Evaluation von Assessment Centern – Ein Wegweiser zur Berechnung der Kriteriumsvalidität

Anlass von Swiss Assessment

Dr. Pia Ingold

Arbeits- und Organisationspsychologie

Universität Zürich

Wegweiser

- Einstieg: Evaluation von ACs
  - Was versteht man darunter?
  - Welche Kriterien?
  - Welches Design?
- Erhebung von Kriteriumsdaten
- Berechnung der Kriteriumsvalidität
- Übliche Höhe der Kriteriumsvalidität
- Mögliche Verzerrungen der Kriteriumsvalidität und Umgang damit

Vermittlungsziele

- grundsätzliche Vorgehensweise zur Bestimmung der Kriteriumsvalidität
- relevante Aspekte für die Wahl von Kriterium und Evaluationsdesign
- wie kann man mit Einflussfaktoren umgehen kann, die die Kriteriumsvalidität beeinflussen

Indikatoren für die Güte eines Auswahlverfahrens

- Validität
  - Kriteriumsvalidität
  - Konstruktvalidität
  - Inhaltsvalidität
- Reliabilität
  - Retest-Reliabilität
  - Beurteiler-Reliabilität
- Fairness (= gleiche Chance für Mitglieder unterschiedlicher Bewerbergruppen)

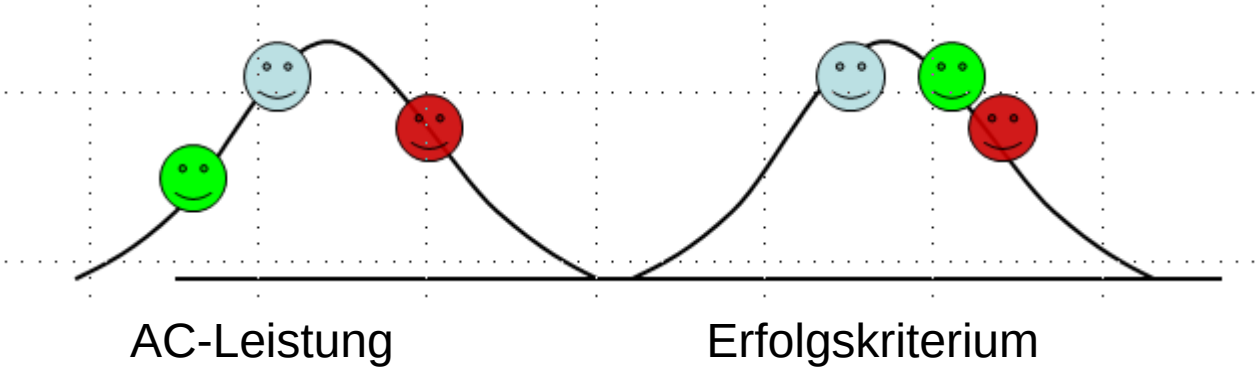
Kriteriumsvalidität

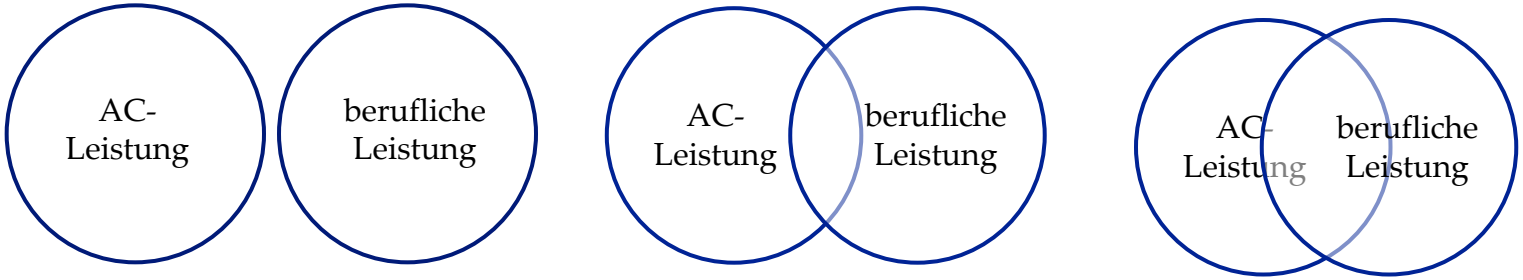
= Korrelation zwischen Leistung im Auswahlverfahren und Erfolgskriterium

