

Evaluating the quality of assessment centers used in employee selection – Development of a Benchmark for Assessment Center Diagnostics (BACDi)

Franziska Schoelmerich*, Jens Nachtwei* & Carsten C. Schermuly*

*Humboldt University Berlin, *Technical University Braunschweig

To cite this article: Schoelmerich, F., Nachtwei, J., & Schermuly C.C. (in press). Evaluating the quality of assessment centers used in employee selection – Development of a Benchmark for Assessment Center Diagnostics (BACDi). In B. Krause, & P. Metzler (Eds.). *Empirische Evaluationsmethoden*. ZeE Verlag: Berlin.

Keywords: assessment center, employee selection, AC quality, evaluation, predictive validity, benchmark

Abstract

Assessment Centers (AC) generally yield an acceptable predictive validity, but the range of validity coefficients varies depending on the study. One reason for these heterogeneous results might be the lack of commonly accepted quality criteria for the use of ACs in organizational practice. The project Benchmark for Assessment Center Diagnostics (BACDi) aims to systematically evaluate AC quality focusing on improving their predictive validity. For this reason, the BACDi-instrument was developed in cooperation with AC experts from research and organizational practice. The instrument provides an evaluation of AC quality according to 67 quality criteria based on scientific evidence or recommendations relating to predictive validity. Furthermore, a norm is created that allows organizations to compare their results to those of similar organizations within a benchmark. This article describes the theoretical background of the

BACDi-project as well as the development of the BACDi-instrument. In addition, the current application and evaluation of the BACDi-instrument and perspectives of the BACDi-project are presented.

Assessment centers in organizational practice

As Klehe (2004, p. 329) points out, “longterm economic fitness is a major objective of most noncharitable organizations”. This objective can be supported by the human resource (HR) department by coordinating employee selection so as to select those applicants that might contribute best to the economic fitness of the organization (Koenig, Klehe, Berchtold, & Kleinmann, 2010). Results of a metaanalysis conducted by Schmidt and Hunter (1986) indicated, that the use of instruments within employee selection yielding a high accuracy in predicting job performance enhances the selection of employees who perform satisfac-

torily in their future job and consequently contribute to the economic success of the organization. For this reason, the accuracy of predicting job performance yielded by selection instruments is of utmost importance and constitutes the most relevant indicator referred to when evaluating these instruments.

A popular instrument used within employee selection to determine applicants' future job performance is the assessment center (AC), which is widely used in organizations all over the world. In a current study by Schuler, Hell, Trapmann, Schaar and Boramir (2007), 57.6% of the German organizations included in the sample reported the application of ACs to select or further develop their employees. The results of a different study show that even 73.4% of the DAX-100-organizations in Germany use ACs (Obermann, Hoefft, & Janke, 2008). Much lower results were yielded by a study of Nachtwei and Schermuly (2009), who found 14% of small

and medium sized companies and 21% of large companies using ACs. The difference between these results might be due to the bias of selfselection. The first two studies asked organizations via e-mail or letter whether they use AC. In general, organizations interested in ACs participate more likely in surveys like that, therefore high results are obtained that might lead to an overestimation of the spreading of ACs. In the third study however, HR professionals were contacted per e-mail or telephone until they provided an answer about their use of AC. In this way, more appropriate results might be obtained.

The AC definition of Nerdinger, Blickle and Schaper (2008) is adopted in this article. The authors describe an AC as a combination of different behavior-oriented simulation exercises (e.g. roleplay, group discussion, presentation), in which several trained assessors observe and rate participants at the same time according to previously defined dimensions.

