



Forschungsnetzwerk Arbeit und Bildung

Kontextanalysen im KOMET- Forschungsprojekt: Erfassen der Testmotivation

Nele Bachmann, Jenny Frenzel, Felix Rauner

Hrsg.:

Universität Bremen
FG Berufsbildungsforschung (:BB)

KIT – Karlsruher Institut für Technologie
Institut für Berufspädagogik und
Allgemeine Pädagogik

Carl von Ossietzky Universität Oldenburg
Institut für Physik/Technische Bildung

Pädagogische Hochschule Weingarten
Professur für Technikdidaktik

**A + B
Forschungsberichte**

16

Nele Bachmann, Jenny Frenzel, Felix Rauner

Kontextanalysen im KOMET-Forschungsprojekt: Erfassen der Testmotivation

A+B Forschungsberichte Nr. 16/2014

Bremen, Oldenburg, Karlsruhe, Weingarten: A+B Forschungsnetzwerk

In den A+B Forschungsberichten werden aktuelle Forschungsberichte aus der Arbeits- und Bildungsforschung veröffentlicht. Arbeit und Bildung verweist auf die vorberufliche und die berufliche Bildung sowie auf die berufliche Weiterbildung. Diese Form der online-Publikation erlaubt es, Forschungsergebnisse zu einem frühen Zeitpunkt zugänglich zu machen.

Jeder Forschungsbericht durchläuft ein internes Reviewverfahren. Die Reihe A+B Forschungsberichte ist auch offen für externe Autoren, die dem Forschungsnetzwerk durch ihre Forschungsarbeiten verbunden sind. Die Verantwortung für den Inhalt der Beiträge liegt bei den Autoren.

A+B Forschungsberichte is a series where topical results of the current research on labour and education are being published. Labour and education refers to pre-vocational education, vocational education and training as well as continuing vocational education. In order to assure a high degree of topicality, A+B Forschungsberichte is published online. Quality is guaranteed by an internal review process involving several researchers. A+B Forschungsberichte offers a platform also for external researchers, who are linked to the Forschungsnetzwerk via their own research in the field of labour and education. The authors are responsible for the content of their contributions.

A + B Forschungsberichte erscheinen online unter:

www.ibb.uni-bremen.de

www.ibap.kit.edu

www.uni-oldenburg.de

www.ph-weingarten.de

ISSN 1867-9277

Redaktion: apl. Prof. Dr. Walter Jungmann

Kontakt: walter.jungmann@kit.edu

Institut für Berufspädagogik und Allgemeine Pädagogik
KIT – Karlsruher Institut für Technologie

(Universität des Landes Baden-Württemberg und nationales Forschungszentrum in der Helmholtz-Gemeinschaft)

© 2014, A+B Forschungsnetzwerk

Universität Bremen
FG Berufsbildungsforschung
(i:BB)

Leobener Straße/NW 2
28359 Bremen
Tel. +49 421 218-4634
E-Mail: rauner@uni-bremen.de

KIT – Karlsruher Institut für
Technologie
Institut für Berufspädagogik
und Allgemeine Pädagogik

Hertzstr. 16
76187 Karlsruhe
Tel.: +49 721 608-43690
Fax: +49 721 608-46104
E-Mail: m.fischer@kit.edu

Universität Oldenburg
Institut für Physik
/Technische Bildung

Ammerländer Heerstr. 114-
118 – 26111 Oldenburg
Tel.: +49 441 798-2966
Fax: +49 441 798-2967
E-Mail: peter.roeben@uni-oldenburg.de

Pädagogische Hochschule
Weingarten
Professur Technikdidaktik

Kirchplatz 2
88250 Weingarten
Tel.: +49 751 501-8273
Fax: +49 751 501-8200
E-Mail: haasler@ph-weingarten.de

Kontextanalysen im KOMET-Forschungsprojekt: Erfassen der Testmotivation

A+B Forschungsberichte Nr. 16/2014

Zusammenfassung:

Die Auswirkung der Testmotivation auf die Testergebnisse im Rahmen von Kompetenztests galt als eine vernachlässigbare Variable in der Kompetenzdiagnostik. Die Testpraxis des KOMET-Projektes zeigte schon bald, dass es erhebliche berufsspezifische und länderspezifische Unterschiede in der Testmotivation gibt und dass ein signifikanter Zusammenhang zwischen den Kompetenzniveaus von Testgruppen und deren Testmotivation besteht. Darüber hinaus beeinflusst das Testarrangement – zum Beispiel die Dauer der Tests – die Testmotivation. Die Einführung eines Tests zur Erfassung der Testmotivation trägt zur vertieften Interpretation der Testergebnisse bei. Vor allem bei berufs- und bildungsgangübergreifenden sowie international vergleichenden Projekten ist das Erfassen der Testmotivation eine notwendige Voraussetzung für die Analyse und Interpretation der Testergebnisse.

Schlüsselbegriffe: Testmotivation, Erfassen der Testmotivation, Bearbeitungszeit von Testaufgaben, Testverweigerer

Abstract:

The effect of test motivation on the test results in the context of competency tests was considered as a neglected variable in the competence diagnostics. The test practice of KOMET project quickly proved that there are significant job-specific and country-specific differences in the test motivation and that there is a significant relationship between the level of competence of test groups and their test motivation. In addition, the test arrangement influences—for example, the duration of the test—the test motivation. The introduction of a test for the detection of test motivation contributes to a carefully interpretation of the test results. Especially in occupational and educational background transition as well as in international comparative projects the detecting of test motivation is a necessary requirement for the analysis and interpretation of test results.

Inhaltsverzeichnis

Einleitung	6
Kontexterhebungen und -analysen	6
1. Erfassen der Testmotivation	7
2. Das Testverfahren	12
4. Testmotivation von Elektronikern	15
3. Explorative Faktorenanalyse zum Zusammenhang zwischen Testmotivation und Testleistung	16
4. Testmotivation der Pflege-Studierenden (Schweiz)	22
5. Die kulturelle Dimension der Testmotivation	27
6. Fazit	29
Literaturangaben	32

Einleitung

Der Ertrag der Kompetenzdiagnostik in der beruflichen Bildung hängt nicht nur von der Qualität des Testverfahrens und der Testaufgaben ab, sondern ebenso von den Kontext-Analysen. Sie sind daher Bestandteil aller KOMET-Projekte.

Die *Testergebnisse* zur Entwicklung und Ausprägung der

- beruflichen Kompetenz der Testgruppen: der Kompetenzniveaus und der Kompetenz-Profile,
- Kooperationsfähigkeit und -erfahrung (als zentrale Dimensionen sozialer Kompetenz),
- beruflichen und betrieblichen Identität und der darauf basierenden Leistungsbereitschaft (commitment)

sind weitgehend selbsterklärend, soweit es um die „gemessenen“ Werte geht. Die empirischen Testergebnisse repräsentieren *objektive* Gegebenheiten. Sie erschließen sich den Adressaten der Berichte umso eher, je detaillierter diese die Einzelbefunde miteinander in Beziehung setzen können. So zeigen die KOMET-Testergebnisse zum Beispiel, dass sich das Kompetenzniveau der Testteilnehmer des zweiten und dritten Ausbildungsjahres häufig nicht voneinander unterscheidet. Über die Ursache(n) für diesen (objektiven) Befund sagen diese Daten jedoch noch nichts aus.

Die *subjektive* Komponente der „selbsterklärenden Testergebnisse“ basiert auf dem Kontextwissen der Rezipienten. Der Grad an Kenntnissen über die Testteilnehmer und die Ausbildungsbedingungen bestimmt die Tiefe, mit der die Testergebnisse interpretiert werden können. In der Regel verfügen Lehrer und Ausbilder über Erfahrungen und Kenntnisse des Lern- bzw. Ausbildungsverhaltens „ihrer“ Auszubildenden sowie die Kontextbedingungen der Ausbildung, deren Qualität mit einer Kontexterhebung nicht erschöpfend erfasst werden kann. Für die Durchführung von Kontextanalysen bedeutet dies, die Potenziale der subjektiven und objektiven Kontextanalysen zu nutzen, indem zum Beispiel die Ergebnisse von quantitativen Kontext-Analysen in Feedback-Workshops mit den beteiligten Lehrern bzw. Ausbildern reflektiert werden.

Kontexterhebungen und -analysen

lassen sich in der Form eines Schalenmodells darstellen. Danach enthält die

Schale I: Daten, die der Beschreibung der Stichprobe dienen: Angaben zur Person sowie formale Angaben zur Ausbildung. Bereits auf dieser ersten Ebene stellt sich die Frage nach den zu erfassenden Merkmalen. Davon hängt ab, auf welchem Differenzierungsniveau die Teilergebnisse ausgewertet, analysiert und dokumentiert werden können. Dabei ist zu prüfen, ob und gegebenenfalls welche Hypothesen auf dieser Analyseebene untersucht werden sollen; zum Beispiel die Hypothese, dass sich die schulische Vorbildung auf die Kompetenzausprägung auswirkt.

Schale II: Erfassen der Testmotivation und anderer Persönlichkeitsmerkmale wie Intelligenz, Wunschberuf, sprachliche und mathematische Fähigkeiten usw.

Schale III: Erfassen der Kontextbedingungen der Ausbildung

- a) aus der Sicht der Testpersonen
- b) aus der Sicht der Lehrer/Ausbilder.

Schale IV: die Berufsausbildung als Schritt im Prozess des lebenslangen Lernens. Hier geht es um das Erfassen der beruflichen Karriereabsichten und -orientierungen der Testpersonen. Es ist zum Beispiel von erheblichem Interesse für die Analyse der Testergebnisse, ob die Testteilnehmer eine Berufsausbildung absolvieren als eine Zugangsvoraussetzung für ein darauf aufbauendes Studium, als Voraussetzung für eine Meisterprüfung oder ob der Wunsch im Vordergrund steht, den zu erlernenden Beruf auszuüben.

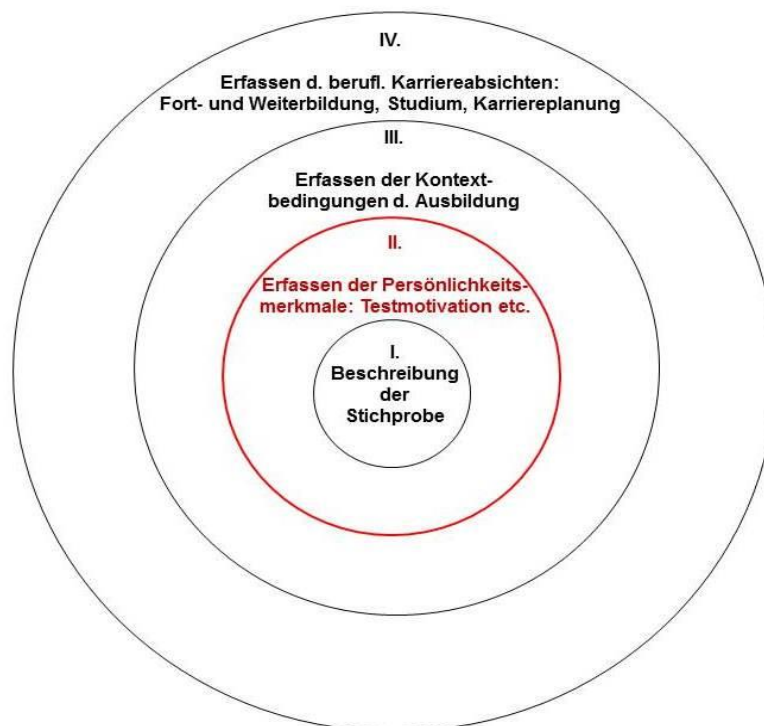


Abb. 1: Die Kontexterhebungen als Schalenmodell

In diesem Forschungsbericht werden die Fragestellungen, Methoden und ausgewählte Ergebnisse zur Testmotivation als eine Determinante der Entwicklung und Ausprägung beruflicher Kompetenz dargestellt.

1. Erfassen der Testmotivation

Zum Einfluss der Testmotivation auf das Testverhalten und die Testergebnisse liegen in der einschlägigen Forschung unterschiedliche und zum Teil sich widersprechende Untersuchungsergebnisse vor. Im Rahmen des ersten PISA-Tests (2000) wurde in Deutschland eine experimentelle Zusatzuntersuchung durchgeführt, um Erkenntnisse über die Testmotivation unterschiedlicher Schülergruppen (differenziert nach Schulformen) sowie den Einfluss verschiedener externer Anreize auf die Testmotivation zu gewinnen. Das zusammenfassende Ergebnis des Experimentes lautet, dass sich die unterschiedlichen Versuchsgruppen nicht im Hinblick auf ihre Anstrengungsbereitschaft unterscheiden (BAUMERT u. a.: Pisa 2000, S. 60).

Das Untersuchungsergebnis legt außerdem nahe, nicht auf externe Anreize zu setzen, da diese Effekte zu vernachlässigen seien. Unterschiede zwischen Hauptschülern und Gymnasiasten in der Testmotivation konnten in dieser Untersuchung nicht festgestellt werden. Unabhängig von der Schulform war die Anstrengungsbereitschaft durchgängig relativ hoch. *Die verschiedenen Anreize hatten keinen signifikanten Einfluss auf die Testergebnisse* (ebd., S. 27 ff.).

