

Vortrag anlässlich der Vereinsversammlung von
Swiss Assessment
Zürich, 15. Januar 2019

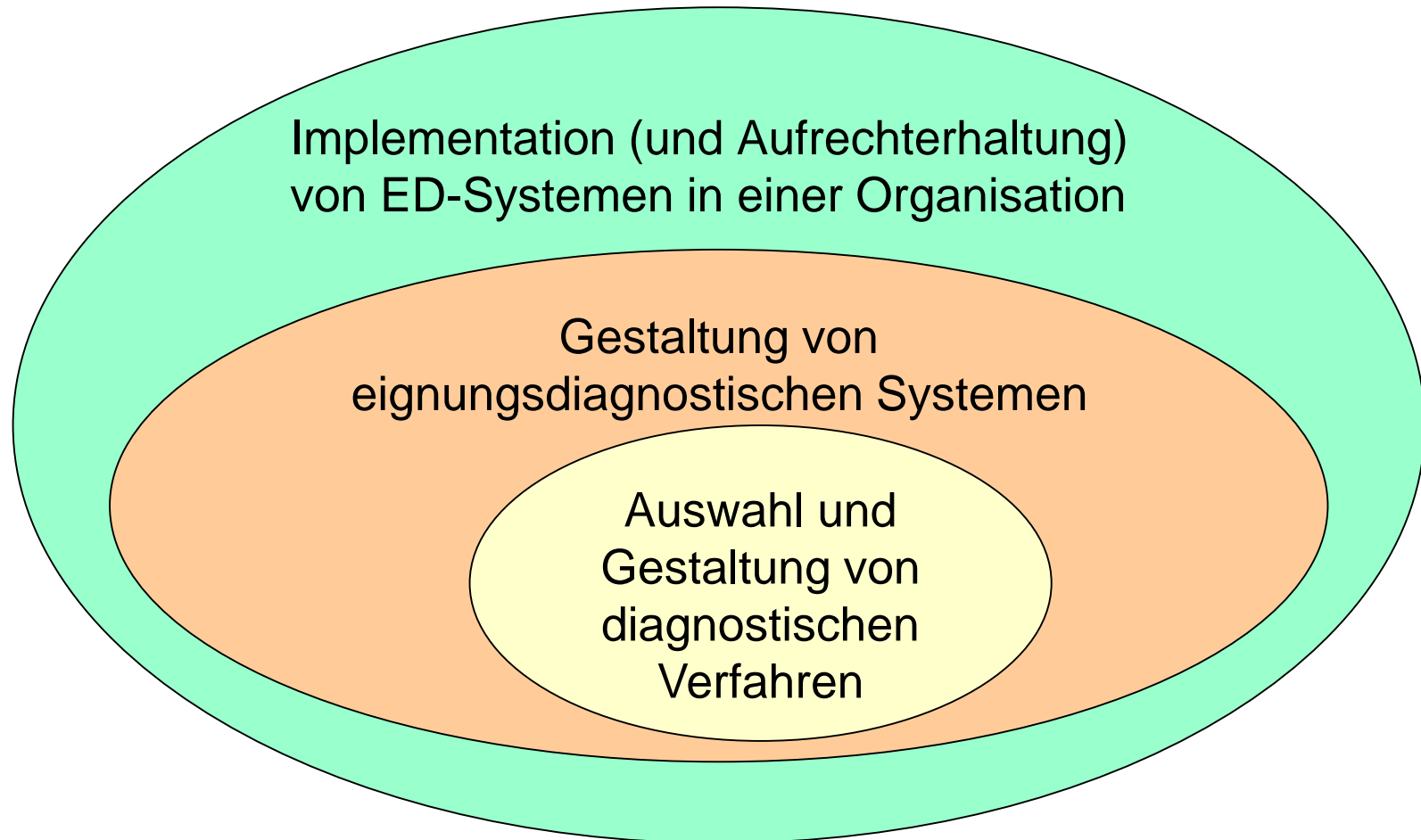


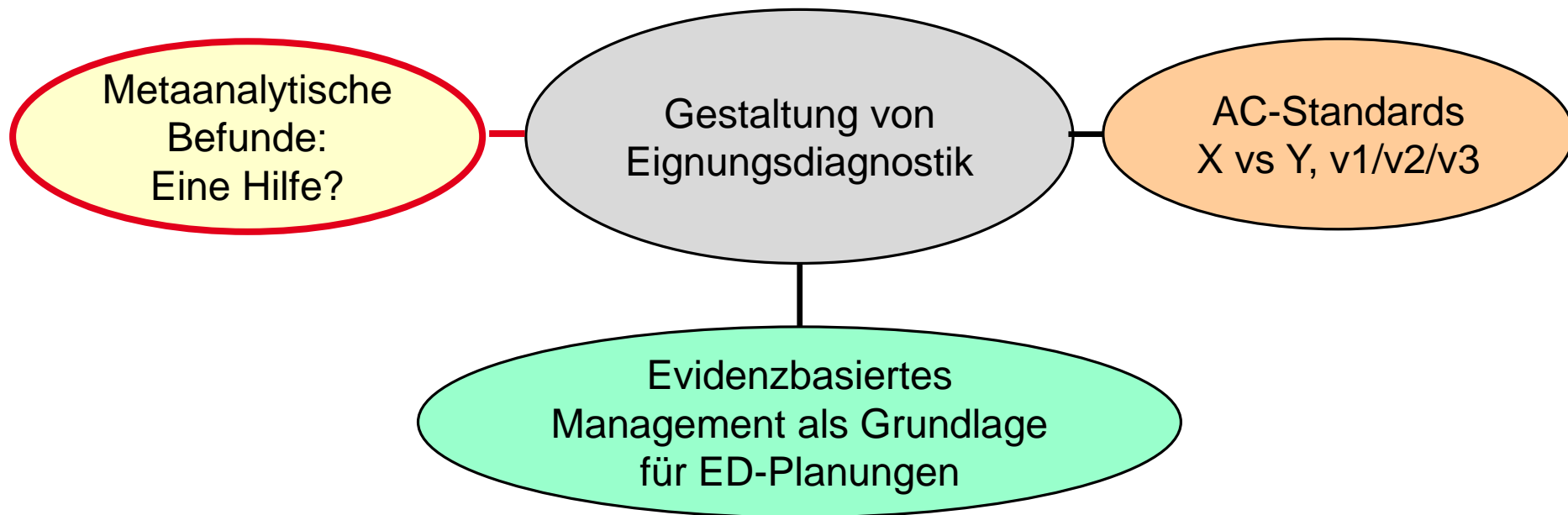
Standards für Assessment Center

Qualitätskriterien zur Gestaltung von Eignungsdiagnostik

Prof. Dr. Stefan Höft
stefan.hoeft@hdba.de







NÜTZLICHE WISSENSCHAFT?

Neuere metaanalytische Befunde zur Vorhersagegüte eignungsdiagnostischer Verfahrensklassen

Einen dosierten Einstieg in die Materie bietet:

Höft, S. (2014). Erfolgsüberprüfung personalpsychologischer Arbeit. In H. Schuler & U. P. Kanning (Hrsg.), *Lehrbuch der Personalpsychologie* (3., vollständig überarbeitete Auflage) (S. 1081-1136). Göttingen: Hogrefe.

PUBLIKUMSFRAGE:

Welches Verfahren würden Sie für die Personalauswahl einer Traineeestelle einsetzen, wenn Sie nur ein einziges Verfahren einsetzen dürfen?

- Wählen Sie mit Ihrem Smartphone bitte die folgende Seite an:
www.menti.com
- Geben Sie den Code **45 07 64** an.
- Wählen Sie eine der drei Optionen aus.

Einer der einflussreichsten Artikel der letzten 30 Jahre zum Generalthema „Validität eignungsdiagnostischer Verfahren“

Psychological Bulletin
1998, Vol. 124, No. 2, 262–274

Copyright 1998 by the American Psychological Association, Inc.
0033-2909/98/\$3.00

The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of Research Findings

Frank L. Schmidt
University of Iowa

John E. Hunter
Michigan State University

This article summarizes the practical and theoretical implications of 85 years of research in personnel selection. On the basis of meta-analytic findings, this article presents the validity of 19 selection procedures for predicting job performance and training performance and the validity of paired combinations of general mental ability (GMA) and the 18 other selection procedures. Overall, the 3 combinations with the highest multivariate validity and utility for job performance were GMA plus a work sample test (mean validity of .63), GMA plus an integrity test (mean validity of .65), and GMA plus a structured interview (mean validity of .63). A further advantage of the latter 2 combinations is that they can be used for both entry level selection and selection of experienced employees. The practical utility implications of these summary findings are substantial. The implications of these research findings for the development of theories of job performance are discussed.