Predictive validities of assessment centers

The application of ACs in employee selection requires more effort than using instruments such as interviews or tests of general mental ability. According to Schermuly and Nachtwei (2010), German organizations invest between 400€ and 2000€ into each AC participant. Despite high costs, ACs enjoy increasing popularity which is partly due to the fact that a large number of studies demonstrate their ability to predict future job performance. The accuracy of prediction, also referred to as predictive validity,

indicates the relationship between a predictor (latent construct) and a manifest criterion that is measured at a later point in time (Buehner, 2006). In this case, the qualification of an applicant measured by his performance in simulated jobrelevant tasks is considered as a predictor and his performance on the job, measured later by indicators such as supervisor ratings or the number of promotions, constitutes the criterion.

One of the first metaanalyses examining the predictive validities of ACs by Gaugler, Rosenthal, Thornton and Bentson (1987) yielded an average predictive validity coefficient of .37. A similar coefficient ($r = .35$) was found in a longitudinal study over the course of 13 years by Jansen and Vinkenburg (2006). Compared to these results, Schmitt, Gooding, Noe and Kirsch (1984) detected a higher validity coefficient of $r = .41$ whereas a metaanalysis by Aamodt (2004) yielded a lower relationship ($r = .22$) between the overall AC ratings and workplace outcomes. Results demonstrate the variations in predictive validity depending on the study, which might be partly due to the heterogeneity of AC applications described in the next section. According to Klink and Hilke (2006), carefully developed and conducted ACs can reach acceptable to good validities, which justify the elevated financial and personnel investments. In order to reduce the heterogeneity of AC applications and consequently enhance their predictive validity, commonly accepted quality criteria have to be established serving as guidelines for HR professionals in organizational practice when developing, conducting or evaluating

ACs. With predictive validity being the most important indicator of AC quality, quality criteria are expected to enhance AC quality. AC quality is defined as the extent to which the development, conduction and evaluation of an AC is consistent with specific quality criteria and consequently leads to the selection of adequate applicants in an organization, thus enhancing predictive validity.

Quality criteria for assessment centers

The use of ACs in organizational practice leads to a permanent discussion about the appropriate way to conduct ACs or the elements that should be included in an AC. In particular, a proposition of quality criteria for the use of AC in organizations to increase their overall quality by Kanning, Poettker and Gelléri (2007) has been met by severe criticism from scientists as well as practitioners. They criticized the rigid quality criteria as well as the idea of comparing the current state of organizational practice with a target state based on scientific results. Moreover, the importance of dialogues and discussions between researchers and practitioners rather than supporting indoctrinations from researchers directed at practitioners was stressed (Hoeft, 2009; Kersting, 2009; Krause, 2009). Most scientists agree on the fact that the predictive validity largely depends on the way of developing, conducting and evaluating ACs. Until now, several groups of experts from research and organizational practice presented recommendations for the use of ACs to increase their predictive validity and quality.

The most important ones are briefly presented below.

The DIN 33430 (2002) is a German industry norm that contains quality criteria as well as standards for aptitude testing in organizational practice and includes one section dealing with general recommendations for the use of ACs. Moreover, a German group of HR practitioners and scientists called Arbeitskreis fuer Assessment Center (2004) (Research Group on Assessment Center) and the International Task Force on Assessment Center (2009) published several recommendations dealing specifically with the use of ACs in organizations. Similar guidelines were published by the South African Assessment Centre Study Group (2007) as well as the Indonesian Task Force on Assessment Center Guidelines (2002). However, these recommendations are not thoroughly based on the results of scientific research and should only be considered as general guidelines that provide orientation for practitioners that use ACs in organizations (Schuler et al., 2007).

So far, there is no commonly accepted list of quality criteria. The debate mentioned above reveals the demand for a standardized and more detailed instrument including quality criteria to evaluate AC quality. In order to take specific issues of organizational practice stronger into consideration, AC quality should be compared to scientific standards as well as to a norm of organizations in order to enhance predictive validity of ACs.