Im ersten KOMET-Projekt wurde daher auf einen Motivationstest verzichtet. Die Testpraxis legte dann jedoch nahe, im Rahmen der Längsschnittuntersuchung bereits beim zweiten Test die Testmotivation als eine Kontextvariable zu berücksichtigen (Tab. 1).

Testinstrument	Einsatz ab Testzeitpunkt
Offene Testaufgaben	t ₁ (2008)
Kontextfragebogen	t ₁ (2008)
Fragebogen zur Testmotivation für die Auszubildenden	t ₂ (2009)
Lehrerfragebogen zur Testmotivation der Auszubildenden	t ₂ (2009)
Raterfragebogen zur Gewichtung der Kompetenzkriterien	t ₂ (2009)
Test kognitiver Grundfähigkeit (Untertest „Figurenalogien“ des Kognitiven Fähigkeits-Test (KFT))	t ₂ (2009)

Tab. 1: Testinstrumentarium zum 1. und 2. Testzeitpunkt des ersten KOMET-Projektes (RAUNER u.a. 2009, S. 41)

Der zeitliche Testumfang als Einflussgröße auf die Testmotivation

Ausgehend von den Erfahrungen der Prüfungspraxis in der dualen Berufsausbildung sowie bei Abschlussprüfungen an Fachschulen wurde im KOMET-Projekt Elektroniker (Bremen, Hessen 2007-2009) eine Längsschnittuntersuchung mit einem Cross-Over-Design gewählt (Abb. 2). Danach waren von jedem Testteilnehmer zu jedem der beiden Testzeitpunkte je zwei komplexe Testaufgaben zu lösen.

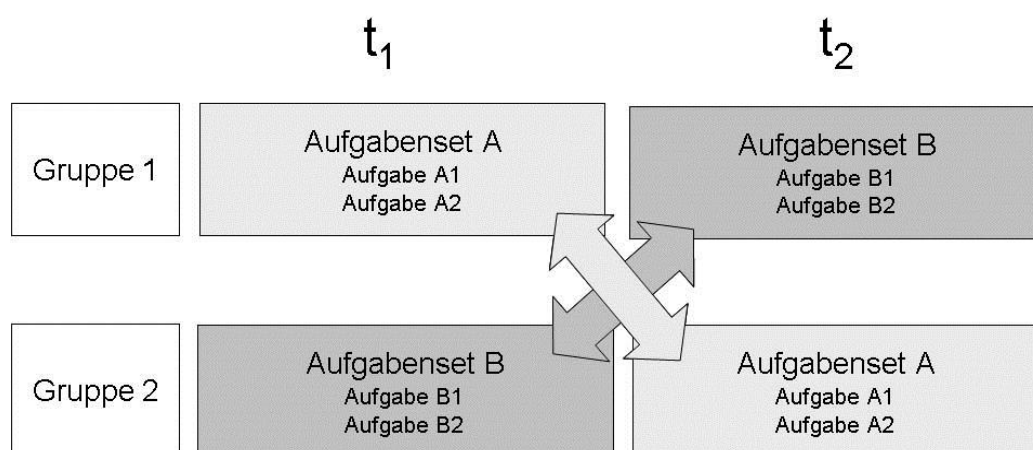


Abb. 2: Cross-Over-Design zum Einsatz von Testaufgaben im Längsschnitt (RAUNER, HAASLER, HEINEMANN, GROLLMANN 2009, 144 f.)

Der Test umfasste insgesamt vier komplexe Testaufgaben (RAUNER u. a. 2009, S.144 f.). Jeder Testteilnehmer hatte zwei Testaufgaben zu lösen. Für die Lösung jeder Testaufgabe standen (maximal) 120 Minuten zur Verfügung.

In einem Pretest wurde zunächst untersucht, ob es einen systematischen Leistungsabfall bei der Bearbeitung der zweiten komplexe Testaufgaben gibt. Bei der Auswertung der Testergebnisse wurde differenziert nach Auszubildenden des zweiten und dritten Ausbildungsjahres der Elektroniker für Betriebstechnik (E-B) und der der Fachrichtung Energie- und Gebäudetechnik (E-EG) sowie nach Fachschulstudierenden.

Die Erfahrungen der ersten Untersuchungsphase legten bereits die Vermutung nahe, dass die Motivation zur Bearbeitung der Testaufgaben bei den verschiedenen Gruppen von Auszubildenden eine größere Rolle spielt als angenommen. Zum einen ließen Berichte der Lehrer darauf schließen, dass die Testmotivation bei Berufsschülern unterschiedlich stark ausgeprägt ist. Zum anderen hat ein Teil der Auszubildenden die Testzeit von 2 x 120 Minuten nur zum Teil ausgeschöpft; einige Schüler haben die Testaufgaben nicht ernsthaft bearbeitet, sie bilden die Gruppe der Testverweigerer¹.

Ausgehend von diesen Erkenntnissen wurde zum zweiten Testzeitpunkt (März/April 2009) die Testmotivation sowie das Testverhalten erfasst, wobei die Formulierung der Fragen an die PISA-Testpraxis angelehnt ist (KUNTER 2002). Des Weiteren haben die die Tests durchführenden Lehrkräfte Fragen zur Testmotivation in der Klasse sowie zur Arbeitsatmosphäre beantwortet; die Ergebnisse können beim Vergleich auf Klassenebene herangezogen werden. Darüber hinaus lässt der Vergleich von Bearbeitungszeit und Testergebnis der ersten mit der zweiten Testaufgabe Rückschlüsse auf den Motivationsverlauf während des Testverlaufs zu.

Ergebnisse der Befragung zur Testmotivation (vgl. HEINEMANN, MAURER, RAUNER 2011, Kap. 6.7)

Das generelle Interesse der Auszubildenden und Studierenden an den KOMET-Testaufgaben ist unterschiedlich stark ausgeprägt. Die Bearbeitung der ersten Testaufgabe fand gut die Hälfte der E-B *interessant* bis *sehr interessant* (55%). Dieser Wert fällt bei den E-EG (61%) und den Fachschulstudierenden (60%) noch höher aus.

Insgesamt haben alle Testgruppen die erste Aufgabe eher konzentriert (73%) und sorgfältig (65%) bearbeitet. Auffällig ist dagegen, dass knapp jeder zweite E-B angibt, für die Bearbeitung der zweiten Aufgabe weniger motiviert zu sein als für die erste Aufgabe; bei den E-EG und den Fachschulstudierenden gibt dies nur jeder Fünfte an (Abb. 3).

¹ Als Testverweigerer gelten Teilnehmer, die einen Gesamtpunktwert unter fünf erzielt haben oder die beide Testaufgaben zusammen weniger als 60 Minuten bearbeitet haben. In Hessen wurden 2009 nach dieser Festlegung 24 Teilnehmer als Verweigerer identifiziert, davon zehn Auszubildende E-EG (7%) und vierzehn Auszubildende E-B (6%).

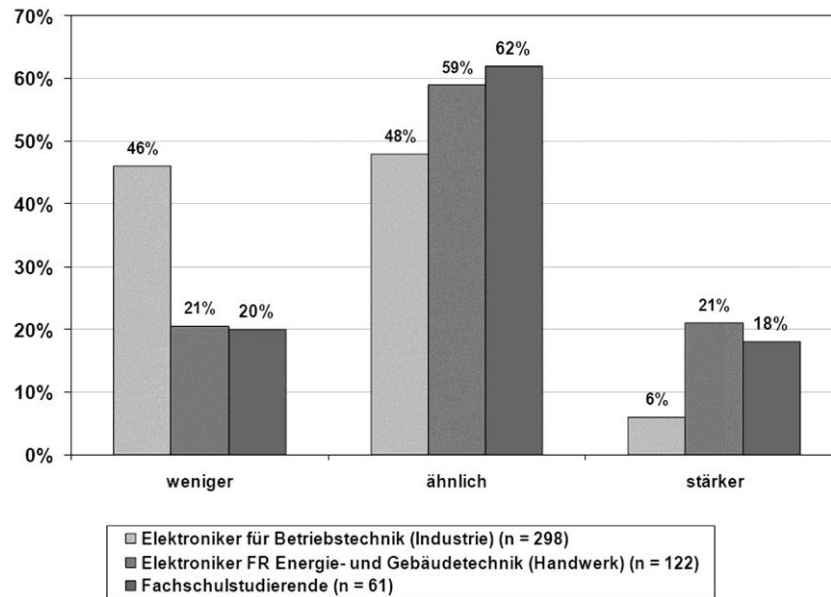


Abb. 3: Häufigkeitsverteilung nach Testgruppen: „Für die Bearbeitung der zweiten Testaufgabe bin ich ... motiviert als für die erste Testaufgabe.“

Vergleich der Ergebnisse der ersten und zweiten Testaufgabe

In diesem Zusammenhang lässt sich durch den Vergleich der Testergebnisse der beiden Testphasen (2 x 120 Minuten für zwei Testaufgaben) untersuchen, ob und bei welchen Testgruppen es signifikante Unterschiede in den Testergebnissen zwischen der ersten und zweiten Testaufgabe gibt. Fällt das Testergebnis einer Testgruppe bei der zweiten Testaufgabe schlechter aus, kann dies als ein Indiz für eine nachlassende Testmotivation gewertet werden.

Der Effekt einer nachlassenden Testmotivation liegt nicht für alle Testgruppen vor. Bei den Auszubildenden E-EB lässt sich kein Unterschied zwischen den Ergebnissen der ersten und der zweiten Testaufgabe feststellen. Dies kann darin begründet liegen, dass diese Gruppe ein insgesamt eher niedriges Testniveau aufweist. Die Auszubildenden Elektroniker für Betriebstechnik erzielen bei der ersten Aufgabe ein signifikant besseres Ergebnis als bei der zweiten Aufgabe: 15% erreichen auf Basis der ersten Testaufgabe das höchste Kompetenzniveau, auf Basis der zweiten Testaufgabe sind es nur 6% (vgl. Abb. 4). Dies entspricht auch der oben beschriebenen geringeren Motivation dieser Testgruppe für die zweite Testaufgabe. Bei der Gruppe der Fachschulstudierenden zeigt sich ein weiterer Effekt: Bei der ersten Aufgabe beträgt die Risikogruppe lediglich 10%, bei der zweiten Aufgabe steigt dieser Wert auf 23% (vgl. Abb. 5). Auch hier ergibt ein Test auf Mittelwertsunterschiede², dass der durchschnittliche Gesamtpunktwert bei der ersten Aufgabe signifikant höher ausfällt als bei der zweiten Aufgabe.

² t-Test für abhängige Stichproben

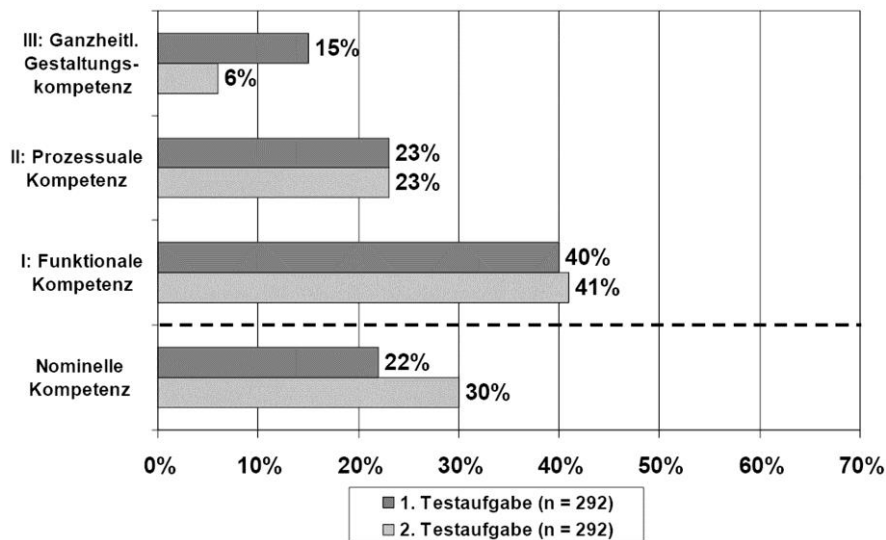


Abb. 4: Kompetenzniveaueverteilung der Gruppe Elektroniker/innen für Betriebstechnik (Hessen), Vergleich der Ergebnisse auf Grundlage der ersten und der zweiten Testaufgabe 2009

Dabei kommt es vor allem bei den leistungsschwächeren Studierenden zu einem erheblichen Motivationsverlust. Leistungsstärkere Studierende verbessern sich von der ersten zur zweiten Testaufgabe. Die Mehrheit schneidet jedoch schlechter ab als bei der ersten Aufgabe. Abb. 6 veranschaulicht diesen Effekt: Jedes Kreuz im Diagramm stellt einen Auszubildenden dar, auf den Achsen sind die jeweils erreichten Gesamtpunktwerte bei den beiden bearbeiteten Aufgaben abzulesen. Die diagonale Linie teilt die Grafik in zwei Teile. Im oberen Bereich (A) befinden sich die Auszubildenden, die bei der zweiten Aufgabe besser abgeschnitten haben als bei der ersten Aufgabe, im unteren Bereich (B) diejenigen, die sich bei der zweiten Aufgabe verschlechtert haben. Deutlich mehr Teilnehmer (63%) befinden sich in Bereich B.

