



Dunning-Kruger-Effekt in der Personalarbeit

Uwe Peter Kanning, Luisa Habacker & Luisa Kestermann

Hochschule Osnabrück

ZUSAMMENFASSUNG

Der Dunning-Kruger-Effekt (Kruger & Dunning, 1999) besagt, dass Menschen die besonders geringe Kompetenzen in einem bestimmten Bereich besitzen, ihre eigenen Kompetenzen signifikant überschätzen, während Menschen, die besonders hohe Kompetenzen aufweisen, ihre eigenen Kompetenzen signifikant unterschätzen. In zwei Experimenten wird dieser Effekt im Arbeitsfeld der Personal- und Organisationspsychologie in jeweils zwei Stichproben (Laien/innen vs. Berufspraktikern/innen) untersucht. Es zeigt sich weitestgehend der Dunning-Kruger-Effekt. Im Falle der Berufspraktiker/innen lässt sich der Effekt fast durchgängig darüber erklären, dass sich die Personen mit der geringsten Kompetenz ihrer eigenen Schwächen nicht bewusst sind (unskilled-and-unaware). Die Ergebnisse werden im Hinblick auf die vielfach beschriebene Diskrepanz zwischen Forschung und Praxis diskutiert.

Schlüsselbegriffe: Dunning-Kruger-Effekt, Forschungs-Praxis-Diskrepanz, Selbsteinschätzung

1 Einführung

Zahlreiche Publikationen beschäftigen sich mit der Diskrepanz zwischen Forschung und Praxis in der (AOW-)Psychologie (z. B. Banks et al., 2016; Kanning et al., 2010; Meyer, 2013). Viele praxisrelevante Erkenntnisse der Forschung scheinen Arbeitgebern entweder nicht bekannt zu sein und/oder werden hier nicht zur Anwendung gebracht. Die Gründe für diesen Missstand erscheinen vielfältig. Manche Publikationen fokussieren die Art der Forschung bzw. die Form, in der praxisrelevante Erkenntnisse aus der Forschung heraus kommuniziert werden (z. B. v. Rosenstiel, 2004). Andere bemängeln eine psychologieferne Aus- und Weiterbildung der Menschen, die Personalverantwortung tragen (z. B. Kanning, 2015; Kanning & Thielsch, 2015). Wieder andere kritisieren Marketingstrategien, mit denen Beratungsunternehmen fragwürdige Methoden betreiben (z. B. Kanning, 2023; Hossiep et al., 2015) oder untersuchen die Bedeutung struktureller Faktoren, die es qualifizierten Menschen im Personalwesen schwer machen, ihre Expertise in praktisches Handeln umzusetzen (König et al., 2010).

Im Folgenden werden zwei Studien vorgestellt, die das Wissen über praxisrelevante Erkenntnisse aus der Personal- und Organisationspsychologie untersuchen. Befragt werden auf der einen Seite Praktiker/innen, die sich in ihrer beruflichen Arbeit mit Themen der Personalauswahl, der Personalentwicklung oder der Führung beschäftigen. Auf der anderen Seite geht es um das Wissen von Menschen, die beruflich nichts mit entsprechenden Themen zu tun haben (Laien/innen). Dabei wird der Dunning-Kruger-Effekt (Dunning, 2011; Kruger & Dunning, 1999) untersucht. Dieser Effekt beschäftigt sich mit der Frage, wie

Menschen ihre eigenen Kompetenzen in einem bestimmten Wissensbereich einschätzen.

2 Forschungs-Praxis-Diskrepanz

Diskrepanzen zwischen den Empfehlungen der Forschung und der gelebten Praxis in Unternehmen sind in vielfältiger Weise belegt (Banks et al., 2016; Kanning, 2015; Meyer, 2013). Als Beispiel kann an dieser Stelle die nach wie vor sehr starke Verbreitung weitgehend unstrukturierter Einstellungsinterviews (Kanning, 2016) genannt werden, obwohl Metaanalysen seit Jahrzehnten zeigen, dass strukturierte Interviews um ein Vielfaches besser die Leistung im Arbeitsleben prognostizieren können (Conway et al., 1995; Huffcutt & Arthur, 1994; Huffcutt et al., 2014; Klehe & Latham, 2005; McDaniel et al., 1994; Sackett et al., 2022; Schmidt & Hunter, 1998; Wiesner & Cronshaw, 1988). Ganz ähnlich verhält es sich mit dem Einsatz des Intelligenztests. Obwohl Intelligenztests im Allgemeinen eine sehr gute prognostische Validität besitzen (Hülshager et al., 2007; Nye et al., 2022; Sackett et al., 2022; Schmidt & Hunter, 1998) und im Besonderen bei der Besetzung intellektuell anspruchsvoller Arbeitsplätze zu empfehlen sind (Salgado et al., 2003), finden sie in Deutschland bestenfalls in Auswahlverfahren für Auszubildende eine nennenswerte Verbreitung (Armoneit et al., 2020). Zudem zeigen mehrere Studien, dass berufliche Erfahrung im Bereich der Personalauswahl nicht dazu führt, dass die Verantwortlichen vor klassischen Urteilsfehlern besser geschützt sind als Personen ohne einschlägige Berufserfahrung (zusammenfassend: Kanning, 2021).

Es erscheint plausibel, dass diese und ähnlich gelagerte Probleme zunächst einmal darauf zurückzuführen sind, dass die Verantwortlichen einschlägige Forschungsergebnisse nicht kennen. Hierfür gibt es einige Belege. So zeigt beispielsweise eine Befragung von mehr als 600 Personalpraktikern/innen, dass die großen Unterschiede zwischen unstrukturierten und strukturierten Interviews hinsichtlich der prognostischen Validität in der Praxis kaum bekannt sind (Varelmann & Kanning, 2018). Ebenso werden Kriterien der Sichtung von Bewerbungsunterlagen, die bestenfalls eine fragwürdige Validität besitzen – beispielsweise die Dauer der Berufserfahrung in Jahren (Quinones et al., 1995; van Iddekinge et al., 2019) –, als sehr valide eingeschätzt (siehe auch Arnoneit et al., 2020). Eine Studie von Kanning und Thielsch (2015) zeigt zudem, dass wissenschaftliche Inhalte in der fachlichen Weiterbildung eine untergeordnete Rolle spielen. Menschen aus dem Berufsfeld präferieren eher Texte, die von anderen Praktikern/innen verfasst werden als solche, die von Personen aus der Wissenschaft stammen. Liegen in der Praxis erst einmal falsche Überzeugungen vor, so werden diese zudem durch einen Confirmation Bias aufrechterhalten und verfestigt (Kanning et al., 2023). Die Betroffenen suchen bevorzugt solche Informationen auf, die ihre eigenen Überzeugungen bestätigen und meiden konträre Informationen. Ist es nicht möglich unerwünschten Informationen aus dem Weg zu gehen, werden sie als weniger nützlich, weniger überzeugend und als unglaubwürdig abgewertet (Kanning et al., 2023).

Ein weiterer, wissensbezogener Grund für die Diskrepanz zwischen Forschung und Praxis mag darin liegen, dass viele ihr eigenes Fachwissen überschätzen und sich ihrer eigenen Schwächen in diesem Bereich nicht bewusst sind. Diese Überlegung führt zum Dunning-Kruger-Effekt (Dunning, 2011; Kruger & Dunning, 1999).

3 Dunning-Kruger-Effekt

Der Dunning-Kruger-Effekt (Dunning, 2011; Kruger & Dunning, 1999) ist eingebettet in die umfangreiche Forschung zum Thema Selbstüberschätzung. Eine Selbstüberschätzung liegt vor, wenn Menschen ihre eigenen Kompetenzen höher einschätzen als sie nachweislich sind (overestimation), sie fälschlicherweise glauben, besser zu sein als andere (overplacement) oder sie eine ungerechtfertigte Gewissheit verspüren, dass ihre eigenen Überzeugungen und Annahmen besonders zutreffend seien (overprecision; Moore & Healy, 2008; Moore & Schatz, 2017). Selbstüberschätzung ist ein weit verbreitetes Phänomen, das vielfach Gegenstand der Forschung war (z. B. Ehrlinger & Dunning, 2003; Gabriel et al., 1994; Invernizzi, 2018; Zell & Krizan, 2014) und dabei häufig als eine Möglichkeit angesehen wird, den eigenen Selbstwert zu steigern bzw. einen positiven Selbstwert aufrechtzuerhalten (Gabriel et al., 1994; Kanning, 2001; Wills, 1981; Taylor & Brown, 1988; Tajfel, 1978).

Während sehr viele Studien immer wieder zeigen konnten, dass Menschen grundlegend zur Selbstüberschätzung neigen (z. B. Gabriel et al., 1994; Invernizzi, 2018; Zell & Krizan, 2014), differenzieren Dunning und Kruger (Dunning, 2011; Kruger & Dunning, 1999) zwischen verschiedenen Kompetenzniveaus und betonen, dass nicht alle

Menschen sich in gleichem Maß überschätzen, sondern dass es auch eine Unterschätzung der eigenen Kompetenzen geben kann. Dies wiederum hängt mit den individuellen Kompetenzniveaus der untersuchten Personen zusammen. Personen mit einer sehr geringen Kompetenz neigen zu einer signifikanten Überschätzung, während sich bei Personen mit einem besonders hohen Kompetenzniveau eher eine Unterschätzung finden lässt. Dieses Phänomen, das den Kern des Dunning-Kruger-Effektes beschreibt, wurde bislang in verschiedenen Kompetenzbereichen belegt, so etwa bezogen auf die kognitive Leistungsfähigkeit, Grammatik, Humor (Kruger & Dunning, 1999), Mathematik (Mattern et al., 2010), Wissen über Waffen (Ehrlinger et al., 2008) und medizinisches Know-how (Hodges et al., 2001).

