

Kompetenzzuwachs im Selbstbericht: Adaptation eines Instrumentes zur Evaluation von Führungskräfte trainings

Tobias Ringeisen, Barbara Neubach, Mareike Florax und Martin Kersting

Traditionell wird der erzielte Kompetenzzuwachs durch Führungskräfte trainings (FKTs) anhand von Fremdbereichten evaluiert. Ein geeignetes Instrument zur Selbsteinschätzung fehlt bislang, obwohl aus der Bildungsforschung ermutigende Ergebnisse für einen derartigen Evaluationsansatz vorliegen. Um diese Lücke zu schließen, wurde ein Instrument zur Evaluation der Hochschullehre – das "Berliner Evaluationsinstrument für selbst eingeschätzte studentische Kompetenzen" (BEvaKomp; Braun, Gusy, Leidner & Hannover, 2008) für den Kontext von FKTs adaptiert. Von den sechs Subskalen wurden diejenigen vier ausgewählt, auf die FKTs im Allgemeinen abzielen (Fach-, Methoden-, Kommunikations- und Selbstkompetenz; vgl. Sonntag & Schaper, 2006). Die modifizierten Skalen wurden N = 390 Führungskräften einer Bundesverwaltung nach einem FKT vorgelegt. Konfirmatorische Faktorenanalysen und Skalenanalysen weisen auf eine befriedigende Konstruktvalidität hin. Die Interkorrelationen zwischen den einzelnen Kompetenzfacetten, Zufriedenheit und wahrgenommenem Nutzen sprechen für eine ausreichende diskriminante Validität.

Theoretischer und empirischer Hintergrund

Führungskräfte trainings (FKTs) haben zum Ziel, den teilnehmenden Führungskräften Wissen und Fertigkeiten zum Umgang mit einer Vielzahl von Situationen aus dem Führungsalltag zu vermitteln. Abhängig vom Inhalt des jeweiligen Trainings werden verschiedene Kompetenzfacetten in unterschiedlichem Ausmaß gefördert. Diese umfassen im Allgemeinen Fach-, Methoden-, Kommunikations- und Selbstkompetenz (Sonntag & Schäfer-Rauser, 1993; Sonntag & Schaper, 2006).

In der Weiterbildungsforschung existiert eine Reihe von Ansätzen, um die Effektivität von FKTs anhand des erzielten Kompetenzzuwachses zu evaluieren (vgl. Erpenbeck & von Rosenstiel, 2007; Kaufhold, 2006). Die Mehrheit dieser Ansätze kombiniert Selbstbericht, Fremdbereicht und verhaltensbasierte Verfahren, um die Wirksamkeit eines FKT auf verschiedenen Ebenen durch die teilnehmenden Führungskräfte, ihre Mitarbeiter und/oder Vorgesetzten einschätzen zu lassen. Um Kosten und Aufwand zu reduzieren, konzentriert sich die Evaluation in der Praxis auf Selbsteinschätzungen der Trainingsteilnehmer (Bogner, 2007). Erho-

ben werden im Allgemeinen Daten zur Organisation und der Struktur der Veranstaltung (Strukturevaluation), zum Umsetzungsprozess und zur Didaktik (Prozessevaluation) sowie dem Lernerfolg (Ergebnisevaluation; Braun & Gusy, 2006). Um die Effektivität eines FKT einschätzen zu können, sollte der Schwerpunkt der Evaluation auf die Ergebnisevaluation gelegt werden (vgl. Kaufhold, 2006).

Offensichtlich besteht eine Lücke zwischen Forschungsanspruch und Praxisnutzung, die mit einem kurzen aber reliablen Instrument geschlossen werden könnte, das den simultanen Zuwachs in mehreren Kompetenzbereichen im Selbstbericht abbilden kann. Im Bereich der Persönlichkeits- und der Bildungsforschung haben sich seit längerem Instrumente zur Selbsteinschätzung als geeignete Methode etabliert, um Kompetenzveränderungen valide und reliabel zu messen (Gosling, John, Craik & Robins, 1998; Lucas & Baird, 2006; Klieme, Artelt & Stanat, 2002). Im Bereich der Erwachsenenbildung werden Selbsteinschätzungen zwar häufig zur Evaluation von FKTs genutzt, doch fehlt für viele der eingesetzten Instrumente ein Nachweis ausreichender Validität sowie Reliabilität. Kritisiert werden die mangelnde theoretische Fundierung wie auch ein unzureichender Nachweis psychometrischer Güte (Kauffeld, 2006; Rowold, Hochholdinger & Schaper, 2008; Sonntag & Schäfer-Rauser, 1993).

Es bedarf also eines kurzen aber reliablen Instrumentes, das den simultanen Zuwachs in mehreren Kompetenzbereichen durch Selbsteinschätzung erfassen kann. Um dieses Ziel zu erreichen, wurde ein Instrument zur Evaluation von Hochschullehre – das "Berliner Evaluationsinstrument für selbst eingeschätzte studentische Kompetenzen" (BEvaKomp; Braun et al., 2008) für den Kontext von FKTs adaptiert und seine psychometrischen Eigenschaften analysiert.

