

Wie wirkt sich die transparente Gestaltung von Assessment Centern auf die Konstruktvalidität aus?

Dipl.-Psych. Sybille Lerche

*„For all we do know about assessment centers, we don't know enough“
(Klimoski & Brickner, 1987, S. 256)*

Ausgangspunkt der Arbeit. Assessment Center gehören zu den häufigst verwendeten, aber auch zu den kostenintensivsten Personalauswahl und -entwicklungsverfahren. Auf ihren Urteilen gründen sich Personalentscheidungen von großer Tragweite. Für derart bedeutungsträchtige Verfahren ist eine hohe Qualität unerlässlich. Obwohl Assessment Center zahlreich genutzt werden und nachweislich prognostisch valide sind, sind sie nicht ohne Kritik geblieben: Der Aspekt, der selbst nach 50 Jahren Forschung noch ungeklärt ist, ist die Konstruktvalidität des Verfahrens. Was wird mit dem Assessment Center erfasst? Auf welcher Grundlage beruhen die Prognosen? Seit Sackett und Dreher (1982) vor nunmehr über 20 Jahren die Problematik der Konstruktvalidität angesprochen haben, ist das Puzzle der Konstruktvalidität (Klimoski & Brickner, 1987) noch immer nicht vollständig. Es wurden bisher zahlreiche Einflussfaktoren auf die Konstruktvalidität untersucht. Die Transparenz des Assessment Centers wird als ein Aspekt diskutiert (Kleinmann, 1997a; Kolk, Born & van der Flier, 2003), eine eindeutige Klärung steht jedoch noch aus. An diesem Punkt setzt die Arbeit an, um ein weiteres Puzzlestück beizutragen.

Anliegen der Arbeit. Es wird der Einfluss der transparenten Gestaltung eines Assessment Centers auf die Konstruktvalidität untersucht. Hierbei werden sowohl die konvergente als auch die diskriminante Validität analysiert. Des Weiteren versucht die Arbeit zu klären, ob das eingesetzte Assessment Center Leistungsmotivation und Intelligenz als Personenmerkmale erfasst.

Methodisches Vorgehen. Es wurden 11 Assessment Center mit 60 studentischen Teilnehmern wirtschaftswissenschaftlicher Studiengänge durchgeführt. Das Verfahren umfasste eine Selbstpräsentation, eine Gruppendiskussion, ein Rollenspiel und ein Interview. Den Teilnehmern wurde das AC kostenpflichtig als Bewerbungstraining angeboten, der jeweils beste Teilnehmer jedes ACs erhielt einen Buchgutschein. Experimentell unterschieden wurden die Bedingungen Transparenz und Intransparenz (jeweils $N = 30$). Zur Auswertung fanden hauptsächlich Korrelationen, Vergleiche der Korrelationskoeffizienten und die Multidimensionale Skalierung Anwendung.

Ergebnisse und Diskussion – konvergente und diskriminante Validität. Die Analyse der Multitrait-Multimethod Matrix (MTMM-Matrix) ergab hinsichtlich der konvergenten und diskriminanten Validität keine statistisch bedeutsamen Unterschiede zwischen den experimentellen Bedingungen Intransparenz und Transparenz. Werden die Maße der konver-

