

SWISS   
assessment

**VORGABEN ZUR  
EVALUATION**

## Ausgangslage und Zielsetzung

### **Standard 9, Checkpoint 19**

Der AC-Anbieter evaluiert seine AC pro Zertifizierungsperiode hinsichtlich Sozialer Validität (z.B. Akzeptanz des Verfahrens bei den Teilnehmenden, Assessoren, weiteren Beteiligten; Kundenzufriedenheit; etc.) oder Kriteriumsvalidität. Damit diese Anforderung nicht zu starken Verzögerungen bei der Erstzertifizierung führt, kann zuerst ein Evaluationsbericht zur Sozialen Validität vorgelegt werden. Die Studie zur Kriteriumsvalidität wird in diesem Fall dann anlässlich der Rezertifizierung fällig. Danach geht es auf diese Weise alternierend weiter.

### **Soziale Validität**

Soziale Validität wird als Sammelbezeichnung dessen definiert, was die eignungsdiagnostische Situation zu einer für die Beteiligten akzeptablen Situation macht. Die vier hauptsächlichen Komponenten der Sozialen Validität sind Information, Partizipation/Kontrolle, Transparenz und Urteilkommunikation/Feedback.

### **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.

### **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 den Kunden, 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 Assessoren und/oder Auftraggeber/Kunden ergänzt werden. Damit die Ergebnisse eine bestimmte statistische Bedeutsamkeit haben, muss die Anzahl der befragten Personen mindestens 40 betragen. Bezüglich Wahl des Erhebungsinstruments gibt es keine Vorgaben. Es müssen eine oder mehrere Komponenten der Sozialen Validität erfasst und 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 aus den einzelnen Dimensionen sowie 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 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 des Kandidaten).
- Zur Berechnung der Kriteriumsvalidität müssen die Daten letztlich in quantitativer Form vorliegen.
- Die Stichprobe muss mindestens 40 Kandidaten/Kandidatinnen 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 Kandidaten/des Kandidaten 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 Kriteriumsvalidität**

Der schriftliche Validierungsbericht im Umfang von ca. 10 Seiten ist wie folgt zu strukturieren:

- 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. Wahl und Definition des Kriteriums, Datenerhebung bzw. Aufarbeiten bereits vorliegender Daten, Instrument zur Datenerhebung, Stichprobe
- Ergebnisse: Deskriptive Kennwerte; Korrelation(en) zwischen AC-Ergebnissen und Aussenkriterium resp. Aussenkriterien; Interpretation und Diskussion der Ergebnisse
- Konsequenzen: Ableiten konkreter Massnahmen für die Weiterentwicklung/Optimierung des/der betreffenden AC

Der **Bericht zur sozialen Validität** orientiert sich an derselben Struktur (Ausgangslage – Methode – Ergebnisse – Konsequenzen).

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 Unternehmung wahrnehmen möchte, dem sei der AKZEPT-AC Fragebogen von Kersting empfohlen. Der aktuelle Bogen kann bei [Martin.Kersting@psychol.uni-giessen.de](mailto: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 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 zur Validierung am Arbeitsplatz). 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 bezogen auf die Stellenanforderungen bezüglich 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 Leistungsbeurteilung, 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**

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), in den Augen der Befragten die Anonymität gut gewährleistet ist, die Daten in Echtzeit vorliegen und rasch ausgewertet werden können. 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.