- „Berufliche Leistung“ als häufigstes Erfolgskriterium zur Bestimmung der Kriteriumsvalidität
- In der Unternehmenspraxis Kriteriumsvalidität oft als wichtigstes Gütekriterium eines Auswahlverfahrens

Kriteriumsvalidität

= Korrelation zwischen Leistung im Auswahlverfahren und Erfolgskriterium





Kriteriumsvalidität

- Korrelation kann Werte zwischen -1 und 1 annehmen und ist ein Mass des Zusammenhangs

a) positiver Zusammenhang (ab 0 bis 1)

= Je höher die AC-Leistung eines Kandidaten, desto besser ist seine berufliche Leistung.

b) kein Zusammenhang (0)

= Wie die AC-Leistung ausfällt, hängt überhaupt nicht mit der beruflichen Leistung des Kandidaten zusammen.

c) -1 bis 0: **negativer Zusammenhang (-1 bis 0)**

= Je höher AC-Leistung eines Kandidaten, desto schlechter ist seine berufliche Leistung.

Wegweiser

- Einstieg: Evaluation von ACs
  - Was versteht man darunter?
  - Welche Kriterien?
  - Welches Design?
- Erhebung von Kriteriumsdaten
- Berechnung der Kriteriumsvalidität
- Übliche Höhe der Kriteriumsvalidität
- Mögliche Verzerrungen der Kriteriumsvalidität und Umgang damit

Auswahl der Kriterien

Was soll eigentlich vorhergesagt werden?

- Berufliche Leistung:
  - Aufgabenleistung (= Task Performance) generell oder auf spezif. Dimensionen?
  - Leistung über die eigentliche Tätigkeit hinaus
  - subjektive oder objektive Leistung
- Beruflicher Erfolg:
  - Beförderungen, Einkommen …
- Berufliche Zufriedenheit:
  - Arbeitszufriedenheit, Commitment, kontraproduktives Verhalten…

Generelle Leistung vs. spezifische Dimensionen (I)

Option 1

- Erhebung unterschiedlicher Dimensionen
- Berechnung der Kriteriumsvalidität pro Dimension

Option 2

- Nur Erhebung der generellen Leistung oder
- Berechnung eines Gesamtwertes über einzelne Dimensionen hinweg

Generelle Leistung vs. spezifische Dimensionen (II)

Argumente, die für Option 2 sprechen:

- In vielen Fällen wird ein Gesamtwert benötigt (z. B. Beförderungsentscheidungen)
- Metaanalyse von Viswesvaran et al. (2005):
  - i. d. R. bedeutsame positive Korrelationen verschiedener Beurteilungsdimensionen
  - In Korrelationen lässt sich ein Globalfaktor finden
  - Dieser klärt etwa 50% der Varianz in den Leistungsbeurteilungen auf
- Gesamtwert hat höhere Reliabilität als spezifische Einzeldimensionen

Leistung über die eigentliche Tätigkeit hinaus

- Organizational Citizenship Behavior (OCB)
  - OCB gegenüber der Organisation (OCB-O: Gewissenhaftigkeit …) und gegenüber individuellen Personen (OCB-I: Hilfsbereitschaft…)

- i. d. R. korrelieren beide OCB-Aspekte jedoch substantiell miteinander und mit Aufgabenleistung (Hoffman et al., 2007)

Subjektive vs. objektive Indikatoren von Leistung

- Viele objektive Leistungs-Indikatoren sind denkbar: Umsatz, Anzahl Neukunden, Anzahl Beschwerden,…

- Problem: Objektive Indikatoren nicht immer vorhanden oder stark von externen Einflüssen betroffen
- Zusammenhang zw. subjektiven und objektiven Leistungs-Indikatoren (Meta-Analyse Bommer et al., 1995) nicht allzu hoch

