

Einsatz und Auswahl von Persönlichkeitstests in der Personalauswahl und die Research-Practice Gap

Dissertation

zur Erlangung des akademischen Grades eines

Doktors der Philosophie

an der Fakultät HW

Bereich Empirische Humanwissenschaften

der Universität des Saarlandes

vorgelegt von

Janina Diekmann

aus Ostercappeln

Saarbrücken, 2019

Dekan:

Prof. Dr. Stefan Strohmeier

Berichterstatter:

Prof. Dr. Cornelius J. König, Universität des Saarlandes

Prof. Dr. Frank Spinath, Universität des Saarlandes

Tag der Disputation:

28. Januar 2019

DANKSAGUNG

Die Fertigstellung dieser Arbeit habe ich vielen Menschen zu verdanken, die mich in den vergangenen Jahren auf vielerlei Weisen unterstützt haben und ohne die ich möglicherweise nie fertig geworden wäre. Mein Dank gilt ...

... an erster Stelle meinem Doktorvater Prof. Dr. König für die umfassende Betreuung während der gesamten Promotionsphase – von der Themenfindung bis zum Abschluss. Ich verdanke dir viel Unterstützung und zahlreiche hilfreiche Anregungen. Besonders bedanken möchte ich mich für deine Geduld mit der untreuen Seele, die deinen Lehrstuhl noch vor Beendigung der Promotion verlassen hat, um in die weite Welt zu ziehen!

... Herrn Prof. Dr. Spinath für Ihre Bereitschaft, die Zweitbegutachtung meiner Arbeit zu übernehmen.

... allen Mitarbeitern des Lehrstuhls für Arbeits- und Organisationspsychologie, die ihr stets als Sparringspartner zur Verfügung standet und gute Laune in die Arbeitseinheit gebracht habt.

... allen Co-Autoren, Hilfskräften und Praktikant/innen für eure Unterstützung bei der Datenerhebung und -auswertung.

... den vielen Menschen, die an meinen Befragungen teilgenommen haben. Danke, dass Sie sich die Zeit für meine Forschung genommen haben.

... meinen Eltern. Ohne euch wäre nichts davon möglich gewesen. Ihr habt mich unterstützt ohne meine Entscheidungen in Frage zu stellen und ihr seid immer da, wenn ich euch brauche. Ich mache es kurz: Ihr seid einfach die besten Eltern der Welt!

... vier ganz besonderen Freunden. Manu, ohne dich hätte ich wahrscheinlich schon aufgegeben, bevor ich wirklich angefangen habe. Danke für fünf wunderschöne, lustige, wahnsinnige gemeinsame Jahre im Saarland. André und Domenico – ohne Worte. Anne, danke für dein offenes Ohr zu meinem Gejammer und den letzten „Tritt“, der mich nach Berlin geführt hat.

ALLGEMEINE ZUSAMMENFASSUNG

Die Auswahl des „richtigen“ Personals ist für Unternehmen ein essentielles Erfolgskriterium und schon lange auch ein zentrales Thema wissenschaftlicher Forschung. Dabei gehen die Meinungen bezüglich des Nutzens von Persönlichkeitsverfahren in der Personalauswahl auseinander. Häufig wird aufgrund geringerer Validitätswerte der Einsatz von Persönlichkeitstests nicht empfohlen, dennoch werden diese in der Praxis verwendet. Hier gibt es eine so genannte Research-Practice Gap, d. h. wissenschaftliche Empfehlungen und Management-Praxis weichen voneinander ab. Es ist wenig bekannt darüber, welche Verfahren in der Praxis tatsächlich zum Einsatz kommen, auch wenn immer wieder einzelne Verfahren, wie z. B. der MBTI genannt werden, die aus wissenschaftlicher Sicht weniger im Fokus stehen, wenn es um Personalauswahl geht. Auch hier zeigt sich somit möglicherweise eine Research-Practice Gap. Darüber hinaus ist unklar, welche Kriterien Personaler bei der Auswahl eines Persönlichkeitsverfahrens zum Zwecke der Personalauswahl heranziehen. Die vorliegende Dissertation adressiert diese offenen Fragen, untersucht die Anwendungspraxis von Persönlichkeitstests in der Personalauswahl, leitet Hypothesen für mögliche Auswahlkriterien ab und testet diese.

In der ersten Studie (Kapitel 2) wurden 166 Praktiker in Deutschland dazu befragt, welche Auswahl- und insbesondere Persönlichkeitsverfahren Sie tatsächlich einsetzen, wie sie den Nutzen einschätzen und welche Präferenzen sie in Bezug auf ausgewählte Testcharakteristika haben. Wie erwartet zeigte sich, dass Persönlichkeitstests eher selten überhaupt zum Zwecke der Personalauswahl eingesetzt werden (nur 15% der Befragten nutzen diese) und dass gleichzeitig eine Vielzahl verschiedener Tests verwendet werden (alleine 15 verschiedene Testverfahren in dieser Befragung, wobei einige Befragte mehrere Verfahren nannten). Während die Wissenschaft sich hier besonders auf die Big Five bezieht, messen die in der Praxis verwendeten Verfahren in der Regel andere Eigenschaften. Die Befragten befanden die Nützlichkeit des Einsatzes als deutlich höher für die Auswahl auf Führungsebene

als auf Mitarbeiterebene. In Bezug auf bestimmte Testcharakteristika ergaben sich beispielsweise Präferenzen zu typenbasierten Tests, statistischer Testentwicklung und computerbasierter Anwendung. Außerdem präferieren Praktiker Verfahren, die auch von anderen Unternehmen genutzt werden und suchen diese vor allem über das Internet oder Flyer. Diese Ergebnisse verdeutlichen die Research-Practice Gap beim Einsatz von Persönlichkeitstests in der Personalauswahl und geben erste Hinweise auf mögliche Kriterien bei der Auswahl solcher Verfahren.

In der zweiten Studie (Kapitel 3) untersuchten wir (basierend auf der Erkenntnis in Kapitel 2, dass Praktiker vor allem im Internet nach Persönlichkeitstests suchen), ob eine spezielle Werbestrategie, nämlich die Verbindung von (fiktiven) neurowissenschaftlichen Informationen mit einem Persönlichkeitstest die Attraktivität dieses Tests steigert und somit die Entscheidung für eben diesen beeinflussen könnte. Vergangene Studien hatten den positiven Einfluss solcher Informationen auf die Bewertung von Erklärungen für bestimmte Phänomene gezeigt. Dazu befragten wir HR Praktiker aus dem Sparkassen-Verband, BWL-Studenten und Psychologie-Studenten mithilfe eines Flyers zu einem fiktiven Persönlichkeitsverfahren. Hier zeigte sich entgegen der Erwartung, dass die Nutzung neurowissenschaftlicher Informationen eher einen gegenteiligen Effekt aufweist, die Bewertung des Tests negativer ausfiel. Dieser Effekt ist bei den HR Praktikern am stärksten ausgeprägt, bei den Psychologie-Studenten am wenigsten. Obwohl der Effekt dieser Werbestrategie negativ war, zeigte sich doch ein deutlicher Einfluss auf die Bewertung des Persönlichkeitstests und somit die Bedeutung von Werbung bzw. einer Unique Selling Proposition auf das Entscheidungsverhalten.

Die dritte Studie (Kapitel 4) besteht aus mehreren Experimenten, in denen der Einfluss von zwei weiteren Kriterien auf die Bewertung eines Persönlichkeitsverfahrens bewertet wurde. Experiment 1 bezog sich dabei auf die Frage, ob die Unterscheidung zwischen typen- oder dimensionsbasierter Ergebnisdarstellung die Attraktivität eines Tests beeinflusst. Dabei konnte weder eine Flyer-Manipulation bei HR Praktikern analog zur zweiten Studie (Experiment 1a)

noch eine Ergebnisbericht-Manipulation bei BWL-Studenten (Experiment 1b) signifikante Unterschiede in den Bewertungen eines fiktiven Testverfahrens hervorrufen. In Experiment 2 untersuchten wir das Kriterium einer vorwiegend statistisch-faktoranalytischen Entwicklung vs. einer theoriebasierten Entwicklung des Testverfahrens. Auch hier konnte bei keiner der drei Teilstichproben (HR-Praktiker, BWL-Studenten, Psychologie-Studenten) ein signifikanter Unterschied in der Beurteilung des fiktiven Persönlichkeitsverfahrens anhand eines Flyers gefunden werden.

In der vierten Studie (Kapitel 5) wurden die impliziten Leistungstheorien von 43 Studenten und 42 erfahrenen Führungskräften analysiert. Es ging um eine Erfassung der Eigenschaften, Kompetenzen und Fähigkeiten, von denen Menschen annehmen, dass sie beruflichen Erfolg beeinflussen. In einer qualitativen Analyse wurden die gesammelten Begrifflichkeiten und Informationen zu insgesamt 29 Kategorien zusammengefasst. Innerhalb dieser Bandbreite an erfolgsrelevanten Eigenschaften gab es größere Unterschiede in Bezug auf die genannten Häufigkeiten von zwei Nennungen bis hin zu 39 Nennungen. Auch die Wichtigkeit der genannten Eigenschaften wurde unterschiedlich eingeschätzt. Insgesamt zeigt die Vielzahl an Kategorien, dass eher ein Bedarf an differenzierten Persönlichkeitsverfahren zu erwarten ist, die auf die konkreten Bedarfe zugeschnitten werden können.

Alles in allem zeigen die Ergebnisse, dass die tatsächliche Testpraxis und wissenschaftliche Empfehlungen und Vorgehensweisen nicht immer zusammenpassen. Es gilt die Ansprüche, Bedarfe und Entscheidungsprozesse von Praktikern besser zu verstehen, Lösungen, d. h. Persönlichkeitsverfahren zu entwickeln, die sowohl wissenschaftliche Kriterien als auch praktischen Ansprüchen genügen, die Kommunikation zu Praktikern zu verbessern und Forschung anwendungsorientiert zu gestalten.

INHALTSVERZEICHNIS

DANKSAGUNG.....	iii
ALLGEMEINE ZUSAMMENFASSUNG	iv
INHALTSVERZEICHNIS.....	vii
LISTE DER PUBLIKATIONEN.....	x
ABBILDUNGSVERZEICHNIS	xii
TABELLENVERZEICHNIS	xiii
KAPITEL 1: ALLGEMEINE EINLEITUNG	1
KAPITEL 2: PERSONALITY TESTING IN PERSONNEL SELECTION: LOVE IT? LEAVE IT? UNDERSTAND IT!	11
Understand it: The practice of personality test use	14
Personality tests by comparison: What’s it all about?.....	19
Test characteristics and presentation of results	19
Mode of delivery	24
Declaration and description of quality criteria	26
Finding a personality test	27
Future prospects and Conclusions.....	28
KAPITEL 3: THE ROLE OF NEUROSCIENCE INFORMATION IN CHOOSING A PERSONALITY TEST: NOT AS SEDUCTIVE AS EXPECTED.....	31
Abstract	32
Introduction	33
Background	34
Method	39
Participants	39
Material	40
Dependent variables	41
Results	42
Discussion	44
KAPITEL 4: FINDING THE RIGHT (TEST) TYPE: ON THE DIFFERENCES BETWEEN TYPE- VS. DIMENSION-BASED PERSONALITY TESTS AND BETWEEN STATISTICS- VS. THEORY BASED PERSONALITY TESTS WHEN DECIDING FOR OR AGAINST A TEST IN PERSONNEL SELECTION.	49
Abstract	50
Introduction	51
Background	51

Experiment 1a (Type).....	56
Method	56
Results and Discussion.....	58
Experiment 1b (Type)	59
Method	59
Results and Discussion.....	62
Experiment 2 (Theory)	62
Method	63
Dependent variables	65
Results and Discussion.....	65
Overall Discussion	66
KAPITEL 5: IMPLIZITE LEISTUNGSTHEORIEN IN DER PERSONALAUSWAHL	70
Zusammenfassung	70
Einleitung	71
Theoretischer Hintergrund	72
Methode.....	77
Design und Vorgehen.....	77
Teilnehmer	78
Datenanalyse	79
Ergebnisse	80
Diskussion	81
KAPITEL 6: GESAMTDISKUSSION	90
Zusammenfassung der Ergebnisse	90
Erkenntnisse und Limitationen.....	93
Hinweise für zukünftige Forschung	97
Implikationen für die Praxis	98
Fazit.....	100
LITERATURVERZEICHNIS	101
ANHÄNGE	131
Anhang A (zu Kapitel 3)	131
Flyer mit neurowissenschaftlicher Information	131
Flyer ohne neurowissenschaftliche Information	132
Anhang B (zu Kapitel 4)	133
Flyer Experiment 1a Typen.....	133

Flyer Experiment 1a Dimensionen.....	134
Ergebnisbericht Experiment 1b Typen.....	135
Ergebnisbericht Experiment 1b Dimensionen.....	137
Flyer Experiment 2 Statistik.....	139
Flyer Experiment 2 Theorie	140
CURRICULUM VITAE	141

LISTE DER PUBLIKATIONEN

Meine Promotion basiert auf einem veröffentlichten Buchkapitel, das eine Einleitung in das Gesamtthema gibt, sowie zwei veröffentlichten Artikeln und einem unveröffentlichten Artikel zu spezifischen Teilaspekten. Alle diese Abhandlungen wurden mithilfe von Ko-Autoren geschrieben, die im Folgenden benannt sind.

Das Buchkapitel und die Artikel wurden jeweils an die Formatierung dieser Arbeit angepasst. Der Wortlaut sowie die inhaltliche Gliederung entspricht jeweils der Veröffentlichung (sofern vorhanden), so dass jeder Artikel eine eigene Einleitung, Methodendarstellung, Ergebnisse und Diskussion hat (das Buchkapitel folgt einer anderen Gliederung). Die Abbildungen und Tabellen sowie die Literaturangaben aus allen Artikeln und Buchkapiteln (Kapitel 2 bis 5, siehe Referenzen unten) werden fortlaufend über Kapitel 2 bis 5 nummeriert und in gemeinsamen Verzeichnissen dargestellt.

Kapitel 2: Diekmann, J., & König, C. J. (2015). Personality testing in personnel selection: Love it? Leave it? Change it! In I. Nikolaou & J. K. Oostrom (Hrsg.), *Employee recruitment, selection, and assessment: Contemporary issues for theory and practice* (S. 117–135). Hove, UK: Psychology Press.

Das Buchkapitel wird außerdem in einer neuen Buchserie von Routledge Books, Key Issues in Work and Organizational Psychology, veröffentlicht.

Kapitel 3: Diekmann, J., König, C. J., & Alles, J. (2015). The role of neuroscience information for choosing a personality test: Not as seductive as expected. *International Journal of Selection and Assessment*, 23, 99-108. doi:10.1111/ijsa.12099

Kapitel 4: Diekmann, J., König, C. J., Schadow, T., Knab, Y., & Harrison, O. V. (2016). Finding the right (test) type: On the differences between type-

vs. dimension-based personality tests and between statistics- vs. theory-based personality tests when deciding for or against a test in personnel selection. *Zeitschrift für Sozialmanagement/Journal of Social Management*, 14, 83-102.

Kapitel 5: Diekmann, J., König, C. J., Bachner, J., & Langer, M., Becker, J., Kammerhoff, J., Knab, Y., & Seo, H.-N. (2018). *Implizite Leistungstheorien in der Personalauswahl: Welche Persönlichkeitsfaktoren halten Laien und Führungskräfte für wichtig?* Unveröffentlichter Artikel, Universität des Saarlandes.

ABBILDUNGSVERZEICHNIS

Figure 1	Frequency of selection methods used in percent (N = 166 German companies) in chapter 2.	17
Figure 2	Evaluation of the benefit of personality tests for different purposes in chapter 2. Agreement regarding usefulness was given on a seven-point scale (1 = no agreement to 7 = full agreement). All differences between users and non-users were significant ($p < 0.01$, all t 's < -2.6).	17
Figure 3	Personality tests used in Germany (in frequencies) in chapter 2. (BIP = Bochum Inventory for profession-related personality description; PPA = Thomas Assessment / Personal Profile Analysis; HPI = Hogan Personality Inventory; GEVA = the GEVA institute is a German consulting company specialized in behavioral analysis and evaluation tools; OPQ = Occupational Personality Questionnaire).	18
Figure 4	Preferences of different criteria that distinguish personality tests in chapter 2. Error bars indicate 95% confidence interval. Preferences were rated on a six-point semantic differential scale.	22
Figure 5	fMRI image used to illustrate the test version with neuroscience information in chapter 3. From "Brain activation for the movement of fingers measured in an fMRI experiment" by K. Tabelow, 2009, http://www.mathematik.hu-berlin.de/~tdm/2009/abstract.php?name=tabelow&id=uf5pYyDxgBv5I . Copyright 2009 by Karsten Tabelow. Reprinted with permission.	41
Figure 6	Overall test rating separately for sample and subsamples in chapter 3. Error bars indicate 95% confidence interval.	43
Figure 7	Representation of traits in the flyer of the type-based condition in chapter 4.	57
Figure 8	Representation of traits in the flyer of the dimension-based condition in chapter 4.	58
Figure 9	Representation of traits in the results report of the dimension-based condition in chapter 4.	61
Figure 10	Representation of traits in the results report of the type-based condition in chapter 4.	61

TABELLENVERZEICHNIS

Table 1	Overall test rating, standard deviations and effects size for sample and subsamples in chapter 3.	42
Table 2	Overall test rating, standard deviations, and effect sizes for sample and subsamples in chapter 4.	66
Tabelle 3	Kategorien der impliziten Leistungstheorie und ihre Definitionen, sortiert nach genannter Häufigkeit in Kapitel 5.	82
Tabelle 4	Vergleich der Kategorien mit den Skalen des Bochumer Inventars zur berufsbezogenen Persönlichkeitsbeschreibung (BIP) (Hossiep & Paschen, 2003) in Kapitel 5	86

KAPITEL 1

ALLGEMEINE EINLEITUNG

Bei der betrieblichen Personalauswahl und Eignungsdiagnostik handelt es sich um ein Kernthema sowohl für die arbeits- und organisationspsychologische Forschung (z. B. zum Nutzen unterschiedlicher Auswahlinstrumente wie Interviews, Assessment Centern oder Testverfahren, siehe z. B. Barrick, Mount, & Judge, 2001; Kuncel, Hezlett, & Ones, 2004; Schmidt & Hunter, 1998) als auch für die betriebliche Praxis von Personalern. Denn: Die Auswahl der passenden Mitarbeiter trägt wesentlich zum unternehmerischen Erfolg bei (Huselid, 1995; Lado & Wilson, 1994; Terpstra & Rozell, 1993). Somit ist – wie auch in anderen Teilbereichen anwendungsorientierter Forschung – zum einen ein hohes Interesse von Personalern an den Ergebnissen dieser Forschung zu erwarten und zum anderen ein hohes Interesse von Forschern, sich an realen Fragestellungen und Bedürfnissen zu orientieren.

Die Realität sieht jedoch manchmal anders aus: Viele Autoren haben in den letzten Jahren über eine zunehmende Research-Practice Gap – eine Kluft zwischen Forschung und Praxis – im Human Resources Management berichtet und über Ursachen und Lösungswege diskutiert (N. Anderson, Herriot, & Hodgkinson, 2001; Buckley, Ferris, Bernardin, & Harvey, 1998; Gelade, 2006a, 2006b; Hodgkinson, 2006; Hodgkinson & Herriot, 2002; Hodgkinson, Herriot, & Anderson, 2001; Klehe, 2004; Rice, 1997; Rynes, Bartunek, & Daft, 2001; Rynes, Giluk, & Brown, 2007; Rynes, Brown, & Colbert, 2002; Starkey & Madan, 2001; Subramony, 2006; Tkachenko, Hahn, & Peterson, 2017; Van de Ven & Johnson, 2006; Weick, 2001). Die Problematik besteht laut den verschiedenen Autoren darin, dass die Erkenntnisse der Wissenschaft im Bereich Human Resources Management, d. h. auf Theorien und empirischen Hinweisen basierende Empfehlungen für „gute“ Managementtechniken, von Praktikern nicht immer übernommen werden (Johns, 1993; Rowe, Williams, & Day, 1994; Terpstra & Rozell, 1993, 1997). Dabei sind sich sowohl Forscher als auch Praktiker bewusst, dass ein Übersetzungsproblem von wissenschaftlichen Befunden in die Praxis besteht und die

wissenschaftliche Wissensproduktion für Praktiker nicht immer ausreichend relevant ist (Shapiro, Kirkman, & Courtney, 2007). Es scheint also einen Bruch zwischen den Anforderungen der Praktiker bzw. Human Resources Manager auf der einen und der Forschungstätigkeit der Wissenschaftler auf der anderen Seite zu geben. Dies führt dazu, dass Forscher und ihre Bemühungen möglicherweise nicht so viel Einfluss außerhalb der akademischen Welt ausüben, wie es möglich wäre (Abrahamson & Eisenman, 2001).

Es wurden bereits verschiedene (möglicherweise z. T. parallel und/oder additiv wirkende) Ursachen diskutiert, die zur Entstehung und Aufrechterhaltung der Research-Practice Gap beitragen (Van de Ven & Johnson, 2006; siehe auch Tkachenko et al, 2017, für einen Überblick). Diese können in drei Kategorien eingeteilt werden.

1. Ursachen auf Seiten der Wissenschaft: Den Wissenschaftlern, die sich mit Human Resources Management befassen, wird teilweise vorgeworfen, ihre anwendungsbezogene Forschung weise zu wenig Anwendungsbezug auf: Nach Kersting (2008) werden die Grundlagenwissenschaft und Laborsituationen innerhalb der Psychologie zu stark betont. Da Untersuchungen im Feld schwerer generalisierbar sind, werden sie manchmal als zweitrangig betrachtet. D. h. die Forschung findet eher auf einer abstrakten Ebene statt, wobei die Bedürfnisse der Praktiker sowie die Rahmenbedingungen und Einschränkungen, mit denen diese sich konfrontiert sehen, wenig Berücksichtigung finden (Ryan & Tippins, 2004; Van de Ven & Johnson, 2006). In der Folge erscheinen die Bemühungen der Wissenschaft für Praktiker als zu wenig relevant (Buckley et al., 1998; Gelade, 2006a). HR Manager nehmen sie eher als abstrakt und für die Praxis wenig geeignet wahr (siehe z. B. Gelade, 2006a; Huff, 2000).

2. Ursachen auf Seiten der Praxis: Praktiker wollen oder können die Erkenntnisse der Wissenschaftler (seien sie nun abstrakt oder praxisbezogen) teilweise nicht anwenden (siehe z. B. Buckley et al., 1998). Dies kann verschiedene Ursachen haben. Zunächst haben Organisationen häufig das Bedürfnis, ihre Einzigartigkeit hervorzuheben. Sie wollen somit keine standardisierten Instrumente bzw. bezweifeln ihre Anwendungsfähigkeit im speziellen

Kontext des Unternehmens (Rousseau, 2006). Nach Subramony (2006) spielt zudem eine Rolle, ob die HR-Aktivitäten zur Gesamtunternehmensstrategie passen. Außerdem spielen für Unternehmen in der Regel finanzielle Aspekte eine wichtige Rolle (Ryan & Tippins, 2004). Auch wenn standardisierte, valide Auswahlverfahren langfristig einen deutlichen Beitrag zum Erfolg eines Unternehmens leisten, ist der kurzfristige finanzielle Erfolg mit (in den Augen der Praktiker) bewährten und weniger aufwändigen Auswahlverfahren größer (Klehe, 2004).

3. Ursachen an der Schnittstelle zwischen Wissenschaft und Praxis: Die (möglichweise ausreichend anwendungsbezogenen) Befunde der Wissenschaft werden von den (möglicherweise durchaus zur Anwendung motivierten) Praktikern nicht ausreichend wahrgenommen und verstanden: Es besteht ein Kommunikationsproblem (siehe z. B. Kersting, 2008). Rynes, Giluk und Brown (2007) untersuchten die Verbreitung und inhaltliche Darstellung von drei aktuell in der Wissenschaft diskutierten Themen (Zusammenhang zwischen Leistung und Intelligenz, Persönlichkeit/ Integrität sowie Zielvereinbarungen) in drei praktikernahen Zeitschriften (HR Magazine, Human Resources Management und Harvard Business Review). Ihre Analyse ergab, dass weniger als 1,2% der Artikel zwischen 2000 und 2005 sich mit diesen Themen befassten, wobei nur die Artikel im HRM Bezug auf wissenschaftliche Erkenntnisse nahmen. Im HR Magazine sowie im HBR wurden Nachweise dagegen in Form von Interviews mit Managern oder Beratern erbracht. Ryan und Tippins (2004) weisen dementsprechend darauf hin, dass Forschungsergebnisse nicht angemessen verbreitet, sondern nur in hochspezialisierten psychologischen Zeitschriften veröffentlicht werden (da sie dort zu größerer wissenschaftlicher Reputation führen). Darüber hinaus sind die Vorgehensweisen unterschiedlich. Eine Trennung findet laut Buckley et al. (1998) in Bezug auf verschiedene Aspekte statt: Wissenschaftler arbeiten eher theoretisch, mit der Unterstützung durch Daten, mit wissenschaftlichen Methoden, akademikerorientiert und auf eine Festanstellung (tenure) ausgerichtet, Praktiker arbeiten dagegen pragmatisch, logisch, mit Fallbeispielen, praxisorientiert und auf organisationalen Erfolg ausgerichtet.

Ursachen und Lösungsmöglichkeiten wurden auf dieser globalen Ebene also bereits umfassend untersucht. In der Regel wird diese Kluft in Bezug auf die gesamte HR-Forschung diskutiert, betrifft somit aber auch den engeren Bereich der beruflichen Eignungsdiagnostik: „Clearly, this is also the case in personnel selection because personnel practices that are often not well supported by empirical research are often very popular whereas personnel practices that have been shown to be effective are less frequently used” (N. Anderson, Lievens, van Dam, & Ryan, 2004, p. 495; siehe auch Buckley et al., 1998; Terpstra, 1996).

Dabei sind unstrukturierte Verfahren bei Praktikern sehr beliebt, obwohl die Forschung strukturierte Verfahren empfiehlt (Highhouse, 2008). Ein gutes Beispiel hierfür ist das Einstellungsgespräch, das auf sehr unterschiedliche Weisen durchgeführt werden kann: Unstrukturiert, teilstrukturiert, vollständig strukturiert, telefonisch oder direkt, von der Fachabteilung oder der Personalabteilung. Die Forschung empfiehlt hier im Sinne einer höheren Validität das standardisierte Einstellungsgespräch (Kepes, Banks, McDaniel, & Whetzel, 2012; McDaniel, Whetzel, Schmidt, & Maurer, 1994). Immer wieder zeigt sich jedoch, dass eine unstrukturierte Version in der Praxis sehr beliebt ist (Lievens & De Paepe, 2004; Stephan & Westhoff, 2002), wobei einzelne Studien auch einen umgekehrten Trend (häufigere Nutzung der standardisierten Version) zeigen (Schuler, Hell, Trapmann, Schaar, & Boramir, 2007). Ein anderes Beispiel betrifft Intelligenztests: Obwohl die hohe Vorhersagequalität schon 1998 (Schmidt & Hunter, 1998) aufgezeigt und seitdem vielfach bestätigt wurde, ist die tatsächliche Nutzungshäufigkeit bei Einstellungsentscheidungen vergleichsweise gering (siehe z. B. Ryan, McFarland, Baron, & Page, 1999; Schuler et al., 2007). Die Gründe für den Einsatz unstrukturierter Verfahren können vielfältig sein und von berechtigten rechtlichen Bedenken beim Einsatz von z. B. Intelligenztests (diese sind in Deutschland unzulässig, sofern sie nur einen Intelligenzquotienten ohne konkreten Bezug zum Arbeitsplatz bieten, siehe Wietfeld, 2016) bis hin zum wahrscheinlich unberechtigten sehr hohen Vertrauen in die eigene Intuition und Urteilskraft (Highhouse, 2008) reichen. Gerade der letzte Aspekt spielt nicht nur bei der

Entscheidung für oder gegen einen Kandidaten eine Rolle. Die menschliche Urteilsfähigkeit (und deren Unzulänglichkeit) wird mit hoher Wahrscheinlichkeit auch die Entscheidung für oder gegen ein Auswahlverfahren bzw. die konkrete Ausgestaltung desselben beeinflussen. In dieser Arbeit soll (unter anderem) diese genauer untersucht werden, und zwar vor dem Hintergrund eines spezifischen Teilbereichs der Eignungsdiagnostik: Der Verwendung von Persönlichkeitstests.

Persönlichkeitstests sind in der Praxis nicht unbekannt (Furnham, 2008). Oakland (2004) betont den enormen Zuwachs an Tests im Laufe des 20. Jahrhunderts und schätzt die Anzahl von Verfahren, die vorwiegend in England und Westeuropa konstruiert wurden, auf 5000. Alleine in den USA gibt es nach Hough und Oswald (2005) ungefähr 2500 Testherausgeber sowie diverse, frei übers Internet verfügbare Persönlichkeitsmaße. Dennoch gab es bisher wenig Versuche herauszufinden, welche Tests im Kontext der Personalauswahl tatsächlich angewendet werden und warum welche Tests verwendet werden. Daher ist es das Ziel dieser Arbeit herauszufinden, welche Tests in der Praxis tatsächlich zum Zweck der Personalauswahl eingesetzt werden, was diese Tests ausmacht und warum sich Personaler dafür entscheiden.

Zunächst einmal ist zu sagen, dass der Nutzen von Persönlichkeitstests für die Personalauswahl auch innerhalb der Psychologie umstritten ist. Murphy und Dzieweczynski (2005) weisen darauf hin, dass erst seit den 90er Jahren aufgrund der zunehmenden Akzeptanz des Fünffaktorenmodells, den positiven Entwicklungen in Bezug auf die Messung von Persönlichkeit sowie der Veröffentlichung verschiedener Metaanalysen die Grundstimmung wieder optimistischer geworden ist (siehe auch Salgado, Viswesvaran, & Ones, 2006). Doch die Kontroverse ist weiterhin aktuell, z. B. aufgrund der Faking-Problematik, dem Mangel an Theorien, welche Persönlichkeit und Arbeitsleistung verbinden, den insgesamt eher geringen Validitätskoeffizienten, methodischen Problemen der Validierungsforschung oder auch der Frage, welche Persönlichkeitsdimensionen angemessen sind (Murphy & Dzieweczynski, 2005;

Rothstein & Goffin, 2000). Obwohl die Big Five die in der Wissenschaft favorisierten Persönlichkeitsdimensionen auch bei der Vorhersage von Berufserfolg sind, ist die Meinung innerhalb der Wissenschaft nicht einhellig und wurde deutlich kritisiert (z. B. Rothstein & Goffin, 2000). Diese kritische Haltung gegenüber Persönlichkeitstests spiegelt sich in der deutschen Praxis wider: Verschiedene weitere Studien zeigen, dass die Einsatzhäufigkeit von Persönlichkeitstests in Deutschland zu personalbezogenen Zwecken mit jeweils um 20% eher gering ist (Piotrowski & Armstrong, 2006; Schuler, Frier, & Kauffmann, 1993; Schuler et al., 2007; Shackleton & Newell, 1994). Insgesamt werden Tests in Deutschland im Vergleich mit anderen europäischen Ländern sehr selten genutzt (Salgado & Anderson, 2002).

Eine Kluft zwischen Wissenschaft und Praxis findet sich aber bei der konkreten Auswahl von Persönlichkeitstests für Personalauswahlzwecke. Verschiedene Autoren (z. B. Hough & Oswald, 2005; Hülshager & Maier, 2008; Johns, 1993; Klehe, 2004; Murphy & Dzieweczynski, 2005) haben darauf hingewiesen, dass die in der Praxis angewendeten Persönlichkeitstests in der Regel nicht die in der Wissenschaft verbreiteten Big Five messen. Viele der veröffentlichten Tests befassen sich zudem nicht näher mit den verschiedenen Formen der Validität und sind somit keine gute Wahl für Unternehmen (Hough & Oswald, 2005; siehe auch Tewes, 1998). In der Praxis häufig angewendete Verfahren wie der MBTI oder selbst entwickelte Persönlichkeitsfragebögen erfahren somit „von wissenschaftlicher Seite geringe Unterstützung“ (Hülshager & Maier, 2008, p. 111).

Hogan (2005) glaubt, dass Psychologen – statt Praktikern Hinweise auf angemessene Testverfahren zu geben – Studenten und anderen Akademikern gegenüber Anti-Persönlichkeitsargumente äußern, so dass die Verbreitung schlecht bzw. gar nicht validierter Testverfahren in der Praxis nicht überrascht. Hough und Oswald (2005) sehen hier unbedingten Handlungsbedarf, um Validitäten bzw. die Bedeutung von Reliabilität und Validität für die Vorhersageleistung eines Tests an Praktiker zu übermitteln und verständlich zu machen. Auf der anderen Seite scheinen Praktiker, auch wenn ihnen aussagekräftige Befunde zur Güte von

Testverfahren vorgelegt werden, unstrukturierte Verfahren zu bevorzugen (Ryan & Tippins, 2004). Terpstra (1996) vermutet fehlende Augenscheinvalidität hinter der Zurückhaltung in Bezug auf einige Testverfahren. Nach Kersting (2008) bieten unstandardisierte Testverfahren außerdem den „Vorteil“, dass implizite Ziele bzw. inoffizielle Regeln in Unternehmen besser durchgesetzt werden können (d. h. die Bewerber können besser im Sinne der unternehmensinternen Wünsche ausgesucht werden). Darüber hinaus haben Praktiker Schwierigkeiten, sich auf dem intransparenten Testmarkt für einen Test zu entscheiden (Göhs & Dick, 2001).

Studien dazu, welche Tests dann tatsächlich im Kontext der Personalauswahl eingesetzt werden, gibt es jedoch kaum (Ausnahmen sind Berchtold, 2005; Di Milia, 2004). Meist wird recht plakativ bemängelt, dass die tatsächlich ausgewählten und bei Auswahlverfahren verwendeten Testverfahren aus wissenschaftlicher Sicht nicht empfohlen werden können. Dennoch werden häufig nur wenige Beispiele wie der MBTI genannt. Daher ergeben sich zwei relevante Forschungsfragen:

Forschungsfrage 1: Welche Tests werden eigentlich in der Praxis eingesetzt und worin unterscheiden sich diese voneinander?

Forschungsfrage 2: Wenn die Entscheidung für einen Test so schwierig ist, an welchen Kriterien orientieren sich Praktiker dann tatsächlich?

Die folgenden Studien gehen diesen beiden Fragen auf den Grund. In Kapitel 2 wird die tatsächliche Testpraxis mithilfe von Befragungen unter Personal-Praktikern erforscht und relevante Kriterien, die für Praktiker entscheidungsrelevant sein können, abgeleitet. In Kapitel 3 und 4 werden drei dieser Kriterien experimentell untersucht und in Kapitel 5 schließlich wird ein weiterer Forschungsansatz, die Bedeutung von impliziten Theorien von Praktikern für die Auswahl eines Persönlichkeitstests, dargestellt.

Kapitel 2 bietet zunächst einen Einstieg in das Forschungsfeld und stellt den Stand der Validitätsforschung und die zugehörigen Forscher-Perspektiven dar. So gibt es sowohl

Verfechter als auch Gegner des Einsatzes von Persönlichkeitstest im Kontext der Personalauswahl. Unabhängig davon zeigen jedoch verschiedene Studien, dass solche Tests in der Praxis angewendet werden, aber nur sehr wenige beschäftigen sich damit, welche Verfahren genau. Aus diesem Grund wird eine Befragung deutscher Praktiker dargestellt, in denen diese Frage beantwortet wird. Darauf aufbauend werden prägnante Unterschiede zwischen verschiedenen Testverfahren herausgestellt, die möglicherweise die Entscheidungen von Praktikern für oder gegen ein Persönlichkeitsverfahren beeinflussen könnten.

In Kapitel 3 und 4 werden drei dieser Kriterien experimentell untersucht, nämlich 1. die Werbestrategie der Verknüpfung von Persönlichkeitstests mit neurowissenschaftlichen Erkenntnissen, 2. die Art des Persönlichkeitstests (Typentest vs. dimensionaler Test) und 3. der Entwicklungshintergrund des Tests (theoriebasiert vs. statistisch orientiert). Kapitel 3 fokussiert sich dabei auf die Werbestrategie. Die Ergebnisse aus Kapitel 2 zeigen, dass Praktiker eine Vielzahl von Persönlichkeitstests nutzen und diese vorwiegend im Internet suchen oder sich Flyer anschauen. Damit spielt Werbung eine bedeutende Rolle bei der Entscheidung für so ein Verfahren. Die Tatsache, dass es viele Verfahren auf dem Markt gibt, bedeutet auch, dass kommerzielle Anbieter sich von ihrer Konkurrenz abheben müssen und möglichst eine „Unique Selling Proposition“, ein einzigartiges Verkaufsargument finden müssen. In dieser Studie liegt der Fokus auf der Verknüpfung von Persönlichkeit mit neurowissenschaftlichen Erkenntnissen, in Anlehnung an das Herrmann Brain Dominance Instrument, ein Verfahren, das diese Verbindung (zumindest metaphorisch) nutzt sowie aufbauend auf verschiedenen Studien, die eine positive Verbindung zwischen Bewertungen und der Nutzung neurowissenschaftlicher Erklärungen herstellen konnten.

Kapitel 4 analysiert den Einfluss von zwei spezifischen Kriterien, in denen sich Tests unterscheiden. Das erste Kriterium betrifft die Frage, ob ein Test typenbasiert oder dimensionsbasiert ist. Die Wissenschaft konzentriert sich häufig eher auf dimensionsbasierte Tests, wie z. B. beim NEO FFI, bei denen das Ergebnis der Testdurchführung eine Einordnung

des Testanden auf mehreren Dimensionen darstellt. In der Praxis beliebte Testverfahren, wie Sie in Kapitel 2 betrachtet werden, sind allerdings häufig typenbasiert, d. h. das Ergebnis der Testdurchführung ist die Zuordnung der Person zu einem bestimmten Typen. Hier ist die Frage, ob Praktiker möglicherweise eine Testart bevorzugen. Das zweite Kriterium betrifft die Frage, auf welcher Basis ein Persönlichkeitsverfahren entwickelt wurde: basierend auf statistischen bzw. faktoranalytischen Methoden oder eher theoriebasiert. Diese Fragestellung ist möglicherweise nicht vollständig unabhängig von der ersten, da einige häufig verwendete Typentests eher theoriebasiert entwickelt wurden (was nicht heißt, dass keine statistischen Methoden zum Einsatz kamen, sondern dass der Entwicklung eine Persönlichkeitstheorie zugrunde lag). In Kapitel 4 überprüfen wir, ob diese beiden Kriterien die Bewertung und damit auch die Entscheidung für oder gegen ein Persönlichkeitsverfahren beeinflussen könnten.