Dunning und Kruger (Dunning et al., 2003; Dunning, 2005) gehen aber noch einen Schritt weiter, indem sie das Phänomen der overestimation mit dem der overprecision verbinden. So stellen sie beispielsweise fest, dass Personen mit sehr geringer Kompetenz nicht nur eine signifikante Selbstüberschätzung zeigen (overestimation), sie sind sich darüber hinaus auch noch besonders sicher, dass ihre eigene Selbsteinschätzung zutreffend ist (overprecision; Kruger & Dunning, 1999). Beides gemeinsam trägt dazu bei, dass die Betroffenen ihre eigenen Defizite als solche nicht erkennen. Die Gruppe der Überschätzer/innen wird daher auch als „unskilled-and-unaware“ beschrieben (Dunning et al., 2003; Dunning, 2005). Dunning und Kruger (Kruger & Dunning, 1999; Dunning, 2005) sprechen in diesem Zusammenhang davon, dass Personen mit besonders geringer Kompetenz zwei Lasten zu tragen haben: Zunächst sind sie so wenig kompetent, dass sie ihre eigenen Kompetenzdefizite nicht erkennen können. Zum anderen sind sie sich in ihrer Selbstüberschätzung so sicher, dass sie keinen Grund dafür sehen, an ihren Kompetenzdefiziten arbeiten zu müssen. Werden die Betroffenen im Rahmen eines Trainings dazu gebracht, ihre eigenen Kompetenzen zu erweitern, reduziert sich die Überschätzung (Kruger & Dunning, 1999). Dies spricht dafür, dass die Betroffenen ein Mindestmaß an Kompetenz benötigen, um ihre eigenen Schwächen als solche erkennen zu können. Personen mit einer sehr hohen Kompetenzausprägung zeigen demgegenüber eine Selbstunterschätzung. Sie sind sich möglicherweise der Komplexität der Materie eher bewusst und sind daher zurückhaltender in ihrer Selbsteinschätzung (Kruger & Dunning, 1999).

Im Gegensatz zum Kern des Dunning-Kruger Effektes, der mehrfach repliziert werden konnte, gilt dies für das Phänomen „unskilled-and-unaware“ nicht. Hier gibt es widersprüchlichen Befunde. Während Dunning und Kruger (Dunning et al., 2003; Dunning, 2005; Kruger & Dunning, 1999) den Effekt finden, zeigen Studien von Händel und Fritsche (2016) sowie Miller und Geraci (2011), dass Personen mit sehr geringen Kompetenzen durchaus ein Bewusstsein für ihre Defizite haben können („unskilled-but-subjectively-aware“; Händel und Fritsche, 2016). Sie überschätzen sich zwar weiterhin signifikant, zeigen sich in ihrer Selbsteinschätzung aber unsicher. Händel und Fritsche (2016) unterscheiden in diesem Zusammenhang zwischen einer lokalen und einer globalen Sicherheitseinschätzung. Die lokale bezieht sich auf die Frage, inwieweit die Betroffenen in einem laufenden Test, Item für Item

glauben, eine Aufgabe richtig gelöst zu haben. Die globale Sicherheitseinschätzung bezieht sich auf die Frage, inwieweit sie nach dem Test davon ausgehen, dass sie ihre Gesamtleistung im Test zutreffend einschätzen. In beiden Einschätzungen zeigt sich bei Personen mit sehr geringer Kompetenz eine signifikant größere Unsicherheit als in der Gruppe der Personen mit sehr hohem Kompetenzniveau (Händel & Fritsche, 2016).

4 Hypothesen

Vor dem Hintergrund der bisherigen Befunde zum Kern des Dunning-Kruger Effektes (Dunning et al., 2003; Ehrlinger et al., 2008; Hodges et al., 2001; Kruger & Dunning, 1999; Mattern et al., 2010) werden zwei Hypothesen formuliert.

Hypothese 1: Personen, mit einer sehr geringen Kompetenzausprägung im Bereich des Wissens um personal- und organisationspsychologische Erkenntnisse, überschätzen ihre eigene Leistung signifikant.

Hypothese 2: Personen, mit einer sehr hohen Kompetenzausprägung im Bereich des Wissens um personal- und organisationspsychologische Erkenntnisse, unterschätzen ihre eigene Leistung signifikant.

Da die Befundlage im Hinblick auf den „unskilled-and-unaware“-Effekt sehr gering und zudem widersprüchlich ausfällt (Kruger & Dunning, 1999; Händel & Bukowski, 2019; Miller & Geraci, 2011), wird an dieser Stelle keine Hypothese, sondern eine explorative Frage formuliert.

Frage 1: Liegt in der Gruppe, der geringsten Kompetenzausprägung, eine größere Unsicherheit im Hinblick auf die Einschätzung der eigenen Leistung vor als in der Gruppe mit der höchsten Kompetenzausprägung?

Darüber hinaus gehen wir explorativ der Frage nach, inwieweit Praktiker/innen und Laien/innen sich signifikant in ihrer Kompetenzausprägung, ihrer Selbsteinschätzung sowie ihrer Sicherheit in der Selbsteinschätzung unterscheiden.

Frage 2: Unterscheiden sich Praktiker/innen in ihrer Kompetenzausprägung, ihrer Selbsteinschätzung und ihrer Sicherheit bei der Selbsteinschätzung von Laien/innen?

5 Studie 1: Personal- und organisationspsychologisches Wissen

Studie 1 untersucht den Dunning-Kruger-Effekt in Bezug auf personal- und organisations-psychologisches Wissen.

5.1 Methode

Stichprobe: An der Online-Studie beteiligten sich 392 Personen, von denen sich 131 Personen beruflich mit Fragen der Personal- und Organisationspsychologie beschäftigten und 261 nicht einschlägig beruflich beschäftigt waren. Erstere werden im Folgenden als „Praktiker/innen“, letztere als „Laien/innen“ bezeichnet. Die Befragten waren zu 54.6 % weiblichen und zu 45.4 % männlichen Geschlechts. Das Durchschnittsalter lag bei 36.63 Jahren (Stdv = 12.56; Range von 18 – 65 Jahren), wobei Perso-

nen mit einem höheren Bildungsgrad in der Stichprobe besonders stark vertreten waren (1.8 % Hauptschulabschluss, 1.2 % Realschulabschluss, 34.7 % (Fach-)Abitur, 20.7 % Bachelor-Abschluss, 28.3 % Master- oder Diplom-Abschluss, 3.3 % Promotion).

Datenerhebung: Die Anwerbung der Befragten erfolgte über soziale Netzwerke (Facebook, Xing, LinkedIn), Schneeballsystem und Aushänge an einer Hochschule. Zur Anwerbung von Praktikern/innen wurde der Link zur Umfrage insbesondere in Gruppen platziert, in denen sich Menschen austauschten, die im Personalwesen tätig waren. Die Teilnahme an der Studie war freiwillig, anonym und unentgeltlich. Interessierte Personen wurden nach Beendigung des Fragebogens über die richtigen Antworten zu den 20 Wissensfragen informiert. Drei Personen gaben bei der Abschlussfrage an, dass sie nicht wollten, dass ihre Daten in die Studie einfließen. Sie wurden nach der Datenerhebung aus dem Datensatz entfernt und flossen nicht in die Stichprobe ein.

Untersuchungsdesign & Messinstrument: Die Studie folgte einem quasi-experimentellen Untersuchungsdesign. Auf der Basis der Anzahl richtig beantworteter Wissensfragen wurde die Gesamtstichprobe in vier Kompetenzgruppen eingeteilt. Die vorliegende Studie folgte in diesem Punkt den Vorgängerstudien zum Dunning-Kruger-Effekt. Als Orientierung diente dabei die Einteilung in Quartile (Kruger & Dunning, 1999). Da sich in den vorliegenden Daten allerdings über die Anzahl der richtig beantworteten Fragen keine vier gleichgroßen Gruppen mit jeweils 25 % der Befragten ergaben, wurde eine abweichende Einteilung gewählt. Im Folgenden wird daher auch nicht von Kompetenzquartilen sondern von Kompetenzgruppen gesprochen. Gruppe 1 umfasste die Personen mit der geringsten Kompetenz (0 bis 9 richtig gelöste Aufgaben = 19.4 % der Stichprobe), Gruppe 2 repräsentierte die nächst höher qualifizierte Gruppe (10 bis 11 gelöste Aufgaben = 24.5 %) gefolgt von Gruppe 3 (12 bis 13 gelöste Aufgaben = 32.1 %) und schließlich Gruppe 4 (14 bis 20 gelöste Aufgaben = 24 %). Die Zuordnung zu den vier Gruppen bildete später die unabhängige Variable im Untersuchungsdesign.

Der Online-Fragebogen gliederte sich in drei Blöcke. Der erste Block umfasste 20 Aufgaben mit deren Hilfe das Ausmaß einschlägigen Fachwissens erfasst wurde. Jede Aufgabe bestand darin, eine vorgegebene Aussage dahingehend einzuschätzen, ob sie zutreffend war oder nicht (Auswahlkategorien: „falsch“ vs. „richtig“). Die Auswahl der Fragen erfolgte durch die Autor/innen vor dem Hintergrund einschlägiger Forschungsergebnisse aus der Psychologie. Die Liste aller Aufgaben inklusive Lösungsschlüssel findet sich in Anhang 1. Unmittelbar nach der Bearbeitung jeder Aufgabe sollten die Befragten auf einer fünfstufigen Skala einschätzen, wie sicher sie sich bei ihrer Entscheidung waren (1 = „sehr unsicher“, 5 = „sehr sicher“). Der Mittelwert dieser Einzeleinschätzungen wird im Folgenden „lokale Sicherheitseinschätzung“ (Händel & Fritsche, 2016) bezeichnet.

Im zweiten Block wurden die Befragten gebeten, die Anzahl der von ihnen richtig gelösten Aufgaben einzuschätzen (= Leistungseinschätzung). Es folgte auch hier eine

Einschätzung der Sicherheit, mit der diese Schätzung vorgenommen wurden (fünfstufige Skala, s.o.). Diese Variable wird im Folgenden als „globale Sicherheitseinschätzung“ bezeichnet (Händel & Fritsche, 2016).

Im dritten Block folgten Fragen zur Demographie: Geschlecht (männlich, weiblich, divers), Alter in Jahren (offene Angabe) und Bildung (kein Abschluss, Hauptschulabschluss, Realschulabschluss, (Fach-)Abitur, Bachelor-Abschluss, Master-Abschluss, Promotion). Den Abschluss

bildete ein Item, mit dem die Befragten ihre Daten für die Auswertung freigeben konnten.

5.2 Ergebnisse

Die Überprüfung der Hypothesen erfolgte getrennt für die Gruppe der Laien/innen und die Gruppe der Praktiker/innen. Die Mittelwerte und Standardabweichungen aller Variablen, sowie die Interkorrelationen finden sich in Tabelle 1.

