

The Relationship Between Assessment Centre Outcomes and Personality Traits

A Confirmation of Nomological Validity

Zsolt Péter Szabó^{1*}, Evelin Simon², Máté Pinczés³, Róbert Laczai³, Ágota Kun¹

¹ Department of Ergonomics and Psychology, Faculty of Economics and Social Sciences, Budapest University of Technology and Economics, H-1117 Budapest, P. O. B. 91, Hungary

² Department of Organisational and Leadership Psychology, Faculty of Education and Psychology, Eötvös Loránd University, H-1075 Budapest, Kazinczy u. 23-27, Hungary

³ At Work Consulting and Service Ltd., H-1093 Budapest, Lónyay u. 41, Hungary

* Corresponding author, e-mail: szabo.zsolt.peter@gtk.bme.hu

Received: 05 November 2020, Accepted: 10 November 2021, Published online: 21 March 2022

Abstract

The present study aimed to assess the nomological validity of the Assessment Centre (AC). More specifically, the authors investigated the relationship between theoretically related personality variables and AC ratings in a personnel selection procedure. Previous studies have reported contradictory findings regarding the nomological validity of the Assessment Centre. We identified shortcomings in the previous studies and demonstrated the nomological validity of the AC in a relatively large sample of line manager candidates ($N = 314$). Our results are in harmony with trait activation theory, and this confirms the nomological validity of Assessment Centres. The theoretical and practical implications of these findings are discussed.

Keywords

Assessment Centre, nomological validity, digitalised testing, personnel selection

1 Introduction

Assessment Centres (AC) are powerful tools for evaluating candidates for selection and promotion purposes. Ideally, ACs are designed to measure job-related behaviours for effective performance in the focal work role; the behavioural ratings are provided by multiple trained assessors; and multiple job-related simulation exercises are used. There are three types of ratings in an AC: dimensional ratings, exercise ratings, and overall assessment rating. Dimensional ratings focus on cross-situational competencies of the participants such as communication, persuasion skills, organising and planning skills, and so on. Exercise ratings focus on situation-specific, exercise effects, in other words the participant's behavioural pattern when solving a situation (e.g., feedback to a supervisor, feedback to a subordinate, and so on). The third type of evaluation provides information on the assessee's general performance as measured by their overall assessment rating (OAR) (Christiansen et al., 2013). Over the past years, research on ACs has primarily focused on the construct validity of ACs. The main question was whether ACs primarily measure the cross-situational

competencies of participants (as originally intended), or the participant's ability to perform effectively in certain situations (Lievens, 2009). Simonenko et al. (2013) point out that in addition to studying the internal structure, more attention should be paid to the nomological validity of the AC. Nomological validity is another type of construct validity that is met when the theoretically expected relationships between the studied variables are confirmed by a nomological network (Weiber and Mühlhaus, 2010). In the present context, nomological validity would be confirmed if the personality traits theoretically related to the variables measured by this AC showed the predicted pattern irrespective of the magnitude of the associations. The aim of our study was to assess the nomological validity of the AC using different types of AC evaluations.

1.1 The relationship between AC outcomes and personality traits

The assumption that AC outcomes and personality traits are related is inherent to the dimensional approach, which

dominated the early history of the AC (Bray et al., 1974). However, the introduction of exercise-related measures and the OAR in the methodological repertoire cast doubt on the position assuming direct relationships between personality traits and the measures provided by a typical AC procedure. In the next section, we will summarise previous findings on the associations between personality traits and AC ratings. We will mainly focus on the relationship between AC ratings and Big Five traits, as our personality measures also employed the Big Five terminology. The Big Five model of personality postulates five broad domains: Emotional stability, Extraversion, Openness, Agreeableness, and Conscientiousness. Emotional stability indicates the extent to which someone experiences negative emotions. Openness indicates how innovative and open-minded someone is. Extraversion indicates the extent to which someone is sociable, and socially active. Agreeableness indicates the extent to which someone values getting along with others. Conscientiousness indicates how organised, goal-directed, and disciplined someone is (Costa and McCrae, 1992).

Inconsistent findings have been reported on the relationships between the AC ratings and personality traits. The OAR has most frequently shown a positive association with extraversion: higher levels of extraversion were in most cases associated with better overall AC ratings (Christiansen et al., 2013; Collins et al., 2003; Scholz and Schuler, 1993; Höft and Schuler, 2001; for an exception see Lance et al., 2007).