genten Validität genauer betrachtet, so ergibt sich kein konsistentes Bild über eine Minderung oder Erhöhung der Zusammenhänge gleicher Anforderungen in unterschiedlichen Übungen auf Grundlage der variierten Bedingungen. Das Wissen der Teilnehmer um die Anforderungen hat demnach in dieser Arbeit keinen systematischen Einfluss auf die konvergente Validität. Dies widerspricht den Arbeiten von Kleinmann et al. (1996) und Kleinmann (1997b). Dass keine bedeutsamen Unterschiede resultierten, kann auch an der geringeren Stichprobengröße von $N = 30$ je Bedingung liegen. Nach Cohens (1977) „sample size“-Tafeln könnten einige dieser Unterschiede spätestens bei $N = 40$ signifikant werden. Die Betrachtung der Maße für die diskriminante Validität zeigt hingegen in 9 von 12 Fällen eine Verbesserung. Das heißt, die Zusammenhänge zwischen den unterschiedlichen Anforderungen einer Übung wurden durch die transparente Gestaltung des Assessment Centers geringer. Es ist davon auszugehen, dass bei einer größeren Teilnehmerzahl pro Bedingung weitere Unterschiede, die kleinere Effektstärken aufwiesen, statistisch bedeutsam werden. Dieser Trend einer Verbesserung zeigte sich über alle Übungen hinweg (gemittelte Übungskoeffizienten). Eine Ausnahme hierbei war die Gruppendiskussion. In Hinsicht auf diese Übung konnte keine Verbesserung der gemittelten diskriminanten Validität festgestellt werden. Die anderen drei Übungen – Selbstpräsentation, Rollenspiel und Interview – zeigten Effektstärken hinsichtlich des Vergleiches zwischen Intransparenz und Transparenz in einer Größenordnung von .21 bis .33. Die Teilnehmer verhalten sich demnach spezifischer hinsichtlich der Anforderungen, und die Beobachter können in der Bedingung Transparenz differenziertere Urteile auf den Anforderungen abgeben. Dies bestätigt die Befunde von Kolk et al. (2003). Ebenso bestätigen die vorliegenden Ergebnisse die Schlussfolgerung von Kolk et al., dass die Transparenz einen relativ kleinen Effekt auf Teilnehmer und ihre Leistung in Gruppenübungen hat, hingegen bei Individualübungen einen größeren Effekt aufweist. Bei einer Mittelung der Maße der diskriminanten Validität pro Übung war die Gruppendiskussion die einzige Übung, in der sich keine Minderung der Zusammenhänge zwischen den Anforderungen einer Übung ergab. Dies bestätigt auch die Arbeiten von Kleinmann et al. (1996) und Kleinmann (1997b), die bei Verwendung von Gruppenübungen keinen Einfluss auf die diskriminante Validität fanden.

Die Multidimensionale Skalierung bestätigt die Ergebnisse der MTMM-Matrix. Es gab keine Verbesserung hinsichtlich der konvergenten Validität, die Verbesserung der diskriminanten Validität ist jedoch deutlich zu erkennen. In der Transparenzbedingung zeigte sich ein deutlich differenzierteres Bild für die einzelnen Anforderungen. Es zeigten sich dennoch in beiden Bedingungen Übungseffekte. Interessant ist, dass die Übungseffekte in der Transparenzbedingung wesentlich strukturierter sind und zum Teil deutlicher voneinander abgegrenzt. Was erklärt nun die Unterschiede zwischen den Bedingungen? Warum zeigten sich die Übungseffekte in der Bedingung Transparenz deutlicher und abgegrenzter, obwohl sie theoretisch durch die transparente Gestaltung des Assessment Centers zugun-

ten der Anforderungseffekte vermindert werden sollten? Der klassische Erklärungsansatz trifft hier offensichtlich nicht zu. Die Teilnehmer würden dabei bei transparenter Gestaltung des Assessment Centers hinsichtlich der Anforderungen konsistenter Auftreten, dies führe zu einer verbesserten konvergenten und diskriminanten Validität und zugleich zu deutlicheren Anforderungseffekten. Dies trifft hier nicht zu. Daher müssen die alternativen Konzepte zur Konstruktvalidierung herangezogen werden. Es ist zu vermuten, dass bei Kenntnis der Anforderungen die Teilnehmer ihr Verhalten an die Anforderung *in der jeweiligen Übung* anpassen. Jede Übung hat letztlich ihre eigenen Anforderungsspezifika, wie auch jede Situation in der Realität. Wie Obermann (2002) bereits anbringt, trifft das klassische Konzept der Konstruktvalidierung, in welchem Übungseffekte als Fehler betrachtet werden, vermutlich nicht zu. Die Ergebnisse der vorliegenden Arbeit unterstützen die Annahme, dass das Assessment Center anstelle von stabilen Personenmerkmalen berufliche Situationen und Rollen erfasst (Lance et al., 2000). Die transparente Gestaltung des Assessment Centers fördert demnach die Erfassung von Teilnehmerverhalten in spezifischen Situationen. Die Beobachter geben differenzierte Beurteilungen für jede Anforderung in jeder Übung ab.