BACDi-project

In consideration of the reflections and discussions mentioned above,

the project 'Benchmark for Assessment Center Diagnostics' (BACDi) was created in 2009 at the Psychological Department of the Humboldt-University in Berlin under the leadership of Jens Nachtwei and Carsten C. Schermuly. Several objectives are pursued within the BACDi-project:

1. Selecting quality criteria based on scientific evidence or recommendations by researchers that demonstrate an influence on the predictive validity of ACs used within employee selection. In order to ensure transparency, all sources of criteria are available.
2. Encouraging the dialogue between professionals from research and organizational practice by considering their knowledge and experience while composing the quality criteria included in the BACDi-instrument.
3. Considering psychometric standards during the development of the BACDi-instrument and revising it in regular intervals in order to ensure diagnostic quality.
4. Standardizing the BACDi-results by comparing AC quality not only to scientific standards but also to a norm including other organizations, thus creating a benchmark of AC quality. Moreover, ratings of appropriateness and practicability by scientists and practitioners are presented to complete the benchmark.
5. Focusing on specific rather than global quality criteria that allow for a detailed feedback to facilitate the improvement of AC quality.
6. Implementing these quality criteria in organizational practice to enhance the quality of developing, conducting and evaluating ACs.

Development of BACDi-instrument

The development and evaluation of the BACDi-instrument includes seven stages (Figure 1): First of all, an extensive amount of scientific literature relating to ACs was inspected and analyzed, focusing on research about the predictive validity of ACs used in employee selection. As a result, 103 quality criteria were selected that significantly influenced the predictive validity of ACs used in employee selection. In several meetings of the BACDi-project-team, redundancies in these quality criteria as well as inconsistent empirical or theoretical findings were identified which led to an elimination of 33 criteria. The 70 remaining quality criteria were assigned to the AC phases development, conduction and postprocessing, the latter including the analysis of AC results as well as AC evaluation.

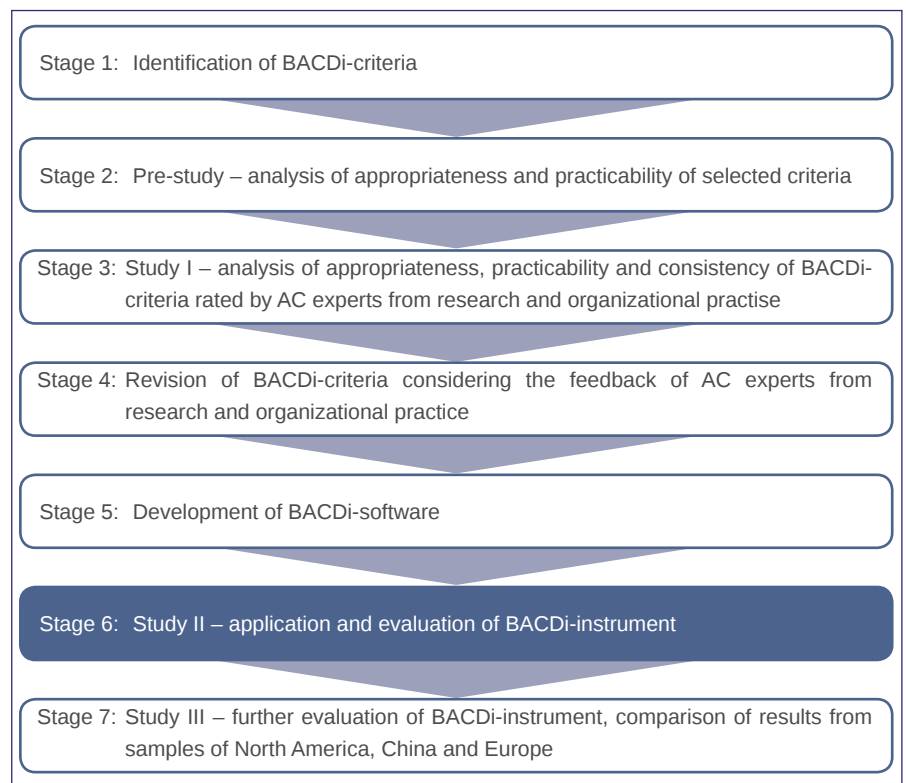
Subsequently, a preliminary study was conducted to verify whether a sufficient number of HR practitioners is interested in the BACDi-project and whether their answers yield enough variance to detect statistical differences. In addition to questions about the consideration of quality criteria in organizational practice, HR professionals that use ACs in their organizations (henceforth referred to as 'AC experts') were asked to rate the quality criteria in terms of their appropriateness as well as their practicability. In total, AC experts of 55 organizations and consultancies answered selected questions concerning the consideration of quality criteria in organizational practice and rated the appropriateness and practicability of these criteria. The majority of the selected criteria referred