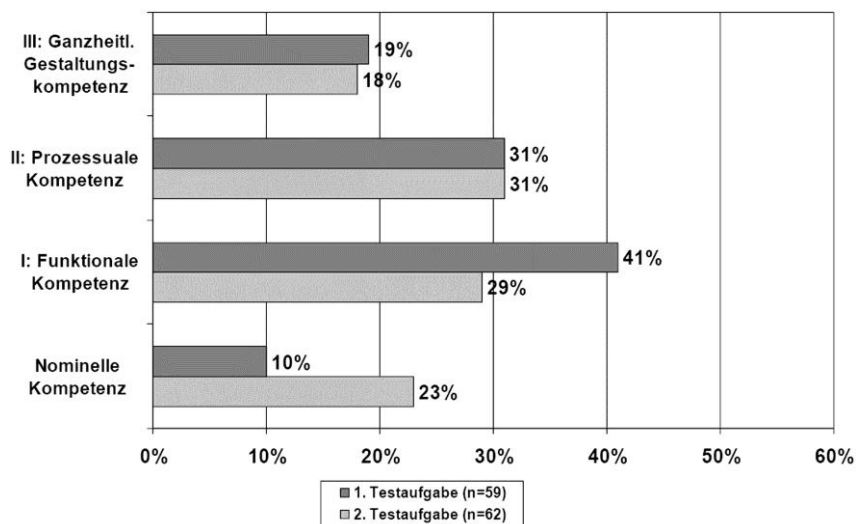


Abb. 5: Kompetenzniveaueverteilung der Gruppe Fachschulstudierende (Hessen), Vergleich der Ergebnisse auf Grundlage der ersten und der zweiten Testaufgabe 2009

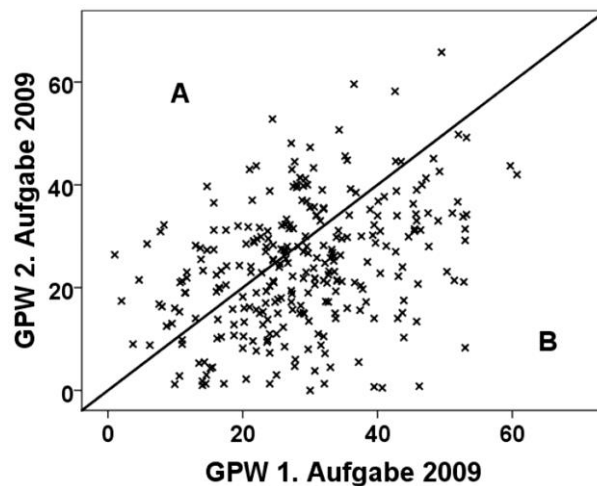


Abb. 6: Streudiagramm zum Vergleich des Ergebnisses bei der ersten und zweiten Aufgabe (Hessen, Elektroniker/innen für Betriebstechnik, 2009, $n = 297$)

Vergleich der Bearbeitungszeit bei der ersten und zweiten Testaufgabe

Die Erfassung der Bearbeitungszeit erlaubt ebenfalls eine Einschätzung zur Frage, inwieweit die Motivation im Verlauf der Testung sinkt. Für die erste Testaufgabe arbeiteten die Testteilnehmer durchschnittlich 100 Minuten, für die zweite Testaufgabe nur noch 83 Minuten. Dies kann als Ermüdungserscheinung bzw. nachlassende Motivation interpretiert werden.

Dieser Pretest zum Zusammenhang von Testmotivation, Bearbeitungszeit und Testergebnissen führte zur Entscheidung, den Testumfang für jeden Testteilnehmer auf die Bearbeitung einer Testaufgabe zu reduzieren. Erst bei den nachfolgenden Projekten wurde die Testmotivation als eine Determinante der Kompetenzentwicklung umfassender in die Kontextanalyse einbezogen.

2. Das Testverfahren

Bei den Items zum Erfassen der Testmotivation unterscheiden wir zwischen

1. den *primären* Motivationsaspekten der Testmotivation

- der Berufsbezogenheit der Testaufgaben
- der Nützlichkeit der Testaufgaben und
- dem Interesse an der Testbearbeitung sowie

2. den *sekundären* Motivationsaspekten

- der Konzentration
- Mühe
- Sorgfalt und
- der Anstrengung bei der Aufgabenlösung.

(vgl. dazu das Testmotivationsmodell Abb. 8).

Die primären Motivationsaspekte repräsentieren die Bewertung der Testaufgaben als relevant für die Berufsausbildung, ohne dass damit bereits die Bereitschaft verbunden ist, sich bei der Bearbeitung der Testaufgaben anzustrengen. Wenn zum Beispiel ein Test in zeitlicher Nähe zu einer Abschlussprüfung stattfindet, kann dies ein Desinteresse am Test auslösen, da dieser als eine Störung in der Prüfungsvorbereitung empfunden wird. Die Testmotivation wird dann beeinträchtigt, ohne dass die Bewertung des Berufsbezugs der Testaufgaben und ihre grundsätzliche Nützlichkeit für die Berufsausbildung davon tangiert werden. Die Bewertung des Interesses an den Testaufgaben bzw. dem Test ergibt sich aus dem Berufsbezug und zugleich aus dem Nutzen des Tests für die Ausbildung sowie gegebenenfalls auch für die Prüfungsvorbereitung.

Die sekundären Motivationsaspekte ergeben sich aus der primären Motivationsdimension. Die vier sekundären Motivationsaspekte repräsentieren verschiedene Teilaspekte der Anstrengungsbereitschaft.

Die Erfassung der Bearbeitungszeit für die Lösung der Testaufgaben kann als eine Dimension der Testmotivation betrachtet werden, wie die oben zitierte Untersuchung zeigt. Zugleich ist unmittelbar evident, dass die Bearbeitungszeit auch ein Indikator für das Kompetenzniveau ist. Leistungsstärkere Testteilnehmer, das zeigen die Testergebnisse, nutzen die zur Verfügung stehende Testzeit (120 Minuten) zu einem höheren Grad aus als leistungsschwächere Testteilnehmer. Die Bearbeitungszeit gilt daher sowohl als ein Indikator für das Kompetenzniveau als auch für die Testmotivation.

In diesem Forschungsbericht werden die Ergebnisse zur Testmotivation von Elektronikern des ersten Haupttestes zusammengefasst. Anschließend werden die Ergebnisse zu einer explorativen Faktorenanalyse zum Modell der Motivationsstruktur am Beispiel des COMET-Projektes „Pflegeausbildung an Höheren Fachschulen“ (Schweiz) dargestellt und auf den Einfluss der Bearbeitungsdauer auf die Testleistung eingegangen. Abschließend werden die Untersuchungsergebnisse zum Zusammenhang zwischen der Testmotivation und der Entwicklung beruflicher Kompetenz dargestellt. Dabei wird auch nach den Motivationsaspekten differenziert.

Fragebogen: Fragen an die Auszubildenden

Name, Vorname: _____

Liebe Studierende,

wir möchten von Ihnen gerne erfahren, wie Sie die von Ihnen bearbeitete Testaufgabe einschätzen. Dazu bitten wir Sie um einige Angaben. Anschließend legen Sie diesen Bogen bitte zu Ihrer Aufgabenlösung in den bereitgelegten Umschlag.

Vielen Dank für Ihre Mitarbeit!

Wie lange haben Sie an der Testaufgabe gearbeitet?

- weniger als ½ Stunde
- ½ - 1 Stunde
- 1 – 1½ Stunden
- 1½ - 2 Stunden

	<i>stimme überhaupt nicht zu</i> (1)	<i>stimme eher nicht zu</i> (2)	<i>unent- schieden</i> (3)	<i>stimme eher zu</i> (4)	<i>stimme voll und ganz zu</i> (5)
Die Bearbeitung der Testaufgabe fand ich sehr interessant.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Diese Art von Testaufgaben finde ich sehr nützlich.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die Testaufgabe hat sehr viel mit meinem Beruf zu tun.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ich habe die Testaufgabe sehr konzentriert bearbeitet.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ich habe die Testaufgabe sehr sorgfältig bearbeitet.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Bei der Bearbeitung der Testaufgabe habe ich mir sehr viel Mühe gegeben.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Bei Dingen, die Ihnen persönlich sehr wichtig sind, strengen Sie sich auch einmal besonders an und geben Ihr Bestes (z.B. Sport, Hobbys,...).

Wie sehr haben Sie sich im Vergleich dazu bei der Testaufgabe angestrengt? (Bitte ankreuzen!)

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

1 = minimale Anstrengung

10 = maximale Anstrengung

Abb. 7: Itemstruktur und Testmotivation

4. Testmotivation von Elektronikern

Im Rahmen des ersten Haupttests der Elektroniker für Betriebstechnik (E-B) sowie der Elektroniker der Fachrichtung Energie- und Gebäudetechnik (E-EG) (NRW) wurde ebenfalls die Testmotivation erfasst. Vergleicht man die Testmotivation der beiden Testgruppen miteinander, dann ergeben sich deutliche Unterschiede in der Motivationsausprägung. (Abb. 8 bis Abb. 10).

Abweichend von der verbreiteten These, dass es keinen signifikanten Zusammenhang zwischen den Kompetenzniveaus und der Testmotivation gebe, zeigt sich bei diesen beiden Testgruppen (EB: $N = 170$; EEG: $N = 192$) ein ausgeprägter Zusammenhang zwischen Testmotivation und Kompetenzausprägung. Besonders deutlich ist dieser Zusammenhang bei den EEG-A für das Item „Interesse an der Aufgabenbearbeitung“ (Abb. 8).

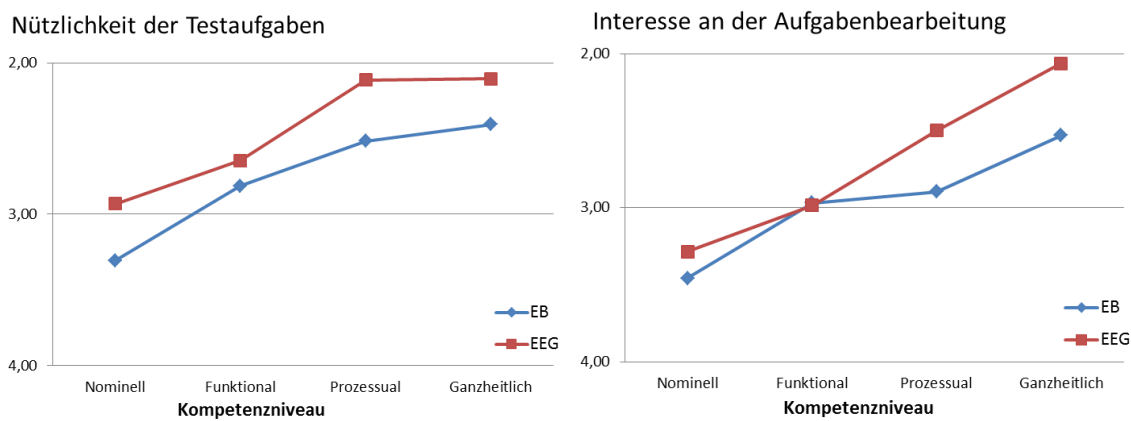


Abb. 8: Vergleich von EB und EEG nach Interesse und Bewertung der Nützlichkeit- Legende: 2= stimme eher zu / 3= unentschieden / 4= stimme eher nicht zu (Kategorien 1= stimme voll und ganz zu und 5= stimme überhaupt nicht zu sind nicht belegt)

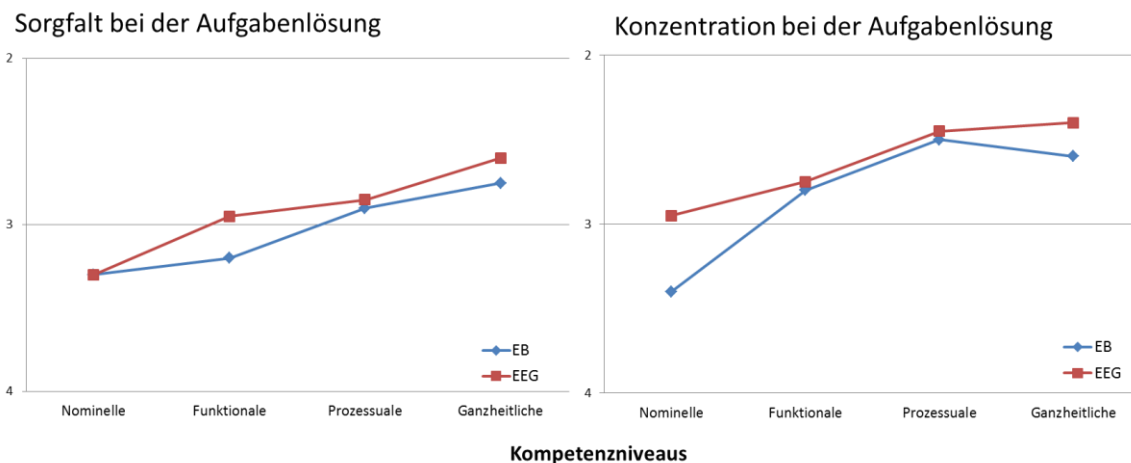


Abb. 9: Vergleich von EB und EEG nach Sorgfalt und Konzentration - Legende: 2= stimme eher zu / 3= unentschieden / 4= stimme eher nicht zu (Kategorien 1= stimme voll und ganz zu und 5= stimme überhaupt nicht zu sind nicht belegt)