Hypothetical links between exercise ratings and personality traits have been suggested in the interactional approach of the trait activation theory (TAT; Tett and Guterman, 2000). TAT suggests that the activation of a trait depends on the extent to which the specific trait is relevant to the situation, and on the strength of the situation. A trait is relevant to a given situation if there are cues that activate the trait, while situational strength determines whether the specific trait or the lack thereof has an impact on the outcome of the situation. Strong situations are defined as situations with a high degree of structure that provide salient cues for behaviour. Weak situations are defined as situations with a low degree of structure that provide no such cues for behaviour. Strong situations elicit similar behaviours from most individuals regardless of their stable traits, while weak situations leave more scope for individual differences. Previous studies found that extraversion and openness were weakly but significantly correlated with participants' performance in leaderless group discussions (LGDs), in role-play exercises, and

on oral presentation tasks (Craik et al., 2002; Hoffman et al., 2015). Conscientiousness, agreeableness, and emotional stability showed weak or no correlation with performance on AC exercises, although some studies revealed an association between conscientiousness and performance in in-basket exercises (Craik et al., 2002; Hoffman et al., 2015). This latter finding along with the relationship between extraversion and performance in leaderless group discussions supports the nomological validity of the AC (Lievens and Christiansen, 2012; Hoffman et al., 2015).

The existing findings on the associations between personality traits and dimension ratings are likewise inconsistent. Lievens and Christiansen (2012) point out that there are at least as many studies that did not find a relationship between AC dimensions and conceptually related personality traits (e.g., Chan, 1996; Fleenor, 1996; Goffin et al., 1996) as there are studies that did (e.g., Dilchert and Ones, 2009; Shore et al., 1990; Thornton et al., 1997). A meta-analysis published by Meriac and colleagues (Meriac et al., 2008) revealed that in most studies weak or non-significant associations were found between AC dimensions and personality traits. Furthermore, while some of the reviewed findings supported nomological validity (e.g., the association between extraversion and the ability to exert social influence), other results pointed to nomological invalidity (e.g., a non-significant relationship between conscientiousness and organising and planning skills). In a subsequent study, Meriac, Hoffman and Woehr (Meriac et al., 2014) found that dimensional AC outcomes showed significant positive associations with the Big Five markers of "getting ahead" behaviours (extraversion, and openness to a lesser extent) but not with the markers of getting along (conscientiousness, agreeableness, and emotional stability).

Some of the above-mentioned inconsistent findings can be explained by the problematic nature of the available evidence. Most primary studies about the relationship between AC evaluations and personality traits often relate to one-time personnel selection projects involving only a small number of candidates. According to Schönbrodt and Perugini (2013) correlational coefficients are often inaccurate in small samples, and in typical scenarios the sample size should approach 250 participants for stable estimates. Unfortunately, most studies do not meet this recommendation (e.g., Chan, 1996: $N = 46$; Craik et al., 2002: $N = 119$; Fleenor, 1996: $N = 102$; Goffin et al., 1996: $N = 68$; Haaland and Christiansen, 2002: $N = 79$; for exceptions see Dilchert and Ones, 2009; Shore et al., 1990; Spector et al., 2000).

In our study, we investigated the relationship between personality variables and AC ratings in a large sample of Hungarian frontline managers. We briefly turn our attention to the context of our study.

1.2 The context of the present study

The empirical data were collected as part of a selection programme developed for a Hungarian-based manufacturing company. Candidates for line manager positions were selected from the company's internal talent pool. During the programme, which was run continuously for three years, a total of 314 candidates participated in an assessment procedure. In the first phase, participants completed a test battery containing personality measures relevant to the frontline manager position, in the second phase, groups of 3 to 5 participants were assessed in an AC procedure comprising three exercises. An initial job analysis preceded the programme.

Based on the preliminary job analysis, the most difficult challenges were the following. First, line managers acted as the main communicational bridge between the different hierarchical levels of the organisation. They needed communicational and social skills to handle conflicts, persuade and convince supervisors and subordinates, and work together with their peers. They had to analyse and solve problems. Another challenge was that the frontline managers were usually promoted from the employee level, creating role conflict and role ambiguity. They needed to manage former peers, and the ability to distance from former co-workers was considered essential. The ideal frontline manager was defined in six dimensional terms:

1. problem-solving and decision-making;
2. effective communication;
3. conflict-handling;
4. persuasion;
5. responsibility;
6. and outcome-orientation.

