

SWISS 
assessment

**VORGABEN ZUR
EVALUATION**

Ausgangslage und Zielsetzung

Standard 9, Checkpoint 19

Der AC-Anbieter evaluiert seine AC in jeder Zertifizierungsperiode. Zur Evaluation kann ein Bericht hinsichtlich Kriteriumsvalidität (d.h. dem Zusammenhang der AC-Messungen mit relevanten Erfolgskriterien) oder Sozialer Validität (z.B. Akzeptanz des Verfahrens bei den Teilnehmenden) vorgelegt werden. Es muss aber innerhalb von zwei Zertifizierungsperioden (d.h. sechs Jahren; gezählt ab der Erstzertifizierung) mindestens ein Evaluationsbericht zur Kriteriumsvalidität vorgelegt werden. Der AC-Anbieter kann entscheiden, ob er über zwei Zertifizierungsperioden hinweg zwei Mal einen Bericht zur Kriteriumsvalidität verfasst oder je einen Bericht zur Sozialen Validität und einen Bericht zur Kriteriumsvalidität vorlegt. Wenn beispielsweise die Erstzertifizierung im Jahr 2020 stattgefunden hat, dann muss entweder bei der Erstzertifizierung im Jahr 2020 oder der Rezertifizierung im Jahr 2023 ein Bericht zur Kriteriumsvalidität vorgelegt werden; die nächsten beiden Zertifizierungsperioden umfassen dann die Rezertifizierung im Jahr 2026 sowie die Rezertifizierung im Jahr 2029 usw.

Soziale Validität

Die Soziale Validität beschreibt die Wahrnehmung eines diagnostischen Verfahrens aus Perspektive der Teilnehmenden. Nach Schuler (1990) umfasst die Soziale Validität vier Komponenten: Information (z.B. vorab kommunizierte Informationen, Möglichkeit zur Vorbereitung auf das AC), Partizipation (z.B. Möglichkeit Kontrolle über die Situation auszuüben), Transparenz (z.B. klare Beurteilungskriterien) und Urteilskommunikation (z.B. ausführliches, verständliches Feedback). Gilliland (1993) beschreibt darüber hinaus folgende Komponenten, welche die Fairnesswahrnehmung beeinflussen: (1) die Jobbezogenheit der verwendeten Verfahren, (2) die Möglichkeit eigene Fähigkeiten zu präsentieren, (3) die wahrgenommene interpersonelle Kommunikation (z.B. wertschätzende Kommunikation, Möglichkeit Rückfragen zu stellen) und (4) die Angemessenheit der verwendeten Fragen (z.B. keine persönlichen, diskriminierenden oder in anderer Weise für die Auswahl nicht relevanten oder unangemessenen Inhalte). Zur Evaluation der sozialen Validität muss die wahrgenommene Fairness mindestens aus Perspektive der AC-Teilnehmenden erfasst werden (siehe auch «Minimalstandards» in diesem Dokument).

Schuler, H. (1990). Personenauswahl aus der Sicht der Bewerber: Zum Erleben eignungsdiagnostischer Situationen. *Zeitschrift für Arbeits- und Organisationspsychologie*, 34(4), 184-191.

Gilliland, S. W. (1993). The perceived fairness of selection systems: An organizational justice perspective. *Academy of management review*, 18(4), 694-734.

Kriteriumsvalidität

Unter Kriteriumsvalidität wird allgemein die Korrelation zwischen der Leistung im Auswahlverfahren (Beurteilung im AC) und einem bestimmten Erfolgskriterium (z.B. berufliche Leistung, Beurteilung durch Vorgesetzte oder Unterstellte, Karriereentwicklung) verstanden. Das ausgewählte Erfolgskriterium muss für die zu besetzende(n) Stelle(n) oder Funktion(en) relevant sein.

Ziel

Mit dem regelmässigen Ausarbeiten von Evaluationsstudien wird eine Kultur der Qualitätsüberprüfung und eines langfristigen Qualitätsmanagements etabliert, die eine fortlaufende Verbesserung der Anwendung von Assessment Centern zur Folge hat. Die Evaluationsstudien sind ein sichtbares Zeichen gegenüber der Kundschaft, dass zertifizierte AC-Anbieter ihr Verfahren regelmässig einer systematischen Überprüfung unterziehen und daraus die notwendigen Massnahmen ableiten.