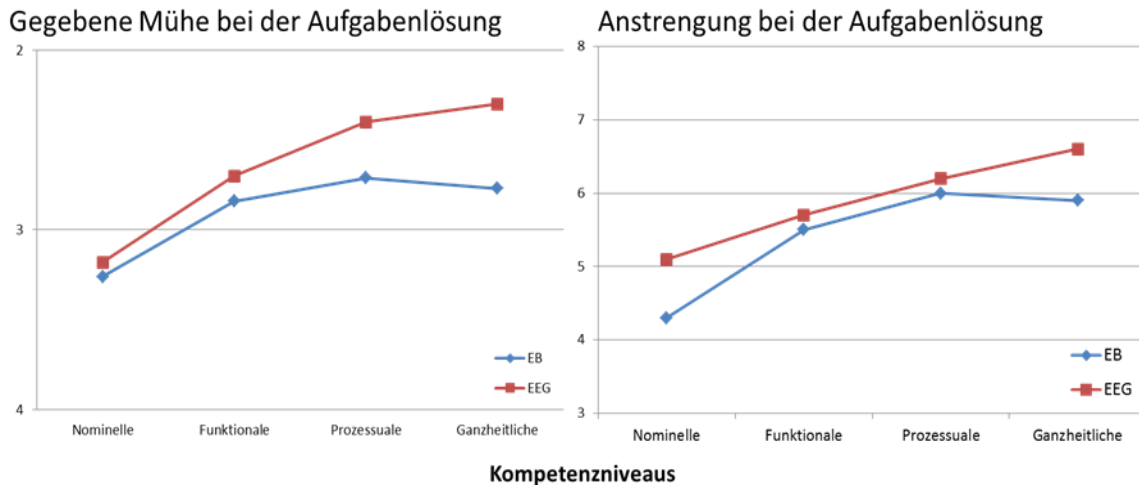


Abb. 10: Vergleich von EB und EEG nach gegebener Mühe und Anstrengung bei der Lösung der Testaufgaben
 Legende: 2= stimme eher zu / 3= unentschieden / 4= stimme eher nicht zu (Kategorien 1= stimme voll und ganz zu und 5= stimme überhaupt nicht zu sind nicht belegt)

3. Explorative Faktorenanalyse zum Zusammenhang zwischen Testmotivation und Testleistung

Das Testmotivationsmodell: Datenstruktur der motivationalen Variablen im KOMET-Testverfahren

Bei den bisherigen KOMET-Projekten wurde die Testmotivation auf der Grundlage der einzelnen Items (Abb. 7) analysiert. Um zu überprüfen, ob sich Zusammenhänge zwischen den beobachtbaren Motivationsaspekten durch übergeordnete Dimensionen erklären lassen, wurde eine exploratorische Faktorenanalyse durchgeführt. Diese erlaubt es, nicht-beobachtbare (latente) Dimensionen aufzudecken, die den beobachtbaren Items übergeordnet werden können.

Das Modell der Testmotivation basiert auf der Hypothese, dass ein von den Auszubildenden wahrgenommener Berufsbezug, Nutzen und Interesse in den vorgelegten Testaufgaben zu einer sorgfältigen, konzentrierten Bearbeitung führt.

Stichprobe:

Die Daten wurden im Rahmen einer KOMET-Testung von Auszubildenden des 2. und 3. Ausbildungsjahres zur Pflegefachkraft aus insgesamt sechs Standorten in der Schweiz erfasst (Standorte: Aarau, Basel, Bern, Luzern, Solothurn, Zürich/Winterthur). An der Befragung nahmen insgesamt N=477 Personen teil, wovon 87 Prozent der Befragten weiblich (n=417) waren.

Die eingesetzten Items des Motivationsfragebogens (Abb. 7) werden folglich einer explorativen Faktorenanalyse unterzogen, welche die den Items zugrundeliegende Strukturen aufdecken soll. Aufgrund des korrelativen Charakters der Items wird die Annahme verfolgt, dass mögliche Faktoren ebenso miteinander korrelieren. Entsprechend wird eine oblimine (direk-

te) Rotation für die Faktorenanalyse angewandt, die einen möglichen Zusammenhang zwischen den Faktoren weithin zulässt.

Im Ergebnis können zwei Faktoren bzw. Motivationsdimensionen extrahiert werden. Die Faktorladungen finden sich in Tab. 2.

Item	Itemkennwerte			Faktorladungen	
	M	SD	r_{it}	1	2
Mühe	2,48	1,00	0,78	0,89	0,01
Sorgfalt	2,44	0,95	0,74	0,87	-0,01
Konzentration	2,58	1,01	0,71	0,82	-0,03
Anstrengung	4,86	2,22	0,67	0,74	0,04
Interesse	2,44	1,04	0,65	0,03	0,83
Nutzen	2,37	1,02	0,6	0,02	0,77
Berufsbezug	2,99	1,11	0,36	-0,02	0,49

Anmerkungen: Faktorladungen mit $>.30$ sind fett markiert; Bartlett-Test: $\chi^2=1709.19(df=21)$, $p<.001$; Kaiser-Meyer-Olkin-Maß=.86. N = Stichprobengröße; M = Mittelwert, SD = Standardabweichung; r_{it} =Trennschärfe.

Tab. 2: Ergebnisse der Faktorladungen der motivationalen Items auf die extrahierten Faktoren (Daten: Pflegefachkräfte Schweiz 2014, $N=477$)

Faktor 1 setzt sich dabei aus den Items Mühe, Sorgfalt, Konzentration und Anstrengung zusammen. Auf den Faktor 2 laden die Aussagen zum Interesse, Nutzen und Berufsbezug. Mit Blick auf die inhaltliche Formulierung der Items (vgl. Abb. 7) lässt sich Faktor 2 als *Sinnhaftigkeit (primäre Motivation)* erkennen. Der Faktor beschreibt den in den Testaufgaben erkannten Nutzen für die berufliche Zukunft und verknüpft ein Interesse damit. Faktor 1 beschreibt das Verhalten bei der Bearbeitung der Testaufgaben und wird mit *Investition (sekundäre Motivation)* bezeichnet. Der Begriff *Investition* bezieht sich dabei auf die während der Bearbeitung eingesetzten motivationalen Fähigkeiten.

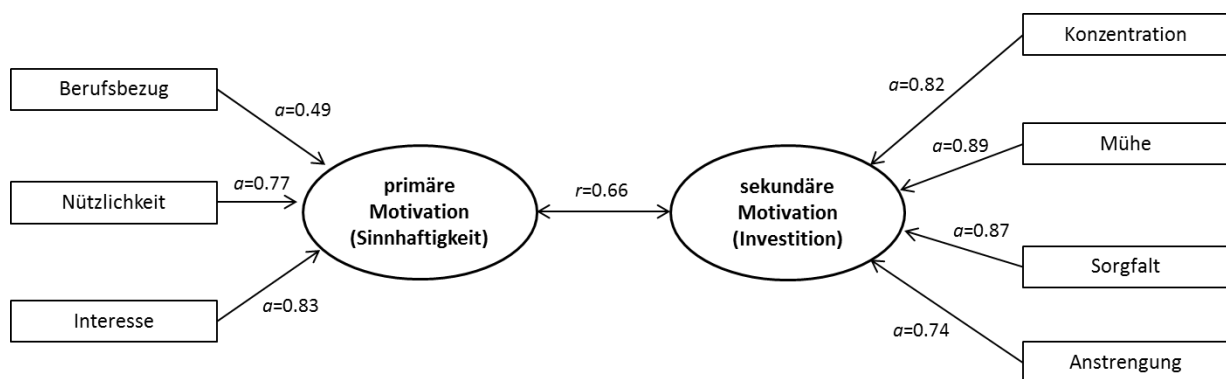


Abb. 11: Ergebnisse der explorativen Faktorenanalyse (Daten: Pflegefachkräfte Schweiz 2014, $N=477$). r = Korrelationskoeffizient; a = Faktorladung.

Die definierten Faktoren hängen positiv mittelstark zusammen ($r=0.66$). Beide Faktoren erklären insgesamt eine Varianz von 62 Prozent.

Zum Zusammenhang zwischen Testmotivation und Testleistung

Bereits die oben zitierten Untersuchungen zeigten Hinweise auf einen Zusammenhang zwischen Testmotivation und -leistung. Mit Blick auf die inhaltliche Dimension der beiden Faktoren liegt die Vermutung nahe, dass die Sinnhaftigkeit als primäre Motivationsdimension und die Investition als sekundäre Motivationsdimension fungieren, da vermutlich Personen, die in der Testung Sinn empfinden auch mehr in die Testbearbeitung investieren. Um diesen Zusammenhang zwischen den beiden extrahierten Motivationsdimensionen sowie den Zusammenhang dieser mit der Testleistung weiterfassend zu untersuchen, wurde die im Folgenden beschriebene Mediatoranalyse durchgeführt.

1. Fragestellung und Hypothese

Auf der Grundlage des Ergebnisses der oben beschriebenen Faktorenanalyse wurde untersucht, ob die beiden Motivationsdimensionen einen Erklärungsbeitrag für die KOMET-Testleistung (gemessen als Gesamtpunktwert (GPW)) leisten.

Zur statistischen Untersuchung der Fragestellung wurde folgende Hypothese aufgestellt:

Die Motivationsfaktoren Sinnhaftigkeit und Investition erklären kausal die Kompetenzleistung im KOMET-Testverfahren. Der Faktor Investition vermittelt dabei als Mediator den Zusammenhang zwischen Faktor Sinnhaftigkeit und Leistung (GPW) im KOMET-Testverfahren.

2. Methode

Um die Hypothese statistisch zu untersuchen wurde mit den drei Variablen Sinnhaftigkeit als unabhängige Variable, Investition als Mediatorvariable und GPW als abhängige Variable eine Mediatoranalyse durchgeführt. Die Analyse erfolgte in den für eine Mediatoranalyse üblichen vier Schritten (PREACHER & HAYES 2004), in denen verschiedene lineare Regressionsanalysen berechnet wurden. In Schritt eins wurde die Regression von GPW auf Sinnhaftigkeit untersucht. Im zweiten Schritt der Analyse wurde eine Regression von Investition auf Sinnhaftigkeit untersucht und im dritten Schritte die Regression von Gesamtpunktwert auf Investition. Im letzten Schritt der Analyse wurde eine multiple Regression von Gesamtpunktwert auf Sinnhaftigkeit und Investition untersucht.

Zur Überprüfung der statistischen Signifikanz eines vorgefundenen Mediatoreffektes wurde zusätzlich der Sobeltest durchgeführt (PREACHER & HAYES 2004).

3. Ergebnisse

In Abb. 12 finden sich die Ergebnisse der Mediatoranalyse. Die Analyse ergab, dass mit einem *korrigierten* $R^2 = .043$ durch die Variable Sinnhaftigkeit 4.3 Prozent der Varianz im Gesamtpunktwert erklärt werden kann. Auch wenn dieser Varianzanteil gering ausfällt, wird das Regressionsmodell aus Schritt eins mit $F = 22.13(1;475)$, $p < .001$ signifikant. Der Zusammenhang von Sinnhaftigkeit und Gesamtpunktwert erweist sich in diesem Modell als signifikant und positiv ($b_1 = 4.314.70$; $p < .001$). Die Ergebnisse der Regression aus Schritt zwei zeigen an, dass mit $F = 210.48(1;477)$ und $p < .001$ eine signifikante Regression vorliegt. Dieses Modell erklärt 30.5 Prozent der Varianz von Investition auf. Der Zusammenhang von Sinnhaftigkeit und Investition ist positiv und signifikant ($b_2 = .60$; $p < .001$). Die Regression

des dritten Analyseschrittes ergab mit $F = 32,53(1;475)$ und $p < .001$ ein signifikantes Ergebnis. Das Modell erklärt 6.4 Prozent der Varianz des Gesamtpunktwertes. Der Zusammenhang von Investition und Gesamtpunkt看wert wird ebenfalls signifikant und fällt positiv aus ($b_3 = 4.79$; $p < .001$).