Kapitel 5 verfolgt eine weitere Forschungsfrage im Themengebiet Persönlichkeitsverfahren in der Personalauswahl und Research-Practice Gap. Der Fokus liegt hier auf der inhaltlichen Passung der impliziten Theorien von Personen zu relevanten Eigenschaften im Berufsleben und vorhandenen Testverfahren. Dabei wurden die impliziten Leistungstheorien von Studenten und Führungskräften untersucht. Implizite Theorien sind innere Überzeugungen zu verschiedensten inhaltlichen Gebieten, die alle Menschen in irgendeiner Form zu den Themen entwickeln, die für ihr Leben relevant sind. Wenn diese Vorstellungen und Überzeugungen in Bezug auf die Frage, welche Eigenschaften, Kompetenzen und Fähigkeiten für beruflichen Erfolg relevant sind, wesentlich von den Strukturen und erfassten Kompetenzen gängiger Persönlichkeitsverfahren abweichen, zeigt sich auch hier eine Research-Practice Gap, die es zu überwinden gilt. In der letzten Studie ging es daher darum, diese implizite Theorie zu erfassen und der aktuellen Test- und Forschungspraxis gegenüberzustellen.

Im letzten Kapitel, der allgemeinen Diskussion, werden die Gesamtergebnisse der vorgestellten Studien zusammengefasst, Stärken und Grenzen der vorliegenden Forschung aufgezeigt und Implikationen für zukünftige Forschung und die Praxis diskutiert.

KAPITEL 2**PERSONALITY TESTING IN PERSONNEL SELECTION: LOVE IT? LEAVE IT?****UNDERSTAND IT!**

This Book Chapter has been published in *Employee recruitment, selection, and assessment: Contemporary issues for theory and practice* (Nikolaou & Oostrom, 2015) and may not exactly replicate the authoritative document published there. It is not the copy of record. The authoritative document can be found at the following source:

Diekmann, J., & König, C. J. (2015). Personality testing in personnel selection: Love it? Leave it? Change it! In I. Nikolaou & J. K. Oostrom (Eds.), *Employee recruitment, selection, and assessment: Contemporary issues for theory and practice* (pp. 117–135). Hove, UK: Psychology Press.

The quality of selection procedures is judged primarily by looking at predictive validity results, as the prediction of performance at work is clearly the most important issue for the practice of personnel selection. Based on these results, researchers have made recommendations to improve methods such as the interview (Kepes et al., 2012; McDaniel et al., 1994) or have contributed to the abandonment of methods with no predictive quality, such as graphology (Driver, Buckley, & Frink, 1996).

Although most established selection methods such as mental ability tests or assessment centers have been found to be valid, the situation is significantly different with regard to personality testing. Discussions about whether personality tests are valid instruments began 60 years ago, with studies finding moderate but profession-dependent results at best (Ghiselli & Barthol, 1953), and generally troubling results at worst (Guion & Gottier, 1965). This discussion was intensified when, in their Big Five meta-analysis, Barrick and Mount (1991) found conscientiousness to be the only trait that was generally and at least moderately predictive of work performance, whereas the other four Big Five traits showed only small correlations which varied between different occupations. Currently, the debate about whether one should use personality tests in personnel selection procedures is dominated by two perspectives, both of which are supported by good arguments.

On the one hand, there are those advocates of personality tests who “love it” (e.g., Bartram, 2004; Ones, Viswesvaran, & Dilchert, 2005): The findings of Barrick and Mount (1991) as well as further meta-analyses (Hurtz & Donovan, 2000; Salgado, 1997) and a second-order meta-analysis (Barrick et al., 2001) are used to argue that there are consistent correlations and to support the central role of conscientiousness and (in part) of emotional stability in predicting job performance. Although the other Big Five traits were not related to overall work performance, they were able to predict performance in specific professions or criteria. Numerous studies and meta-analyses explored the personality-performance relationship. For example, a number of researchers examined the longitudinal impact of the Big Five on career

success (Judge, Higgins, Thoresen, & Barrick, 1999) using specific criteria such as job satisfaction (Judge, Heller, & Mount, 2002), or specific occupations or roles such as social professions (Blickle & Kramer, 2012) or leadership roles (Judge, Bono, Ilies, & Gerhardt, 2002). In these studies and meta-analyses, researchers frequently found high criterion-related validities (for a detailed overview of research, see Rothstein & Goffin, 2006). Moreover, advocates of personality measures in personnel selection argue that personality traits particularly predict typical performance, whereas general mental ability particularly predicts maximum performance (e.g., Marcus, Goffin, Johnston, & Rothstein, 2007).

Some debate within the “love it” group concerns the preference for broad or narrow personality traits: While some researchers recommend using all relevant personality traits together to maximize validity (Barrick & Mount, 2005) or using so-called compound personality traits (Ones & Viswesvaran, 1996; Ones et al., 2005) to predict overall job performance, others believe that narrow traits (and specific criteria) with well-considered theoretical assumptions of the trait-performance relationship will lead to better predictions (Dudley, Orvis, Lebiecki, & Cortina, 2006; J. Hogan & Holland, 2003; Tett, Steele, & Beauregard, 2003). Nevertheless, there is a group of advocates of personality testing who feel that “personality matters” (Barrick & Mount, 2005, p. 359).

On the other hand, there are researchers (e.g., Murphy & Dzieweczynski, 2005) who are more drawn to a “leave it” position. They argue that the correlations found in the above-mentioned meta-analyses are quite small and that there is a lack of convincing general theories that relate personality constructs to job performance (Murphy & Dzieweczynski, 2005). Even those who see themselves as more or less impartial (Morgeson et al., 2007a) are concerned about the low validity, which is sometimes “pimped” by corrections for predictor unreliability (Campion in Morgeson et al., 2007a). They therefore advise against the use of most personality tests in personnel selection contexts or recommend the additional use of tests of general mental ability. In addition to this validity issue, critics often also point to the problem of faking. There

is little doubt that applicants can, and actually do, fake answers when completing a personality test (e.g., Birkeland, Manson, Kisamore, Brannick, & Smith, 2006). Although some researchers consider this to be unproblematic (e.g., J. Hogan, Barrett, & Hogan, 2007), faking does seem to change rank orders and therefore affects actual selection decisions (Stewart, Darnold, Zimmerman, Parks, & Dustin, 2010). Common correction methods such as lie scales do not provide a satisfactory solution to the problem either (e.g., Campion, Dipboye and Schmitt in Morgeson et al., 2007a), although assessors believe that they do (Robie, Tuzinski, & Bly, 2006).

As this “love it or leave it” debate continues, so too does the use of personality tests (Bartram, 2004). Research clearly shows that organizations use personality tests: Personality testing is quite popular in Belgium, France, Greece, Ireland, The Netherlands, Portugal, Spain and Britain (Bruchon-Schweitzer & Ferrieux, 1991; Eleftheriou & Robertson, 1999; Hodgkinson, Daley, & Payne, 1995; Hodgkinson & Payne, 1998; Ryan et al., 1999; Schuler et al., 1993; Shackleton & Newell, 1994; Williams, 1992; Zibarras & Woods, 2010). It is also known to be a regularly used instrument in several other countries such as Germany, Italy, Scotland, and the USA (Harris, Dworkin, & Park, 1990; Piotrowski & Armstrong, 2006; Ryan et al., 1999; Rynes, Orlitzky, & Bretz, 1997; Scholarios & Lockyer, 1999; Schuler et al., 1993, 2007; Shackleton & Newell, 1994).

Understand it: The practice of personality test use

Against this background, we believe that it is time to set out on a new research path that concentrates on the practice of personality test use in organizational settings. Apart from the highly important questions of validity and faking, research should find out which tests are being used in which ways and for what reasons in order to optimize our recommendations to practitioners.

To our knowledge, only few authors have been interested in which tests are actually used by organizations or (Industrial and Organizational) psychologists in general (e.g., Brown, 1999; Evers et al., 2012; Furnham, 2008; Muñiz & Fernández-Hermida, 2010; Muñiz, Prieto,

Almeida, & Bartram, 1999; Ryan & Sackett, 1987, 1992; Sneath, Thakur, & Madjuck, 1976; Steck, 1997). Even fewer have explored which tests are used for personnel selection in particular (Berchtold, 2005; Di Milia, 2004), even though the criticism has been raised that personality tests are “poorly chosen” (Murphy & Dzieweczynski, 2005, p. 343).

At first, we will have a look on general test use in business contexts. Taking into account those studies which survey general test use in organizations and those conducted by Industrial and Organizational psychologists without a specific focus on selection (Berchtold, 2005; Brown, 1999; Di Milia, 2004; Furnham, 2008; Muñiz & Fernández-Hermida, 2010; Ryan & Sackett, 1987, 1992), the evidence so far shows that the tests most frequently mentioned across studies are the 16 Personality Factor Questionnaire (16 PF), the Myers-Briggs Type Indicator (MBTI), the Occupational Personality Questionnaire (OPQ), the Minnesota Multiphasic Personality Inventory (MMPI), the Big Five Personality Inventory (NEO), the California Psychological Inventory (CPI), and the Thomas Assessment/ Personal Profile Analysis (PPA). This is in line with information from job websites or free personality test websites listing the supposed main personality tests (Donston-Miller, n.d.; Free Personality Test, n.d.). However, many more tests are mentioned in these studies, reflecting the huge variety of tests which exist (there are an estimated 2,500 publishers in the United States alone, see R. Hogan, 2005; Hough & Oswald, 2005; Psychometric Success, 2013), operating in a \$500 million industry (Psychometric Success, 2013).

A closer look at the two studies that exclusively considered tests used in personnel selection procedures (Berchtold, 2005; Di Milia, 2004) reveals that there may be differences in test use that could be due to regional preferences or the fact that some tests have only a national range. Examining personality test use in selection procedures of Australian organizations, Di Milia (2004) found not only the OPQ, MBTI, NEO and 16PF to be frequently used, but also questionnaires, such as the Personal Characteristics Inventory (PCI), the Fifteen Factor Questionnaire, the Occupational Personality Profile (OPP) and the DISC (standing for

Dominance, Influence, Steadiness, and Conscientiousness). Swiss organizations (Berchtold, 2005) also use the MBTI, 16PF, Thomas Assessment, OPQ and NEO, supplemented by tests like the Master Person Analysis (MPA), Insights Discovery or MDI, the Bochum Inventory for profession-related personality description (BIP), the DISG (the German version of the DISC) or the Herrmann Brain Dominance Instrument (HBDI). All in all, 173 companies were found to use 52 different personality tests for selection purposes in Switzerland.

To complement the existing studies and to survey the current state of personality testing in Germany, we conducted our own study, questioning HR practitioners in companies of all sizes across Germany¹. We found that personality tests were used in 15.1% of the surveyed companies (see Figure 1 for the application frequency of all selection methods). This is slightly less than the 20% which has usually been found in Germany over the last twenty years (Schuler et al., 2007) but can probably be explained by the fact that we also had smaller companies in our sample (41.6% had fewer than 500 employees). Respondents found personality tests to be moderately useful for promotion, planning of personnel development activities, assistance in team development activities and for personnel selection at the employee level, and to be somewhat more useful for personnel selection at the management level (see Figure 2). Actual test users found personality tests to be significantly more useful for all purposes than did non-test users. Concerning the question of which personality tests were used, in accordance with the

¹ We randomly called 769 companies; in 605, we were able to talk to employees or managers who worked in conducting the selection process. 403 people were interested in participating in the study and were invited to take part in the online survey by e-mail. A total of 166 persons (37.3% male, 56.6% female, 6.0% did not specify their gender) actually completed the whole survey (292 dropped out). Respondents had been in their current jobs for an average of 12.7 years ($SD = 8.8$) and most (71.7%) had a university education, with the majority being trained in business administration (58.0%) and only 5.0% in psychology. On average, they had been involved in 41.3 selection procedures during the last year ($SD = 111.9$), and a total of 77.1% had decision-making rights concerning the choice of selection methods. Companies had approximately 904.4 ($SD = 1608.9$) employees (7.8% had up to 50 employees, 16.3% between 51 and 250, 24.1% between 251 and 500 and 39.2% had over 500 employees; 12.7% did not answer this question), 72.3% were operating internationally, mostly in manufacturing, wholesale and the retail trade, financial and insurance activities or personnel services. The survey consisted of three main parts: First, we wanted to know which selection methods the companies used. Second, we asked participants about the purposes for which they found personality tests to be useful. Third, we concentrated on personality test use in personnel selection and asked for preferences of 15 different criteria that can be used to distinguish these tests.

studies mentioned above, we found a huge variety of different methods, including Insights Discovery or MDI, the BIP, the PPA, the 16 PF, the DISC, the Hogan Personality Inventory (HPI) and the Predictive Index (PI). For an overview of all mentioned tests, see Figure 3.

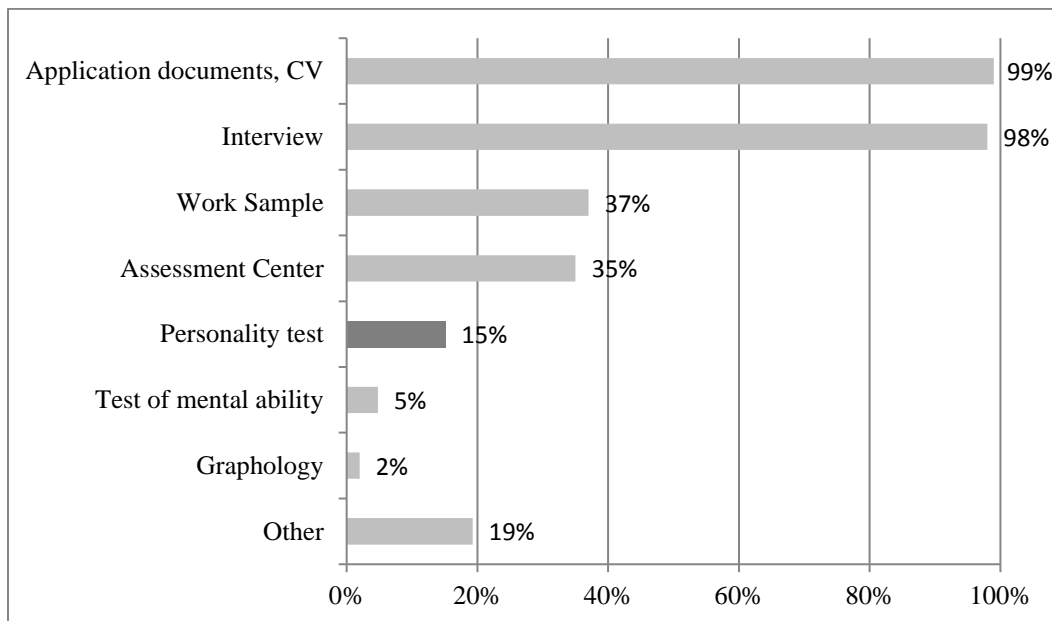


Figure 1. Frequency of selection methods used in percent (N = 166 German companies) in chapter 2.

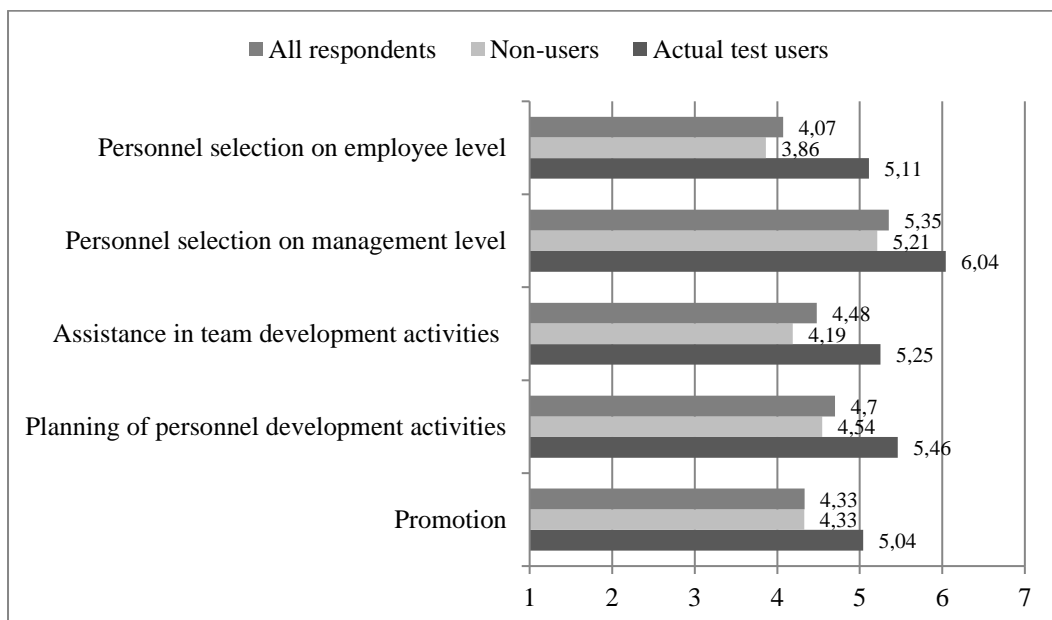


Figure 2. Evaluation of the benefit of personality tests for different purposes in chapter 2. Agreement regarding usefulness was given on a seven-point scale (1 = no agreement to 7 = full agreement). All differences between users and non-users were significant ($p < 0.01$, all t 's < -2.6).

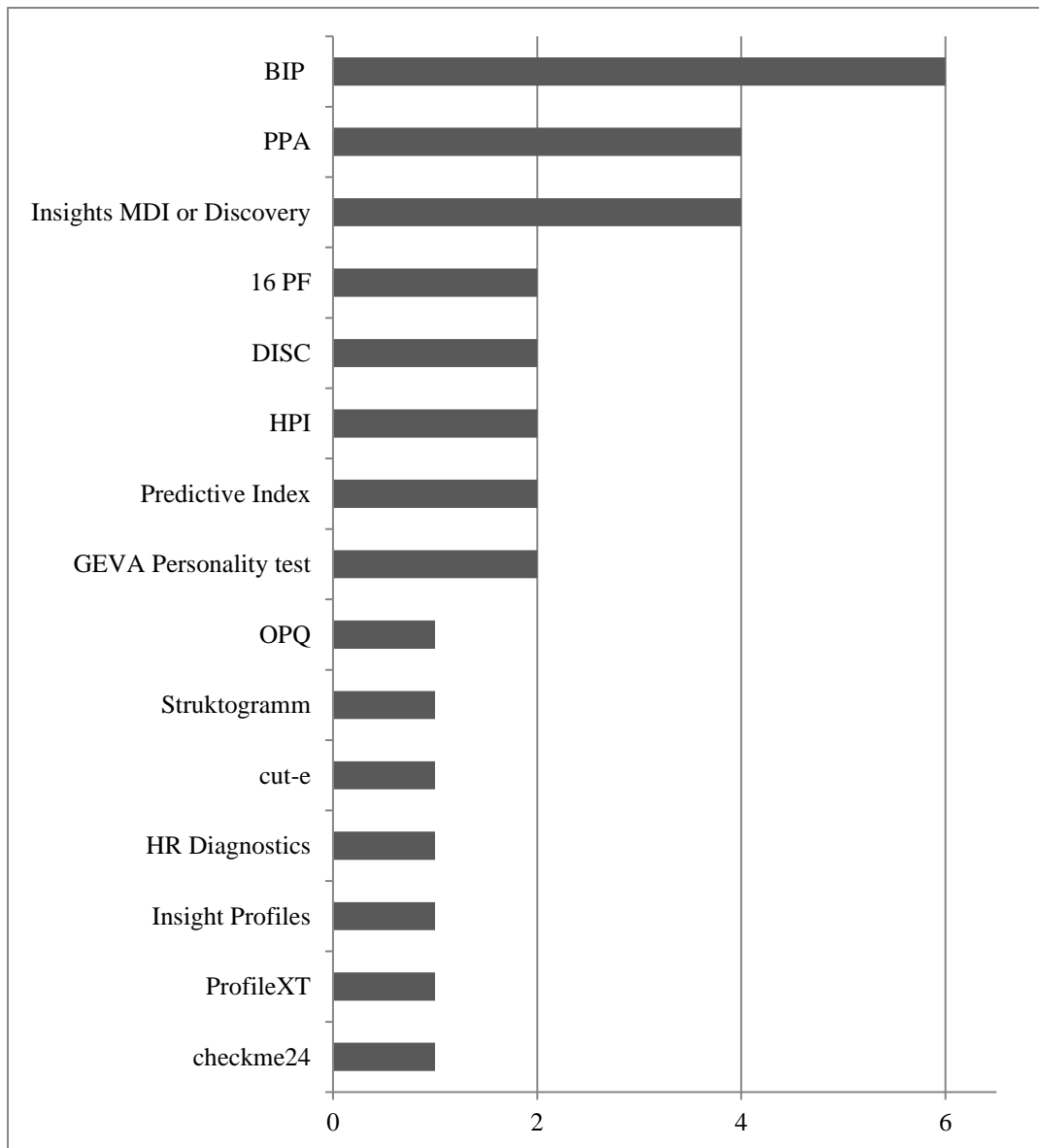


Figure 3. Personality tests used in Germany (in frequencies) in chapter 2. (BIP = Bochum Inventory for profession-related personality description; PPA = Thomas Assessment / Personal Profile Analysis; HPI = Hogan Personality Inventory; GEVA = the GEVA institute is a German consulting company specialized in behavioral analysis and evaluation tools; OPQ = Occupational Personality Questionnaire).

This study and the two previous ones (Berchtold, 2005; Di Milia, 2004) provide a first impression of the world of selection by personality testing. The MBTI is clearly one of the most frequently used personality tests; it is not only mentioned in various studies but is also high in the rank order of frequently used tests within these studies. Although the NEO personality inventory is also used in several countries, it generally ranks (far) below the MBTI (Berchtold,

2005; Di Milia, 2004; Furnham, 2008). This points towards the so-called research-practice gap in personnel selection, which describes the fact that research contents and recommendations of researchers are not always in line with the current implementation practice (e.g., Rynes et al., 2007): While we as researchers focus very much on the Big Five and instruments measuring these personality traits, practitioners seem to prefer other instruments like the MBTI although there is great doubt about its validity (e.g., Ones et al., 2005). Moreover, the three studies concentrating on selection (Berchtold, 2005; Di Milia, 2004; and our own study) clearly show that there is much more to personality testing than the MBTI and NEO (surprisingly, neither the MBTI nor the NEO are among the tests used in Germany). These three studies demonstrate the vast variety of personality tests in existence and use, some of which are restricted to certain countries/languages (for example the BIP, which was developed in Germany) and some of which are probably not appropriate in selection procedures.

Personality tests by comparison: What's it all about?

Let's take a closer look at the above-mentioned personality questionnaires: In the following section, we describe and discuss several important criteria beyond standard criteria such as reliability and validity (because literature was able to show that these criteria, as used for example in the Dutch COTAN system, are not the only criteria important to practitioners, e. g., König, Klehe, Berchtold, & Kleinmann, 2010) that concern characteristics of the personality test and its presentation of results, aspects of application, description of quality criteria and the process of finding a personality test that might influence the allure of often-used personality tests for practitioners.

Test characteristics and presentation of results

A first distinguishing criterion is whether the test results in a personality type (e.g., MBTI, DISC, HBDI) or in a dimensional personality profile (e.g., 16 PF, NEO, BIP, MPA). Whereas dimensions reflect the idea that a person usually shows all traits to a certain degree on a continuous scale, types group people into discrete classes (Gangestad & Snyder, 1985). The

measurement of dimensions is widespread in psychological research, but there seems to be a nagging distrust of types, which are often seen as an (over)simplification, a trigger of stereotyped thinking, or even pure invention (e.g., Gangestad & Snyder, 1985). Moreover, it is often difficult to decide where to set theoretically or empirically meaningful cut-off points that assign a person to one type or the other without misclassification, and there is the general question of whether a person can exclusively be assigned to one type (Robins, John, & Caspi, 1998; York & John, 1992). Even defenders of the MBTI believe that people can belong to more than one type and that the test alone will not find the “right” type, but that one needs to talk to the test taker (Bayne, 2005). At the same time, type-tests may have advantages over dimensional personality tests. For example, the reduction of information and complexity into one type may be easier to interpret and therefore more appealing. Whereas a dimension-based test reports many scales with a person varying on all of these scales, a type includes all information in an economical manner and makes it easier to differentiate between applicants. A schema-like categorization system may also better match the human knowledge structure of cognitive schemata (Smith & Queller, 2008) and limitations of cognitive capacity (Tversky & Kahneman, 1974).

Both approaches to personality testing may thus have their advantages and disadvantages (and may not only co-exist but even benefit from each other; Robins & Tracy, 2003). In our survey, we also asked the practitioners whether they preferred dimension-based personality tests or type tests and whether they preferred the results to be aggregated into one comparable value or to be presented in multiple comparable facets² (see Figure 4). Results of one-sample *t*-tests, testing for differences to the scale middle of 3.5, showed a significant³ preference for types rather than dimensions ($M = 3.89$, $SD = 1.53$) on the one hand and a

² Each preference item had two poles on a one- to six-point scale, e.g. “Would you prefer...” and “... a dimensional representation of measured traits” on one pole and “... the aggregation of measured traits in types” on the other pole.

³ Whenever we speak of significance, we mean at least $p < .05$.

significant preference for facets rather than an aggregation to one value ($M = 3.97$, $SD = 1.51$) on the other. This indicates that a mixture of both types of results may be most attractive. Interestingly, actual test users ($n = 28$, $M = 3.21$, $SD = 1.62$) preferred dimensions, whereas those who did not use personality tests ($n = 138$, $M = 4.03$, $SD = 1.48$) showed a strong preference for types. This suggests that a certain expertise concerning personality tests leads to a difference in preferences (but given the small sample of test users, this result should be treated with caution).

A second criterion concerns the report. The user is confronted with a type or a profile that she or he needs to interpret and compare with an ideal type or profile and/or other applicants. On the one hand, this compact alternative has the advantage that the user does not have to read a long report but can focus on the aspects that are important to him or her. On the other hand, if a practitioner is interested in an interpretation, he or she is left alone with this task. That can be a considerable problem if he or she is not a psychologist with appropriate training in test interpretation. A manual can be very helpful but may not always be easy to understand. The other option, which is usually provided automatically with online test versions, consists of detailed narrative reports, which offer the advantage of an extensive, easy-to-understand and quick evaluation that is less prone to mistakes regarding subjectivity and the difficult task of simultaneously processing several variables (Bartram, 1995; Snyder, 2000). So-called computer-based test interpretations have been used and discussed for decades now, especially in clinical psychology (e.g., Butcher, Perry, & Dean, 2009; Fowler, 1985). They are almost standard in reports of commercial test publishers as well as in science-based personality tests like the NEO (at least in some versions: in Germany, a narrative report is available for the NEO-PI R+, but not for the NEO FFI) and the BIP. These narrative reports can differ in terms of various aspects, for example the extent to which text and graphs are integrated, the involvement of interpretation of configurations and interactions, or the possibility to adapt a test to the context (e.g., development or selection) (Bartram, 1995).

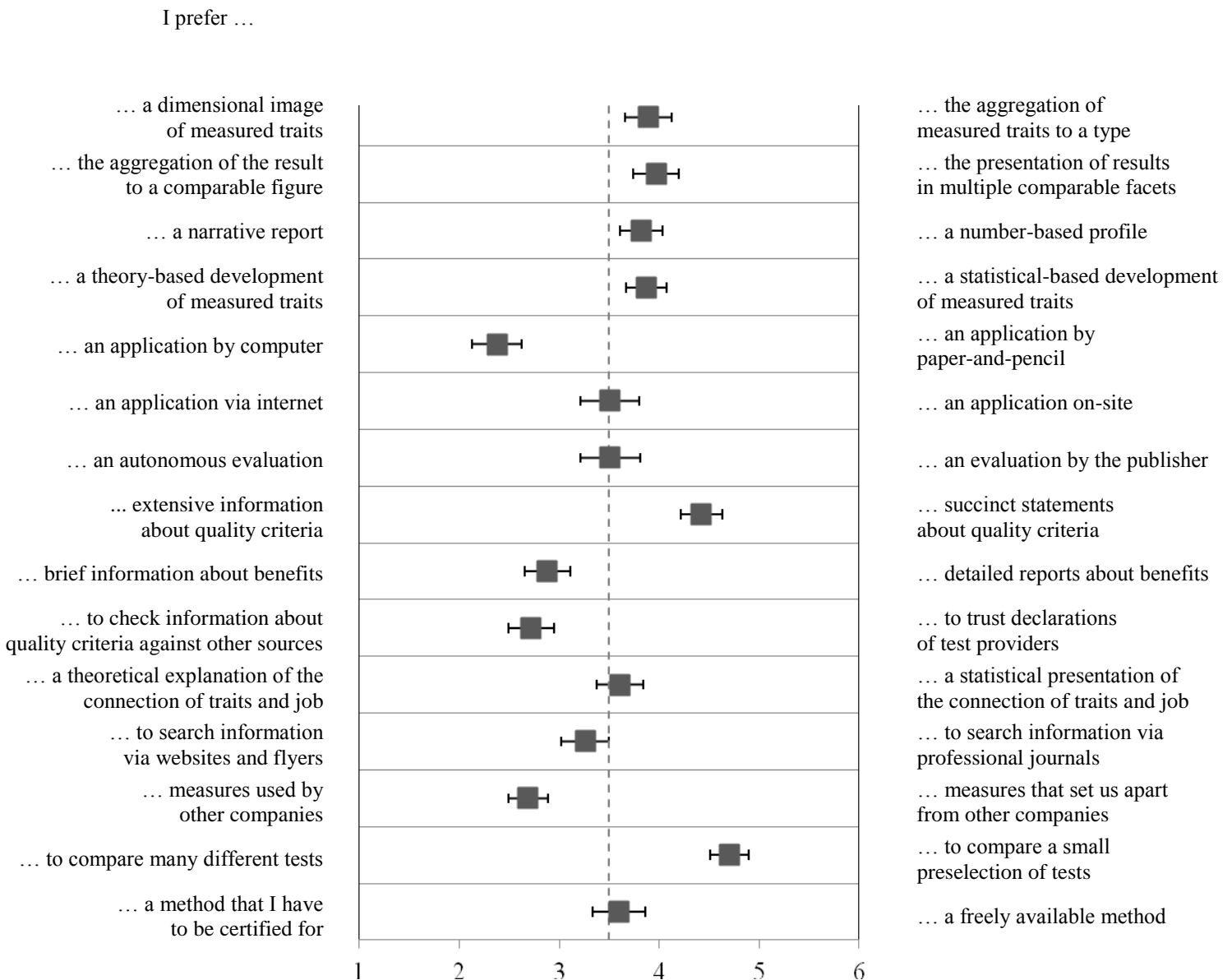


Figure 4. Preferences of different criteria that distinguish personality tests in chapter 2. Error bars indicate 95% confidence interval. Preferences were rated on a six-point semantic differential scale.

The gain of being provided with an interpretation is often bought with the uncertainty about accuracy and validity of these interpretations, and narrative reports of different tests probably differ in their accuracy (Kellett, McCahon, & James, 1991). Especially in the case of tests from commercial publishers, it is often difficult to evaluate how these interpretations are generated, which statistical methods and which interpretive rules or algorithms are used to combine test results and text modules, or how these text modules were developed. Frequently, the report cannot be modified or adapted to the current test context (Bartram, 1995), and even

if this were the case, it is questionable whether non-trained personnel staff would be able to do so appropriately. Some reports may even take advantage (knowingly or not) of the Barnum effect: They make such broad statements that people usually feel that the report is accurate, scientifically precise and offers good reasons for decisions, but it is actually too general for a practitioner to make well-grounded judgments (Guastello, Guastello, & Craft, 1989; Guastello & Rieke, 1990; Snyder, 2000). Unfortunately, there is barely any research concerning the issue of narrative reports in an organizational context or addressing the huge variety of tests in use. Our survey found a significant preference for a profile ($M = 3.82$, $SD = 1.41$) rather than a narrative report. Perhaps there is a stronger need for quick comparisons in selection procedures, making narrative reports less important than, for example, in consulting and development activities.

Another criterion concerns the development and background of a test. Although test development can have different backgrounds, there seem to be two major variations: A personality test can be based on a personality theory or on a statistical approach. The MBTI, for instance, is an example of the theory-based approach. It was developed by Katherine Briggs and her daughter Isabel Briggs Myers, under the influence of C. G. Jung's typology (Briggs Myers & Myers, 1993; Jung, 1960). Another influential theory concerns William Marston's (1979) behavioral types – originally called Dominance (D), Inducement (I), Submission (S), and Compliance (C). This led not only to the DISC assessment but also to the development of other personality tests such as the Personal Profile Analysis. Usually, these tests use an adapted version of Marston's original types, called Dominance (D), Influence (I), Steadiness (S) and Conscientiousness (C). Insights MDI used both models as a background (Euteneier & Scheelen, 2010). The HBDI, by contrast, was developed by Ned Herrmann (1989), taking into account brain hemispheres theory (e.g., Mintzberg, 1976) and MacLean's (1985) theory of the "triune brain." It results in four thinking styles, reflected by a four-quadrant brain model. Another (main) way of developing a test, which is favored by most scientists, is based on a statistical

approach. The NEO, for example, has such a statistical, non-theoretical background. It is based on the so-called lexical approach, and the Five Factors measured in this test were developed through factor analytical methods (e.g., McCrae & Costa, 1997).

There are good reasons why practitioners might be attracted by both approaches. On the one hand, the statistical, factor analytical method is an empirical one. This alone may give a personality test a serious appearance, meeting needs of legal security. On the other hand, people have a strong need for explanations, in particular explanations of human behavior (Keil, 2006; Lombrozo, 2006; Malle, 2004), and although the above-mentioned theories probably do not deliver such an explanation, they may serve as compensation. At least they suggest that there is more to a test than just a description of traits, and people may usually not require a scientifically tested theory (Keil, 2003, 2006; Rozenblit & Keil, 2002; Wilson & Keil, 1998). Moreover, such a general structure as derived in the NEO may not meet practitioners' requirements, as it does not refer to work-related applications such as personnel selection (Hough & Oswald, 2005). In our study, we also asked practitioners whether they preferred a theory-based or statistically-based development of traits. Results indicate that practitioners significantly favored a statistically-based development ($M = 3.87$, $SD = 1.32$). It thus appears that practitioners do understand the importance of a scientific approach.

Mode of delivery

The most apparent point of application concerns the presentation of the test: the "classic" paper-and-pencil form and the application at the computer with a local test system or via the internet. The advantages of an electronic application are obvious: The testing material as well as test and response time can be controlled, items can be easily adapted, application and evaluation of results are highly objective, printing costs and unwieldy paper copies are eliminated and feedback is available in an instant (Bartram, 2000; Lievens & Harris, 2003). What is more, the internet provides a high flexibility, as applicants can be tested independently of place and time (Lievens & Harris, 2003). At the same time, there are some difficulties that

have to be faced, which have been discussed to different degrees in the literature: Problems such as connection problems during internet testing or a lack of computer or internet access are likely not as serious as they were a couple of years ago but are probably still an issue. Moreover, practitioners should keep in mind that people have different levels of affinity to computers and the internet, which might lead to discrimination of some groups such as older people or ethnic minorities (Bartram, 2000). The ethical question of security of data transfer and confidential management of test results also remains important. A further question concerns the transferability of paper-and-pencil tests to the computer format. Currently, computer-based tests are usually still the same as their paper-and-pencil predecessors (Bartram, 2000). However, it is necessary to ensure that the psychometric properties are the same for two reasons: First, companies may use both versions and compare applicants undergoing paper-and-pencil and computer-based assessments, and second, equivalent scores are required in order to use the norms traditionally gleaned from the paper-and-pencil version (Meade, Michels, & Lautenschlager, 2007). Most studies found encouraging results (Bartram & Brown, 2004; Chuah, Drasgow, & Roberts, 2006; Salgado & Moscoso, 2003) and even some benefits of web-based testing (e.g., more normal distribution or higher reliabilities, Ployhart, Weekley, Holtz, & Kemp, 2003). Nevertheless, there are differences (e.g., concerning means, Ployhart et al., 2003), and Meade et al. (2007) warn that comparability cannot be taken for granted. Practitioners in our sample strongly preferred a computer application over a paper-and-pencil application ($M = 2.38$, $SD = 1.62$), but were indifferent as to whether the test should be applied via the internet or on-site ($M = 3.51$, $SD = 1.93$). Moreover, there is no preference regarding who (the company/ the practitioner or the test publisher) evaluates test results ($M = 3.51$, $SD = 1.97$), meaning that the focus seems to be on an automated process and not on the way in which this automation is delivered (by an external provider, on-site or via the internet). On the other hand, actual test users do prefer an application by internet ($n = 28$, $M = 2.61$, $SD = 1.77$)

compared to non-users ($n = 138$, $M = 3.70$, $SD = 1.87$), meaning that people who already use personality tests seem to perceive the advantages of this medium.

Declaration and description of quality criteria

Quality criteria, especially measures of reliability and validity (which we will subsume with the term quality criteria in the following), are very important to researchers, who consequently present these measures in extensive test manuals, as do some commercial test publishers. However, considering the huge amount of personality tests available, the extent to which publishers are interested in measuring and providing quality criteria likely varies. Besides, the existence of quality criteria does not mean that practitioners have access to such information before buying a test. There is a huge variety of ways in which quality criteria can be reported: According to our experience, information on publishers' or distributors' websites is (a) seldom extensive, (b) often only brief, (c) sometimes only available on demand or by buying the manual, or (d) not available at all. A brief description of quality criteria may be an alternative which is more convenient to practitioners, as they probably do not have the time to read long manuals. In our study, practitioners significantly preferred succinct statements about quality criteria rather than extensive information ($M = 4.42$, $SD = 1.36$), and brief information about benefits rather than detailed reports ($M = 2.88$, $SD = 1.49$). Nevertheless, they do not seem to be naïve in terms of believing these statements, as they strongly prefer to check this information rather than trusting the declarations of the author ($M = 2.72$, $SD = 1.49$). At the same time, actual test users significantly preferred more detailed reports about the benefits of a certain test ($n = 28$, $M = 3.57$, $SD = 1.69$) compared to non-users ($n = 138$, $M = 2.74$, $SD = 1.41$), whereas there was no difference concerning the length of quality criteria information. Consequently, there is perhaps more to selling personality tests than numerical criteria. Moreover, no significant results were found regarding the question of whether practitioners would prefer a theoretical explanation of why the measured traits should be important for their employees' professional performance compared to statistical measures ($M = 3.61$, $SD = 1.53$).

Once again, this indicates that both kinds of information are needed, and more is needed to convince practitioners of the benefit of personality tests in personnel selection than the scientists' mere focus on proving validity data.