Passung von Prädiktor und Kriterium

- **Arbeitsleistung** lässt sich besser durch Prädiktoren vorhersagen, die vor allem **Leistung und tätigkeitsrelevante Kenntnisse** erfassen
- **Zufriedenheit** und Korrelate von Zufriedenheit (Absentismus, kontraproduktives Verhalten, Kündigung, …) lassen sich besser durch Prädiktoren vorhersagen, die **Motivation und Interessen** erfassen

Ratschlag Kriteriumsauswahl

- Kriterium sollte immer Bezug zum Erfolg bei der Tätigkeit aufweisen (vgl. SIOP Guidelines 2003)
- aufgabenbezogene Leistung Klassiker als Kriterium→ empfehlenswert über Tätigkeiten hinweg
  - Aufteilung nach Dimensionen nur bei spez. Interesse

Ratschlag Kriteriumsauswahl

- weitere Kriterien:
  - unternehmensinterne Daten
    - objektive Daten (z.B. Neukunden im Verkaufsbereich, Stückzahlen)
    - Fortbildungen, Trainings etc.
    - Erfolg: Beförderungen, Gehalt etc.
  - eher optional: Arbeitszufriedenheitsmessungen
- generell: keine hohe Übereinstimmung zw. Ergebnissen der Kriteriumsvalidität mit subj. vs. obj. Daten zu erwarten

Wegweiser

- Einstieg: Evaluation von ACs
  - Was versteht man darunter?
  - Welche Kriterien?
  - Welches Design?
- Erhebung von Kriteriumsdaten
- Berechnung der Kriteriumsvalidität
- Übliche Höhe der Kriteriumsvalidität
- Mögliche Verzerrungen der Kriteriumsvalidität und Umgang damit

Relevante Aspekte des Evaluationsdesigns

- Stelleninhaber vs. Bewerber als Evaluationsstichprobe
- Konkurrente vs. prädiktive Validität
- Kriteriumsmessung speziell für die Evaluationsstudie?

Stelleninhaber vs. Bewerber

- Stelleninhaber z. B. gut verfügbar für Vortest bei Neuentwicklung eines Verfahren

Aber: Stelleninhaber…

- … verfügen über bessere Tätigkeitskenntnisse
- … strengen sich ggf. weniger an
- … verstellen sich weniger

→ ihre Leistung im AC spiegelt unter Umständen eher Tätigkeitskenntnisse oder typisches Verhalten wider

Konkurrente vs. prädiktive Validität

- In der Regel fallen konkurrente Validitäten höher aus als prädiktive Validitäten (Van Iddekinge & Ployhart, 2008)

- Mögliche Ursachen:
  - Merkmalsinstabilität
  - Veränderung der Bedeutung eines Prädiktors

Kriteriumsmessung speziell für Evaluationsstudie?

2 Optionen:

- Verwendung vorliegender Leistungsbeurteilungen, die für administrative Zwecke erhoben wurden (Jahresbeurteilungen o. ä.)
  - Beurteilungen sind „öffentlich“ und relevant für Beförderungen, Gehalt etc.

  → i. d. Regel erhebliche Mildefehler
- Erhebung speziell für Evaluationsstudie
  - Beurteilungen geheim und ohne Relevanz für Beurteilte

  → i. d. Regel relativ angemessene Beurteilungen und damit brauchbarer für Evaluationsstudien

Ratschlag Evaluationsdesign

- wenn Möglichkeit vorhanden, Daten neu erheben
- konkurrentes Design potentiell höherer Zusammenhang, aber nach Umsetzbarkeit auswählen
  - z. B. nach dem AC bei Vorgesetzten Kriteriumsdaten erheben (z.B. Kundenpflege, Qualitätssicherung)
  - vor/nach Feedbackgespräch zum AC mit dem Bewerber potentiell Möglichkeit zur Erhebung von Selbstberichtdaten (z.B. Akzeptanz des Verfahrens)

Wegweiser

- Einstieg: Evaluation von ACs
  - Was versteht man darunter?
  - Welche Kriterien?
  - Welches Design?
- Erhebung von Kriteriumsdaten
- Berechnung der Kriteriumsvalidität
- Übliche Höhe der Kriteriumsvalidität
- Mögliche Verzerrungen der Kriteriumsvalidität und Umgang damit