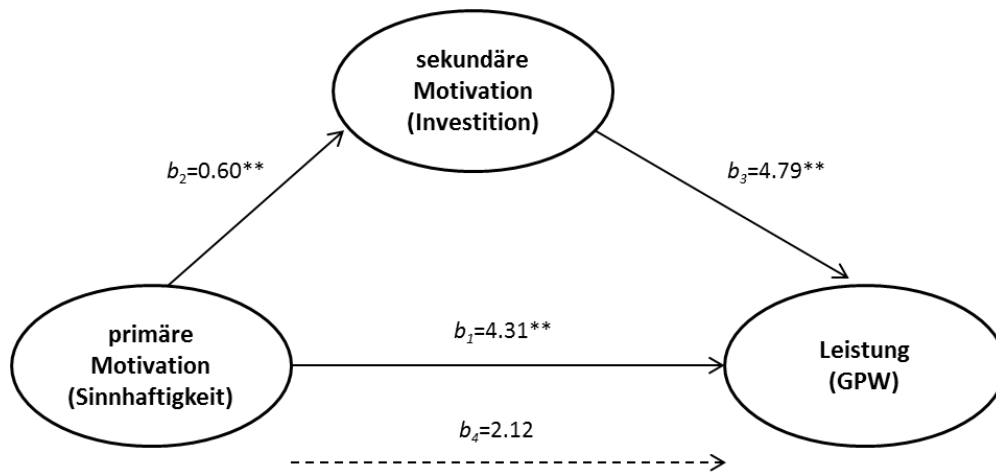


Abb. 12: Mediatoranalyse zur Bestimmung des Zusammenhangs von primärer und sekundärer Motivation und der Leistung im KOMET Testverfahren (Daten: Pflegefachkräfte Schweiz 2014; $N=477$). b_i = Regressionskoeffizient; * $p < 0.05$; ** $p < 0.01$.

Die Regression aus Schritt vier zeigt an, dass mit 6.8 Prozent ein signifikantes Maß an Gesamtvarianz des Gesamtpunktwertes durch die beiden Variablen Sinnhaftigkeit und Investition erklärt werden kann ($F = 18.28(2;474)$, $p < .001$). Der Zusammenhang zwischen Investition und Gesamtpunkt看wert bleibt positiv und signifikant ($b_4 = 3.72$; $p < .001$), während der Zusammenhang zwischen Sinnhaftigkeit und Gesamtpunkt看wert nicht mehr signifikant wird ($b = 2.12$; $p = .051$). Der Durbin-Watson-Wert dieses Modells liegt bei 0.74, somit muss von einer starken positiven Autokorrelation (1. Ordnung) der Residuen ausgegangen werden (BROSIOUS 2013).

Der Sobel Test wurde mit Sobel-Teststatistik = 5.31, $p < .001$ in der vorliegenden Untersuchung signifikant.

4. Diskussion

Die Ergebnisse der Mediatoranalyse weisen darauf hin, dass die Investition den Zusammenhang zwischen Sinnhaftigkeit und Leistung vollständig mediiert. Das Ergebnis des Sobel Tests bestätigt dieses Resultat. Dies bedeutet, dass Personen, die mehr Sinn im KOMET-Testverfahren sehen, mehr investieren und auf diese Weise eine bessere Leistung erzielen. Somit kann die Investition als primäre Motivation und die Investition als sekundäre Motivation bestätigt werden. Die in früheren KOMET-Untersuchungen entstandene Annahme, dass stärker motivierte Personen auch bessere Testleistungen erzielen, wird hierdurch bestätigt. Allerdings wird durch die Analyse deutlich, dass die Bereitschaft zur Investition in die Testaufgabe davon abhängt, wie stark die Sinnhaftigkeit der Testung empfunden wird. Für die Zukunft sollte daher den Testteilnehmern vor Testdurchführung die Sinnhaftigkeit des KOMET-Testverfahrens nahegelegt werden, um schlechte Testleistungen zu vermeiden, die aufgrund von mangelhaft empfundener Sinnhaftigkeit resultieren.

Die relativ geringe Varianzaufklärung des Gesamtmodells von 6.8 Prozent zeigt an, dass ein Großteil der Varianz der Leistung durch das vorliegende Modell noch unerklärt bleibt. Dies ist ein Hinweis darauf, dass neben den Motivationskomponenten noch andere Faktoren einen Einfluss auf die Testleistung haben. Dies ist unmittelbar einleuchtend, da Leistungen selbstverständlich vor allem vom Wissen sowie den Fähigkeiten und Fertigkeiten abhängen. Einen weiteren Hinweis für diese Annahme liefert das Ergebnis des Durbin-Watson-Tests, welches auf eine starke positive Autokorrelation (1. Ordnung) der Residuen hinweist. Dies könnte ein Hinweis dafür sein, dass wichtige erklärende Variablen im berechneten Modell fehlen. Die Folge von starker Autokorrelation der Residuen kann sein, dass die wahren Standardfehler unterschätzt werden. Dies verzerrt weiterhin die Ergebnisse der Signifikanztests (BROSIUS 2013). Die vorliegenden Ergebnisse müssen entsprechend mit Vorsicht interpretiert werden.

Der Einfluss der Bearbeitungsdauer auf die Gesamtleistung

Ausgehend von der relativ geringen Varianzaufklärung der Leistung durch die beiden Motivationsfaktoren, welche sich in der Mediatoranalyse gezeigt hat, könnte die Variable der Bearbeitungsdauer zu einer Steigerung der Varianzaufklärung führen. Ausgehend von der Hypothese, dass sich eine umfassende, reflektierte Aufgabenlösung mit ausführlichen Begründungen (entsprechend der KOMET-Aufgabenstellung) zwangsläufig in einer höheren Bearbeitungsdauer niederschlägt, wird die Bearbeitungszeit auch nach Kürzung des KOMET-Tests um eine Aufgabe erfasst, um diesen Aspekt als Indikator der Testmotivation weiterhin untersuchen zu können.

Die auf einer vierstufigen Skala erfragte Bearbeitungszeit für den Test stellt sich in der vorliegenden Stichprobe folgendermaßen dar:

<u>Wie lange haben Sie an der Testaufgabe gearbeitet?</u>		
	Anzahl	Prozent
(1) weniger als 1/2 Stunde	16	3,4
(2) 1/2 - 1 Stunde	83	17,4
(3) 1 1/2 Stunde	127	26,6
(4) 1 1/2 - 2 Stunden	236	49,5
fehlend	15	3,1
gesamt	477	100

Es zeigt sich, dass nahezu die Hälfte der Auszubildenden die Bearbeitungszeit ausschöpfte. Eine Korrelation der Bearbeitungsdauer mit dem Gesamtpunktwert von $r=0.53$ ergab ein signifikantes Ergebnis. Damit liegt ein mittelstarker positiver Zusammenhang zwischen der Bearbeitungsdauer des KOMET-Tests und dem erreichten Gesamtpunktwert vor.

Vertiefend soll nun der Frage des Zusammenspiels zwischen den motivationalen Aspekten der Sinnhaftigkeit und Investition, der Bearbeitungsdauer und dem Gesamtpunktwert nachgegangen werden. Das Forschungsinteresse zielt hier insbesondere auf die Untersuchung des Einflusses der Bearbeitungsdauer neben den beiden Motivationsdimensionen auf die Leistung der Auszubildenden.

Dazu werden die Faktoren Investition und Sinnhaftigkeit sowie die Bearbeitungsdauer in einer hierarchischen Regressionsanalyse mit dem Gesamtpunktwert als abhängige Variable untersucht. Im Ergebnis zeigt sich eine bedeutende zusätzliche Varianzaufklärung von 22% durch die Variable Bearbeitungsdauer (vgl. Tab. 3).

Variable	B	SE	β	R ²	Korrigiertes R ²	p
<i>1. Schritt</i>						
Sinnhaftigkeit	2.47	1.12	0.12	0.077	0.073	.028
Investition	3.62	1.03	0.19			.001
<i>2. Schritt</i>						
Sinnhaftigkeit	1.30	0.99	0.06	0.296	0.291	.187
Investition	2.69	0.91	0.14			.003
Bearbeitungsdauer	19.02	1.60	0.48			< .001

Tab. 3: Zusammenfassung der hierarchischen Regressionsanalyse zur Vorhersage der Variable „Gesamtpunktwert“ (n=462).

Anmerkung: abhängige Variable: Gesamtpunktwert; Variable Dauer wurde dichotomisiert: (0)=0-1 Stunde, (1)=1-2 Stunden).

Der Mittelwert der Leistung (Gesamtpunktwert) liegt in der vorliegenden Stichprobe bei $M=46,9$ ($SD=16,5$).

Tab. 3 zeigt anhand der korrigierten R²-Werte den Anteil der aufgeklärten Varianz des Leistungswertes durch die beeinflussenden Faktoren. Zu beobachten ist eine sehr kleine Varianzaufklärung des Gesamtpunktwertes durch die Faktoren Sinnhaftigkeit und Investition (R²= .07). Dennoch zeigt das Regressionsmodell aus Schritt 1 eine statistische Signifikanz mit $F(2, 459)= 19.22$; $p < .001$). Die Aufnahme der dummy-kodierten Variablen Bearbeitungsdauer im Modell des Schrittes 2 führt zu einem bedeutenden Zuwachs an Varianzaufklärung in Höhe von 22 Prozent. Jedoch verliert der Faktor Sinnhaftigkeit dadurch an Bedeutung und kann keinen signifikanten Erklärungsbeitrag in diesem Modell leisten. Auch für das Regressionsmodell im Schritt 2 ist eine statistische Signifikanz festzustellen ($F(2, 459)=64.09$; $p < .001$).

Erhöht sich der Faktor Investition um eine Einheit bei gleichzeitiger Kontrolle der Faktoren Sinnhaftigkeit und Bearbeitungsdauer, dann erhöht sich der Gesamtpunktwert um 2,7 Punkte. Folglich führt eine Erhöhung der sekundären Motivationsaspekte (Faktor Investition) zu höheren Leistungswerten. Für die Variable Bearbeitungsdauer kann aus dem Modell 2 formuliert werden: Die Auszubildenden, die 1-2 Stunden für die Bearbeitung der Testaufgabe investierten, erreichten bei Kontrolle der Faktoren Sinnhaftigkeit und Investition einen um 19 Punkte höheren Gesamtpunktwert als Auszubildende, die sich weniger als 1 Stunde in der Bearbeitung der Testaufgabe engagierten.

Die Dauer der Bearbeitungszeit ist damit zu einem wichtigen Einflussfaktor auf den Leistungswert im Zusammenspiel mit dem Faktor Investition geworden.

4. Testmotivation der Pflege-Studierenden (Schweiz)

Die Testmotivation der Pflege-Studierenden ist überdurchschnittlich hoch und variiert zwischen den Standorten (Bildungszentren) in Bezug auf das Kriterium der „Anstrengung“ zwischen 6,1 und 7,5 (bei einer Skala von 1-10) (Abb. 13).

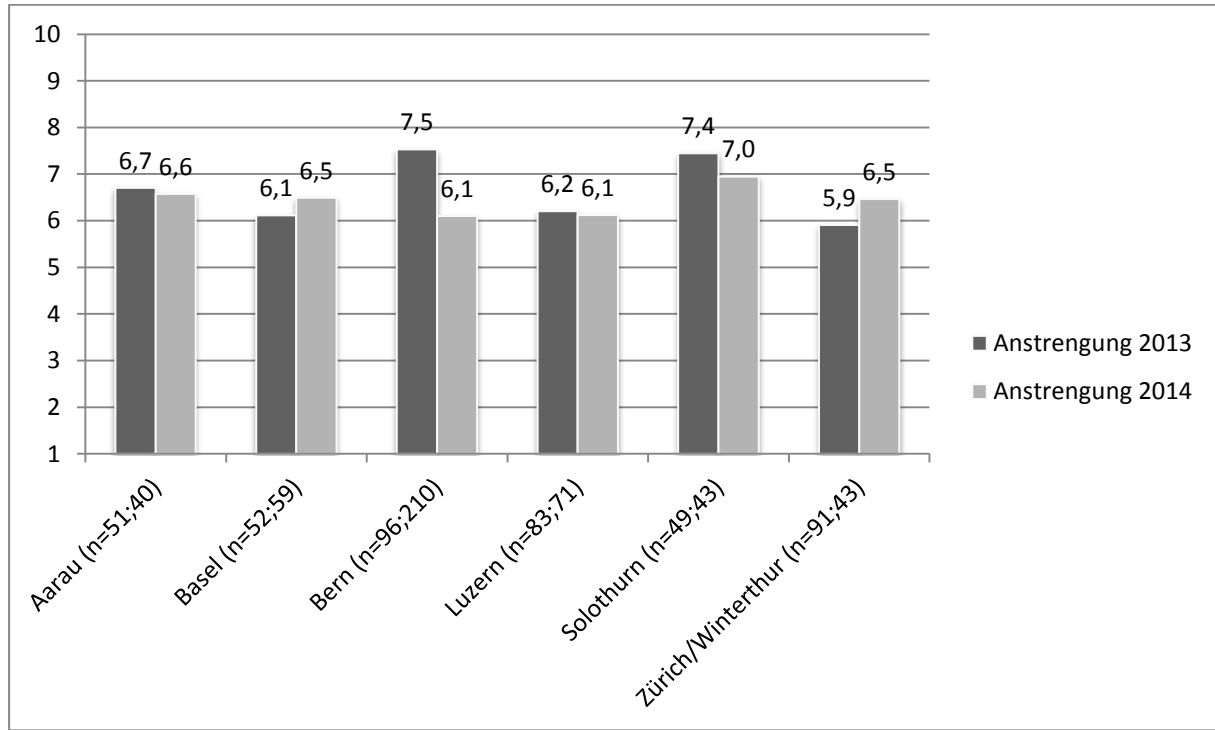


Abb. 13: Kriterium "Anstrengung" - Testmotivation der Pflege-Studierenden

Dabei zeigt sich, dass sich die Anstrengungsbereitschaft an den einzelnen Standorten im Laufe eines Jahres bei den Testteilnehmern zwischen den Bildungszentren verschoben hat.

