

Assessment Center-Mythen entlarvt: Wie Sie häufige Designfehler vermeiden und wissenschaftliche Erkenntnisse besser nutzen

Von Kompetenz-Dschungeln und Dimensions-Fallen - Ein Praxis-Leitfaden für valide und effiziente Assessment Center auf Basis aktueller Forschung

Dr. Walter Lieberei

01.06.2025

Die hartnäckige Kluft zwischen Wissenschaft und HR-Praxis

Im Personalwesen (HR) klafft oft eine bemerkenswerte Lücke zwischen gesicherten wissenschaftlichen Erkenntnissen und der gelebten betrieblichen Praxis. Dieses Phänomen, bekannt als "Wissenschafts-Praxis-Lücke" oder "Science-Practice Gap", führt dazu, dass in Unternehmen Methoden und Instrumente eingesetzt werden, deren Wirksamkeit und Validität von der Forschung längst in Frage gestellt oder sogar widerlegt wurden. Die Gründe hierfür sind vielfältig: Mangelnder Transfer von Forschungsergebnissen in die Praxis, Präferenz für "Best Practices" ohne kritische Prüfung, Zeitdruck oder schlichtweg fehlendes Wissen über aktuelle Studien.

Hell, Päßler und Nido (2023) beschreiben im Fall der Eignungsdiagnostik treffend, wie wissenschaftliche Erkenntnisse der HR-Forschung oft nicht oder nur unzureichend in der betrieblichen Praxis angewendet werden. Diese Kluft manifestiert sich in verschiedenen Dimensionen: Einerseits kommen wissenschaftliche Befunde nicht in der Praxis an, andererseits werden praktische Herausforderungen nicht ausreichend von der Wissenschaft beforscht. Das Problem betrifft alle Facetten im Personalbereich, ist international und führt dazu, dass HR-Professionals in kritischen Bereichen wie Personalauswahl, Leistungsmanagement und Kompensation häufig suboptimale Praktiken anwenden, die dem aktuellen Forschungsstand widersprechen.

Typische Beispiele für diese Diskrepanz in der HR-Praxis sind:

- **Personalauswahl:** Unstrukturierte Interviews, deren geringe prognostische Validität seit Jahrzehnten bekannt ist, werden weiterhin präferiert. Grafologische Gutachten oder nicht validierte Persönlichkeitstests finden ebenfalls noch Anwendung.
- **Leistungsmanagement:** Traditionelle, jährliche Mitarbeitergespräche mit starren Zielvorgaben und numerischen Bewertungen werden beibehalten, obwohl agile und kontinuierliche Feedbacksysteme oft bessere Ergebnisse liefern.
- **Vergütung:** Komplexe Bonussysteme, deren motivierende Wirkung oft überschätzt wird und die sogar kontraproduktiv sein können, sind weit verbreitet.
- **Training & Entwicklung:** Standardisierte Trainings ohne Bezug zu individuellen Bedürfnissen oder Transferüberlegungen ("Gießkannenprinzip").

Ein besonders prägnantes Beispiel für diesen Gap findet sich im Bereich der Assessment Center (AC). Obwohl ACs als Goldstandard der Eignungsdiagnostik gelten, werden sie in der

Praxis oft auf eine Weise konzipiert und durchgeführt, die grundlegende wissenschaftliche Erkenntnisse der letzten Jahrzehnte ignoriert. Dieser Artikel widmet sich zwei zentralen "Mythen" oder Praxisirrtümern im AC-Design und zeigt auf, wie eine wissenschaftlich fundierte Herangehensweise zu valideren und effizienteren Ergebnissen führen kann.

Mythos 1: Der Kompetenz-Dschungel - Weniger ist oft mehr

Ein häufig anzutreffendes Phänomen in der AC-Praxis ist die willkürliche und oft unsystematische Zusammenstellung einer Vielzahl von Eignungsmerkmalen, den sogenannten Kompetenzen oder Dimensionen. Nicht selten umfassen Anforderungsprofile 10, 15 oder sogar mehr Kompetenzen, die im AC beobachtet und bewertet werden sollen. Diese reichen von Klassikern wie "Kommunikationsfähigkeit" und "Teamfähigkeit" bis hin zu spezifischeren oder modischen Begriffen wie "Disruptives Denken" oder "Agile Mindset".