Finding a personality test

An additional criterion that distinguishes personality tests is where and how practitioners can find information about them. In our survey, practitioners significantly preferred to inform themselves by searching websites and flyers rather than professional journals and magazines⁴ ($M = 3.26$, $SD = 1.56$). Answers to an open question concerning sources revealed that most used the internet (35.5%), information and recommendations from their personal network (12.7%), and professional (HR-related) magazines (12.0%). They strongly favored tests used by many companies rather than tests that set them apart from other companies ($M = 2.69$, $SD = 1.29$), a confirmation of the finding of König et al. (2010). Most commercial publishers seem to take advantage of this practice of using recommendations, by citing referees who predominantly work in well-known companies on their websites. These references do not necessarily contain any information about the frequency and reason of use in the respective company. Moreover, our sample preferred to compare a small pre-selection of tests rather than many different tests ($M = 4.70$, $SD = 1.28$), even more so when they were not currently using a personality test ($n = 138$, $M = 4.80$, $SD = 1.22$) than when they were already using one ($n = 28$, $M = 4.21$, $SD = 1.45$), which might not be too surprising considering the huge amount of tests available.

Another criterion that may affect the selection of a personality test is whether practitioners have to gain a certificate to use a special test (i.e., some publishers do not sell their inventories or at least part of them to people who are not trained and certified, and others offer training as an additional service, i.e. the MBTI or the HBDI certification). Practitioners in our

⁴ In our survey we used the German word „Fachzeitschrift“ that includes professional and peer-reviewed journals as well as magazines.

sample did not have a particular preference for or against certification ($M = 3.60$, $SD = 1.72$), although actual test users prefer certification ($n = 28$, $M = 2.89$, $SD = 1.77$) compared to non-users ($n = 138$, $M = 3.74$, $SD = 1.68$). Offering training seems reasonable, at least for non-psychologists, who have probably not had such training during their education, because otherwise, there is no guarantee that users are really informed about the proper application and interpretation of results.

An additional factor which is important in the decision-making process but is not covered in our survey⁵ concerns the promotion of personality tests. Promotion strategies may differ to various extents: For example, there is “classic” advertisement in HR journals or stands at HR fairs. In addition, some may rely on a factual strategy, while others may (consciously or unconsciously) emphasize special characteristics of their tests in the sense of a unique or emotional selling proposition (e.g., the HBDI stresses a metaphorical connection to the brain) (Herrmann International, *n.d.*) or point out the model of personality upon which the test is based (i.e., they can highlight that their tests rely on well-established models, for instance the MBTI on Jungian theory). Whatever their strategy, commercial test publishers probably invest a lot in their promotion strategies in order to stand out from the crowd of personality tests.

Future prospects and Conclusions

We were able to show that – at least in Germany – there is definitely more to personality testing than just the Big Five or MBTI, and we believe that it is necessary to gain a broader overview, an international appraisal of actual personality test use rather than to focus solely on particular single measures. Not only is there a large range of personality tests offered to practitioners, but many of them are also in use in the context of personnel selection. We discussed the influence of different criteria on the decision-making process, such as certain test characteristics, the different ways of presenting results or aspects of application. We believe

⁵ Practitioners probably do not know anything about promotion strategies and they cannot consciously evaluate the effect of promotion on their decision.

that a deeper understanding needs to be gained of this decision-making process, the requirements and needs of practitioners and the advantages and disadvantages of the manifold alternatives. For instance, we know nothing about the quality and actual handling of narrative reports in the selection process. Moreover, we concentrate strongly on dimension-based tests without even considering whether types might somehow meet practitioners' needs. Although the development of the Big Five certainly has great advantages in terms of comparability, it may not fit with categories of practitioners in personnel selection. Other traits or competencies may be more important to them because they are meaningful in terms of showing an intuitive theoretical relation to job performance. We need to find out a lot more about how personality tests are actually used, what may influence the decision for implementing personality tests in the selection process, and how attitudes to personality tests may change before and after this implementation. Our survey was only a first attempt to learn something about practitioners' needs and requirements concerning the use of personality testing in personnel selection and to initiate a change in perspectives – away from believing that reliability and validity are the only criteria important to practitioners towards an understanding of the existence of multiple influences. Possibly, there may be many more criteria according to which personality tests can be differentiated (e.g., whether items relate to organizational contexts, to clinical contexts or neither, the costs of one or several applications, the number of dimensions or types measured, item format, how dimensions and types are named, whether they are special tests for different roles like leaders or salesmen, etc.). It will be the task of future research to use this new perspective to develop arguments for propositions and specific hypotheses concerning the influence of different criteria to the decision-making process of practitioners.

Moreover, it is not enough merely to survey practitioners, as questionnaires are prone to socially desirable responding (as it probably happened in our question whether practitioners would prefer to check quality information or trust declarations of the author). Rather, practitioners' decision making needs to be experimentally analyzed.

In addition, it may be necessary to take a step away from pure research and to try to diminish the research-practice gap in personnel selection. One such step may be to simplify the search and comparison of different personality tests by setting up national websites that list personality tests categorized according to their benefit for different purposes (e.g., development, selection, general assessment of personality, ...) and provide the most important information and professional and independent evaluations of common tests. Another step may be to develop training programs for different personality tests in different organizational contexts in order to improve actual test use.

As personality tests continue to be used – no matter how scientists evaluate this – it is important to understand this use and make adequate recommendations and offers to practitioners. Thus, not only do scientists need to be better in explaining validity to inform practitioners about its value, they should also not ignore the needs and requirements of practitioners and should therefore try to adapt their research priorities accordingly.

KAPITEL 3**THE ROLE OF NEUROSCIENCE INFORMATION IN CHOOSING A
PERSONALITY TEST: NOT AS SEDUCTIVE AS EXPECTED**

This article has been published in the *International Journal of Selection and Assessment* and may not exactly replicate the authoritative document published there. It is not the copy of record. The authoritative document can be found at the following source:

Diekmann, J., König, C. J., & Alles, J. (2015). The role of neuroscience information in choosing a personality test: Not as seductive as expected. *International Journal of Selection and Assessment*, 23, 99–108. <https://doi.org/10.1111/ijsa.12099>

Abstract

There is a large number of different personality tests used by practitioners for selection purposes, many of which are promoted by commercial test publishers and do not measure the Big Five. The present study examined one particular promotion factor used by at least one successful personality test: the influence of a link between the brain and personality on the decision for or against a personality test. This factor was chosen as past research has demonstrated the seductive appeal of neuroscience information on judgments in other areas. Three samples consisting of human resource practitioners, business management students, and psychology students rated two versions of a fictitious personality test, one with neuroscience information and one without. Contrary to our expectations, the personality test with neuroscience information was rated more negatively than the same test without this information. Human resource practitioners in particular reacted negatively, whereas psychology students were not influenced by our manipulation.

Introduction

In recent years, many authors have referred to a persistent, possibly even increasing gap between science and practice (e.g., N. Anderson et al., 2001; Rynes et al., 2007). The field of industrial and organizational psychology faces the particularly troubling problem that scientific findings in human resources (HR) management and recommendations based on these findings are not used in practice (at least not as much as we would hope). A classic example can be found in the field of personnel selection (e.g., N. Anderson et al., 2004; Klehe, 2004; Terpstra & Rozell, 1997): Here, there still seems to be a preference for unstructured procedures like the unstructured interview (e.g., Stephan & Westhoff, 2002), and when structured methods like personality tests are used, these tests often seem to be “poorly chosen” (Hough & Oswald, 2005, p. 375).

Although many authors (e.g., Murphy & Dzieweczynski, 2005) have expressed their unease regarding the success of non-scientific tests and have identified the need to provide practitioners with better advice (Hogan, 2005), almost no research has explored why personality tests with questionable qualities seem to be attractive to practitioners. Many nonresearch-based personality tests (as well as some research-based tests) are sold by commercial test publishers, which have to promote their tests, and this entails making a choice with regard to positioning and finding a unique and/or emotional selling proposition (Barrena & Sánchez, 2009; Heath, Brandt, & Nairn, 2006; Ries & Trout, 2001; Rossiter & Bellman, 2012). This is especially necessary because, at least in the United States and Western Europe, the test market seems to be highly saturated (Hough & Oswald, 2005; Oakland, 2004). One particularly striking promotion argument is the link between a personality test and the brain, which is used, for example, by the Herrmann Brain Dominance Instrument (HBDI), or even neuroscience information. Such a promotion strategy is consistent with research showing that localization hypotheses attract not only scientists (e.g., DeYoung et al., 2010) but also the popular press (Racine, Bar-Ilan, & Illes, 2005), and that the brain and neuroscience information seem to be

especially seductive (Weisberg, Keil, Goodstein, Rawson, & Gray, 2007). The main purpose of the experiments reported here was to show that a personality test which is promoted with neuroscience information is more attractive than the same test without such a link with the brain.

Background

The gap between research and practice is a well-known phenomenon in industrial and organizational psychology: scientific findings in HR management – recommendations of management techniques based on established theories and advice – are seldom used in practice, whereas some methods, which actually are in use, are not well supported by scientific research (Johns, 1993; Rowe et al., 1994; Terpstra & Rozell, 1993, 1997). Scientists and practitioners seem to agree that (a) scientists' research is not always adjusted to practitioners' needs and (b) there is a gap in terms of transferring scientific findings to practice (Shapiro et al., 2007). Concerns have been raised that this gap is growing, and might be hard to reduce (N. Anderson et al., 2004; Hodgkinson et al., 2001; Rynes et al., 2007). This implies the troubling effect that academics, and their efforts, may only have a moderate impact outside the world of science and its journals (Abrahamson & Eisenman, 2001).

One special part of the research–practice gap refers to personnel selection (e.g., N. Anderson et al., 2004; Terpstra, 1996). The literature indicates a striking popularity of unstructured procedures in personnel selection, combined with a high confidence in one's own experience and intuition (Highhouse, 2008). One prominent example is the pronounced preference for the unstructured interview (Lievens & De Paepe, 2004; Stephan & Westhoff, 2002), even though research has shown that standardization leads to higher validity (Kepes et al., 2012; McDaniel et al., 1994). The role of general mental ability in personnel selection, as another example, has been emphasized for decades now (Schmidt & Hunter, 1998), but tests of cognitive ability are used only occasionally (e.g., companies surveyed by Ryan, McFarland, Baron, & Page, 1999, indicated the extend of use of cognitive ability tests with 21-50%, and

only 30% of German companies use cognitive ability tests with at least some applicants, Schuler, Hell, Trapmann, Schaar, & Boramir, 2007). On the other hand, even methods like graphology, which have been found to be inappropriate in selection procedures (Driver et al., 1996), still find some supporters (Berchtold, 2005; Di Milia, Smith, & Brown, 1994; Ryan et al., 1999; Shackleton & Newell, 1994), although this is not as widespread as is sometimes believed (Bangerter, König, Blatti, & Salvisberg, 2009).

Another gap can be found when considering the use of personality tests in personnel selection. The Big Five personality dimensions are highly accepted among most researchers, and studies and discussions about the validity of personality measures very often focus on these five dimensions (e.g., Barrick et al., 2001; Hurtz & Donovan, 2000; Morgeson et al., 2007b; Ones, Dilchert, Viswesvaran, & Judge, 2007; Tett & Christiansen, 2007). However, personality testing consists of more than just the Big Five: Hogan (2005) reports that there are about 2,500 test publishers in the United States and Oakland (2004) believes that about 5,000 standardized tests are developed, sold and used in Western countries. These tests differ in content (e.g., traits, motivation, emotions), the extent of validity, type (e.g., psychometric vs. projective, dimension vs. type) and other effects (for an overview of some frequently used tests, their differences and characteristics, see Diekmann & König, 2015). Many authors (e.g., R. Hogan, 2005; Hough & Oswald, 2005; Hülshager & Maier, 2008; Johns, 1993; Klehe, 2004; Murphy & Dzieweczynski, 2005; Tewes, 1998) have argued that most of these personality tests do not measure the Big Five personality dimensions, and have not been constructed on scientifically based concepts or validated based on external performance criteria. In particular, the Myers–Briggs Type Indicator (MBTI), as one of the most famous and most widely used personality tests in the world, has often been criticized and found to be inappropriate for applications in organizational settings (e.g., Murphy & Dzieweczynski, 2005; Pittenger, 2005).

As many personality tests are sold by commercial test publishers, which promote their tests on their websites, in flyers and in contacts with potential clients, there is likely more to a

decision for or against a special personality test than “just” scientific quality criteria. This paper seeks to prove the effect of one possible promotion attribute: the link between personality and the brain, as implied by the HBDI. The HBDI is a US product that is used all over the world (Herrmann International, 2013) and that has been translated into 18 languages (Herrmann International, 2015b). According to the various local websites, it has been completed by more than two million people worldwide and is used by up to 70% of the Fortune 500 companies (e.g., Herrmann International, 2015b; Herrmann International UK, 2015). Although none of these websites recommend using the test in personnel selection, there is evidence that the HBDI is used in selection contexts (Berchtold, 2005).

The HBDI was constructed by personnel development specialist Ned Herrmann (1989) and is based on the idea of hemispheric specialization. It assigns people to four different types, which correspond to four brain quadrants made up of two dimensions (left hemisphere vs. right hemisphere and cerebral vs. limbic). Although historically speaking, it began with neuroscience findings, the test publisher now underlines the metaphorical link to the brain, which only serves as a model (Herrmann, 1989). Nevertheless, a link to the brain is pointed out (Herrmann, 1989). This special characteristic may well serve as an emotional selling proposition and therefore be a reason for the test’s diffusiveness (cf. Barrena & Sánchez, 2009; Heath et al., 2006; Ries & Trout, 2001; Rossiter & Bellman, 2012).

The link between the brain and personality serves as an emotional selling proposition and people may be attracted to tests using this link. This can be demonstrated by the increasing influence of neuroscience information in the public domain and by studies concerning the effects neuroscience has on people. Neuroscience has gained a great deal of influence in the past few years, both in scientific journals and among the general public (Canli & Amin, 2002; Illes, Kirschen, & Gabrieli, 2003; Poldrack, 2008; Racine et al., 2005; Rose, 2003; Roskies, 2008). Articles usually regard matters of neuroscience and the underlying techniques such as functional magnetic resonance imaging (fMRI) positively (Racine et al., 2005; Racine, Bar-

Ilan, & Illes, 2006), even though neuroimages have to be interpreted and therefore proneness to error has to be considered (Farah & Wolpe, 2004; Klein, 2010; Kretschmann & Weinrich, 2007; Poldrack, 2006). Especially when it comes to localization hypotheses, a deep understanding of the technique is necessary to understand the results, but usually this is not transported through public media, which tends to simplify and sometimes even misinterpret results (Beck, 2010; Cacioppo et al., 2003; Farah, 2005; McCabe & Castel, 2008). Not surprisingly, laypersons often misunderstand findings or overestimate them (Dumit, 2004; Weisberg, 2008). Furthermore, not only are a huge number of such findings made public, they are also very convincing to layperson readers (Dumit, 2004; Poldrack, 2008). Weisberg et al. (2007) were able to show that poor explanations of different psychological phenomena were rated as more satisfying when supported with (logically irrelevant) neuroscience information. McCabe and Castel (2008) found that agreement with the conclusions of studies concerning neuroscience was higher when there was a brain image than when no such image was provided. In line with other research showing that people sometimes use some kind of heuristics to evaluate explanations (e.g. Keil, 2006; Lombrozo, 2006; Trout, 2002), Weisberg et al. (2007) believe that neuroscience might be such a heuristic. Moreover, in public media, neuroscience information frequently serves as explanation of human behavior and may therefore be seen as explanatory information (Beck, 2010; Racine et al., 2005).

Given people's fascination with neuroscience research and information (as reflected in the high presence of such information in the popular press) and its "seductive allure" (Weisberg et al., 2007, p. 470), as reflected in higher ratings of information with than without neuroscience explanations or images, personality tests that refer to neuroscience information should be more compelling than tests without. More formally, we hypothesize:

H1: A test that refers to neuroscience information will be evaluated as more attractive than the same test without such information.

Weisberg et al. (2007) found different effects for laypersons and experts of neuroscience: laypersons rated poor explanations with neuroscience information as more satisfying, whereas no such main effect was found for experts of neuroscience. This finding leads to the question of whether different samples react differently to neuroscience information. In contrast to Weisberg et al. (2007), we are interested not only in differential effects of neuroscience expertise, but also in possible effects of HR and personality expertise because this kind of expertise might have an impact on decisions in practice. Therefore, we acquired three samples: first, a sample of HR experts with extensive knowledge about personnel selection; second, a sample of business management students with basic knowledge about personnel selection but probably no experience in neuroscience; and third, a sample of psychology students who should have basic knowledge in both areas.

The view of HR practitioners is considered as the most important when addressing such an issue as the research–practice gap. HR experts – at least in Germany – very often have a business management education but are seldom trained in psychology (Deller, Süßmair, Albrecht, & Bruchmüller, 2005; Haufe eCampus Redaktion, 2012; for evidence in other countries see, e.g., Hoque & Noon, 2001; König et al., 2010), so they probably know little about neuroscience. At the same time, they are very likely to be in contact with consultants and test publishers, meaning that they may be accustomed to extensive promotion messages. We therefore expect a moderate effect of neuroscience information in HR practitioners. The business management students, with only basic knowledge about personnel selection and no knowledge of neuroscience, should show a stronger effect as they have no experience with personality tests or the advertisement of personality tests. Psychology students, in contrast, have basic knowledge in personnel selection as well as personality theory and neuroscience. Moreover, they should be trained in the critical evaluation of neuroscience findings. We therefore expect only a small or even negative effect of neuroscience information in this group. Accordingly, we hypothesize:

H2: The positive effect of neuroscience information on the judgment of personality tests should be strong for business management students, moderate for HR experts, and small for psychology students.

Method

Participants

HR experts. Participants were personnel managers in the German financial industry. They all were employed by so-called “Sparkassen” – public banks that work as independent and decentralized institutions under local management with own personnel staff but that all belong to the same umbrella organization with 422 sites across Germany (as of June 2013). They were contacted by email and sent an invitation to an online survey operated via Unipark (QuestBack GmbH). No compensation was offered. A total of 109 HR experts (52.3% male, 34.9% female, 12.8% did not specify their sex) participated in this experiment. Almost half of them (44%) had completed an academic education, mostly having studied business management (50%), with a focus on business studies (22.9%) or psychology (12.5%). They had been working in the field of HR for an average of 14.5 years ($SD = 9.7$); 83.5% were involved in personnel selection procedures and 70.6% in decisions about which selection procedures are used. Personality tests were already used by 24.8% of the respondents. Once they had started the survey, participants were welcomed and were given some information about informed consent conditions. They were asked to think of a situation in which they wanted to supplement their selection procedure with a personality test and were randomly assigned to the with ($n = 64$) or without ($n = 45$) neuroscience condition, which differed in the subsequent material presented.

Business management students. Potential participants were contacted by email distribution lists with business management students of different German universities. All participants had the chance to win one of five 10€ Amazon gift cards. Of the initial sample, only those studying business management or similar were included in the analysis. A total of

108 participants (35.2% male, 61.1% female, 3.7% did not specify their sex) remained, the majority of whom were still studying in a bachelor's degree program (70.4%). Some participants (47.2%) already had some practical experience with personnel selection (e.g., during an internship), but only 9.3% had experience with personality tests. In this group, basic knowledge concerning selection procedures as well as sufficient inexperience with personality tests can be assumed. Again, participants were randomly assigned to the with ($n = 50$) and without ($n = 58$) neuroscience condition.

Psychology students. Participants were recruited from an introductory class of industrial and organizational psychology (which also covered a basic introduction to personnel selection) at a German university in the middle of a semester. No incentive or compensation was provided. According to the module manual, these second-year students were supposed to have an introductory class of clinical neuroscience in the same semester and to have completed two classes of biopsychology (covering the neuroscience side of psychology) in the previous year and should therefore have some first impressions about personnel selection as well as neuroscience. Participants were randomly assigned to the with ($n = 54$) and without ($n = 47$) neuroscience condition. After all of the students who were not studying psychology had been excluded, 101 participants (21.8% male, 77.2% female, 1.0% did not specify their sex) remained and were included in analysis. Some participants (19.8%) already had practical experience with personnel selection (e.g., during an internship), but only 6.9% had experience with personality tests.

Overall sample. In total, 318 people (36.8% male, 57.2% female, 6.0% did not specify their sex) participated, 168 of whom completed the with neuroscience condition and 150 the without neuroscience condition.

Material

To avoid bias concerning a real test publisher that might be known by participants (especially the HR experts) and to avoid copyright problems, we created material that refers to

a fictitious personality test. Based on websites and flyers for different personality tests (e.g., HBDI, Thomas International, and MBTI), a flyer for a personality test called Personality at Work Inventory (P-WIN) was created. This flyer explained the benefits of capturing personality in personnel selection, described twelve factors or dimensions which the test was supposed to measure as well as some details on application and reporting. We developed two versions of this flyer, which differed in their reference to neuroscience. The first version had no relation to neuroscience, stating that the test was developed in cooperation with scientists based on new scientific findings and showing a picture of three figures climbing ladders. The second version stated that the test was developed in cooperation with scientists based on new neuroscience findings and that personality traits can be detected through the activation of different brain areas. This version was illustrated by an fMRI image (see Figure 5). See also Appendix A for the description of the scenario and the two version of flyers.

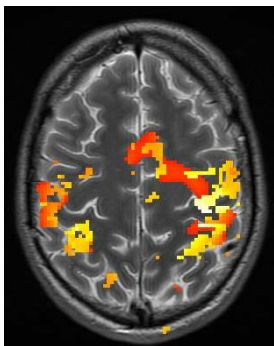


Figure 5. fMRI image used to illustrate the test version with neuroscience information in chapter 3. From “Brain activation for the movement of fingers measured in an fMRI experiment” by K. Tabelow, 2009, <http://www.mathematik.hu-berlin.de/~tdm/2009/abstract.php?name=tabelow&id=uf5pYyDxgBv5I>. Copyright 2009 by Karsten Tabelow. Reprinted with permission.

Dependent variables

Participants were asked to evaluate the described personality tests on six items using a 7-point rating scale (strongly disagree to strongly agree) referring to adequacy (Item 1) and objectiveness (Item 2) of the test in selection contexts, content (Item 3), and design (Item 4) of

the flyer, interest in additional information on the test (Item 5) and overall liking (Item 6). Beyond this, participants were able to comment on what had affected their ratings using an open-format item. The six items showed a good reliability (Cronbach's $\alpha = .87$, which was comparable over all three subsamples: HR experts $\alpha = .89$; business management students $\alpha = .89$; psychology students $\alpha = .83$) and were therefore combined to form a scale serving as a measure of overall test rating.

Results

Means and standard deviations of all groups can be found in Table 1. Please note that Levene's test for equality of variances indicated that the variability of each group was approximately equal, $F(5, 312) = 1.49$. Furthermore, the effects are visualized in Figure 6.

Table 1
Overall test rating, standard deviations and effects size for sample and subsamples in chapter 3.

Sample	Test version	<i>n</i>	<i>M</i>	<i>SD</i>	<i>d</i>
Overall	with	168	4.05	1.11	-0.29
	without	150	4.39	1.15	
HR experts	with	64	3.79	1.17	-0.41
	without	45	4.28	1.25	
Business management students	with	50	4.15	1.67	-0.31
	without	58	4.51	1.17	
Psychology students	with	54	4.27	0.91	-0.07
	without	47	4.34	1.04	

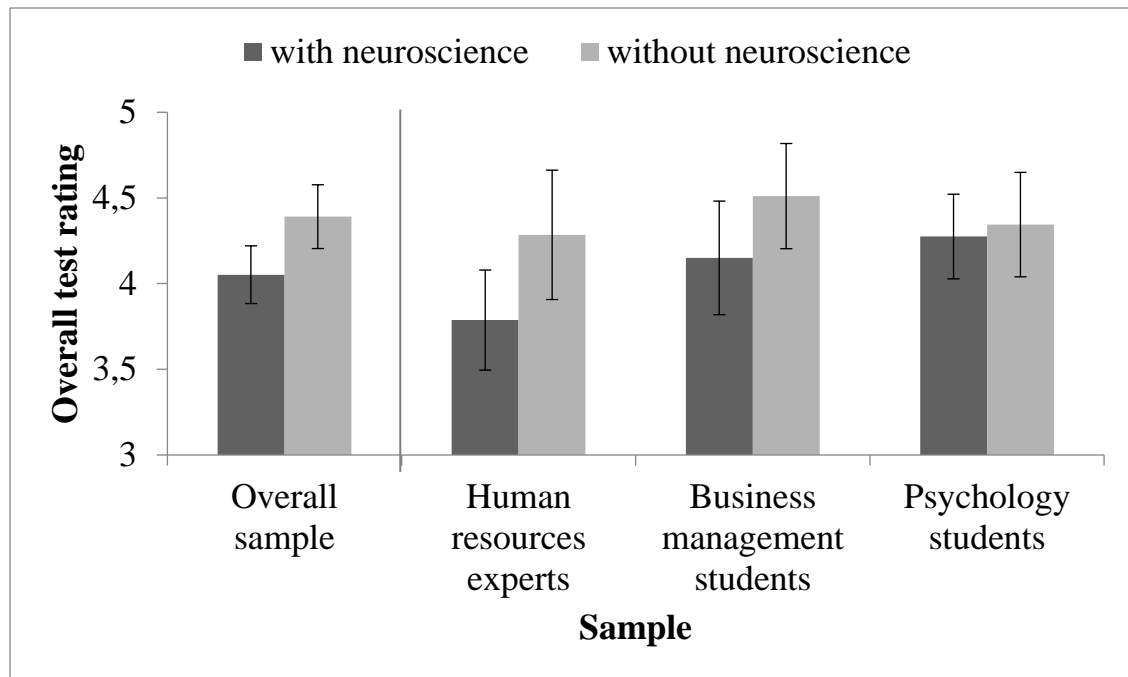


Figure 6. Overall test rating separately for sample and subsamples in chapter 3. Error bars indicate 95% confidence interval.

To find out whether the addition of neuroscience information to the flyer of a personality test has any effect on the assessment of the fictitious P-WIN test (*H1*), we analyzed the ratings using a 2 (with vs. without neuroscience) \times 3 (HR experts vs. business management students vs. psychology students) analysis of variance. There was a significant main effect of flyer version, $F(1,312) = 5.94$, $p < .05$, indicating that neuroscience information has an impact on ratings. Contrary to our expectations, the flyer with information (overall $M = 4.05$, $SD = 1.11$) was rated more negatively than the flyer without neuroscience information (overall $M = 4.39$, $SD = 1.15$); thus, our hypothesis was not confirmed, and was instead disproved (see Figure 6, Table 1). There was no main effect of subgroups, $F(2, 312) = 2.25$, $p = .11$, indicating that ratings between the three subgroups were equal. There was no interaction between flyer version and subgroups, $F(2, 312) = .97$, $p = .38$.

Although the effect was contrary to *H1*, we can still analyze whether there are differences between groups concerning the size of effect (*H2*). The descriptive statistics (see Table 1) and a visual inspection of results (see Figure 6) indicate that the (negative) effect of

neuroscience information was greatest for HR experts, moderate for business management students and small for psychology students. We tested this adapted *H2* with linear contrasts, and it showed a significant difference, $t(315) = 2.01$, $p < .05$, indicating that neuroscience information had differential effects on subsamples.

Discussion

The research-practice gap in personnel selection has often been mentioned (e.g., N. Anderson et al., 2004; Klehe, 2004; Terpstra & Rozell, 1997), and many authors (e.g., Hough & Oswald, 2000; Murphy & Dzieweczynski, 2005) have lamented in particular why personality tests used in personnel selection practice quite often do not meet scientific criteria. The present study was designed to investigate whether a personality test using neuroscience information in its promotional flyer is more attractive than the same test without such information. Contrary to expectations, we found that neuroscience information served as a negative cue: test flyers with this information received lower ratings than test flyers without this information. An additional contrast analysis revealed that HR experts were most negatively impacted by neuroscience information, whereas psychology students were hardly affected.

These findings reveal some good and some bad news. The bad news is that our results show how easily decision processes regarding selection tools can be influenced. Our manipulation was quite small, differing only in a picture and two sentences at one point in the flyer – nevertheless, it affected HR experts and, to a lesser extent, business management students. This is alarming as it shows (once more) that quality criteria and the fitting of dimensions to requirements are not the only aspects which play a role in the decision for a personality test (cf. König et al., 2010). Moreover, it may be possible that such small differences in promotion might be even more influential than differences in quality criteria if such a small manipulation works. As commercial test publishers probably invest a lot more time and money in the development of a beneficial promotion strategy than scientists, our findings can

contribute to the question of the research-practice gap and why some personality tests used in practice are not what we expect them to be.

The good news is that particularly, HR experts did not fall into the trap and not only resisted but even contradicted the assumed seductive allure of neuroscience. This is good news at least for the field of personnel selection, since it suggests that HR practitioners cannot be easily entrapped with some nice brain scan pictures and the like, in marked contrast to the perception that a huge amount of public literature typically reports positively about neuroscience findings (e.g., Pillay, 2011; Racine et al., 2005; Racine, Waldman, Rosenberg, & Illes, 2010; Weisberg, 2008). One reason for our finding might be that HR experts are frequently exposed to personality test promotion in the form of information flyers or via telephone calls from commercial test providers. Therefore, they could be sensitized to emotional selling propositions of this kind and react critically if they judge certain statements as questionable. Another or additional reason could be that the huge amount of articles about neuroscience findings in the popular press has the same effect: People might have become more reticent regarding the issue, possibly because the technique is no longer new and exciting, and therefore, results are seen more impartially.

Although HR experts and, at least in part, business management students reacted negatively toward using a neuroimage for promotional purposes, psychology students were not particularly critical toward the manipulation (despite knowledge about neuroscience). It seems that the impact of neuroscience information is not a general one but depends on one's training and employment background. Our results also suggest that training in the field of neuroscience does not lead to a critical attitude toward this subject. Weisberg et al. (2007) even found that their students (recruited from a neuroscience course) rated explanations with neuroscience information even more positively than explanations without. Thus, training in neuroscience and its limitations seems to lead to a favorable attitude toward the technique of fMRI, at least as long as people are not neuroscience experts (Weisberg et al., 2007). Therefore, the question

arises of whether we educate our students appropriately. Our passion for our subject, regardless of whether it is neuroscience or personnel selection, probably affects our way of teaching. An advocate of personality tests in personnel selection may communicate a different view to his or her students than an opponent. A researcher in neuroscience may impart a positive attitude of his subject to his or her students. Perhaps we sometimes need to take a step aside and reconsider the content of our classes and the way in which we impart our knowledge.

At the same time, our results also question the general “seductive allure of neuroscience explanations” (Weisberg et al., 2007, p. 470) and imply that findings from the general literature about the effects of neuroscience information (Weisberg et al., 2007) have to be reconsidered accordingly. In fact, our study is in contrast with Weisberg et al. (2007) and more consistent with very recent research that found no effects of neuroimages (e.g., Hook & Farah, 2013; Schweitzer, Baker, & Risko, 2013). The differential results of our study suggest that experience and training background play an important role in the decision process, at least concerning the influence of neuroscience information as part of the promotion strategy. This is in sharp contrast to Weisberg et al. (2007), who found positive effects of neuroscience information on poor explanations for neuroscience particularly for laypersons (like HR experts).

More generally, this study suggests that it is beneficial to search for tailored solutions to specific facets of the research-practice gap. This gap concerns many different aspects of HR instruments, with personnel selection and personality tests being one part of it. Although many general solutions have already been suggested, in particular on how to conduct scientific research or how to communicate scientific findings to practitioners (Buckley et al., 1998; Gelade, 2006a; Huff, 2000; Van de Ven & Johnson, 2006), facets of the research-practice gap seem to be better understood if an additional, more focused perspective is additionally chosen, such as a close look at how tests are promoted.

As always, our study is not free from limitations. First, it should be mentioned that the setting was a simulation, as participants were not in a real situation of a decision process

resulting in the choice for a personality test. However, to test our causal hypotheses, we needed an experimental design, and convincing test publishers to manipulate their promotion material was deemed impossible. Furthermore, personality tests are not used that often (e.g., Diekmann & König, 2015; Schuler et al., 2007), making the decision for or against a test a rare event. Therefore, the disadvantage of the use of a hypothetical scenario must be considered together with the advantage of being able to draw causal conclusions. Moreover, there is a high correlation of laboratory and field effects in industrial and organizational psychology (G. Mitchell, 2012), suggesting that lab studies generally generalize fairly well to the field.

Second, a decision process is likely more complex than in our simulation. We did not simulate a complete decision process in which different alternatives are likely compared and in which more variables – individual as well as organizational and situational factors – come into play. Previous research has already shown that aspects such as validity, anticipated applicant reactions, and costs are generally important for decision makers (König et al., 2010), and there might be even more aspects (see Diekmann & König, 2015) because decision making research shows that rationality in decision making is bounded (Simon, 1972, 2000) and heuristics are used (Tversky & Kahneman, 1974). The impact of these features in general and especially in combination with differences in validity aspects has yet to be explored. Such exploration could be achieved using both experimental and correlational designs, with the former offering causal explanations in a controlled setting and the latter offering the inclusion of other potential variables such as the general attitude toward personality tests.

Third, we conducted our study with German participants. In the past, a greater reservation concerning personality testing in Germany than in other countries has been mentioned (Schuler et al., 1993), and nothing is known about the coverage of neuroscience information in public media in Germany in comparison to other countries. Therefore, replication in other countries seems necessary.

Future research should also explore the (perhaps not so) seductive allure of neuroscience in subfields of industrial and organizational psychology other than selection. For example, neuroscience approaches have been used to study leadership (for a critical review of this, see Lindebaum & Zundel, 2013). Although our data showed a negative reaction of HR experts to neuroscience information, it remains to be tested how HR practitioners react to neuroscience explanations regarding leadership issues.

This study has important implications for commercial test publishers and for other institutions selling scientifically constructed tests. On the one hand, our research shows – in line with others (see Fitzsimons et al., 2002) – that some nonconscious aspects like certain promotion features have an impact on potential customers. On the other hand, relatively cheap tricks to sell a test (e.g., adding a brain scan) will likely not do the job, as HR experts will react negatively.

To conclude, the research–practice gap concerning the use of personality tests in personnel selection has often been mentioned (e.g., Murphy & Dzieweczynski, 2005). To our knowledge, only few authors have dealt with the question of which personality tests are actually used (see Diekmann & König, 2015, for an overview) and why. We showed that using neuroscience information is likely a bad idea for promoting a personality test, and we hope that more research will explore other aspects that are also important in the decision process of choosing a personality test for selection purposes.

KAPITEL 4

**FINDING THE RIGHT (TEST) TYPE: ON THE DIFFERENCES BETWEEN TYPE-
VS. DIMENSION-BASED PERSONALITY TESTS AND BETWEEN STATISTICS-
VS. THEORY BASED PERSONALITY TESTS WHEN DECIDING FOR OR
AGAINST A TEST IN PERSONNEL SELECTION.**

This article has been published in the *Zeitschrift für Sozialmanagement/ Journal of Social Management* and may not exactly replicate the authoritative document published there. It is not the copy of record. The authoritative document can be found at the following source:

Diekmann, J., König, C., Schadow, T., Knab, Y., & Harrison, O. V. (2016). Finding the right (test) type: On the differences between type- vs. dimension-based personality tests and between statistics- vs. theory-based personality tests when deciding for or against a test in personnel selection. *Zeitschrift für Sozialmanagement/Journal of Social Management*, 14, 83–102.

Abstract

The personality test market offers a wider range of different tests that human resource (HR) practitioners may use for personnel selection. The decision for or against a specific test is likely affected by different criteria. The current three studies examine two such criteria that have previously been mentioned in the literature – whether a test is type-based or dimension-based, and whether it has a theoretical or a statistical development background. Using different versions of a fictitious personality test, we examined the attractiveness of these conditions with different subsamples of HR practitioners and business management and psychology students. We did not find differences between conditions in any of the studies. Implications of these null findings for selection scientists and practitioners are discussed.

Introduction

Practitioners use a wide range of different personality tests (Berchtold, 2005; Di Milia, 2004). Indeed, their choices are not always easy to understand, as they also use tests that do not seem to fit the purpose of selection well (Diekmann & König, 2015). For example, when the first author of this article applied for a job with a large German company in the automotive industry, a personality test was applied for preselection. Although the company was clearly using a Big Five personality inventory, its test only covered four of the Big Five – and the dimension missing was the one known to have the best predictive validity, conscientiousness (Barrick et al., 2001). It seems that validity was not the only selection criterion.

Diekmann and König (2015) discussed some possible criteria that might affect practitioners' decision for or against a particular personality test beyond the criterion of validity. In particular, they discussed the attractiveness of type-based personality tests (a test that groups people into different classes) compared to dimension-based personality tests (a test that places people on continua) and considered a theoretical development background (a personality test based on a personality theory) compared to a statistical background (a personality test based on factor analytic procedures). They found that HR practitioners described type-based tests and a statistical background as more attractive. The purpose of the experiments reported here was to extend these results and to examine these two decision criteria experimentally.

Background

Several surveys have provided an overview of practitioners' use of personnel selection methods. They show that some kind of interview is almost always used during the selection process, in management as well as non-management positions and with graduates as well as experienced employees (e.g., Carless, 2007; Chartered Institute of Personnel and Development, 2009; Di Milia, 2004; Diekmann & König, 2015; Ryan et al., 1999; Schuler et al., 2007; Taylor, Keelty, & McDonnell, 2002). The use of all other selection methods varies considerably in different countries: Whereas personal references are quite often used in Sweden or Malaysia,

they play only a minor role in the Netherlands or Spain (Ryan et al., 1999). While cognitive ability or personality tests are quite common in the United Kingdom or Australia, they are less popular in Germany and the USA (Carless, 2007; Chartered Institute of Personnel and Development, 2009; Diekmann & König, 2015; Ryan et al., 1999; Schuler et al., 2007).

The situation becomes even more complicated if one focuses on the use of personality tests in personnel selection. There are many different tests used in practice in the process of personnel selection (Berchtold, 2005; Di Milia, 2004; Diekmann & König, 2015), for example the 16 Personality Factor Questionnaire (16PF; Conn & Rieke, 1994), the Herrmann Brain Dominance Instrument (HBDI; Herrmann International, 2015a), the Hogan Personality Inventory (HPI; R. Hogan, 1986), the Myers-Briggs Type Indicator (MBTI; Briggs Myers, McCaulley, Quenk, & Hammer, 1998), the DISC (Marston, 1979) and the Big Five Personality Inventory (NEO; Costa & McCrae, 1992). It is likely that only few of these tests were designed to be applied in selection processes. Indeed, some of the test publishers explicitly recommend not using them for selection purposes and instead stress the benefit of their use for personnel development (Herrmann International, 2015a; The Insights Group Limited, 2015).