Verbindet man die Werte der primären und sekundären Motivationsaspekte zu einem Motivationsprofil, dann fallen vier Besonderheiten auf.

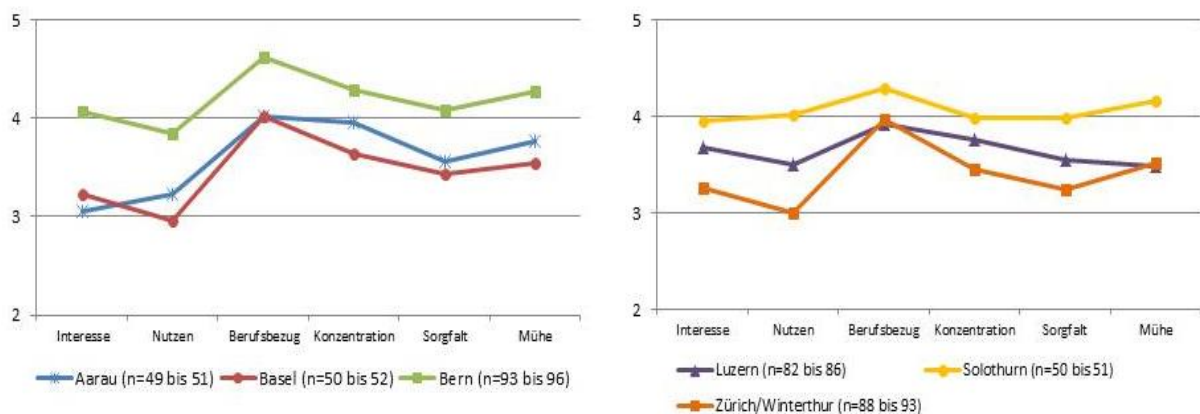


Abb. 14: Testmotivation der unterschiedlichen Standorte, 1. Haupttest 2013.

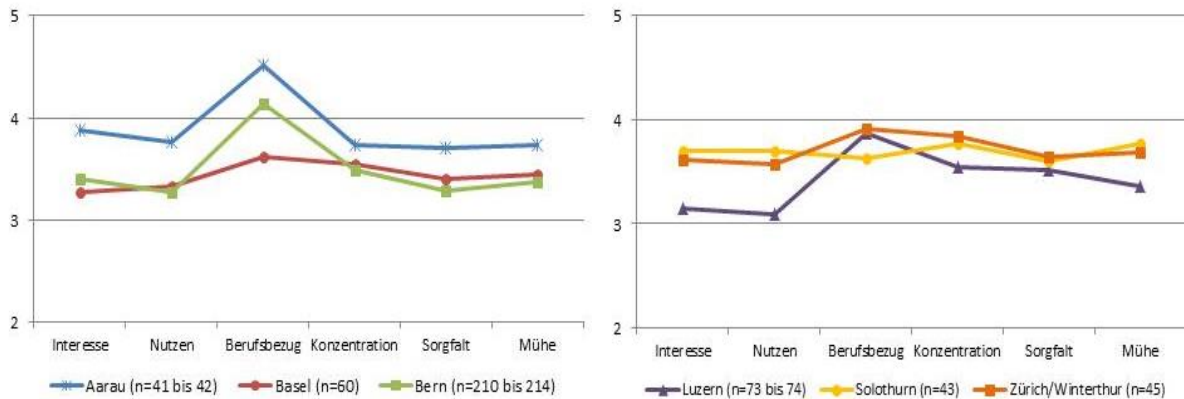


Abb. 15: Testmotivation der unterschiedlichen Standorte, 2. Haupttest 2014

- An vielen Standorten bewerten die Studierenden den *Berufsbezug* der Testaufgaben als hoch bis sehr hoch.
- Es überrascht daher, dass der *Nutzen* des Tests teilweise als deutlich niedriger eingeschätzt wird. Unterstellt man, dass diese Studierenden den Nutzen des Tests (die Bearbeitung der Testaufgaben) auf Ihre Ausbildungssituation beziehen, dann stützt dies die These, dass die Studierenden deutlich unterscheiden zwischen ihrer Ausbildung und dem zu erlernenden Beruf. Eine Aufklärung dieser Differenz kann sich in einem ersten Schritt auf die Auswertung der Kontextdaten zur Ausbildungsqualität stützen (s. u.).
- Das Interesse am Test scheint sich aus dem Berufsbezug und dem Nutzen zu speisen: Die Werte zum Interesse-Aspekt der Motivation liegen daher in vielen Fällen zwischen den Bewertungen dieser beiden anderen primären Motivationsaspekte.
- Die sekundären Motivationsaspekte repräsentieren ganz offensichtlich in großer Übereinstimmung dieselbe Motivationsdimension (s. u.).
- Am Bildungszentrum A ist im Laufe eines Jahres ein deutlicher Anstieg der Testmotivation zu verzeichnen. Dagegen ist die Testmotivation an einem anderen Bildungszentrum (L) abgefallen. Anhand der Kontextdaten war es den Lehrern/-innen möglich im Rahmen eines Feedback Workshops diese Entwicklung aufzuklären. Hier wurden beispielsweise Konflikte innerhalb von Studiengängen oder die Identifikation mit dem Testverfahren als ursächlich für die unterschiedliche Testmotivation der Teilnehmer/innen diskutiert (Evaluationsworkshop der Pflege Schweiz, Januar 2015).

Repräsentation der Faktorenwerte in der Form einer Matrix

Auf der Grundlage der oben beschriebenen Faktorenanalyse lassen sich die Ergebnisse zur Testmotivation in einer Faktor-Matrix zusammenfassen (Abb. 16). Die Faktorwerte geben Aufschluss darüber, ob bei der Testgruppe im Verhältnis zur untersuchten Stichprobe über- oder unterdurchschnittliche Werte erzielt wurden. Dabei repräsentieren Werte >0 überdurchschnittliche und Werte <0 unterdurchschnittliche Motivationsniveaus. **Abb. 16** zeigt, dass die Testteilnehmer der Bildungszentren C und E 2013 eine überdurchschnittliche Testmotivation

– in Bezug auf beide Motivationsdimensionen – haben. Komplementär dazu weisen die Testteilnehmer der Bildungszentren F und B 2013 eine unterdurchschnittliche Testmotivation auf.

Für die beteiligten Bildungszentren ist interessant, wie sich ihre Position in der Faktormatrix vom 1. zum 2. Testzeitpunkt – im Laufe eines Jahres – verändert hat (Abb. 17). Dabei zeigt sich, dass die Heterogenität der Testmotivation insgesamt leicht abgenommen hat und dass sich die Testmotivation an einigen Zentren deutlich verändert hat.

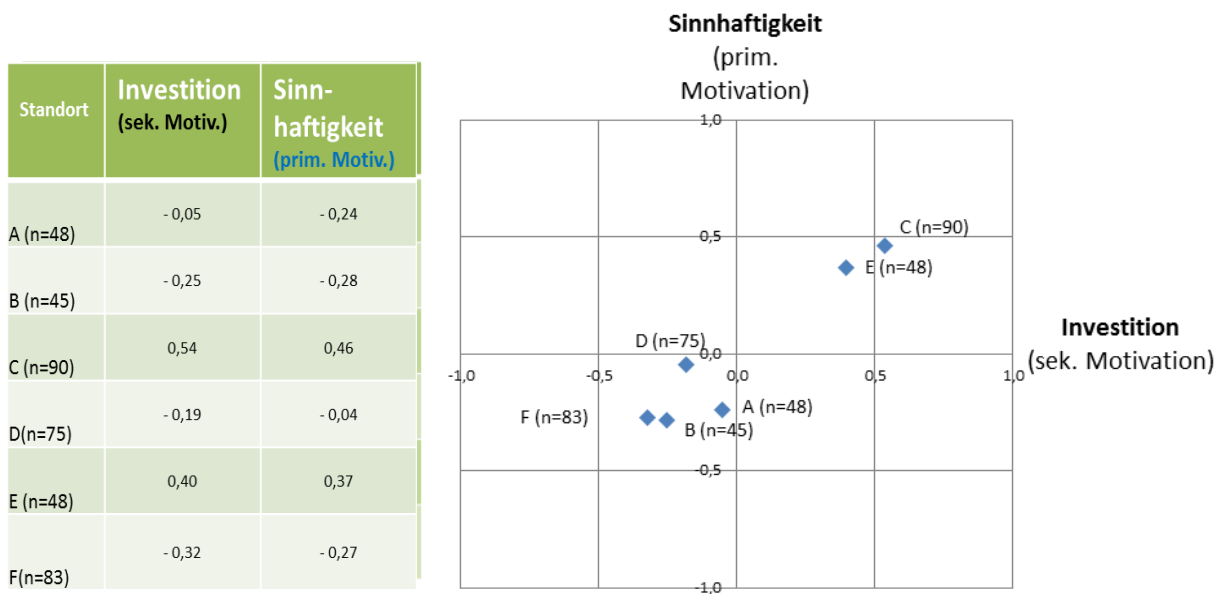


Abb. 16: Faktormatrix Pflegeausbildung Schweiz 2013

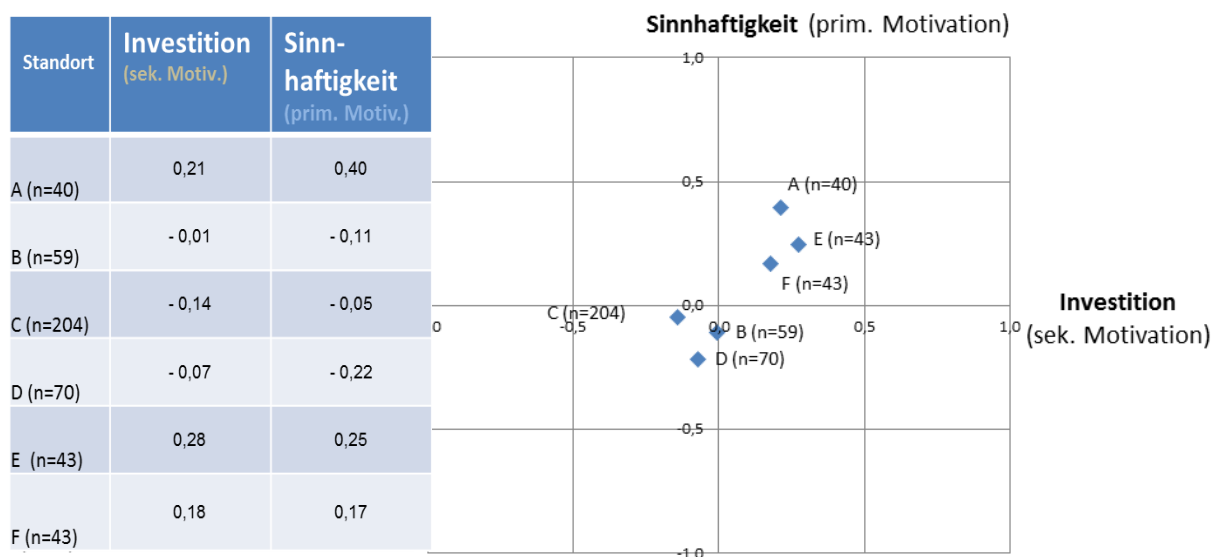


Abb. 17: Faktormatrix Pflegeausbildung Schweiz 2014

Unterschiede zwischen der (KOMET) Test- und der Prüfungsmotivation

Wie oben aufgeführt, wurden Zusammenhänge zwischen Motivation und Testleistung innerhalb des KOMET-Projektes untersucht und die Daten zur Testmotivation standen bei der Interpretation der KOMET-Testergebnisse zur Verfügung. Im Rahmen der ebenfalls in die-

sem Text erwähnten Feedback-Workshops mit den berufsbezogenen Projektgruppen wurden immer wieder Vergleiche zwischen der Motivation bei Prüfungen – einerseits – sowie bei der Beteiligung an KOMET-Tests – andererseits – angestellt. Dabei begründeten erfahrene Prüfer ihre Vermutung, dass die Prüfungsmotivation bei den Auszubildenden deutlich höher sei als die Testmotivation bei den KOMET-Testverfahren. Ein gutes Prüfungsergebnis käme schließlich auch der beruflichen Karriere zugute, während die Beteiligung an einem KOMET-Test allenfalls Einsichten in den Stand der Kompetenzentwicklung vermittele.

Anhand einer Stichprobe von N = 34 bis 35 Prüfungsteilnehmern (Kfz-Mechatroniker), die auch an dem KOMET-Test für KFZ-Mechatroniker mit insgesamt N = 400 Auszubildenden teilgenommen hatten, konnte die Testmotivation mit der Prüfungsmotivation verglichen werden.