Minimalstandards

Soziale Validität

Befragt werden Direktbeteiligte der von der Unternehmung durchgeführten AC. In der Regel sind dies die AC-Teilnehmenden, je nach Fragestellungen kann die Studie mit den Sichtweisen der Assessor*innen und/oder Auftraggeber*innen/Kund*innen ergänzt werden. Damit die Ergebnisse eine bestimmte statistische Bedeutsamkeit haben, muss die Anzahl der befragten AC-Teilnehmenden mindestens 40 betragen. Bezüglich der Wahl des Erhebungsinstrumentes gibt es keine Vorgaben. Es müssen jedoch eine oder mehrere Komponenten der Sozialen Validität (z.B. Information, Partizipation/Kontrolle, Transparenz und Urteilkommunikation/Feedback, s.o.) erfasst sowie die Ergebnisse der Auswertung nachvollziehbar dargestellt und diskutiert werden (siehe dazu auch die Hinweise zur Erstellung des Berichts).

Kriteriumsvalidität

Angesichts der unterschiedlichen Tätigkeitsbereiche, Funktionen und Hierarchiestufen, auf welche die AC der Swiss Assessment-Mitglieder ausgerichtet sind, werden für die Erhebung der Kriteriumsvalidität Minimalstandards definiert. Diese sind unabdingbar, um im Rahmen des Zertifizierungsprozesses das Einhalten dieses Qualitätsstandards möglichst objektiv überprüfen zu können. Obwohl beim Festlegen dieser Vorgaben darauf geachtet wurde, dass eine solche Studie auch im Praxisumfeld machbar ist, dürfte vor allem bei der erstmaligen Durchführung dieser Evaluation für die betreffende Unternehmung ein bedeutsamer Mehraufwand entstehen.

Daten zur Kriteriumsvalidität

Aus dem Assessment Center werden die Bewertungen der einzelnen Dimensionen, der einzelnen Übungen und/oder ein daraus berechneter ungewichteter Summenwert oder Mittelwert (AC-Gesamtbeurteilung; Overall Assessment Rating, OAR) als Prädiktor zur Evaluation verwendet.

Das Aussenkriterium muss einen deutlich erkennbaren Bezug zu den AC-Beurteilungen aufweisen (z.B. berufliche Leistung generell oder bezogen auf spezifische Leistungsdimensionen) und/oder hinsichtlich des betreffenden Selektionsprozesses von Bedeutung sein (z.B. beruflicher Erfolg, Zufriedenheit der Vorgesetzten/Unterstellten, objektive Leistungsdaten).

Bezüglich der Datenerhebung gelten zudem folgende Vorgaben:

- Es können bereits vorliegende Daten verwendet (z.B. Leistungsbeurteilung, objektive Daten) oder gezielt welche erhoben werden (z.B. Befragung von Vorgesetzten der Kandidatin/des Kandidaten).
- Zur Berechnung der Kriteriumsvalidität müssen sowohl die Prädiktordaten als auch die Kriteriumsdaten letztlich in quantitativer Form vorliegen.
- Die Stichprobe muss mindestens 40 Kandidat*innen umfassen. Ist die Stichprobe für ein einzelnes AC kleiner, können ähnliche AC (z.B. mit vergleichbaren Leistungsdimensionen) zusammengefasst werden, damit auf diese Weise eine geschlossene Validierungsstudie mit der erforderlichen Stichprobengrösse vorgelegt werden kann.
- Der Zeitraum zwischen dem Stellenantritt der Kandidat*innen bzw. des Kandidaten/der Kandidatin und der Erhebung des Kriteriums muss zwischen einem halben Jahr und max. drei Jahren betragen.

- Personen, die am Arbeitsplatz Einschätzungen der Kandidatin/des Kandidaten vornehmen, müssen mindestens seit einem halben Jahr mit ihr/ihm zusammengearbeitet haben.