Die Intention dahinter ist oft nachvollziehbar: Man möchte ein möglichst umfassendes Bild der Kandidaten erhalten. Die Praxis zeigt jedoch, dass eine solche Fülle an Dimensionen mehrere Probleme mit sich bringt:

1. **Kognitive Überlastung der Beobachter:** Assessoren können nur eine begrenzte Anzahl von Dimensionen gleichzeitig reliabel beobachten und bewerten. Eine zu hohe Anzahl führt zu oberflächlichen Beobachtungen und ungenauen Urteilen.
2. **Mangelnde Trennschärfe:** Viele der definierten Kompetenzen überlappen inhaltlich stark (z.B. "Durchsetzungsvermögen" und "Konfliktfähigkeit") oder sind schwer voneinander abzugrenzen. Dies führt zu redundanten Bewertungen und einem "Halo-Effekt", bei dem die Bewertung einer Dimension die anderen beeinflusst.
3. **Aufwand und Komplexität:** Die Entwicklung von Übungen, Beobachtertrainings und Auswertungsverfahren für eine große Anzahl von Dimensionen ist extrem aufwendig und kostenintensiv.
4. **Fragliche Konstruktvalidität:** Wie wir später noch sehen werden, ist es ohnehin fraglich, ob ACs tatsächlich all diese fein zisierten Dimensionen trennscharf messen können.

Die wissenschaftliche Forschung liefert hierzu klare Hinweise. Eine Meta-Analyse von Meriac, Hoffman & Woehr (2014) zur Struktur von Assessment Center-Dimensionen zeigt eindrücklich, dass bereits eine deutlich reduzierte Anzahl von übergeordneten Faktoren einen Großteil der Varianz in den AC-Bewertungen erklären kann. Die Autoren untersuchten verschiedene Modelle und fanden, dass selbst eine **Sechs-Faktoren-Lösung**, basierend auf der Arbeit von Arthur et al. (2003), in der Lage ist, eine Vielzahl der sonst üblichen, spezifischeren Kompetenzen zu integrieren und abzudecken. (Hinweis: Meriac et al. (2014) kommen in ihrer eigenen Meta-Analyse letztlich zu einer noch stärker aggregierten 3-Faktoren-Lösung: "Administrative Skills", "Relational Skills" und "Drive". Für die praktische Anwendbarkeit und die Integration vieler gängiger Kompetenzen ist jedoch die hier dargestellte 6-Faktoren-Struktur, wie sie auch in Meriac et al. (2014, Figure 1, S. 1275) als Basis für ihre Analyse diente, oft ein guter Ausgangspunkt).

Tabelle 1: Das 6-Faktoren-Modell der Assessment Center-Dimensionen (basierend auf Arthur et al., 2003, wie in Meriac et al., 2014, Figure 1 dargestellt)