Metaanalytische Befundlage zur Validität eignungsdiagnostischer Verfahren

		<i>Schmidt & Hunter (1998)</i>		
Prädiktor		Val	mR	inkrV
Eigenschaft	Allgemeine kognitive Fähigkeitstests	.51		
	“Integrity“-Tests	.41	.65	.14
	Gewissenhaftigkeitstests	.31	.60	.09
	Fachkenntnistests	.48	.58	.07
	Interessen	.10	.52	.01
Simulation	Probezeit	.44	.58	.07
	Arbeitsproben	.54	.63	.12
	Situational Judgement Tests			
	Assessment Center	.37	.53	.02
Biografie	Graphologie	.02	.51	.00
	Biografische Daten	.35	.52	.01
	Unstrukturiertes Einstellungsgespräch	.36	.55	.04
	Strukturiertes Einstellungsgespräch	.51	.63	.12

Anmerkungen: Als Artefakte wurden die Kriteriumsreliabilität und Varianzeinschränkungen korrigiert.

Val: kriteriumsbezogene Validität (Kriterium: Leistungsbeurteilung)

MR: geschätzte kombinierte Vorhersage in einer multiplen Regression

inkrV: inkrementelle Validität, d.h. der Validitätszuwachs durch Hinzunahme des zweiten Prädiktors zstl zum allgemeinen kognitive Fähigkeitstest

Kasten 4:

Zusammenfassung der metaanalytischen Befundlage zur Validität eignungsdiagnostischer Verfahren (verkürzt nach Schmidt & Hunter, 1998, S. 22)

Prädiktor	Val	mR	inkrV	%Zuw	Quelle
allgemeine kognitive Fähigkeitstests	.51				Abs. 2.2.1
Arbeitsproben	.54	.63	.12	24%	Kap. 6
„Integrity“-Tests	.41	.65	.14	27%	Abs. 3.1
Gewissenhaftigkeitstests	.31	.60	.09	18%	Abs. 2.2.3
strukturiertes Einstellungsgespräch	.51	.63	.12	24%	Kap. 7
unstrukturiertes Einstellungsgespräch	.38	.55	.04	8%	Kap. 7
Fachkenntnistests	.48	.58	.07	14%	Abs. 2.2.2
Probezeit	.44	.58	.07	14%	
Biographische Daten	.35	.52	.01	2%	Kap. 7
Assessment Center	.37	.53	.02	4%	Kap. 6
Interessen	.10	.52	.01	2%	Abs. 2.2.5
Graphologie	.02	.51	.00	0%	Kap. 7

Als Artefakte wurden die Kriteriumsreliabilität und Varianzeinschränkungen korrigiert.

Folgende Abkürzungen werden verwendet:

Val: Kriteriumsbezogene Validität (Kriterium: Leistungsbeurteilung)

mR: multiple Korrelation mit allgemeiner kognitiver Fähigkeit als erster und mit dem jeweiligen Prädiktor als zweiter Variable

inkrV: inkrementelle Validität, d.h. der Validitätszuwachs durch Hinzunahme des zweiten Prädiktors

%Zuw: Prozentualer Zuwachs der Validität durch Hinzunahme des zweiten Prädiktors

Quelle: Hier findet sich ein Hinweis, in welchem Kapitel (Kap.) des Lehrbuchs bzw. in welchem Abschnitt innerhalb dieses Kapitels (Abs.) das entsprechende Verfahren vorgestellt wird.

Hinweis:

Diese Validitätszusammenstellung von Schmidt und Hunter (1998) kann nicht unkommentiert bleiben:

- Bedingt durch andere gesetzliche Rahmenbedingungen, andere Berufsordnungen usw. sind nicht alle Befunde direkt auf deutsche Verhältnisse übertragbar. So gibt es z.B. in Deutschland keinen regulär in der Praxis eingesetzten „Integrity“-Test.
- Die Autoren haben bei konkurrierenden metaanalytischen Befunden eine „wohlwollende Auswahl“ getroffen, indem sie die jeweils höchste gefundene Validität in die Übersicht aufgenommen haben (vgl. als Beispiel Abschnitt 2.2.3.2 zu persönlichkeitsorientierten Verfahren).
- Die Koeffizienten stammen aus unterschiedlichen Metaanalysen. Dadurch konnten einige Randvariable nicht parallelisiert werden, z. B. die Kriterien, nach denen Primärstudien in die Metaanalyse einbezogen (oder verworfen) wurden.
- Als Berufserfolgskriterium diente in den meisten Fällen eine globale Leistungsbeurteilung durch den Vorgesetzten. Wie Kapitel 15 zeigt, wird damit nicht der gesamte Leistungsbereich abgedeckt.
- Generell dienen die angegebenen Koeffizienten nur als grobe Orientierungsgröße. Zu den Validitätsbefunden nicht-kognitiver Verfahren sei explizit auf die restlichen Abschnitte in diesem Kapitel sowie auf die Kapitel 6 und 7 verwiesen. Die hohe Validität kognitiver Fähigkeitstests setzt voraus, daß keine starke Vorselektion in bezug auf Intelligenz existiert.

Zusammenfassung der aktuellen metaanalytischen Befundlage zur Validität eignungsdiagnostischer Verfahren

Die Ergebnisübersicht von Schmidt und Hunter (1998) sollte nicht unreflektiert übernommen werden.

Studiendarstellung bei Schuler & Höft (2001) im Lehrbuch Personalpsychologie (1. Auflage)

Bitte des Kleingedruckte beachten!

- Übersicht mit selektiver Auswahl, **keine Metaanalyse**
- Hauptsächlich US-Befunde
- Keine Parallelisierung der unterschiedlichen Metaanalysen
- Indirekte Schätzung der inkrementellen Validität
- Empfehlung von Schuler & Höft: „Ergebnisse nur als grobe Orientierung verwenden“

Zwischenzeitliche Aktualisierung

		<i>Schmidt & Hunter (1998)</i>			<i>Update nach Schuler, Höft & Hell (2014)</i>	
Prädiktor		Val	mR	inkrV	Val	Neuere Metaanalyse
Eigenschaft	Allgemeine kognitive Fähigkeitstests	.51			.62	Salgado et al. (2003)
	“Integrity“-Tests	.41	.65	.14	.18	Van Iddekinge, Roth, Raymark & Odle-Dusseau (2012)
	Gewissenhaftigkeitstests	.31	.60	.09	.31	Barrick, Mount & Judge (2001)
	Fachkenntnistests	.48	.58	.07		
	Interessen	.10	.52	.01	.23	Van Iddekinge, Roth, Putka & Lanivich (2011)
Simulation	Probezeit	.44	.58	.07		
	Arbeitsproben	.54	.63	.12	.39	Roth, Bobko & McFarland (2005)
	Situational Judgement Tests				.34	McDaniel, Morgeson, Bruhn Finnegan & Campion (2001)
	Assessment Center	.37	.53	.02	.28	Hermelin, Lievens & Robertson (2007)
Biografie	Graphologie	.02	.51	.00		
	Biografische Daten	.35	.52	.01	.32	Bliesener (1996)
	Unstrukturiertes Einstellungsgespräch	.36	.55	.04	.20	Huffcutt & Arthur (1994)
	Strukturiertes Einstellungsgespräch	.51	.63	.12	.40	Schmidt & Rader (1999)

Anmerkungen: Als Artefakte wurden die Kriteriumsreliabilität und Varianzeinschränkungen korrigiert.

Val: kriteriumsbezogene Validität (Kriterium: Leistungsbeurteilung)

MR: geschätzte kombinierte Vorhersage in einer multiplen Regression

inkrV: inkrementelle Validität, d.h. der Validitätszuwachs durch Hinzunahme des zweiten Prädiktors zstl zum allgemeinen kognitive Fähigkeitstest

„Aktualisierung“ von Schmidt et al. zu S&H (1998) führt zu Diskussionen

Working Paper

The Validity and Utility of Selection Methods in Personnel Psychology: Practical and
Theoretical Implications of 100 Years of Research Findings

Frank L. Schmidt

University of Iowa

In-Sue Oh

Temple University

Jonathan A. Shaffer

West Texas A&M University

October
2016

Achtung!
Kein Peer-Review

* This working paper was prepared as an update to Schmidt and Hunter (1998). We have made it available on Research Gate and SSRN because of repeated requests for such updates from practitioners and researchers from all over the world. We intend to publish a version of this paper later.

Schmidt et al.-Update (2016) im Detail

		Schmidt & Hunter (1998)			Update nach Schmidt, Oh & Shaffer (2016)		
Prädiktor		Val	mR	inkrV	Val	mR	inkrV
Eigenschaft	Allgemeine kognitive Fähigkeitstests	.51			.65		
	“Integrity“-Tests	.41	.65	.14	.46	.78	.13
	Gewissenhaftigkeitstests	.31	.60	.09	.22	.70	.05
	Fachkenntnistests	.48	.58	.07	.48	.65	.00
	Interessen	.10	.52	.01	.31	.71	.06
Simulation	Probezeit	.44	.58	.07	.44	.65	.00
	Arbeitsproben	.54	.63	.12	.33	.65	.00
	Situational Judgement Tests				.26	.65	.00
	Assessment Center	.37	.53	.02	.36	.66	.01
Biografie	Graphologie	.02	.51	.00	.02	.65	.00
	Biografische Daten	.35	.52	.01	.35	.68	.04
	Unstrukturiertes Einstellungsgespräch	.36	.55	.04	.58	.73	.09
	Strukturiertes Einstellungsgespräch	.51	.63	.12	.58	.76	.12

Anmerkungen: Als Artefakte wurden die Kriteriumsreliabilität und Varianzeinschränkungen korrigiert.