Nevertheless, research has focused almost exclusively on the Big Five model, and has even tried to analyze the relationship of other personality tests and models to the Big Five scheme (e.g., Furnham, Moutafi, & Crump, 2003; John, Naumann, & Soto, 2008). There is extensive research concentrating on the predictive validity of these five personality factors concerning different performance criteria such as academic success (Poropat, 2009; Trapmann, Hell, Hirn, & Schuler, 2007), job satisfaction (Judge, Heller, et al., 2002) or job performance (Barrick & Mount, 1991; Hurtz & Donovan, 2000; Salgado, 1997). A stable finding of these meta-analyses is the predictive validity of conscientiousness. The predictive quality of the other four factors differs depending on criteria or occupations but seems to be generally lower than that of conscientiousness.

Due to the often lamented scientist-practitioner gap (e.g., Shapiro et al., 2007), practitioners' knowledge of these research findings remains questionable. In particular, of the American practitioners surveyed by Rynes, Colbert and Brown (2002), only half correctly disagreed with the statement that there are only four personality dimensions, as measured by the MBTI, and even fewer correctly disagreed with the statement that there is little difference in the predictive validity of different personality tests. These results have been replicated for Australia (Carless, Rasiah, & Irmer, 2009), the Netherlands (Sanders, van Riemsdijk, & Groen, 2008) as well as Finland, South Korea and Spain (Tenhiälä et al., 2016). Thus, there seems to be a persistent gap in practitioners' knowledge of research results in personnel selection, implying that reasons other than scientific findings influence the decision to use a particular selection method and a specific personality test.

To better understand practitioners and their use of personality tests for selection purposes, Diekmann and König (2015) took a closer look at the differences between personality tests in use. They reviewed the literature on the use of personality tests for personnel selection, focusing on practitioners' potential reasons for using one or the other test, and also complemented their arguments with a survey on practitioners' preferences. This led them to suggest several hypotheses, two of which we discuss and experimentally test in this article.

The first hypothesis refers to the question of whether dimension- or type-based personality tests are more appealing to practitioners, because the tests actually used by practitioners represent both types of test; for example, the MBTI and the DISC are type-based tests, while the 16PF and the NEO use dimensions. The survey conducted by Diekmann and König (2015) provided first hints regarding the relevance of the type- vs. dimension-based differentiation: HR practitioners in their sample stated that they preferred the aggregation of traits to a type over the dimensional representation of traits. There are several reasons for this preference. A type-based personality test, in which the dimensional information of a scale is reduced to distinct types (e.g., whether a person is introverted or extroverted), may be more

appealing because it is easier to interpret and to compare with a defined requirements specification (e.g., whether the applicant is supposed to be introverted or extroverted) or with other applicants. Moreover, the requirements specification itself may be easier: One only needs to decide whether the applicant is supposed to be one type or the other, while it is not necessary to determine the degree of (for example) introversion and extroversion. This corresponds to the limitation of cognitive capacity (Tversky & Kahneman, 1974) and the human tendency to think in schemata (Smith & Queller, 2008). Therefore, the appealing effect of types may positively influence practitioners' decision in favor of a respective personality test, although psychological researchers mostly adopt a more critical stance (e.g., concerning simplification, cutoff points and uniqueness of assignment to a type; Gangestad & Snyder, 1985; Robins et al., 1998; York & John, 1992) when it comes to the type-based personality tests. However, most HR practitioners are not psychologists (e.g., only 5% of the Diekmann & König, 2015, sample of German HR practitioners) and are thus unlikely to be aware of these critical issues, and they have often not been trained to apply and interpret personality tests. Accordingly, we hypothesize (and test in Experiments 1a and 1b) that:

H1: A personality test based on types will be evaluated more positively than the same test based on dimensions.

Our second hypothesis concerns the development of a given personality test, more precisely the question of whether a more theoretical or a more statistical background affects the decision for or against a personality test. The development of the various tests was influenced by different personality theories, for example the typology of C. G. Jung (Jung, 1960) or the behavioral types of William Marston (Marston, 1979). Jung's theory, consisting of two attitudes (extraversion and introversion) as well as two functions (sensing vs. intuitive perception and thinking vs. feeling judgment) was the starting point for the development of the MBTI (Briggs Myers et al., 1998; Briggs Myers & Myers, 1993), one of the most frequently used tests in business contexts (e.g., Brown, 1999; Furnham, 2008; Muñiz & Fernández-Hermida, 2010;

Ryan & Sackett, 1987, 1992). Marston's types Dominance (D), Inducement (I), Submission (S), and Compliance (C) inspired the development of personality tests like the DISC Analysis. Other tests have a more statistical background, for example the NEO (e.g., Costa & McCrae, 1992; McCrae & Costa, 1997), which is based on the factor-analytically derived Five Factor Model of personality. Cattell's 16PF is also based on factor analytic methods (Cattell, 1978; Conn & Rieke, 1994). Often, personality tests use both approaches: Personality tests that were inspired by theory use statistical methods to evaluate, for example, factor structure and measurement properties, and personality tests based on factor analysis use theories, for example, to construct items. However, whereas some tests, like the NEO, stress the statistical background (Costa & McCrae, 1992), others, like the MBTI, focus on the personality theory (Briggs Myers et al., 1998).

Both approaches may have their attractive aspects for practitioners (Diekmann & König, 2015). On the one hand, the statistical methodology may be appealing as the empirical approach gives the test and its results the appearance of scientific and thus also legal legitimacy, which is also an important decision criterion (Klehe, 2004; König et al., 2010). On the other hand, a test based on factor analysis of extensive universal personality traits does not focus on the work context and may therefore be too general for practitioners. A theory-based personality test may appear well-conceived: The mentioned theories are not new but have been known for years, they may be easily recognized and perceived as well proven, and they may serve the human need for explanation (Keil, 2006; Lombrozo, 2006; Malle, 2004) better than statistics. In the survey by Diekmann and König (2015), practitioners preferred the statistics-based development. We wish to test this expressed preference experimentally (in Experiment 2), and therefore hypothesize that:

H2: A statistics-based personality test will be evaluated more positively than the same test based on a personality theory.

Experiment 1a (Type)

Method

Participants. Participants were HR employees in German companies, found from different listings of companies with more than 500 employees, as we expected mainly large companies to have HR staff involved in personnel selection. We contacted HR departments by telephone and asked for the appropriate contact person concerning personnel selection. If this person agreed to participate, they were invited by e-mail to an online survey operated via Unipark (QuestBack GmbH). As compensation, they were offered a summary of the study results.

In total, 115 HR experts (48.7% female, 47.8% male, 3.5% did not indicate their gender) participated in the first experiment. They had an average of about 13 years of professional experience in HR ($M = 12.9$, $SD = 9.2$). Almost all (92.2%) were personally involved in the selection process; 78.8% had the authority to decide which selection methods were to be used. Approximately two thirds (67.8%) had completed an academic education, mostly in business management ($n = 37$), personnel management ($n = 8$), and psychology ($n = 8$). About one third (33.9%) had completed vocational training, in the majority of cases as industrial clerk ($n = 9$), merchant ($n = 7$), personnel officer ($n = 5$), and office clerk ($n = 4$). In 26.1 % of the companies, personality tests were used in selection procedures, which lies above the 15 % reported by Diekmann and König (2015).

Once they had begun the survey, participants were welcomed and given some information about informed consent. They were asked to think of a situation in which they wanted to supplement their selection procedure with a personality test and were randomly assigned to the type ($n = 53$) or the dimension ($n = 62$) condition, which differed with regard to the subsequent material presented.

Material. The material consisted of a one-page personality test flyer. The design of this flyer was based on websites and flyers of popular personality tests, but the flyer was completely

fictitious as we wished to avoid copyright issues and distortion based on connections to real test publishers that might be known by HR experts. Analogous to Diekmann, König and Alles (2015), we called the personality test the Personality at Work Inventory (P-WIN), and the structure of the flyer was also similar. The flyer was divided into three parts. On the left-hand side, it explained the benefits of using a personality test in personnel selection in general, on the right-hand side it clarified the advantages of this particular personality test, and in the middle it told the reader what P-WIN was supposed to measure. This part differed in the type and the dimension condition. The dimension condition emphasized the existence of eight basic personality dimensions. The word “dimension” was used twice, both times printed in bold. A graph showed the eight dimensions (see Figure 7). Similarly, in the type condition, the existence of eight basic personality types was pointed out and the word “type” was used twice, printed in bold. The graph showed the eight types, which were similar to the dimensions with respect to content (see Figure 8). The content of the eight traits was adapted from various personality tests, but the exact wording was completely invented for this study (see also Appendix B for complete material).

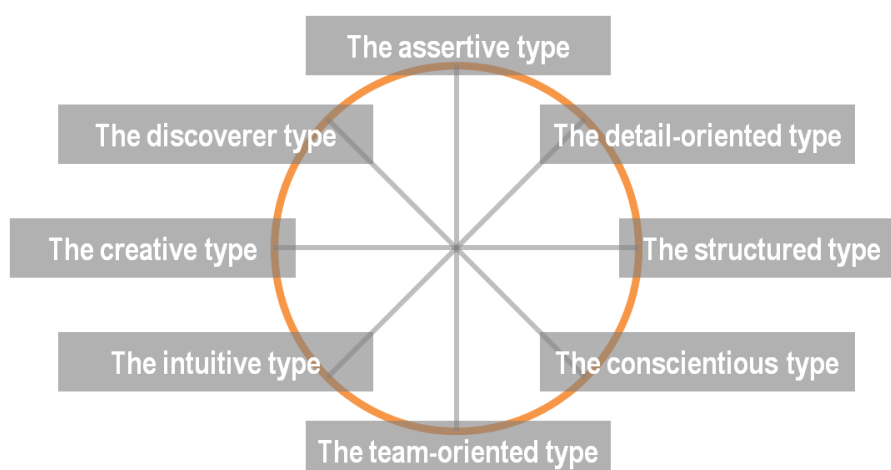


Figure 7. Representation of traits in the flyer of the type-based condition in chapter 4.

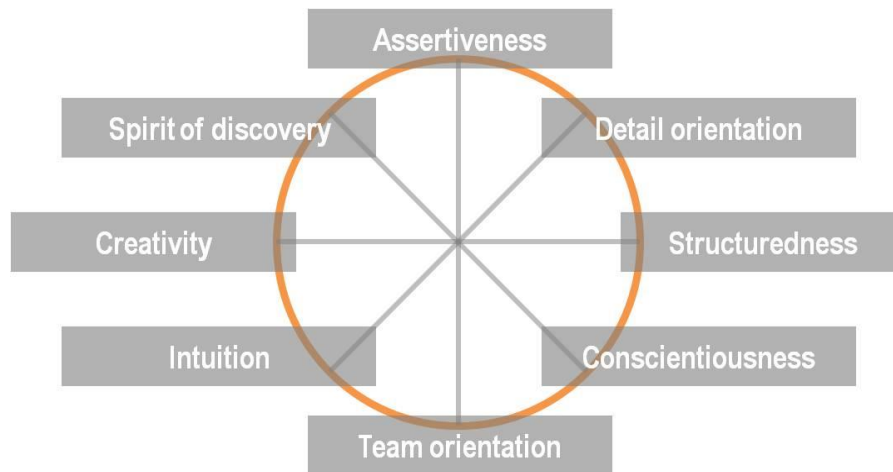


Figure 8. Representation of traits in the flyer of the dimension-based condition in chapter 4.

Dependent variables. After reading the flyer for P-WIN, participants were asked to evaluate this personality test. For this, we used seven items with a 7-point rating scale (1 = *strongly disagree* to 7 = *strongly agree*) as well as one open-format item in which participants were able to comment on which aspects concerning content or design of the flyer had affected their ratings. The seven rating items concerned the suitability of the instrument for personnel selection (item 1), contribution to the objectivity of the selection decision (item 2), attractiveness of the content (item 3), extent of gain of information compared with other selection methods (item 4), interest in more information about P-WIN (item 5), whether the participant would use this test in his or her company (item 6) and the overall liking of the personality test (item 7). As these seven items showed a very good reliability (Cronbach's $\alpha = .91$), they were combined to form a scale of overall test rating.

Results and Discussion

We used a t-test for independent samples to test our hypothesis that a type-based test is more attractive to HR experts than a dimension-based personality test. Contrary to expectation, we found no significant effect, $t(114) = -0.18$, between the type ($M = 3.72$, $SD = 1.29$) and dimension ($M = 3.72$, $SD = 1.18$) condition. Hypothesis 1 was therefore not confirmed.

A possible explanation may lie in the potency of our manipulation. As the two versions of our flyer were fairly similar, our participants may not have been aware of the fact that they were evaluating a type-based personality test or a dimension-based personality test. Overall, 58.3% of the respondents used the possibility to write a comment using the open-format item, but only 3.5% mentioned the word “type” and only 2.6% the word “dimension” (both values increase marginally to 4.3% when including the word “category” for type and the words “component” and “personality factor” for dimension). On the other hand, in the study by Diekmann et al. (2015), a comparable manipulation did work. However, to enhance the strength of our manipulation, we designed a second experiment in which we did not use a flyer, but a results report with two different versions. Our hypothesis that type-based tests are generally more attractive than dimension-based tests remained the same and was tested with business management students as participants.

Experiment 1b (Type)

Method

Participants. The recruitment of participants was twofold: One part (27.5%) was contacted during two lectures at a local technical college in Germany. They were given a paper-and-pencil version of the material and answered the questions during the lectures. No compensation was offered. Another part (72.5%) was contacted by email distribution lists comprising business management students of different German universities. They were offered the chance to win one of five 10€ Amazon gift cards. Of the initial sample, only those studying business management or similar were included in the analysis.

A total of 160 students (66.3% male, 33.8% female) participated in this experiment. Of these, 78.1% were studying at a technical college and 20.6% at a university. Most of them had a study focus on personnel management (16.3%), controlling (15.0%) or logistics (13.1%) and were heading for a bachelor’s degree (79.4%). One third (36.9%) had experience in personnel

selection (for example by having been an applicant in a selection process) and of these, 3.8% stated having experience with personality tests.

Participants using the paper-and-pencil version were verbally welcomed and told about informed consent conditions. On the first page of the material, they were asked to imagine being a human resources employee who wants to supplement their selection procedure with a personality test. On the next page, they were randomly assigned to the type or the dimension condition. Participants of the online version were welcomed and told about informed consent conditions in written form. The subsequent information matched the paper-and-pencil version. Altogether, $n = 78$ participants were assigned to the type condition and $n = 82$ participants to the dimension condition.

Material. For this study, we designed a fictitious results report of the P-WIN personality test in a type-based version and a dimension-based version. The report consisted of two pages. On the first page, there was a logo of P-WIN, the information that this was the “results profile of John Doe, sample company”, as well as some basic information about P-WIN and the “fact” that there are eight basic dimensions or types of personality, respectively. On the second page, there was a description of the eight dimensions or types (which were almost the same as in Experiment 1a – we merely replaced structuredness with composure and spirit of discovery with performance orientation) as well as a pictorial representation of the resulting personality profile. This differed markedly in the two versions: In the dimension-based report version, there was a point on a scale representing the manifestation on each of the eight dimensions (see Figure 9); in the type-based version, there were only two of the eight types marked as primary and secondary preference (see Figure 10). Therefore, there was a stark difference between the two results reports. In the dimension-based version, John Doe was self-evaluated on all eight dimensions; in the type-based version, there were preferences for only two types (see also Appendix B for complete material).

The P-WIN® profile of John Doe

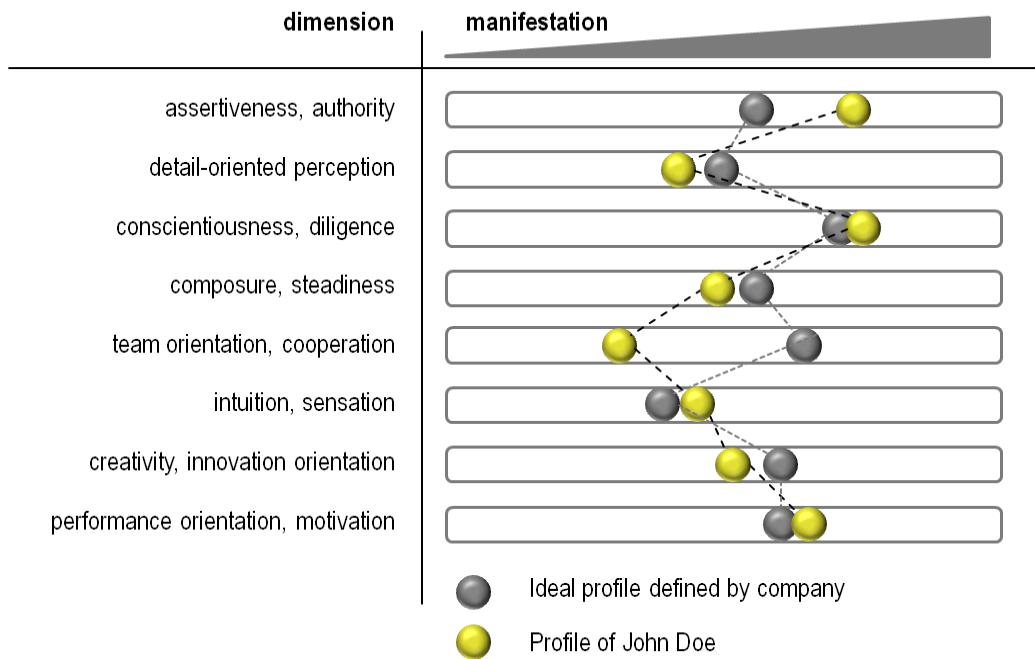
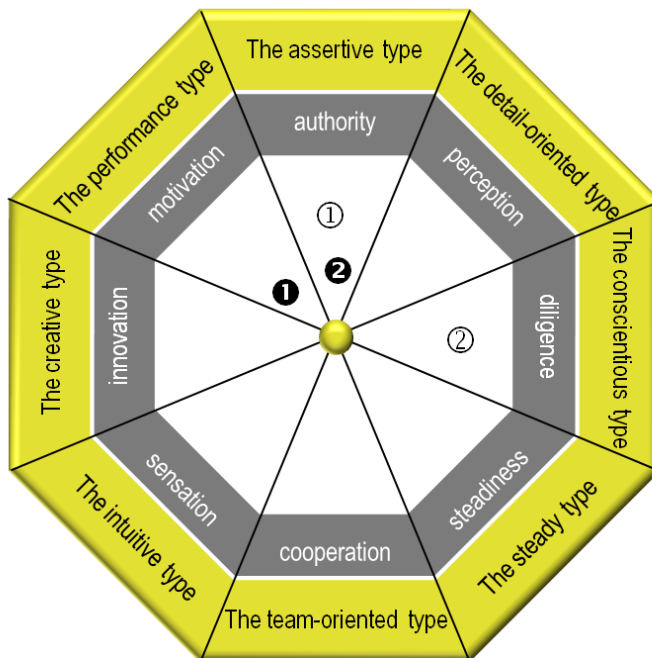


Figure 9. Representation of traits in the results report of the dimension-based condition in chapter 4.

The P-WIN® profile of John Doe



① Primary preference: This type conforms most with the personality; it symbolizes the natural, inherent and learned behavior patterns.

② Secondary preference: This type characterizes the behavior in stressful situations. There may be a difference to the primary preference (but it does not have to be).

① / ② Ideal preferences defined by company.

Figure 10. Representation of traits in the results report of the type-based condition in chapter 4.

Dependent variables. The evaluation of the results reports began immediately after reading the two pages of report. We shortened the number of items to five, concerning the suitability of the instrument for personnel selection (item 1), contribution to the objectivity of the selection decision (item 2), attractiveness of the results report (item 3), interest in more information about P-WIN (item 4), and the overall liking of the personality test (item 5). Again, we used a 7-point rating scale (1 = *strongly disagree* to 7 = *strongly agree*) and added the same open-format question as in the first experiment. The five items showed an acceptable reliability (Cronbach's $\alpha = .79$) and were combined to form a scale of overall test rating.

Results and Discussion

Again, we used a t-test for independent samples to test our hypothesis that the type-based report is more attractive than the dimension-based personality report. However, we found no significant effect, $t(158) = 0.92$, between the type ($M = 4.57$, $SD = 1.07$) and dimension ($M = 4.42$, $SD = 0.99$) condition. Only 15.6% of the participants used the chance to explain their ratings in the open-format question, and just 3.1% mentioned the word "type" and 0.6% the word "dimension". When including the words "category" and "aspect" for type, and "trait", "component", "aspect" or "points" for dimension, 10.0% commented on the type aspect and 3.8% on the dimension aspect. Despite the more obvious difference between the two conditions, our hypothesis was again not confirmed. [Furthermore, taking the different modalities (paper-and-pencil vs. online) into account did not change these results, with the corresponding ANCOVA showing an $F(157) = 1.17$.]

Therefore, the results of Experiment 1a were replicated: In both experiments, participants did not show a preference for a type-based test over a dimension-based test.

Experiment 2 (Theory)

In our second experiment we wanted to test our hypothesis that a statistics-based personality test will be evaluated more positively than its theory-driven counterpart. We decided to use the flyer manipulation again, as it did work in the study by Diekmann et al.

(2015) and we could not transfer the results report design to this hypothesis. Furthermore, we decided to work with three groups of participants. As in Experiment 1, we asked HR experts and business management students, as most of the HR experts had studied business management and we wanted to see whether the difference in experience accounted for differences in evaluations. As a third group, we included psychology student because they should know personality tests from their studies, and we wished to analyze possible differences between the different training backgrounds.

Method

Participants

HR experts. The recruitment procedure was the same as in Experiment 1a. A total of 89 German HR experts (48.3% female, 41.6% male, 10.1% did not indicate their gender) participated in the second experiment. They had been working as HR employees for an average of 11.4 years ($SD = 9.4$). Overall, 82.0% were personally involved in the selection process, and 61.8% had decision-making power concerning the choice of selection methods. The majority (71.9%) had completed an academic education, mostly having studied business management ($n = 27$) or psychology ($n = 18$). About one third (34.8%) had completed vocational training, mostly as industrial clerk ($n = 14$) or as banker ($n = 6$). In 27.0 % of the companies, personality tests were used in selection procedures. Participants were randomly assigned to the theory ($n = 44$) or the statistics ($n = 45$) condition.

Business management students. Business management students were recruited during a business management lecture at a local university in Germany. They were given a paper-and-pencil version of the material and answered the questions during the lecture (without any compensation). Of the initial sample, only those studying business management or similar were included.

Overall, 97 students (37.1% male, 62.9% female) participated. Most of them focused on personnel management (18.6%) and marketing (12.4%) and were aiming for a Master degree

(71.5%). Almost half (42.3%) stated that they had experience in personnel selection; 12.5% of these had experience with personality tests. Participants were verbally welcomed and told about informed consent conditions. On the first page of the material, they were asked to imagine being an employee in human resources who wants to supplement the selection process with a personality test. On the next page, they were randomly assigned to the theory ($n = 44$) or the statistics ($n = 53$) condition.

Psychology students. Psychology students were recruited via e-mail distribution lists with psychology students of different German universities. They were offered the chance to win one of five 10€ Amazon gift cards. Of the initial sample, only those studying psychology were included in the analysis.

A total of 93 psychology students (80.6% female, 18.3% male, 1.1% did not indicate their gender) participated. They were mostly heading for a Bachelor (67.7%) or Master degree (20.4%) with a focus on clinical psychology (33.3%) or work and organizational psychology (21.5%). Only 23.7% had experience in personnel selection, and of these, 22.7% had experience with personality tests. The procedure was the same as in Experiment 1. Participants were again randomly assigned to the theory ($n = 48$) or the statistics ($n = 45$) condition.

Overall sample. Altogether, 279 individuals participated in Experiment 2, with 48.7% in the theory and 51.3% in the statistics condition.

Material

Building on Diekmann et al. (2015), we used the flyer material from Experiment 1. The structure was the same, with the benefits of generally using a personality test explained on the left-hand side and the advantages of using this personality test in particular outlined on the right-hand side. The middle part informed the reader about the construction background of P-WIN, which differed in the theory and the statistics condition. Both cases involved a fictitious American professor named Smith. In the statistics condition, the reader was informed that this professor discovered the existence of eight personality traits using statistical methods. The use

of statistics was emphasized three times. In the theory condition, it was stated that the same professor developed a personality theory instead. Again, the word “theory” was used three times. The personality traits were the same in both conditions and printed below this information (see also Appendix B for complete material).

Dependent variables

We used the same six items as in Diekmann et al. (2015), with the 7-point rating scale (1 = *strongly disagree* to 7 = *strongly agree*) plus the open-format item. The items referred to the suitability of the instrument for personnel selection (item 1), contribution to the objectivity of the selection contexts (item 2), attractiveness of content (item 3) and design (item 4) of the flyer, interest in additional information about P-WIN (item 5) and overall test liking (item 6). As the six items showed a good reliability (Cronbach’s $\alpha = .87$; also in the three subsamples: HR experts $\alpha = .86$, business management students $\alpha = .83$; psychology students $\alpha = .88$), they were combined to form a scale of overall test rating.

Results and Discussion

To analyze whether people would prefer a test developed by statistical methods or a test designed on the basis of a psychological theory, we used a 2 (statistics vs. theory) \times 3 (HR experts vs. business management students vs. psychology students) analysis of variance. Levene’s test for equality of variances indicated that the variability of each group was approximately equal, $F(5,273) = 1.10$. Means and standard deviations of all groups can be found in Table 2. Contrary to our expectations, an ANOVA showed no significant main effect of test version, $F(1,273) = 0.58$. Thus, there was no support for our hypothesis that the perception of tests with a statistical background differed from that of tests with a theoretical background. There was, however, a significant main effect of subsample, $F(2,273) = 14.85$, $p < .01$, indicating a difference of ratings between subsamples. Descriptive results indicated that psychology students ($M = 4.46$, $SD = 1.18$) generally rated our test better than business management students ($M = 3.86$, $SD = 0.98$) or HR practitioners ($M = 3.60$, $SD = 1.10$). A post

hoc Scheffé test revealed significant differences between psychology students and business management students, $p < .01$, and between psychology students and HR practitioners, $p < .01$. No interaction between subsample and test version could be found, $F(2,273) = 0.17$.

Table 2.
Overall test rating, standard deviations, and effect sizes for sample and subsamples in chapter 4.

Sample	Test version	<i>n</i>	<i>M</i>	<i>SD</i>	<i>d</i>
Overall	Statistics	142	3.92	1.06	0.10
	Theory	137	4.03	1.22	
HR experts	Statistics	44	3.58	1.06	0.04
	Theory	45	3.62	1.15	
Business management students	Statistics	53	3.83	0.95	0.05
	Theory	44	3.88	1.02	
Psychology students	Statistics	45	4.35	1.06	0.18
	Theory	48	4.56	1.29	

Overall Discussion

In the present studies, we analyzed two possible decision criteria of practitioners when choosing a personality test for specific use in employee selection. Contrary to our first hypothesis, there was no significant difference between a type-based and a dimension-based personality test in our experimental design, although Diekmann and König (2015) found a preference for types in their survey. There may be different explanations for this: First, although we performed our experimental design with a second test material stressing the difference between types and dimensions (the results report in contrast to the flyer), the manipulation may still not have been sufficiently strong. Second, HR managers in the survey by Diekmann and König (2015) likely had particular tests in mind when thinking about dimension- or type-based tests. Possibly, some of them worked with (or at least knew) some particular tests in personnel selection or personnel development which were type-based, therefore triggering a preference

for types. Our experiment was not based on a specific personality test (either type-based or dimension-based) and therefore eliminated this bias.

Likewise, we did not find a significant difference between the evaluations of theory-based and statistics-based personality tests, although Diekmann and König (2015) found a (self-stated) preference for the statistics-based development of personality tests among practitioners. Again, the survey by Diekmann and König (2015) may have entailed a bias based on knowledge: The HR experts surveyed may have heard that a statistics-based development is more scientific and answered accordingly due to social desirability in this study conducted by university researchers.

Although there were no significant differences between conditions, we should not underestimate the implications of our results – finding no effects also matters for this field. Psychological science pays great attention to dimension-based personality tests. Research about types is limited, and the most widespread type-based model focuses on three types of children (resilient, under- and overcontroller; e.g. Van den Akker, Deković, Asscher, Shiner, & Prinzie, 2013). Researchers should ask themselves whether they are meeting practitioners' needs when they choose to ignore type-based personality models in organizational contexts. If a considerable number of practitioners like (or at least do not dislike) a type-based approach to personality, ignoring types might create another research-practice gap. Instead, researchers should try to find out what causes this interest and what practitioners really want when using a personality test. Researchers could either adjust their focus and try to develop more scientifically sound type models or increase their efforts to give good advice concerning the benefit of dimension-based personality tests.

Similarly, researchers should reconsider a purely statistical approach to the construction of personality tests. More effort should be undertaken to develop a theoretical grounding of our personality models and/or it should be argued more explicitly what the importance of empirically founded personality traits is. Furthermore, Diekmann et al. (2015) recommend

more tailored solutions to the research-practice gap and its specific aspects – in this case the use of personality tests in personnel selection. Our results suggest the same: Researchers should be more open to the needs of practitioners, and conscientiously develop new and agreeable solutions to establish an emotionally stable relationship with our practitioner partners.

Thus, these studies show that there are also merits in null findings. In our case, there were no significant differences between conditions, but this also means a lack of effect in a direction science may desire (e.g., a preference for dimensions over types). Although psychology seems to have a bias towards publishing significant results (e.g., Fanelli, 2010a, 2010b, 2011), this study shows the advantages of the recent trend to overcome this bias (e.g., Landis, James, Lance, Pierce, & Rogelberg, 2014; van Dick, 2015; see also Kundoor & Mueen, 2010).

As in every study, there are limitations that need to be addressed. The experiments in this paper were simulations using a fictitious personality test in a laboratory framework. The fact that we used a fictitious test is based on practical reasons: We could not compare a real dimension-based personality test to a real type-based personality test because to our knowledge, there is no such test existing in both versions. However, since Diekmann et al. (2015) already used this design and found effects, we could be fairly certain that the design of the flyer would work. The laboratory design was used to enable us to experimentally test our hypotheses without confounding variables affecting our results. In order to draw causal conclusions, we did not simulate a real-world decision, but rather simplified the design and reduced the variables. Although this might constitute a limitation to the external validity of our results, it is crucial to keep in mind that laboratory and field effects correlate rather highly (G. Mitchell, 2012).

Our experiments are hopefully just the beginning of a new line of research (see also Diekmann & König, 2015): To better understand which criteria have an impact on practitioners' decisions for or against a personality test, further research is clearly needed. This should be

accomplished with an application-oriented objective – to deliver better solutions to HR practitioners that meet scientific standards as well as practical requirements.

KAPITEL 5

IMPLIZITE LEISTUNGSTHEORIEN IN DER PERSONALAUSWAHL

Zusammenfassung

Implizite Theorien darüber, warum Menschen erfolgreich sind und welche Eigenschaften und Kompetenzen sie dabei mitbringen müssen, spielen wahrscheinlich auch bei der Personalauswahl und somit bei der Auswahl von Persönlichkeitstests zu diesem Zweck eine Rolle. Die vorliegende Studie untersucht, welche Eigenschaften von Berufstätigen und Studenten als wichtig für beruflichen Erfolg erachtet werden. Die qualitative Datenanalyse ergibt, dass insgesamt 29 Kategorien von Eigenschaften mit unterschiedlicher Häufigkeit genannt werden. Diese große Bandbreite steht damit möglicherweise im Widerspruch zu dem Versuch von Persönlichkeitsmodellen und auch von Testverfahren, die Vielfalt von Eigenschaften und Kompetenzen auf möglichst wenige Dimensionen zu reduzieren. Vergleiche mit zwei bekannten Testverfahren zeigen sowohl die Unterschiede in der Anzahl von Kategorien als auch in deren Bezeichnungen, die für die Augenscheinvalidität und somit für die Auswahl von Testverfahren ebenfalls entscheidend sein können.

Einleitung

Menschen haben ein starkes Bedürfnis, die komplexe Welt, in der sie leben, zu verstehen und zu erklären (siehe z. B. Kelemen, 1999): Sie wollen verstehen, wie und warum die Dinge funktionieren, wie unterschiedliche Systeme miteinander verbunden sind und warum Menschen handeln, wie sie es tun. Wir können nicht Expertinnen und Experten in allen Aspekten unseres Lebens sein oder jede wissenschaftlich fundierte Theorie verstehen. Dennoch haben wir eine Vorstellung, eine so genannte implizite Theorie, über viele Bereiche des täglichen Lebens, egal ob es sich um die richtige Ernährung, eine gesunde Lebensführung, die Erziehung unserer Kinder oder das Verhalten von Menschen und ihre Beziehungen zu einander handelt. Obwohl wir nicht gleichzeitig Ernährungswissenschaftler, Ärzte, Erzieher oder Psychologen sind, haben wir für uns ein Verständnis von diesen Dingen (das nicht immer objektiv richtig sein muss). Dieses Wissen entwickelt sich aus und wächst mit der Erfahrung (C. A. Anderson & Lindsay, 1998), z. B. haben Eltern ein besseres Verständnis von der Kindesentwicklung als Menschen ohne Kinder (Furnham & Weir, 1996). Das Alltagswissen über soziale und psychologische Zusammenhänge sollte besonders ausgeprägt sein, da Menschen jeden Tag miteinander interagieren und ihre individuellen Aktionen und Reaktionen beobachten können (Heider, 1958).

Da implizite Theorien verbreitet sind, sollten sie ebenfalls in dem speziellen Feld der Personalauswahl existieren. Praktiker, HR Manager oder Führungskräfte, die in Personalauswahlprozesse involviert sind, haben wahrscheinlich eigene Vorstellungen davon, wie man den „richtigen“ Mitarbeiter auswählt und wie der „richtige“ Mitarbeiter sein muss, insbesondere wenn sie keine spezielle Ausbildung im Feld Personalauswahl haben. Ihre implizite Leistungstheorie – ihre Theorie darüber, welche Eigenschaften und Kompetenzen einen erfolgreichen Mitarbeiter ausmachen – sollten nicht nur die Auswahl selbst beeinflussen, sondern auch den gesamten Personalauswahlprozess, z. B. welcher Persönlichkeitstest genutzt wird oder welche Kompetenzen im Assessment Center beobachtet werden sollen. Eine

genauere Untersuchung der impliziten Leistungstheorie sollte daher einen Beitrag zur so genannten Research-Practice Gap in der Personalauswahl leisten können, welche die hohe Popularität von unstrukturierten und von der Wissenschaft nicht empfohlenen Verfahren betrifft (siehe z. B. N. Anderson et al., 2004).

Das Hauptanliegen dieser Studie war es herauszufinden, welche Eigenschaften aus Laiensicht ein erfolgreicher Mitarbeiter im Kontext der Personalauswahl haben sollte und warum Führungskräfte glauben, dass diese Eigenschaften zu persönlichem Erfolg führen. Ein besseres Verständnis der impliziten Leistungstheorie kann helfen, unsere Empfehlungen an Praktiker in Bezug auf die Methoden der Personalauswahl und Beobachtertrainings zu verbessern und unsere Forschung nicht nur auf etablierte Persönlichkeitsmodelle und Kompetenzen zu fokussieren, sondern diese Modelle so zu überdenken, dass sie zu den Bedürfnissen der Praktiker passen.

Theoretischer Hintergrund

Menschen haben ein starkes Bedürfnis, ihre physische, biologische und besonders ihre soziale Umwelt zu verstehen und zu erklären (Heider, 1958; Kelemen, 1999; Kelley, 1973). Ein grundlegendes Verständnis der Welt und ihrer Funktionsweise hilft uns, zukünftige Ereignisse und das Verhalten anderer Menschen vorherzusehen und sagt uns, wie wir diese Dinge gegebenenfalls beeinflussen können (C. A. Anderson & Lindsay, 1998). Diese Tendenz und unser Bedürfnis, unsere Umwelt zu kontrollieren, führt dazu, dass Menschen implizite Theorien über mögliche Zusammenhänge entwickeln (Heider, 1958; Hong, Levy, & Chiu, 2001; Kelemen, 1999), um ein allgemeines Gefühl der Kontrolle über ihr Leben zu bekommen. Vor diesem Hintergrund wurden Menschen als (naive) Wissenschaftler (Kelley, 1972; Kelly, 1955) oder – im Kontext psychologischer Phänomene – als naive Psychologen bezeichnet (Heider, 1958).

Implizite Theorien wurden z. B. auch als naive, intuitive oder Laien-Theorien bezeichnet (Hong et al., 2001; Wegener & Petty, 1998). Im Allgemeinen werden implizite

Theorien als organisierte, Schema-ähnliche Wissensstrukturen oder Überzeugungen über bestimmte Themen verstanden (Guzzo, Wagner, Maguire, Herr, & Hawley, 1986; Hong et al., 2001; Ross, 1989). Es geht also um die Einteilung von Informationen in bedeutungsvolle Kategorien. Darüber hinaus haben implizite Theorien eine kausale Komponente: Die Informationen innerhalb und zwischen den Themen sind durch kausale Annahmen miteinander verbunden, mithilfe derer es möglich ist, Vergangenes zu erklären und zukünftige Ereignisse vorherzusehen (C. A. Anderson & Lindsay, 1998). Implizite Theorien entwickeln sich aufgrund direkter Erfahrung (d. h. der selbst gemachten Erfahrungen), durch indirekte Erfahrung (kommuniziert durch andere Menschen oder auch durch Medien) oder durch den aktuellen Bedarf, ein bestimmtes Ereignis zu verstehen oder vorherzusehen (C. A. Anderson & Lindsay, 1998). Die Theorien entwickeln sich also mit wachsender Erfahrung weiter, was ein Grund für die unterschiedliche Qualität und Quantität dieser Theorien zwischen verschiedenen Menschen sein könnte (Furnham, 1988). Im Personalkontext sollten sowohl HR Praktiker, aber insbesondere auch Führungskräfte recht ausgeprägte implizite Theorien darüber haben, was Menschen erfolgreich macht.