Übergeordneter Faktor	Definition (Exemplarisch)	Integrierte Kompetenzen (Beispiele)
1. Problemlösung	Fähigkeit, komplexe Situationen zu analysieren, Probleme zu identifizieren, relevante Informationen zu sammeln und zu bewerten sowie effektive und kreative Lösungen zu entwickeln und umzusetzen.	Analytische Fähigkeiten, Urteilsvermögen, Entscheidungsfähigkeit, Kreativität, Innovatives Denken, Konzeptionelles Denken, Strategisches Denken, Lernfähigkeit.
2. Planung & Organisation	Fähigkeit, Aufgaben und Projekte systematisch zu strukturieren, Prioritäten zu setzen, Ressourcen effizient einzusetzen, Zeitpläne zu erstellen und einzuhalten sowie vorausschauend zu agieren.	Organisationsfähigkeit, Planungsfähigkeit, Zeitmanagement, Selbstmanagement, Sorgfalt, Ergebnisorientierung, Prozessorientierung, Delegationsfähigkeit (teilweise).
3. Antrieb (Drive)	Innere Motivation, Engagement und Ausdauer, um Ziele zu erreichen, Herausforderungen anzunehmen, Initiative zu zeigen und hohe Leistungsstandards zu setzen und zu halten.	Leistungsbereitschaft, Eigeninitiative, Ergebnisorientierung, Beharrlichkeit, Engagement, Belastbarkeit (teilweise), Ehrgeiz, Zielorientierung.
4. Kommunikation	Fähigkeit, Informationen mündlich und schriftlich klar, präzise, verständlich und adressatengerecht zu vermitteln sowie aktiv zuzuhören und effektiv auf Fragen und Argumente einzugehen.	Mündliche Kommunikationsfähigkeit, Schriftliche Kommunikationsfähigkeit, Präsentationsfähigkeit, Überzeugungskraft (teilweise), Aktives Zuhören, Argumentationsfähigkeit.
5. Rücksichtnahme/ Bewusstsein für Andere	Fähigkeit, die Bedürfnisse, Perspektiven und Gefühle anderer Personen wahrzunehmen, zu verstehen und angemessen darauf zu reagieren sowie konstruktive Beziehungen aufzubauen und zu pflegen.	Empathie, Soziale Kompetenz, Teamfähigkeit, Interpersonelle Sensibilität, Kooperationsbereitschaft, Kundenorientierung (teilweise), Konfliktfähigkeit (teilweise), Wertschätzung.
6. Beeinflussung Anderer	Fähigkeit, andere von eigenen Ideen und Zielen zu überzeugen, sie zu motivieren und für gemeinsame Aufgaben zu gewinnen, sowie Führung zu übernehmen und richtungweisend zu agieren.	Überzeugungskraft, Verhandlungsgeschick, Durchsetzungsvermögen, Führungskompetenz, Motivationsfähigkeit, Inspirationsfähigkeit, Konfliktfähigkeit (teilweise), Veränderungsbereitschaft initiieren.

Übergeordneter Faktor	Definition (Exemplarisch)	Integrierte Kompetenzen (Beispiele)
(Optional: 7. Stresstoleranz)	Fähigkeit, auch unter Druck, bei Unsicherheit oder in komplexen Situationen handlungsfähig und leistungsstark zu bleiben und Emotionen angemessen zu regulieren.	Belastbarkeit, Frustrationstoleranz, Resilienz, Umgang mit Ambiguität, emotionale Stabilität. (Arthur et al. (2003) identifizierten diesen Faktor ursprünglich; Meriac et al. (2014) ließen ihn in ihrer 6-Faktoren-Darstellung oft weg).

Wie diese Erkenntnis die AC-Praxis optimieren kann:

Die Reduktion auf ein Set von beispielsweise sechs robusten, übergeordneten Dimensionen hat weitreichende positive Konsequenzen für die AC-Praxis:

- 1. Fokussierung und Effizienz:** Unternehmen können sich auf die wirklich erfolgskritischen übergeordneten Fähigkeiten konzentrieren. Dies vereinfacht die Konzeption der Übungen, da nicht mehr jede Nuance einer langen Kompetenzliste abgebildet werden muss.
- 2. Verbesserte Beobachterreliabilität:** Assessoren müssen weniger Dimensionen gleichzeitig im Blick haben, was ihre kognitive Belastung reduziert und zu genaueren, konsistenteren Bewertungen führt. Das Training der Beobachter kann intensiver auf diese Kernfaktoren ausgerichtet werden.
- 3. Erhöhte Trennschärfe:** Die übergeordneten Faktoren sind per Definition breiter angelegt und weisen geringere inhaltliche Überschneidungen auf als viele feingranulare Kompetenzen.
- 4. Bessere Vergleichbarkeit:** Die Verwendung eines standardisierten, wissenschaftlich fundierten Rahmenwerks erleichtert die Vergleichbarkeit von AC-Ergebnissen über verschiedene Abteilungen, Hierarchieebenen oder sogar Unternehmen hinweg (bei entsprechender Kalibrierung).
- 5. Reduzierter Aufwand:** Weniger Dimensionen bedeuten weniger aufwendige Definitionsprozesse, weniger komplexe Beobachtungsbögen und potenziell schlankere Auswertungs- und Feedbackprozesse.
- 6. Fundierte Personalentwicklung:** Feedback und Entwicklungsmaßnahmen können sich auf grundlegende Fähigkeitsbereiche konzentrieren, was oft nachhaltiger ist als die Arbeit an sehr spezifischen Einzelkompetenzen, deren Transferwert begrenzt sein kann.