Ergebnisse der Vergleichsuntersuchung: Test- und Prüfungsmotivation bei Kfz-Auszubildenden

Abb. 18 zeigt einen auffälligen Unterschied zwischen der Test- und Prüfungsmotivation.

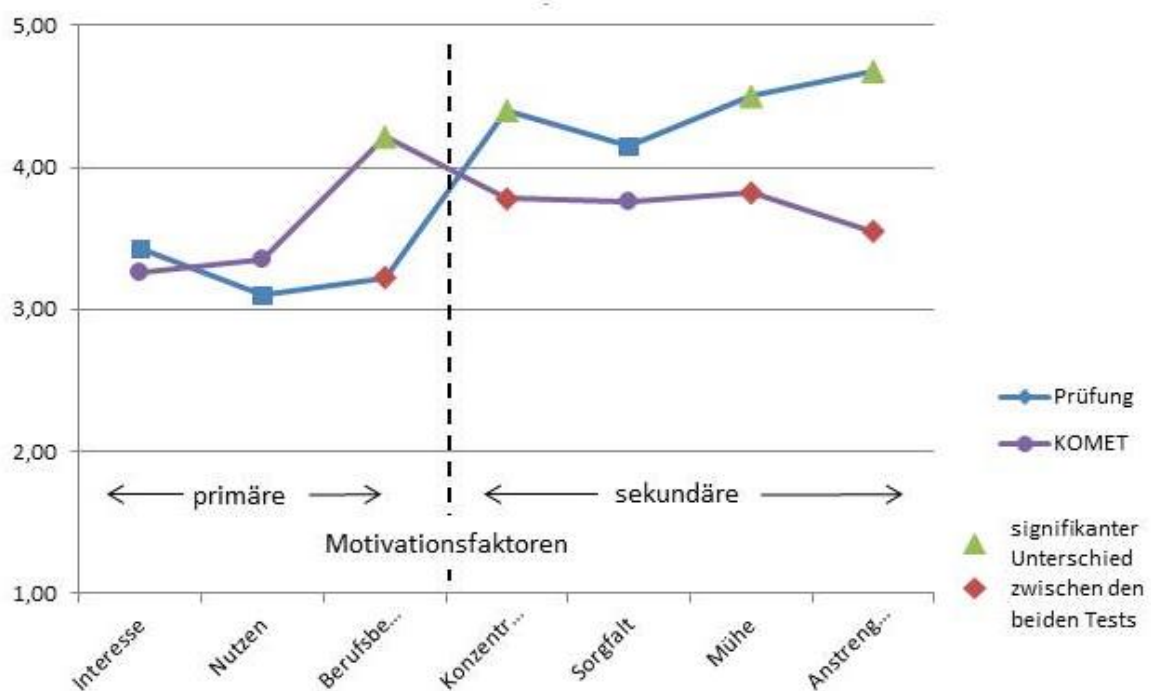


Abb. 18: Primäre und sekundäre Motivationsfaktoren im Vergleich von Prüfungen und KOMET-Tests

Die Bedeutung der primären Motivationsaspekte *Berufsbezug* und *Nutzen* wird für das KOMET-Test Verfahren (deutlich) höher bewertet als für Prüfungen. Besonders auffällig ist die große Differenz für das primäre Motivationskriterien des Berufsbezuges. Die mittlere Bewertung des Berufsbezug wird mit einem Mittelwert von MW = 4,2 für den KOMET-Test dabei signifikant höher eingeschätzt als für den Berufsbezug der Prüfungen (MW = 3,2). Die Bewertung von Prüfungen unter dem Aspekt des Nutzens fällt indifferent mit einem Mittelwert von MW = 3,0 aus. Die Abschlussprüfung hat danach für die Kfz-Auszubildenden weder ei-

nen hohen Berufsbezug noch einen hohen Nutzen (für die Ausbildung und die konkrete Berufstätigkeit).

Komplementär dazu fallen die Werte für die sekundären Motivationsaspekte aus. Bei Prüfungen ist die Anstrengungsbereitschaft der Prüflinge signifikant höher als bei Tests, die im Rahmen der Kompetenzdiagnostik durchgeführt werden (Abb. 18 und Abb. 19).

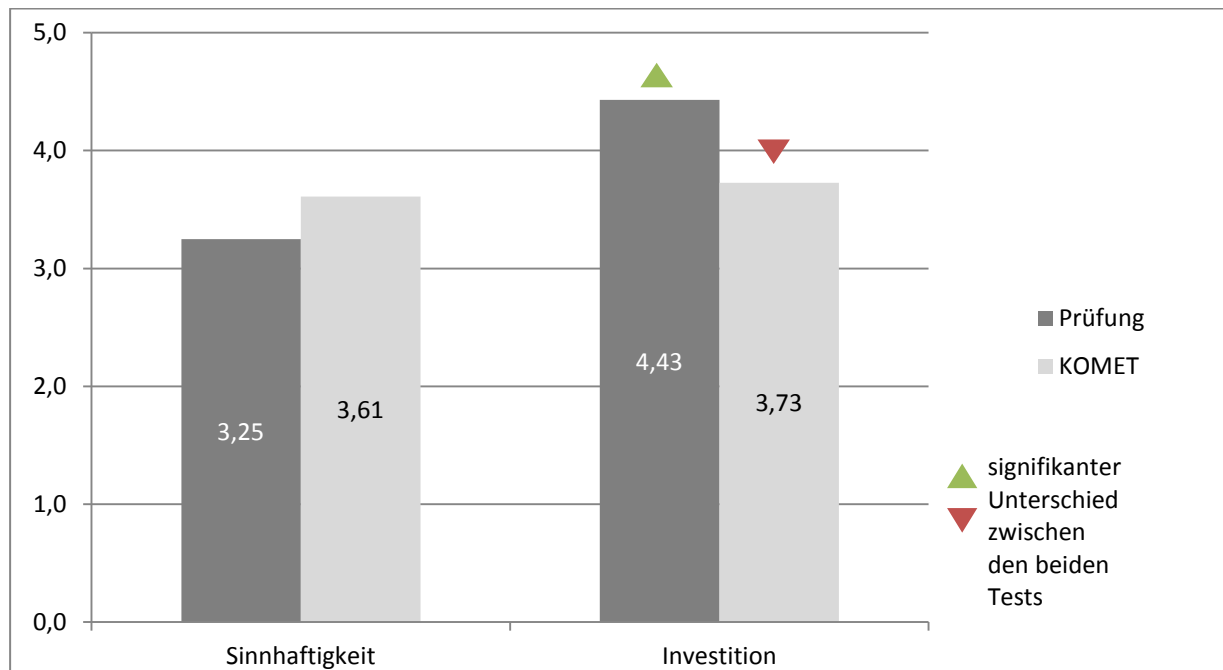


Abb. 19: Komplementäre Motivationsstruktur im Vergleich von Prüfungen und KOMET-Tests

Insgesamt ergibt sich eine *komplementäre Motivationsstruktur* für KOMET-Tests und Abschlussprüfungen. Die primären Motivationsfaktoren (ausgenommen das Interesse) werden für das KOMET-Testverfahren höher bewertet. Darauf basiert eine leicht überdurchschnittliche Motivation in Bezug auf die sekundären Motivationsfaktoren. Mit anderen Worten: die Testmotivation ergibt sich aus der Einschätzung der Testteilnehmer, dass die Testaufgaben einen hohen Berufsbezug und daher auch einen hohen Nutzen für die Ausbildung haben.

Ganz anders fällt das Motivationsprofil für Prüfungen aus. Die sekundären Motivationsfaktoren werden hoch bewertet: im Mittel zwischen $MW = 4,2 - 4,5$. Diese hohen Motivationswerte für die Indikatoren Sorgfalt, Konzentration, Mühe und Anstrengung stehen in einem deutlichen Widerspruch zu den indifferenten Bewertungen der primären Motivationsfaktoren. Die Einsicht der Auszubildenden, dass ein gutes Prüfungsergebnis für die berufliche Karriere vorteilhaft ist, prägt ganz offensichtlich die hohe Prüfungsmotivation. Die hohe Prüfungsmotivation wird auch dadurch nicht gemindert, dass der Berufsbezug und der Nutzen der Prüfung als indifferent eingeschätzt werden.

Diese Fallstudie in einem stark nachgefragten Ausbildungsberuf stützt das Interesse der Organisationen der Arbeitswelt an einer Anwendung des KOMET-Kompetenz- und Messmodells für die Verfahren der Abschlussprüfungen.

5. Die kulturelle Dimension der Testmotivation

Am Beispiel des KOMET-Projektes Südafrika (Elektroniker) lässt sich zeigen, dass die nationalen Arbeitsmarkt- und Ausbildungsstrukturen – der kulturelle Kontext – eine ganz entscheidende Determinante der Test- und Ausbildungsmotivation der Auszubildenden ist. Das Beispiel KOMET (Elektroniker) Südafrika bestätigt diese These deshalb besonders eindrucksvoll, da alle beteiligten Bildungsgänge ein – im internationalen Vergleich (Deutschland, China, Südafrika) – nur niedriges Kompetenzniveau erreicht haben (KOMET RSA Study 2013). Zugleich waren die Testteilnehmer sehr hoch motiviert. Zusammenfassend heißt es im Projektbericht:

„The test takers were highly motivated and interested in the test tasks. Still, the test results are often below the level of functional competence. Professional and holistic shaping competence have rarely been reached. On the other hand, the South African learners were very motivated to take the test and are very committed to their learning in general, as well.“ (ebd., 44)

Sowohl die Berufsbezogenheit und die Nützlichkeit der Testaufgaben wurden von den Testteilnehmern hoch eingeschätzt (Abb. 20 und Abb. 21).

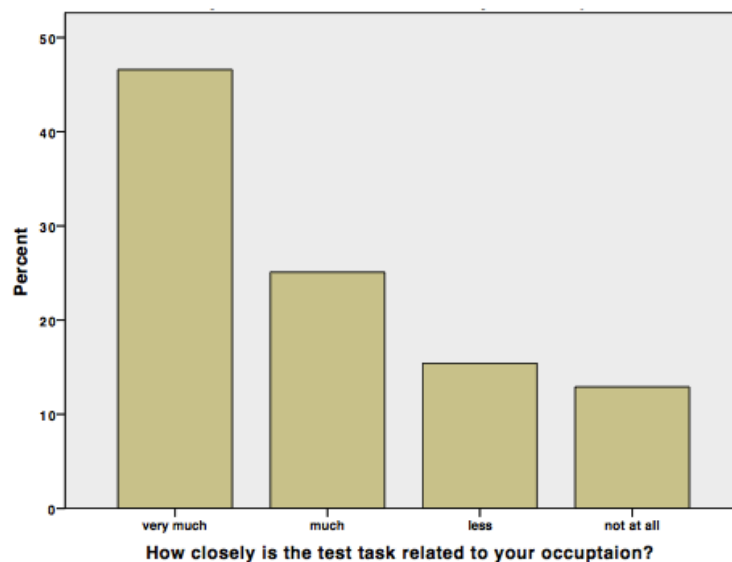


Abb. 20: Bewertung des Berufsbezugs der Testaufgaben (KOMET RSA 2013)

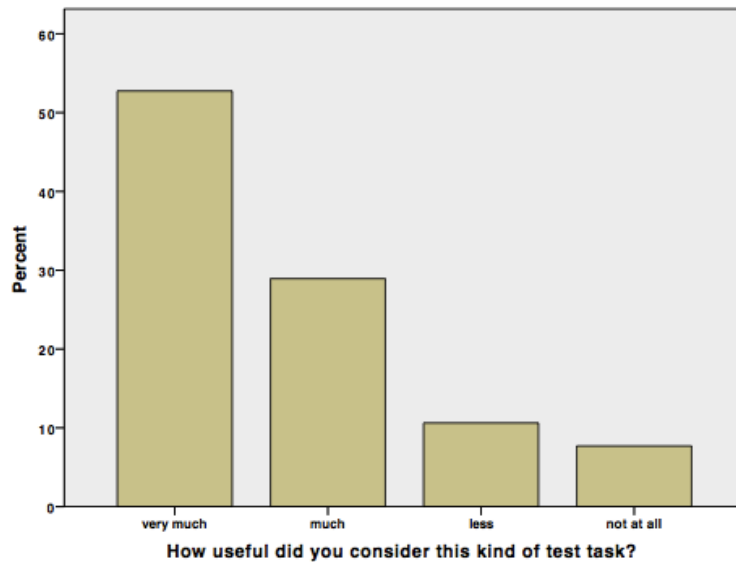


Abb. 21: Bewertung der Nützlichkeit der Testaufgaben (KOMET RSA 2013)

Bei der Frage nach dem Interesse an der Bearbeitung der Testaufgaben gaben 53 % der Befragten an, ein sehr hohes und 34 % ein hohes Interesse zu haben. Nur 10 % gaben an, ein weniger hohes und weitere 10 % kein Interesse an den Testaufgaben zu haben.

Dementsprechend hoch war die Anstrengung, mit der die Testaufgaben bearbeitet wurden.