Methode

Stichprobe

Die Adaptation des BEvaKomp erfolgte auf Basis der Daten, die an einer Stichprobe von 390 Führungskräften ($n = 120$ Frauen, $n = 270$ Männer) einer Bundesverwaltung erhoben wurden. Aus Gründen der Anonymität wurden die personenbezogenen Daten nur in Form von kategorialen Angaben erfasst. Ein knappes Viertel der Stichprobe war jünger als 40 Jahre, ein weiteres Viertel älter als 54 Jahre. Mehr als die Hälfte der Stichprobe deckte die Spanne von 41 bis 53 Jahren ab. Im Hinblick auf die Berufserfahrung zeigte sich eine ähnlich breite Streuung: Jeweils knapp 30% berichteten zwischen 10 und 20 bzw. 20 und 30 Jahren Berufserfahrung. Knapp ein Fünftel hatte mehr als 30 Jahre Erfahrung. Das restliche Viertel der Stichprobe gab zu gleichen Anteilen 0 bis 5 Jahre bzw. 5 bis 10 Jahre Erfahrung an. Mehr als

die Hälfte der Führungskräfte trugen für weniger als 50 Beschäftigte Verantwortung, knapp 20% für 50 bis 100 Mitarbeiter. Das verbliebene Drittel wies große Führungsspannen auf: Jeweils ca. 10% der Führungskräfte hatten für 101 bis 200, für 201 bis 500 bzw. für mehr als 500 Beschäftigte Verantwortung.

Instrumente

Um den Zuwachs an Kompetenzen im Selbstbericht erfassen zu können, wurde das „Berliner Evaluationsinstrument für selbst eingeschätzte studentische Kompetenzen“ für eine Adaptation ausgewählt (BEvaKomp; Braun et al., 2008). Das BEvaKomp stellt ein standardisiertes und psychometrisch gut abgesichertes Instrument zur Messung von Kompetenzzuwachs durch Selbsteinschätzung dar. Es wurde auf Basis eines mehrfaktoriellen Modells der Handlungskompetenz entwickelt und erlaubt eine Abbildung des Zuwachses für die Facetten Fach-, Methoden-, Präsentations-, Kommunikations-, Kooperations- sowie Personalkompetenz. In der Literatur hat sich für Personalkompetenz mehrheitlich der Begriff Selbstkompetenz durchgesetzt, weshalb wir diese Bezeichnung im Folgenden nutzen (vgl. Sonntag & Schäfer-Rauser, 1993; Sonntag & Schaper, 2006).

Das Original des BEvaKomp umfasst 26 Items, die Subskalen haben 2 bis 6 Items. Die psychometrische Güte des Instruments konnte in mehreren Studien zur Evaluation von universitären Lehrveranstaltungen bestätigt werden (z.B. Braun & Hannover, im Druck; Braun & Leidner, 2009). Das BEvaKomp wurde für den Hochschulbereich entwickelt, doch kann es auch fachübergreifend für die Evaluation von unterschiedlichen Veranstaltungsformen in der Aus- und Weiterbildung eingesetzt werden (vgl. Braun et al., 2008, S. 38/39). Die Items sind so formuliert, dass in der ersten Person Singular („Ich“) explizit ein Zuwachs an Kompetenz durch den Besuch einer Lehrveranstaltung erhoben wird (z.B. für den Bereich Kommunikationskompetenz Item Komm5: „In Folge dieser Lehrveranstaltung kann ich besser Gespräche führen.“). Eine Ausnahme stellt die Subskala Fachkompetenz dar, die die Kompetenzausprägung zum Zeitpunkt der Befragung und nicht den Kompetenzzuwachs erhebt (z.B. Item Fach1: „Ich kann wichtige Begriffe/Sachverhalte aus dieser Lehrveranstaltung wiedergeben.“).

Das BEvaKomp wird den Teilnehmern nach Abschluss einer Veranstaltung vorgelegt. Diese geben durch den Grad Ihrer Zustimmung an, inwieweit der jeweilige Kompetenzbereich ausgeprägt ist (Fachkompetenz) bzw. einen Zuwachs erfahren hat (restliche Subskalen). Die zugehörige 5-stufige Ratingskala reicht von „1 = trifft nicht zu“ bis „5 = trifft völlig zu“. In der vorliegenden Studie füllten die teilnehmenden Führungskräfte das adaptierte

BEvaKomp nach Abschluss eines FKT zum Thema Kommunikation („Methoden und Gesprächsgestaltung“) aus. Um die diskriminante Validität einschätzen zu können, wurden zusätzlich der wahrgenommene Nutzen (5-stufige Ratingskala von „1 = trifft nicht zu“ bis „5 = trifft völlig zu“) und die Zufriedenheit mit dem Seminar (7-stufige Ratingskala von „1 = sehr gering“ bis „7 = sehr hoch“) über je ein Item erhoben. Ferner wurden einige wenige demographische Daten erfragt (siehe oben). Zur Messung von Daten der Prozess- und Strukturevaluation wurden Items mit binärer Ratingskala (ja/nein) verwendet, welche durch ein Antwortfeld für freie Kommentare ergänzt wurden (vgl. Kauffeld, 2006; Rowold et al., 2008). Daten zur Struktur- bzw. Prozessevaluation wurden für die vorliegende Studie nicht ausgewertet und werden daher nicht berichtet.

Vorgehen

Von den sechs Subskalen des BEvaKomp wurden vier für die Adaptation ausgewählt. