on ACs.

[1]For those interested in reading Lance's focal article and the responses to the article, Volume 1, Issue 1 of IOP is available online: https://www.cambridge.org/core/journals/industrial-and-organizational-psychology/issue/F8319F5B9E1B45CC024A74BE3AFEBB01

© The Author(s), 2024. Published by Cambridge University Press on behalf of Society for Industrial and Organizational Psychology. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

https://doi.org/10.1017/iop.2024.8 Published online by Cambridge University Press

During the 15 years that have passed since the Lance (2008) focal article and its responses, perspectives on ACs have continued to develop and have benefitted from knowledge generated from ongoing research. This special issue on ACs provides an opportunity to reflect again on what conceptual perspectives on ACs have prevailed and emerged since 2008 as well as an opportunity to explore and showcase recent insightful empirical AC research. It moreover provides insights of value to practice and on how ACs can continue to provide an abundant source of information for organizational research and for decision makers.

Overview of papers in the special issue

Dewberry (2024) reviews research literature concerned with whether ACs assess dimensions (or competencies, e.g., communication skills, tolerance) reliably and as intended. Fifteen years after Lance's focal article, Dewberry reviews more recent research on whether ACs measure dimensions, particularly research utilizing generalizability theory (G theory), which provides statistically controlled estimates of dimension effects. He concludes that evidence derived from G theory research confirms that ACs do not measure dimensions and concurs with Lance that attempts to measure dimensions with ACs should be abandoned. Dewberry moreover presents an argument against interactionist perspectives on ACs (e.g., trait activation theory, the mixed-model perspective), suggesting that some of the patterns used to support these perspectives may simply reflect artifacts of the AC measurement design. This work raises considerations for practitioners about whether they should cease attempts to use ACs to measure dimensions, scoring them instead only in relation to exercises and/or overall performance.

In a conceptual paper, Nottingham and Rupp (2024) propose that ACs could be used to serve the emerging aim of selecting and developing inclusive leaders in organizations. In contrast to the Dewberry paper, using a dimension approach, Nottingham and Rupp argue that measuring inclusive leadership may add incremental validity to overall assessment ratings (OARs). Specifically, Nottingham and Rupp suggest assessing inclusive leadership proficiency as a behavioral leadership dimension and they develop specific propositions about its relationship with leadership and follower performance and diversity among followers. Given that the evaluation of leaders and the fostering of diversity in organizations are key considerations for both research and practice, this article provides valuable guidance on how to potentially optimize evaluation criteria when aiming to assess inclusive leadership.

Prior research has illustrated that AC ratings are dependent on impressions that assessors form of candidates (Ingold et al., 2018; Lance et al., 2004). Yet it is unknown to what degree these impressions remain consistent in ACs and if the impact of these impressions on AC ratings changes across and within AC exercises. Building on the thin-slice paradigm, Ingold et al. (in press) address this topic and investigate the consistency of assessor candidate impressions by using different slices of video material from the beginning, middle, and end of three AC exercises. Results suggest that the impressions participants convey across different time points of the AC (i.e., the beginning, middle, and end of each AC, but also across exercises) are consistent. Moreover, their findings suggest that these impressions predict AC performance and can also relate to participant's job performance. This study offers research insights into the relevance of assessor impressions and into the development of our understanding of assessee behavior.

The topic of assessor training for ACs, which can be positioned in the broad research area of frame-of-reference training (Roch et al., 2012; Woehr & Huffcutt, 1994), is addressed by Gorman et al. (2024). Applied to ACs, frame-of-reference training provides assessors with a common set of standards to evaluate performance with the goal of increasing rater effectiveness and consistency. In their study, Gorman et al. provide a perspective on the multifaceted structure of frame-of-reference training. The authors found that assessor training was of most assistance in the identification of low-performing AC participants. They moreover found that ratings for assessors

https://doi.org/10.1017/iop.2024.8 Published online by Cambridge University Press

who were not trained were associated with larger proportions of residual error than for those who were trained. This study contributes knowledge to research and practice associated with AC training and provides insights into how the variance profile of AC ratings depends on whether assessors have been trained.