The AC procedure comprised three exercises to assess these dimensions: a leaderless group discussion (LGD), a potentially conflictual situation involving feedback to a superior (role play), and a potentially conflictual situation involving feedback to a subordinate (role play).

1.3 The aim and hypotheses of the present study

The present study was primarily aimed at verifying the nomological validity of the AC. Our expectation was that extraversion would be positively associated with OAR, the

exercise ratings of LGD and role-plays, and dimension ratings related to social skills and social influence. We also expected a negative relationship between agreeableness and exercise and dimension ratings related to conflict, especially with subordinates. We also expected that conscientiousness would be related to problem-solving and decision making in the LGD, and to ratings in exercises which had a strong problem-solving element.

2 Study

2.1 Participants and procedure

The overall sample included 314 participants (237 males, 75.5%). The mean age of the sample was 37.5 years ($SD = 8.39$), and participants' age varied between 23 and 65 years.

Personality tests were administered in an electronic format via the computerised Vienna Test System (VTS). ACs were administered in groups of 3 to 5 participants. They received written information both on the digital tests and on the AC procedure and gave their written consent to participation prior to starting the assessment.

2.2 Measures

2.2.1 Assessment Centre: tasks

The AC procedure comprised three exercises, which are briefly presented below.

Task 1. Leaderless group discussion. In this exercise, participants cooperated with each other in a group setting. They received written instructions informing them that their superior was taken to hospital due to an accident, and he would be absent from work for a month, during which period they had to fulfil his duties. Each participant had 25 minutes for individual preparation, then they had 30 minutes to make a joint decision on the actions to take. This task was designed to measure the following dimensions: problem solving and decision making; outcome orientation; communication.

Task 2. Feedback to superior. In this exercise, each participant had a one-on-one meeting with the site director of their company played by a confederate. According to the instructions, participants were hired by the company a few months before, and the CEO routinely asked new managers for feedback on their first experiences at the company. Participants received written information on the points to be addressed in their feedback, some of which involved potential conflict with their superior, while others referred to issues whose solution would considerably improve the efficiency of the organisation. Each participant had 20 minutes

for individual preparation, and 15 minutes to give feedback to their superior. This task was designed to measure the following dimensions: problem solving and decision making; responsibility; conflict management; persuasion skills.

Task 3. Feedback to subordinate in a lower managerial position. In this exercise, participants had a one-on-one meeting with a subordinate in a lower managerial position, who was played by a confederate. The instructions informed them that several problems had arisen with their subordinate recently, despite their having been a productive member of the organisation for many years. They were then instructed to meet their subordinate to clarify the issues with him in person. Each participant had 15 minutes for individual preparation, then they had 20 minutes to conduct the meeting. This task was designed to measure the following dimensions: communication skills; responsibility; conflict management; persuasion skills.

Individual participants' performance in each exercise was evaluated on an evaluation sheet, which contained a rating scale assigned to each measured dimension and a list of the behavioural indicators of each dimension. Each rating scale ranged from 1 ("The dimension is not or minimally demonstrated by the assessee") to 5 ("Demonstration of the dimension is outstanding and exemplary"). Each participant was assessed in each exercise by at least two evaluators. The AC procedure was designed in such a way that each measured dimension was assessed twice during the three exercises (outcome orientation was assessed only once). Participants' performance in the AC exercises was evaluated by expert assessors. The performance of the participants was rated in three different ways in the AC: OAR; exercise ratings; and dimension ratings.

2.2.2 Personality measures

The test battery used to assess participants' personality traits comprised three self-reported personality tests. The tests were administered via the VTS: participants were presented with the instructions on a computer screen and their responses were processed automatically. In the VTS, individual assessee's test results are expressed in a percentile rank (PR) on each dimension, which ranges from 0 to 100 and indicates the assessee's rank as compared to the norm group. Unfortunately, the VTS does not provide access to the raw scores, thus reliability measures of the test variables could not be calculated. For the sake of clarity, the variables of the three personality tests were put together in an exploratory factor analytic procedure.