Val: kriteriumsbezogene Validität (Kriterium: Leistungsbeurteilung)

MR: geschätzte kombinierte Vorhersage in einer multiplen Regression

inkrV: inkrementelle Validität, d.h. der Validitätszuwachs durch Hinzunahme des zweiten Prädiktors
zstl zum allgemeinen kognitive Fähigkeitstest

Schnelle Resonanz im (inter-)nationalen Netz ...



HOLTMEIER & FRIENDS BLOG

Metaanalyse

Methoden der Personalauswahl

Eine Metaanalyse zur Validität von Personalauswahlverfahren, die jeder Personaler kennen sollte!

Viele unserer Leser kennen die wichtige Meta-Analyse von Schmidt und Hunter, in denen die Validitäten verschiedener Personalauswahlverfahren auf Basis von Metaanalysen bestimmt wurden. Bis heute werden wichtige Entscheidungen für den Personalmanagement basierend auf diesen Ergebnissen getroffen. Nun gibt es jedoch eine Neuauflage (Schmidt et al., 2017), die wir durchaus als *disruptiv* für die Branche bezeichnen dürfen.

Recrutainment Blog

Online-Assessment, Berufsorientierung & Employer Branding



Powered by **CYQUEST**
The Recruitment Company

Berufs- & Studienorientierung | Online-Assessment & SelfAssessment | Employer Branding | Personalmarketing

eRecruiting | Mobile Recruiting | Demografischer Wandel | Fundstücke | Vorträge & Veranstaltungen

Wie gut sagen verschiedene Personalauswahlverfahren eigentlich Berufserfolg voraus? Diese Methoden bringen es wirklich...


© 5. Juni 2017 | Berufsorientierung, eAssessment, Eignungsdiagnostische Tests, eRecruiting, ...

Azubi Recruiting



Predicting Job Performance: Unstructured Interviews Second Only to Cognitive Ability

By Paul Barrett on May 23, 2017



©Photographer.eu/Fotolia

Well, this result has been available on the web in a research paper since 2013, yet I still see many trumpeting structured interviews and personality assessment as superior to the humble unstructured interview.

The [current version](#) of the Schmidt, Oh, and Shaffer working paper is freely available for download from the Social Science Research Network site. It updates the prediction validities of the now famous Schmidt and Hunter (1998) article, listing the validities for a variety of assessment attributes as predictors of job performance.

Wo kommen die neuen Befunde zur Validität von (un-)strukturierten Interviews her?

CHAPTER 12

RETHINKING THE VALIDITY OF INTERVIEWS FOR EMPLOYMENT DECISION MAKING

Implications of Recent Developments in Meta-analysis

In-Sue Oh, Bennett E. Postlethwaite, and Frank L. Schmidt

*Received Wisdom, Kernels of Truth, and Boundary Conditions in
Organizational Studies, pp. 297–329*

Buchkapitel
(2013)

*Achtung!
Kein Peer-
Review*

Oh, Postlethwaite & Schmidt (2013): Wie kommt es zu solchen drastisch geänderten Schätzungen?

- **Neuerung 1:**
Korrektur von angenommenen Publikationsbiases (Nicht-Veröffentlichung von nicht signifikanten Studienergebnissen) mithilfe neuer Schätztechniken (Trim-and-fill-Ansatz von Duval & Tweedie, 2000)

PERSONNEL PSYCHOLOGY
2006, 59, 927–953

PUBLICATION BIAS: A CASE STUDY OF FOUR TEST VENDORS

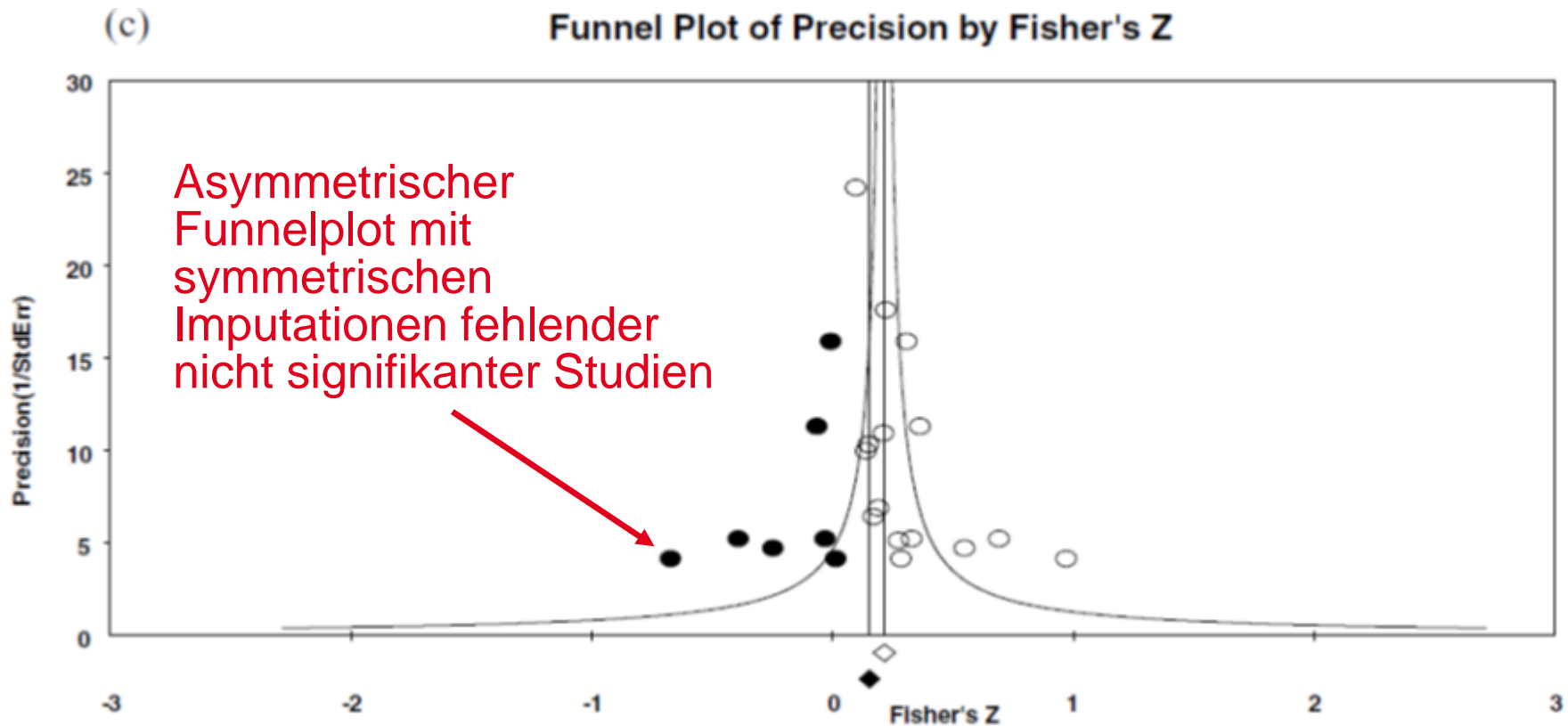
MICHAEL A. McDANIEL
Virginia Commonwealth University

HANNAH R. ROTHSTEIN
Baruch College

DEBORAH L. WHETZEL
U.S. Postal Service

Grundprinzip des Trim-and-fill-Ansatzes

- Auf Grundlage der Verteilungsform der publizierten Primärstudien wird geschätzt, wie die nicht publizierten (weil nicht signifikanten) Primärstudien ausgefallen sein könnte



Oh, Postlethwaite & Schmidt (2013): Wie kommt es zu solchen drastisch geänderten Schätzungen?

- Neuerung 2:
Berücksichtigung von (in-)direkten Varianzeinschränkungen in den Primärstudien der Original-Metaanalyse von McDaniel et al. (1994)
 - **Direkte Varianzeinschränkung** tritt auf, wenn Personen auf der Grundlage des Interviewergebnisses eingestellt wurden und nur für diese Kriteriumsdaten vorliegen
(Korrektur mit Thorndikes (1949) Case II-Formel)
 - **Indirekte Varianzeinschränkung**, wenn zusätzliche Informationen (= weitere eignungsdiagnostische Verfahren) die Selektionsentscheidung beeinflussen und/oder Personen unter verschiedenen Jobangeboten auswählen können
(Korrektur mit Schmidt et al. (2006) Case IV-Formel)

Huffcutt et al. (2014):

Komplexere Herangehensweise an mögliche Varianten der (in-)direkten Varianzeinschränkung

International Journal of Selection and Assessment

Volume 22 Number 3 September 2014

Moving Forward Indirectly: Reanalyzing the validity of employment interviews with indirect range restriction methodology

Allen I. Huffcutt*, **Satoris S. Culbertson**** and **William S. Weyhrauch*****

*Department of Psychology, Bradley University, Peoria, IL 61625, USA. huffcutt@fsmail.bradley.edu

**Management Department, Kansas State University, Manhattan, KS 66506, USA

***Department of Psychology, Kansas State University, Manhattan, KS 66506, USA

Genauere Spezifikation der mögliche Formen der (in-)direkten Varianzeinschränkung

Category	Description	RR Mechanism	No. Studies	%
1	All candidates who were interviewed were hired	None	3	2.3
2	Candidates were hired based on the interview	Direct	3	2.3
3	Selection data were collected from incumbents	Indirect	68	51.1
4	Interview given, but hiring based on another predictor	Indirect	5	3.8
5	Hiring based on a composite of the interview and another predictor	Indirect	33	24.8
6	Hiring based on another predictor first, and then on the interview	Indirect then direct	19	14.3
7	Hiring based on the interview first, and then on another predictor	Direct then indirect	2	1.5
		Total	133	

Note: RR = range restriction.