Verschiedene Studien haben den Inhalt von Laintheorien untersucht, indem sie Menschen dazu befragt haben. So gibt es beispielsweise Laintheorien zur Intelligenz (siehe z. B. Sternberg, 1985), darüber, warum Menschen im sich im Arbeitskontext (kritisch) äußern (implicit voice theories; Detert & Edmondson, 2011), über die Ursachen, Zusammenhänge und Vermeidung von Arbeitsstress (siehe z. B. Furnham, 1997), über die Ursachen, Erklärungen und Reduktion von Arbeitslosigkeit (siehe z. B. Furnham, 1982; Lewis & Furnham, 1986) oder über die Eigenschaften von Führungskräften (siehe z. B. Schyns & Schilling, 2011). Andere Studien haben die impliziten Theorien von Menschen über experimentelle Designs untersucht. So hat die Forschung beispielsweise gezeigt, dass Menschen Urteile über soziale Beziehungen in Gruppen fällen können, wenn sie Beschreibungen über die grundlegenden Eigenschaften dieser Gruppe vorliegen haben. Dies weist auf eine implizite Theorie über Typen von Gruppen

hin (Lickel, Hamilton, & Sherman, 2001). Eine besonders gut untersuchte implizite Theorie im Feld der Arbeits- und Organisationspsychologie betrifft die implizite Führungstheorie. Die dazu gehörige Forschung startete mit einer Studie von Eden und Leviatan (1975): Sie fanden heraus, dass es möglich ist, die Faktorstruktur einer Führungsskala zu reproduzieren, wenn die urteilenden Personen lediglich eine Beschreibung der Organisation und keine Informationen über die Führungskultur vorliegen hatten. Die Autoren folgerten daraus, dass eine implizite Führungstheorie, in der Führungsverhalten mit bestimmten Verhaltensweisen verbunden wird, für die Faktorstruktur verantwortlich ist. Dieses Ergebnis konnte vielfach reproduziert (siehe z. B. Bryman, 1987, 1987; Lord, Binning, Rush, & Thomas, 1978; Phillips & Lord, 1986; Rush, Phillips, & Lord, 1981; Rush, Thomas, & Lord, 1977) und durch weitere Forschung ergänzt werden – z. B. besteht die implizite Überzeugung, dass gute Leistung durch gute Führung beeinflusst wird (siehe z. B. Bryman, 1987; Larson, 1982; Lord et al., 1978; T. R. Mitchell, Larson Jr., & Green, 1977; Nye, 2005)

Implizite Theorien spielen auch in der Forschung zum so genannten Leader-Member-Exchange (LMX) eine wichtige Rolle. LMX beschreibt die dyadische Beziehung zwischen Führungskräften und ihren Mitarbeitern. So untersuchten Engle und Lord (1997) die impliziten Leistungstheorien von Führungskräften und Mitarbeitern in ihrer LMX-Forschung. Sie ließen die Prototypikalität von Eigenschaften und Verhaltensweisen beurteilen, die Leistungsträger auf der Basis von Wernimont (1971) beschreiben. Die Ergebnisse in Bezug auf die Eigenschaften guter Mitarbeiter beinhalteten Eigenschaften wie Gewissenhaftigkeit, Ehrlichkeit, Sorgfalt, Selbstmotivation, effektive Kommunikation oder Zuverlässigkeit. Die Untersuchung einer allgemeineren impliziten Mitarbeitertheorie (ohne den Leistungsfaktor) fand Eigenschaften wie Fleiß, Inkompetenz, Konformität, ein Team-Player sein, Enthusiasmus, Ungehorsam, Flexibilität, kommunikative Fähigkeiten, Loyalität oder Unternehmenszugehörigkeit (citizenship) als Charakteristika von Mitarbeitern (Carsten, Uhl-Bien, West, Patera, & McGregor, 2010; Sy, 2010).

Implizite Führungstheorien beziehen sich auf die “Bottom Up” Perspektive, die Wahrnehmung der Führungskräfte durch ihre Mitarbeiter. Implizite Leistungstheorien fokussieren stattdessen auf die andere, „Top Down“ Perspektive von Führungskräften, die ihre eigenen Vorstellungen von notwendigen Eigenschaften, Kompetenzen und Verhaltensweisen haben, die für eine Arbeit notwendig sind (Swee, 2009). So wie implizite Führungstheorien die Bewertungen von Führungskräften beeinflussen, werden wahrscheinlich auch Führungskräfte auf implizite Leistungstheorien zurückgreifen, wenn sie – zukünftige oder bestehende – Mitarbeiter beurteilen (Engle & Lord, 1997). Obwohl es Forschung zu Mitarbeiter-Charakteristika im Allgemeinen (Carsten et al., 2010; Sy, 2010) und umfangreiche Literatur zum Thema Leistung gibt (Sonntag & Frese, 2002), haben sich kaum Wissenschaftler mit der impliziten Leistungstheorie und ihrem Inhalt beschäftigt (mit Ausnahme von Engle & Lord, 1997; und Wernimont, 1971).

Dies überrascht, wenn man bedenkt, dass diese implizite Theorie die Bewertung von Mitarbeitern (z. B. im jährlichen Beurteilungsgespräch) sowie den Personalauswahlprozess beeinflussen könnte. Darüber hinaus bietet es einen neuen Ansatz für die so genannte Research-Practice Gap im Bereich der Personalauswahl. Diese Gap („Kluft“) bezeichnet die Abweichung zwischen dem, was in der wissenschaftlichen Forschung untersucht und empfohlen wird und dem, was Praktiker tatsächlich anwenden (siehe z. B. N. Anderson et al., 2001). Sie existiert in vielen Forschungsbereichen und ebenfalls im Bereich der Personalauswahl (siehe z. B. N. Anderson et al., 2004) und zeigt sich beispielsweise in der hohen Popularität von unstrukturierten Interviews durch Praktiker (siehe z. B. Macan, 2009; Stephan & Westhoff, 2002), während die Forschung ein strukturiertes Vorgehen empfiehlt (siehe z. B. McDaniel et al., 1994).

Praktiker scheinen an Persönlichkeitseigenschaften und sozialen Fähigkeiten für den Zweck der Personalauswahl interessiert zu sein und sie in Interviews auch messen zu wollen (Huffcutt, Conway, Roth, & Stone, 2001), obwohl Tests ein strukturiertes und damit

objektiveres und psychometrisch überlegenes Vorgehen bieten. Darüber hinaus gibt es eine unüberschaubar Anzahl von Persönlichkeitstests auf dem Markt (Oakland, 2004), und viele davon werden im organisationalen Kontext verwendet (Berchtold, 2005; Brown, 1999; Di Milia, 2004; Diekmann & König, 2015; Furnham, 2008; Muñiz & Fernández-Hermida, 2010; Ryan & Sackett, 1987, 1992), wobei – wie schon in Kapitel 2 dargestellt – nicht alle davon für einen Einsatz in der Personalauswahl geeignet zu sein scheinen. Darüber hinaus scheint sich die Forschung meist ausschließlich auf das Fünffaktorenmodell (siehe z. B. Barrick et al., 2001; Ones et al., 2007; Tett & Christiansen, 2007) zu konzentrieren. Es gibt wenig Testverfahren, die auch soziale Fähigkeiten erfassen (eine Ausnahme stellt z. B. das Inventar sozialer Kompetenzen dar, siehe Kanning, 2009), und sogar die Forschung über angewandte soziale Fähigkeiten ist selten (Huffcutt et al., 2001). Es fehlt an wissenschaftlichen Leitlinien, die Orientierung bieten, welche sozialen Fähigkeiten eine gute prädiktive Validität haben. So ist es kein Wunder, dass Praktiker eher Methoden wie das Interview nutzen und sich möglicherweise bei den Inhalten oder den Bewertungen im Assessment Center an ihren impliziten Theorien orientieren (siehe z. B. Uggerslev & Sulsky, 2008).

Um besser auf die Bedarfe von Praktikern eingehen zu können, müssen wir (in Ergänzung zu der in den Kapitel 3 bis 5 dargestellten Hypothesen und Studien) wissen, welche Eigenschaften für diese wichtig und bedeutsam sind um dann entweder bestehende Verfahren weiterzuentwickeln, neue Inventare zu entwickeln (auch vor dem Hintergrund der Augenscheinvalidität, deren Bedeutung immer mehr erkannt wird, siehe z. B. Connell et al., 2018; Hardesty & Bearden, 2004) oder zu verdeutlichen, dass diese Eigenschaften wenig prädiktive Validität aufweisen und die Nutzung nicht sinnvoll ist. Dabei sollten wir nicht nur Eigenschaften im engeren Sinne, sondern auch Kompetenzen betrachten. Es ist davon auszugehen, dass die wissenschaftliche Trennung für Praktiker weniger relevant und verständlich ist.

Vor diesem Hintergrund wurde die folgende Fragestellung aufgestellt: *Welche Konstrukte – welche Eigenschaften und Kompetenzen – werden aus Laiensicht als relevant für den beruflichen Erfolg eines Menschen eingestuft?*

Methode

Design und Vorgehen

Ziel dieser Studie ist es herauszufinden, welche Eigenschaften aus Laiensicht einen Einfluss auf den beruflichen Erfolg haben. Dabei wird „Einfluss“ im ersten Schritt neutral verstanden, d. h. es geht um Eigenschaften und Kompetenzen, die sowohl positiven als auch negativen Einfluss haben können. Design und Vorgehen wurden an die Arbeiten zur impliziten Führungstheorie von Schyns und Schilling (2011) sowie Offerman (1994) angelehnt, da es große Parallelen in Bezug auf das Forschungsziel gibt, auch wenn der Inhalt ein anderer ist. Dabei wurden zwei unterschiedliche Stichproben befragt: Die erste Stichprobe besteht aus Studenten, die bisher wenig Erfahrung mit dieser Fragestellung haben dürfte und somit auf eine eher grundlegend ausgeprägte implizite Theorie zurückgreifen muss. Die zweite Stichprobe besteht aus Personen mit Berufs- und Führungserfahrung, deren implizite Theorie aufgrund der höheren Erfahrung ausgeprägter und ggf. durch explizite Theorien angereichert sein dürfte. Dieses Vorgehen erlaubt es, diese unterschiedlichen Vorstellungen miteinander zu vergleichen.

Beide Stichproben wurden (in Anlehnung an Schyns & Schilling, 2011) gebeten, sich eine Person vorzustellen, die sich für einen neuen Job bewirbt. Beispielsweise jemanden, den sie aus einem beruflichen oder privaten Kontext kennen. Sie wurden angehalten, sich vorzustellen, welche Eigenschaften kritisch für den zukünftigen beruflichen Erfolg sein würden, wobei sowohl positive als auch negative Eigenschaften den Erfolg beeinflussen könnten. Sie konnten bis zu sechs Eigenschaften nennen. Auf diese Weise waren die Teilnehmer gezwungen, sich auf die für sie wichtigsten Eigenschaften zu konzentrieren, ohne eine zu starke Einschränkung vornehmen zu müssen. Jede Nennung konnte durch Verhaltensbeispiele erklärt bzw. definiert werden. Dieses Vorgehen hat zwei Vorteile: Zum

einen gibt es bei der Auswertung ein besseres Verständnis, was mit einer bestimmten Eigenschaft gemeint ist; zum anderen wurden die Teilnehmer dadurch angeregt, Eigenschaften bzw. Kompetenzen von Verhalten zu trennen.

Teilnehmer

Die Umfrage wurde mithilfe des Anbieters Unipark über eine Online-Befragung durchgeführt. Die Befragung der beiden Stichproben fand über zwei getrennte, aber inhaltlich identische Fragebögen statt. Die Studentenchprobe wurde durch soziale Netzwerke, E-Mail-Verteiler und über die Weiterleitung des Online-Links an verschiedene Fachschaften im deutschsprachigen Raum akquiriert. Bei der Auswahl der studentischen Teilnehmer wurde darauf geachtet, dass sie sich später in der freien Wirtschaft bewerben müssen, d. h. es wurden gezielt keine Lehramtsstudenten oder Mediziner angesprochen. Die Stichprobe der Berufstätigen wurde über das soziale Netzwerk „Xing“ durch das Suchwort „Führungserfahrung“ ausgewählt und über dieses Netzwerk auch angesprochen.

Insgesamt nahmen an der Studie 417 Personen teil, welche die Befragung starteten. Beendet wurde der Onlinefragebogen von insgesamt 111 Personen (26.6%), 85 davon (20.4%) gaben mindestens eine Eigenschaft an. Diese 85 Personen werden im Folgenden als endgültige Stichprobe betrachtet. In der Studentenchprobe gaben 43 Personen (58.0% Frauen, 29.5%, Männer, eine fehlende Angabe) mindestens eine Eigenschaft an. In der Berufstätigenstichprobe haben 42 Personen (28.6% Frauen, 61.9% Männer, drei fehlende Angaben) aktiv teilgenommen. Das Durchschnittsalter der Studentenchprobe betrug 24 Jahre ($SD = 4.2$), das der Berufstätigen 46 Jahre ($SD = 8.9$). Die Berufstätigen hatten eine mittlere Arbeitserfahrung von 21 Jahren ($SD = 9.6$) und eine durchschnittliche Führungserfahrung von 13 Jahren ($SD = 8.0$). Sie stammten aus verschiedensten Beschäftigungssparten, so reichte die Stichprobe von Regionalleitern der Pharmaindustrie bis hin zu Leitern der Aus- und Weiterbildung. Elf Berufstätige studierten vor ihrer Karriere Wirtschaftsstudiengänge, womit sie mit 26.2% den

größten Anteil ausmachten. Auch bei den Studenten machte der Anteil an Wirtschaftsstudenten mit 23.3% den größten Anteil aus.

Datenanalyse

Die vorhandenen Daten wurden in Anlehnung an Mayring (2010) qualitativ wie folgt ausgewertet:

1. **Material:** Aus den Fragebögen wurden die Eigenschaftsbegriffe und ggf. die Spezifikation durch Verhaltensbeispiele als Grundlage der Datenanalyse entnommen.
2. **Ziel der Analyse:** Das Ziel der qualitativen Analyse war es, die Eigenschaften zu sinnvollen Einheiten zusammenzufassen um eine Komplexitätsreduktion zu erreichen.
3. **Selektionskriterium:** Es wurden alle Nennungen entfernt, die nicht zur oben genannten Definition des Eigenschaftsbegriffes passten, also z. B. Nennungen wie Fachwissen oder fachliche Kompetenz und Intelligenz, Hinweise auf den Notendurchschnitt und das Aussehen oder reine Verhaltensbeschreibungen. Eigenschaften, Kompetenzen und Fähigkeiten in positiver oder negativer Ausprägung wurden in die Analyse aufgenommen.
4. **Synonym- und Ähnlichkeits-Kategorisierung:** Für die Kategorisierung wurden vor der ersten Kategorisierung und ergänzend in der folgenden Diskussion (siehe Punkt 5) folgende Vereinbarungen getroffen:
 - Die Eigenschaften werden nach semantischer Ähnlichkeit zusammengefasst.
 - Mit Hilfe der Verhaltensbeispiele wird darauf geachtet, dass gleich benannte Eigenschaften auch dasselbe bedeuten.
 - Sollte das Verhaltensbeispiel nicht mit dem Begriff in Einklang zu bringen sein, gilt der Eigenschaftsbegriff.
 - Werden zwei widersprüchliche Verhaltensweisen angegeben, gilt die erste.
 - Negative Eigenschaften werden mit ihrem positiven Gegenpart in derselben Kategorie zusammengefasst.

5. Materialbearbeitung: Bei der Kategorisierung wurden zunächst die beiden Stichproben durch zwei wissenschaftliche Hilfskräfte getrennt voneinander kategorisiert. Anschließend wurden in einer Diskussionsrunde gemeinsam mit der Erstautorin die entstandenen Kategorien und Ergebnisse verglichen, kritische Fälle diskutiert und ggf. Kategorien noch einmal aufgeteilt oder zusammengefasst, so dass schließlich einheitliche Kategorien für beide Stichproben entstanden.
6. Revision der Daten: Nach der Einigung wurde zur Überprüfung die Kategorisierung erneut durch eine Forschungspraktikantin vorgenommen und ein als gut zu bewertender Übereinstimmungskoeffizient (Cohens Kappa) von 0.698 berechnet (für Einschätzungen hinsichtlich der Höhe eines angemessenen Kappa-Wertes siehe z. B. Wirtz & Caspar, 2002). Dennoch wurden durch die Erstautorin auf Basis der Ergebnisse noch einmal kleinere Anpassungen bei den Kategorien vorgenommen und anschließend eine weitere Überprüfung durch eine weitere wissenschaftliche Hilfskraft vorgenommen. Der berechnete Übereinstimmungskoeffizient (Cohens Kappa) auf der Grundlage dieser finalen Kategorisierung stieg auf einen sehr guten Wert von 0.792.
7. Interpretation der Daten: Hier wurde schließlich im Ergebnis betrachtet, welche Kategorien entstanden sind und welche Unterschiede es zwischen Berufstätigen und Studenten gibt.

Ergebnisse

Insgesamt wurden 29 Kategorien bestehend aus insgesamt 351 (188 aus der Studentenchprobe, 163 aus der Berufstätigenstichprobe) genannten Eigenschaften und Kompetenzen gefunden. Im Mittel bestanden die Kategorien aus 12.10 Eigenschaften ($SD = 10.25$), bei den Studenten aus 6.48 Eigenschaften ($SD = 6.54$) und bei den Berufstätigen aus 5.62 Eigenschaften ($SD = 4.44$). In Tabelle 3 sind die Kategorien, die jeweilige Definition basierend auf den zugeordneten Eigenschaften und Kompetenzen, sowie die Häufigkeit der Nennung einer Eigenschaft innerhalb dieser Kategorie dargestellt. Den größten Anteil machen die Kategorien „soziale Kompetenz“ (39 Nennungen), „Zuverlässigkeit“ (37 Nennungen),

„Teamfähigkeit“ (32 Nennungen), „Selbstmotivation“ (27 Nennungen) und „Leistungsbereitschaft“ (25 Nennungen) aus. Diese Kategorien liegen auch in der Studentenchprobe ganz vorne in der Rangfolge. In der Berufstätigenstichprobe fällt dagegen statt „Selbstmotivation“ die Kategorie „Kommunikationsfähigkeit“ (12 Nennungen) unter die Top Fünf. Trotz einiger kleinerer Unterschiede in den genannten Häufigkeiten, die zumindest teilweise darauf zurückzuführen sind, dass die Berufstätigen insgesamt weniger Eigenschaften und Kompetenzen benannt haben, ist das Bild zwischen den beiden Stichproben relativ homogen. „Teamfähigkeit“, „Selbstmotivation“ und „Zuverlässigkeit“ wurden von der Studentenchprobe um zwölf, elf und sieben Nennungen häufiger als in der Berufstätigenstichprobe genannt. „Kommunikationsfähigkeit“, „Sorgfalt“ und „Durchhaltevermögen“ wurden mit acht, fünf und vier Nennungen Abstand dagegen häufiger in der Berufstätigenstichprobe genannt.

Diskussion

Insgesamt zeigt die hohe Bandbreite an entstandenen Kategorien, dass in den Augen von Laien sehr viele Eigenschaften, Fähigkeiten und Kompetenzen relevant für beruflichen Erfolg sind. Die unterschiedliche Nennhäufigkeit zeigt auch, dass in den impliziten Theorien der Menschen einige wenige Eigenschaften wie z. B. soziale Kompetenz, Zuverlässigkeit und Teamfähigkeit, von extrem hoher Bedeutung sind, während viele andere Eigenschaften wie z. B. die Fähigkeit zum Netzwerken, Ordentlichkeit und Motivationsfähigkeit nur bei wenigen Personen einen zentralen Teil der impliziten Theorie darstellen. Dabei zeigt die geringere Standardabweichung, dass bei den Berufstätigen die Verteilung etwas ausgeglichener ist. Dies könnte ein Anhaltspunkt dafür sein, dass die implizite Theorie stärker ausgeprägt ist und durch die Erfahrung die Relevanz weiterer Eigenschaften festgestellt wurde.

Tabelle 3
Kategorien der impliziten Leistungstheorie und ihre Definitionen, sortiert nach genannter Häufigkeit in Kapitel 5.

Kategorie	Definition	Genannte Häufigkeit Gesamt	Häufigkeit Studentenstrichprobe	Häufigkeit Berufstätigenstrichprobe	Genannte Wichtigkeit (M)	Genannte Wichtigkeit (SD)
soziale Kompetenz	offener, freundlicher, respektvoller Umgang mit Menschen, sich einfühlen können	39 (11.11%)	19 (10.11%)	20 (12.27%)	7.92	2.05
Zuverlässigkeit	sich pünktlich an Absprachen halten/ Deadlines einhalten	37 (10.54%)	22 (11.70%)	15 (9.20%)	9.12	1.04
Teamfähigkeit	mit anderen zusammen arbeiten können	32 (9.12%)	22 (11.70%)	10 (6.13%)	7.13	2.08
Selbstmotivation	Eigenantrieb und Leidenschaft für die Arbeit	27 (7.69%)	19 (10.11%)	8 (4.91%)	8.38	1.45
Leistungsbereitschaft	über das "Normale" hinaus Leistung zeigen	25 (7.12%)	15 (7.98%)	10 (6.13%)	8.65	2.12
kreatives Problemlösen	kreativ und sein und Probleme lösen	15 (4.27%)	10 (5.32%)	5 (3.07%)	6.73	1.56
Flexibilität	zeitliche, örtliche, situationale und aufgabenbezogene Flexibilität	15 (4.27%)	7 (3.72%)	8 (4.91%)	8.31	1.32
Selbstbewusstsein	von sich selbst überzeugt, sicheres Auftreten	15 (4.27%)	8 (4.26%)	7 (4.29%)	8.31	2.39
Zielstrebigkeit	Ziele setzen und verfolgen	14 (3.99%)	6 (3.19%)	8 (4.91%)	8.00	1.77
Kommunikationsfähigkeit	Offene und verständliche Kommunikation	12 (3.42%)	2 (1.06%)	10 (6.13%)	8.58	1.00
Durchsetzungsvermögen	eigene Ideen/ Interessen durchsetzen	11 (3.13%)	6 (3.19%)	5 (3.07%)	6.92	1.17
Sorgfalt	Aufgaben genau, sorgfältig, gewissenhaft erledigen	11 (3.13%)	3 (1.60%)	8 (4.91%)	7.92	1.51
Offenheit für Neues	Allgemeine Offenheit für Neues und Anderes	10 (2.85%)	5 (2.66%)	5 (3.07%)	6.85	1.73
Stressresistenz	bei Stress/ Mehrarbeit Ruhe bewahren	9 (2.56%)	4 (2.13%)	5 (3.07%)	8.62	0.87
Lernbereitschaft	Bereitschaft zur Weiterentwicklung	9 (2.56%)	7 (3.72%)	2 (1.23%)	7.54	2.10
Durchhaltevermögen	bei Stress/ Widerständen dran bleiben, durchhalten	8 (2.28%)	2 (1.06%)	6 (3.68%)	8.92	1.15
Selbstorganisation	Strukturiertes und organisiertes Arbeiten	8 (2.28%)	4 (2.13%)	4 (2.45%)	8.15	1.14
Ehrlichkeit	ehrlich und authentisch sein (in Bezug auf Aufgaben und sich selbst)	7 (1.99%)	6 (3.19%)	1 (0.61%)	8.35	1.38
(eigene) Kritikfähigkeit	Kritik annehmen können	7 (1.99%)	2 (1.06%)	5 (3.07%)	8.08	1.94
Unternehmerisches Denken	Ausrichtung des eigenen Handelns am Interesse des Unternehmens	6 (1.71%)	4 (2.13%)	2 (1.23%)	7.38	2.57
Selbstständigkeit	selbstständig Probleme angehen, Aufgaben lösen	6 (1.71%)	2 (1.06%)	4 (2.45%)	7.08	1.40
Work-Life-Balance	Ausgleich zwischen Beruf und Familie	5 (1.42%)	2 (1.06%)	3 (1.85%)	7.50	2.18
Selbstmarketing	sich selbst gut verkaufen	5 (1.42%)	3 (1.60%)	2 (1.23%)	7.13	2.38
Fähigkeit, andere zu kritisieren	konstruktive Kritik üben	4 (1.14%)	3 (1.60%)	1 (0.61%)	6.19	2.50
Optimismus	Positive Lebenseinstellung, auch nach Rückschlägen	3 (0.85%)	3 (1.60%)	0 (0.00%)	8.46	1.71
Verantwortungsbewusstsein	Verantwortung für das eigene Handeln übernehmen	3 (0.85%)	0 (0.00%)	3 (1.85%)	7.85	2.04
Motivationsfähigkeit	andere motivieren und begeistern	3 (0.85%)	1 (0.53%)	2 (1.23%)	7.92	1.35
Ordentlich	Aufräumen, Ordnung halten	3 (0.85%)	1 (0.53%)	2 (1.23%)	4.88	2.22
Netzwerken	Netzwerke aufbauen und nutzen	2 (0.57%)	0 (0.00%)	2 (1.23%)	8.00	1.59
Gesamt		351	188	163		

Anmerkung. Die Angaben zur Wichtigkeit wurden in einer Folgestudie von Kneifel (2013) erhoben und hier ergänzt.

Unsere Studie zeigt einige Übereinstimmungen zu den oben genannten Studien von Wernimont (1971), Engle und Lord (1997), Carsten et al. (2010) und Sy (2010). So finden sich z. B. die Kategorien „Zuverlässigkeit“, „Teamfähigkeit“, „Leistungsbereitschaft“, „Kommunikationsfähigkeit“, „Ehrlichkeit“ und „Selbstständigkeit“ zumeist auch in diesen Studien. Es gibt allerdings auch Unterschiede. So weisen die anderen Studien grundsätzlich weniger Kategorien auf. Obwohl es sicher auch in dieser Studie möglich gewesen wäre, weitere Kategorien zusammenzufassen, sollte ein möglichst umfassendes Bild erarbeitet werden. Daher wäre eine weitere Verdichtung hier nicht sinnvoll gewesen. Zudem wurden durch die Verhaltensbeschreibungen doch klare Unterschiede zwischen den Kategorien deutlich und der hohe Übereinstimmungskoeffizient, den wir zum Schluss erreichen konnten, weist außerdem auf die Trennschärfe der Kategorien hin. Die Kategorien „soziale Kompetenz“, „kreatives Problemlösen“, „Selbstbewusstsein“, „Zielstrebigkeit“, „Durchsetzungsvermögen“, „Stressresistenz“, „Lernbereitschaft“, „Work-Life Balance“ und „Selbstmarketing“ finden sich in der Form bei den anderen Studien nicht oder die entsprechenden Eigenschaften sind in anderen Kategorien aufgegangen. Zwei Aspekte finden sich dafür in den anderen Studien, aber nicht unserer, nämlich die Themen Loyalität und Gehorsam gegenüber dem Vorgesetzten und Fähigkeiten/ Qualifizierung. Letzteres hatten wir bewusst aus unserer Analyse ausgeschlossen, da wir es nicht als Eigenschaft oder Kompetenz definiert haben. Ansonsten wäre möglicherweise eine solche Kategorie entstanden. Ersteres hat sich möglicherweise aufgrund der zugrundeliegenden Fragestellung bei uns nicht ergeben. In den anderen Studien ging es z. B. auch um Erwartungen von Vorgesetzten an ihre Mitarbeiter, während bei uns nach Eigenschaften, die den beruflichen Erfolg beeinflussen, gefragt wurde.

Insgesamt zeigt der Vergleich, dass unsere Kategorien eine recht hohe Übereinstimmung mit vorherigen Ergebnissen aufweisen, aber – ggf. aufgrund unterschiedlicher Auswertungsstrategien – breiter gefächert sind. Für Forschungszwecke, bei der Entwicklung von Persönlichkeitsmodellen und Kategoriensystemen scheint es immer

einfacher, die Datenbasis so gut und kompakt wie möglich zusammenzufassen. Möglicherweise ist dies in der Anwendungspraxis aber nicht immer hilfreich, da für Praktiker einzelne Facetten wichtig sein könnten, nicht die Komprimierung. Nach der deutschen Rechtsprechung muss beispielsweise bei Fragen nach Persönlichkeitseigenschaften der Bezug zur Tätigkeit und zum Arbeitsplatz vorhanden sein. Daher werden in der Regel allgemeine Intelligenz- oder Persönlichkeitstests als unzulässig angesehen (Wietfeld, 2016). Dann stellt möglicherweise die Auswahl aus verschiedenen Facetten ein notwendiges Angebot dar. Der Vergleich zeigt außerdem, dass es nicht nur um Eigenschaften geht, sondern vielfach auch um Kompetenzen. Für Laien scheint die Unterscheidung keine wesentliche Rolle zu spielen. Obwohl wir ausschließlich nach Eigenschaften gefragt haben, wurden doch zu einem großen Teil Kompetenzen genannt. Dies ist ein wichtiger Ansatzpunkt für die Research-Practice Gap. Möglicherweise greifen viele Praktiker auf wenig anerkannte Verfahren zurück (siehe auch Ergebnisse aus Kapitel 2), weil sie für sie (zumindest scheinbar) einen größeren Anwendungsbezug, eine höhere Augenscheinvalidität haben.

Vor diesem Hintergrund ist ein Vergleich mit vorhanden Persönlichkeitsmodellen bzw. den in Testverfahren gemessenen Kompetenzen interessant, um einen Eindruck darüber zu bekommen, ob unsere gefundenen Kompetenzen dort überhaupt Anwendung finden, in welchem Ausmaß und möglicherweise mit welcher Augenscheinvalidität. Ein Vergleich mit dem Persönlichkeitsmodell, das in der Wissenschaft wohl die größte Aufmerksamkeit bekommt – dem Big Five Modell bzw. dem NEO-PI-R (Ostendorf & Angleitner, 2004) als entsprechendem Messverfahren – zeigt einzelne Übereinstimmungen auf der Ebene der Facetten. So zählen die Facetten N4 (Soziale Befangenheit; Neurotizismus), A1 (Vertrauen; Verträglichkeit), A3 (Altruismus; Verträglichkeit) und A4 (Entgegenkommen; Verträglichkeit) sicher auf unsere Kategorie „soziale Kompetenz“ ein. Zu den Kategorien „Zuverlässigkeit“, „Sorgfalt“, „Ordentlich“ und „Leistungsbereitschaft“ passen die Facetten der Skala Gewissenhaftigkeit (z. B. C2 Ordnungsliebe, C3 Pflichtbewusstsein, C4 Leistungsstreben, C5

Selbstdisziplin). Die Kategorie „Offenheit für Neues“ findet Übereinstimmungen in der Skala Offenheit für Erfahrungen und einige Facetten der Skala Extraversion spielen eine Rolle für die Kategorien „Teamfähigkeit“ und „Durchsetzungsvermögen“ (z. B. E2 Geselligkeit, E3 Durchsetzungsfähigkeit). Allerdings könnte der NEO-PI-R aus verschiedenen Gründen dennoch nicht interessant sein für Praktiker: Zum einen werden z. T. vielleicht teilweise nur Skalen bekannt, die eben eine stärkere Zusammenfassung darstellen. Zum anderen passen die Bezeichnungen möglicherweise nicht zu den Begrifflichkeiten der Praktiker. Die Augenscheinvalidität für die Anwendung im beruflichen Kontext wird dadurch eingeschränkt.

Ein Vergleich mit dem Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung (Hossiep & Paschen, 2003) zeigt, dass insbesondere die häufig genannten Kategorien gut abgedeckt werden (vergleiche Tabelle 4). Dabei geht das Inventar explizit auch auf Kompetenzen ein. Die Skalen sind zumindest teilweise so benannt, dass wahrscheinlich eine höhere Augenscheinvalidität vorliegt (die Skalen Sensitivität und Soziabilität könnten allerdings ebenfalls für Laien schwer zu verstehen sein). Damit bietet der BIP Praktikern eine gute Möglichkeit, die für sie interessantesten Eigenschaften und Kompetenzen mit einem sorgfältig entwickelten und validierten Verfahren zu messen. In einer Studie von Diekmann und König (2015) war es bereits das in Deutschland am häufigsten genutzte Persönlichkeitsverfahren. Es gilt, solche Instrumente gezielt zu entwickeln bzw. weiterzuentwickeln und vor allem sie zielgerichtet an Praktiker zu vermarkten, um die Nutzung validierter Verfahren im Vergleich zu anderen Tests zu erhöhen.

Tabelle 4

Vergleich der Kategorien mit den Skalen des Bochumer Inventars zur berufsbezogenen Persönlichkeitsbeschreibung (BIP) (Hossiep & Paschen, 2003) in Kapitel 5

Kategorie	BIP Skalen
soziale Kompetenz	Sensitivität & Soziabilität
Zuverlässigkeit	Gewissenhaftigkeit
Teamfähigkeit	Kontaktfähigkeit & Teamorientierung
Selbstmotivation	Leistungsmotivation & Begeisterungsfähigkeit
Leistungsbereitschaft	Leistungsmotivation
kreatives Problemlösen	Gestaltungsmotivation
Flexibilität	Flexibilität
Selbstbewusstsein	Selbstbewusstsein
Zielstrebigkeit	Handlungsorientierung
Kommunikationsfähigkeit	
Durchsetzungsvermögen	Durchsetzungsstärke
Sorgfalt	Gewissenhaftigkeit
Offenheit für Neues	
Stressresistenz	Belastbarkeit
Lernbereitschaft	
Durchhaltevermögen	Emotionale Stabilität
Selbstorganisation	
Ehrlichkeit	
(eigene) Kritikfähigkeit	
Unternehmerisches Denken	
Selbstständigkeit	
Work-Life-Balance	
Selbstmarketing	
Fähigkeit, andere zu kritisieren	
Optimismus	
Verantwortungsbewusstsein	
Motivationsfähigkeit	Begeisterungsfähigkeit
Ordentlich	
Netzwerken	

Dazu ist es wichtig, die Bedarfe und (impliziten) Gedanken der Praktiker noch besser zu verstehen, um zielgerichtet darauf reagieren zu können. In einer unveröffentlichten Bachelorarbeit (Kneifel, 2013) unter Anleitung der Autorin dieser Dissertation wurde dieser Gedanke weiter verfolgt. Hier wurden insgesamt 21 Führungskräfte in halb- bis einstündigen halbstrukturierten Interviews befragt, wie die in der aktuellen Studie genannten Eigenschaften und Kompetenzen zum beruflichen Erfolg von Menschen beitragen, also welche tatsächlichen Theorien die Führungskräfte dazu haben und für wie wichtig sie die Eigenschaften jeweils halten. Die Interviews wurden transkribiert und ebenfalls angelehnt an Mayring (2010) eine

qualitative Vorgehensweise gewählt, um Ähnlichkeiten in den Aussagen bzw. zugrundeliegenden Theorien zu finden. Hier zeigt sich, dass es bei einigen Kategorien sehr hohe Übereinstimmungen zu einzelnen Aussagen gab und auch eine Verbindung zwischen verschiedenen Kategorien. Beispielsweise nannten die Befragten bei der Kategorie „Zuverlässigkeit“ mit einer sehr hohen Übereinstimmung den Einfluss von zuverlässigem Verhalten auf die Arbeit des Teams bzw. die Effizienz und Effektivität des Systems. Dazu gehören beispielsweise die Annahmen, dass

- ein Team nur dann leistungsstark ist, wenn alle an einem Strang ziehen,
- das Gesamtsystem bei unzuverlässigen Individuen nicht mehr funktioniert,
- Unzuverlässigkeit durch andere Teammitglieder oder die Führungskraft ausgeglichen werden muss, was wiederum Auswirkungen auf das Stresslevel, das Arbeitsklima und die Zusammenarbeit hat.

Die Befragten wiesen bei einigen Kategorien (z. B. Selbstorganisation oder Sorgfalt) auch auf die Branchen- bzw. Berufsabhängigkeit hin. Dies ist wiederum ein wichtiger Hinweis darauf, dass möglicherweise Persönlichkeitsinventare mit wenigen breiten Eigenschaften weniger praxistauglich sind als ein modulhafter Aufbau mit vielen Facetten (eine Idee, die wir in Kapitel 2 noch nicht betrachtet haben und die somit einen weiteren Ansatzpunkt für weitere Forschung bietet). Wieder andere Kategorien wurden kritisch betrachtet. Bei der Kategorie „Offenheit für Neues“ herrschte beispielsweise eine hohe Übereinstimmung dahingehend, dass – z. T. in Abhängigkeit von der Branche – Strukturen und Entscheidungen vom Unternehmen vorgegeben seien. In so einem Fall müsse man sich fügen und die Relevanz für den beruflichen Erfolg sei eher gering. Bei wieder anderen Kategorien wurde auf die Vor- und Nachteile verwiesen. Beispielsweise sei bei „Durchsetzungsvermögen“ entscheidend, ob jemand das Ellenbogenprinzip nutze oder diplomatisch vorgehe. Hier wären also bei einem Persönlichkeitsinventar nicht nur die Kategorien, sondern auch zugrundeliegenden Items entscheidend.

Eine weitere Betrachtungsweise ist auch die Beurteilung der Wichtigkeit der Kategorien für den beruflichen Erfolg. So wurde die zweithäufigste Kategorie „Zuverlässigkeit“ mit $M = 9.12$ ($SD = 1.04$) als am wichtigsten angesehen. Die „soziale Kompetenz“ und die „Teamfähigkeit“ wurden zwar sehr häufig genannt, werden mit $M = 7.92$ ($SD = 2.05$) und $M = 7.13$ ($SD = 2.08$) jedoch als weniger wichtig eingeschätzt als viele deutlich weniger häufig genannte Eigenschaften. Möglicherweise werden einige Eigenschaften schnell erinnert (weil sie zu einer gemeinsamen impliziten Theorie gehören), werden aber nicht immer als sehr relevant eingeschätzt (siehe Tabelle 3 für alle Einschätzungen der Wichtigkeit aus der ergänzenden Bachelorarbeit).

Diese Befragung stellt eine wichtige Ergänzung und Anregung für weitere Forschung zur aktuellen Studie dar. So scheinen nicht alle Kategorien gleich bedeutsam für jede Branche bzw. jeden Beruf zu sein. Darüber hinaus scheinen der Titel und die Beschreibung einer Kategorie relevant zu sein, damit der Nutzen für Praktiker erkennbar ist. Die weitere Forschung kann hier weitere Einblicke liefern um Nutzenargumente zur Überzeugung von Praktikern zu liefern und darüber hinaus tiefere Erkenntnisse in deren Denkweise für Forscher zu schaffen. In jedem Fall bietet diese Studie weitere neue Ansatzpunkte, welche die Hypothesen, die von Diekmann und König (2015), bzw. in Kapitel 2, aufgestellt und in den folgenden Kapiteln teilweise untersucht wurden, ergänzen.

Selbstverständlich hat auch diese Studie ihre Einschränkungen. Beispielsweise sind die beiden genannten Teilstichproben nicht sehr groß und nicht unbedingt repräsentativ für die Gesamtverteilung von Berufen und Unternehmensgrößen in Deutschland. Zudem handelt es sich ausschließlich um deutsche Teilnehmer. Eine uneingeschränkte Generalisierung der Ergebnisse ist daher nicht möglich. Gerade das Thema Persönlichkeit in der Personalauswahl ist möglicherweise kulturell geprägt und in Deutschland aufgrund der Rechtsprechung und wegen Vorbehalten von Teilnehmern und Unternehmen nicht ganz einfach. Vielleicht ist hier eine Diskussion über verschiedene Kulturkreise hinweg auch weniger sinnvoll, als eine lokal

ausgeprägte Forschung und Entwicklung von Lösungen. Die aktuelle und auch die Folgestudie folgen einem qualitativen Auswertungsschema. Dies scheint hier die richtige Lösung zu sein, um das Forschungsfeld zunächst zu erschließen. Allerdings sollte diese mittel- bis langfristig auch durch quantitative Auswertungen, z. B. faktoranalytische Auswertungen oder die Untersuchungen von Zusammenhängen/ Unterschieden zwischen Branchen oder Persönlichkeit der Befragten ergänzt werden.