The *Big Five Structure Inventory* (BFSI) (Arendasy et al., 2011) assesses the Big Five personality traits. For the purposes of the present study, 12 out of the 30 subscales were selected based on the results of the preliminary job analysis. Each subscale was measured with 10 items. Five subscales were related to conscientiousness, three were related to emotional stability, two were related to openness, and one was related to agreeableness and extraversion each. Participants rated adjectives according to the extent to which each applied to them. Each adjective was rated on a four-point scale ranging from 1 ("untypical for me") to 4 ("typical for me"). Reliability (α) for the 12 BFSI subscales varied between 0.77 and 0.92 with a mean value of 0.84 in previous studies (Arendasy et al., 2011).

The *Inventory for Personality Assessment in Situations* (IPS) (Schaarschmidt and Fischer, 2011) is a personality test that requires respondents to appraise behaviours and experiences typically occurring in everyday life. The IPS presents the assessed trait-relevant behaviours in specific situational contexts. The test comprises 15 situation descriptions, each of which is followed by several related statements, and the respondent is requested to indicate the extent to which each statement applies to them. The 15 situations are associated with a total of 80 statements, which compose 15 variables. Respondents rate each statement on a four-point scale ranging from 1 ("definitely true") to 4 ("not true at all"). For the purposes of the present study, 8 variables were included in the data analysis:

- Assertiveness,
- Tendency to confrontation,
- Considerateness,
- Communicational activity,
- Inertia when change is required,
- Stability in a stressful situation, Active recreation behaviour,
- Ability to relax.

Reliability (α) for the 15 variables based on normative data varied between 0.71 and 0.91 with a mean value of 0.82 (Schaarschmidt and Fischer, 2011).

The *Management Potential Analysis* (MAP; Sonnenberg and Wottowa, 2013) is a 109-item questionnaire that assesses leadership and work-related attitudes and personality characteristics on 12 dimensions. For the purposes of the present study, 5 out of the 12 variables of the MAP were included in the data analysis, all of which were relevant to the activities associated with the positions offered.

The 5 variables were: Open reaction to frustration and criticism, Empathic attitude to others, Striving for contact, Willingness to work under pressure, and Planned working style. Respondents rate each statement on a four-point scale ranging from 1 ("is not true") to 4 ("is always true"). Reliability (α) for the 12 dimensions of the MAP varied between 0.65 and 0.85 with a mean value of 0.76 (Sonnenberg and Wottowa, 2013).

2.3 Results

AC measures were calculated from individual participants' AC evaluations at three levels (dimension ratings, exercise ratings, OAR). For data reduction purposes, all 25 personality test variables were entered into an exploratory factor analysis using maximum likelihood extraction and promax rotation. The sample met the statistical requirements for factor analysis (KMO = .91, Bartlett's $\chi^2 < .001$). The eigenvalues and the scree plot revealed 4 factors accounting for 49.5% of the total variance.

The first factor (VAR% = 31.34%) included measures related to conscientiousness: self-control, sense of duty, caution, love of order, and discipline. All these variables were measured with the respective BFSI subscales of conscientiousness.

The second factor (VAR% = 6.84%) included measures related to extraversion and the social aspects of coping with stress. Extraversion was measured with the two variables tapping assertive communication (one from the BFSI and IPS each) and with activity in a familiar communicative situation (IPS) and striving for contact (MAP). The social aspects of coping with stress were measured with

competence (BFSI) and with stability in a stressful situation, ability to relax after the working day, and low inertia when change is required (IPS).

The third factor (VAR% = 6.32%) included measures related to emotional stability: equanimity, open reaction to frustration and criticism, low tendency to confrontation, and emotional robustness. Equanimity and emotional robustness were measured with the BFSI, confrontation tendency with the IPS, and reactions to frustration and criticism with the MAP.

The fourth factor (VAR% = 4.96%) included measures related to agreeableness: helpfulness and openness to feelings were measured with the BFSI, supportive communication with the IPS, empathic attitude to others with the MAP.

Descriptive statistics, reliabilities and correlations for the main variables are shown in Table 1. Table 1 shows that the OAR was positively associated with extraversion and negatively with agreeableness. Overall performance in each of the three exercises showed a significant positive relationship with extraversion, while overall performance in each of the two feedback exercises correlated negatively with agreeableness. All dimensional measures were associated positively with extraversion, and negatively with agreeableness. Emotional stability and Conscientiousness were not associated with any of the AC rating.