Tabelle 1: Korrelation und Deskriptivstatistiken sämtlicher Variablen aus Studie 1

	1	2	3	4	5	6	7	M (Stdv.) Laien/innen
1. Geschlecht ^a		-.21	-.20	-.04	-.21	-.20	-.17	1.58 (.49)
2. Alter	-.24		.24	.03	-.28	.26	.27	35.39 (12.75)
3. Bildung ^b	.01	.30		-.17	.18	.05	.10	3.59 (1.09)
4. Leistung	-.19	.13	-.10		.04	.13	.05	11.80 (2.38)
5. eingeschätzte Leistung	-.19	.12	-.05	.18*		.61	.44	12.42 (3.36)
6. lokale Sicherheit bei Einschätzung	-.20	.18	-.04	.07	.54		.53	3.65 (.47)
7. globale Sicherheit bei Einschätzung	-.14	.19	.05	.09	.48	.53		3.35 (.85)
M (Stdv.) Praktiker/innen	1.47 (.50)	39.08 (11.87)	3.99 (1.19)	11.63 (2.66)	13.98 (3.37)	4.01 (.42)	3.75 (.72)	

Erläuterung: dargestellt werden Produkt-Moment-Korrelationen nach Pearson sowie das arithmetische Mittel und in Klammern die Standardabweichung; Ergebnisse oberhalb der Diagonale beziehen sich auf die Teilstichprobe der Laien/innen und unterhalb der Diagonale auf die Teilstichprobe der Praktiker/innen; ^amännlich = 1, weiblich = 2; ^bHauptschulabschluss = 1, Realschulabschluss = 2, (Fach-)Hochschulreife = 3, Bachelor = 4, Master/Diplom = 5, Promotion = 6; * $p \leq .05$ ** $p \leq .01$ *** $p \leq .001$

In einem ersten Schritt wurde in der *Stichprobe der Laien/innen* untersucht, inwieweit sich Personen aus den vier Kompetenzgruppen in ihrer Leistung bzw. in ihrer selbst eingeschätzten Leistung voneinander unterscheiden. Hierzu wurde eine multivariate Varianzanalyse durchgeführt. Die vier Kompetenzgruppen bildeten dabei die Stufen der unabhängigen Variable, die tatsächliche Leistung und die Selbsteinschätzung der Leistung die beiden abhängigen Variablen. Es ergab sich ein signifikanter, multivariater Haupteffekt ($F_{(6/514)} = 71.51$ $p \leq .001$ $h^2 = .46$). Univariat ergab sich jedoch nur bezogen auf die tatsächliche Leistung ein signifikanter Unterschied ($F_{(3/258)} = 735.78$ $p \leq .001$ $h^2 = .90$). Alle vier Gruppen unterschieden sich hinsichtlich der tatsächlichen Leistung signifikant voneinander jedoch nicht im Hinblick auf die Selbsteinschätzung der Leistung (Tab. 2). Die tatsächliche Leistung stieg von Gruppe 1 zu Gruppe 4 jeweils signifikant an.

In einem zweiten Schritt wurde überprüft, ob sich innerhalb der vier Kompetenzgruppen die tatsächliche Leistung, von der selbst eingeschätzten Leistung unterschied. Zum Einsatz kamen dabei t-Tests für gepaarte Stichproben. Dabei ergaben sich signifikante Unterschiede für die Gruppen 1 ($t_{(45)} = -6.54$ $p \leq .001$), 2 ($t_{(72)} = -5.79$ $p \leq .001$)

und 4 ($t_{(64)} = 7.76$ $p \leq .001$). Bei den Gruppen 1 und 2 lag eine signifikante Überschätzung der eigenen Leistung und bei Gruppe 4 eine signifikante Unterschätzung der eigenen Leistung vor (Abb. 1, obere Hälfte). Das Ergebnismuster bestätigte die Hypothesen 1 und 2.

In einem letzten Schritt wurde überprüft, inwieweit sich die vier Gruppen dahingehend unterschieden, wie sicher sie sich bei der Einschätzung der eigenen Leistung waren. Hierzu standen zwei Operationalisierungen zur Verfügung. Zum einen der Mittelwert über die Sicherheitseinschätzung nach jeder Leistungsaufgabe (= „lokale Sicherheit bei der Schätzung“), zum anderen die globale Sicherheitseinschätzung nachdem alle Aufgaben bearbeitet wurden. Berechnet wurde eine multivariate Varianzanalyse. Für die Stichprobe der Laien/innen zeigte sich im Ergebnis kein Effekt der unabhängigen Variable ($F_{(6/514)} = 1.00$ $p < .05$ $h^2 = .01$). Die vier Leistungsgruppen unterschieden sich nicht hinsichtlich der Sicherheit, mit der sie ihre eigene Leistung einschätzten (Tab. 2). Frage 1 musste somit für die Stichprobe der Laien/innen verneint werden.

Tabelle 2: Ergebnisse Studie 1

	Kompetenzgruppen			
	1 (n=45) (n=31)	2 (n=72) (n=24)	3 (n=80) (n=46)	4 (n=64) (n=30)
Anzahl gelöster Items	8.22 ^a (.97)	10.54 ^b (.50)	12.55 ^c (.50)	14.81 ^d (1.08)

	7.94 ^a (1.24)	10.54 ^b (.51)	12.57 ^c (.50)	14.90 ^d (1.18)
Schätzung der Anzahl gelöster Items	11.80 (3.53) 12.90 ^a (3.75)	12.99 (3.53) 14.75 (2.67)	12.30 (3.65) 13.52 (3.56)	12.37 (2.55) 15.17 ^b (2.74)
lokale Sicherheit bei der Schätzung	3.61 (.44) 3.95 (.49)	3.58 (.43) 4.06 (.33)	3.65 (.50) 4.02 (.46)	3.74 (.47) 4.00 (.34)
globale Sicherheit bei der Schätzung	3.27 (.92) 3.65 (.84)	3.37 (.90) 3.75 (.74)	3.36 (.92) 3.72 (.72)	3.38 (.68) 3.90 (.54)

Erläuterung: dargestellt werden Mittelwerte und in Klammern Standardabweichungen; die obere Zeile bezieht sich auf die Teilstichprobe der Laien/innen, die untere auf die Teilstichprobe der Praktiker/innen; Die Mittelwerte unterscheiden sich zeilenweise, wenn sie unterschiedliche Indizes aufweisen $p \leq .05$; Gruppe 1 umfasst Personen mit der geringsten und Gruppe 4 Personen mit der höchsten Leistung.

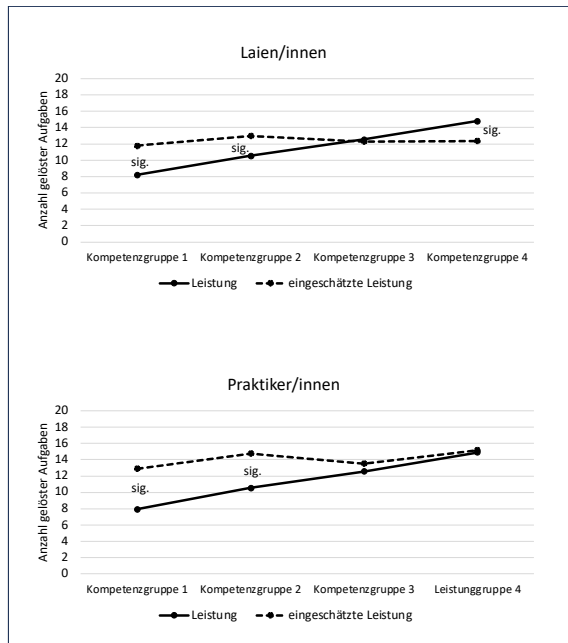


Abbildung 1: Dunning-Kruger-Effekt in Studie 1

Die zuvor beschriebene Analyse wurde in gleicher Weise auch für die *Stichprobe der Praktiker/innen* durchgeführt. Zunächst konnte gefunden werden, dass sich die vier Kompetenzgruppen hinsichtlich des tatsächlichen Leistungsniveaus und der selbst eingeschätzten Leistung multivariat signifikant voneinander unterscheiden ($F_{(6/254)} = 36.35 p \leq .001 h^2=.46$). Bezogen auf die tatsächliche Leistung ergab sich für alle Einzelvergleiche zwischen den Gruppen ein signifikanter univariater Effekt ($F_{(3/128)} = 331,29 p \leq .001 h^2=.89$; Tab. 2). Hinsichtlich der selbst eingeschätzten Leistung unterschieden sich lediglich die Gruppen 1 und 4 voneinander ($F_{(3/128)} = 3,15 p \leq .05 h^2=.07$). In der leistungsstärksten Gruppe wurde die eigene Leistung signifikant höher eingeschätzt als in der leistungsschwächsten Gruppe (Tab. 2).

Die Überprüfung des Dunning-Kruger-Effektes ergab signifikante Unterschiede zwischen Leistung und Selbsteinschätzung für die Leistungsgruppe 1 und 2 ($t_{(31)} = -6.95 p \leq .001 / (t_{(24)} = -7.50 p \leq .001)$). Personen mit sehr geringer bzw. geringer Kompetenz überschätzten ihre Leistung demnach signifikant. Ersteres stand in Übereinstimmung zu Hypothese 1. Die beiden übrigen Gruppen wiesen eine realistische Einschätzung ihrer Leistung auf (Abb. 1, untere Hälfte). Der Befund zu Gruppe 4 stand im Widerspruch zu Hypothese 2.

Bei einem Vergleich der vier Gruppen in Bezug auf die lokale bzw. globale Sicherheit bei der Einschätzung der eigenen Leistung zeigten sich keine signifikanten Unter-

schiede ($F_{(6/254)} = .57 p > .05 h^2=.01$; Tab. 2). Auch für die Gruppe der Praktiker/innen musste Frage 1 somit verneint werden.

Abschließend wurde mit Hilfe einer multivariaten Varianzanalyse explorativ untersucht, ob sich Laien/innen und Praktiker/innen in den vier abhängigen Variablen signifikant unterschieden (Frage 2). Multivariat ergab sich dabei ein signifikanter Haupteffekt ($F_{(4/387)} = 14.32 p \leq .001 h^2=.13$). Univariat unterschieden sich die beiden Gruppen nicht hinsichtlich der tatsächlichen Leistung ($M_{\text{Laien/innen}}=11.80 SD=2.38$; $M_{\text{Praktiker/innen}}=11.63 SD=2.66$; $F_{(1/391)} = .42 p > .05 h^2=.0$), wohl aber im Hinblick auf alle übrigen Variablen. Praktiker/innen schätzen demnach ihre Leistung signifikant höher ein ($M_{\text{Laien/innen}}=12.42 SD=3.36$; $M_{\text{Praktiker/innen}}=13.98 SD=3.37$; $F_{(1/391)} = 18.66 p \leq .001 h^2=.05$). Gleichzeitig erleben sie mehr Sicherheit bei der Einschätzung der eigenen Leistung. Dies gilt sowohl für die lokale Sicherheitseinschätzung ($M_{\text{Laien/innen}}=3.65 SD=.47$; $M_{\text{Praktiker/innen}}=4.01 SD=.42$; $F_{(1/391)} = 55.02 p \leq .001 h^2=.12$) als auch für die globale Sicherheitseinschätzung ($M_{\text{Laien/innen}}=3.35 SD=.85$; $M_{\text{Praktiker/innen}}=3.75 SD=.72$; $F_{(1/391)} = 20.77 p \leq .001 h^2=.05$).