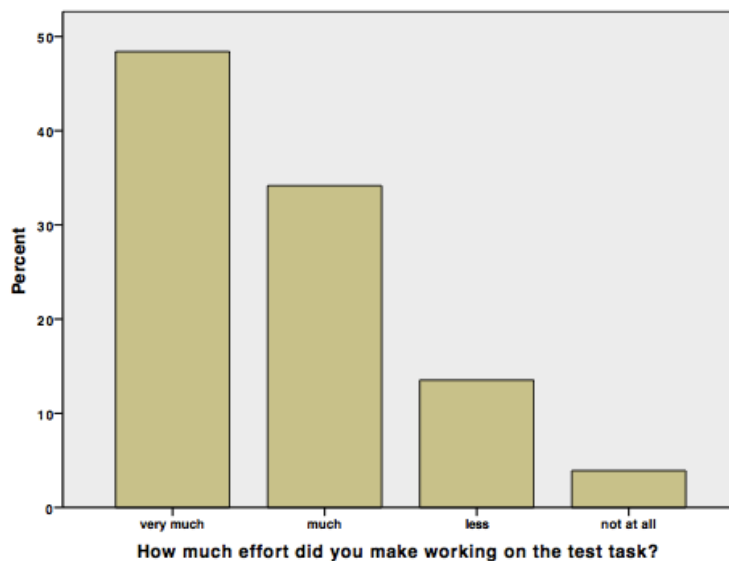


Abb. 22: Bewertung der aufgewandten Anstrengung für die Testaufgaben (KOMET RSA 2013)

Diese hohen Werte der Testmotivation korrelieren mit der Identifizierung mit den Ausbildungsbetrieben.

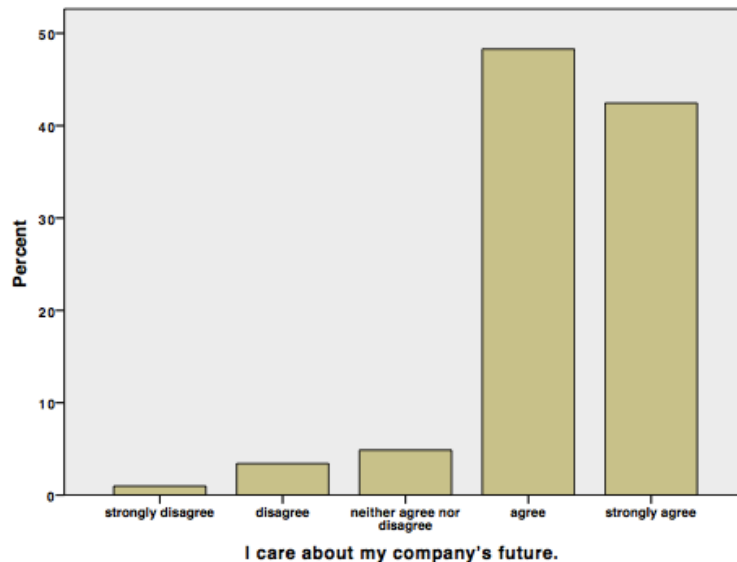


Abb. 23: Identifikation mit den Ausbildungsbetrieben

Die hohe Jugendarbeitslosigkeit in Südafrika gilt als die entscheidende Ursache für die hohe betriebliche Identität und die sehr hohe Ausbildungsmotivation der Jugendlichen, die über einen Ausbildungsvertrag verfügen. Diese überträgt sich auch auf die Testmotivation.

6. Fazit

In diesem Forschungsfeld der Kompetenzdiagnostik befindet sich die Grundlagenforschung noch ganz am Anfang. Das in diesem Bericht vorgestellte Motivationsmodell hat sich in einem ersten Schritt als eine die Fragestellungen, Methoden und Ergebnisse differenzierende Struktur erwiesen. Es konnten dabei primäre und sekundäre Motivationsaspekte unterschieden und aufeinander bezogen werden und in ihrem Einfluss auf die Testleistung bestimmt werden.

Die vertiefende Analyse der einzelnen Motivationsaspekte macht deutlich, dass die Zusammenhänge zwischen den einzelnen Motivationsitems durch zwei zugrundeliegende Dimensionen erklärt werden können: Sinnhaftigkeit und Investition.

Es hat sich gezeigt, dass die Testleistung der Auszubildenden von drei Faktoren bestimmt wird: Wird die Sinnhaftigkeit (primäre Motivation) der Testaufgaben hoch eingeschätzt, dann führt dies i.d.R. auch zu einer engagierten Bearbeitung der Testaufgaben (sekundäre Motivation/Investition). Beide Motivationsdimensionen haben einen signifikanten Einfluss auf die Testleistung. So erzielen höher motivierte Testteilnehmer bessere Leistungen. Die durch die Motivationsdimensionen aufgeklärte Varianz der Testleistung ist jedoch nicht erschöpfend. Als weiterer signifikanter, starker Einflussfaktor auf die Testleistung hat sich die Dauer der Testbearbeitung erwiesen.

Die diskutierten Ergebnisse und die Beforschung der Testmotivation eröffnen auch die Möglichkeit, die berufsspezifischen Unterschiede in der Testmotivation zu untersuchen.

Ein Vergleich der Motivation in Bezug auf die Bearbeitung von KOMET-Testaufgaben und Prüfungsaufgaben ergab ein konträres Motivationsbild. Während die Einschätzung der primären Motivationsaspekte bezüglich der KOMET-Aufgabe bei den Auszubildenden (signifikant) höher ausfiel, so fiel im Kontrast dazu die Einschätzung der sekundären Motivationsaspekte bezüglich der Prüfungsaufgaben signifikant höher aus. Die Ergebnisse legen die Vermutung nahe, dass die Tatsache, dass Prüfungen einen Einfluss auf die berufliche Karriere haben, während die KOMET-Testung keinen solchen Einfluss hat, als externer Anreiz funktioniert. Auszubildende werden in der Prüfungssituation hierdurch extrinsisch motiviert, während es im Zuge des KOMET-Testverfahrens zu keiner solchen Motivation kommt. Die extrinsische Motivation äußert sich im Ergebnis darin, dass Auszubildende stark in die Bearbeitung der Prüfungsaufgabe investieren, obwohl sie den Berufsbezug und den Nutzen der Prüfungsaufgabe eher indifferent einschätzen und kein übermäßiges Interesse an der Prüfungsaufgabe haben. Entsprechend stellt die nur unzureichende Einbeziehung des KOMET-Tests in die Feedbackstruktur an der Berufsschule ein großes Problem dar. Hier kann auch ein Grund für die teilweise niedrige Testmotivation liegen. Die Auswirkungen eines guten Feedbacks auf den „Lernerfolg“ von Schülern/Auszubildenden/Studierenden durch die Lehrer, Ausbilder oder Dozenten z.B. auf der Grundlage von Tests oder anderer Formen der Evaluation der Kompetenzentwicklung werden in der empirischen Bildungsforschung übereinstimmend als sehr hoch gewertet. Feedback in der Form von Lern- und Ausbildungsberatung gilt zu Recht als Dreh- und Angelpunkt einer guten Lernkultur (WEINERT 1996).

Die KOMET-Testteilnehmer erhalten zwar eine individuelle Rückmeldung zu ihrem Testergebnis, sie sind sich jedoch bewusst, dass die Testergebnisse nicht notenrelevant sind. Immerhin gab es ein durchgängiges Interesse an einer Rückmeldung der Testergebnisse. Darin äußert sich das prinzipielle Interesse am Feedback. Solange jedoch diese Form der Kompetenzdiagnostik in ihrem Potential für die Ausbildungs- und Lernberatung nicht genutzt und nicht systematisch in eine neue Feedbackkultur integriert wird, so lange wird auch die Testmotivation für einen Teil der Testteilnehmer gering bleiben. Verstärkt wird dieser Effekt dadurch, dass die Auszubildenden auf die zwei punktuellen Prüfungszeitpunkte (Teil 1 und Teil 2 der Prüfung) fixiert sind und alle anderen Formen der Überprüfung ihres Ausbildungserfolges deutlich geringer gewichten. Dies gilt durchgängig auch für alle Formen schulischer Leistungsmessung. Teil 1 der Prüfung findet nach 18 Monaten und Teil 2 der Prüfung am Ende der Ausbildung – nach drei bis dreieinhalb Jahren – statt.

Es kann davon ausgegangen werden, dass das Interesse an dieser Form der Kompetenzdiagnostik zunimmt, wenn die Testteilnehmer die Kompetenzdiagnostik als ein Evaluationsinstrumentarium erleben, dem ein hoher diagnostischer Stellenwert zur Bewertung und damit auch zur Verbesserung der Ausbildungsqualität zukommt.

In Einzelfällen wird die KOMET-Kompetenzerhebung jedoch aktuell sogar als Störung der regulären Ausbildung erlebt. Die Bearbeitung der Testaufgaben ist mit einer erheblichen Anstrengung verbunden, wenn sie konzentriert und engagiert wahrgenommen wird. Dies bewerten die Auszubildenden in ihrer überwiegenden Mehrzahl ganz offensichtlich als eine zusätzlich zu erbringende Leistung. Wie stark diese Einstellung sich auf das Testergebnis auswirkt, hängt eng mit dem Zeitpunkt der Testung zusammen. Eine Fallstudie, bei der eine Gruppe von Auszubildenden des vierten Ausbildungsjahres in zeitlicher Nähe zur Abschlussprüfung (Teil 2) getestet wurde, zeigt, dass die Testmotivation zu diesem Zeitpunkt so gering war, dass die Ergebnisse insgesamt deutlich schlechter ausfielen als aufgrund der

bereits fortgeschrittenen Ausbildung zu erwarten war. Die Angaben zur Motivation der Auszubildenden bestätigen dies. Verursacht wird dies durch die Konzentration der Auszubildenden auf ein gutes Bestehen der Prüfung. Ein Test, der nicht eingebunden ist in die Prüfungsvorbereitung oder in die Prüfung selbst, wird daher in Einzelfällen als Störung der Ausbildung empfunden.

Neben diesen Aspekten entstand des Weiteren die Hypothese, dass kontextbezogene Faktoren wie beispielsweise das Schulklima einen Einfluss auf die Testmotivation haben. Dieser Hypothese sollte in Folgeuntersuchungen nachgegangen werden. Eine Untersuchung des Zusammenhanges zwischen der Entwicklung beruflicher und betrieblicher Identität, des arbeitsbezogenen Engagements sowie der Testmotivation steht ebenso noch aus.

Ein neues Forschungsfeld ist in diesem Zusammenhang das Erfassen der Testmotivation sowie des Ausbildungsengagements unter den Bedingungen unterschiedlicher nationaler Ausbildungstraditionen und -kulturen. So konnte zum Beispiel die ungewöhnlich hohe Ausbildungs- und Testmotivation von südafrikanischen Auszubildenden und Studierenden darauf zurückgeführt werden, dass diese zur Minderheit der Jugendlichen gehörten, die mit dem erfolgreichen Abschluss der Ausbildung mit einer Beschäftigung in den ausbildenden Unternehmen rechnen konnten. Eine vergleichsweise hohe Testmotivation wurde auch bei chinesischen Auszubildenden und Studierenden der höheren Fachschulen gemessen. Die Hypothese des chinesischen KOMET-Konsortiums, dass dies (auch) auf eine gesellschaftliche Norm zurückzuführen sei, nach der man sich bei „offiziellen“ Befragungen positiv zu den Verhältnissen äußert, konnte bisher empirisch noch nicht bestätigt werden.

Der hier dargestellte Ansatz zur Erfassung der Testmotivation sowie die Testergebnisse zeigen, dass die Erforschung der Testmotivation für die Durchführung von Projekten der Kompetenzdiagnostik im Bereich der beruflichen Bildung eine notwendige Voraussetzung für die Analyse der Testergebnisse ist.

Literaturangaben

Baumert, J. u.a. (2001): PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske + Budrich.

Brosius, F. (2013): *SPSS 21* (1. Aufl.). Heidelberg: mitp.

Heinemann, L.; Maurer, A.; Rauner, F. (2011): Modellversuchsergebnisse im Überblick. In: Rauner, F.; Heinemann, L.; Maurer, A.; Li Ji; Zhao, Z.: Messen beruflicher Kompetenz. Band III. Drei Jahre KOMET Testerfahrung. Münster: LIT, S. 150-209.

Kuhler, M. u.a. (2002): PISA 2000 – Dokumentation der Erhebungsinstrumente. Berlin: MPI für Bildungsforschung.

Preacher, K. J. & Hayes, A. F. (2004): SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, 36 (4), 717-731.

Rauner, F.; Haasler, B.; Heinemann, L.; Grollmann, P. (2009): Messen beruflicher Kompetenz. Band I. Grundlagen und Konzeption des KOMET-Projektes (2. Auflage). Münster: LIT.

Rauner, F. u.a. (2009): Messen beruflicher Kompetenz. Band II. Ergebnisse KOMET 2008 (2. Auflage). Münster: LIT.

Rauner, F.; Heinemann, L.; Hauschildt, U.; Piening, D. (2012): COMET South Africa: Implementing of learning areas: testing a competence model as basis for test tasks and learning tasks – a pilot project of large-scale competence diagnostics in South African schools and VET colleges. Eds.: merSETA in collaboration with i:BB Universität Bremen.