Konkretere Planung der Erhebung von Kriteriumsdaten

- Festlegung der Kriterien
  - Welches Kriterium ist relevant für die Evaluation?
  - Welche Fragen können zur Messung eingesetzt werden?
- Festlegung der zu befragenden Personen
  - Vorgesetzte
  - Mitarbeiter, Kunden etc.
- Wahl der Erhebungsmethode
  - PaperPencil
  - Onlinesurvey
  - ggf. Gespräche

Messung der Kriterien

- Höhe der maximal erreichbaren Kriteriumsvalidität wird durch die Reliabilität der Messung beeinflusst
- Reliabilität = Genauigkeit, mit der ein Test eine Merkmalsdimension erfasst
- Unreliable Messung des Kriteriums → schlechtere Kriteriumsvalidität
- Deshalb: etablierte Messinstrumente mit hoher Reliabilität benutzen

Beispiel Messung Aufgabenbezogene Arbeitsleistung

Im Vergleich zu Arbeitskollegen bzw. Arbeitskolleginnen...

... erreicht sie die Zielsetzungen bei ihrer Arbeit.

... zeigt sie Fachkenntnisse in allen arbeitsrelevanten Aufgaben.

... übernimmt sie mehr Verantwortung als für ihre Arbeit erforderlich ist.

... ist sie in allen Arbeitsbereichen kompetent.

... plant und organisiert sie das Vorgehen, um ihre Arbeitsziele zu erreichen.

Antwortformat von 1 = “trifft überhaupt nicht zu” zu 7 = “trifft voll und ganz zu”

Fragebogenskala task-based performance von Bott, Svyantek, Goodman, and Bernal (2003) übersetzt nach Jansen et al. (2013)

Beispiel Messung Organizational Citizenship Behavior

Der Mitarbeiter/ die Mitarbeiterin...

…hilft anderen, wenn diese mit Arbeit überlastet sind.

… kommt immer pünktlich zur Arbeit.

... verbringt viel Zeit damit, sich über Belanglosigkeiten zu beklagen.

... wirkt bei auftretenden Meinungsverschiedenheiten ausgleichend auf Kollegen/Kolleginnen ein.

... informiert frühzeitig, wenn er nicht zur Arbeit kommen kann.

... neigt dazu, aus einer Mücke einen Elefanten zu machen.

... ergreift freiwillig die Initiative, neuen Kollegen/Kolleginnen bei der Einarbeitung zu helfen.

Quelle: Staufenbiel, T., & Hartz, C. (2000). Organizational Citizenship Behavior: Entwicklung und erste Validierung eines Meßinstruments [Organizational citizenship behavior: Development and validation of a measurement instrument]. *Diagnostica, 46, 73-83. doi:10.1026//0012-1924.46.2.73*

Teildarstellung des Messinstruments auf den Folien (Itemzahl insgesamt 20)

Beispiel Messung Organizational Citizenship Behavior

Der Mitarbeiter/ die Mitarbeiterin...

... zeichnet sich durch besonders wenige Fehlzeiten aus.

... sieht alles, was das Unternehmen macht, als falsch an.

... bemüht sich aktiv darum, Schwierigkeiten mit Kollegen/Kolleginnen vorzubeugen.

... beachtet Vorschriften und Arbeitsanweisungen mit grösster Sorgfalt.

... kritisiert häufig an Kollegen/Kolleginnen herum.

... ermuntert Kollegen/Kolleginnen, wenn diese niedergeschlagen sind.

... nimmt sich nur in äusserst dringenden Fällen frei.

... äussert Vorbehalte gegenüber jeglichen Veränderungen im Unternehmen.

Quelle: Staufenbiel, T., & Hartz, C. (2000). Organizational Citizenship Behavior: Entwicklung und erste Validierung eines Meßinstruments [Organizational citizenship behavior: Development and validation of a measurement instrument]. *Diagnostica, 46, 73-83. doi:10.1026//0012-1924.46.2.73*

Weitere Kriterien

- objektive Kriterien: Anzahl Beförderungen, Anzahl Neukunden, Einkommen etc.