Leistungsmotivation und Intelligenz. Für eine ansatzweise umfassende Analyse des Assessment Centers wurden Zusammenhänge mit Leistungsmotivation und allgemeiner Intelligenz untersucht. Das Leistungsmotivationsinventar (LMI, Schuler & Prochaska, 2001) wurde zur Erhebung der Leistungsmotivation der Teilnehmer eingesetzt und mit den Beobachterurteilen in Zusammenhang gebracht. Es zeigte sich, dass die Anforderungen Zielorientierung/Leistungsmotivation und Auftreten des Assessment Centers in Zusammenhang mit den LMI-Dimensionen standen, diese beiden Anforderungen also tatsächlich erfasst wurden. Mit diesen Ergebnissen konnten die Arbeiten von Scholz und Schuler (1993) sowie Kelbetz und Schuler (2002) hinsichtlich der Leistungsmotivation bestätigt werden. Bei einer Unterscheidung der experimentellen Bedingungen ist auffällig, dass in der Bedingung Intransparenz deutlich mehr Zusammenhänge zwischen den Anforderungen und dem LMI vorzufinden waren. In der Transparenzbedingung zeigten sich deutlich weniger Zusammenhänge. Dies kann unter Umständen erklärt werden durch einen Haloeffekt der Anforderung Zielorientierung/Leistungsmotivation in der Bedingung Intransparenz. Die Teilnehmer haben die Anforderungen möglicherweise nicht erkannt und wurden entsprechend dann besser bewertet, wenn sie mit einer hohen Leistungsmotivation an die Aufgaben herangingen. Die transparente Gestaltung des Assessment Centers schaltet den Haloeffekt nahezu aus. Eine differenziertere Erfassung der Anforderungen ist daher möglich. Als Indikator für die Allgemeine Intelligenz diente in der vorliegenden Arbeit der Abiturdurchschnitt der Teilnehmer. Das Assessment Center wies kleinere bis mittlere Zusammenhänge mit dem Abiturdurchschnitt der Teilnehmer auf. Zum Großteil sind die Zusammenhänge zwischen dem Abiturdurchschnitt als Indikator für Intelligenz und dem Assessment Center als klein einzustufen ($r < .30$). Werden die Ergeb-

nisse getrennt nach Bedingungen betrachtet, so zeigen die Ergebnisse, dass es hier wie auch schon beim Leistungsmotivationsinventar (LMI, Schuler & Prochaska, 2001) Unterschiede in der Richtung der Zusammenhänge gibt. Die Zusammenhänge zeigten sich in der Transparenzbedingung klarer und sinnvoller, eine höhere Intelligenz geht mit einem besseren Abschneiden im Assessment Center einher. Es lässt sich vermuten, dass bei Kenntnis der Anforderungen die Teilnehmer sich darauf konzentrieren, anforderungsrelevantes Verhalten einzusetzen und dabei ihre vorhandenen Kompetenzen wie die Intelligenz optimal nutzen können. Sofern sie die Anforderungen nicht wissen, müssen sie sich als erstes entscheiden hinsichtlich welcher Anforderungen sie sich verhalten. Wenn sie auf eine nicht erfasste Anforderung setzten, kann ihr Potential nicht richtig zum Ausdruck kommen.

Zusammenfassung. Das Gesamtbild der Ergebnisse weist in eine Richtung, die sich von dem klassischen Verständnis der Konstruktvalidität entfernt. Das Assessment Center erfasst Übungen anstelle von Anforderungen. Die vorliegende Arbeit unterstützt damit die Annahmen von Obermann (2002) und Lance et al. (2000). Die Transparenz, als untersuchter Einflussfaktor auf die Konstruktvalidität, verringert die Korrelationen der Anforderungen innerhalb einer Übung, d. h. erhöht die diskriminante Validität innerhalb der MTMM. Die Multidimensionale Skalierung lässt deutlich erkennen, dass bei einer transparenten Gestaltung des Assessment Centers die Übungseffekte ein klareres Muster der Aufgabentypen aufweisen.

Hinsichtlich des LMI ergaben sich für die Transparenzbedingung deutlich weniger Zusammenhänge, was darauf hinweist, dass die Teilnehmer anforderungsspezifischer bewertet wurden. Für den Abiturdurchschnitt zeigte sich, dass die Zusammenhänge in der Transparenzbedingung sinnvoller und logischer waren, zudem erwiesen sich die Korrelationen für die Gesamtwerte Anforderung und Übung als höher im Vergleich zur Intransparenzbedingung. Transparenz führte somit zu einem besseren Zusammenhang der Assessment Center-Ergebnisse mit dem Abiturdurchschnitt.

Dies bedeutet die Beobachter urteilen nicht global über die einzelnen Übungen hinweg, sondern wiesen in der Transparenzbedingung eine wesentlich bessere Differenzierung für die einzelnen Anforderungen in den einzelnen Übungen auf. Das Assessment Center erfasst die Anforderungen übungsspezifisch nicht aber übungsübergreifend.