3 Discussion

The present study investigated the nomological validity of the AC method in a relatively large group of frontline manager candidates. The obtained findings mostly confirm the nomological validity of the AC in the sense that

Table 1 Means, Standard Deviations, Reliabilities and Pearson's Correlation for the AC Ratings and Psychological Test Variables ($N = 314$)

	Mean (SD)	α	<i>C</i>	<i>X</i>	<i>E</i>	<i>A</i>
OAR	31.54 (11.91)	0.94	0.07	0.19**	0.03	-0.17**
LGD	2.77 (1.26)	0.93	0.09	0.21**	0.02	-0.09
Feedback to Superior	3.01 (1.25)	0.95	0.02	0.15**	0.03	-0.18**
Feedback to Subordinate	2.80 (1.29)	0.95	0.07	0.14*	0.03	-0.15**
Problem Solving	2.92 (1.14)	0.70	0.06	0.20**	0.03	-0.14*
Communication	2.91 (1.11)	0.67	0.09	0.20**	0.01	-0.12*
Responsibility	3.08 (1.19)	0.67	0.05	0.16**	0.04	-0.15**
Conflict Management	2.75 (1.21)	0.68	0.04	0.16**	0.05	-0.17**
Persuasion Skills	2.74 (1.25)	0.75	0.06	0.15*	0.01	-0.20**
Outcome Orientation	2.72 (1.43)	-	0.07	0.19**	0.02	-0.11†

Notes: † $p < .10$, * $p < .05$, ** $p < .01$. OAR = Overall Assessment Rating. OAR was a summation of the dimension ratings. The minimal value of OAR was 5, the maximum value was 55. The dimension ratings were in a 5-point scale. Outcome orientation was assessed only in the Leaderless Group Discussion (LGD). Psychological Test Variables were measured in Percentile Rank, minimal PR value was 0 and maximum value was 100. *C* = Conscientiousness, *X* = Extraversion, *E* = Emotional Stability, *A* = Agreeableness

most of the observed correlations met the predictions (see e.g., Christiansen et al., 2013; Lance, 2008; Lievens, 2009; Hoffman et al., 2015; Merkulova et al., 2016; Jackson et al., 2016). The significant associations between overall performance in the three AC exercises and personality measures may be adequately explained in the framework of trait activation theory (Tett and Guterman, 2000), since the findings on extraversion and agreeableness are clearly consistent with the situational contexts of the exercises. All three AC exercises mobilise social skills; thus, it is not a surprise that overall performance on these exercises positively correlated with extraversion. Both one-on-one situations required efficient advocacy of personal and organisational interests and hence facing potential conflicts with the interactional partner, which may explain why a purely supportive approach did not help participants perform well in these situations. Highly agreeable participants often chose to give up their leadership role, especially when giving feedback to a subordinate in a lower managerial position: they acted as a partner much more efficiently than they fulfilled the control function associated with their leadership role. This also explains why agreeableness significantly negatively correlated with outcome orientation in the leaderless group discussion.

Our study also suggests that separate AC projects have their own nomological networks based on the specific design of the given AC. For instance, agreeableness is generally thought to be not connected to AC performance (e.g., Lievens and Christiansen, 2012; Meriac et al., 2014; Hoffman et al., 2015), while in the present AC, agreeableness was negatively related to many AC ratings. From the perspective of a nomological network, it is not surprising that high levels of agreeableness hindered the performance of our candidates in situations with high conflict potential (see also Jensen-Campbell and Graziano, 2001).

The present study contributes to the literature in numerous ways. Most importantly, the present study investigates the relationship between personality traits and AC ratings in a relatively large, homogenous sample. Prior empirical support for the nomological validity of the AC is weak because studies have either struggled with inadequate sample sizes. Our study belongs to the rare exceptions that have sufficient statistical power and a homogenous

sample (see also Dilchert and Ones, 2009; Shore et al., 1990; Spector et al., 2000). Our study also investigates the relationship between stable personality differences and all types of AC ratings.

The observed low magnitude of the predicted correlations between the personality traits and the AC measures has an important practical implication besides its theoretical importance for the nomological validity of the AC. Namely, this observation suggests that the two types of measures provide – at least partly – different types of information on assesseees (i.e. they are non-redundant measures), meaning that it is advisable to use them in combination for selection purposes (see also Christiansen et al., 2013). One of the most important reasons for the weak associations is that the typically used AC measures (exercises, dimensions, OAR) not only reflect personality traits but also various skills, expertise, and experiences.