Huffcutt et al. (2014, p. 298)

Ergebnis nach zusätzlicher Korrektur von (in-)direkten Varianzeinschränkungen

Table 3. Combined results for the criterion-related validity of employment interviews

Category	k	N	ρ
Overall	92	7,389	.67
Level 1	3	592	.20
Level 2	20	2,002	.46
Level 3	15	1,413	.71
Level 4	54	3,382	.70

Note: In the above table, k is the number of studies, N is the total sample size, and ρ is the fully corrected mean validity.

Huffcutt et al. (2014, p. 300)

- Numerisch noch einmal höhere Validitätsbefunde als bei Schmidt et al. (2016)
- Deutliche Validitätssteigerung bereits durch einfache Formen der Standardisierung: Deckeneffekt ab Strukturierungslevel 3 (>= Fragenkatalog, >= dimensionale Bewertungen)

Sackett, Shewach & Keiser (2017): Eine Metaanalyse zur Ehrenrettung des Assessment Centers

Journal of Applied Psychology
2017, Vol. 102, No. 10, 1435–1447

© 2017 American Psychological Association
0021-9010/17/\$12.00 <http://dx.doi.org/10.1037/apl0000236>

RESEARCH REPORT

Assessment Centers Versus Cognitive Ability Tests: Challenging the Conventional Wisdom on Criterion-Related Validity

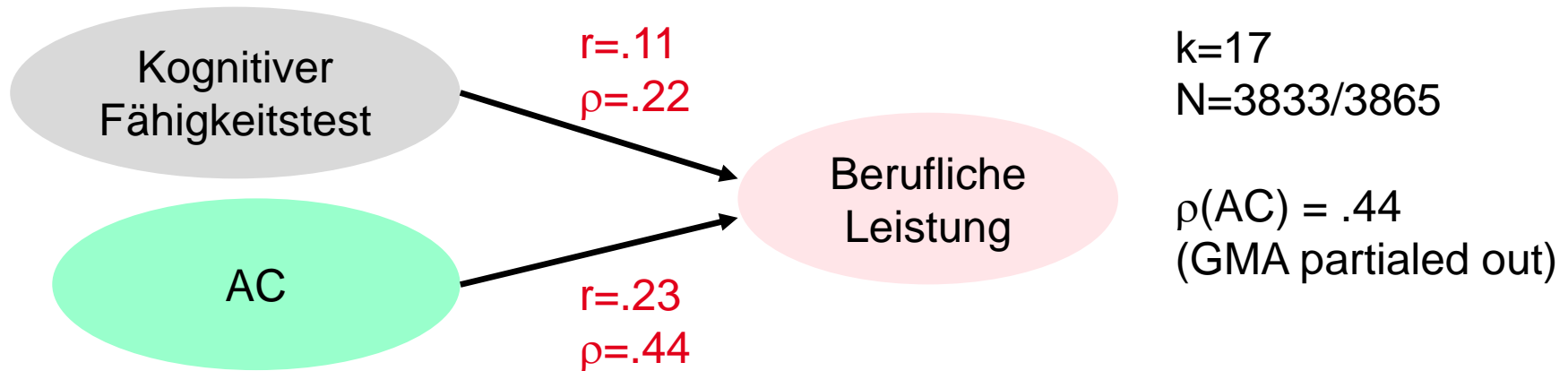
Paul R. Sackett, Oren R. Shewach, and Heidi N. Keiser
University of Minnesota

- Grundidee: Nur Berücksichtigung von Studien, in denen AC und Kognitive Fähigkeitstests parallel nebeneinander eingesetzt wurden.

Dadurch:

- Homogenisierung der Teilnehmerstichprobe und Konstanthaltung des Einsatzkontextes (=gleiche Auswahlstufe)
- Zusammenhangsanalyse mit identischen Berufserfolgskriterien

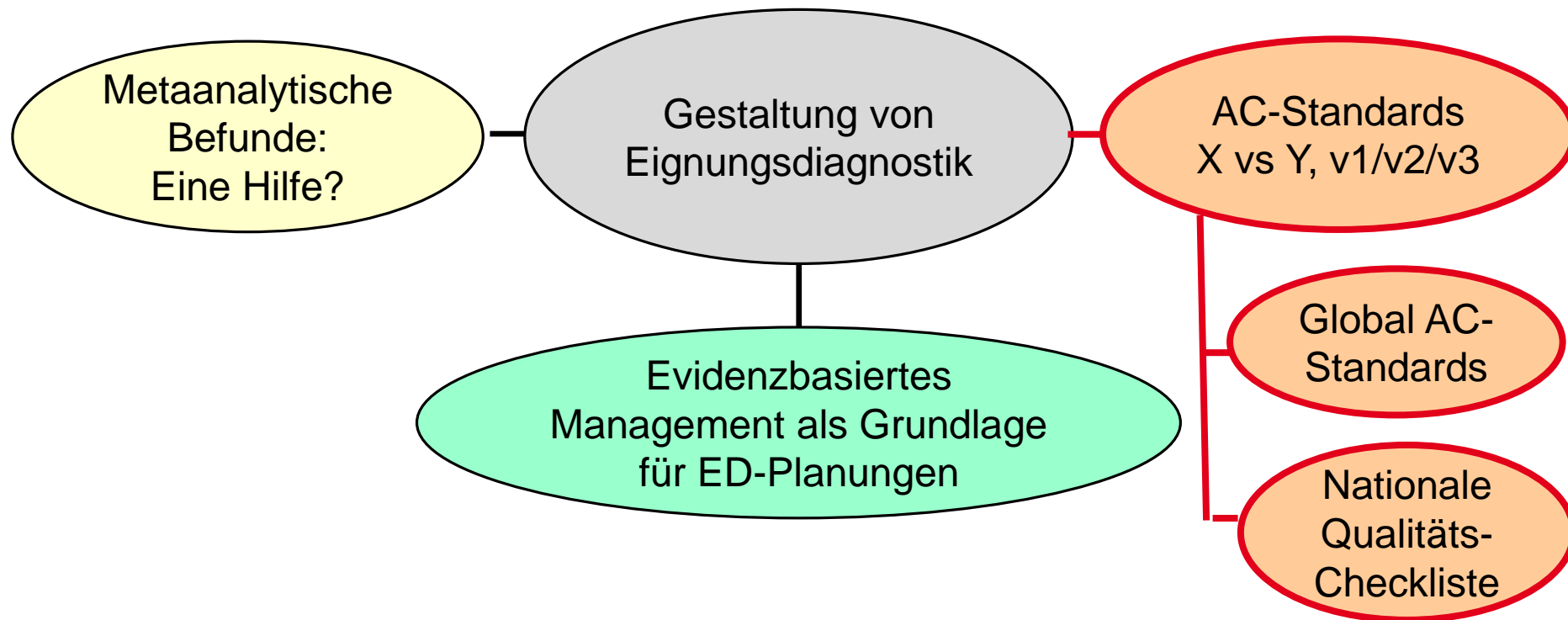
Bei einem kombinierten Einsatz erzielen ACs bessere Vorhersagewerte als kognitive Fähigkeitstests



Mögliche Gründe für die Befundlage:

- Varianzeinschränkung der kognitiven Fähigkeitswerte durch Bedeutsamkeit bei der vorlaufenden Teilnehmerauswahl
Für eine vollständige Kompensation müsste die vorlaufende Selektionsquote allerdings 5% betragen, um die üblichen metaanalytischen Befundlage zu replizieren.
- Berufliche Leistungskriterien sind in den betreffenden Studien eher breiter und kontextualer definiert

- Gesamthafte Validitätsübersichten (Schmidt & Hunter, 1998; Schmidt et al., 2016) sind irreführend, da die selektiv berichteten Befunde nicht hinsichtlich der durchgeführten Artefaktkorrekturen angeglichen sind.
- Neuere Metaanalysen (>2006) implementieren elaboriertere Korrekturtechniken (Varianzeinschränkung, Publikationsbias), die bei älteren Studien noch fehlen: Deren Aussagegehalt muss dadurch z.T. kritisch hinterfragt werden.
- Sinnvoller: Themenspezifische Metanalysen detailliert betrachten.
- Die Metaanalyse von Sackett et al. (2017) hinterfragt deutlich, inwieweit eine unspezifische Validitätsgeneralisierung Relevanz für die reale Anwendungssituation hat.



Standards als Weg der Qualitätssicherung: Warum?

Wie begründet beispielsweise die DIN 33430 ihr Existenzrecht?

- Qualitätsstandards formulieren Regeln und Prinzipien für einen Anwendungsbereich.
- Gütekriterien wie Validität und Reliabilität lassen sich in der Praxis z.B. aufgrund zu geringer Fallzahlen und/oder zu kurzer Verweildauer des Personals häufig nicht direkt bestimmen.