5.3 Diskussion

Studie 1 bestätigte sowohl für die Gruppe der Laien/innen als auch für die Gruppe der Praktiker/innen weitestgehend den Dunning-Kruger-Effekt (Dunning et al., 2011; Kruger & Dunning, 1999). Demnach überschätzten Personen in der schwächsten Kompetenzgruppe ihre Leistung signifikant (Laien/innen und Praktiker/innen), während Personen aus der Gruppe der Praktiker/innen in der stärksten Kompetenzgruppe ihre Leistung signifikant unterschätzten (Hypothese 1 und 2). Im Gegensatz zu den Ergebnissen von Händel und Fritsche (2016) konnte nicht belegt werden, dass Personen mit der geringsten Kompetenz in ihrer Selbsteinschätzung mehr Unsicherheit verspürten als Personen in der stärksten Kompetenzgruppe. Dies spricht für das Phänomen „unskilled-and-unaware“, wie es ursprünglich von Kruger und Dunning (1999) formuliert wurde. Da dies auch in der Gruppe der Praktiker/innen zu beobachten ist, deutet sich hier eine Teilerklärung für mangelnde Anwendung wissenschaftlicher Erkenntnisse in der Praxis an: Ein Teil der handelnden Menschen – in der vorliegenden Studie sind dies mehr als 40 % der befragten Praktiker/innen (Gruppe 1 + Gruppe 2) – zeigen ein eher niedriges Fachwissen und sind sich dessen kaum bewusst. Folgerichtig liegt auch kein Grund vor, warum sie

sich entsprechend weiterqualifizieren sollten, denn aus ihrer subjektiven Sicht liegt kein Defizit vor.

Bedenklich stimmt, dass Praktiker/innen keine höhere Leistung erbrachten als Laien/innen. Gleichwohl schätzten sie ihre Leistung signifikant höher ein und waren sich auch signifikant sicherer, dass ihre Einschätzung zutreffend sei. Insofern erlagen in dieser Studie die untersuchten Praktiker/innen einer größeren selbstwertdienlichen Verzerrung als die untersuchten Laien/innen. Unter dem Gesichtspunkt der praktischen Relevanz des abgefragten Wissens ist dieses Ergebnis betrüblich, weil es die befragten Praktiker/innen neben dem Dunning-Kruger-Effekt daran hindert, eigene Defizite aktiv anzugehen.

Da die Ergebnisse der Studien davon abhängen, welche Fachfragen gestellt werden, sollen die Befunde in einer zweiten Studie mit anderen Fachfragen zusätzlich abgesichert werden.

6 Studie 2: Wissen zu Personalauswahlmethoden

Studie 2 untersucht den Dunning-Kruger-Effekt in Bezug auf Wissen zum Thema Personalauswahl.

6.1 Methode

Stichprobe: An der Online-Studie beteiligten sich 658 Personen. Hiervon beschäftigten sich 265 Personen beruflich mit Fragen der Personalauswahl (= „Praktiker/innen“), während für weitere 393 Personen dies nicht galt (= „Laien/innen“). Das Alter der Befragten variierte zwischen 18 und 65 Jahren, mit einem Durchschnittsalter von 31.55 Jahren (*Stdv* = 12.56). 68.7 % waren Frauen und 31.3 % Männer. Auch in der zweiten Studie stellten Menschen mit hohem Bildungsgrad den Großteil der Stichprobe dar (0.6 % Hauptschulabschluss, 5.8 % Realschulabschluss, 48.3 (Fach-)Abitur, 44.4 % Bachelor-, Master- oder Diplom-Abschluss, 0.9 % Promotion).

Datenerhebung: Die Anwerbung der Befragten erfolgte vergleichbar zu Studie 1 insbesondere über soziale Netzwerke. Personen mit potentieller Berufserfahrung in der Personalauswahl wurden über spezifische Gruppen in Xing und LinkedIn angesprochen. Personen ohne potentielle Berufserfahrung erhielten Einladungen über WhatsApp und Instagram und wurden zusätzlich über Aushänge an einer Hochschule auf die Studie aufmerksam gemacht. Die Teilnahme an der Studie war freiwillig, anonym und unentgeltlich. Drei Personen wurden vor der abschließenden Stichprobenbildung aus dem Datensatz entfernt, weil sie bei der Abschlussfrage angaben, dass sie ihre Daten nicht für die Studie freigeben wollten. Sie waren in der oben beschriebenen Stichprobe daher nicht enthalten.

Untersuchungsdesign & Messinstrument: Auch Studie 2 folgte einem quasi-experimentellen Untersuchungsdesign in dem vier Kompetenzgruppen miteinander verglichen

wurden. In Abhängigkeit von der Anzahl der richtig beantworteten Wissensfragen wurden die Befragten in vier Gruppen eingeteilt. Als Orientierung diente erneut die Einteilung in Quartile, wie in den Studien von Kruger und Dunning (1999). Wie in Studie 1 waren die Gruppen jedoch nicht im Sinne von Quartilen gleich groß. Die Leistungsgruppen wurden daher wie folgt gebildet: In Gruppe 1 befanden sich die Personen mit der geringsten Kompetenz (0 bis 9 richtig gelöste Aufgaben = 19.9 % der Stichprobe), Gruppe 2 umfasste Personen aus der nächsthöheren Kompetenzgruppe (10 bis 11 gelöste Aufgaben = 22.9 %), gefolgt von Gruppe 3 (12 bis 14 gelöste Aufgaben = 32.8 %) und schließlich Gruppe 4 (15 bis 20 gelöste Aufgaben = 24,4 %).

Der Online-Fragebogen umfasste wieder drei Blöcke. Der erste Block beinhaltete 20 Aufgaben zur Erfassung des Fachwissens im Bereich der Personalauswahl. Die einzelne Aufgabe bestand in der Einschätzung, ob eine bestimmte Aussage zutreffend war oder nicht (Items, siehe Anhang 2). Direkt nach der Bearbeitung einer Aufgabe mussten die Befragten auf einer fünfstufigen Skala angeben, wie sicher sie sich bei ihrer Antwort waren (1 = „sehr unsicher“, 5 = „sehr sicher“). Über den Mittelwert dieser Schätzungen wurde später die Variable „lokale Sicherheitseinschätzung“ (Händel & Fritsche, 2016) berechnet.

Der zweite Block bezog sich auf die Selbsteinschätzung der eigenen Leistung. Zunächst wurden die Befragten gebeten, einzuschätzen, wie viele der 20 Aufgaben sie richtig gelöst hatten (= „Leistungseinschätzung“). Anschließend sollten sie auf einer fünfstufigen Skala (s.o.) einschätzen, wie sicher sie sich bei dieser Einschätzung waren (= „globale Sicherheitseinschätzung“; Händel & Fritsche, 2016).

Im dritten Block folgten Fragen zur Demographie: Geschlecht (männlich, weiblich, divers), Alter in Jahren (offene Angabe) und Bildung (kein Abschluss, Hauptschulabschluss, Realschulabschluss, (Fach-)Abitur, Bachelor-Abschluss, Master-Abschluss, Promotion). Den Abschluss bildete ein Item, mit dem die Befragten ihre Daten für die Auswertung freigeben konnten.

6.2 Ergebnisse

Die Datenanalyse erfolgte wieder getrennt für Personen mit und ohne berufliche Erfahrung im Bereich der Personalauswahl. In Tabelle 3 finden sich die Mittelwerte, Standardabweichungen und Interkorrelationen der Variablen.

Tabelle 3: Korrelation und Deskriptivstatistiken sämtlicher Variablen aus Studie 2

	1	2	3	4	5	6	7	<i>M (Stdv.) Laien/innen</i>
1. Geschlecht ^a		-.01	.02	.02	-.19	-.09	-.10	1.73 (.45)
2. Alter	-.27		-.07	-.06	-.05	.06	.05	26.95

3. Bildung ^b	.00	-.01		.05	.02	-.03	.01	(9.89)
4. Leistung	-.10	-.05	.18		.38	.45	.30	3.22 (.56)
5. eingeschätzte Leistung	-.08	.12*	-.04	.28		.67	.55	12.13 (3.46)
6. lokale Sicherheit bei Einschätzung	-.14	.14*	-.02	.31	.58		.50	11.64 (4.13)
7. globale Sicherheit bei Einschätzung	-.14	.15*	-.04	.16	.50	.60		3.72 (.56)
M (Stdv.) Praktiker/innen	1.63 (.48)	38.37 (11.52)	3.65 (.66)	13.02 (3.26)	13.96 (3.34)	4.04 (.48)	3.67 (.75)	3.16 (.93)

Erläuterung: dargestellt werden Produkt-Moment-Korrelationen nach Pearson sowie das arithmetische Mittel und in Klammern die Standardabweichung; Ergebnisse oberhalb der Diagonale beziehen sich auf die Teilstichprobe der Laien/innen und unterhalb der Diagonale auf die Teilstichprobe der Praktiker/innen; ^amännlich = 1, weiblich = 2; ^bHauptschulabschluss = 1, Realschulabschluss = 2, (Fach-)Hochschulreife = 3, Hochschulabschluss = 4, Promotion = 5; * $p \leq .05$ ** $p \leq .01$ *** $p \leq .001$

Analog zur ersten Studie wurde zunächst für die Gruppe der Laien/innen mit Hilfe einer multivariaten Varianzanalyse überprüft, ob sich die tatsächliche Leistung sowie die selbst eingeschätzte Leistung in den vier Kompetenzgruppen voneinander unterschieden. Die unabhängige Variable bildeten die vier Kompetenzgruppen, die beiden abhängigen Variablen die tatsächliche Leistung und die selbst eingeschätzte Leistung. Es ergab sich ein signifikanter multivariater Effekt ($F_{(6/778)} = 120.09 p \leq .001 h^2 = .48$) und darüber hinaus sowohl für die tatsächliche Leistung ($F_{(3/389)} = 1400.23 p \leq .001 h^2 = .90$) als auch für die selbst eingeschätzte Leistung ($F_{(3/389)} = 28.84 p \leq .001 h^2 = .19$) ein signifikanter univariater Effekt. Hinsichtlich der tatsächlichen Leistung unterschieden sich alle vier Gruppen signifikant voneinander. Die tatsächliche Leistung stieg von Gruppe 1 zu Gruppe 4 jeweils signifikant an (Tab. 4). Hinsichtlich der selbst eingeschätzten Leistung zeigte sich lediglich bei Gruppe 4 ein signifikanter Unterschied zu allen übrigen Gruppen. Personen aus Gruppe 4 schätzten ihre Leistung signifikant höher ein (Tab. 4). Innerhalb jeder Gruppe wurde in einem weiteren Berechnungsschritt mit Hilfe von t-Tests für gepaarte Stichproben untersucht, ob sich die tatsächliche Leistung, von der selbst eingeschätzten Leistung unterschied. Dies war in drei von vier Gruppen der Fall (Abb. 2, obere Hälfte). In der Gruppe mit der geringsten Leistung (Gruppe 1) konnte in Übereinstimmung mit Hypothese 1 eine signifikante Überschätzung der eigenen Leistung belegt werden ($t_{(91)} = -6.20 p \leq .001$). In den Gruppen 3 und 4 fand sich hingegen eine signifikante Unterschätzung der eigenen Leistung ($t_{(125)} = 6.45 p \leq .001$ bzw. $t_{(78)} = 6.98 p \leq .001$). Das Ergebnis für Gruppe 4 stand in Einklang mit Hypothese 2.