Bericht zur Sozialen Validität oder Kriteriumsvalidität

Die Berichte zur Sozialen Validität sowie zur Kriteriumsvalidität orientieren sich an derselben Struktur (Ausgangslage – Methode – Ergebnisse – Konsequenzen). Der schriftliche Validierungsbericht im Umfang von ca. 10 Seiten kann wie folgt strukturiert werden:

- *Ausgangslage*: Beschreibung des/der zu evaluierenden AC; kurze Übersicht zu den Übungen und Dimensionen; Einsatz als Selektions- oder Entwicklungsinstrument; Zielsetzungen/Erwartungen des Kunden an das AC
- *Methode*: Vorgehen bei der Validierungsstudie; Begründung des Vorgehens, d.h. Datenerhebung bzw. Aufarbeiten bereits vorliegender Daten, Instrument zur Datenerhebung, Beschreibung der Stichprobe, sowie bei Berichten zur Kriteriumsvalidität die Wahl und Definition des Kriteriums.
- *Ergebnisse*: Deskriptive Kennwerte; bei Berichten zur Kriteriumsvalidität: Korrelation(en) zwischen AC-Ergebnissen und Aussenkriterium resp. Aussenkriterien; bei allen Berichten: Interpretation und Diskussion der Ergebnisse (ggf. auch Vergleich zu vorangegangenen Evaluationsergebnissen)
- *Konsequenzen*: Ableiten konkreter Massnahmen für die Weiterentwicklung/Optimierung des/der betreffenden AC

Liegt aus dem für die Zertifizierung relevanten Zeitraum eine anerkannte wissenschaftliche Studie (z.B. Masterarbeit; Dissertation; Journalartikel) zu einem oder mehreren von der Unternehmung durchgeführten AC vor, so gilt die betreffende Arbeit als Bericht und es sind zusätzlich nur noch die Konsequenzen aufzuführen.

Empfehlungen

Soziale Validität

In der Regel kann man sich diesbezüglich auf unternehmensintern bestehende Evaluationen zur Kundenzufriedenheit stützen und das Vorgehen mit spezifischen Fragestellungen und einer detaillierteren Datenauswertung erweitern.

Wer ein breiter abgestütztes Erhebungsinstrument benutzen und damit die Gelegenheit zum Vergleich mit anderen Unternehmungen wahrnehmen möchte, dem sei der AKZEPT-AC Fragebogen von Kersting empfohlen. Der aktuelle Bogen kann bei Martin.Kersting@psychol.uni-giessen.de direkt angefordert werden. Mit diesem Vorgehen wird sichergestellt, dass man die jeweils aktuelle und passende Version erhält. Ausserdem bekommt man auch gleich eine SPSS-Auswertungssyntax. Das sichert die Auswertungsqualität und erleichtert eine allfällige Zusammenführung der Daten.

Kersting, M. (2010). Akzeptanz von Assessment Center: Was kommt an und worauf kommt es an? *Wirtschaftspsychologie*, 12, 58-65.

Kriteriumsvalidität

Bezüglich der Wahl des Aussenkriteriums gibt es zahlreiche Möglichkeiten. Die Verwendung bereits vorliegender Leistungsbeurteilungen erscheint naheliegend, diese weisen aber – in der Regel durch den Mildefehler bedingt – oft eine geringere Varianz auf. Es kann deshalb alternativ ratsam sein, speziell für die Evaluationsstudie Daten zu erheben. Da die betreffenden Beurteilungen anonym durchgeführt werden können und weder für die Beurteilenden noch die Beurteilten von Relevanz sind, fallen sie angemessener aus. Nachfolgend sind zur besseren Orientierung einige Beispiele aufgeführt, wie resp. womit die Daten erhoben werden könnten.

1:1-Vergleich auf der Basis des AC-Anforderungsprofils resp. der AC-Dimensionen

Erhebung am Arbeitsplatz der Kandidatin/des Kandidaten zu den im AC erhobenen Dimensionen (z.B. Planung und Organisation im AC erhoben und dann am Arbeitsplatz zwecks Validierung). Die Befragung der betreffenden Vorgesetzten und/oder Arbeitskolleg*innen und/oder Unterstellten könnte wie folgt gestaltet werden:

Stufen Sie das Verhalten des Mitarbeiters/der Mitarbeiterin zu den Stellenanforderungen hinsichtlich der nachfolgend aufgeführten Dimensionen ein.

Planung und Organisation

Ist im täglichen Ablauf gut organisiert. Nutzt Planungshilfen. Kann auch unter Unsicherheit entscheiden.

Gestaltungsvermögen

Erkennt sich verändernde Rahmenbedingungen und gestaltet diese aktiv. Versucht Hindernisse zu beeinflussen. Setzt seine Ideen rasch um. Handelt autonom.

Gewissenhaftigkeit

Arbeitet genau und berücksichtigt die Details. Handelt zuverlässig und verantwortungsbewusst.

etc.