→ ggf. zugänglich, auch bei externen Anbietern

- Commitment, Absenzen des Mitarbeiters, kontraproduktives Arbeitsverhalten

Beispiel Messung Kontraproduktives Arbeitsverhalten

Bitte geben Sie an, wie häufig Sie in den vergangenen zwölf Monaten bestimmte Verhaltensweisen am Arbeitsplatz gezeigt haben (Antwortkategorien: nie, einmal im Jahr, zweimal im Jahr, mehrere Male im Jahr, monatlich, wöchentlich, täglich)

10. Ich verfälschte Abrechnungen über Geschäftsausgaben zu eigenen Gunsten.
11. Ich legte zu viele oder zu lange (Arbeits-)Pausen ein.
12. Ich erschien ohne Erlaubnis verspätet zur Arbeit.
13. Ich hinterließ meine Arbeitsumgebung unordentlich/dreckig.
14. Ich befolgte Anweisungen meiner/s Vorgesetzten nicht.
15. Ich arbeitete absichtlich langsamer, als ich es hätte tun können.
16. Ich sprach mit Unbefugten über vertrauliche (Firmen-)Informationen.
17. Ich nahm Drogen oder konsumierte Alkohol während der Arbeit.
18. Ich investierte wenig Engagement in meine Arbeit.
19. Ich schob meine Arbeit auf, um Überstunden anzusammeln.

Deutsche Übersetzung der Items von Bennett & Robinson (2000)

Quelle: Zettler, I. (2009). *Politische Fähigkeiten und berufliche Leistung: Ein kurvilinearer Zusammenhang.* Dissertation, Rheinisch-Westfälische Technische Hochschule Aachen.

Bsp. Onlineumfrage Vorgesetzte

- Onlineumfrage (Questback, Surveymonkey etc.) mit personalisierter Einladung per Mail

Fragebogen zur Arbeitsleistung

Im Folgenden finden Sie eine Reihe von Aussagen über die Arbeitsleistung Ihres Mitarbeiters #u_firstname# #u_name#. Bitte geben Sie für jede Aussage an, in welchem Ausmass die Aussage auf Ihren Mitarbeiter zutrifft.

Beachten Sie bitte, dass alle Aussagen im Vergleich zu Arbeitskollegen/-innen von #u_firstname# #u_name# einzuschätzen sind, die eine vergleichbare Position haben bzw. eine vergleichbare Tätigkeit ausüben.

Sind momentan keine Arbeitskollegen/-innen in einer vergleichbaren Position vorhanden, können Sie auch frühere Mitarbeiter/-innen als Vergleichspersonen heranziehen.

Die Daten werden rein zu Forschungszwecken verwendet und weder an Ihren beurteilten Mitarbeiter noch an Dritte weitergegeben.

Zurück | Weiter

Bsp. Onlineumfrage Vorgesetzte

Bitte geben Sie für jede Aussage an, in welchem Ausmass die Aussage auf Ihren Mitarbeiter #u_firstname# #u_name# zutrifft.

Beachten Sie bitte, dass alle Aussagen im Vergleich zu Arbeitskollegen/-innen Ihres Mitarbeiters einzuschätzen sind, die eine vergleichbare Position haben bzw. eine vergleichbare Tätigkeit ausüben.

Sind momentan keine Arbeitskollegen/-innen in einer vergleichbaren Position vorhanden, können Sie auch frühere Mitarbeiter/-innen als Vergleichspersonen heranziehen.