Die Umstellung erfordert ein Umdenken: Weg von der Illusion der vollständigen Erfassung aller denkbaren Einzelaspekte, hin zu einer Konzentration auf die wissenschaftlich belegten Kernfaktoren erfolgreichen Verhaltens. Dies bedeutet nicht, dass spezifische Verhaltensanker unwichtig werden – im Gegenteil. Diese Verhaltensanker sollten dann aber diesen übergeordneten Faktoren zugeordnet werden und als Indikatoren für deren Ausprägung dienen.

Mythos 2: Die Dimensions-Falle - Was messen wir wirklich im AC?

Ein weiterer hartnäckiger Irrtum in der AC-Praxis betrifft die Interpretation der Ergebnisse. Traditionell werden die Leistungen der Kandidaten in den einzelnen Übungen anhand der vordefinierten Dimensionen bewertet. Am Ende des ACs wird dann versucht, für jede Dimension eine übergreifende Einschätzung über alle Übungen hinweg zu bilden. Die aggregierten Ergebnisse werden oft in einer Matrix dargestellt: Dimensionen in den Zeilen, Übungen in den Spalten (oder umgekehrt). Die Hoffnung ist, dass sich konsistente Muster für die einzelnen Dimensionen *über die Übungen hinweg* zeigen.

Bereits 1982 zeigten jedoch Sackett & Dreher in einer bahnbrechenden Studie (die in der AC-Literatur immer wieder zitiert wird, z. B. in Jackson, Lance & Hoffman, 2012, S. 68-69, oder Meriac et al., 2014, S. 1271), dass dieser Ansatz problematisch ist. Ihre Analysen und zahlreiche Folgestudien belegen immer wieder das sogenannte **“Exercise Effect”**:

- Die Korrelationen der Bewertungen *derselben Dimension (Traits oder Fähigkeiten)* über *verschiedene Übungen hinweg* sind oft niedrig. Das heißt, ein Kandidat, der in Übung A als “sehr kommunikationsstark” bewertet wird, muss nicht zwangsläufig auch in Übung B als “sehr kommunikationsstark” eingeschätzt werden.
- Im Gegensatz dazu sind die Korrelationen der Bewertungen *verschiedener Dimensionen innerhalb derselben Übung* oft hoch. Das heißt, wenn ein Kandidat in Übung A gut abschneidet, wird er tendenziell in vielen Dimensionen (Traits oder Fähigkeiten) für diese Übung gut bewertet.

Dieses Muster legt nahe, dass die Übungen selbst einen stärkeren Einfluss auf die Bewertung haben als die postulierten, übergreifenden Dimensionen (Traits oder latenten Fähigkeiten). Die Varianz in den AC-Ratings wird also stärker durch die spezifische Übung erklärt als durch die angenommene stabile Ausprägung einer Dimension oder Fähigkeit bei einem Kandidaten.

Illustratives Beispiel (Fiktive Ergebnismatrix):

Tabelle 2: Fiktive Ergebnismatrix eines Assessment Centers (Bewertungsskala 1-5)

Dimension / Übung	Rollenspiel (RS)	Gruppendiskussion (GD)	Präsentation (P)	Ø Dimension
Kommunikationsfähigkeit	5	2	4	3.67
Problemlösung	4	2	3	3.00
Teamfähigkeit	(nicht relevant)	1	(nicht relevant)	1.00
Ø Übung	4.5	1.67	3.5	

Beobachtung: Kandidat X zeigt im Rollenspiel hohe Werte in Kommunikation und Problemlösung. In der Gruppendiskussion sind die Werte für dieselben Dimensionen niedrig, Teamfähigkeit ist ebenfalls niedrig. Die Präsentation zeigt wieder bessere Werte. Die Leistung scheint stark von der Übung abzuhängen.

Illustratives Beispiel (Faktoranalytische Darstellung – konzeptionell):

Wenn man die Bewertungen aller Kandidaten einer Faktorenanalyse unterzieht, würden sich typischerweise Faktoren bilden, die den Übungen entsprechen, nicht den Dimensionen.