Nevertheless, the present study has several limitations. The sample and the AC procedure on which the findings are based are specific in that all data were collected from one organisation and all participants were assessed for lower managerial positions. These circumstances considerably restrict the generalisability of the findings across various types of organisations and scopes of activities. The personality tests used in the study are problematic in two respects: they are self-report measures, and they are – apart from the MAP test – not workplace-specific measures of personality (see the importance of workplace-specific measures here: Ziegler, 2014).

4 Conclusion

The present study confirmed the nomological validity of the Assessment Centre. Our study provides additional evidence to the existing literature, confirming the positive relationship between extraversion and AC ratings, and demonstrating that traits such as agreeableness might also relate to AC performance (negatively in our case). The main conclusion of the present study is that there are theoretically predictable associations between personality test measures and AC measures obtained in a selection process assessing candidates for lower managerial positions at a Hungary-based large company.

References

- Arendasy, M., Sommer, M., Feldhammer, M. (2011) "Manual Big-Five Structure Inventory BFSI", Schuhfried, Mödling, Austria.
- Bray, D. W., Campbell, R. J., Grant, D. L. (1974) "Formative years in business: A long term AT & T study of managerial lives", Wiley, New York, NY, USA.
- Chan, D. (1996) "Criterion and construct validation of an assessment centre", *Journal of Occupational and Organizational Psychology*, 69(2), pp. 167–181.
<https://doi.org/10.1111/j.2044-8325.1996.tb00608.x>
- Christiansen, N. D., Hoffman, B. J., Lievens, F., Speer, A. B. (2013) "Assessment Centers and the Measurement of Personality", In: Christiansen, N., Tett, R. (eds.) *Handbook of Personality at Work*, Routledge, New York, NY, USA, pp. 477–497.
- Collins, J. M., Schmidt, F. L., Sanchez-Ku, M., Thomas, L., McDaniel, M. A., Le, H. (2003) "Can Basic Individual Differences Shed Light on the Construct Meaning of Assessment Center Evaluations?", *International Journal of Selection and Assessment*, 11(1), pp. 17–29.
<https://doi.org/10.1111/1468-2389.00223>
- Costa, P. T., McCrae, R. R. (1992) "Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI)", Psychological Assessment Resources, Odessa, Florida, CA, USA.
- Craik, K. H., Ware, A. P., Kamp, J., O'Reilly III, C., Staw, B., Zedeck, S. (2002) "Explorations of construct validity in a combined managerial and personality assessment programme", *Journal of Occupational and Organizational Psychology*, 75(2), pp. 171–193.
<https://doi.org/10.1348/09631790260098758>
- Dilchert, S., Ones, D. S. (2009) "Assessment Center Dimensions: Individual differences correlates and meta-analytic incremental validity", *International Journal of Selection and Assessment*, 17(3), pp. 254–270.
<https://doi.org/10.1111/j.1468-2389.2009.00468.x>
- Fleener, J. W. (1996) "Constructs and developmental assessment centers: Further troubling empirical findings", *Journal of Business and Psychology*, 10(3), pp. 319–335.
<https://doi.org/10.1007/BF02249606>
- Goffin, R. D., Rothstein, M. G., Johnston, N. G. (1996) "Personality testing and the assessment center: Incremental validity for managerial selection", *Journal of Applied Psychology*, 81(6), pp. 746–756.
<https://doi.org/10.1037/0021-9010.81.6.746>
- Haaland, S., Christiansen, N. D. (2002) "Implications of trait-activation theory for evaluating the construct validity of assessment center ratings", *Personnel Psychology*, 55(1), pp. 137–163.
<https://doi.org/10.1111/j.1744-6570.2002.tb00106.x>
- 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(4), pp. 1143–1168.
<https://doi.org/10.1037/a0038707>
- Höft, S., Schuler, H. (2001) "The Conceptual Basis of Assessment Centre Ratings", *International Journal of Selection and Assessment*, 9(1–2), pp. 114–123.
<https://doi.org/10.1111/1468-2389.00168>
- Jackson, D. J., Michaelides, G., Dewberry, C., Kim, Y. J. (2016) "Everything that you have ever been told about assessment center ratings is confounded", *Journal of Applied Psychology*, 101(7), pp. 976–994.
<https://doi.org/10.1037/apl0000102>
- Jensen-Campbell, L. A., Graziano, W. G. (2001) "Agreeableness as a Moderator of Interpersonal Conflict", *Journal of Personality*, 69(2), pp. 323–362.
<https://doi.org/10.1111/1467-6494.00148>
- 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), pp. 84–97.
<https://doi.org/10.1111/j.1754-9434.2007.00017.x>
- Lance, C. E., Foster, M. R., Nemeth, Y. M., Gentry, W. A., Drollinger, S. (2007) "Extending the Nomological Network of Assessment Center Construct Validity: Prediction of Cross-Situationally Consistent and Specific Aspects of Assessment Center Performance", *Human Performance*, 20(4), pp. 345–362.
- Lievens, F. (2009) "Assessment centres: A tale about dimensions, exercises, and dancing bears", *European Journal of Work and Organizational Psychology*, 18(1), pp. 102–121.
<https://doi.org/10.1080/13594320802058997>
- Lievens, F., Christiansen, N. D. (2012) "Core Debates in Assessment Center Research: Dimensions Versus Exercises", In: Jackson, D. J. R., Lance, C. E., Hoffman, B. J. (eds.) *The Psychology of Assessment Centers*, Routledge, New York, NY, USA, pp. 68–93.
- Meriac, J. P., Hoffman, B. J., Woehr, D. J., Fleisher, M. S. (2008) "Further evidence for the validity of assessment center dimensions: A meta-analysis of the incremental criterion-related validity of dimension ratings", *Journal of Applied Psychology*, 93(5), pp. 1042–1052.
<https://doi.org/10.1037/0021-9010.93.5.1042>
- 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(5), pp. 1269–1296.
<https://doi.org/10.1177/0149206314522299>
- Merkulova, N., Melchers, K. G., Kleinmann, M., Annen, H., Szvircsev Tresch, T. (2016) "A test of the generalizability of a recently suggested conceptual model for assessment center ratings", *Human Performance*, 29(3), pp. 226–250.
<https://doi.org/10.1080/08959285.2016.1160093>
- Schaarschmidt, U., Fischer, A. W. (2011) "Manual Inventory for Personality Assessment in Situations (IPS)", Schuhfried, Mödling, Austria.
- Scholz, G., Schuler, H. (1993) "Das nomologische Netzwerk des Assessment Centers: eine metaanalyse" (The nomological network of the assessment center: A metaanalysis), *Zeitschrift für Arbeits- und Organisationspsychologie*, 37, pp 73–85. (in German)
- Schönbrodt, F. D., Perugini, M. (2013) "At what sample size do correlations stabilize?", *Journal of Research in Personality*, 47(5), pp. 609–612.
<https://doi.org/10.1016/j.jrp.2013.05.009>