Der Mitarbeiter...

	trifft gar nicht zu	trifft nicht zu	trifft eher nicht zu	trifft teil-weise zu	trifft eher zu	trifft zu	trifft voll und ganz zu
... ist in allen Arbeitsbereichen kompetent.	○	○	○	○	○	○	○
... plant und organisiert das Vorgehen, um seine Arbeitsziele zu erreichen.	○	○	○	○	○	○	○
... erfüllt die gesetzten Leistungsanforderungen an seine Position.	○	○	○	○	○	○	○
... kommt den in den Arbeitsplatzbeschreibungen festgelegten Verpflichtungen nach.	○	○	○	○	○	○	○

Ratschlag Evaluationsplanung

- etablierte Messinstrumente mit hoher Reliabilität benutzen, sonst geringere Validität
- externe AC-Anbieter:
  - Qualitätssicherung/ Optimierung des ACs als Vorteil für Kunden herausstreichen
  - Vertrauliche Verwendung der Daten betonen
  - Zeitlichen Aufwand für Befragung genau prüfen und eingrenzen
- ggf. mehrere Quellen einplanen (Sicherheitsnetz bei geringem Commitment der Befragten, vorhandene Daten nutzen)
- Pros/ Cons der Art der Befragung abwägen (z.B. online kann anonymer als PaperPencil erscheinen)

Wegweiser

- Einstieg: Evaluation von ACs
  - Was versteht man darunter?
  - Welche Kriterien?
  - Welches Design?
- Erhebung von Kriteriumsdaten
- Berechnung der Kriteriumsvalidität
- Übliche Höhe der Kriteriumsvalidität
- Mögliche Verzerrungen der Kriteriumsvalidität und Umgang damit

Zwischenschritte vor Berechnung

- Umkodierung von invers kodierten Items
  - Bsp.: ...erfüllt die Aufgaben nicht vs. erfüllt die Aufgaben
- Berechnung der Reliabilität der Kriteriumsdaten
- Berechnung von Mittelwerten der Skala für jeden Teilnehmer

✎ **Achtung: geringe Reliabilität**
→ geringere Kriteriumsvalidität

Berechnung der Kriteriumsvalidität

- Zentrales Prädiktor-Mass: Gesamtleistung im AC (englisch: Overall Assessment Rating, OAR)
- Zentrales Kriteriumsmass: Aufgabenbezogene Leistung

Bsp. Excelfile

Kriteriumsvalidität berechnet…

und nun?

Kriteriumsvalidität von ACs – Was sind übliche Werte?

- Einige Studien finden beeindruckende Kriteriumsvaliditäten von ACs bzgl. der Vorhersage beruflicher Leistung und Karriereentwicklung
- Metaanalysen finden eher mittlere Effekte bzgl. der Kriteriumsvalidität

Kriteriumsvalidität von ACs – Was sind übliche Werte?

Ergebnisse aus metaanalytischen Studien:

Kriterium	*r*	ρ	Quelle
Vorgesetztenbeurteilungen	.25	.36	Gaugler et al. (1987)
	.17	.28	Hermelin et al. (2007)
	.20	.28[a]	Hardison & Sackett (2007)
Trainingsleistung	.30	.35	Gaugler et al. (1987)
	.31	.35[a]	Hardison & Sackett (2007)
Karrierefortschritt	.30	.36	Gaugler et al. (1987)
Beförderung	.27		Hardison & Sackett (2007)

[a] keine Korrektur für Varianzeinschränkung

Glaubwürdigkeitsintervalle Hardison & Sackett (2004)

Meta-analysis results for current study, Gaugler et al. (1987) and combined studies.

						corrected for criterion unreliability		corrected for criterion unreliability and range restriction		90% Credibility intervals	
	Sample size	Number of r's	Number of studies (k) [a]	Average r	SDr	ρ	SDρ	ρ	SDρ	Lower	Upper
Total											
Combined	23371	213	87	.26	.15	.30	.13	-	-	.04	.56
Current	11136	106	40	.22	.14	.26	.13	-	-	-.01	.52
Gaugler	12235	107	47	.29	.15	.34	.13	.37	.13	.08(.12) [b]	.59(.62) [b]
Performance											
Combined	8378	93	58	.22	.12	.31	.12		-	.07	.55
Current	4198	49	29	.20	.08	.28	.00	-	-	.28	.28
Gaugler	4180	44	29	.25	.15	.35	.17	.36	.14	.01(.09) [b]	.68(.63) [b]
Training											
Combined	4565	23	18	.31	.11	.35	.11	-	-	.14	.55
Current	3503	15	10	.31	.10	.35	.09	-	-	.16	.53
Gaugler	1062	8	8	.30	.15	.34	.14	.35	.14	.06(.08) [b]	.61(.62) [b]