to the dimensions examined in an AC. The results of the preliminary study are consistent with those of Kanning et al. (2007). Both studies show that the majority of criteria is considered in organizations. However, in certain domains such as the statistical evaluation of ACs, quality improvement is needed. For example, a large number of organizations included in the sample did not statistically evaluate their ACs at all. As expected, the criteria's appropriateness was rated higher than their practicability by organizations as well as consultancies.

In the next stage, AC experts from research as well as organizational practice rated the complete set of 70 quality criteria included in the BACDi-instrument. The total sample of study I consisted of 185 practitioners, who rated the appropriateness and the practicability of the criteria as well as 41 scientists, who rated the appropriateness and the consistency of the criteria. Furthermore, participants were encouraged to suggest improvements regarding the criteria's wording and to rate the completeness of the list of criteria. This procedure ensured that both scientists and practitioners equally participated in the wording and weighting of the quality criteria. The large number of participants in the study reflects widespread concern about AC quality among scientists and practitioners and reinforces the demand for an instrument that enables organizations to evaluate the quality of their ACs.

Results show that scientists and practitioners alike accepted most of the criteria with a mean of 5.79 (SD = .73) on a scale from 1 to 7 (Schermuly, Nachtwei, & Meissner, submitted). Significant differences

Figure 1. Stages of the BACDi-project



between subgroups were not identified, which might be due to the fact that more than 80% of scientists claim to have had practical experience with ACs during their career. According to these results, the gap between scientists and practitioners in the sample is not very large and practitioners seem to be informed about how to improve quality and predictive validity of ACs (Schermuly et al., submitted) thus an exchange of knowledge exists. Nevertheless, differences between the ratings of criteria could be determined as displayed in Table 1, which shows the average ratings of appropriateness, practicability and consistency by scientists and practitioners on a scale from 1 to 7.

Taking into account the AC experts' feedback of study I, the list of criteria was revised once again. In order to be included in the final list of quality criteria, the AC experts' ratings of appropriateness were of most

importance. Scientists and practitioners rated the appropriateness of quality criteria in study I on a scale from 1 (not reasonable at all) to 7 (very reasonable). Ratings were subsequently transferred into a percent scale (0-100%). Quality criteria yielding an appropriateness rating by scientists, practitioners or both groups lower than 50% (indicating medium or less appropriateness) were excluded from the final list of criteria. This was the case for three criteria. The remaining quality criteria were revised according to the AC experts' feedback of study I and transferred into questions, since questions are preferable to statements to explore precise circumstances (Bortz & Doering, 2006). Finally, 67 items were included in the BACDi-instrument. 35 items were assigned to the AC phase development, 14 items present the AC phase conduction and 18 items were classed within the AC phase

Table 1. Most and least accepted quality criteria (Schermuly, Nachtwei, & Meissner, submitted)