Insgesamt dient die aktuelle Studie somit als guter Startpunkt, um implizite Leistungstheorien im Bereich der Personalauswahl zu erforschen und ein neues Licht auf die Research-Practice Gap werfen bzw. eine neue Perspektive zu eröffnen. Ziel sollte es sein, die „Schlucht“ zunehmend zu schließen, ein besseres gegenseitiges Verständnis für die Sichtweisen und Herausforderungen sowohl auf Praktiker- als auch auf Forscherseite zu entwickeln und schlussendlich praxistaugliche Lösungen zu entwickeln, welche die wichtigen Erkenntnisse insbesondere zur Validität von Kategorien und Verfahren in der Personalauswahl berücksichtigt.

KAPITEL 6

GESAMTDISKUSSION

Hauptziel dieser Dissertation war es, das Forschungsfeld der Research-Practice Gap in der Personalauswahl zu erfassen und im Speziellen die Frage nach dem Einsatz und Auswahlkriterien für Persönlichkeitstests in der Personalauswahl näher zu beleuchten. Aus diesem Grund wurden mehrere Studien durchgeführt. In Kapitel 2 wurde eine theoretische Einführung in den Themenbereich gegeben und eine Studie zum tatsächlichen Einsatz von Persönlichkeitsverfahren in der Personalauswahl durchgeführt. In Kapitel 3 und 4 wurden einzelne Hypothesen zu Einflusskriterien auf die Auswahl von Persönlichkeitstests getestet, nämlich die Verbindung eines fiktiven Testverfahrens mit neurowissenschaftlichen Befunden als Werbestrategie (Kapitel 3), einer typenbasierten vs. dimensionsbasierten Testart (Kapitel 4, Studie 1) und einer statistischen vs. theoriebasierten Entwicklungsweise (Kapitel 4, Studie 2). Zuletzt wurde in Kapitel 5 der Inhalt impliziter Leistungstheorien analysiert, also welche Eigenschaften, Fähigkeiten und Kompetenzen Menschen als relevant für beruflichen Erfolg erachten, sowie der Zusammenhang mit den erfassten Eigenschaften von aktuellen Persönlichkeitsverfahren. Die Ergebnisse werden nun zusammengefasst, Stärken und Grenzen diskutiert und Hinweise für weitere Forschung und praktische Implikationen erörtert.

Zusammenfassung der Ergebnisse

Anderson et al. (2001) starteten vor nun fast zwei Jahrzehnten mit dem Hinweis auf eine zunehmende Kluft zwischen Wissenschaftlern und Praktikern, der so genannten Research-Practice Gap, ein Thema, das seitdem vielfach von verschiedensten Wissenschaftler diskutiert wurde und weiterhin aktuell ist (siehe auch Tkachenko et al., 2017, für einen Überblick). Diese Kluft bezieht sich auf verschiedenste Aspekte im Bereich Human Resources Management im Allgemeinen, aber auch sehr speziell auf das Thema Personalauswahl (N. Anderson et al., 2004) und die Tatsache, dass für diesen Zweck häufig keine geeigneten Persönlichkeitstest ausgewählt werden (Murphy & Dzieweczynski, 2005), sofern sie überhaupt eingesetzt werden. Um diesen

Ausschnitt näher zu beleuchten wurde in Kapitel 2 zunächst ein Überblick über die grundsätzliche Debatte zum Nutzen bzw. der Validität von Persönlichkeitstests gegeben, vor allem aber eine Befragung von HR Praktikern durchgeführt, um die tatsächliche Nutzung von Persönlichkeitsverfahren im Rahmen der Personalauswahl besser zu verstehen. Hier zeigte sich, dass diese Tests in 15% der befragten Unternehmen eingesetzt wurden und dabei deren Nützlichkeit von den Praktikern vor allem bei der Auswahl von Führungskräften gesehen wurde. Ein Hauptergebnis dieser Befragung war die Tatsache, dass sehr viele verschiedene Testverfahren genutzt werden. Dies steht im deutlichen Gegensatz zu der Fokussierung vieler wissenschaftlicher Beiträge auf das Fünffaktorenmodell, den so genannten Big Five der Persönlichkeit und liefert so gleich zwei Hinweise auf das Bestehen einer Research-Practice Gap: Erstens nutzen Praktiker Verfahren, die nicht unbedingt dem Modell entsprechen, das von der Wissenschaft seit Jahren favorisiert wird. Zweitens zeigen sich in der Bandbreite an Testverfahren auch solche, die für Forscher möglicherweise wenig präsent sind und nicht unbedingt für den Zweck der Personalauswahl empfohlen werden. Damit liefert diese Studie einen (bisher selten geleisteten) Nachweis für diesen speziellen Teilbereich der Research-Practice Gap. Um nun Ableitungen für eine Annäherung zu Wissenschaft und Praxis ziehen zu können, sollte außerdem untersucht werden, worin sich verschiedene Persönlichkeitsverfahren eigentlich unterscheiden und welche dieser Ausprägungen für Praktiker möglicherweise besonders entscheidend oder attraktiv sein könnten. In Kapitel 2 wurden verschiedene Unterscheidungen angesprochen und bei der Befragung der Praktiker ebenfalls einbezogen. Hier zeigte sich unter anderem, dass Praktiker Typen bevorzugen, allerdings gleichzeitig die Ergebnisdarstellung in verschiedenen Facetten (wobei sie ein Zahlenprofil einem narrativen Bericht vorziehen), dass sie eine statistische der theoriebasierten Testentwicklung vorziehen und eine Testdurchführung am Computer bevorzugen. Wenn sie einen Test auswählen, bekommen sie lieber kurze und unabhängige Informationen über Qualitätskriterien und

Nutzenbeschreibungen, haben eine kleine Vorauswahl von Verfahren, die sie über Webseiten und Flyer suchen und die bevorzugt auch von anderen Unternehmen genutzt werden.

Aus diesen ersten Aussagen von Praktikern, die sicher durch die direkte Befragungsweise anfällig für soziale Erwünschtheit sind, wurden Hypothesen abgeleitet und in den Studien der Kapitel 3 und 4 getestet. In Kapitel 3, ging es – vor dem Hintergrund, dass erstens Praktiker vor allem über Webseiten und Flyer nach Persönlichkeitsverfahren suchen und damit Werbung ein wichtiger Einflussfaktor sein könnte und zweitens Studien gezeigt haben, dass eine neurowissenschaftliche Erklärung einen positiven Einfluss auf die Wahrnehmung von Aussagen hat – um die Frage, ob so eine neurowissenschaftliche Verbindung mit einem Persönlichkeitsverfahren dessen Bewertung beeinflusst. Entgegen der Erwartungen wurde ein fiktiver Persönlichkeitstest anhand eines Flyers signifikant weniger positiv beurteilt, wenn eine Verknüpfung zu neurowissenschaftlichen Erkenntnissen hergestellt wurde. Dieser Effekt war bei HR Praktikern am stärksten ausgeprägt, bei BWL-Studenten weniger stark und bei Psychologie-Studenten am geringsten. Unabhängig von der Richtung des Effekts zeigt diese Studie, dass (möglicherweise für die Qualität eines Verfahrens irrelevante Aspekte) deutlichen Einfluss auf die Auswahl desselben haben können.

In den Studien aus Kapitel 4 wurden zwei weitere Hypothesen aus dem zweiten Kapitel weiter getestet. In der ersten Hypothese ging es um die Frage, ob typen- oder dimensionsbasierte Persönlichkeitsverfahren von Praktikern als attraktiver wahrgenommen werden. In der Befragung aus Kapitel 2 gaben sie an, dass sie Typentests bevorzugen, aber auch unterschiedliche Facetten im Ergebnis dargestellt haben möchten. Um diese etwas widersprüchlichen Informationen experimentell zu untersuchen, wurde in einer Studie der Flyer aus Kapitel 3 entsprechend der Fragestellung angepasst und mit HR-Praktikern getestet und in einer zweiten Studie fiktive Ergebnisberichte erstellt, welche die Unterscheidung zwischen den beiden Bedingungen noch deutlicher hervorheben sollten, und mit BWL-Studenten geprüft. In keiner der beiden Studien ergab sich ein signifikanter Unterschied zwischen den beiden

Bedingungen. In der zweiten Hypothese ging es um die Frage, ob eine theoriebasierte oder statistische Entwicklungsmethode die Attraktivität eines Verfahrens beeinflusst, wobei in der Befragung aus Kapitel 2 die HR Praktiker die zweite Entwicklungsmethode bevorzugten. Mithilfe des Flyers wurde diese Hypothese mit drei Teilstichproben bestehend aus Praktikern, BWL- und Psychologie-Studenten getestet. Auch hier gab es keine signifikanten Unterschiede zwischen den beiden Bedingungen.

In den Studien aus Kapitel 5 wurde der Fokus auf die inhaltliche Gestaltung von Persönlichkeitsverfahren gelegt. Die Akzeptanz von Persönlichkeitsverfahren bei Praktikern sollte nicht nur von „harten“ Validitätskennwerten zur Vorhersagequalität abhängen, sondern auch von der Augenscheinvalidität, deren Bedeutung bei der Entwicklung von Verfahren immer mehr erkannt wird (Connell et al., 2018; Hardesty & Bearden, 2004). Auf der Basis der Forschung zu impliziten Theorien wurden hier Führungskräfte und Studenten befragt, welche Persönlichkeitseigenschaften und Kompetenzen ihrer Ansicht nach relevant für beruflichen Erfolg sind. Nach einer qualitativen Analyse fanden sich insgesamt 29 Kategorien von Eigenschaften und Kompetenzen, die mit unterschiedlicher Häufigkeit genannt wurden. Diese Anzahl ist deutlich höher, als die Anzahl an Eigenschaften, die in den meisten Persönlichkeitsverfahren erhoben wird. Ein Vergleich zum NEO-PI-R (Ostendorf & Angleitner, 2004) zeigt Gemeinsamkeiten und Unterschiede zwischen den Facetten und den gefundenen Kategorien auf, die möglicherweise Einfluss auf die Augenscheinvalidität und damit die Auswahl von Persönlichkeitsverfahren haben.

Erkenntnisse und Limitationen

Die dargestellten Studien haben einige Einsichten in ein bisher sehr wenig untersuchtes Forschungsfeld erbracht. Obwohl die Research-Practice Gap in den vergangenen Jahren bereits viel diskutiert wurde – die spezielle Kluft im Bereich Personalauswahl und Persönlichkeitstests wurde bisher kaum tiefergehend betrachtet. Es wird zwar zum Teil darauf hingewiesen, dass einige Persönlichkeitsverfahren, wie beispielsweise der MBTI, genutzt werden, obgleich Sie

aus Sicht der Forscher eher wenig empfehlenswert sind (siehe z. B. Hough & Oswald, 2005). Dennoch gab es bisher wenig Forschung dazu, welche Tests tatsächlich eingesetzt werden. Die Tatsache, dass die entsprechende Studie in Kapitel 2 diese Aussagen stützen konnte, mag zwar im Nachhinein wenig überraschen, dennoch gibt die Vielzahl an Tests bzw. Persönlichkeitsverfahren, die zum Teil in der Forschung möglicherweise bisher wenig bis gar nicht bekannt waren, klaren Anlass zu einer vertieften Analyse des Themas. Der zweite Teil der Befragung aus Kapitel 2, in der die HR Praktiker nach ihren Präferenzen in Bezug auf Unterschiede von Tests und deren Durchführung befragt wurden, lieferte Hinweise für die weitere Forschung und die Frage, welche Merkmale ein Persönlichkeitsverfahren eigentlich attraktiv für Praktiker macht. Hier sind jedoch auch klare Limitationen der Befragungsmethode erkennbar: Neben möglichen Einflüssen sozialer Erwünschtheit (z. B. bei der Frage, ob man die Angaben zu Qualitätskriterien über andere Quellen prüft oder den Angaben des Anbieters vertraut), fehlt die Möglichkeit zu Verständnisfragen (verstehen z. B. die Praktiker unter einem Typentest und Facetten dasselbe wie wir). Darüber hinaus spielt der Zeitaspekt bei Praktiker-Befragungen eine entscheidende Rolle bei der Teilnehmer-Akquise: Es ist kaum möglich, detaillierte Teilaspekte zu erfragen, da eine längere Befragungsdauer die Teilnahme immer unwahrscheinlicher macht. Die dargestellten Kriterien decken sicher nur einen kleinen Teil an Möglichkeiten ab.

In den Studien aus Kapitel 3 und 4 wurde versucht, einige Teilaspekte experimentell tiefergehend zu untersuchen. Hierfür nutzten wir in den meisten Studien eine verhältnismäßig einfache Manipulation: Den Flyer zu einem fiktiven Persönlichkeitstest, der je nach Fragestellung in zwei unterschiedlichen Versionen entwickelt wurde. Diese einfache Manipulation führte in der Studie aus Kapitel 3 zum Einfluss von neurowissenschaftlichen Informationen zu einem klaren (wenn auch unterwarteten) Effekt. Hier zeigt sich, dass möglicherweise nicht immer die Inhalte, die erfassten Merkmale allein, entscheidungsrelevant sind, sondern dass bereits kleine, werbewirksame Hinweise auf bestimmte

Entstehungsmerkmale, Zusammenhänge oder Testcharakteristika – in unserem Fall die neurowissenschaftlichen Informationen – entscheidungsrelevant sein können. Dies ist ein Aspekt, der für Forscher bei der Entwicklung von Verfahren vermutlich eher als weniger relevant eingeschätzt wird. In dem Versuch, objektive Kriterien zu entwickeln, anhand derer man Menschen unterscheiden und möglichst noch beruflichen Erfolg vorhersagen kann, wird die emotionale Komponente, die für den Vertrieb bzw. die Anwendung dieses Verfahrens in der Praxis nicht unerheblich ist, möglicherweise weniger betrachtet, als bei Anbietern auf dem freien Markt. Darüber hinaus gibt es eine zweite wichtige Erkenntnis: Erfahrung und Ausbildung beeinflussen die Sichtweise auf solche Aspekte: Während die Praktiker (mit vorwiegend betriebswirtschaftlichem Hintergrund) und BWL-Studenten die Verwendung des Verfahrens mit neurowissenschaftlichen Informationen eher abgelehnt haben, waren Psychologie-Studenten, die vermutlich in ihrem Studium mit den Neurowissenschaften in Berührung gekommen sind, neutral. Hier zeigt sich möglicherweise eine (allgemeine) Skepsis gegenüber dem Unbekannten, die dann entscheidungsrelevant werden kann.

In den Studien aus Kapiteln 4 zur Unterscheidung typenbasierter oder dimensionsbasierter und theoriebasierter oder statistisch entwickelter Test ergaben unsere Manipulationen keine signifikanten Unterschiede, obwohl die HR Praktiker aus der Studie in Kapitel 2 hier unterschiedlichen Präferenzen angaben. Dennoch bietet uns auch dieses Ergebnis eine wichtige Aussage: Während beispielsweise die Forschung sich klar auf dimensionsbasierte Testverfahren fokussiert, sehen zumindest einige Praktiker den Vorteil auch in einer typenbasierten Darstellung. Die Tatsache, dass wir keine signifikanten Unterschiede in den Bewertungen gefunden haben, könnte eben auch darauf hinweisen, dass einige Praktiker das eine und andere das andere bevorzugen. Dann wäre diese unterschiedliche persönliche Präferenz ebenfalls entscheidungsrelevant.

Bei der Betrachtung der Ergebnisse der Studien in Kapitel 3 und 4 gilt zu beachten: Um subjektive Einflüsse und sozial erwünschte Antworten auszuschließen sollte ein

experimentelles Design durchgeführt werden. Dies konnte aber nur in einer laborähnlichen simulierten Szenerie stattfinden. Aufgrund der Seltenheit so eines Entscheidungsprozesses ist die Beobachtung in der Praxis kaum möglich (aber wahrscheinlich auch nicht unbedingt nötig, siehe G. Mitchell, 2012). Dennoch ist so ein Entscheidungsprozess in der Praxis grundsätzlich deutlich komplexer. Ein Forschungsansatz, in dem mehrere Verfahren zur Auswahl vorgestellt werden, würde diese Komplexität sicher deutlich besser darstellen, dennoch würde dies zum einen mehr Zeit benötigen und damit (wie bereits oben beschrieben) die Teilnahmequote der HR Praktiker vermutlich deutlich senken und zudem den direkten Vergleich der verschiedenen Manipulationen deutlich schwieriger machen. Wir haben in unseren Studien versucht, die Wirksamkeitskontrolle der Manipulation (d. h. ob den Teilnehmern die für uns relevanten Aspekte aufgefallen sind) über ein offenes Fragefeld, in dem die Teilnehmer angeben konnten, welche Inhalte oder Darstellungen ihre Bewertung beeinflusst haben, abzubilden. Tatsächlich wurde hier selten ein Teil unserer Manipulation explizit genannt. Dennoch weißt das deutlich signifikante Ergebnis der Studie aus Kapitel 3 aus unserer Sicht auf die Effektivität der Manipulation hin.

In Kapitel 5 wurde ein etwas anderer Forschungsansatz verfolgt. Zum einen lag hier der Fokus auf den Inhalten, also den erfassten Eigenschaften an sich, zum anderen wurden die Befragungsergebnisse qualitativ analysiert. Die Ergebnisse liefern gleich mehrere relevante Erkenntnisse: Zunächst einmal weisen die insgesamt 29 Kategorien darauf hin, dass viele Eigenschaften und Kompetenzen für die Praxis relevant sein könnten und eine zu starke Verdichtung auf wenige Faktoren hier möglicherweise nicht den Bedarf abdeckt. Die in der Diskussion zu Kapitel 5 angesprochene Folgeuntersuchung (Kneifel, 2013) in Interviewform zeigt außerdem, dass Führungskräfte verschiedene Eigenschaften für verschiedene Branchen oder Berufe wichtiger oder weniger wichtiger finden und somit vermutlich ein modularer Aufbau eines Verfahrens attraktiv sein könnte. Die gefundenen Kategorien zeigen außerdem, dass Menschen nicht nur Eigenschaften im engeren Sinne, sondern auch Kompetenzen und

Fähigkeiten als entscheidend für beruflichen Erfolg erachten. Die Unterscheidung zwischen diesen beiden Bereichen könnte hier für Praktiker unerheblich sein, da sie Bewerber wollen, die sowohl extrovertiert (Eigenschaft) als auch teamfähig (Kompetenz) sind. Ein Verfahren, das beides abbildet, sollte demnach deutlich attraktiver sein als eines, das sich ausschließlich auf Eigenschaften fokussiert. In der vorliegenden Studie wurden Führungskräfte und Studenten befragt. Um die Generalisierbarkeit der Ergebnisse zu erhöhen, wäre eine größere Stichprobe und die Ergänzung um die Perspektive von Recruitern bzw. HR Praktikern wünschenswert. Für alle dargestellten Studien gilt außerdem: Die Stichproben bestanden immer aus deutschen Teilnehmern. Gerade beim Thema Persönlichkeit und Persönlichkeitsverfahren könnte hier eine deutsche Skepsis vorhanden sein, welche die Ergebnisse beeinflusst hat. Auch verbietet die deutsche Rechtsprechung die Anwendung von Persönlichkeitsverfahren, die nicht berufsbezogen sind, was den meisten HR Praktikern bewusst sein dürfte und unsere Befragungsergebnisse beeinflussen könnte.

Hinweise für zukünftige Forschung

Eine weitere Generalisierung würde daher zunächst die Überprüfung an internationalen Stichproben voraussetzen. Auch in Bezug auf die Frage, welche Persönlichkeitsverfahren genutzt werden, ist es sinnvoll, internationale Studien durchzuführen. Hier gilt es z. B. größere Stichproben zu befragen, die beispielsweise explizit Unternehmen unterschiedlicher Größe einbeziehen, und zu fragen, welche Verfahren im Rahmen der Personalauswahl und der Personalentwicklung eingesetzt werden und für welche Führungs- und Mitarbeitererebenen. Möglicherweise gibt es hier Unterschiede oder aber es zeigt sich, dass Unternehmen das Verfahren, das sie kennen, in allen möglichen Kontexten nutzen (wie externe und interne Personalauswahl oder Personalentwicklung). Aus möglichen internationalen Unterschieden lassen sich dann wieder nationale Fragestellungen und Lösungen ableiten.

Darüber hinaus könnte weitere Forschung die bisherigen Erkenntnisse aus der Entscheidungsforschung einbeziehen und versuchen, tatsächliche Entscheidungsprozesse zu

begleiten bzw. über Interviews genauer zu erfragen. Wie bereits an verschiedenen Stellen dieser Arbeit angedeutet, sind Entscheidungsprozesse deutlich komplexer, als sie in den vorliegenden Experimenten abgebildet werden konnten. Gerade im Rahmen von Interviews lassen sich möglicherweise weitere, wichtige Entscheidungskriterien erfragen, die wiederum weitere Hinweise auf Ansätze zur Reduktion der Research-Practice Gap bieten. Auch vor dem Hintergrund der Erkenntnisse aus Kapitel 5 könnten Praktiker-Interviews dazu dienen, die tatsächlichen Bedarfe von HR Praktikern im Auswahlprozess noch besser zu ergründen.

Als anwendungsorientiertes Forschungsfeld gilt es eine Balance zu finden zwischen den Ansprüchen der Wissenschaft sowohl an die Forschung als auch an die Testentwicklung auf der einen, als auch die Orientierung an der Praxis auf der anderen Seite – über Aufklärungsarbeit, aber auch die Entwicklung valider und passgenauer Lösungen. Die Forschung in den letzten Jahren hat sich viel mit den verschiedenen Auswahlritten und -methoden und deren Validität sowie mit Reaktionen von Bewerbern und Faking-Möglichkeiten beschäftigt und dort wichtige Erkenntnisse gewonnen (siehe z. B. Robertson & Smith, 2001; Salgado, 2017). Für die zukünftige Forschung gilt es, diese Perspektive um die Perspektive der Anwender, der HR-Praktiker und deren Bedarfe zu ergänzen, um einen Beitrag zur Reduzierung der Research-Practice Gap leisten zu können.

Implikationen für die Praxis

Zielsetzung der dargestellten Forschung war es, mehr über die Research-Practice Gap im Bereich Persönlichkeitstests und Personalauswahl herauszufinden und letztlich daraus erste Ideen für eine Verringerung dieser Gap ableiten zu können. In Kapitel 2 wurde bereits eine solche Idee angesprochen: Wenn die Anzahl von Testverfahren auf dem Markt so groß ist (Hough & Oswald, 2005), viele davon auch genutzt werden und Praktiker zumeist über das Internet danach suchen, wäre die Entwicklung einer öffentlichen Datenbank mit einer Liste von Persönlichkeitsverfahren sinnvoll, in der wichtige Daten z. B. zu Validität und erfassten Eigenschaften und/oder Kompetenzen angegeben sind, Empfehlungen zur Nutzung hinterlegt

sind und die Möglichkeit zur Bewertung der Verfahren durch Praktiker, Forscher und Testanbieter besteht. Dies würde die Informationssuche für Praktiker deutlich erleichtern und gegebenenfalls zur Nutzung von valideren Verfahren beitragen. Darüber hinaus sollte auch uns Forschern bewusst sein, dass die Vermarktung von Verfahren, die einen tatsächlichen Beitrag zur Vorhersage von Berufserfolg leisten, eine Rolle bei der Häufigkeit der Anwendung spielt. Viele Testverfahren in Deutschland werden über die Testzentrale des Hogrefe Verlags verkauft. Diese kommt zwar schon einer Datenbank nahe und gibt einen guten Überblick über Qualitätskriterien, doch die sehr sachliche Darstellungsweise könnte weniger ansprechend sein als bei Verfahren, die über andere Anbieter vertrieben werden, die ausschließlich ihr eigenes Produkt vermarkten.

Insgesamt weisen unsere Studien auch darauf hin, dass mehr Aufklärungsarbeit, wenn möglich schon im Studium, notwendig ist. Die Unterschiede in den Bewertungen des fiktiven Testverfahrens mit und ohne neurowissenschaftliche Informationen in Kapitel 3 zeigen, dass unterschiedliches Vorwissen die Bewertung und Auswahl eines Persönlichkeitsverfahrens klar beeinflussen kann. Hier gilt es, in allen Ausbildungs- und Studiengängen, die für eine HR Laufbahn relevant sind (wie die dargestellten Studien zeigen, sind dies in Deutschland häufig betriebswirtschaftliche Hintergründe), entsprechende Grundlagen zu setzen und zum Beispiel ein Verständnis für die Validität von Auswahlverfahren herzustellen, das bei Arbeits- und Organisationspsychologen deutlich besser ausgeprägt ist, als bei anderen Ausbildungshintergründen (Jackson, Dewberry, Gallagher, & Close, 2018). Aber auch nach der Ausbildung gilt es, HR Praktiker auf dem Laufenden zu halten und wichtige Informationen angemessen weiterzugeben. Hier wird es für viele Forscher eine Herausforderung sein, sich von der reinen Publikation in wissenschaftlichen Fachzeitschriften zu lösen und auf Kommunikationswege zu setzen, die auch von Nicht-Psychologen gelesen werden (Kanning & Thielsch, 2015).

Fazit

Die dargestellten Studien zeigen, dass die Research-Practice Gap im Bereich Persönlichkeitstests und Personalauswahl ein aktuelles Thema ist: Es kommen viele Verfahren zum Einsatz, welche die Forschung bisher kaum betrachtet hat. Hier gilt es für die Zukunft 1. ein besseres Verständnis für die Bedarfe von Praktikern zu bekommen, 2. die Kommunikation von Lösungen an die Praktiker zu verbessern und 3. passgenaue und valide Lösungen für die Praxis zu entwickeln.

LITERATURVERZEICHNIS

- Abrahamson, E., & Eisenman, M. (2001). Why management scholars must intervene strategically in the management knowledge market. *Human Relations*, *54*, 67–75. <https://doi.org/10.1177/0018726701541009>
- Anderson, C. A., & Lindsay, J. J. (1998). The development, perseverance, and change of naive theories. *Social Cognition*, *16*, 8–30. <https://doi.org/10.1521/soco.1998.16.1.8>
- Anderson, N., Herriot, P., & Hodgkinson, G. P. (2001). The practitioner-researcher divide in Industrial, Work and Organizational (IWO) psychology: Where are we now, and where do we go from here? *Journal of Occupational and Organizational Psychology*, *74*, 391–412. <https://doi.org/10.1348/096317901167451>
- Anderson, N., Lievens, F., van Dam, K., & Ryan, A. M. (2004). Future perspectives on employee selection: Key directions for future research and practice. *Applied Psychology: An International Review*, *53*, 487–501. <https://doi.org/10.1111/j.1464-0597.2004.00183.x>
- Asendorpf, J. (2012). *Persönlichkeitspsychologie für Bachelor [Personality psychology for bachelor students]*. Berlin: Springer.
- Bangerter, A., König, C. J., Blatti, S., & Salvisberg, A. (2009). How widespread is graphology in personnel selection practice? A case study of a job market myth. *International Journal of Selection and Assessment*, *17*, 219–230. <https://doi.org/10.1111/j.1468-2389.2009.00464.x>
- Barrena, R., & Sánchez, M. (2009). Using emotional benefits as a differentiation strategy in saturated markets. *Psychology and Marketing*, *26*, 1002–1030. <https://doi.org/10.1002/mar.20310>
- Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, *44*, 1–26. <https://doi.org/10.1111/j.1744-6570.1991.tb00688.x>

- Barrick, M. R., & Mount, M. K. (2005). Yes, personality matters: Moving on to more important matters. *Human Performance*, *18*, 359–372.
https://doi.org/10.1207/s15327043hup1804_3
- Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment*, *9*, 9–30.
<https://doi.org/10.1111/1468-2389.00160>
- Bartram, D. (1995). The role of computer-based test interpretation (CBTI) in occupational assessment. *International Journal of Selection and Assessment*, *3*, 178–185.
<https://doi.org/10.1111/j.1468-2389.1995.tb00025.x>
- Bartram, D. (2000). Internet recruitment and selection: Kissing frogs to find princes. *International Journal of Selection and Assessment*, *8*, 261–274.
<https://doi.org/10.1111/1468-2389.00155>
- Bartram, D. (2004). Assessment in organisations. *Applied Psychology: An International Review*, *53*, 237–259. <https://doi.org/10.1111/j.1464-0597.2004.00170.x>
- Bartram, D., & Brown, A. (2004). Online testing: Mode of administration and the stability of OPQ 32i scores. *International Journal of Selection and Assessment*, *12*, 278–284.
https://doi.org/10.1111/j.0965-075X.2004.282_1.x
- Bayne, R. (2005). *Ideas and evidence: Critical reflections on MBTI theory and practice* (1. ed.). Gainesville, Fla.: CAPT.
- Beck, D. M. (2010). The appeal of the brain in the popular press. *Perspectives on Psychological Science*, *5*, 762–766. <https://doi.org/10.1177/1745691610388779>
- Berchtold, M. (2005). Häufigste Auswahlverfahren in der Personalselektion [Most frequently used selection methods in personnel selection]. *HR Today*, *12/2005*, 37–39.
- Birkeland, S. A., Manson, T. M., Kisamore, J. L., Brannick, M. T., & Smith, M. A. (2006). A meta-analytic investigation of job applicant faking on personality measures.

- International Journal of Selection and Assessment*, 14, 317–335.
<https://doi.org/10.1111/j.1468-2389.2006.00354.x>
- Blickle, G., & Kramer, J. (2012). Intelligenz, Persönlichkeit, Einkommen und Fremdbeurteilungen der Leistung in sozialen Berufen [Intelligence, personality, income, and job performance assessments by others in social welfare jobs: A validation study]. *Zeitschrift für Arbeits- und Organisationspsychologie*, 56, 14–23.
<https://doi.org/10.1026/0932-4089/a000070>
- Briggs Myers, I., McCaulley, M. H., Quenk, N. L., & Hammer, A. L. (1998). *MBTI manual: A guide to the development and use of the Myers-Briggs Type Indicator* (3rd ed.). Palo Alto, CA: Consulting Psychologists Press.
- Briggs Myers, I., & Myers, P. B. (1993). *Gifts differing: Understanding personality type*. Palo Alto, CA: Consulting Psychologists Press.
- Brown, R. (1999). The use of personality tests: A survey of usage and practice in the UK. *Selection and Development Review*, 15, 3–8.
- Bruchon-Schweitzer, M., & Ferrieux, D. (1991). Une enquête sur le recrutement en France [An inquiry on personnel recruitment in France]. *European Review of Applied Psychology / Revue Européenne de Psychologie Appliquée*, 41, 9–17.
- Bryman, A. (1987). The generalizability of implicit leadership theory. *Journal of Social Psychology*, 127, 129–141. <https://doi.org/10.1080/00224545.1987.9713672>
- Buckley, M. R., Ferris, G. R., Bernardin, H. J., & Harvey, M. G. (1998). The disconnect between the science and practice of management. *Business Horizons*, 41, 31–38.
[https://doi.org/16/S0007-6813\(98\)90032-5](https://doi.org/16/S0007-6813(98)90032-5)
- Butcher, J. N., Perry, J. N., & Dean, B. L. (2009). How to use computer-based reports. In J. N. Butcher (Ed.), *Oxford Handbook of Personality Assessment* (pp. 693–706). Oxford, UK: Oxford University Press.

- Cacioppo, J. T., Berntson, G. G., Lorig, T. S., Norris, C. J., Rickett, E., & Nusbaum, H. (2003). Just because you're imaging the brain doesn't mean you can stop using your head: A primer and set of first principles. *Journal of Personality and Social Psychology*, *85*, 650–661. <https://doi.org/10.1037/0022-3514.85.4.650>
- Canli, T., & Amin, Z. (2002). Neuroimaging of emotion and personality: Scientific evidence and ethical considerations. *Brain and Cognition*, *50*, 414–431. [https://doi.org/10.1016/S0278-2626\(02\)00517-1](https://doi.org/10.1016/S0278-2626(02)00517-1)
- Carless, S. A. (2007). Graduate recruitment and selection in Australia. *International Journal of Selection and Assessment*, *15*, 153–166. <https://doi.org/10.1111/j.1468-2389.2007.00377.x>
- Carless, S. A., Rasiah, J., & Irmer, B. E. (2009). Discrepancy between human resource research and practice: Comparison of industrial/organisational psychologists and human resource practitioners' beliefs. *Australian Psychologist*, *44*, 105–111. <https://doi.org/10.1080/00050060802630015>
- Carsten, M. K., Uhl-Bien, M., West, B. J., Patera, J. L., & McGregor, R. (2010). Exploring social constructions of followership: A qualitative study. *The Leadership Quarterly*, *21*, 543–562. <https://doi.org/10.1016/j.leaqua.2010.03.015>
- Cattell, R. B. (1978). *The scientific use of factor analysis in behavioral and life sciences*. New York: Plenum Press.
- Chartered Institute of Personnel and Development. (2009). *Recruitment, retention and turnover. Annual survey report 2009* (pp. 1–42).
- Chuah, S. C., Drasgow, F., & Roberts, B. W. (2006). Personality assessment: Does the medium matter? No. *Journal of Research in Personality*, *40*, 359–376. <https://doi.org/10.1016/j.jrp.2005.01.006>
- Conn, S. R., & Rieke, M. L. (Eds.). (1994). *The 16PF fifth edition technical manual*. Champaign, IL: Institute for Personality and Ability Testing.