Die Übungseffekte stehen, wie Lance et al. (2000) anführten, für die Situationspezifität des Verhaltens und definieren unterschiedliche Aufgaben und Rollen des beruflichen Lebens. Um diese Annahme zu rechtfertigen, fanden sie einen Zusammenhang von Übungseffekten mit externen Variablen wie der Intelligenz und dem Berufserfolg. Hinsichtlich der Intelligenz ließ sich auch in dieser Studie ein höherer Zusammenhang ($q = -.11$) finden zwischen dem Gesamtwert der Übungen und dem Abiturdurchschnitt im Vergleich zum Gesamtwert der Anforderungen. Wenn nun also davon ausgegangen wird, dass das

Assessment Center Übungen statt Anforderungen erfasst, dann führte die transparente Gestaltung des Assessment Centers zu einer Verbesserung der Qualität des Verfahrens.

Da bisher zu dieser Sichtweise vergleichsweise wenige Befunde vorliegen, ist eine Replikation der vorhandenen Ergebnisse sinnvoll. Sollten sich die Ergebnisse weiterhin bestätigen ist das Konzept der Konstruktvalidität neu zu formulieren. Wie sind Assessment Center in Zukunft zu gestalten und noch wichtiger, auf welcher Grundlage sind die Teilnehmer der Assessment Center zu bewerten? Die Ergebnisse dieser Arbeit unterstützen die Forderung nach einem transparenten Vorgehen sowohl im ethischen Sinne als auch zur Sicherung der Qualität des Verfahrens. Letztlich sollte darüber nachgedacht werden, die Bewertung und Urteilsfindung weniger anforderungs- sondern vielmehr situationsspezifisch vorzunehmen.

Literatur

- Brocke, B. (1995). Intelligenz: Struktur und Prozeß. In W. Sarges, *Management Diagnostik* (S. 225-240). Göttingen: Hogrefe.
- Höft, S. & Bolz, C. (2004). Zwei Seiten derselben Medaille? Empirische Überlappungen zwischen Persönlichkeitseigenschaften und Assessment Center-Anforderungsdimensionen. *Zeitschrift für Personalpsychologie*, 3, 6-23.
- Kelbetz, G. & Schuler, H. (2002). Verbessert Vorerfahrung die Leistung im Assessment Center? *Zeitschrift für Personalpsychologie*, 1, 4-18.
- Kleinmann, M. (1997a). *Assessment Center. Stand der Forschung – Konsequenzen für die Praxis*. Göttingen: Verlag für Angewandte Psychologie.
- Kleinmann, M. (1997b). Transparenz der Anforderungsdimensionen: Ein Moderator der Konstrukt- und Kriteriumsvalidität des Assessment-Centers. *Zeitschrift für Arbeits- und Organisationspsychologie*, 41, 171-181.
- Kleinmann, M., Kuptsch, C. & Köller, O. (1996). Transparency: A Necessary Requirement for the Construct Validity of Assessment Centres. *Applied Psychology: In International Review*, 45, 67-84.
- Klimoski, R. & Brickner, M. (1987). Why do Assessment Centers work? The Puzzle of Assessment Center Validity. *Personnel Psychology*, 40, 243-260.
- Kolk, N. J., Born, M. & Flier, H. van der (2003). The Transparent Assessment Center: The Effekt of Revealing Dimensions to Candidates. *Applied Psychology: An International Review*, 52, 648-668.
- Lance, C. E., Newbolt, W. H., Gatewood, R. D., Foster, M. R., French, N. R. & Smith, D. E. (2000). Assessment Center Exercise Factors Represent Cross-Situational Specificity, Not Method Bias. *Human Performance*, 13, 323-353.
- Obermann, C. (2002). *Assessment Center. Entwicklung, Durchführung, Trends*. Wiesbaden: Gabler.
- Sackett, P. R. & Dreher, G. F. (1982). Construct and Assessment Center Dimensions: Some Troubling Empirical Findings. *Journal of Applied Psychology*, 67, 401-410.
- Scholz, G. & Schuler, H. (1993). Das nomologische Netzwerk des Assessment Centers: eine Metaanalyse. *Zeitschrift für Arbeits- und Organisationspsychologie*, 37, 73-85.
- Schuler, H. & Prochaska, M. (2001). *LMI Leistungsmotivationsinventar: Dimensionen berufsbezogener Leistungsorientierung*. Göttingen: Hogrefe.