In her article, Roch (2024) provides evidence for perceptual differences among applicants relating to AC exercises and an ability test. Different applicant perceptions were found for different exercise types. Whether the AC was rated live or via a recording had implications for fairness perceptions. Moreover, Roch found that whether an assessee had previous experience on an AC influenced levels of perceived self-efficacy. This study contributes to knowledge on applicant reactions that can influence a practitioners' choice of exercises and psychometric tests.

Procedural justice is conceptually related to considerations of ethics, which is a topic of major consequence to organizations. Fostering just, moral, and ethical behavior is paramount: not only for organizations but also for the wider development of society. d'Amato et al. (2024) address this issue in their paper and raise questions about how leaders can develop ethical and moral behavior using the AC method. They provide initial findings suggesting that the development of ethical leadership attitudes may result in negative, backlash-oriented repercussions. For research, this study offers insights into the application of ACs in the context of ethical leadership. For practice, it provides early warnings about some of the pitfalls of attempting to develop attitudes with ACs.

Organizational decision making based on AC results is complex and, as Rupp et al. (2024) suggest, it can involve the decision maker in a consideration of theory, empirical contributions, and best-practice considerations. In their paper, Rupp et al. present an epistemology for the integration of these three factors. They apply their framework to present a perspective on assessment and development that is directly relevant to considerations of ACs. They conclude that there are areas of alignment among theory, empirical contributions, and best practice. However, they also highlight key gaps and areas for further development. The Rupp et al. framework could be applied by both researchers and practitioners to assist in furthering research to help ensure a better integration across theory, research, and practice in assessment and development and also in other areas of complex workplace interventions.

Conclusions

This special issue offers an exploration into existing, new, and alternative lines of inquiry that showcases the progressive, enterprising, and current nature of AC research and development. In our view, modern ACs possess a vast wealth of knowledge and value to offer both individuals and organizations and will continue to stimulate research as they have before and following the first issue of *IOP*. In our reading, the current collection of works suggests that researchers and practitioners are served best by exploring, debating, and engaging with areas for development, challenges, and controversies associated with ACs. We propose that learning more about such issues, stimulating debate around them, and allowing different perspectives to be heard is how a research area can grow and develop. We hope that the current set of papers will contribute to the AC debate, will reignite unresolved controversies, and will stimulate new lines of enquiry. It is through such debates and discussions that we can, as researchers and practitioners, further strive toward an enhanced understanding of the AC method to the continued benefit of individuals and organizations.

Competing interests. None.

https://doi.org/10.1017/iop.2024.8 Published online by Cambridge University Press

References

Arthur, W., Jr., Day, E. A., & Woehr, D. J. (2008). Mend it, don't end it: An alternate view of assessment center construct-related validity evidence. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, **1**, 105–111. https://doi.org/10.1111/j.1754-9434.2007.00019.x

Breil, S. M., Lievens, F., Forthmann, B., & Back, M. D. (2023). Interpersonal behavior in assessment center role-play exercises: Investigating structure, consistency, and effectiveness. *Personnel Psychology*, **76**(3), 759–795. https://doi.org/10.1111/peps.12507

d'Amato, A., Murugavel, V., Mereiros, K., & Watts, L. L. (2024). An ethical leadership assessment center pilot: Assessing and developing moral person and moral manager dimensions. *Industrial and Organizational Psychology*, **17**.

Dewberry, C. (2024). Assessment centers do not measure stable competencies: Why this is now beyond reasonable doubt. *Industrial and Organizational Psychology*, **17**.

Dimotakis, N., Mitchell, D., & Maurer, T. J. (2017). Positive and negative assessment center feedback in relation to development self-efficacy, feedback seeking, and promotion. *Journal of Applied Psychology*, **102**(11), 1514–1527. https://doi.org/10.1037/apl0000228

Gorman, C. A., Jackson, D. J. R., Meriac, J. P., & Himmler, J. R. (2024). Unpacking frame-of-reference assessor training effectiveness. *Industrial and Organizational Psychology*, **17**.

Handyside, J. D., & Duncan, D. C. (1954). Four years later: A follow-up of an experiment in selecting supervisors. *Occupational Psychology*, **28**, 9–23.