Most accepted quality criteria				
		Appropriateness ¹ M (SD)	Practicability ² M (SD)	Consistency ³ M (SD)
1	Requirement profile is created prior to AC construction	6.81 (.50)	6.38 (1.02)	6.10 (1.69)
2	Observers are familiarized with AC dimensions during observer training	6.78 (.55)	6.45 (1.04)	6.33 (1.42)
3	An equal amount of time is provided for all participants to prepare for and conduct AC exercises	6.74 (.81)	6.53 (.96)	6.23 (1.51)
Least accepted quality criteria				
		Appropriateness ¹ M(SD)	Practicability ² M(SD)	Consistency ³ M(SD)
1	Eye contact with participants is avoided by observers	2.91 (1.98)	3.00 (1.78)	3.20 (2.15)
2	Ratings of other participants (peer-ratings) are used in group exercises	3.76 (2.04)	4.32 (1.91)	4.03 (2.12)
3	No information about AC dimensions is received by participants	4.07 (2.37)	5.40 (1.85)	4.22 (2.30)

Annotations: ¹ averaged ratings of appropriateness by scientists and practitioners; ² averaged ratings of practicability by practitioners, ³ averaged ratings of consistency by scientists; M = arithmetic mean, SD = standard deviation.

postprocessing. Participants are expected to need approximately 40 minutes to complete the BACDi-instrument.

To establish objectivity of the BACDi-instrument, specific answering options were assigned to each item. According to Bortz and Doering (2006), neutral answering options such as 'I don't know' or 'no information available' should be avoided because the interpretation of results becomes more difficult as the number of participants selecting these answering option increases. Thus, each item requires participants to report specific information about the AC used in their organization.

Furthermore, an explicit rule of interpretation is assigned to each answering option. This rule defines the points a participant receives for selecting a particular answering option. A participant can maximally receive one point for each item.

Consequently, the overall raw score of the BACDi-instrument is 67. The sum of raw scores of each AC phase (development, conduction, postprocessing) and the overall raw score are transferred to a percent scale ranging from 0% to 100% in order to create standardized indices of AC quality. In other words, the BACDi-result contains a standardized index of AC quality for each AC phase as well as a standardized overall index referring to the quality of the entire AC, the BACDi overall quality index.

Moreover, a software for the BACDi-instrument was programmed and examined in several testruns focusing on the usability of data-entry as well as the correctness of the dataoutput in an excel file. Each item is presented on a single page as displayed in Figure 2, which presents items 2, 37 and 67 assigned to the AC phases development, conduction and postprocessing, re-

spectively. After having completed the BACDi-instrument, participants receive a feedback indicating their BACDi-result compared to the anonymous results of other organizations. In addition to this, scientists' as well as practitioners' ratings of the criteria's appropriateness and practicability are displayed. This manner of representing the BACDi-result enables HR professionals to derive improvement suggestions to enhance the predictive validity of their ACs. Moreover, ratings of appropriateness indicate the criteria implying the highest demand for action and ratings of practicability show HR professionals, which criteria are easily implemented.

Evaluation of BACDi-instrument

In the current stage of this project, the BACDi-instrument is applied for the first time in organizational practice and evaluated by analy-

zing its reliability and validity. For this reason, several practitioners responsible for development, conduction or postprocessing of ACs in their organization use the BACDi-instrument within study II. Results provide the first standardization of the BACDi-instrument allowing future participants to compare their results to the anonymous results of former participants. More specifically, results can be compared not only to the scientific standard, but also to results of organizations similar in area of business and number of employees. Furthermore, study II examines whether AC-experts consider the BACDi-feedback as useful and which indicators of job performance are registered in the participating organizations in order to prepare a further analysis of the relationship between predictive validities of ACs and BACDi-results in study III. In this way, the BACDi-instrument is further evaluated according to a more objective measure of predictive validity.

When evaluating an instrument, the following three indicators are most important: objectivity, reliability and validity (Bortz & Doering, 2006). Due to standardized instructions, clearly defined answeringoptions and the standardization of results, the objectivity of the conduction, analysis and evaluation of the BACDi-instrument can be considered as given. The reliability of an instrument refers to its accuracy in measuring a specific attribute (Buehner, 2006). Study II is designed to analyze the reliability of the BACDi-instrument by means of the agreement between two HR practitioners who independently evaluate the same AC using the BACDi-instrument. The extent of rater agreement indicates

Item 2: Was a requirement profile created prior to AC construction?