	Erfüllt die Anforderungen deutlich nicht	Erfüllt die Anforderungen knapp nicht	Erfüllt die Anforderungen	Übertrifft die Anforderungen leicht	Übertrifft die Anforderungen deutlich	Nicht beurteilbar
Planung und Organisation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Gestaltungsvermögen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Gewissenhaftigkeit	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Anhand dieser Daten können die einzelnen Dimensionen sowie die berechneten Gesamtwerte korreliert werden.

Messung der aufgabenbezogenen Arbeitsleistung

Die Messung der aufgabenbezogenen beruflichen Leistung hat den Vorteil, dass sie für viele verschiedene Tätigkeiten mit unterschiedlichen Inhalten angewandt werden kann und etablierte Messinstrumente bestehen, die entsprechend reliabel sind.

Ein einfaches und breitflächig verwendbares Instrument zur Messung der Arbeitsleistung ist die Fragebogenskala ‚task-based performance‘ (Bott, Svyantek, Goodman & Bernal, 2003), deren deutsche Übersetzung (Jansen et al., 2013) hier aufgeführt ist.

Im Vergleich zu Arbeitskollegen bzw. Arbeitskolleginnen ...	Trifft überhaupt nicht zu	Trifft nicht zu	Trifft eher nicht zu	Weder noch	Trifft eher zu	Trifft zu	Trifft voll und ganz zu
... erreicht er/sie die Zielsetzungen bei seiner/ihrer Arbeit.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... zeigt er/sie Fachkenntnisse in allen arbeitsrelevanten Aufgaben.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... übernimmt er/sie mehr Verantwortung als für seine/ihre Arbeit erforderlich ist.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... ist er/sie in allen Arbeitsbereichen kompetent.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
... plant und organisiert er/sie das Vorgehen, um seine/ihre Arbeitsziele zu erreichen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Bott, J.P., Svyantek, D.J., Goodman, S.A., & Bernal, D.S. (2003). Expanding the performance domain: Who says nice guys finish last? *The International Journal of Organizational Analysis*, 11, 137-152 .

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.

Weitere Hinweise zur Kriteriumsauswahl

Generell ist darauf zu achten, dass das Kriterium einen erkennbaren Bezug zum Erfolg in der Tätigkeit aufweist. Hier bietet sich wie erwähnt die Messung der aufgabenbezogenen Leistung (allgemein oder dimensionsspezifisch) an.

Zudem können bereits vorliegende unternehmensinterne Daten verwendet werden, z.B. standardisierte Leistungsbeurteilungen, Beförderungen, Gehaltsentwicklung oder weitere objektive Daten wie Umsatz, Anzahl Beschwerden, Neukunden im Verkaufsbereich etc. Es sollte hierbei beachtet werden, dass objektive Daten auch von externen Faktoren (z.B. konjunkturbedingte Einflüsse) abhängen, die nur bedingt vom Kandidaten/von der Kandidatin beeinflusst werden können. Der Zusammenhang zwischen AC-Leistung und objektiven Daten kann deshalb anders ausfallen als der Zusammenhang zwischen AC-Leistung und subjektiven Daten (z.B. aufgabenbezogene Leistung).

Datenerhebung zur Sozialen Validität oder Kriteriumsvalidität

Als Erhebungsmethoden bieten sich Fragebogen, Online-Befragungen oder gegebenenfalls Interviews an. Zu empfehlen ist eine Online-Befragung, da heutzutage einfache Tools im Internet greifbar sind (z.B. SurveyMonkey, LimeSurvey, Questback, Google Formulare, Doodle, etc.), das Ganze einfach administriert werden kann (Versand von E-Mail mit entsprechendem Link) und in den Augen der Befragten die Anonymität gut gewährleistet ist. Zudem liegen die Daten in Echtzeit vor und können rasch ausgewertet werden. Die Möglichkeit von Interviews ist nicht auszuschliessen, allerdings muss man sich bewusst sein, dass die erhaltenen Informationen im Hinblick auf die Berechnung der Korrelation in quantitativer Form vorliegen müssen.

Weitere relevante Hinweise zur Erstellung einer Evaluationsstudie zur Kriteriumsvalidität finden sich im Foliensatz „**Evaluation von Assessment Centern – Ein Wegweiser zur Berechnung der Kriteriumsvalidität**“ von Dr. Pia Ingold. Dieses Dokument ist auf der Website von Swiss Assessment unter der Rubrik ‚Publikationen/Fachbeiträge‘ greifbar.