Heimann, A. L., Ingold, P. V., Lievens, F., Melchers, K. G., Keen, G., & Kleinmann, M. (2022). Actions define a character: Assessment centers as behavior-focused personality measures. *Personnel Psychology*, **75**(3), 675–705. https://doi.org/10.1111/peps.12478

Hickman, L., Herde, C. N., Lievens, F., & Tay, L. (2023). Automatic scoring of speeded interpersonal assessment center exercises via machine learning: Initial psychometric evidence and practical guidelines. *International Journal of Selection and Assessment*. https://doi.org/10.1111/ijsa.12418.

Highhouse, S., & Nolan, K. P. (2012). One history of the assessment center. In D. J. R. Jackson, C. E. Lance, & B. J. Hoffman (Ed.), *The psychology of assessment centers* (pp. 25–44). Routledge/Taylor & Francis Group.

Hoffman, B. J., Kennedy, C. L., LoPilato, A. C., Monahan, E. L., & Lance, C. E. (2015). A review of the content, criterion-related, and construct-related validity of assessment center exercises. *Journal of Applied Psychology*, **100**, 1143–1168. https://doi.org/10.1037/a0038707

Howard, A. (2008). Making assessment centers work the way they are supposed to. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, **1**, 98–104. https://doi.org/10.1111/j.1754-9434.2007.00018.x

Ingold, P. V., Dönni, M., & Lievens, F. (2018). A dual-process theory perspective to better understand judgments in assessment centers: The role of initial impressions for dimension ratings and validity. *Journal of Applied Psychology*, **103**(12), 1367–1378. https://doi.org/10.1037/apl0000333

Ingold, P. V., Heimann, A. L., & Breil, S. M. (in press). Any slice is predictive? On the consistency of impressions from the beginning, middle, and end of assessment center exercises and their relation to performance. *Industrial and Organizational Psychology: Perspectives on Science and Practice.*

Ingold, P. V., Kleinmann, M., Konig, C. J., & Melchers, K. G. (2016). Transparency of assessment centers: Lower criterion-related validity but greater opportunity to perform? *Personnel Psychology*, **69**(2), 467–497. https://doi.org/10.1111/peps.12105

Jackson, D. J. R., Michaelides, G., Dewberry, C., Nelson, J., & Stephens, C. (2022). Reliability in assessment centres depends on general and exercise performance, but not on dimensions. *Journal of Occupational and Organizational Psychology*, **95**(4), 739–757. https://doi.org/10.1111/joop.12398

Jackson, D. J. R., Michaelides, M., Dewberry, C., & Kim, Y. (2016). Everything that you have ever been told about assessment center ratings is confounded. *Journal of Applied Psychology*, **101**(7), 976–994. https://doi.org/10.1037/apl0000102

Jansen, A., Melchers, K. G., Lievens, F., Kleinmann, M., Brandli, M., Fraefel, L., & Konig, 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**(2), 326–341. https://doi.org/10.1037/a0031257

Kleinmann, M., & Ingold, P. V. (2019). Toward a better understanding of assessment centers: A conceptual review. *Annual Review of Organizational Psychology and Organizational Behavior*, **6**, 349–372. https://doi.org/10.1146/annurev-orgpsych-012218-014955

Kuncel, N. R., & Sackett, P. R. (2014). Resolving the assessment center construct validity problem (as we know it). *Journal of Applied Psychology*, **99**(1), 38–47. https://doi.org/10.1037/a0034147

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**(1), 84–97. https://doi.org/10.1111/j.1754-9434.2007.00017.x

Lance, C. E., Foster, C., Gentry, W. A., & Thoresen, J. D. (2004). Assessor cognitive processes in an operational assessment center. *Journal of Applied Psychology*, **89**, 22–35.

https://doi.org/10.1017/iop.2024.8 Published online by Cambridge University Press

Lievens, F. (2009). Assessment centres: A tale about dimensions, exercises, and dancing bears. *European Journal of Work and Organizational Psychology*, **18**, 102–121.