Konzeption
Seite 2 von 87

► Wurde vor der AC-Konstruktion ein Anforderungsprofil* erstellt?

Nein
 Ja

* fasst alle Anforderungsdimensionen zusammen und definiert, welche Fähigkeiten und Eigenschaften ein Teilnehmer in welchem Ausmaß besitzen sollte, um im AC und später im Beruf erfolgreich abzuschneiden

Item 37: Which written documents were available for observers and moderators?



Durchführung
Seite 37 von 87

► Die schriftlichen Unterlagen für Beobachter und Moderatoren beinhalten: (Mehrfachnennungen möglich)

Anforderungsdimensionen-Übungs-Beobachtungs-Matrix*
 alle Übungsunterlagen
 Teilnehmer- bzw. Mitarbeiterinstruktionen (Beobachter, Moderatoren)
 andere Unterlagen

* in Tabellenform werden Dimensionen, Übungen und Teilnehmerbeobachtungen angeordnet; zeigt u.a. in welcher Übung welche Anforderungsdimensionen beobachtet werden

whether the BACDi-instrument reliably measures AC quality.

Validity is the most important determinant of an instrument's quality and indicates whether an instrument measures the construct it is supposed to measure (Bortz & Doering, 2006) and describes the extent of adequacy of conclusions derived from an instrument's results (Kleinmann & Strauss, 1996). A distinction is drawn between content,

construct and criterion validity. Content validity is given if an instrument thoroughly measures the most important aspects of a specific construct. According to Moosbrugger and Kelava (2008), expert ratings are important when determining content validity. AC experts from research and organizational practice were involved throughout the whole development of the BACDi-instrument. They were asked to rate the

criteria's completeness regarding AC quality. Due to the consideration of the experts' suggestions, the content validity of the BACDi-instrument can be regarded as satisfying. Construct validity of an instrument is at hand, if hypotheses derived from the construct of interest can be verified based on an instrument's results. Rather than considering a single criterion, a net of hypotheses is developed dealing with the relationship between the construct and other manifest or latent variables (Bortz & Doering, 2006). As part of the construct validity, the convergent validity describes the accordance of an instrument's results with the results of a different instrument that measures the same or a similar construct (Moosbrugger & Kelava, 2008). Until now, a comparable instrument measuring AC quality is not available; therefore the convergent validity of the BACDi-instrument cannot be determined. Nevertheless, its construct validity is examined according to several previously determined hypotheses. Criterion validity indicates the extent to which the results of an instrument measuring a latent construct match the results of an instrument measuring a corresponding manifest attribute (Bortz & Doering, 2006). An important indicator of criterion validity is the predictive validity, which analyses the relationship between a predictor (latent construct) and a manifest criterion that is measured at a later point in time (Buehner, 2006), as mentioned above. The identification of an adequate criterion is often difficult; therefore an instrument should be evaluated by means of several criteria. In study II, the validity of the BACDi-instrument is examined according to as-

Item 67: How often is every exercise checked for its actuality and adapted if necessary?

Nachbereitung
Seite 67 von 87

► Wie oft wird jede Übung auf ihre Aktualität geprüft und gegebenenfalls neu angepasst? (falls Sie die Übungen nicht überprüfen, tragen Sie hier bitte eine 0 ein)

Alle Jahre (nur Zahlen)

sumptions derived from scientific literature and recommendations relating specifically to AC quality. This includes, for instance, the analysis of relationships between subjective ratings of ACs' predictive validities by AC experts and the BACDi-result or between financial and personnel investments into ACs and the BACDi-result.