Aus Qualitätsstandards lassen sich konkrete Qualitätsforderungen für den Einzelfall ableiten, so dass

- die Erfüllung oder Nichterfüllung im Sinne eines unmittelbaren Feedbacks bestimmt werden kann und
- Maßnahmen zur Qualitätsoptimierung ergriffen werden können

Qualitätsstandards sind somit Leistungsstandards, die ein gerichtetes Verhalten motivieren.

(Kersting, 2017)

Am Anfang standen: The International Standards

Journal of Management
Vol. 41 No. 4, May 2015 1244–1273
DOI: 10.1177/0149206314567780
© The Author(s) 2015
Reprints and permissions:
sagepub.com/journalsPermissions.nav

Public Guidelines for Assessment Center Operations

Guidelines and Ethical Considerations for Assessment Center Operations

International Taskforce on Assessment Center Guidelines

Keywords: *assessment center; assessment centers; assessment; behavioral assessment; assessment technology; assessors; validation; assessor training; developmental assessment centers*

Table of Contents

- I. Purpose
- II. History of Guidelines
- III. Assessment Center Defined
- IV. Non-Assessment Center Activities
- V. Assessment Centers for Different Purposes
- VI. Assessment Center Policy Document
- VII. Assessor Training
- VIII. Training and Qualifications of Other Assessment Center Staff
- IX. Validation Issues
- X. Technology
- XI. Ethics, Legal Compliance, and Social Responsibility

Taskforce Members: Deborah E. Rupp, Purdue University, USA, Chair; Brian J. Hoffman, University of Georgia, USA, Co-Chair; David Buckhof, Deloitte, South Africa, Co-Chair; William Byham, Development Dimensions International, USA; Lynn Collins, BTS, USA; Alyssa Gibbons, Colorado State University, USA; Shinichi Hirose, International University of Japan, Japan; Martin Kleinmann, University of Zurich, Switzerland; Jeffrey D. Rudolph, University of Maryland, USA; Martin Lunik, Pitsigle, USA; Duncan J. R. Jackson, Birkbeck, University of London, UK; Myungjoon Kim, Assesa, South Korea; Filip Lievens, Ghent University, Belgium; Deon Meiring, University of Pretoria, South Africa; Klaus G. Meichers, Universitäts Ufm, Germany; Vina G. Penla, Daya Dimensi, Indonesia; Dan J. Puka, Human Resources Research Organisation, USA; Nigel Posak, Assessment and Development Consultants, UK; Doug Reynolds, Development Dimensions International, USA; Sandra Soltébusch, LEMASA, South Africa; John Scou, APTMetrics, USA; Svetlana Simonenko, Desec, Russia; George Thornton, Colorado State University, USA

Corresponding author: Deborah Rupp, Purdue Department of Psychological Sciences, Purdue University, 703 Third Street, West Lafayette, IN 47907, USA.

E-mail: ru.ppd@purdue.edu

Guidelines and Ethical Considerations for Assessment Center Operations

entwickelt von der
International Taskforce

- 1st version: 1975
- 6th edition: 2014

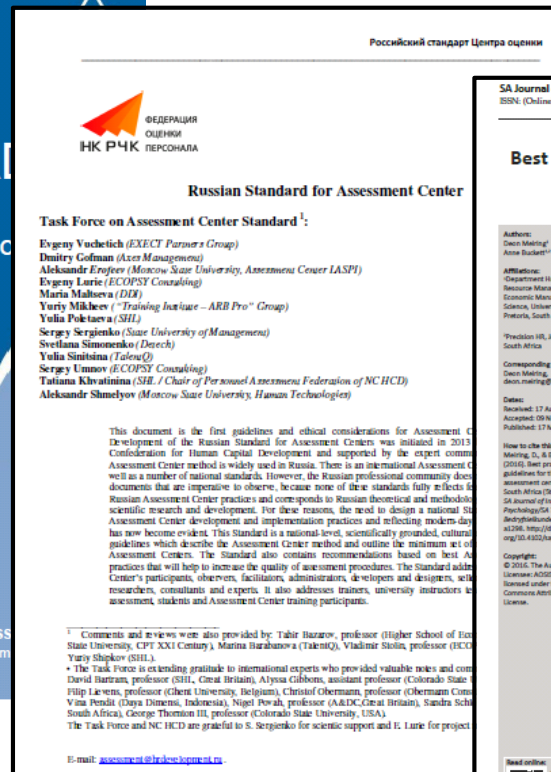
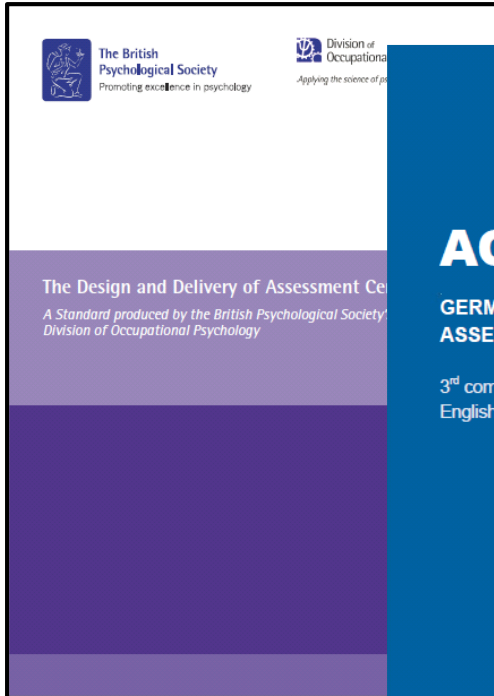
Weitere nationale Standards

British
Guidelines 2005
Standards 2015

German
1st version 1992
3rd version 2016

Russian
1st version 2014

South-African
1st version 1987
5th version 2016



+ Indonesian version 2002
+ Swiss version 2007 + Turkish version 2015

Warum gibt es überhaupt nationale AC-Standards?

International Guidelines (p. 1266)

The international guidelines present broad, universal guidance to which all nations (and national standards) need comply, whereas the national standards provide more detailed recommendations based on local legal and cultural environments.

Sürmeli (2015): (Erweiterte) Liste möglicher Gründe

- Länderspezifische kulturelle Spezifika
- Rechtliche Rahmenbedingungen vor Ort
- Akademische Infrastruktur & Forschungseinrichtungen
- Vorherrschende Unternehmenskulturen
- Ziele von Talentmanagement (Auswahl vs. Entwicklung)
- (Vor-)Geschichte von Tests&Assessments im Land/in den Organisationen
- Kooperation mit Ländern, in denen bereits nationale Standards/Erfahrungen existieren
- Schlüsselliteratur (dominante Lehrbücher, prominente Organisationen)
- Die Entwicklung nationaler Standards ist identitätsstiftend für nationale Vereinigungen

Zielsetzung des Projektes „Globale AC-Standards“

(Initiiert von Sürmeli & Höft)

- **Identifikation von “Konsensgebiete” in allen (inter-)nationalen Standards:** Was sind allgemein anerkannte Kriterien bei der Gestaltung von Acs und welche Qualitätskriterien müssen nicht diskutiert werden (‘Muss’ Kriterien)? Welche Aspekte sind wünschenswert, können aber nicht immer implementiert werden (‘Soll’-Kriterien)?
- **Diskussion nationaler Kriterien:** Was sind spezifische Forderungen in den nationalen Standards, die von den internationalen Vorgaben Abweichen? Was sind mögliche Ursachen dafür?

High Agreement in Most Standards – 10-9-8 Points

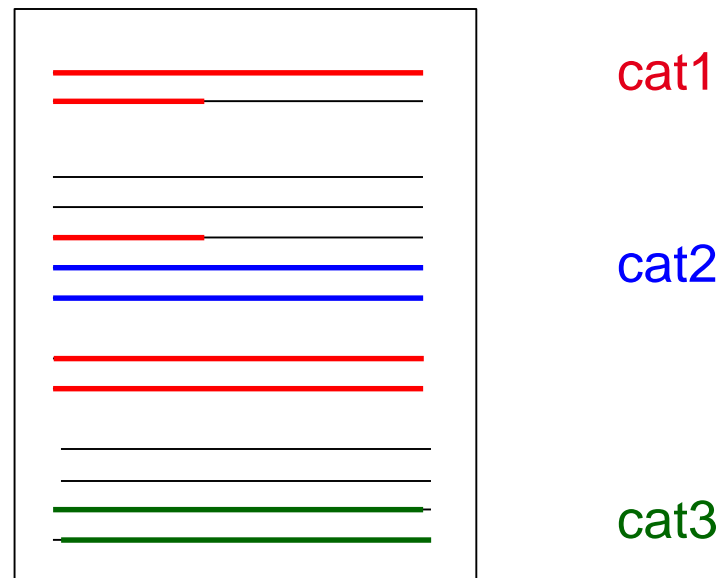
CRITERIA – 10-9-8 Points	International 2015	German 2016	Russia 2013	South Africa 2016	United Kingdom 2015	Turkey 2015
Assessment Center Purposes (AC & DC)	●	○	●	●	○	●
Validation Issues	●	●	●	●	●	-
Technology	●	-	●	●	●	●
Glossary of Relevant Terms	●	●	●	-	●	●
Multiple Assessors	●	●	○	●	○	●
Using psychometric tests in Acs	○	●	●	○	●	●
Mode & Content of feedback	○	●	○	○	●	●
Evaluating the Centre process	○	●	●	●	●	-
Assesse / Participant Rights	●	○	●	●	●	-
Non-Assessment Center Activities	●	-	●	●	-	●
Making the decision to use an AC	○	○	●	-	●	●
Development Plans	○	●	○	●	-	●
Recency of Training and Experience	●	●	○	○	●	-
Monitoring of Assessor Performance	●	●	○	○	●	-
Centre Designers	-	-	●	●	●	●

Zielsetzung der ersten Teilstudie

- Schwächen der globalen Sürmeli-Ratings:
Unklare Relibailität und Validität, die die Beurteilungsgrundlage der Ratings unbekannt ist. Keine Beobachterübereinstimmungsmaße.