Die Analyse zum Vergleich der vier Gruppen im Hinblick auf ihre Sicherheit bei der Einschätzung der eigenen Leistung erfolgte wieder mit Hilfe einer multivariaten Varianzanalyse. Als unabhängige Variable wurden die vier Leistungsgruppen und als abhängige Variablen die lokale sowie die globale Einschätzung der Sicherheit eingesetzt. Es zeigte sich zunächst ein signifikanter multivariater Effekt ($F_{(6/778)} = 18.18 p \leq .001 h^2 = .12$) sowie jeweils ein signifikanter univariater Effekt in Bezug auf die lokale Sicherheit ($F_{(3/389)} = 41.09 p \leq .001 h^2 = .24$) sowie die globale Sicherheit ($F_{(3/389)} = 13.65 p \leq .001 h^2 = .10$) der eigenen Einschätzung (Tab. 4). In Bezug auf Frage 1 zeigten Laien/innen aus der leistungsstärksten Gruppe mehr Sicherheit bei der Einschätzung ihrer Leistung, und zwar sowohl im Hinblick auf die lokale als auch im Hinblick auf die globale Sicherheit.

Tabelle 4: Ergebnisse Studie 2

	Kompetenzgruppen			
	1 (n=91) (n=40)	2 (n=97) (n=54)	3 (n=126) (n=90)	4 (n=79) (n=81)
Anzahl gelöster Items	8.12 ^a (1.01) 8.43 ^a (.98)	10.56 ^b (.50) 10.57 ^b (.50)	12.76 ^c (.83) 12.88 ^c (.83)	17.70 ^d (1.90) 17.06 ^d (1.61)
Schätzung der Anzahl gelöster Items	10.81 ^a (3.90) 12.45 ^a (2.80)	10.85 ^a (3.75) 13.74 (3.07)	10.65 ^a (3.53) 13.87 (3.80)	15.14 ^b (3.89) 14.95 ^b (2.92)
lokale Sicherheit bei der Schätzung	3.58 ^a (.51) 3.80 ^a (.41)	3.60 ^a (.50) 3.97 ^a (.48)	3.57 ^a (.47) 4.01 ^a (.50)	4.27 ^b (.50) 4.24 ^b (.43)
globale Sicherheit bei der Schätzung	3.03 ^a (.92) 3.60 (.67)	3.03 ^a (.86) 3.59 (.79)	3.00 ^a (.89) 3.58 (.81)	3.73 ^b (.87) 3.85 (.65)

Erläuterung: dargestellt werden Mittelwerte und in Klammern Standardabweichungen; die obere Zeile bezieht sich auf die Teilstichprobe der Laien/innen, die untere auf die Teilstichprobe der Praktiker/innen; Die Mittelwerte unterscheiden sich zeilenweise, wenn sie unterschiedliche Indizes aufweisen $p \leq .05$; Gruppe 1 umfasst Personen mit der geringsten und Gruppe 4 Personen mit der höchsten Leistung.

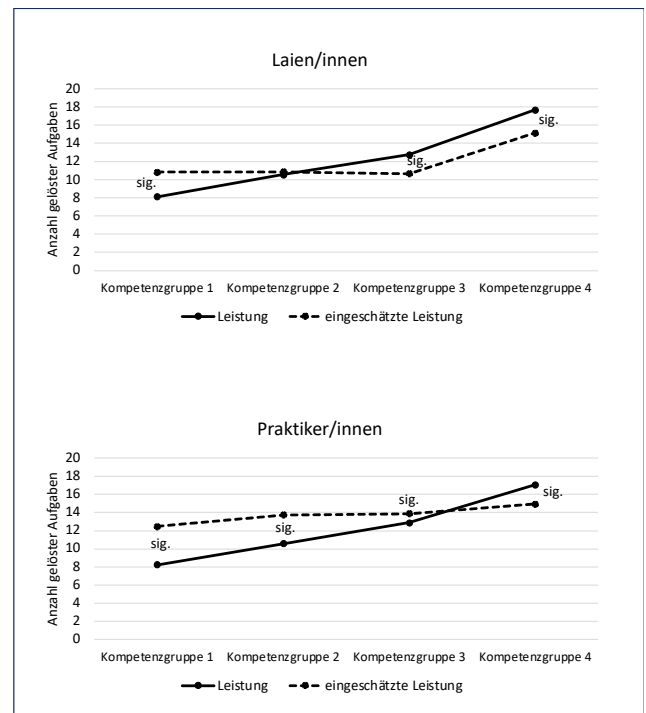


Abbildung 2: Dunning-Kruger-Effekt in Studie 2

In gleicher Weise wurde bei der Analyse der Daten der *Praktiker/innen* verfahren. Die multivariate Varianzanalyse zum Vergleich der vier Kompetenzgruppen im Hinblick auf die tatsächlich sowie die selbst eingeschätzte Leistung ergab einen signifikanten multivariaten Effekt ($F_{(6/522)} = 70.13$ $p \leq .001$ $h^2 = .45$). Beide univariaten Tests bezogen auf die Leistung bzw. die eingeschätzte Leistung waren ebenfalls signifikant ($F_{(3/261)} = 679.11$ $p \leq .001$ $h^2 = .89$ bzw. $F_{(3/261)} = 5.46$ $p \leq .01$ $h^2 = .06$). Alle vier Kompetenzgruppen unterschieden sich in ihrer tatsächlichen Leistung signifikant voneinander, wobei die Leistung von Gruppe 1 zu Gruppe 4 jeweils signifikant anstieg (Tab. 4). Im Hinblick auf die selbst eingeschätzte Leistung konnte lediglich ein signifikanter Unterschied zwischen Gruppe 1 und Gruppe 4 gefunden werden. Personen aus Gruppe 1 schätzten demnach ihre Leistung signifikant niedriger ein als Personen aus Gruppe 4. Zur Überprüfung der Hypothesen 1 und 2 wurde innerhalb jeder Leistungsgruppe mit Hilfe von t-Tests ein Vergleich zwischen tatsächlicher und eingeschätzter Leistung vorgenommen. In allen vier Fällen zeigten sich signifikante Effekte (Abb. 2, untere Hälfte). In den Kompetenzgruppen 1, 2 und 3 konnten jeweils eine Überschätzung der eigenen Leistung belegt werden ($t_{(39)} = -9.19$ $p \leq .001$ bzw. $t_{(53)} = -7.80$ $p \leq .001$ bzw. $t_{(89)} = -2.40$ $p \leq .05$). Die signifikante Überschätzung in Gruppe 1 bestätigt Hypothese 1. In der leistungsstärksten Gruppe fand sich hingegen eine signifikante Unterschätzung der eigenen Leistung ($t_{(80)} = 6.90$ $p \leq .001$). Dies bestätigt Hypothese 2.

Der Vergleich zwischen den vier Kompetenzgruppen im Hinblick auf die lokale und globale Sicherheit bei der Einschätzung der eigenen Leistung ergab einen multivariaten Effekt ($F_{(6/522)} = 5.17$ $p \leq .001$ $h^2 = .06$) und einen signifikanten univariaten Effekt bezogen auf die lokale Sicherheit ($F_{(3/261)} = 9.46$ $p \leq .001$ $h^2 = .10$). Dabei unterschied sich die Gruppe 4 signifikant von allen anderen Gruppen. Sie wies eine signifikant höhere lokale Sicherheit auf (Tab. 4). Frage 1 musste in der Gruppe der *Praktiker/innen* somit in Bezug auf die lokale Sicherheit, nicht aber in Bezug auf die globale Sicherheit mit ja beantwortet werden.

Mit Hilfe einer abschließenden multivariaten Varianzanalyse wurde untersucht, ob sich die Gruppe der *Laien/innen* hinsichtlich der vier abhängigen Variablen (tatsächliche Leistung, eingeschätzte Leistung, lokale und globale Sicherheit bei der Leistungseinschätzung) von der Gruppe der *Praktiker/innen* signifikant unterschied (Frage 2). Im Ergebnis zeigte sich ein multivariater Haupteffekt ($F_{(4/653)} = 19.73$ $p \leq .001$ $h^2 = .11$). Univariat ergaben sich für alle vier abhängigen Variablen Unterschiede zwischen beiden Gruppen. *Praktiker/innen* lösten im Wissenstest signifikant mehr Aufgaben richtig als *Laien/innen* ($M_{\text{Laien/innen}} = 12.13$ $SD = 3.46$; $M_{\text{Praktiker/innen}} = 13.02$ $SD = 3.26$; $F_{(1/656)} = 10.73$ $p \leq .01$ $h^2 = .02$). Sie schätzten ihre Leistung signifikant höher ein ($M_{\text{Laien/innen}} = 11.64$ $SD = 4.13$; $M_{\text{Praktiker/innen}} = 13.96$ $SD = 3.34$; $F_{(1/656)} = 58.09$ $p \leq .001$ $h^2 = .08$). Zudem wiesen sie eine höhere Sicherheit auf, wenn es um die Einschätzung ihrer Leistung ging. Dies galt sowohl für die lokale Sicherheit ($M_{\text{Laien/innen}} = 3.72$ $SD = 0.56$; $M_{\text{Praktiker/innen}} = 4.04$ $SD = .48$; $F_{(1/656)} = 56.70$ $p \leq .001$ $h^2 = .08$), als auch für die globale Sicherheit ($M_{\text{Laien/innen}} = 3.16$

$SD = .93$; $M_{\text{Praktiker/innen}} = 3.67$ $SD = .75$; $F_{(1/656)} = 54.046$ $p \leq .001$ $h^2 = .08$).