Lievens, F., Schollaert, E., & Keen, G. (2015). The interplay of elicitation and evaluation of trait-expressive behavior: Evidence in assessment center exercises. *Journal of Applied Psychology*, **100**(4), 1169–1188. https://doi.org/10.1037/apl0000004

Meriac, J. P., Hoffman, B. J., & Woehr, D. J. (2014). A conceptual and empirical review of the structure of assessment center dimensions. *Journal of Management*, **40**, 1269–1296. https://doi.org/10.1177/0149206314522299

Monahan, E. L., Hoffman, B. J., Lance, C. E., Jackson, D. J. R., & Foster, M. R. (2013). Now you see them, now you do not: The influence of indicator-factor ratio on support for assessment center dimensions. *Personnel Psychology*, **66**, 1009–1047. https://doi.org/10.1111/peps.12049

Nottingham, A., & Rupp, D. E. (2024). Inclusive leadership as a valid assessment center dimension. *Industrial and Organizational Psychology*, **17**.

Oliver, T., Hausdorf, P., Lievens, F., & Conlon, P. (2016). Interpersonal dynamics in assessment center exercises: Effects of role player portrayed disposition. *Journal of Management*, **42**(7), 1992–2017. https://doi.org/10.1177/0149206314525207

Putka, D. J., & Hoffman, B. J. (2013). Clarifying the contribution of assessee-, dimension-, exercise-, and assessor-related effects to reliable and unreliable variance in assessment center ratings. *Journal of Applied Psychology*, **98**(1), 114–133. https://doi.org/10.1037/a0030887

Roch, S. G. (2024). Perceptions of assessment center exercises: Between exercises differences and interventions. *Industrial and Organizational Psychology*, **17**.

Roch, S. G., Woehr, D. J., Mishra, V., & Kieszczynska, U. (2012). Rater training revisited: An updated meta-analytic review of frame-of-reference training. *(Journal of Occupational Psychology), Journal of Occupational and Organizational Psychology*, **85**(2), 370–395. https://doi.org/10.1111/j.2044-8325.2011.02045.x

Rupp, D. E., Thornton, G. C., III., Bisbey, T. M., Nottingham, A., Salas, E., & Murphy, K. R. (2024). An epistemology for assessment and development: How do we know what we know? *Industrial and Organizational Psychology*, **17**.

Rupp, D. E., Thornton, G. C., & Gibbons, A. M. (2008). The construct validity of the assessment center method and usefulness of dimensions as focal constructs. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, **1**, 116–120. https://doi.org/10.1111/j.1754-9434.2007.00021.x

Sackett, P. R., Shewach, O. R., & Keiser, H. N. (2017). Assessment centers versus cognitive ability tests: Challenging the conventional wisdom on criterion-related validity. *Journal of Applied Psychology*, **102**(10), 1435–1447, https://doi.org/10.1037/apl0000236

Speer, A. B., Christiansen, N. D., Goffin, R. D., & Goff, M. (2014). Situational bandwidth and the criterion-related validity of assessment center ratings: Is cross-exercise convergence always desirable? *Journal of Applied Psychology*, **99**, 282–295, https://doi.org/10.1037/a0035213

Thornton, G. C., Rupp, D. E., Gibbons, A. M., & Vanhove, A. J. (2019). Same-gender and same-race bias in assessment center ratings: A rating error approach to understanding subgroup differences. *International Journal of Selection and Assessment*, **27**(1), 54–71. https://doi.org/10.1111/ijsa.12229

Wirz, A., Melchers, K. G., Kleinmann, M., Lievens, F., Annen, H., Blum, U., & Ingold, P. V. (2020). Do overall dimension ratings from assessment centres show external construct-related validity? *European Journal of Work and Organizational Psychology*, **29**(3), 405–420. https://doi.org/10.1080/1359432X.2020.1714593

Woehr, D. J., & Huffcutt, A. I. (1994). Rater training for performance-appraisal—a quantitative review. *Journal of Occupational and Organizational Psychology*, **67**, 189–205.

Cite this article: Jackson, D. J. R., Blair, M. D., & Ingold, P. V. (2024). Assessment centers: Reflections, developments, and empirical insights. *Industrial and Organizational Psychology* **17**, 149–153. https://doi.org/10.1017/iop.2024.8

https://doi.org/10.1017/iop.2024.8 Published online by Cambridge University Press