Ziele:

- Qualitative Hinterlegung der holistischen Urteile
- Identifikation relevanter Textstellen für die vorliegende Kategorisierung



Screenshot der MaxQDA-Kodiersoftware HdBA

The screenshot displays the MaxQDA software interface with the following components:

- Start Menu:** Includes options like 'Neues Projekt', 'Projekt öffnen', 'Liste der Dokumente', 'Liste der Codes', 'Dokument Browser', 'Liste der Codings', 'Logbuch', 'Benutzerverwaltung', 'Projekt speichern unter', 'Projekt anonymisiert speichern', 'Projekt aus aktivierten Dokumenten', 'Projekte zusammenführen', 'Exchange-Datei öffnen', 'Exchange-Datei exportieren', and 'Externe Dateien'.
- Document Browser:** Shows a tree view of documents under 'Dokumente' and 'Sets'. The 'RussianAC-Standards' document is selected.
- Code List:** A hierarchical list of codes such as 'Purpose', 'Technology', 'Training and Qualifications of other Assessment C...', 'Validation Issues', 'Essential Elements of an AC', 'Assessor Training', 'Behavioral Classification', 'Competency Dimension / AC Component Matrix', 'Data Integration', 'Job Analysis before AC', 'Multiple Assessment Techniques', 'Multiple Assessors', 'Recording and Scoring of Behaviors', 'Simulations/Exercises', 'Standardization', 'Preparation for an AC', 'AC Program Development', 'Accommodating of Disabled Persons', 'Briefing Candidates', 'Contingency Planning', 'Decision to use AC', 'Development and Selection of Simulation Exercises', 'Environment', 'Identification of Behavioral Constructs/Compete...', 'Job Analysis', 'Materials', 'Obtaining Consent of Participants', 'Participant Registration and Preparation', 'Stakeholder Involvement', 'Technical Support', 'Types of Simulations', 'Using Psychometric Tests in ACs', and 'AC Delivery'.
- Document Viewer:** Displays the text of the 'RussianAC-Standards' document. The text includes sections on 'Multiple Assessors', '2.3. AC purpose and objectives', and '2.4. Assessment method'. A 'Codings' window is overlaid on the text, showing a table of codes applied to the selected text.
- Codings Window:** A pop-up window titled 'Codings' showing 'Code: Multiple Assessors' and '6 Codings aus 5 Dokumenten und 0 Dokumentgruppen'. It contains a table with the following data:

Dokumentname	Code
<input type="radio"/> InternationalACGuidelines_JOM	Multiple Assessors
<input type="radio"/> AkAC Standards 2016_englishv0-4	Multiple Assessors
<input type="radio"/> AkAC Standards 2016_englishv0-4	Multiple Assessors
<input type="radio"/> SouthAfricaStandards	Multiple Assessors
<input checked="" type="radio"/> RussianAC-Standards	Multiple Assessors
<input type="radio"/> BrithACGuidelines	Multiple Assessors

The text in the document viewer includes:

psychometric tests, questionnaires.

6. It is necessary to use interactive simulation exercises (exercises reproducing the most essential behavioral aspects of collaborative work in a job).

7. **Assignment of scores on each of the competencies should be based on observations of at least two expert observers who have taken a special training course.**

8. Determination of the demonstrated level of a competency is the main result of observers' work.

9. All aggregate scores should be agreed upon during data integration by means of consensual discussion of observers or obtained by a justified statistical integration process.

2.3. AC purpose and objectives

The main purpose of ACs is to estimate potential success within a job (professional success). Jobs may be determined specifically (e.g. a position within an organization) or by type (e.g. a certain management level).

Participants' job performance process. Even in the case of evaluated not prior results.

Today ACs are used to manage:

1. Selection (e.g. hiring, promotion)
2. Determination of individual (PDPs), career guidance
3. Determination of direction building a managerial
4. Training within the context adaptation). It is important.

AC objectives are not limited to the same time, all the AC objectives should not impair participants' performance.

2.4. Assessment method

An assessment is not an assessment:

1. It includes only one aspect
2. It does not contain multiple
3. It does not contain interactive aspects of collaborative
4. It only consists of a test
5. It only includes one interview
6. It is based on the judgment used;
7. It does not include a decision and several observers.

3.1. Preparation for an AC

3.1.1. Making the decision to use an AC

Kodierungssystem der Inhaltsanalyse

- Das erste Kodierungsschema von Surmeli (2015) war stark am Grundaufbau der Internationalen Standards orientiert
- V2.0: 10 Kategorien mit 96 Subkategorien:

1. *General topics*
2. *Essential elements of an AC*
3. *Preparation for an AC*
4. *AC delivery*
5. *Issues related to AC results*
6. *Post-center review*
7. *Training issues*
8. *AC roles*
9. *Ethical issues*
10. *Legal issues*

Certification
Monitoring
Recency
Training
Training I

- AC policy document
- Assessment Center purpose (AD and DC)
- Cross-cultural issues
- Definition of AC
- Ethics, legal compliance, social responsibility
- Glossary of relevant terms
- History of guidelines
- Index
- Intended audience
- Key sources and references
- Non-assessment center activities
- Purpose
- Technology
- Training and qualifications
- Validation issues

atrix

Ergebnisse zur Kodierungsabdeckung

		International	British	German	Russian	South-African					
1	General topics	67761	65,6	45514	24,7	19414	33,1	24232	36,5	66529	74,6
2	Essential elements of an AC	25537	24,7	16653	9,0	12437	21,2	7582	11,4	13090	14,7
3	Preparation for an AC	12450	12,0	38718	21,0	10001	17,0	13705	20,7	8234	9,2
4	AC delivery	4572	4,4	14821	8,1	7822	13,3	8261	12,5	2119	2,4
5	Issues related to AC results	627	0,6	7032	3,8	6703	11,4	6453	9,7	2496	2,8
6	Post-center review	931	0,9	16677	9,1	4738	8,1	2597	3,9	480	0,5
7	Training issues	9798	9,5	13147	7,1	1649	2,8	8472	12,8	1648	1,8
8	AC roles	4492	4,3	13770	7,5	1397	2,4	6840	10,3	6511	7,3
9	Ethical issues	5535	5,4	4928	2,7	0	0,0	1556	2,3	7258	8,1
10	Legal issues	1392	1,3	40740	22,1	255	0,4	1842	2,8	7411	8,3
	Total length	98640		175506		56174		58700		83993	

Similarity matrix measuring category pattern equivalence
(proximity scale, simple match, all 96 subcategories)

	International	British	German	Russian
International	1			
British	0,6	1		
German	0,57	0,66	1	
Russian	0,64	0,74	0,67	1
South-African	0,72	0,72	0,63	0,69

Details about subcategories in cat5 „Issues related to AC results“

	International		British		German		Russian		South-African	
5_Customer Feedback							426	0,6		
5_Development Plans			1267	2,2			406	0,6	144	0,2
5_Feedback for AC Participants			3496	1,9	2853	4,9	1600	2,4	282	0,3
5_Feedback Providers			393	0,2			213	0,3		
5_Follow-Up Sessions	53	0,1							764	0,9
5_Linkage between AC and HR Processes			409	0,2	561	1,0	913	1,4	66	0,1
5_Mode and Content of Feedback			1519	0,8	2022	3,4	1434	2,2	376	0,4
5_Usage of AC Results	574	0,6	1215	0,7			1461	2,2	864	1,0

Exemplarische Inhaltsanalyse: Ratio of assessors to assessees

International (p. 1250) = South African (p. 5)

The maximum ratio of assessees to assessors is a function of several variables,

The ratio of assessees to assessors should be minimized where practicable in the interests of reducing cognitive load,

British
(p. 29)

5.21 There shall be at least one Assessor for every three Participants as a minimum. Ideally there should be at least one Assessor for every two Participants. Where the Centre design allows a lower ratio the reasons why this is appropriate should be explicitly noted in the Centre documentation.

German
(p. 11)

The actions of the participants in the behavior simulations and in the interview are recorded and assessed independently of each other by at least two assessors or interviewers who have been trained for this task and have specifically prepared for the respective procedure, exercise, and assessment system. The quality of assessment of the behavior in group tasks should not be impaired by the participation and observation of too many candidates (max. 5-6 participants in one exercise, ideally 2-3 max to observe).