6.3 Diskussion

In Studie 2 konnte der Dunning-Kruger-Effekt (Kruger & Dunning, 1999) vollständig bestätigt werden. Sowohl in der Stichprobe der *Laien/innen* als auch in der Stichprobe der *Praktiker/innen* überschätzten Menschen aus der schwächsten Gruppe ihre eigenen Kompetenzen signifikant, während Menschen aus der stärksten Gruppe ihre Leistung signifikant unterschätzten. In der Gruppe der *Laien/innen* ging dieses Phänomen durchgängig mit einer größeren Unsicherheit in der Selbsteinschätzung einher. Die Betroffenen waren sich mithin zumindest teilweise ihrer eigenen Defizite bewusst („unskilled-but-subjectively-aware“; Händel & Fritsche, 2016; Miller & Geraci, 2011). In der Gruppe der *Praktiker/innen* galt dies nur bezogen auf die lokale, nicht aber in Bezug auf die globale Sicherheit. *Praktiker/innen* wiesen im Vergleich zu *Laien/innen* eine höhere Kompetenz auf und schätzten ihre Leistung auch zutreffend höher ein. Hierzu passte, dass sie auch eine signifikant höhere Sicherheit bei der Einschätzung der eigenen Leistung erlebten.

7 Gesamtdiskussion

In zwei Studien konnte erstmals der Dunning-Kruger-Effekt (Kruger & Dunning, 1999) in Bezug auf Wissen aus dem Bereich der Personal- und Organisationspsychologie weitestgehend belegt werden. Dies galt für *Laien/innen*, wie für *Praktiker/innen*, die in ihrem beruflichen Alltag personal- und organisationspsychologisches Wissen nutzenbringend einsetzen könnten. Letzteres liefert eine Teilerklärung für die vielfach beschriebene große Diskrepanz zwischen Forschung und Praxis in der Personalarbeit. Ein Teil der in der Praxis arbeitenden Personen wissen wenig über einschlägige Erkenntnisse aus der Forschung und gleichzeitig überschätzen sie ihr eigenes Wissen in diesem Bereich. In der Stichprobe der *Praktiker/innen* gilt Letzteres für etwa 44 % (Studie 1) bis 75 % (Studie 2) der befragten Personen. Unsere Berechnungen zur Sicherheit der eigenen Selbsteinschätzung belegt in drei von vier Fällen das Vorliegen eines unskilled-and-unaware-Effektes: Personen mit der geringsten oder zweitgeringsten Leistung sind in ihrer Überschätzung keineswegs unsicher, sondern glauben sich selbst ebenso zutreffend einschätzen zu können, wie Personen mit einem hohen und sehr hohen Leistungsniveau. Da sie sich ihrer Defizite nicht bewusst sind, gibt es für sie auch keinen Grund, sich entsprechend fachlich weiterzubilden und das eigene praktische Verhalten kritisch zu reflektieren oder gar zu ändern. In der Überzeugung besonders leistungsstark zu sein, besteht zudem die Gefahr, dass sie im Sinne eines Confirmation-Bias die eigenen fachlichen Defizite dauerhaft aufrechterhalten und sich falsche Überzeugungen festigen (Kanning et al., 2023). Abhilfe könnten hier gezielte Weiterbildungen schaffen. Je mehr die Betroffenen über einschlägiges Wissen verfügen, desto eher verstehen sie die Komplexität des Kompetenzfeldes, erkennen eigene Schwächen und überschätzen sich weniger (Kruger & Dunning, 1999).

Bedenklich ist, dass die untersuchten Praktiker/innen in Studie 1 keine höhere Leistung zeigten als Laien/innen und in Studie 2 im Mittelwert nur knapp eine Aufgabe mehr lösen konnten. Hierin spiegelt sich direkt das Science-Practice-Gap der Personal- und Organisationspsychologie.

Betrachten wir alle vier Teilstichproben in beiden Studien, so fällt auf, dass die Frage, ob die starke Selbstüberschätzung in der Gruppe der Menschen mit geringer Kompetenz den Handelnden bewusst ist oder nicht, nicht klar beantwortet werden kann. Wie auch in zuvor durchgeführten Studien gibt es sowohl Belege für das Prinzip unskilled-and-unaware als auch für das Prinzip unskilled-but-subjectively-aware. Zukünftige Studien sollten der Frage nachgehen, unter welchen Bedingungen eher das eine oder das andere gilt.

Zudem stellt sich die Frage, inwieweit verschiedene Formen der Selbsteinschätzung der eigenen Kompetenzen zu unterschiedlichen Ergebnissen gelangen. In den vorliegenden Studien wurden die Befragten gebeten, die Anzahl der gelösten Aufgaben zu schätzen. In anderen Studien bestand die Aufgabe in einem sozialen Vergleich. Die Befragten sollten einschätzen, wie viel Prozent einer Vergleichsgruppe (hier etwa die Gruppe der Personaler/innen) eine bessere oder schlechtere Leistung im Test erbringen würden (Kruger & Dunning, 1999). Die Einschätzung einer imaginären Bezugsgruppe sollte weitaus mehr Möglichkeiten zur Selbstüberschätzung bieten als die Einschätzung der eigenen Leistung bezogen auf die Testaufgaben. Hier ergibt sich ein weiteres Feld zur Erforschung des Dunning-Kruger-Effektes in Bereich des Personalwesens.

Die Aussagekraft beider Studien ist durch verschiedene Faktoren eingeschränkt. Beide Studien arbeiten nicht mit repräsentativen Stichproben aus den jeweiligen Populationen. Es bleibt daher unklar, inwieweit sich die Ergebnisse verallgemeinern lassen. Ganz ähnlich verhält es sich bei der Auswahl der Leistungsaufgaben zur Festlegung des Kompetenzniveaus. Auch sie sind nicht repräsentativ für das praxisrelevante Wissen aus dem Bereich der Personal- und Organisationspsychologie. In den beiden Gruppen der Berufspraktiker/innen wird nicht näher differenziert, in welchem Ausmaß berufliche Erfahrungen vorliegen, also beispielsweise, seit wie vielen Jahren und wie häufig pro Jahr die Befragten an Personalauswahlverfahren beteiligt sind. Dies weitergehend zu differenzieren wäre eine interessante Aufgabe für zukünftige Studien.

8 Literatur

- Armoneit, C., Schuler, H. & Hell, B. (2020). Nutzung, Validität, Praktikabilität und Akzeptanz psychologischer Personalauswahlverfahren in Deutschland 1985, 1993, 2007, 2019: Fortführung einer Trendstudie. *Zeitschrift für Arbeits- und Organisationspsychologie*, 64, 67-82. <https://doi.org/10.1026/0932-4089/a000311>
- Banks, G. C., Pollack, J. M., Bochantin, J. E., Kirkman, B. L., Whelpley, C. E. & O'Boyle, E. H. (2016). Management's science-practice gap: A grand challenge for all stakeholders. *Academy of Management Journal*, 59(6), 2205-2231. <https://doi.org/10.5465/amj.2015.0728>
- Conway, J. M., Jako, R. A. & Goodman, D. F. (1995). A meta-analysis of the interrater and internal consistency reliability of selection interviews. *Journal of Applied Psychology*, 80, 565-579. <https://doi.org/10.1037/0021-9010.80.5.565>
- Dunning, D. (2005). The Dunning-Kruger effect: On Being ignorant of one's own ignorance. *Advances in Experimental Social Psychology*, 44, 247-296. <https://doi.org/10.1016/b978-0-12-385522-0.00005-6>
- Dunning, D. (2011). The Dunning-Kruger effect: On being ignorant of one's own ignorance. In J. M. Olson & M. P. Zanna (Eds.), *Advances in Experimental Social Psychology*, 44, (pp. 247-296). Academic Press. <https://doi.org/10.1016/B978-0-12-385522-0.00005-6>
- Dunning, D., Johnson, K. Ehrlinger, J. & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, 12(3), 83-87. <https://doi.org/10.1111/1467-8721.01235>
- Ehrlinger, J., Johnson, K., Banner, M., Dunning, D. & Kruger, J. (2008). Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent. *Organizational Behavior and Human Decision Processes*, 105(1), 98-121. <https://doi.org/10.1016/j.obhdp.2007.05.002>
- Gabriel, M. T., Critelli, J. W. & Ee, J. S. (1994). Narcissistic illusions in self-evaluations of intelligence and attractiveness. *Journal of Personality*, 62(1), 143-155. <https://doi.org/10.1111/j.1467-6494.1994.tb00798.x>
- Händel, M. & Fritzsche, E. S. (2016). Unskilled but subjectively aware: Metacognitive monitoring ability and respective awareness in low-performing students. *Memory & Cognition*, 44(2), 229-241. <https://doi.org/10.3758/s13421-015-0552-0>
- Hodges, B., Regehr, G. & Martin, D. (2001). Difficulties in recognizing one's own in-competence: Novice physicians who are unskilled and unaware of it. *Journal of the Association of American Medical Colleges*, 12(76), 87-9. <https://doi.org/10.1097/00001888-200110001-00029>
- Hossiep, R., Schecke, J. & Weiß, S. (2015). Zum Einsatz von persönlichkeitsorientierten Fragebogen – Eine Erhebung unter den 580 größten deutschen Unternehmen. *Psychologische Rundschau*, 127-129. <https://doi.org/10.1026/0033-3042/a000235>
- Hülshager, U. R., Maier, G. W. & Stumpp, T. (2007). Validity of general mental ability for the prediction of job performance and training success in Germany: A meta-analysis. *International Journal of Selection and Assessment*, 15, 3-18. <https://doi.org/10.1111/j.1468-2389.2007.00363.x>