Benefits for organizations using BACDi-instrument

The BACDi-instrument is the only available instrument including an encompassing collection of quality criteria to improve methodological AC quality. All criteria are based on scientific evidence or AC-experts' recommendations and were validated by more than 200 researchers and practitioners. By using the BACDi-software, organizations can easily receive a feedback about the methodological quality of their AC. This feedback contains a comparison with the scientific standard, which indicates the quality criteria needing improvement. Furthermore, a benchmark allows for a com-

parison between individual results and anonymous results of organizations of the same size and area of business. The normative sample referred to in this benchmark is updated automatically, as more and more organizations use the BACDi-instrument. In addition to that, ratings of AC-experts indicating the appropriateness and practicability of the quality criteria are presented in the feedback. Taken together, the information provided in the feedback enables HR professionals to derive improvement suggestions to enhance AC quality and supports them when determining the demand of action concerning each quality criterion. In order to stay competitive, organizations have to improve their employee selection according to the results included in the BACDi-feedback.

Further developments

The next stage of the BACDi-project will focus on internationalizing the normative sample as well as further evaluating the BACDi-instrument within a third study. In order

to compare results from organizations in German speaking countries (Germany, Austria, Switzerland), data will be collected in organizations in North America (USA and Canada), the countries with the longest AC tradition. Chinese organizations will also be studied, where ACs were introduced only recently. In addition to that, the relationship between a more objective measure of the predictive validity of an AC such as the relationship between an indicator of job performance and an applicant's AC result as well as the BACDi-result will be examined in order to increase the reliability of the predictive validity determination. To prepare study III, study II includes a list of possible indicators of job performance that was completed by participants. Participants are free to decide whether they want to provide information about indicators of job performance registered in their organization. If they share information, their e-mail-address is saved exclusively in this part of the survey. This allows for later contact with specific organizations registering the same indicators of job performance in order to collect comparable data.

Conclusion

The BACDi-project was created with the objective to support HR practitioners while improving the quality of AC development, conduction and postprocessing. AC experts from research and organizational practice were involved in the composition of a list of quality criteria that were included in the BACDi-instrument. After completing the instrument, participants receive feedback that allows for the comparison of their

AC quality not only referring to scientific standards but to the anonymous results of other organizations and ratings of AC experts. In a current study, the BACDi-instrument is applied and evaluated. Subsequent stages include a further evaluation of the predictive validity of the BACDi-instrument as well as data collection in North American and Chinese organizations. The participation of additional organizations as well as further research relating to its validity will enhance the establishment of the BACDi-instrument into the future.