Tabelle 3: Darstellung einer faktoranalytischen Struktur von AC-Ergebnissen

(Adaptiert von Sackett, P. R., & Dreher, G. F. (1982))

Dimensi- on	Übung	Faktor 1	Faktor 2	Faktor 3	Faktor 4	Faktor 5	Faktor 6
Mündliche Kommuni- kation	Gruppen- diskussi- on	.77					
Interakti- on mit Manage- ment	Gruppen- diskussi- on	.89					
Organisa- tionsfä- higkeit	Gruppen- diskussi- on	.89					
Analyti- sche Fä- higkeiten	Gruppen- diskussi- on	.80					
Organisa- torisches Geschick	Gruppen- diskussi- on	.85					
Schriftli- che Kom- munikati- on	Schriftli- che Kom- munikati- on		.60				
Organisa- tionsfä- higkeit	Schriftli- che Kom- munikati- on		.87				
Analyti- sche Fä- higkeiten	Schriftli- che Kom- munikati- on		.88				
Organisa- torisches Geschick	Schriftli- che Kom- munikati- on		.80				

Dimensi- on	Übung	Faktor 1	Faktor 2	Faktor 3	Faktor 4	Faktor 5	Faktor 6
Mündliche Kommuni- kation	Postkorb			.62			
Schriftli- che Kom- munikati- on	Postkorb			.64			
Erteilung von Ar- beitsauf- trägen	Postkorb			.70			
Organisa- tionsfä- higkeit	Postkorb			.84			
Analyti- sche Fä- higkeiten	Postkorb			.83			
Organisa- torisches Geschick	Postkorb			.84			
Mündliche Kommuni- kation	Rollen- spiel (1)				.68		
Interper- sonelle Kompe- tenz (Mit- arbeiter)	Rollen- spiel (1)				.83		
Analyti- sche Fä- higkeiten	Rollen- spiel (1)				.86		
Mündliche Kommuni- kation	Rollen- spiel (2)					.60	
Interper- sonelle Kompe- tenz (Mit- arbeiter)	Rollen- spiel (2)					.84	

Dimensi- on	Übung	Faktor 1	Faktor 2	Faktor 3	Faktor 4	Faktor 5	Faktor 6
Erteilung von Ar- beitsauf- trägen	Rollen- spiel (2)					.80	
Analyti- sche Fä- higkeiten	Rollen- spiel (2)					.81	
Mündliche Kommuni- kation	Mündliche Kommuni- kation						.72
Organisa- tionsfä- higkeit	Mündliche Kommuni- kation						.79
Analyti- sche Fä- higkeiten	Mündliche Kommuni- kation						.79
Organisa- torisches Geschick	Mündliche Kommuni- kation						.82

Interpretation: Die Bewertungen innerhalb einer Übung korrelieren so stark miteinander, dass sie gemeinsame Faktoren bilden, die primär die jeweilige Übung repräsentieren. Die übergeordneten Dimensionen treten als eigenständige Faktoren oft nicht oder nur schwach hervor.

Die Lösung: Task-Based Assessment Center (TBAC)

Angesichts dieser robusten Befunde plädieren Forscher wie Jackson, Lance und Hoffman (2012) für einen Paradigmenwechsel hin zum **Task-Based Assessment Center (TBAC)**. Die Kernidee des TBAC ist, den "Exercise Effect" nicht als Problem, sondern als Realität anzuerkennen und die AC-Auswertung darauf auszurichten.

Anstatt zu versuchen, abstrakte, übergreifende Dimensionen zu messen, konzentriert sich das TBAC auf die Bewertung der Leistung in den **konkreten Aufgaben (Tasks) oder Übungen** des ACs.

Die Grundprinzipien eines TBAC sind (Jackson et al., 2012, S. 176, Table 8.1)

1. **Fokus auf Übungen/Tasks:** Die Übungen selbst sind die primären Analyseeinheiten. Sie repräsentieren Ausschnitte erfolgskritischer beruflicher Rollen oder Aufgaben.
2. **Jobrelevanz:** Die Übungen müssen so realitätsnah und jobrelevant wie möglich gestaltet sein (hohe Augenscheinvalidität und Inhaltsvalidität).