- Huffcutt, A. I. & Arthur, W. Jr. (1994). Hunter and Hunter (1984) revisited: Interview validity for entry-level jobs. *Journal of Applied Psychology*, 79, 184-190. <https://doi.org/10.1037/0021-9010.79.2.184>
- Huffcutt, A. I., Culbertson, S. S. & Weyhrauch, W. S. (2014). Moving forward indirectly: Reanalyzing the validity of employment interviews with indirect range restriction methodology. *International Journal of Selection and Assessment*, 22, 297-309. <https://doi.org/10.1111/ijsa.12078>
- Kanning, U. P. (2015). *Personalauswahl zwischen Anspruch und Wirklichkeit – Eine wirtschaftspsychologische Analyse*. Berlin: Springer. <https://doi.org/10.1007/978-3-662-45553-1>
- Kanning, U. P. (2016). Einstellungsinterviews in der Praxis. *Report Psychologie*, 11, 442-450.
- Kanning, U. P. (2021). Wenn Erfahrung nicht vor Torheit schützt – Urteilsfehler in der Personalauswahl. *Skeptiker*, 2, 64-71.
- Kanning, U. P. (2023). *Wider alle Vernunft – Coaching und HR-Management auf Abwegen*. Lengerich: Pabst.
- Kanning, U. P., Carstens, A. & Stenneken, A. (2023). Confirmation Bias in der Bewertung von Personalauswahlmethoden. *Zeitschrift für Arbeits- und Organisationspsychologie*, 67, 201-213. <https://doi.org/10.1026/0932-4089/a000413>
- 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>
- Kanning, U. P., Rosenstiel, L. v. & Schuler, H. (2010). *Jenseits des Elfenbeinturms: Psychologie als nützliche Wissenschaft*. Göttingen: Vandenhoeck & Ruprecht.
- Klehe, U.-C. & Latham, G. P. (2005). The predictive and incremental validity of the situational and patterned behavior description interviews for teamplaying behavior. *International Journal of Selection and Assessment*, 13, 108-115. <https://doi.org/10.1111/j.0965-075X.2005.00305.x>
- 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>
- Kruger, J. & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121-1134. <https://doi.org/10.1037/0022-3514.77.6.1121>
- Mattern, K. D., Burrus, J. & Shaw, E. (2010). When both the skilled and unskilled are unaware: Consequences for academic performance. *Self and Identity*, 9(2), 129-141. <https://doi.org/10.1080/15298860802618963>
- McDaniel, M. A., Whetzel, D. L., Schmidt, F. L., Hunter, J. E., Maurer, S. D. & Russel, J. (1994). The validity of employment interviews: A review and meta-analysis. *Journal of Applied Psychology*, 79, 599-602. <https://doi.org/10.1037/0021-9010.79.4.599>
- Meyer, J. P. (2013). The science-practice gap and employee engagement: It's a matter of principle. *Canadian Psychology*, 54(4), 235-245. <https://doi.org/10.1037/a0034521>
- Miller, T. M. & Geraci, L. (2011). Unskilled but aware: Reinterpreting overconfidence in low-performing students. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(2), 502-506. <https://doi.org/10.1037/a0021802>
- Moore, D. A. & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review* 115(2), 502-517. <https://doi.org/10.1037/0033-295x.115.2.502>
- Moore, D. A. & Schatz, D. (2017). The three faces of overconfidence. *Social and Personality Psychology Compass*, 11(8), e12331. <https://doi.org/10.1111/spc3.12331>
- Nye, C. D., Ma, J. & Wee, S. (2022). Cognitive ability and job performance: Meta-analytic evidence for the validity of narrow cognitive abilities. *Journal of Business and Psychology*, 37, 1119-1139. <https://doi.org/10.1007/s10869-022-09796-1>
- Quinones, M. A., Ford, J. K. & Teachout, M. S. (1995). The relationship between work experience and job performance: A conceptual and meta-analytic review. *Personnel Psychology*, 48, 887-910. <https://doi.org/10.1111/j.1744-6570.1995.tb01785.x>
- Rosenstiel, L. v. (2004). Arbeits- und Organisationspsychologie – Wo bleibt der Anwendungsbezug?. *Zeitschrift für Arbeits- und Organisationspsychologie*, 48, 87-94. <https://doi.org/10.1026/0932-4089.48.2.87>
- Sackett, P. R., Zhang, C., Berry, C. M. & Lievens, F. (2022). Revisiting meta-analytic estimates of validity in personnel selection: Addressing systematic overcorrection for restriction of range. *Journal of Applied Psychology*, 107(11), 2040-2068. <https://doi.org/10.1037/apl0000994>
- Salgado, J. F., Moscoso, S., de Fruyt, F., Anderson, N., Bertua, C. & Rolland, J. P. (2003). A meta-analytic study of general mental ability validity for different occupations in the European Community. *Journal of Applied Psychology*, 88, 1068-1081. <https://doi.org/10.1037/0021-9010.88.6.1068>

- Schmidt, F. L. & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practice and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262-274. <https://doi.org/10.1037/0033-2909.124.2.262>
- Tajfel, H. (Ed.). (1978). *Differentiation between social groups*. London: Academic Press.
- Taylor, S. E. & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, 103(2), 193-210. <https://doi.org/10.1037/0033-2909.103.2.193>
- Van Iddekinge, C. H., Arnold, J. D., Frieder, R. E. & Roth, P. L. (2019). A meta-analysis of the criterion-related validity of prehire work experience. *Personnel Psychology*, 72(4), 571-598. <https://doi.org/10.1111/peps.12335>
- Varelmann, L. & Kanning, U. P. (2018). Personalauswahl: Praktiker überschätzen Validität von Auswahlverfahren. *Wirtschaftspsychologie aktuell*, 1, 43-47.
- Wiesner, W. H. & Cronshaw, S. F. (1988). A meta-analytic investigation of the impact of interview format and degree of structure on the validity of the employment interview. *Journal of Occupational Psychology*, 61, 275-290. <https://doi.org/10.1111/j.2044-8325.1988.tb00467.x>
- Wills, T.A. (1991). Similarity and self-esteem in downward comparison. In J. Suls & T.A. Wills (Eds.), *Social comparison. Contemporary theory and research* (pp. 51-87). Hillsdale, NJ: Erlbaum.
- Zell, E. & Krizan, Z. (2014). Do people have insight into their abilities? A Metasynthesis. *Perspectives on Psychological Science*, 9(2), 111-125. <https://doi.org/10.1177/1745691613518075>

Anhang 1: Leistungsaufgaben aus Studie 1

Statement	Lösungsschlüssel	Beispiele für Literaturquellen
1. Ob ein Mitarbeiter eine Personalentwicklungsmaßnahme erfolgreich abschließen kann, hängt u.a. davon ab wie intelligent der Teilnehmer ist.	richtig	Schmidt, F. L. & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practice and theoretical implications of 85 years of research findings. <i>Psychological Bulletin</i> , 124, 262-274.
2. Outdoor-Trainings sind eine nachweislich effektive Methode der Personalentwicklung. Erläuterung: Outdoor-Trainings finden im Freien außerhalb der gewohnten Umgebung der meisten Arbeitnehmer statt, wobei Teams eine Reihe von häufig auch physisch anstrengenden Aufgaben bearbeiten.	falsch	Kanning, U. P. (2013). Wenn Manager auf Bäume klettern: <i>Mythen der Personalentwicklung und Weiterbildung</i> . Pabst.
3. Eine aussagekräftige Evaluation von Trainingsmaßnahmen besteht darin, die Teilnehmer/innen im Anschluss danach zu fragen, wie sie das Training erlebt haben.	falsch	Alliger, G. M., Tannenbaum, S. I., Bennett, W. Jr., Traver, H. & Shotland, A. (1997). A meta-analysis of the relations among training criteria. <i>Personnel Psychology</i> , 50, 341-358.
4. Männer profitieren stärker vom Mentoring als Frauen. Erläuterung: Um jemanden stärker in seiner Karriere zu unterstützen, begleitet ein erfahrener Kollege (Mentor) einen jüngeren Mitarbeiter (Mentee).	falsch	Allen, T. D., Eby, L. T., Poteet, M. L. & Lentz, E. (2004). Career benefit associated with mentoring for protégés: A meta-analysis. <i>Journal of Applied Psychology</i> , 89, 127-136.
5. Die Zielsetzung führt nur dann zu mehr Leistung, wenn die Mitarbeiter die Ziele für sinnvoll halten.	falsch	Klein, H. J., Wesson, M. J., Hollenbeck, J. R. & Alge, B. J. (1999). Goal commitment and the goal setting process. <i>Journal of Applied Psychology</i> , 84, 885- 896.
6. Menschen aus der Generation Y unterscheiden sich in ihren Arbeitsmotiven grundlegend von Menschen vorheriger Generationen. Erläuterung: Unter Generation Y befindet sich die Bevölkerungsgeneration, die im Zeitraum der frühen 1980er bis zu den späten 1990er Jahren geboren wurde.	falsch	Kanning, U. P. & Schmitt, A. (2023). Wie sinnvoll ist es, Menschen in Generationen einzuteilen? <i>Personalwirtschaft</i> , 9, 64-66.
7. Ziele sollten herausfordernd formuliert werden.	richtig	Locke, E. A. & Latham, G. P. (1990). <i>A Theory of goal setting and task performance</i> . Prentice Hall.
8. Beim Feedback ist es wichtig, die Sandwich-Methode einzusetzen. Dabei wird eine negative Botschaft zwischen zwei positive Botschaften gepackt.	falsch	Kanning, U. P. & Berkahn, J. (2020). Feedbackregel – Was wirkt wie? <i>Personalführung</i> , 11, 56-61.
9. Menschen lassen sich durch Geld nicht zu mehr Leistung motivieren.	falsch	Cerasoli, C. P., Nicklin, J. M., & Ford, M. T. (2014). Intrinsic motivation and extrinsic incentives jointly predict performance: A 40-year meta-analysis. <i>Psychological Bulletin</i> , 140(4), 980-1008.
10. Wenn Menschen in Gruppen zusammenarbeiten, führt dies dazu, dass die Gruppenleistung höher ausfällt als die Summe der Einzelleistungen.	falsch	Wegge, J. (2014). Gruppenarbeit und Management von Teams. In H. Schuler und U. P. Kanning (Hrsg.), <i>Lehrbuch der Personalpsychologie</i> (S. 933 – 984). Hogrefe.
11. In der Gruppenarbeit wirken sich Altersunterschiede unter den Teammitgliedern positiv auf die Gruppenleistung aus.	falsch	Joshi, A., & Roh, H. (2009). The role of context in work team diversity research: A meta-analytic review. <i>Academy of Management Journal</i> , 52(3), 599-627.
12. Die berufliche Leistung hängt überwiegend von der Arbeitszufriedenheit ab.	falsch	Sonntag, K., Frieling, E. & Stegmaier, R. (2012). <i>Lehrbuch Arbeitspsychologie</i> . Bern: Huber.
13. Die Mehrheit der Beschäftigten ist unzufrieden mit ihrer Arbeit.	falsch	Sonntag, K., Frieling, E. & Stegmaier, R. (2012). <i>Lehrbuch Arbeitspsychologie</i> . Bern: Huber.
14. Ein Bewerber, der 10 Jahre Berufserfahrung hat und dabei immer die gleichen Tätigkeiten verrichten musste, sollte eher eingestellt werden als jemand mit nur 5 Jahren Berufserfahrung in unterschiedlichen Tätigkeiten.	falsch	Quinones, M. A., Ford, J. K. & Teachout, M. S. (1995). The relationship between work experience and job performance: A conceptual and meta-analytic review. <i>Personnel Psychology</i> , 48, 887-910.
15. Bei der Sichtung von Zeugnissen sind Durchschnittsnoten aussagekräftiger als Einzelnoten.	richtig	Schuler, H., Funke, U. & Baron-Boldt, J. (1990). Predictive validity of school grades – A meta-analysis. <i>Applied Psychology: An International Journal Review</i> , 39, 89-103.
16. Das Neurolinguistische Programmieren ist ein nachweislich wirkungsvoller Ansatz der Personalentwicklung.	falsch	Greif, S. (2018). Woran erkennt man pseudowissenschaftliche Theorien und weshalb sie im Coaching problematisch sind – am Beispiel NLP. <i>Organisationsberatung Supervision Coaching</i> , 25, 371-387.
17. In der Personalauswahl sind allgemeine Intelligenztests sehr gut geeignet, in verschiedensten Berufen den späteren Berufserfolg zu prognostizieren.		Salgado, J. F., Moscoso, S., de Fruyt, F., Anderson, N., Bertua, C. & Rolland, J. P. (2003). A meta-analytic study of general mental ability validity for different occupations in the