References

- Aamodt, M. G. (2004). *Research in law enforcement selection*. Boca Raton, FL: BrownWalker.
- Arbeitskreis Assessment Center, e. V. (2004). Standards der Assessment-Center-Technik. Retrieved from <http://www.arbeitskreis-ac.de/index.php/projekte/standards.html>
- Bortz, J., & Doering, N. (2006). *Forschungsmethoden und Evaluation fuer Human- und Sozialwissenschaftler* (Vol. 4). Heidelberg: Springer.
- Buehner, M. (2006). *Einfuehrung in die Test- und Fragebogenkonstruktion* (Vol. 2). Muenchen: Pearson.
- DIN33430. (2002). DIN33430: Anforderungen an Verfahren und deren Einsatz bei berufsbezogenen Eignungsbeurteilungen. In U. P. Kanning (Ed.), *Standards der Personaldiagnostik* (pp. 569-592). Goettingen: Hogrefe.
- Gaugler, B. B., Rosenthal, D. B., Thornton III, G. C., & Bentson, C. (1987). Meta-Analysis of Assessment Center Validity. *Journal of Applied Psychology*, 72(3), 493-511.
- Guidelines, I. T. F. o. A. C. (2009). Guidelines and Ethical Considerations for Assessment Center Operations. *International Journal of Selection and Assessment*, 17(3), 243-253.
- Hoefl, S. (2009). Gutes AC, schlechtes AC? Ein kritischer Kommentar zu den wissenschaftlichen Qualitätskriterien für Assessment Center von Kanning et al. (2007). *Zeitschrift fuer Arbeits- und Organisationspsychologie*, 53(2), 74-77.
- Jansen, P. G. W., & Vinkenburg, C. J. (2006). Predicting management career success from assessment center data: A longitudinal study. *Journal of Vocational Behavior*, 68, 253-266.
- Kanning, U. P., Poettker, J., & Gelléri, P. (2007). Assessment Center-Praxis in deutschen Großunternehmen Ein Vergleich zwischen wissenschaftlichem Anspruch und Realitaet. *Zeitschrift fuer Arbeits- und Organisationspsychologie*, 51(4), 155-167.
- Kersting, M. (2009). Qualitaet im Assessment Center. *Zeitschrift fuer Arbeits- und Organisationspsychologie*, 53(2), 70-74.
- Klehe, U.-C. (2004). Choosing How to Choose: Institutional Pressures Affecting the Adoption of Personnel Selection Procedures. *International Journal of Selection and Assessment*, 12(4), 327-342.
- Kleinmann, M., & Strauss, B. (1996). Konstrukt- und Kriteriumsvaliditaet des Assessment Centers: Ein Spannungsfeld. In W. Sarges (Ed.), *Weiterentwicklungen der Assessment Center-Methode* (pp. 1-16). Goettingen: Hogrefe.
- Klink, D., & Hilke, R. (2006). Berufspsychologie und Berufseignungsdiagnostik. In K. Pawlik (Ed.), *Handbuch Psychologie*. Heidelberg: Springer.
- Koenig, C., Klehe, U., Berchtold, M., & Kleinmann, M. (2010). Reasons for being selective when choosing personnel selection procedures. *International Journal of Selection and Assessment*, 18(1), 17-27.
- Krause, D. E. (2009). Ob Kritik konstruktiv wird, entscheidet der Kritisierte - Kritische Reflexion der AC-Standards. *Zeitschrift fuer Arbeits- und Organisationspsychologie*, 53(2), 77-79.
- Moosbrugger, H., & Kelava, A. (2008). *Testtheorie und Fragebogenkonstruktion*. Heidelberg: Springer.
- Nachtwei, J., & Schermuly, C. (2009). Acht Mythen über Eignungstests. *Harvard Business Manager*.
- Nerdinger, F. W., Blickle, G., & Schaper, N. (2008). *Arbeits- und Organisationspsychologie*. Heidelberg: Springer.
- Obermann, C., Hoefl, S., & Janke, O. (2008). Was machen die DAX100-Unternehmen. *Die große AkAC-Studie*.
- Schermuly, C., & Nachtwei, J. (2010). Assessment-Center optimieren. *Harvard Business Manager*, 09, 16-17.
- Schermuly, C., Nachtwei, J., & Meissner, M. (submitted). Quality criteria for Assessment Center. *Practical Assessment, Research and Evaluation*.
- Schmidt, F. L., & Hunter, J. E. (1986). The Economic Impact of Job Selection Methods on Size, Productivity, and Payroll Costs of the Federal Work Force: an Empirically Based Demonstration. *Personnel Psychology*, 39, 1-29.
- Schmitt, N., Gooding, R. Z., Noe, R. A., & Kirsch, M. (1984). Metaanalyses of Validity Studies Published Between 1964 and 1982 and the Investigation of Study Characteristics. *Personnel Psychology*, 37(3), 407-422.
- Schuler, H., Hell, B., Trapmann, S., Schaar, H., & Boramir, I. (2007). Die Nutzung psychologischer Verfahren der externen Personalauswahl in deutschen Unternehmen. Ein Vergleich über 20 Jahre. *Zeitschrift fuer Personalpsychologie*, 6(2), 60-70.

Authors:

Franziska Schoelmerich
Humboldt University Berlin

Jens Nachtwei
Humboldt University Berlin

Carsten C. Schermuly
Technical University Braunschweig

To contact the BACDi project team: bacdi@lists.hu-berlin.de
For further information visit: www.bacdi.de