3. **Verhaltensbasierte Checklisten (innerhalb der Übung):** Anstelle von globalen Dimensionsratings werden für jede Übung spezifische, beobachtbare Verhaltensweisen (Fertigkeiten, Skills) oder zu erreichende Teilergebnisse definiert und bewertet. Diese sind "nested in simulation exercises".
4. **Keine Dimensionsaggregation über Übungen hinweg:** Es wird keine übergreifende Bewertung einer Dimension (z.B. "Kommunikationsfähigkeit") über alle Übungen hinweg vorgenommen. Stattdessen gibt es eine Bewertung für "Kommunikation im Rollenspiel", eine für "Kommunikation in der Gruppendiskussion" etc.
5. **Gesamtbewertung pro Übung:** Jede Übung erhält eine Gesamtbewertung, die die Leistung in dieser spezifischen Aufgabe widerspiegelt.
6. **Feedback:** Das Feedback ist konkret, verhaltens- und aufgabenbezogen und damit oft handlungsorientierter für die Kandidaten.

TBACs: Erfassung von Skills in spezifischen Tasks statt abstrakter Dimensionen

Ein wichtiger Punkt ist, dass der TBAC-Ansatz nicht zwangsläufig bedeutet, dass zugrundeliegende Fähigkeiten (Dimensionen) nicht existieren oder irrelevant sind. Vielmehr geht es darum, was im AC *reliabel gemessen* werden kann.

TBACs erfassen primär **konkrete Fertigkeiten (Skills) und Verhaltensweisen, die zur Bewältigung spezifischer Aufgabenstellungen (Tasks) notwendig sind**. Es ist plausibel, dass diese Skills und Verhaltensweisen von übergeordneten Fähigkeiten (Dimensionen) beeinflusst werden. Beispielsweise wird eine Person mit einer hohen grundlegenden Kommunikationsfähigkeit wahrscheinlich in vielen kommunikationsintensiven Tasks besser abschneiden.

Der "Exercise Effect" deutet jedoch darauf hin, dass die spezifischen Anforderungen und der Kontext einer Übung einen so starken Einfluss haben, dass die Messung der *isolierten* Dimension über verschiedene Kontexte hinweg schwierig wird. Eine Übung (z. B. ein Konfliktgespräch im Rollenspiel) ist eben nicht nur eine "Plattform zur Messung von Kommunikationsfähigkeit", sondern eine komplexe Interaktion, in der auch Problemlösung, Empathie, Durchsetzungsvermögen etc. eine Rolle spielen und die spezifische Situation (Art des Konflikts, Persönlichkeit des Rollenspielers etc.) die Ausprägung dieser Fähigkeiten moduliert.

TBACs erkennen an, dass die Leistung in einer bestimmten Übung eine Mischung aus verschiedenen zugrundeliegenden Fähigkeiten darstellt, die durch die spezifischen Anforderungen dieser Übung hervorgerufen und geformt wird. Die Bewertung konzentriert sich auf das *beobachtbare Ergebnis und Verhalten in dieser spezifischen Aufgabe*, nicht auf die Inferenz einer abstrakten, kontextunabhängigen Dimension.

Wie diese Erkenntnis die AC-Praxis optimieren kann:

Die Implementierung eines TBAC-Ansatzes oder zumindest von Elementen daraus kann die AC-Praxis erheblich verbessern:

1. **Valideres Feedback:** Das Feedback wird spezifischer und handlungsorientierter. Anstatt einem Kandidaten zu sagen "Ihre Kommunikationsfähigkeit ist

verbesserungswürdig“, kann man sagen: “In der Konfliktsimulation mit dem schwierigen Kunden haben Sie zwar Ihre Argumente klar dargelegt (positiv), aber Sie haben zu wenig auf die emotionalen Signale des Kunden geachtet (Entwicklungsfeld).”

2. **Höhere Akzeptanz:** Konkretes, aufgabenbezogenes Feedback wird von Kandidaten oft besser angenommen als abstrakte Dimensionsbewertungen.
3. **Einfachere Beobachterkalibrierung:** Es ist oft einfacher, Beobachter auf die Bewertung konkreter Verhaltensweisen innerhalb einer klar definierten Übung zu kalibrieren, als auf die Interpretation abstrakter Dimensionen über verschiedene Kontexte hinweg.
4. **Realistischere Erwartungen:** Man gibt die Illusion auf, im AC eine Art “Persönlichkeitsprofil” anhand stabiler Dimensionen zu erstellen. Stattdessen erhält man ein Profil der Leistungsfähigkeit in erfolgskritischen beruflichen Aufgaben.
5. **Bessere Prognose?** Es gibt Hinweise (Jackson et al., 2012, S. 223-224), dass TBACs mindestens ebenso gute, wenn nicht sogar bessere prädiktive Validität für beruflichen Erfolg aufweisen können, da sie näher am tatsächlichen Arbeitsverhalten sind (Prinzip der Verhaltensähnlichkeit).
6. **Direkterer Bezug zur Personalentwicklung:** Entwicklungsmaßnahmen können direkt an den beobachteten Defiziten in spezifischen Aufgaben ansetzen.