Russian (p. 8)

The observation plan should be made in such a way that a participant's behavior during the AC on the whole is assessed by at least two observers. Observation of a participant in the course of a specific exercise can be carried out by a single observer.

Schlussfolgerung & Ausblick

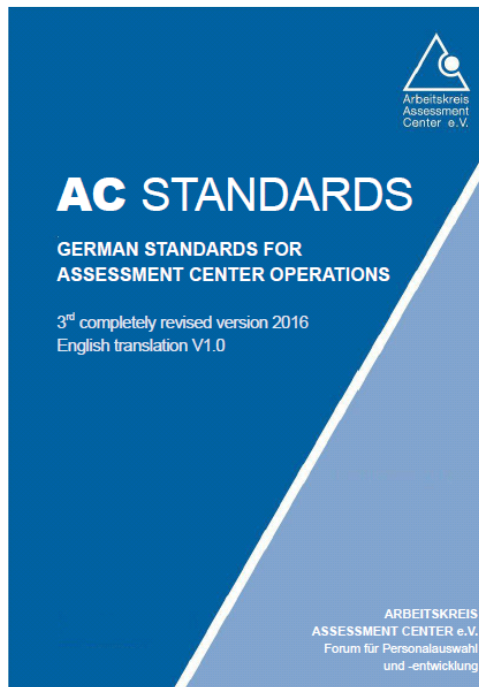
- Qualitative Textanalysen sind ein hilfreiches Tool, um globale Ratings mit konkreten Textaussagen zu verknüpfen
- Die vorgestellten Textlängenanalysen (und nicht berichtete Keyword-Suchen) können nur einen ersten Schritt bei den Analysen darstellen. „Echte“ qualitative Unterschiede bei inhaltlichen Forderungen werden nicht erfasst.

Nächste Schritte

- Überprüfung der Kodierungen und Finalisierung, um das bestehende MaxQDA-Projekt als Referenzierungssystem nutzen zu können.
- Analyse ausgewählter Standards (International/Deutsch/Russisch) auf Einzelaussagenebene, um qualitative Unterschiede zu identifizieren (EAWOP-Präsentation im Mai 2019 in Turin)

Analyse deutscher Standardtexte

AkAC Standards v3.0

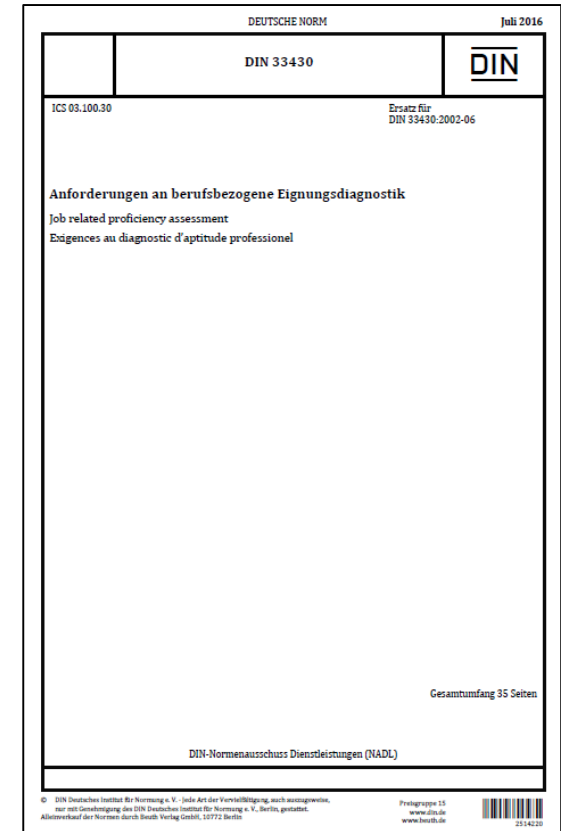


?



- Inhaltliche Abdeckung
- Formuliere Muss/Soll-Anforderungen

DIN33430 v2.0



Beide Standards/Normen wurden 2016 als Überarbeitungen veröffentlicht und beanspruchen beide, Qualitätskriterien für eignungsdiagnostische Prozesse zu definieren

Vorgehensweise: Textabdeckung und Anforderungsableitung

- Für die Entwicklung des Kategoriensystems dient die 10-Phasen-Struktur der AC-Standards. Unterhalb davon wurden insgesamt 63 Subkategorien formativ abgeleitet
- **Analyseebene 1:** MaxQDA-Markierungen relevanter Textabschnitte

Code	Code-Titel	AkAC		DIN	
		#	%	#	%
P0	Allgemeine oder redaktionelle Hinweise	1734	2,82	7033	6,94
P01	Ziele der Standards bzw. der DIN-Norm	391	0,64	1959	1,93
P02	Infos zur Organisation (AkAC/DIN)	1347	2,19	181	0,18
P03	Historie und Änderungen	1343	2,18	1212	1,2
P04	Anwendungsbereiche der Norm/der Standards			1276	1,26
P05	Datenschutz	79	0,13	857	0,85
P06	Berücksichtigung von Behinderungen			2389	2,36
P1	Auftragsklärung und Einbindung	4651	7,56		
P11	Auftragsklärung	2667	4,33	1480	1,46
P12	Integration mit anderen Personalmaßnahmen	877	1,43	50	0,05
P13	Kommunikation	1256	2,04	137	0,14
P14	Verantwortlichkeit Eignungsdiagnostiker			1684	1,66
P15	Hinweise zur Ausschreibung von ED-Prozessen/Verfahren			2972	2,93
P151	Hinweise zur Ausschreibung von ED-Prozessen			3813	3,76
P152	Hinweise zur Ausschreibung von ED-Verfahren			2578	2,54

Erfasste Themengebiete

	Prozentuale Textabdeckung	
	AkAC	DIN
Textumfang (in Zeichen)	59.100	95.381
P0 Allgemeine oder redaktionelle Hinweise	8,0	14,7
P1 Auftragsklärung und Einbindung	15,4	12,6
P2 Festlegung eines Anforderungsprofils	13,7	6,5
P3 Verfahrensauswahl und -entwicklung	17,8	37,8
P4 Auswahl/Vorbereitung Durchführungsbeteiligte	13,4	18,4
P5 Vorauswahl/Vorbereitung Teilnehmer	11,4	1,3
P6 Vorbereitung/Durchführung des Verfahrens	11,6	4,1
P7 Datengewinnung und Bewertung	14,3	2,0
P8 Datenintegration und Ergebnisermittlung	13,5	2,3
P9 Feedback und Folgemaßnahmen	13,3	4,5
PX Evaluation	14,6	4,0
PG Begriffserklärung/Glossar	22,0	9,5

Durch Doppelzuweisung ergeben sich in der Summe >100%

Vorgehensweise: Textabdeckung und Anforderungsableitung

- **Analyseebene 2:** Aus dem Fließtext wurden Soll/Muss-Kriterien für die Gestaltung eignungsdiagnostischer Prozesse abgeleitet

			AkAC		DIN	
			#	%	#	%
P05	Datenschutz	M			4	1,82
P06	Berücksichtigung von Behinderungen	M			4	1,82
P1	Auftragsklärung und Einbindung	M	8	4,734		0,00
		S	3	1,775		0,00
P11	Auftragsklärung	M	2	1,183	4	1,82
		S			1	0,45
P12	Integration mit anderen Personalmaßnahmen	M			1	0,45
P13	Kommunikation	M	3	1,775		0,00
		S			1	0,45
P14	Verantwortlichkeit Eignungsdiagnostiker	M			2	0,91
P151	Hinweise zur Ausschreibung von ED-Prozessen	S			6	2,73
P152	Hinweise zur Ausschreibung von ED-Verfahren	M			2	0,91
		S			2	0,91
P2	Festlegung eines Anforderungsprofils	M	5	2,959	1	0,45
		S	1	0,592		0,00
P21	Anforderungsanalyse durchführen	M	4	2,367	2	0,91
		S	2	1,183	1	0,45
P22	Anforderungsprofil erstellen	M	6	3,55	2	0,91
		S	2	1,183	2	0,91