Wegweiser

- Allgemeine Hintergrundinformationen zur Evaluation von ACs
  - Was versteht man darunter?
  - Welche Kriterien?
  - Welches Design?
- Erhebung von Kriteriumsdaten
- Berechnung der Kriteriumsvalidität
- Übliche Höhe der Kriteriumsvalidität
- Mögliche Verzerrung
  - Korrekturen für Unreliabilität und Varianzeinschränkung

Kriteriumsvalidität berechnet und mit metaanalytischen Werten verglichen

und nun endlich?

Einflussfaktoren, die die Kriteriumsvalidität verzerren

- Es gibt eine Reihe von Faktoren, die die Kriteriumsvalidität systematisch verzerren (vgl. z. B. Hunter & Schmidt, 2004)

In der Personalauswahl von besonderer Relevanz:

- Unreliabilität der Variablen
- Varianzeinschränkung im Prädiktor

Unreliabilität der Variablen

Auswirkung unterschiedlicher Reliabilitäten auf $r_{x_t y_t} = .50$

Table 1. Average Value of r_{xy} as a Function of r_{xx} and r_{yy} When the Actual Correlation between Constructs ($r_{x_t y_t}$) is 0.50

Reliability of Measure y	*Reliability of Measure x*					
	0.40	*0.50*	*0.60*	*0.70*	*0.80*	*0.90*
0.40	0.20					
0.50	0.22	0.25				
0.60	0.25	0.27	0.30			
0.70	0.27	0.30	0.32	0.35		
0.80	0.28	0.32	0.35	0.37	0.40	
0.90	0.30	0.34	0.37	0.40	0.42	0.45

aus Schmidt & Hunter (1999)

Unreliabilität des Kriteriums

Optionen zur Einschätzung:

- Bei Beurteilung durch mehrere Beurteiler:
  Ermittlung der Beurteiler-Reliabilität
- Bei wiederholter Beurteilung:
  Ermittlung der Retest-Reliabilität
- Verwendung von metaanalytisch ermittelten Reliabilitätswerten, z. B. Viswesvaran et al. (1996):
  - z.B. für Korrektur der Reliabilität von Beurteilungen der generellen Arbeitsleistung durch Vorgesetzte
    - .52 als Beurteiler-Reliabilität
    - .81 als Retest-Reliabilität

Einfache Minderungskorrektur

Die Validität kann nur maximal so hoch sein wie die Wurzel der Reliabilität, deshalb:

Minderungskorrektur zur Berechnung der tatsächlichen Validität

Korrektur für Unreliabilität des Kriteriums:

$$r_{xy_t} = \frac{r_{xy}}{\sqrt{Rel_y}}$$

- Regelfall bei Bestimmung der Kriteriumsvalidität von Personalauswahlverfahren
  → nur Korrektur für Unreliabilität der Kriteriums-Variable (z. B. berufl. Leistung)