	richtig	European Community. <i>Journal of Applied Psychology</i> , 88, 1068-1081.
18. Führungserfolg hängt vor allem von der Persönlichkeit der Führungskraft ab.	falsch	Derue, D. S., Nahrgang, J. D., Wellman, N. & Humphrey, S. E. (2011). Trait and behavioral theories of leadership: An integration and meta-analytic test of their relative validity. <i>Personnel Psychology</i> , 64, 7-52.
19. Viele kleine Pausen im Berufsalltag sind effektiver als wenige große.	richtig	Sonntag, K., Frieling, E. & Stegmaier, R. (2012). <i>Lehrbuch Arbeitspsychologie</i> . Bern: Huber.
20. Kurzfristig wirksame Stressbewältigungsstrategien können langfristig schädlich sein.	richtig	Sonntag, K., Frieling, E. & Stegmaier, R. (2012). <i>Lehrbuch Arbeitspsychologie</i> . Bern: Huber.

Anhang 2: Leistungsaufgaben aus Studie 2

Statement	Lösungsschlüssel	Literaturquelle
1. Rechtschreibfehler in den Bewerbungsunterlagen sollten bei der Bewertung der Bewerber/innen ignoriert werden.	richtig	Kanning, U. P., Budde, L. & Hülskötter, M. (2018). Wie valide ist die regelkonforme Gestaltung von Bewerbungsunterlagen? <i>PERSONALquaterly</i> , 4, 38-45.
2. Bei der Sichtung der Bewerbungsunterlagen sollten Freizeitaktivitäten (z. B. Hobbys, Sport) in der Regel nicht interpretiert werden.	richtig	Kanning, U. P. & Wörmann, J. (2018). Ist es sinnvoll, Freizeitaktivitäten in der Personalauswahl zu berücksichtigen? <i>Report Psychologie</i> , 2, 58-66.
3. Im Vergleich zu anderen Methoden (z. B. strukturierte Einstellungsinterviews, Fachwissenstests, Assessment Center) sind die Bewerbungsunterlagen eine besonders aussagekräftige Methode zur Prognose der beruflichen Leistung.	falsch	Van Iddekinge, C. H., Arnold, J. D., Frieder, R. E., & Roth, P. L. (2019). A meta-analysis of the criterion-related validity of prehire work experience. <i>Personnel Psychology</i> , 72(4), 571-598.
4. Arbeitszeugnisse liefern sichere Informationen im Hinblick auf die vergangene Leistung der Bewerber/innen.	falsch	Sende, C. (2016). <i>Arbeitszeugnisse aus eignungsdiagnostischer Perspektive</i> . Dissertation Friedrich-Alexander-Universität Erlangen-Nürnberg.
5. Lücken im Lebenslauf geben in der Regel keinen nennenswerten Aufschluss über die Zielstrebigkeit der Bewerber/innen.	richtig	Frank, F., Wach, D. & Kanning, U. P. (2017). Zusammenhang zwischen Lücken im Lebenslauf und Berufserfolg: Ein Mythos der Personalauswahlpraxis. <i>Zeitschrift für Arbeits- und Organisationspsychologie</i> , 61, 69-80.
6. Strukturierte Einstellungsinterviews eignen sich besser für die Vorhersage der Leistung von Bewerber/innen als unstrukturierte Einstellungsinterviews. Erläuterung: Bei strukturierten Einstellungsinterviews folgen die Fragen einem festgelegten Leitfadens; die Bewertung erfolgt nach verbindlichen Kriterien. Bei unstrukturierten Einstellungsinterviews folgen die Beurteiler/innen ihrer Menschenkenntnis; Fragen folgen keinem festgelegten Leitfadens und die Bewertung erfolgt nicht nach verbindlichen Kriterien.	richtig	Sackett, P. R., Zhang, C., Berry, C. M., & Lievens, F. (2022). Revisiting meta-analytic estimates of validity in personnel selection: Addressing systematic overcorrection for restriction of range. <i>Journal of Applied Psychology</i> , 107(11), 2040-2068.
7. Je häufiger jemand unstrukturierte Einstellungsinterviews durchführt, desto besser wird er/sie in der Einschätzung der Bewerber/innen.	falsch	Kanning, U. P., Book, D. & König, L. (2022). Ähnlichkeits-Attraktivitäts-Effekt in der Personalauswahl – Schützt Berufserfahrung vor Urteilsfehlern? <i>Report Psychologie</i> , 3, 16-24.
8. Die Fragen nach den Stärken und Schwächen der Bewerber/innen in einem Einstellungsinterview geben korrekte Informationen über deren persönlichen Eigenschaften.	falsch	Kanning, U. P. (2017). Strategisches Verhalten in der Personalauswahl – Wie Bewerber versuchen, ein gutes Ergebnis zu erzielen. <i>Zeitschrift für Arbeits- und Organisationspsychologie</i> , 61, 3-17.
9. In einem Einstellungsinterview sollten möglichst viele (anstatt wenige ausgewählte) Eigenschaften und Kompetenzen der Bewerber/innen untersucht werden.	falsch	Schuler, H. (2018). <i>Das Einstellungsinterview</i> . Hogrefe.
10. Mehrere Fragen pro Jobanforderung (z. B. Teamfähigkeit, Kontaktfähigkeit) an die Bewerber/innen erhöhen die Aussagekraft eines Einstellungsinterviews.	richtig	Schuler, H. (2018). <i>Das Einstellungsinterview</i> . Hogrefe.
11. In einem Assessment Center sollten die Beobachter/innen das Verhalten der Bewerber/innen nach jeder Übung diskutieren. Erläuterung: Ein Assessment Center ist eine Auswahlmethode, bei der die Bewerber/innen mehrere unterschiedliche Übungen wie z. B. Rollenspiele oder Arbeitsproben durchlaufen und dabei von Beobachter/innen bewertet werden.	falsch	Poersch, A. & Kanning, U. P. (2022). Wie Führungskräfte die Ergebnisse von Assessment Centern beeinflussen. <i>Personalführung</i> , 4, 10-13.
12. Die Beobachter/innen im Assessment Center sollten über die Bewerbungsunterlagen und die Ergebnisse vorgeschalteter Auswahlverfahren (z. B. Übungsergebnisse) der Bewerber/innen informiert werden.		Wenderdel, M. & Kanning, U. P. (2008). Wer mehr weiß beur-

	falsch	teilt anders. <i>Personalwirtschaft</i> , 8, 52-54.
13. Um die Aussagekraft eines Assessment Centers zu steigern, sollten nur Verhaltensübungen (z. B. Rollenspiele, Gruppendiskussionen) und keine Testverfahren (z. B. Leistungstest) eingesetzt werden.	falsch	Becker, N., Höft, S., Holzenkamp, M. & Spinath, F. M. (2011). The predictive validity of assessment centers in German-speaking regions. <i>Journal of Personnel Psychology</i> , 10, 61-69.
14. Beobachter/innen sollten im Assessment Center (z. B. in den Pausen) persönlichen Kontakt zu den Bewerber/innen haben.	falsch	Kanning, U. P. (2019). <i>Standards der Personaldiagnostik</i> . Hogrefe.
15. Erfahrene Führungskräfte, die in einem Assessment Center als Beobachter/innen fungieren sollen, müssen trotz ihrer Erfahrung ausgiebig geschult werden.	richtig	Kanning, U. P. (2021). Wenn Erfahrung nicht vor Torheit schützt – Urteilsfehler in der Personalauswahl. <i>Skeptiker</i> , 2, 64-71.
16. Bei einer Auswahlentscheidung sollte die Vielfältigkeit der bisherigen beruflichen Aufgaben der Bewerber/innen stärker gewichtet werden als die Dauer der Berufserfahrung.	richtig	Quinones, M. A., Ford, J. K. & Teachout, M. S. (1995). The relationship between work experience and job performance: A conceptual and meta-analytic review. <i>Personnel Psychology</i> , 48, 887-910.
17. Die Qualität der Auswahlentscheidungen ist besonders hoch, wenn erfahrene Entscheidungsträger/innen ihrer Intuition folgen.	falsch	Kanning, U. P. (2021). Wenn Erfahrung nicht vor Torheit schützt – Urteilsfehler in der Personalauswahl. <i>Skeptiker</i> , 2, 64-71.
18. Bei mehrstufigen Auswahlprozessen (z. B. Sichtung der Bewerbungsunterlagen, Einstellungsinterview, Assessment Center o.ä.) sollten in jeder Stufe unterschiedliche Beobachter/innen eingesetzt werden.	richtig	Wenderdel, M. & Kanning, U. P. (2008). Wer mehr weiß beurteilt anders. <i>Personalwirtschaft</i> , 8, 52-54.
19. Es ist wichtig, vor der Durchführung eines Auswahlverfahrens die Anforderungen der Stelle spezifisch zu analysieren (z. B. durch die Befragung von mehreren Arbeitsexperten/innen).	richtig	Huffcutt, A. I., Culbertson, S. S. & Weyhrauch, W. S. (2014). Moving forward indirectly: Reanalyzing the validity of employment interview with indirect range restriction methodology. <i>International Journal of Selection and Assessment</i> , 22, 297-309.
20. Im Durchschnitt sind allgemeine Intelligenztests in der Personalauswahl gut geeignet, um den späteren Berufserfolg vorherzusagen.	richtig	Sackett, P. R., Zhang, C., Berry, C. M., & Lievens, F. (2022). Revisiting meta-analytic estimates of validity in personnel selection: Addressing systematic overcorrection for restriction of range. <i>Journal of Applied Psychology</i> , 107(11), 2040-2068.

Korrespondenzadresse:

Prof. Dr. Uwe P. Kanning
Hochschule Osnabrück
Fakultät Wirtschafts- und Sozialwissenschaften
Caprivistrasse 30a
49076 Osnabrück
Deutschland
U.Kanning@hs-osnabrueck.de