- Shore, T. H., Thornton III, G. C., McFarlane Shore, L. (1990) "Construct validity of two categories of assessment center dimension ratings", *Personnel Psychology*, 43(1), pp. 101–114.
<https://doi.org/10.1111/j.1744-6570.1990.tb02008.x>
- Simonenko, S., Thornton III, G. C., Gibbons, A. M., Kravtsova, A. (2013) "Personality Correlates of Assessment Center Consensus Competency Ratings: Evidence from Russia", *International Journal of Selection and Assessment*, 21(4), pp. 407–418.
<https://doi.org/10.1111/ijsa.12050>
- Sonnenberg, H.-G., Wottowa, H. (2013) "Manual Management Potential Analysis (MAP)", Schuchfried, Mödling, Austria.
- Spector, P. E., Schneider, J. R., Vance, C. A., Hezlett, S. A. (2000) "The Relation of Cognitive Ability and Personality Traits to Assessment Center Performance", *Journal of Applied Social Psychology*, 30(7), pp. 1474–1491.
<https://doi.org/10.1111/j.1559-1816.2000.tb02531.x>
- Tett, R. P., Guterman, H. A. (2000) "Situation Trait Relevance, Trait Expression, and Cross-Situational Consistency: Testing a Principle of Trait Activation", *Journal of Research in Personality*, 34(4), pp. 397–423.
<https://doi.org/10.1006/jrpe.2000.2292>
- Thornton III, G. C., Tziner, A., Dahan, M., Clevenger, J. P., Meir, E. (1997) "Construct validity of assessment centre judgments: Analyses of the behavioral reporting method", *Journal of Social Behavior & Personality*, 12(5), pp. 109–128.
- Weiber, R., Mühlhaus, D. (2010) "Strukturgleichungs-modellierung" (Structural Equation Modeling), Springer, Berlin-Heidelberg, Germany. (in German)
- Ziegler, M. (2014) "Big Five Inventory of Personality in Occupational Situations, Manual", Schuchfried, Mödling, Austria.