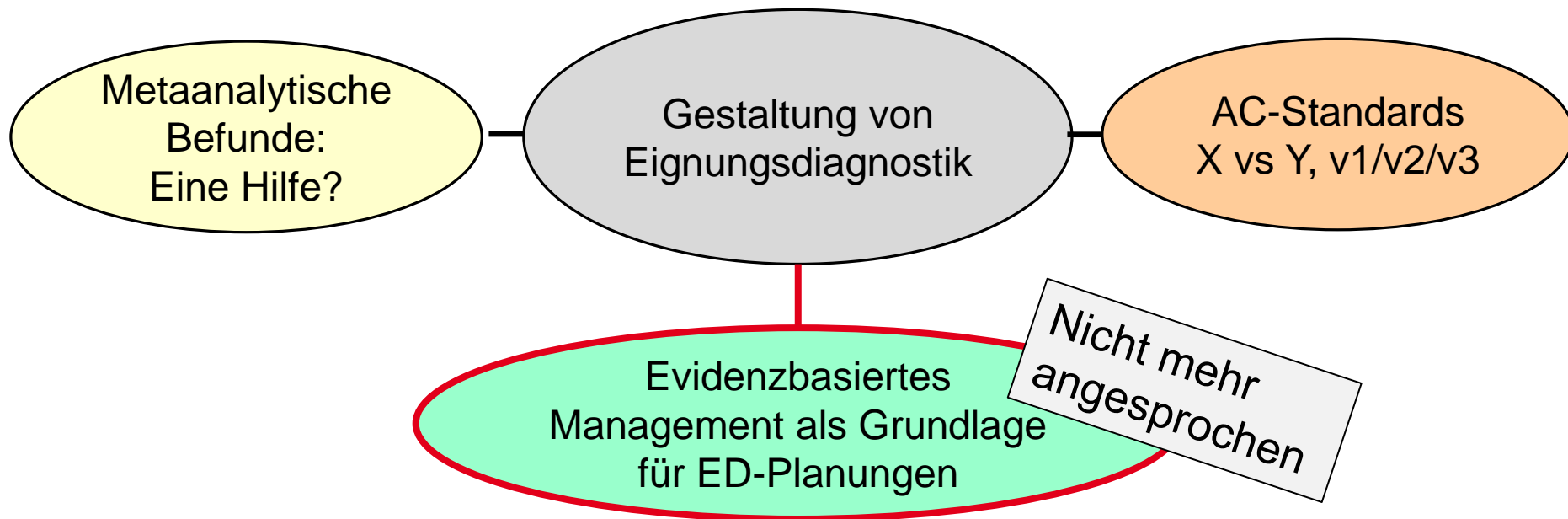
Formulierte Muss/Soll- Anforderungen

		AkAC		DIN	
		#	%	#	%
P0 Allgemeine oder redaktionelle Hinweise	M		0,0	8	3,6
	S		0,0		0,0
P1 Auftragsklärung und Einbindung	M	13	7,7	9	4,1
	S	3	1,8	10	4,5
P2 Festlegung eines Anforderungsprofils	M	15	8,9	5	2,3
	S	5	3,0	3	1,4
P3 Verfahrensauswahl und -entwicklung	M	21	12,4	61	27,7
	S	3	1,8	37	16,8
P4 Auswahl/Vorbereitung Durchführungsbeteiligte	M	13	7,7	19	8,6
	S	4	2,4	1	0,5
P5 Vorauswahl/Vorbereitung Teilnehmer	M	5	3,0	4	1,8
	S	7	4,1		0,0
P6 Vorbereitung/Durchführung des Verfahrens	M	14	8,3	15	6,8
	S	3	1,8	3	1,4
P7 Datengewinnung und Bewertung	M	13	7,7	12	5,5
	S	1	0,6	1	0,5
P8 Datenintegration und Ergebnisermittlung	M	6	3,6	6	2,7
	S	7	4,1		0,0
P9 Feedback und Folgemaßnahmen	M	17	10,1	10	4,5
	S	6	3,6	1	0,5
PX Evaluation	M	7	4,1	6	2,7
	S	6	3,6	9	4,1
		169		220	

Anstehende Arbeitspakete:

1. Erstellung einer gemeinsamen Muss/Soll-Liste AkAC/DIN/..., dabei Reduktion redundanter Anforderungen, Integration konfundierter Kriterien unterschiedlicher Abstraktionsebenen usw.
2. Abgleich mit weiteren relevanten & sichtbaren Kriteriensammlungen
3. Beurteilung der Einzelkriterien von Personalverantwortlichen zu Bedeutsamkeit, Relevanz für die Gesamtqualität, Erfüllungsgrad usw. bei der eigenen eignungsdiagnostischen Arbeit
4. Entwicklung einer Qualitätscheckliste mit einem kriteriums- und normorientiertem Auswertungsschlüssel

Geplante Anwendungsgebiete: Online-Selbstbeurteilung, Hilfsmittel für Evaluationsprojekte von AkAC/FA-Mitgliedern, Grundlage für eine mögliche Zertifizierung



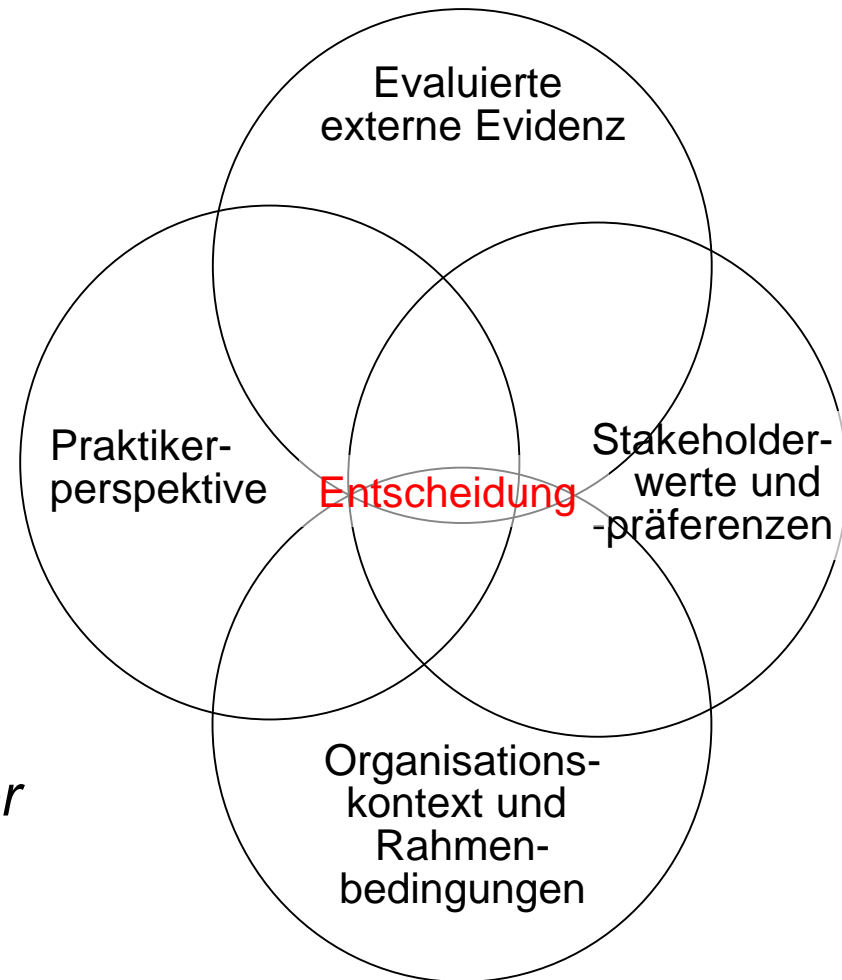
Was ist Evidenzbasiertes Management?

(nach Briner, Denyer, Rousseau, 2009)

**Nicht mehr
angesprochen**
University

Evidenzbasiertes Management stützt sich bei der Entscheidungsfindung auf die gewissenhafte, explizite und stringente Nutzung von vier Informationsquellen:

- *Expertise und Einschätzungen von Praktikern,*
- *Evidenz abgeleitet aus den lokalen Rahmenbedingungen,*
- *Kritische Evaluation der am ehesten passenden wissenschaftliche Befunde*
- *Berücksichtigung der Perspektive der durch die Entscheidung betroffenen Personen*



Typische Fragen bei einer EbM-Entscheidung



Expertise und Einschätzungen von Praktikern

- Habe ich X bereits vorher gesehen? Wo? Wann? Wie?
- Was sind m.E. mögliche Ursachen und Lösungen?
- Wie wurde in der Vergangenheit (erfolgreich) mit X umgegangen?

Evidenz lokaler Rahmenbedingungen

- Wie sieht die genaue Faktenlage aus?
- Welche besonderen Bedingungen liegen vor?
- Welche Kosten/Nutzen-Relation weisen mögliche Interventionen auf?

Typische Fragen bei einer EbM-Entscheidung



Kritische Evaluation der am ehesten passenden wissenschaftliche Befunde

- Ist die Situation ungewöhnlich im Vergleich zum Durchschnitt?
- Welche Aussagen/Empfehlungen ergeben sich auf Basis systematisch zusammengestellter wissenschaftlicher Befunde?
- Inwieweit können die allgemeinen Befund hier angewandt werden?

Perspektive der betroffenen Personen (Stakeholder)

- Wie beurteilen involvierte Personen die möglichen Interventionsalternativen?
- Welche Alternativlösungen haben Andere?

Evidenzbasierte Entscheidungsfindung im (betrieblichen) Alltag



- Die Grundqualifikation der Bewerber/-innen hat sich verschlechtert. Wie kann ich mit den geänderten Rahmenbedingungen umgehen?
- Betriebliche Entscheidungsträger kritisieren das bisherige etablierte Auswahlssystem („zu teuer“, nicht innovativ“). Wie soll ich vorgehen?

FAZIT (vgl. Pfeffer & Sutton, 2006):

Evidenzbasiertes Management versucht, die Entscheidungsfindung bei der Unternehmensführung durch Rückgriff auf bestmögliche wissenschaftliche (wenn möglich empirische) Methoden und (Evaluations-)Befunde zu stützen und abzusichern.

Fragen?

Kontakt: stefan.hoeft@hdba.de