Bsp. Korrigierte Validität

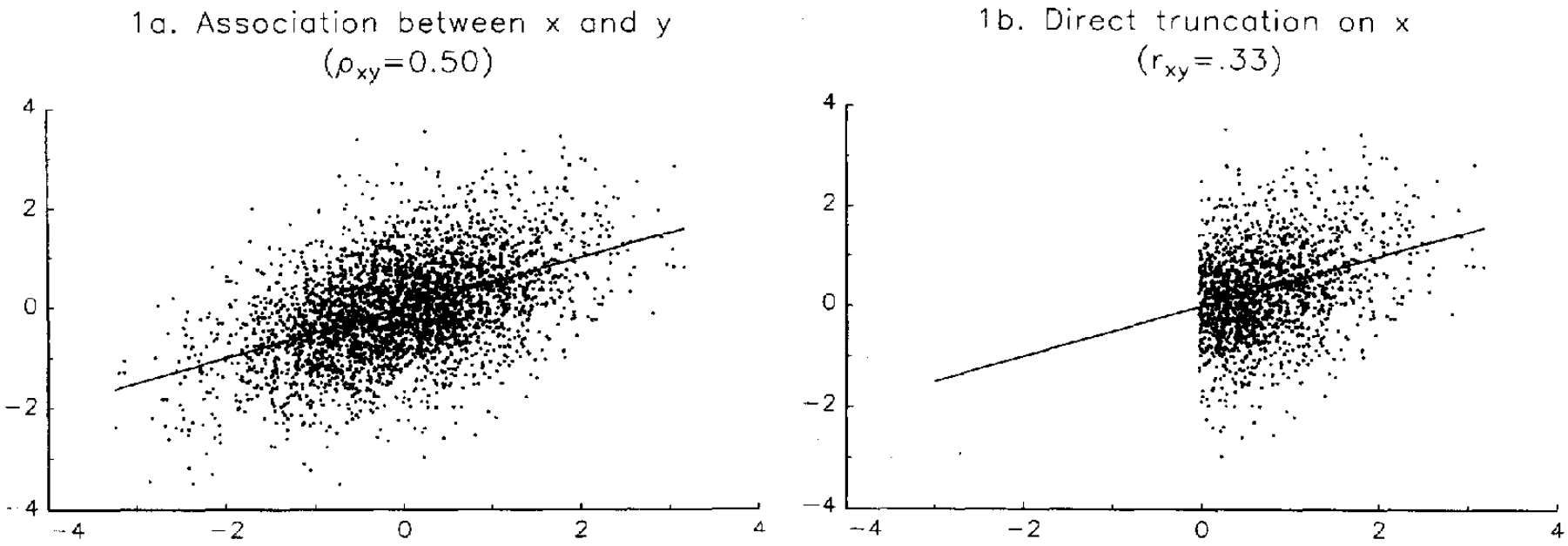
emp. Korrelation	Reliabilität Kriterium	korrig. Korrelation
.24	.52	.33

Varianzeinschränkung

- Korrelation prüft, wie sehr Prädiktor und Kriterium gemeinsam variieren
- Einschränkungen der Varianz im Prädiktor schränken damit die mögliche Höhe einer Korrelation ein
- Typisches Problem in der Personalauswahl:
  - nur Personen über Cut-Off-Wert werden eingestellt

Varianzeinschränkung durch Selektion

Selektive Einstellung von Personen über einem Cut-Off-Wert



aus Sackett & Yang (2000)

Ermittlung der Varianzeinschränkung

- Bestimmung der Standardabweichung im Prädiktor für Gesamtgruppe aller Bewerber (SD_{Ges})
- Bestimmung der Standardabweichung im Prädiktor für Gruppe der eingestellten Bewerber (SD_{Eing})
- Bestimmung des Korrektur-Faktors $u = SD_{\text{Eing}}/SD_{\text{Ges}}$
- Berechnung der korrigierten Korrelation r_{korr} :

$$r_{korr} = \frac{\frac{r_{Eing}}{u}}{\sqrt{1 - r_{Eing}^2 + \frac{r_{Eing}^2}{u^2}}}$$

Zusammenfassung

- Die Wahl des Kriteriums beeinflusst die Höhe der möglichen Kriteriumsvalidität
- Häufiges Kriterium: aufgabenbezogene berufliche Leistung
- Die Höhe der Kriteriumsvalidität wird oft unterschätzt, da sie durch beeinträchtigende Faktoren systematisch nach unten verzerrt werden kann
- Verzerrungen durch Unreliabilität und Varianzeinschränkung können rechnerisch berücksichtigt werden

Implikationen

- individuelle Lösung zur Evaluation für den eigenen Kontext finden
- Empfehlung bei eigener Erhebung:
  - Vorgesetztenurteile online zu aufgabenbezogener beruflicher Leistung durchführen und je nach Interesse 1-2 weitere Masse erheben (max. 7 min Dauer)
  - wenn zugänglich, Daten des Unternehmens (obj. Daten etc.) einbeziehen und bzgl. Ergebnissen vergleichen (zukünftige Entschlackung der eigenen Erhebung)