- Connell, J., Carlton, J., Grundy, A., Buck, E. T., Keetharuth, A. D., Ricketts, T., ... Brazier, J. (2018). The importance of content and face validity in instrument development: lessons learnt from service users when developing the Recovering Quality of Life measure (ReQoL). *Quality of Life Research*, 27, 1893–1902. <https://doi.org/10.1007/s11136-018-1847-y>
- Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.
- Deller, J., Süßmair, A., Albrecht, A.-G., & Bruchmüller, S. (2005). Akademische Ausbildung von Fachkräften für das Personalmanagement: Der HR-Manager als strategischer Partner [Perspectives of Academic Education of HR Specialists]. *Personalführung*, 38, 52–59.
- Detert, J. R., & Edmondson, A. C. (2011). Implicit voice theories: Taken-for-granted rules of self-censorship at work. *Academy of Management Journal*, 54, 461–488. <https://doi.org/10.5465/AMJ.2011.61967925>
- DeYoung, C. G., Hirsh, J. B., Shane, M. S., Papademetris, X., Rajeevan, N., & Gray, J. R. (2010). Testing predictions from personality neuroscience. Brain structure and the Big Five. *Psychological Science*, 21, 820–828. <https://doi.org/10.1177/0956797610370159>
- Di Milia, L. (2004). Australian management selection practices: Closing the gap between research findings and practice. *Asia Pacific Journal of Human Resources*, 42, 214–228. <https://doi.org/10.1177/1038411104045364>
- Di Milia, L., Smith, P. A., & Brown, D. F. (1994). Management selection in Australia: A comparison with British and French findings. *International Journal of Selection and Assessment*, 2, 80–90. <https://doi.org/10.1111/j.1468-2389.1994.tb00152.x>
- Diekmann, J., & König, C. J. (2015). Personality testing in personnel selection: Love it? Leave it? Change it! In I. Nikolaou & J. K. Oostrom (Eds.), *Employee recruitment, selection,*

- and assessment: Contemporary issues for theory and practice* (pp. 117–135). Hove, UK: Psychology Press.
- Diekmann, J., König, C. J., & Alles, J. (2015). The role of neuroscience information for choosing a personality test: Not as seductive as expected. *International Journal of Selection and Assessment*, *23*, 99–108. <https://doi.org/10.1111/ijsa.12099>
- Donston-Miller, D. (n.d.). Top personality tests used in hiring. Retrieved August 26, 2013, from <http://www.theladders.com/career-advice/top-personality-tests-hiring>
- Driver, R. W., Buckley, M. R., & Frink, D. D. (1996). Should we write off graphology? *International Journal of Selection and Assessment*, *4*, 78–86. <https://doi.org/10.1111/j.1468-2389.1996.tb00062.x>
- Dudley, N. M., Orvis, K. A., Lebiecki, J. E., & Cortina, J. M. (2006). A meta-analytic investigation of conscientiousness in the prediction of job performance: Examining the intercorrelations and the incremental validity of narrow traits. *Journal of Applied Psychology*, *91*, 40–57. <https://doi.org/10.1037/0021-9010.91.1.40>
- Dumit, J. (2004). *Picturing personhood: Brain scans and biomedical identity*. Princeton, NJ: Princeton University Press.
- Eden, D., & Leviatan, U. (1975). Implicit leadership theory as a determinant of the factor structure underlying supervisory behavior scales. *Journal of Applied Psychology*, *60*, 736–741. <https://doi.org/10.1037/0021-9010.60.6.736>
- Eleftheriou, A., & Robertson, I. (1999). A survey of management selection practices in Greece. *International Journal of Selection and Assessment*, *7*, 203–208. <https://doi.org/10.1111/1468-2389.00120>
- Engle, E. M., & Lord, R. G. (1997). Implicit theories, self-schemas, and leader-member exchange. *Academy of Management Journal*, *40*, 988–1010. <https://doi.org/10.2307/256956>

- Euteneier, R. J., & Scheelen, F. M. (2010). INSIGHTS MDI® by Scheelen - Verhalten, Werte, Fertigkeiten [INSIGHTS MDI® by Scheelen - behaviour, values, skills]. In W. Simon (Ed.), *Persönlichkeitsmodelle und Persönlichkeitstests [Models of personality and personality tests]* (pp. 19–36). Offenbach: GABAL.
- Evers, A., Muñiz, J., Bartram, D., Boben, D., Egeland, J., Fernández-Hermida, J. R., ... Urbánek, T. (2012). Testing practices in the 21st century. *European Psychologist, 17*, 300–319. <https://doi.org/10.1027/1016-9040/a000102>
- Fanelli, D. (2010a). Do pressures to publish increase scientists' bias? An empirical support from US States data. *PLoS ONE, 5*, e10271. <https://doi.org/10.1371/journal.pone.0010271>
- Fanelli, D. (2010b). “Positive” results increase down the hierarchy of the sciences. *PLoS ONE, 5*, e10068. <https://doi.org/10.1371/journal.pone.0010068>
- Fanelli, D. (2011). Negative results are disappearing from most disciplines and countries. *Scientometrics, 90*, 891–904. <https://doi.org/10.1007/s11192-011-0494-7>
- Farah, M. J. (2005). Neuroethics: The practical and the philosophical. *Trends in Cognitive Sciences, 9*, 34–40. <https://doi.org/10.1016/j.tics.2004.12.001>
- Farah, M. J., & Wolpe, P. R. (2004). Monitoring and manipulating brain function: New neuroscience technologies and their ethical implications. *Hastings Center Report, 34*, 35–45. <https://doi.org/10.2307/3528418>
- Fitzsimons, G. J., Hutchinson, J. W., Williams, P., Alba, J. W., Chartrand, T. L., Huber, J., ... Tavassoli, N. T. (2002). Non-conscious influences on consumer choice. *Marketing Letters, 13*, 269–279. <https://doi.org/10.1023/A:1020313710388>
- Fowler, R. D. (1985). Landmarks in computer-assisted psychological assessment. *Journal of Consulting and Clinical Psychology, 53*, 748–759. <https://doi.org/10.1037/0022-006X.53.6.748>

- Free Personality Test. (n.d.). Personality tests: The main personality tests (Myers Briggs, Big 5, 16 PF ...). Retrieved August 26, 2013, from <http://www.personalitytestfree.net/personality-tests.php>
- Furnham, A. (1982). Explanations for unemployment in Britain. *European Journal of Social Psychology, 12*, 335–352. <https://doi.org/10.1002/ejsp.2420120402>
- Furnham, A. (1988). *Lay theories: Everyday understanding of problems in the social sciences*. Elmsford, NY: Pergamon Press.
- Furnham, A. (1997). Lay theories of work stress. *Work and Stress, 11*, 68–78. <https://doi.org/10.1080/02678379708256823>
- Furnham, A. (2008). HR professionals' beliefs about, and knowledge of, assessment techniques and psychometric tests. *International Journal of Selection and Assessment, 16*, 300–305. <https://doi.org/10.1111/j.1468-2389.2008.00436.x>
- Furnham, A., Moutafi, J., & Crump, J. (2003). The relationship between the revised NEO-personality inventory and the Myers-Briggs Type Indicator. *Social Behavior and Personality: An International Journal, 31*, 577–584. <https://doi.org/10.2224/sbp.2003.31.6.577>
- Furnham, A., & Weir, C. (1996). Lay theories of child development. *Journal of Genetic Psychology, 157*, 211–226. <https://doi.org/10.1080/00221325.1996.9914859>
- Gangestad, S., & Snyder, M. (1985). “To carve nature at its joints”: On the existence of discrete classes in personality. *Psychological Review, 92*, 317–349. <https://doi.org/10.1037/0033-295X.92.3.317>
- Gelade, G. A. (2006a). But what does it mean in practice? The Journal of Occupational and Organizational Psychology from a practitioner perspective. *Journal of Occupational and Organizational Psychology, 79*, 153–160. <https://doi.org/10.1348/096317905X85638>

- Gelade, G. A. (2006b). Wider still and wider: Broadening the readership of the Journal of Occupational and Organizational Psychology. *Journal of Occupational and Organizational Psychology*, 79, 179–181. <https://doi.org/10.1348/096317906X103997>
- Ghiselli, E. E., & Barthol, R. P. (1953). The validity of personality inventories in the selection of employees. *Journal of Applied Psychology*, 37, 18–20. <https://doi.org/10.1037/h0059438>
- Göhs, N., & Dick, J. (2001). Testverfahren bei der Personalauswahl. *Personal - Zeitschrift für Human Resource Management*, 1, 46–48.
- Guastello, S. J., Guastello, D. D., & Craft, L. L. (1989). Assessment of the Barnum effect in computer-based test interpretations. *Journal of Psychology*, 123, 477.
- Guastello, S. J., & Rieke, M. L. (1990). The Barnum effect and validity of computer-based test interpretations: The Human Resource Development Report. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 2, 186–190. <https://doi.org/10.1037/1040-3590.2.2.186>
- Guion, R. M., & Gottier, R. F. (1965). Validity of personality measures in personnel selection. *Personnel Psychology*, 18, 135–164. <https://doi.org/10.1111/j.1744-6570.1965.tb00273.x>
- Guzzo, R. A., Wagner, D. B., Maguire, E., Herr, B., & Hawley, C. (1986). Implicit theories and the evaluation of group process and performance. *Organizational Behavior and Human Decision Processes*, 37, 279–295. [https://doi.org/10.1016/0749-5978\(86\)90056-7](https://doi.org/10.1016/0749-5978(86)90056-7)
- Hardesty, D. M., & Bearden, W. O. (2004). The use of expert judges in scale development: Implications for improving face validity of measures of unobservable constructs. *Journal of Business Research*, 57, 98–107. [https://doi.org/10.1016/S0148-2963\(01\)00295-8](https://doi.org/10.1016/S0148-2963(01)00295-8)

- Harris, M. M., Dworkin, J. B., & Park, J. (1990). Preemployment screening procedures: How human resource managers perceive them. *Journal of Business and Psychology, 4*, 279–292. <https://doi.org/10.1007/BF01125240>
- Haufe eCampus Redaktion. (2012, January 10). Berufswunsch: Personaler – Einstiegswege. Retrieved May 24, 2013, from <http://ecampus.haufe.de/karriere/berufswunsch-personaler-einstiegswege/>
- Heath, R., Brandt, D., & Nairn, A. (2006). Brand relationships: Strengthened by emotion, weakened by attention. *Journal of Advertising Research, 46*, 410–419.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York, NY: Wiley.
- Herrmann International. (2013). Herrmann International. Retrieved May 2, 2013, from <http://www.herrmanninternational.com/home/index.php>
- Herrmann International. (2015a). HBDI Assessment. Retrieved October 20, 2015, from <http://www.herrmannsolutions.com/hbdi-assessment/>
- Herrmann International. (2015b). Tap into our thinking. Retrieved March 30, 2015, from <http://www.herrmannsolutions.com/quick-facts/>
- Herrmann International. (n.d.). A history of innovation in business. Retrieved December 30, 2013, from <http://www.herrmannsolutions.com/our-foundation/>
- Herrmann International UK. (2015). The Originators and Trailblazers of Whole Brain® Thinking. Retrieved March 30, 2015, from <http://www.herrmannsolutions.co.uk/about/>
- Herrmann, N. (1989). *The creative brain*. Lake Lure, NC: Brain Books.
- Highhouse, S. (2008). Stubborn reliance on intuition and subjectivity in employee selection. *Industrial and Organizational Psychology, 1*, 333–342. <https://doi.org/10.1111/j.1754-9434.2008.00058.x>
- Hodgkinson, G. P. (2006). The role of JOOP (and other scientific journals) in bridging the practitioner-researcher divide in industrial, work and organizational (IWO) psychology.

- Journal of Occupational and Organizational Psychology*, 79, 173–178.
<https://doi.org/10.1348/096317906X104013>
- Hodgkinson, G. P., Daley, N., & Payne, R. L. (1995). Knowledge of, and attitudes towards, the demographic time bomb: A survey of its impact on graduate recruitment in the UK. *International Journal of Manpower*, 16, 59–76.
<http://www.emeraldinsight.com/ijm.htm>
- Hodgkinson, G. P., & Herriot, P. (2002). The role of psychologists in enhancing organizational effectiveness. In I. T. Robertson, M. Callinan, & D. Bartram (Eds.), *Organizational effectiveness: The role of psychology* (pp. 45–60). Chichester, UK: Wiley.
- Hodgkinson, G. P., Herriot, P., & Anderson, N. (2001). Re-aligning the stakeholders in management research: Lessons from Industrial, Work and Organizational Psychology. *British Journal of Management*, 12, 41–48. <https://doi.org/10.1111/1467-8551.12.s1.5>
- Hodgkinson, G. P., & Payne, R. L. (1998). Graduate selection in three European countries. *Journal of Occupational and Organizational Psychology*, 71, 359–365.
<https://doi.org/10.1111/j.2044-8325.1998.tb00682.x>
- Hogan, J., Barrett, P., & Hogan, R. (2007). Personality measurement, faking, and employment selection. *Journal of Applied Psychology*, 92, 1270–1285. <https://doi.org/10.1037/0021-9010.92.5.1270>
- Hogan, J., & Holland, B. (2003). Using theory to evaluate personality and job-performance relations: A socioanalytic perspective. *Journal of Applied Psychology*, 88, 100–112.
<https://doi.org/10.1037/0021-9010.88.1.100>
- Hogan, R. (1986). *Manual for the Hogan Personality Inventory*. Minneapolis: National Computer Systems.
- Hogan, R. (2005). In defense of personality measurement: New wine for old whiners. *Human Performance*, 18, 331–341. https://doi.org/10.1207/s15327043hup1804_1

- Hong, Y., Levy, S. R., & Chiu, C. (2001). The contribution of the lay theories approach to the study of groups. *Personality and Social Psychology Review*, 5, 98–106. https://doi.org/10.1207/S15327957PSPR0502_1
- Hook, C. J., & Farah, M. J. (2013). Look again: Effects of brain images and mind - Brain dualism on lay evaluations of research. *Journal of Cognitive Neuroscience*, 25, 1397–1405. https://doi.org/10.1162/jocn_a_00407
- Hoque, K., & Noon, M. (2001). Counting angels: A comparison of personnel and HR specialists. *Human Resource Management Journal*, 11, 5–22. <https://doi.org/10.1111/j.1748-8583.2001.tb00042.x>
- Hossiep, R., & Paschen, M. (2003). *Das Bochumer Inventar zur berufsbezogenen Persönlichkeitsbeschreibung: BIP*. Göttingen: Hogrefe.
- Hough, L. M., & Oswald, F. L. (2000). Personnel selection: Looking toward the future - remembering the past. *Annual Review of Psychology*, 51, 631–664. <https://doi.org/10.1146/annurev.psych.51.1.631>
- Hough, L. M., & Oswald, F. L. (2005). They're right, well ... mostly right: Research evidence and an agenda to rescue personality testing from 1960s insights. *Human Performance*, 18, 373–387. https://doi.org/10.1207/s15327043hup1804_4
- Huff, A. S. (2000). 1999 Presidential Address: Changes in organizational knowledge production. *Academy of Management Review*, 25, 288–293. <https://doi.org/10.5465/AMR.2000.3312916>
- Huffcutt, A. I., Conway, J. M., Roth, P. L., & Stone, N. J. (2001). Identification and meta-analytic assessment of psychological constructs measured in employment interviews. *Journal of Applied Psychology*, 86, 897–913. <https://doi.org/10.1037/0021-9010.86.5.897>
- Hülshager, U. R., & Maier, G. W. (2008). Persönlichkeitseigenschaften, Intelligenz und Erfolg im Beruf: Eine Bestandsaufnahme internationaler und nationaler Forschung.

- [Personality traits, intelligence, and success at work: A review of international and national research]. *Psychologische Rundschau*, 59, 108–122.
<https://doi.org/10.1026/0033-3042.59.2.108>
- Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. *Journal of Applied Psychology*, 85, 869–879. <https://doi.org/10.1037/0021-9010.85.6.869>
- Huselid, M. A. (1995). The impact of human resource management practices on turnover, productivity, and corporate financial performance. *Academy of Management Journal*, 38, 635–672. <https://doi.org/10.2307/256741>
- Illes, J., Kirschen, M. P., & Gabrieli, J. D. E. (2003). From neuroimaging to neuroethics. *Nature Neuroscience*, 6, 205. <https://doi.org/10.1038/nn0303-205>
- Jackson, D. J. R., Dewberry, C., Gallagher, J., & Close, L. (2018). A comparative study of practitioner perceptions of selection methods in the United Kingdom. *Journal of Occupational and Organizational Psychology*, 91, 33–56.
<https://doi.org/10.1111/joop.12187>
- John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114–156). New York, NY: Guilford Press.
- Johns, G. (1993). Constraints on the adoption of psychology-based personnel practices: Lessons from organizational innovation. *Personnel Psychology*, 46, 569–592.
<https://doi.org/10.1111/j.1744-6570.1993.tb00885.x>
- Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology*, 87, 765–780.
<https://doi.org/10.1037/0021-9010.87.4.765>

- Judge, T. A., Heller, D., & Mount, M. K. (2002). Five-factor model of personality and job satisfaction: A meta-analysis. *Journal of Applied Psychology, 87*, 530–541. <https://doi.org/10.1037/0021-9010.87.3.530>
- Judge, T. A., Higgins, C. A., Thoresen, C. J., & Barrick, M. R. (1999). The Big Five personality traits, general mental ability, and career success across the life span. *Personnel Psychology, 52*, 621–652. <https://doi.org/10.1111/j.1744-6570.1999.tb00174.x>
- Jung, C. G. (1960). *Psychologische Typen [Psychological types]* (9th ed.). Zürich: Rascher.
- Kanning, U. P. (2009). *ISK. Inventar sozialer Kompetenzen*. Göttingen: Hogrefe.
- Kanning, U. P., & Thielsch, M. T. (2015). Wie bilden Personalpraktiker/innen sich weiter? *Zeitschrift für Arbeits- und Organisationspsychologie, 59*, 206–214. <https://doi.org/10.1026/0932-4089/a000195>
- Keil, F. C. (2003). Folkscience: Coarse interpretations of a complex reality. *Trends in Cognitive Sciences, 7*, 368. [https://doi.org/10.1016/S1364-6613\(03\)00158-X](https://doi.org/10.1016/S1364-6613(03)00158-X)
- Keil, F. C. (2006). Explanation and understanding. *Annual Review of Psychology, 57*, 227–254. <https://doi.org/10.1146/annurev.psych.57.102904.190100>
- Kelemen, D. (1999). Function, goals and intention: Children's teleological reasoning about objects. *Trends in Cognitive Sciences, 3*, 461–468. [https://doi.org/10.1016/S1364-6613\(99\)01402-3](https://doi.org/10.1016/S1364-6613(99)01402-3)
- Kellett, D., McCahon, S., & James, J. (1991). Preliminary evaluation of five computer-generated narrative reports derived from four different personality questionnaires. *European Work and Organizational Psychologist, 1*, 196–210.
- Kelley, H. H. (1972). Attribution in social interaction. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 1–26). Morristown, NJ: General Learning Press.
- Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist, 28*, 107–128. <https://doi.org/10.1037/h0034225>

- Kelly, G. A. (1955). *The psychology of personal constructs. Volume 1: A theory of personality* (Vol. 1). New York: Norton.
- Kepes, S., Banks, G. C., McDaniel, M., & Whetzel, D. L. (2012). Publication bias in the organizational sciences. *Organizational Research Methods, 15*, 624–662. <https://doi.org/10.1177/1094428112452760>
- Kersting, M. (2008). *Qualität in der Diagnostik und Personalauswahl - der DIN-Ansatz* (1. Aufl.). Hogrefe-Verlag.
- Klehe, U.-C. (2004). Choosing how to choose: Institutional pressures affecting the adoption of personnel selection procedures. *International Journal of Selection and Assessment, 12*, 327–342. <https://doi.org/10.1111/j.0965-075X.2004.00288.x>
- Klein, C. (2010). Images are not the evidence in neuroimaging. *The British Journal for the Philosophy of Science, 61*, 265–278. <https://doi.org/10.1093/bjps/axp035>
- Kneifel, K. (2013). *Implizite Leistungstheorien: Persönlichkeitseigenschaften als Prädiktor für beruflichen Erfolg*. Unveröffentlichte Bachelorarbeit, Universität des Saarlandes.
- König, C. J., Klehe, U.-C., Berchtold, M., & Kleinmann, M. (2010). Reasons for being selective when choosing personnel selection procedures. *International Journal of Selection and Assessment, 18*, 17–27. <https://doi.org/10.1111/j.1468-2389.2010.00485.x>
- Kretschmann, H.-J., & Weinrich, W. (2007). *Klinische Neuroanatomie und kraniale Bilddiagnostik: Atlas der Magnetresonanztomographie und Computertomographie [Clinical neuroanatomy and cranial imaging diagnosis: Atlas of magnet resonance imaging and computed tomography]* (3rd ed.). Stuttgart, Germany: Thieme.
- Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2004). Academic performance, career potential, Creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology, 86*, 148–161. <https://doi.org/10.1037/0022-3514.86.1.148>

- Kundoor, V., & Mueen, A. K. K. (2010). Uncovering negative results: Introducing an open access journal "Journal of Pharmaceutical Negative Results." *Pharmacognosy Journal*, 2, 365–367. [https://doi.org/10.1016/S0975-3575\(10\)80016-1](https://doi.org/10.1016/S0975-3575(10)80016-1)
- Lado, A. A., & Wilson, M. C. (1994). Human resource systems and sustained competitive advantage: A competency-based perspective. *Academy of Management Review*, 19, 699–727. <https://doi.org/10.5465/AMR.1994.9412190216>
- Landis, R. S., James, L. R., Lance, C. E., Pierce, C. A., & Rogelberg, S. G. (2014). When is nothing something? Editorial for the null results special issue of Journal of Business and Psychology. *Journal of Business and Psychology*, 29, 163–167. <https://doi.org/10.1007/s10869-014-9347-8>
- Larson, J. R. J. (1982). Cognitive mechanisms mediating the impact of implicit theories of leader behavior on leader behavior ratings. *Organizational Behavior and Human Performance*, 29, 129–140. [https://doi.org/10.1016/0030-5073\(82\)90245-8](https://doi.org/10.1016/0030-5073(82)90245-8)
- Lewis, A., & Furnham, A. (1986). Reducing unemployment: Lay beliefs about how to reduce current unemployment. *Journal of Economic Psychology*, 7, 75–85. [https://doi.org/10.1016/0167-4870\(86\)90013-9](https://doi.org/10.1016/0167-4870(86)90013-9)
- Lickel, B., Hamilton, D. L., & Sherman, S. J. (2001). Elements of a lay theory of groups: Types of groups, relational styles, and the perception of group entitativity. *Personality and Social Psychology Review*, 5, 129–140. https://doi.org/10.1207/S15327957PSPR0502_4
- Lievens, F., & De Paepe, A. (2004). An empirical investigation of interviewer-related factors that discourage the use of high structure interviews. *Journal of Organizational Behavior*, 25, 29–46. <https://doi.org/10.1002/job.246>
- Lievens, F., & Harris, M. M. (2003). Research on internet recruiting and testing: Current status and future directions. In C. L. Cooper & I. T. Robertson (Eds.), *International review of industrial and organizational psychology* (Vol. 18). John Wiley & Sons, Ltd.

- Lindebaum, D., & Zundel, M. (2013). Not quite a revolution: Scrutinizing organizational neuroscience in leadership studies. *Human Relations*, *66*, 857–877. <https://doi.org/10.1177/0018726713482151>
- Lombrozo, T. (2006). The structure and function of explanations. *Trends in Cognitive Sciences*, *10*, 464–470. <https://doi.org/10.1177/0018726713482151>
- Lord, R. G., Binning, J. F., Rush, M. C., & Thomas, J. C. (1978). The effect of performance cues and leader behavior on questionnaire ratings of leadership behavior. *Organizational Behavior and Human Performance*, *21*, 27–39. [https://doi.org/10.1016/0030-5073\(78\)90036-3](https://doi.org/10.1016/0030-5073(78)90036-3)
- Macan, T. (2009). The employment interview: A review of current studies and directions for future research. *Human Resource Management Review*, *19*, 203–218. <https://doi.org/10.1016/j.hrmr.2009.03.006>
- MacLean, P. D. (1985). Evolutionary psychiatry and the triune brain. *Psychological Medicine*, *15*, 219–221. <https://doi.org/10.1017/S0033291700023485>
- Malle, B. F. (2004). *How the mind explains behavior: Folk explanations, meaning, and social interaction*. Cambridge, Mass.: MIT Press.
- Marcus, B., Goffin, R. D., Johnston, N. G., & Rothstein, M. G. (2007). Personality and cognitive ability as predictors of typical and maximum managerial performance. *Human Performance*, *20*, 275–285. <https://doi.org/10.1080/08959280701333362>
- Marston, W. M. (1979). *Emotions of normal people*. Minneapolis: Persona Press.
- Mayring, P. (2010). Qualitative Inhaltsanalyse. In G. Mey & K. Mruck (Eds.), *Handbuch Qualitative Forschung in der Psychologie* (pp. 601–613). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92052-8_42
- McCabe, D. P., & Castel, A. D. (2008). Seeing is believing: The effect of brain images on judgments of scientific reasoning. *Cognition*, *107*, 343–352. <https://doi.org/10.1016/j.cognition.2007.07.017>

- McCrae, R. R., & Costa, P. T. J. (1997). Personality trait structure as a human universal. *American Psychologist*, *52*, 509–516. <https://doi.org/10.1037/0003-066X.52.5.509>
- McDaniel, M. A., Whetzel, D. L., Schmidt, F. L., & Maurer, S. D. (1994). The validity of employment interviews: A comprehensive review and meta-analysis. *Journal of Applied Psychology*, *79*, 599–616. <https://doi.org/10.1037/0021-9010.79.4.599>
- Meade, A. W., Michels, L. C., & Lautenschlager, G. J. (2007). Are internet and paper-and-pencil personality tests truly comparable? An experimental design measurement invariance study. *Organizational Research Methods*, *10*, 322–345. <https://doi.org/10.1177/1094428106289393>
- Mintzberg, H. (1976). Planning on the left side and managing on the right. *Harvard Business Review*, *54*, 49–58.
- Mitchell, G. (2012). Revisiting truth or triviality: The external validity of research in the psychological laboratory. *Perspectives on Psychological Science*, *7*, 109–117. <https://doi.org/10.1177/1745691611432343>
- Mitchell, T. R., Larson Jr., J. R., & Green, S. G. (1977). Leader behavior, situational moderators, and group performance: An attributional analysis. *Organizational Behavior and Human Performance*, *18*, 254–268. [https://doi.org/10.1016/0030-5073\(77\)90030-7](https://doi.org/10.1016/0030-5073(77)90030-7)
- Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K. R., & Schmitt, N. (2007a). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology*, *60*, 683–729. <https://doi.org/10.1111/j.1744-6570.2007.00089.x>
- Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007b). Are we getting fooled again? Coming to terms with limitations in the use of personality tests for personnel selection. *Personnel Psychology*, *60*, 1029–1049. <https://doi.org/10.1111/j.1744-6570.2007.00100.x>

- Muñiz, J., & Fernández-Hermida, J. R. (2010). La opinión de los psicólogos españoles sobre el uso de los tests [Spanish psychologists opinions on test use]. *Papeles Del Psicólogo*, *31*, 108–121.
- Muñiz, J., Prieto, G., Almeida, L., & Bartram, D. (1999). Test use in Spain, Portugal and Latin American Countries. *European Journal of Psychological Assessment*, *15*, 151–157. <https://doi.org/10.1027//1015-5759.15.2.151>
- Murphy, K. R., & Dzieweczynski, J. L. (2005). Why don't measures of broad dimensions of personality perform better as predictors of job performance? *Human Performance*, *18*, 343–357. https://doi.org/10.1207/s15327043hup1804_2
- Nye, J. L. (2005). Implicit theories and leadership perceptions in the thick of it. The effects of prototype matching, group setbacks, and group outcomes. In B. Schyns & J. R. Meindl (Eds.), *Implicit leadership theories: Essays and explorations* (pp. 39–61). Greenwich, Conn: IAP.
- Oakland, T. (2004). Use of educational and psychological tests internationally. *Applied Psychology*, *53*, 157–172. <https://doi.org/10.1111/j.1464-0597.2004.00166.x>
- Offermann, L. R., Kennedy, J. K., & Wirtz, P. W. (1994). Implicit leadership theories: Content, structure, and generalizability. *Leadership Quarterly*, *5*, 43–58. [https://doi.org/10.1016/1048-9843\(94\)90005-1](https://doi.org/10.1016/1048-9843(94)90005-1)
- Ones, D. S., Dilchert, S., Viswesvaran, C., & Judge, T. A. (2007). In support of personality assessment in organizational settings. *Personnel Psychology*, *60*, 995–1027. <https://doi.org/10.1111/j.1744-6570.2007.00099.x>
- Ones, D. S., & Viswesvaran, C. (1996). Bandwidth–fidelity dilemma in personality measurement for personnel selection. *Journal of Organizational Behavior*, *17*, 609–626. [https://doi.org/10.1002/\(SICI\)1099-1379\(199611\)17:6<609::AID-JOB1828>3.0.CO;2-K](https://doi.org/10.1002/(SICI)1099-1379(199611)17:6<609::AID-JOB1828>3.0.CO;2-K)

- Ones, D. S., Viswesvaran, C., & Dilchert, S. (2005). Personality at work: Raising awareness and correcting misconceptions. *Human Performance, 18*, 389–404. https://doi.org/10.1207/s15327043hup1804_5
- Ostendorf, F., & Angleitner, A. (2004). *NEO-PI-R - NEO Persönlichkeitsinventar nach Costa und McCrae - Revidierte Fassung*. Göttingen: Hogrefe.
- Phillips, J. S., & Lord, R. G. (1986). Notes on the practical and theoretical consequences of implicit leadership theories for the future of leadership measurement. *Journal of Management, 12*, 31–41. <https://doi.org/10.1177/014920638601200104>
- Pillay, S. S. (2011). *Your brain and business: The neuroscience of great leaders*. Upper Saddle River, NJ: Pearson Education, FT Press.
- Piotrowski, C., & Armstrong, T. (2006). Current recruitment and selection practices: A national survey of fortune 1000 firms. *North American Journal of Psychology, 8*, 489–496.
- Pittenger, D. J. (2005). Cautionary comments regarding the Myers-Briggs Type Indicator. *Consulting Psychology Journal: Practice and Research, 57*, 210–221. <https://doi.org/10.1037/1065-9293.57.3.210>
- Ployhart, R. E., Weekley, J. A., Holtz, B. C., & Kemp, C. (2003). Web-based and paper-and-pencil testing of applicants in a proctored setting: Are personality, biodata, and situational judgment tests comparable? *Personnel Psychology, 56*, 733–752. <https://doi.org/10.1111/j.1744-6570.2003.tb00757.x>
- Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences, 10*, 59–63. <https://doi.org/10.1016/j.tics.2005.12.004>
- Poldrack, R. A. (2008). The role of fMRI in cognitive neuroscience: Where do we stand? *Current Opinion in Neurobiology, 18*, 223–227. <https://doi.org/10.1016/j.conb.2008.07.006>
- Poropat, A. E. (2009). A meta-analysis of the Five-Factor model of personality and academic performance. *Psychological Bulletin, 135*, 322–338. <https://doi.org/10.1037/a0014996>

- Psychometric Success. (2013). Understanding the Personality Test Industry. Retrieved August 26, 2013, from <http://www.psychometric-success.com/personality-tests/personality-tests-understanding-industry.htm>
- Racine, E., Bar-Ilan, O., & Illes, J. (2005). Science and society: fMRI in the public eye. *Nature Reviews Neuroscience*, *6*, 159–164. <https://doi.org/10.1038/nrn1609>
- Racine, E., Bar-Ilan, O., & Illes, J. (2006). Brain imaging: A decade of coverage in the print media. *Science Communication*, *28*, 122–143. <https://doi.org/10.1177/1075547006291990>
- Racine, E., Waldman, S., Rosenberg, J., & Illes, J. (2010). Contemporary neuroscience in the media. *Social Science & Medicine*, *71*, 725–733. <https://doi.org/10.1016/j.socscimed.2010.05.017>
- Rice, C. E. (1997). The scientist-practitioner split and the future of psychology. *American Psychologist*, *52*, 1173–1181. <https://doi.org/10.1037/0003-066X.52.11.1173>
- Ries, A., & Trout, J. (2001). *Positioning: The battle for your mind*. New York: McGraw-Hill.
- Robertson, I. T., & Smith, M. (2001). Personnel selection. *Journal of Occupational and Organizational Psychology*, *74*, 441–472. <https://doi.org/10.1348/096317901167479>
- Robie, C., Tuzinski, K. A., & Bly, P. R. (2006). A survey of assessor beliefs and practices related to faking. *Journal of Managerial Psychology*, *21*, 669–681. <https://doi.org/10.1108/02683940610690204>
- Robins, R. W., John, O. P., & Caspi, A. (1998). The typological approach to studying personality. In R. B. Cairns, L. R. Bergman, & J. Kagan (Eds.), *Methods and models for studying the individual* (pp. 135–160). Thousand Oaks, CA, US: Sage Publications, Inc.
- Robins, R. W., & Tracy, J. L. (2003). Setting an agenda for a person-centered approach to personality development. *Monographs of the Society for Research in Child Development*, *68*, 110–122. <https://doi.org/10.1111/j.1540-5834.2003.06801012.x>

- Rose, S. P. R. (2003). How to (or not to) communicate science. *Biochemical Society Transactions*, *31*, 307–312. <https://doi.org/10.1042/BST0310307>
- Roskies, A. L. (2008). Neuroimaging and inferential distance. *Neuroethics*, *1*, 19–30. <https://doi.org/10.1007/s12152-007-9003-3>
- Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review*, *96*, 341–357. <https://doi.org/10.1037/0033-295X.96.2.341>
- Rossiter, J., & Bellman, S. (2012). Emotional branding pays off: How brands meet share of requirements through bonding, companionship, and love. *Journal of Advertising Research*, *52*, 291–296. <https://doi.org/10.2501/JAR-52-3-291-296>
- Rothstein, M. G., & Goffin, R. D. (2000). The assessment of personality constructs in industrial-organizational psychology. In R. D. Goffin & E. Helmes (Eds.), *Problems and solutions in human assessment: Honoring Douglas N. Jackson at seventy*. (pp. 215–248). New York, NY US: Kluwer Academic/Plenum Publishers.
- Rothstein, M. G., & Goffin, R. D. (2006). The use of personality measures in personnel selection: What does current research support? *Human Resource Management Review*, *16*, 155–180. <https://doi.org/10.1016/j.hrmr.2006.03.004>
- Rousseau, D. M. (2006). Is there such a thing as “evidence-based management”? *Academy of Management Review*, *31*, 256–269. <https://doi.org/10.5465/AMR.2006.20208679>
- Rowe, P. M., Williams, M. C., & Day, A. L. (1994). Selection procedures in North America. *International Journal of Selection and Assessment*, *2*, 74–79. <https://doi.org/10.1111/j.1468-2389.1994.tb00153.x>
- Rozenblit, L., & Keil, F. C. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science*, *26*, 521–562. https://doi.org/10.1207/s15516709cog2605_1

- Rush, M. C., Phillips, J. S., & Lord, R. G. (1981). Effects of a temporal delay in rating on leader behavior descriptions: A laboratory investigation. *Journal of Applied Psychology, 66*, 442–450. <https://doi.org/10.1037/0021-9010.66.4.442>
- Rush, M. C., Thomas, J. C., & Lord, R. G. (1977). Implicit leadership theory: A potential threat to the internal validity of leader behavior questionnaires. *Organizational Behavior and Human Performance, 20*, 93–110. [https://doi.org/10.1016/0030-5073\(77\)90046-0](https://doi.org/10.1016/0030-5073(77)90046-0)
- Ryan, A. M., McFarland, L., Baron, H., & Page, R. (1999). An international look at selection practices: Nation and culture as explanations for variability in practice. *Personnel Psychology, 52*, 359–391. <https://doi.org/10.1111/j.1744-6570.1999.tb00165.x>
- Ryan, A. M., & Sackett, P. R. (1987). A survey of individual assessment practices by I/O psychologists. *Personnel Psychology, 40*, 455–488. <https://doi.org/10.1111/j.1744-6570.1987.tb00610.x>
- Ryan, A. M., & Sackett, P. R. (1992). Relationships between graduate training, professional affiliation, and individual psychological assessment practices for personnel decisions. *Personnel Psychology, 45*, 363–387. <https://doi.org/10.1111/j.1744-6570.1992.tb00854.x>
- Ryan, A. M., & Tippins, N. T. (2004). Attracting and selecting: What psychological research tells us. *Human Resource Management, 43*, 305–318.
- Rynes, S. L., Bartunek, J. M., & Daft, R. L. (2001). Across the great divide: Knowledge creation and transfer between practitioners and academics. *Academy of Management Journal, 44*, 340–355.
- Rynes, S. L., Brown, K. G., & Colbert, A. E. (2002). Seven common misconceptions about human resource practices: Research findings versus practitioner beliefs. *Academy of Management Executive, 16*, 92–103.

- Rynes, S. L., Colbert, A. E., & Brown, K. G. (2002). HR Professionals' beliefs about effective human resource practices: Correspondence between research and practice. *Human Resource Management, 41*, 149–174. <https://doi.org/10.1002/hrm.10029>
- Rynes, S. L., Giluk, T. L., & Brown, K. G. (2007). The very separate worlds of academic and practitioner periodicals in human resource management: Implications for evidence-based management. *Academy of Management Journal, 50*, 987–1008. <https://doi.org/10.5465/AMJ.2007.27151939>
- Rynes, S. L., Orlitzky, M. O., & Bretz, R. D. (1997). Experienced hiring versus college recruiting: Practices and emerging trends. *Personnel Psychology, 50*, 309–339. <https://doi.org/10.1111/j.1744-6570.1997.tb00910.x>
- Salgado, J. F. (1997). The Five Factor model of personality and job performance in the European Community. *Journal of Applied Psychology, 82*, 30–43. <https://doi.org/10.1037/0021-9010.82.1.30>
- Salgado, J. F. (2017). Personnel Selection. *Oxford Research Encyclopedia of Psychology*. <https://doi.org/10.1093/acrefore/9780190236557.013.8>
- Salgado, J. F., & Anderson, N. (2002). Cognitive and GMA testing in the european community: Issues and evidence. *Human Performance, 15*, 75–96. <https://doi.org/10.1080/08959285.2002.9668084>
- Salgado, J. F., & Moscoso, S. (2003). Internet-based personality testing: Equivalence of measures and assesses' perceptions and reactions. *International Journal of Selection and Assessment, 11*, 194–205. <https://doi.org/10.1111/1468-2389.00243>
- Salgado, J. F., Viswesvaran, C., & Ones, D. S. (2006). Predictors used for personnel selection: An overview of constructs, methods and techniques. In N. Anderson (Ed.), *Handbook of industrial, work and organizational psychology: Personnel psychology* (pp. 165–199). London, UK: SAGE.

- Sanders, K., van Riemsdijk, M., & Groen, B. (2008). The gap between research and practice: A replication study on the HR professionals' beliefs about effective human resource practices. *International Journal of Human Resource Management*, *19*, 1976–1988. <https://doi.org/10.1080/09585190802324304>
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*, 262–274. <https://doi.org/10.1037/0033-2909.124.2.262>
- Scholarios, D., & Lockyer, C. (1999). Recruiting and selecting professionals: Context, qualities and methods. *International Journal of Selection and Assessment*, *7*, 142–156. <https://doi.org/10.1111/1468-2389.00114>
- Schuler, H., Frier, D., & Kauffmann, M. (1993). *Personalauswahl im europäischen Vergleich [Personnel selection in a European comparison]*. Göttingen: Verlag für Angewandte Psychologie.
- Schuler, H., Hell, B., Trapmann, S., Schaar, H., & Boramir, I. (2007). Die Nutzung psychologischer Verfahren der externen Personalauswahl in deutschen Unternehmen: Ein Vergleich über 20 Jahre [Use of personnel selection instruments in German organizations during the last 20 years]. *Zeitschrift für Personalpsychologie*, *6*, 60–70. <https://doi.org/10.1026/1617-6391.6.2.60>
- Schweitzer, N. J., Baker, D. A., & Risko, E. F. (2013). Fooled by the brain: Re-examining the influence of neuroimages. *Cognition*, *129*, 501–511. <https://doi.org/10.1016/j.cognition.2013.08.009>
- Schyns, B., & Schilling, J. (2011). Implicit leadership theories: Think leader, think effective? *Journal of Management Inquiry*, *20*, 141–150. <https://doi.org/10.1177/1056492610375989>

- Shackleton, V., & Newell, S. (1994). European management selection methods: A comparison of five countries. *International Journal of Selection and Assessment*, 2, 91–102. <https://doi.org/10.1111/j.1468-2389.1994.tb00155.x>
- Shapiro, D. L., Kirkman, B. L., & Courtney, H. G. (2007). Perceived causes and solutions of the translation problem in management research. *Academy of Management Journal*, 50, 249–266. <https://doi.org/10.5465/AMJ.2007.24634433>
- Simon, H. A. (1972). Theories of bounded rationality. In C. McGuire & R. Radner (Eds.), *Decision and organization* (pp. 161–176). Amsterdam: North-Holland.
- Simon, H. A. (2000). Bounded rationality in social science: Today and tomorrow. *Mind & Society*, 1, 25–39. <https://doi.org/10.1007/BF02512227>
- Smith, E. R., & Queller, S. (2008). Mental representations. In A. Tesser & N. Schwarz (Eds.), *Blackwell handbook of social psychology: Intraindividual processes* (pp. 111–133). John Wiley & Sons.
- Sneath, F., Thakur, M., & Madjuck, B. (1976). *Testing people at work* (Information Report No. 24). London: Institute of Personnel Management.
- Snyder, D. K. (2000). Computer-assisted judgment: Defining strengths and liabilities. *Psychological Assessment*, 12, 52–60. <https://doi.org/10.1037/1040-3590.12.1.52>
- Sonnentag, S., & Frese, M. (2002). Performance concepts and performance theory. In S. Sonnentag (Ed.), *Psychological management of individual performance* (pp. 3–26). Chichester, UK: Wiley.
- Starkey, K., & Madan, P. (2001). Bridging the relevance gap: Aligning stakeholders in the future of management research. *British Journal of Management*, 12, 3–26.
- Steck, P. (1997). Psychologische Testverfahren in der Praxis: Ergebnisse einer Umfrage unter Testanwendern. [Psychological tests in practice: A survey among test users.]. *Diagnostica*, 43, 267–284.