Der Übergang zu einem reinen TBAC mag für viele Unternehmen ein großer Schritt sein. Aber bereits das Bewusstsein für den “Exercise Effect” und die kritische Reflexion, ob man wirklich abstrakte Dimensionen oder eher die Bewältigung konkreter Aufgaben misst und bewerten will, kann zu erheblichen Verbesserungen im AC-Design und in der Ergebnisinterpretation führen.

Schlussfolgerung: Den Mut zur evidenzbasierten Praxis im Assessment Center haben

Die Wissenschafts-Praxis-Lücke im HR-Bereich ist ein bekanntes und leider persistentes Problem. Im Kontext von Assessment Centern führt sie oft zu Designs, die an wissenschaftlichen Erkenntnissen der letzten Jahrzehnte vorbeigehen. Die willkürliche Anhäufung von Kompetenzen und der Versuch, abstrakte Dimensionen über verschiedene Übungen hinweg zu messen, sind zwei typische Beispiele.

Die Forschung, insbesondere Arbeiten wie die von Meriac et al. (2014) und Jackson et al. (2012), liefert wertvolle Impulse für eine Optimierung:

- Eine **Reduktion auf wenige, robuste übergeordnete Dimensionen** (z. B. das 6-Faktoren-Modell) kann die Komplexität reduzieren, die Beobachterreliabilität erhöhen und den Fokus schärfen.
- Die Anerkennung des **“Exercise Effects”** und die Hinwendung zu einer stärker **aufgabenbasierten Bewertung (TBAC)** kann zu validerem Feedback und einer realistischeren Einschätzung der Leistungsfähigkeit in erfolgskritischen Situationen führen.

Die Umsetzung dieser Erkenntnisse erfordert Mut und die Bereitschaft, etablierte Praktiken zu hinterfragen. Doch der Lohn ist ein Assessment Center, das nicht nur als "Goldstandard" gilt, sondern diesen Anspruch auch durch wissenschaftliche Fundierung und praktische Wirksamkeit einlöst. Es ist an der Zeit, die Brücke zwischen Wissenschaft und Praxis im AC-Design aktiv zu bauen – für validere Entscheidungen und eine effektivere Personalentwicklung.

Quellenverzeichnis (Auswahl)

- Arthur, W., Jr., Day, E. A., McNelly, T. L., & Edens, P. S. (2003). A meta-analysis of the criterion-related validity of assessment center dimensions. *Personnel Psychology*, 56(1), 125–154. <https://doi.org/10.1111/j.1744-6570.2003.tb00146.x>
- Hell, B., Päßler, K., & Nido, M. (2023). Wissenschafts-Praxis-Gap in der Eignungsdiagnostik: Die Validität von Personalauswahlverfahren aus Sicht von HR-Fachpersonen im Vergleich zu metaanalytischen Erkenntnissen. *Wirtschaftspsychologie*, 25(141-148). <https://doi.org/10.5281/zenodo.13833658>
- Jackson, D. J. R., Lance, C. E., & Hoffman, B. J. (Eds.). (2012). *The psychology of assessment centers*. Routledge.
- 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), 1269–1296. <https://doi.org/10.1177/0149206314522299>
- Povah, N., & Povah, L. (2012). What are assessment centers and how can they enhance organizations? In D. J. R. Jackson, C. E. Lance, & B. J. Hoffman (Eds.), *The psychology of assessment centers* (pp. 3-23). Routledge.
- Sackett, P. R., & Dreher, G. F. (1982). Constructs and assessment center dimensions: Some troubling empirical findings. *Journal of Applied Psychology*, 67(4), 401–410. <https://doi.org/10.1037/0021-9010.67.4.401>