- Stephan, U., & Westhoff, K. (2002). Personalauswahlgespräche im Führungskräftebereich des deutschen Mittelstandes: Bestandsaufnahme und Einsparungspotenzial durch strukturierte Gespräche [Personnel selection interviews in the management sector of German medium-sized business organizations: Current situation and potential improvements through structured interviews]. *Wirtschaftspsychologie*, *4*, 3–17.
- Sternberg, R. J. (1985). Implicit theories of intelligence, creativity, and wisdom. *Journal of Personality and Social Psychology*, *49*, 607–627. <https://doi.org/10.1037/0022-3514.49.3.607>
- Stewart, G. L., Darnold, T. C., Zimmerman, R. D., Parks, L., & Dustin, S. L. (2010). Exploring how response distortion of personality measures affects individuals. *Personality and Individual Differences*, *49*, 622–628. <https://doi.org/10.1016/j.paid.2010.05.035>
- Subramony, M. (2006). Why organizations adopt some human resource management practices and reject others: An exploration of rationales. *Human Resource Management*, *45*, 195–210.
- Swee, H.-Y. (2009). *A cognitive perspective of self-other agreement: A look at outcomes and predictors of shared implicit performance theories* (Doctorial dissertation). University of Akron. Retrieved from http://etd.ohiolink.edu/view.cgi?acc_num=akron1247775372
- Sy, T. (2010). What do you think of followers? Examining the content, structure, and consequences of implicit followership theories. *Organizational Behavior and Human Decision Processes*, *113*, 73–84. <https://doi.org/10.1016/j.obhdp.2010.06.001>
- Taylor, P., Keelty, Y., & McDonnell, B. (2002). Evolving personnel selection practices in New Zealand organisations and recruitment firms. *New Zealand Journal of Psychology*, *31*, 8–18.
- Tenhiälä, A., Giluk, T. L., Kepes, S., Simón, C., Oh, I.-S., & Kim, S. (2016). The research-practice gap in human resource management: A cross-cultural study. *Human Resource Management*, *55*, 179–200. <https://doi.org/10.1002/hrm.21656>

- Terpstra, D. E. (1996). The search for effective methods. *HR Focus*, 73, 16.
- Terpstra, D. E., & Rozell, E. J. (1993). The relationship of staffing practices to organizational level measures of performance. *Personnel Psychology*, 46, 27–48. <https://doi.org/10.1111/j.1744-6570.1993.tb00866.x>
- Terpstra, D. E., & Rozell, E. J. (1997). Why some potentially effective staffing practices are seldom used. *Public Personnel Management*, 26, 483. <https://doi.org/10.1177/009102609702600405>
- Tett, R. P., & Christiansen, N. D. (2007). Personality tests at the crossroads: A response to Morgeson, Campion, Dipboye, Hollenbeck, Murphy, and Schmitt (2007). *Personnel Psychology*, 60, 967–993. <https://doi.org/10.1111/j.1744-6570.2007.00098.x>
- Tett, R. P., Steele, J. R., & Beauregard, R. S. (2003). Broad and narrow measures on both sides of the personality–job performance relationship. *Journal of Organizational Behavior*, 24, 335–356. <https://doi.org/10.1002/job.191>
- Tewes, U. (1998). Qualitätsmanagement in der psychologischen Diagnostik [Quality management in psychological diagnostics]. *Zeitschrift für Medizinische Psychologie*, 7, 114–120.
- The Insights Group Limited. (2015). Insights Discovery Personal Profile. Retrieved October 20, 2015, from <https://www.insights.com/845/insights-discovery-personal-profile.html>
- Tkachenko, O., Hahn, H.-J., & Peterson, S. L. (2017). Research-practice gap in applied fields: An integrative literature review. *Human Resource Development Review*, 16, 235–262. <https://doi.org/10.1177/1534484317707562>
- Trapmann, S., Hell, B., Hirn, J.-O. W., & Schuler, H. (2007). Meta-analysis of the relationship between the Big Five and academic success at university. *Zeitschrift für Psychologie [Journal of Psychology]*, 215, 132–151. <https://doi.org/10.1027/0044-3409.215.2.132>

- Trout, J. D. (2002). Scientific explanation and the sense of understanding. *Philosophy of Science*, *69*, 212–233. <https://doi.org/10.1086/341050>
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131. <https://doi.org/10.1126/science.185.4157.1124>
- Uggerslev, K. L., & Sulsky, L. M. (2008). Using frame-of-reference training to understand the implications of rater idiosyncrasy for rating accuracy. *Journal of Applied Psychology*, *93*, 711–719. <https://doi.org/10.1037/0021-9010.93.3.711>
- Van de Ven, A. H., & Johnson, P. E. (2006). Knowledge for theory and practice. *Academy of Management Review*, *31*, 802–821. <https://doi.org/10.5465/AMR.2006.22527385>
- Van den Akker, A. L., Deković, M., Asscher, J. J., Shiner, R. L., & Prinzie, P. (2013). Personality types in childhood: Relations to latent trajectory classes of problem behavior and overreactive parenting across the transition into adolescence. *Journal of Personality and Social Psychology*, *104*, 750–764. <https://doi.org/10.1037/a0031184>
- van Dick, R. (2015). Registered reports, advance articles online, and the way ahead. *Journal of Personnel Psychology*, *14*, 1–3. <https://doi.org/10.1027/1866-5888/a000140>
- Wegener, D. T., & Petty, R. E. (1998). The naive scientist revisited: Naive theories and social judgment. *Social Cognition*, *16*, 1–7. <https://doi.org/10.1521/soco.1998.16.1.1>
- Weick, K. E. (2001). Gapping the relevance bridge: Fashions meet fundamentals in management research. *British Journal of Management*, *12*, 71–75. <https://doi.org/10.1111/1467-8551.12.s1.9>
- Weisberg, D. S. (2008). Caveat lector: The presentation of neuroscience information in the popular media. *Scientific Review of Mental Health Practice*, *6*, 51–56.
- Weisberg, D. S., Keil, F. C., Goodstein, J., Rawson, E., & Gray, J. R. (2007). The seductive allure of neuroscience explanations. *Journal of Cognitive Neuroscience*, *20*, 470–477. <https://doi.org/10.1162/jocn.2008.20040>

- Wernimont, P. F. (1971). What supervisors and subordinates expect of each other. *Personnel Journal*, 50, 204–208.
- Wietfeld, A. C. (2016). Der Einsatz psychologischer Testverfahren bei der Personalauswahl – Grundlagen und rechtliche Bewertung. *Zeitschrift für Arbeitsrecht*, 47, 215–260.
<https://doi.org/10.1515/zfa-2016-0204>
- Williams, R. S. (1992). Management selection in local government: A survey of practice in England and Wales. *Human Resource Management Journal*, 3, 63–73.
<https://doi.org/10.1111/j.1748-8583.1992.tb00311.x>
- Wilson, R. A., & Keil, F. C. (1998). The shadows and shallows of explanation. *Minds and Machines: Journal for Artificial Intelligence, Philosophy, and Cognitive Science*, 8, 137–159.
- Wirtz, M., & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität*. Göttingen: Hogrefe.
- York, K. L., & John, O. P. (1992). The four faces of Eve: A typological analysis of women's personality at midlife. *Journal of Personality and Social Psychology*, 63, 494–508.
<https://doi.org/10.1037/0022-3514.63.3.494>
- Zibarras, L. D., & Woods, S. A. (2010). A survey of UK selection practices across different organization sizes and industry sectors. *Journal of Occupational and Organizational Psychology*, 83, 499–511. <https://doi.org/10.1348/096317909X425203>

ANHÄNGE

Anhang A (zu Kapitel 3)

Flyer mit neurowissenschaftlicher Information

Personality at Work Inventory (P-WIN®)

Mit Persönlichkeit zum Erfolg.



Nutzen Sie die Erfolgsfaktoren der Persönlichkeit

Die Persönlichkeit Ihrer Mitarbeiter stellt eine bedeutsame Komponente des beruflichen und wirtschaftlichen Erfolges dar. Eines der wichtigsten Themen der Personalarbeit ist daher die Passung zwischen der Persönlichkeit eines Mitarbeiters und den Merkmalen des Unternehmens, der Unternehmenskultur.

Die besten Leistungen werden erzielt, wenn die Persönlichkeitseigenschaften zu den Arbeitsanforderungen passen!

Mit 12 Faktoren zum Erfolg

Das Personality at Work Inventory (P-WIN®) ist speziell für den beruflichen Kontext entwickelt worden. Das Instrument liefert Ihnen mit insgesamt 12 Erfolgsfaktoren eine objektive Entscheidungsgrundlage für die Auswahl des passenden Mitarbeiters. Dabei können bedeutsame Aussagen über die drei berufsrelevanten Bereiche soziales Gespür, Arbeitsweise sowie Emotion & Motivation getroffen werden.

Soziales Gespür

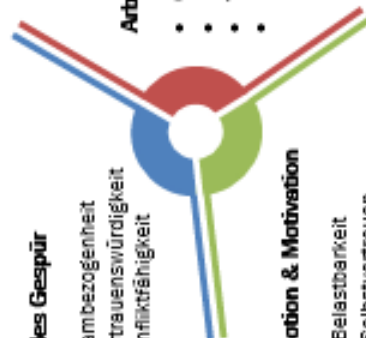
- Teambezogenheit
- Vertrauenswürdigkeit
- Konfliktfähigkeit

Arbeitsweise

- Gewissenhaftigkeit
- Flexibilität
- Verantwortungsbereitschaft
- Leistungsorientierung

Emotion & Motivation

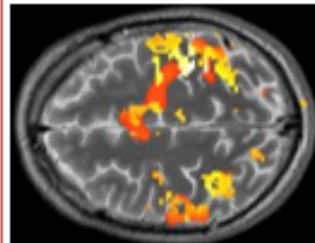
- Belastbarkeit
- Selbstvertrauen
- Selbstdisziplin
- Engagement



Ein innovatives Verfahren für innovative Unternehmen

P-WIN® ist das erste Instrument, das speziell die Erfolgsfaktoren der Persönlichkeit misst.

P-WIN® ist ein innovatives Verfahren. Es wurde in Kooperation mit erfahrenen WissenschaftlerInnen entwickelt und basiert auf neuesten neurowissenschaftlichen Erkenntnissen. Individuelle Persönlichkeitsmerkmale lassen sich durch die Aktivierung unterschiedlicher Gehirnbereiche nachweisen!



Ihr Nutzen

Erfolgreiche Menschen arbeiten in einem Arbeitsumfeld, das genau ihren persönlichen Stärken entspricht. Finden Sie in nur 30 Minuten den Mitarbeiter, dessen Interessen, Werte und Einstellungen am besten zu Ihren Anforderungen passen.

Die Work and Management Consulting Deutschland GmbH bietet Ihnen mit P-WIN® ein professionelles und innovatives Verfahren zur Personalauswahl. P-WIN® wird sowohl als Online-Fragebogen als auch in einer Papier- und Bleistift-Version angeboten.

Die Auswertung erfolgt exklusiv durch die Work and Management Consulting Deutschland GmbH. Sie erhalten ein graphisches Ergebnisprofil mit umfangreichen Erläuterungen, sowie auf Wunsch normative Vergleichsprofile aus den Bereichen Management, Verkauf, Produktion und Verwaltung.

Flyer ohne neurowissenschaftliche Information

Personality at Work Inventory (P-WIN®)

Mit Persönlichkeit zum Erfolg.



Nutzen Sie die Erfolgsfaktoren der Persönlichkeit

Die Persönlichkeit Ihrer Mitarbeiter stellt eine bedeutsame Komponente des beruflichen und wirtschaftlichen Erfolges dar. Eines der wichtigsten Themen der Personalarbeit ist daher die Passung zwischen der Persönlichkeit eines Mitarbeiters und den Merkmalen des Unternehmens, der Unternehmenskultur.

Die besten Leistungen werden erzielt, wenn die Persönlichkeitseigenschaften zu den Arbeitsanforderungen passen!

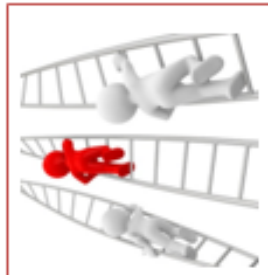
Mit 12 Faktoren zum Erfolg

Das Personality at Work Inventory (P-WIN®) ist speziell für den beruflichen Kontext entwickelt worden. Das Instrument liefert Ihnen mit insgesamt 12 Erfolgsfaktoren eine objektive Entscheidungsgrundlage für die Auswahl des passenden Mitarbeiters. Dabei können bedeutsame Aussagen über die drei berufsrelevanten Bereiche soziales Gespür, Arbeitsweise sowie Emotion & Motivation getroffen werden.



Ein innovatives Verfahren für innovative Unternehmen

P-WIN® ist das erste Instrument, das speziell die Erfolgsfaktoren der Persönlichkeit misst.



P-WIN® ist ein innovatives Verfahren. Es wurde in Kooperation mit erfahrenen WissenschaftlerInnen entwickelt und basiert auf neuesten wissenschaftlichen Erkenntnissen.

Ihr Nutzen


Erfolgreiche Menschen arbeiten in einem Arbeitsumfeld, das genau ihren persönlichen Stärken entspricht. Finden Sie in nur 30 Minuten den Mitarbeiter, dessen Interessen, Werte und Einstellungen am besten zu Ihren Anforderungen passen.

Die Work and Management Consulting Deutschland GmbH bietet Ihnen mit P-WIN® ein professionelles und innovatives Verfahren zur Personalauswahl. P-WIN® wird sowohl als Online-Fragebogen als auch in einer Papier- und Bleistift-Version angeboten.

Die Auswertung erfolgt exklusiv durch die Work and Management Consulting Deutschland GmbH. Sie erhalten ein graphisches Ergebnisprofil mit umfangreichen Erläuterungen, sowie auf Wunsch normative Vergleichsprofile aus den Bereichen Management, Verkauf, Produktion und Verwaltung.

Anhang B (zu Kapitel 4)

Flyer Experiment 1a Typen



P-WIN[®]
PERSONALITY AT WORK INVENTORY

Mit Persönlichkeit zum Erfolg:

Die Persönlichkeit Ihrer Mitarbeiter stellt eine beachtliche Komponente des beruflichen und wirtschaftlichen Erfolgs dar. Eines der wichtigsten Themen der Personalarbeit ist daher die Passung zwischen der Persönlichkeit des Mitarbeiters und den Merkmalen des Unternehmens.


Die besten Leistungen werden erzielt, wenn die Persönlichkeitseigenschaften zu den Arbeitsanforderungen passen!

Nutzen Sie diesen Vorteil. Erheben Sie die für Sie relevanten Persönlichkeitseigenschaften Ihrer Bewerber anhand unseres wissenschaftlich fundierten Verfahrens.

Die 8 Basistypen:

Das Personality at Work Inventory **P-WIN**[®] wurde speziell für den beruflichen Kontext entwickelt. Das Instrument liefert Ihnen eine objektive Entscheidungsgrundlage für die Auswahl des passenden Mitarbeiters.

Wie die Forschung zeigt, gibt es acht **Basistypen** der Persönlichkeit. Diese unterschiedlichen Typen zeichnen sich durch verschiedene Kompetenzbereiche aus. Da je nach Arbeitsprofil unterschiedliche Anforderungen an den Mitarbeiter gestellt werden, ist es wichtig, die Eignung Ihrer Mitarbeiter sicherzustellen.



“Personality variables show what employees ‘will do’, not just what they ‘can do’.” Dr. L. Hough

Ihr Vorteil:


Erfolgreiche Menschen arbeiten in einem Arbeitsumfeld, das genau ihren persönlichen Stärken entspricht. Finden Sie in nur 30 Minuten den Mitarbeiter, dessen Interessen, Werte und Einstellungen am besten zu Ihren Anforderungen passen.

Die **Work and Management Consulting Deutschland GmbH** bietet Ihnen mit **P-WIN**[®] ein professionelles und innovatives Verfahren zur Personalauswahl. **P-WIN**[®] ist sowohl als Online Fragebogen als auch in einer Papier- und Bleistift-Version verfügbar.

Die Auswertung erfolgt exklusiv durch die **Work and Management Consulting Deutschland GmbH**. Sie erhalten ein graphisches Ergebnisprofil mit umfangreichen Erläuterungen sowie normative Vergleichsprofile aus den Bereichen Management, Verkauf, Produktion und Verwaltung.

Work and Management Consulting Deutschland GmbH
 Wilhelm-Leuch-Allee 3
 D-66121 Saarbrücken
 info@pwin.de

Flyer Experiment 1a Dimensionen



P-WIN[®]
PERSONALITY AT WORK INVENTORY

Mit Persönlichkeit zum Erfolg:

Die Persönlichkeit Ihrer Mitarbeiter stellt eine beachtliche Komponente des beruflichen und wirtschaftlichen Erfolgs dar. Eines der wichtigsten Themen der Personalarbeit ist daher die Passung zwischen der Persönlichkeit des Mitarbeiters und den Merkmalen des Unternehmens.

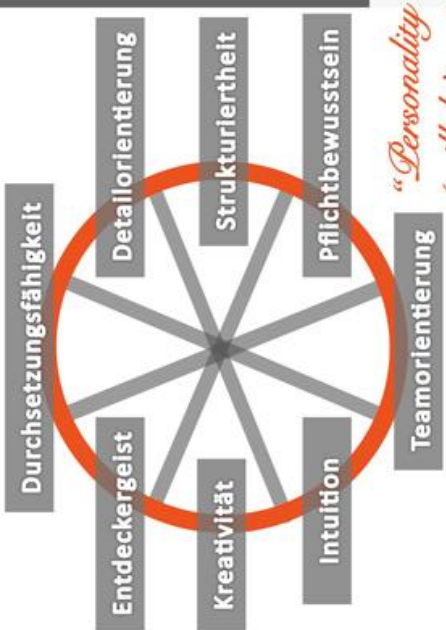
Die besten Leistungen werden erzielt, wenn die Persönlichkeitseigenschaften zu den Arbeitsanforderungen passen!

Nutzen Sie diesen Vorteil. Erheben Sie die für Sie relevanten Persönlichkeitseigenschaften Ihrer Bewerber anhand unseres wissenschaftlich fundierten Verfahrens.

Die 8 Basisdimensionen:

Das Personality at Work INVENTORY P-WIN[®] wurde speziell für den beruflichen Kontext entwickelt. Das Instrument liefert Ihnen eine objektive Entscheidungsgrundlage für die Auswahl des passenden Mitarbeiters.

Wie die Forschung zeigt, gibt es acht **Basisdimensionen** der Persönlichkeit. Diese unterschiedlichen Dimensionen zeichnen sich durch verschiedene Kompetenzbereiche aus. Da je nach Arbeitsprofil unterschiedliche Anforderungen an den Mitarbeiter gestellt werden, ist es wichtig, die Eignung Ihrer Mitarbeiter sicherzustellen.



“Personality variables show what employees ‘will do’, not just what they ‘can do’.” Dr. L. Hough

Ihr Vorteil:

Erfolgreiche Menschen arbeiten in einem Arbeitsumfeld, das genau ihren persönlichen Stärken entspricht. Finden Sie in nur 30 Minuten den Mitarbeiter, dessen Interessen, Werte und Einstellungen am besten zu Ihren Anforderungen passen.

Die **Work and Management Consulting Deutschland GmbH** bietet Ihnen mit **P-WIN**[®] ein professionelles und innovatives Verfahren zur Personalauswahl. **P-WIN**[®] ist sowohl als Online Fragebogen als auch in einer Papier- und-Bleistift-Version verfügbar.

Die Auswertung erfolgt exklusiv durch die **Work and Management Consulting Deutschland GmbH**. Sie erhalten ein graphisches Ergebnisprofil mit umfangreichen Erläuterungen sowie normative Vergleichsprofile aus den Bereichen Management, Verkauf, Produktion und Verwaltung.

Work and Management Consulting Deutschland GmbH

Wilhelm-Leuch-Allee 3

D-66121 Saarbrücken

info@pwin.de

Ergebnisbericht Experiment 1b Typen



Personality at Work Inventory – P-WIN®

Ergebnisprofil für
Max Mustermann,
Beispiel GmbH

Die Work and Management Consulting Deutschland GmbH bietet Ihnen mit P-WIN® ein professionelles und innovatives Verfahren zur Personalauswahl. P-WIN® wurde speziell für den beruflichen Kontext entwickelt. Das Instrument liefert Ihnen eine objektive Entscheidungsgrundlage für die Auswahl des passenden Mitarbeiters.

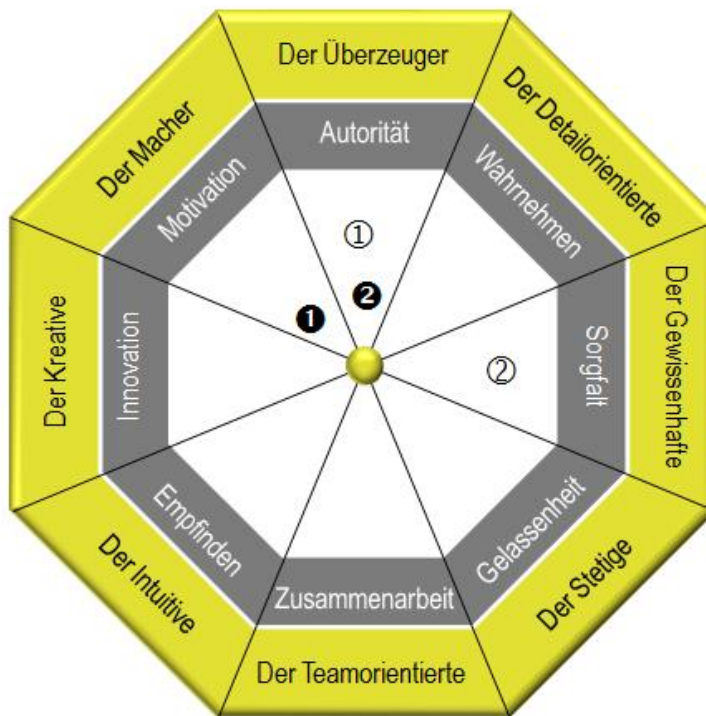
Wie die Forschung zeigt, gibt es acht Basistypen der Persönlichkeit. Diese unterschiedlichen Typen zeichnen sich durch verschiedene Kompetenzbereiche aus. Da je nach Arbeitsprofil unterschiedliche Anforderungen an den Mitarbeiter gestellt werden, ist es wichtig, die Eignung Ihrer Mitarbeiter sicherzustellen.

Work and Management Consulting Deutschland GmbH
Wilhelm-Leuch-Allee 3
D-66121 Saarbrücken

phone (0049) 681 – 111 11 00
info@pwin.de
www.pwin.de



Das P-WIN[®] Profil von Herrn Mustermann



① Primäre Präferenz: Dieser Typ entspricht der Persönlichkeit am ehesten; er symbolisiert die natürlichen, angeborenen und erlernten Verhaltensweisen.

② Sekundäre Präferenz: Dieser Typ kennzeichnet das Verhalten in stressigen Situationen. Hier kann eine Abweichung zur primären Präferenz auftreten (muss aber nicht).

❶ / ❷ Die vom Unternehmen definierten idealen Präferenzen.

Die acht Basistypen der Persönlichkeit

Der Überzeuger: Menschen dieses Typs treten sehr bestimmt auf und schaffen es, ihr Gegenüber mit guten Argumenten und ihrer natürlichen Autorität von ihren Ideen zu überzeugen.

Der Detailorientierte: Menschen dieses Typs richten ihre Wahrnehmung bei ihrer Arbeit immer auch auf die Details und sorgen so dafür, dass alle Aspekte ausreichend berücksichtigt werden.

Der Gewissenhafte: Menschen dieses Typs gehen ihre Aufgaben sehr strukturiert und sorgfältig an und erzeugen so Ordnung und Übersicht für sich und ihre Kollegen.

Der Stetige: Menschen dieses Typs finden sich auch in stressigen Situationen gelassen und schaffen ein stabiles, harmonisches Arbeitsumfeld für sich und ihre Kollegen.

Der Teamorientierte: Menschen dieses Typs richten ihre Denk- und Arbeitsweise optimal am Team aus und sorgen dafür, dass ein Team gemeinsam ziel- und erfolgsorientiert handelt.

Der Intuitive: Menschen dieses Typs sind ganzheitlich orientiert und richten ihr Handeln an den Bedürfnissen und Emotionen ihrer Mitmenschen aus.

Der Kreative: Menschen dieses Typs sind durch großen Einfallsreichtum und hohes Innovationsstreben gekennzeichnet. Kreative Menschen können sich für Veränderungen und alles Neue begeistern.

Der Macher: Menschen dieses Typs haben eine hohe Leistungsmotivation und schaffen es, sie im Sinne des Unternehmens ein- bzw. in ergebnisorientierte Handlungen umzusetzen.

Ergebnisbericht Experiment 1b Dimensionen



Personality at Work Inventory – P-WIN®

Ergebnisprofil für
Max Mustermann,
Beispiel GmbH

Die Work and Management Consulting Deutschland GmbH bietet Ihnen mit P-WIN® ein professionelles und innovatives Verfahren zur Personalauswahl. P-WIN® wurde speziell für den beruflichen Kontext entwickelt. Das Instrument liefert Ihnen eine objektive Entscheidungsgrundlage für die Auswahl des passenden Mitarbeiters.

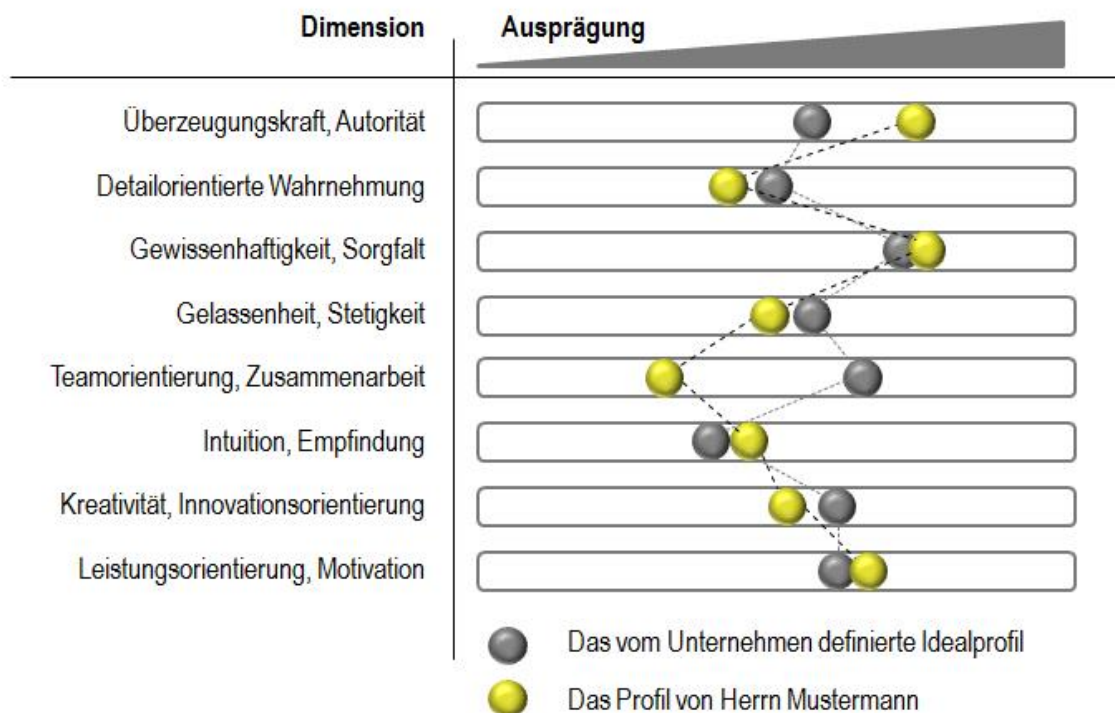
Wie die Forschung zeigt, gibt es acht Basisdimensionen der Persönlichkeit. Diese unterschiedlichen Dimensionen zeichnen sich durch verschiedene Kompetenzbereiche aus. Da je nach Arbeitsprofil unterschiedliche Anforderungen an den Mitarbeiter gestellt werden, ist es wichtig, die Eignung Ihrer Mitarbeiter sicherzustellen.

Work and Management Consulting Deutschland GmbH
Wilhelm-Leuch-Allee 3
D-66121 Saarbrücken

phone (0049) 681 – 111 11 00
info@pwin.de
www.pwin.de



Das P-WIN[®] Profil von Herrn Mustermann



Die acht Basisdimensionen der Persönlichkeit

Überzeugungskraft, Autorität: Menschen mit hoher Ausprägung treten sehr bestimmt auf und schaffen es, ihr Gegenüber mit guten Argumenten und ihrer natürlichen Autorität von ihren Ideen zu überzeugen.

Detailorientierte Wahrnehmung: Menschen mit hoher Ausprägung richten ihre Wahrnehmung bei ihrer Arbeit immer auch auf die Details und sorgen so dafür, dass alle Aspekte ausreichend berücksichtigt werden.

Gewissenhaftigkeit, Sorgfalt: Menschen mit hoher Ausprägung gehen ihre Aufgaben sehr strukturiert und sorgfältig an und erzeugen so Ordnung und Übersicht für sich und ihre Kollegen.

Gelassenheit, Stetigkeit: Menschen mit hoher Ausprägung sind auch in stressigen Situationen gelassenen und schaffen ein stabiles, harmonisches Arbeitsumfeld für sich und ihre Kollegen.


Teamorientierung, Zusammenarbeit: Menschen mit hoher Ausprägung richten ihre Denk- und Arbeitsweise optimal am Team aus und sorgen dafür, dass ein Team gemeinsam ziel- und erfolgsorientiert handelt.

Intuition, Empfindung: Menschen mit hoher Ausprägung sind ganzheitlich orientiert und richten ihr Handeln an den Bedürfnissen und Emotionen ihrer Mitmenschen aus.

Kreativität, Innovationsorientierung: Menschen mit hoher Ausprägung sind durch großen Einfallsreichtum und hohes Innovationsstreben gekennzeichnet. Kreative Menschen können sich für Veränderungen und alles Neue begeistern.

Leistungsorientierung, Motivation: Menschen mit hoher Ausprägung haben eine hohe Leistungsmotivation und schaffen es, sie im Sinne des Unternehmens ein- bzw. in ergebnisorientierte Handlungen umzusetzen.

Flyer Experiment 2 Statistik



Mit Persönlichkeit zum Erfolg

Die Persönlichkeit Ihrer Mitarbeiter stellt eine beachtliche Komponente des beruflichen und wirtschaftlichen Erfolgs dar. Eines der wichtigsten Themen der Personalarbeit ist daher die Passung zwischen der Persönlichkeit des Mitarbeiters und den Merkmalen des Unternehmens.

Die besten Leistungen werden erzielt, wenn die Persönlichkeitseigenschaften zu den Arbeitsanforderungen passen! Nutzen Sie diesen Vorteil. Erheben Sie die für Sie relevanten Persönlichkeitseigenschaften Ihrer Bewerber anhand unseres wissenschaftlich fundierten Verfahrens.

Konstruktion des Tests

Der amerikanische Psychologe Prof. Dr. Smith hat sich in seiner langjährigen Forschung damit beschäftigt, wie man mit statistischen Methoden wichtige Persönlichkeitseigenschaften voneinander unterscheiden kann. Seine statistischen Ergebnisse zeigen, dass Leistung im Beruf insbesondere von acht grundlegenden Persönlichkeitseigenschaften abhängt.


Das auf der Basis dieses statistischen Ansatzes entwickelte *Personality at Work Inventory P-WIN* erfasst diese wichtigen individuellen Unterschiede. Es fußt somit auf anerkannten statistischen Methoden.

Innovationsorientierung

Belastbarkeit

Teamorientierung

Gewissenhaftigkeit



Überzeugungskraft

Gelassenheit

Intuition

Leistungsorientierung

Ihr Vorteil

P-WIN wurde speziell für den beruflichen Kontext entwickelt. Das Instrument liefert Ihnen eine objektive Entscheidungsgrundlage für die Auswahl des passenden Mitarbeiters.

Finden Sie in nur 30 Minuten den Mitarbeiter, dessen Interessen, Werte und Einstellungen am besten zu Ihren Anforderungen passen.

Die **Work and Management Consulting Deutschland GmbH** bietet Ihnen mit **P-WIN** ein professionelles und innovatives Verfahren zur Personalauswahl. **P-WIN** ist sowohl als Online Fragebogen als auch in einer Papier- und Bleistift-Version verfügbar.

Die Auswertung erfolgt exklusiv durch die **Work and Management Consulting Deutschland GmbH**. Sie erhalten ein übersichtliches graphisches Ergebnisprofil mit umfangreichen Erläuterungen sowie normative Vergleichsprofile aus den Bereichen Management, Verkauf, Produktion und Verwaltung.

Work and Management Consulting Deutschland GmbH
Wilhelm-Leuch-Allée 3
D-66121 Saarbrücken
info@pwin.de

Flyer Experiment 2 Theorie



Mit Persönlichkeit zum Erfolg

Die Persönlichkeit Ihrer Mitarbeiter stellt eine beachtliche Komponente des beruflichen und wirtschaftlichen Erfolgs dar. Eines der wichtigsten Themen der Personalarbeit ist daher die Passung zwischen der Persönlichkeit des Mitarbeiters und den Merkmalen des Unternehmens.

Die besten Leistungen werden erzielt, wenn die Persönlichkeitseigenschaften zu den Arbeitsanforderungen passen!

Nutzen Sie diesen Vorteil. Erheben Sie die für Sie relevanten Persönlichkeitseigenschaften Ihrer Bewerber anhand unseres wissenschaftlich fundierten Verfahrens.

Konstruktion des Tests

Der amerikanische Psychologe Prof. Dr. Smith entwickelte auf der Basis langjähriger Forschung eine umfassende Theorie der Persönlichkeit, die berufliche Leistung vorhersagen kann.

Seine Theorie erklärt, wie Leistung im Beruf insbesondere von acht grundlegenden Persönlichkeitseigenschaften abhängt.

Das auf der Basis dieser Theorie entwickelte Personality at Work Inventory **P-WIN** erfasst diese wichtigen individuellen Unterschiede. Es fußt somit auf einer anerkannten Theorie der Persönlichkeit.


Innovationsorientierung

Überzeugungskraft

Belastbarkeit

Teamorientierung

Gewissenhaftigkeit



Gelassenheit

Intuition

Leistungsorientierung

Ihr Vorteil

P-WIN wurde speziell für den beruflichen Kontext entwickelt. Das Instrument liefert Ihnen eine objektive Entscheidungsgrundlage für die Auswahl des passenden Mitarbeiters.

Finden Sie in nur 30 Minuten den Mitarbeiter, dessen Interessen, Werte und Einstellungen am besten zu Ihren Anforderungen passen.

Die **Work and Management Consulting Deutschland GmbH** bietet Ihnen mit **P-WIN** ein professionelles und innovatives Verfahren zur Personalauswahl. **P-WIN** ist sowohl als Online Fragebogen als auch in einer Papier- und Bleistift-Version verfügbar.

Die Auswertung erfolgt exklusiv durch die **Work and Management Consulting Deutschland GmbH**. Sie erhalten ein übersichtliches graphisches Ergebnisprofil mit umfangreichen Erläuterungen sowie normative Vergleichsprofile aus den Bereichen Management, Verkauf, Produktion und Verwaltung.

Work and Management Consulting Deutschland GmbH

Wilhelm-Leuch-Allée 3

D-66121 Saarbrücken

info@pwin.de

CURRICULUM VITAE

Janina Diekmann
 Universität des Saarlandes
 Arbeits- und Organisationspsychologie
 Campus A1 3, 66123 Saarbrücken, Germany
 j.diekmann@mx.uni-saarland.de

STUDIUM

- | | |
|-------------------|--|
| Seit 12/2010 | UNIVERSITÄT DES SAARLANDES, DOKTORANDIN DER PSYCHOLOGIE
Forschungsschwerpunkt: Personalauswahl, Persönlichkeitstests,
Research-Practice Gap |
| 10/2004 – 10/2010 | UNIVERSITÄT OSNABRÜCK, STUDIUM DER PSYCHOLOGIE
Schwerpunkte: Arbeits- und Organisationspsychologie
sowie Gesundheits- und Rehabilitationspsychologie;
Nebenfach: BWL (Interkulturelles Management)
Diplomarbeit: Validierung und Optimierung eines kognitiven
Eignungstests;
Abschluss: Diplom (Note 1,22) |

PRAXISERFAHRUNG

- | | |
|-------------------|---|
| Seit 01/2016 | REFERENTIN/ MANAGERIN PERSONALENTWICKLUNG
CEMEX Deutschland AG |
| 10/2011 – 12/2015 | REFERENTIN PERSONALENTWICKLUNG
Cosmos Lebensversicherungs-AG |
| 12/2010 – 11/2013 | WISSENSCHAFTLICHE MITARBEITERIN
Universität des Saarlandes, Lehrstuhl für Arbeits- und
Organisationspsychologie |
| 01/2009 – 10/2010 | WERKSTUDENTIN
Essenzio Beratungsgesellschaft mbH |
| 03/2008 – 05/2008 | PRAKTIKANTIN
Abteilung für Personalmanagement der MKO |
| 08/2007 – 12/2008 | PRAKTIKANTIN UND WERKSTUDENTIN
Meinsen & Steinhübel Organisationsberatung |
| 12/2006 – 03/2008 | STUDENTISCHE HILFSKRAFT UND TUTORIN
Universität Osnabrück |

PUBLIKATIONEN

PEER-REVIEWED JOURNALS

- Diekmann, J., König, C. J., & Alles, J. (2015). The role of neuroscience information for choosing a personality test: Not as seductive as expected. *International Journal of Selection and Assessment*, 23, 99-108. doi:10.1111/ijsa.12099
- Diekmann, J., König, C. J., Schadow, T., Knab, Y., & Harrison, O. V. (2016). Finding the right (test) type: On the differences between type- vs. dimension-based personality tests and between statistics- vs. theory-based personality tests when deciding for or against a test in personnel selection. *Zeitschrift für Sozialmanagement/Journal of Social Management*, 14(2), 83-102.

BUCHBEITRÄGE

- Diekmann, J., & König, C. J. (2015). Personality testing in personnel selection: Love it? Leave it? Understand it! In I. Nikolaou & J. Ostrom (eds.), *Employee recruitment, selection, and assessment: Contemporary issues for theory and practice*. Hove, UK: Psychology Press.
- Meinsen, S. & Diekmann, J. (2011). Wissensmanagement. In: G. Hensen & P. Hensen (Hrsg.): *Gesundheits- und Sozialmanagement – Leitbegriffe und Grundlagen modernen Managements*. Stuttgart: Kohlhammer.

BETREUUNG VON ABSCHLUSSARBEITEN

DIPLOMARBEITEN

- Alles, J. (2013). *Die Research-Practice-Gap bei der Verwendung von Persönlichkeitstests in der Personalauswahl*. Bachelorarbeit, Arbeitseinheit Arbeits- und Organisationspsychologie, Universität des Saarlandes.

BACHELORARBEITEN

- Kneifel, K. (2013). *Implizite Leistungstheorien – Persönlichkeitseigenschaften als Prädiktor für beruflichen Erfolg*. Bachelorarbeit, Arbeitseinheit Arbeits- und Organisationspsychologie, Universität des Saarlandes.
- Schadow, T. (2013). *Dimensionen- und Typentest der Persönlichkeit im Vergleich: Welcher der beiden Tests ist bei Anwendern beliebter?* Bachelorarbeit, Arbeitseinheit Arbeits- und Organisationspsychologie, Universität des Saarlandes.
- Von Schwänenflug, N. (2013). *Implizite Leistungstheorien*. Bachelorarbeit, Arbeitseinheit Arbeits- und Organisationspsychologie, Universität des Saarlandes.
- Bachner, J., & Langer, Markus (2012). *Erfolgskritische Eigenschaften im Beruf*. Bachelorarbeit, Arbeitseinheit Arbeits- und Organisationspsychologie, Universität des Saarlandes.

LEHRERFAHRUNG

SS 2013	Seminar „Personalauswahl“
WS 2012/2013	Seminar „Organisationsdiagnostik“
SS 2012	Seminar „Personalauswahl“
WS 2011/2012	Seminar „Mitarbeiterbefragung“
SS 2011	Seminar „Kreativitätstechniken und Innovationsmanagement“ Seminar „Personalauswahl“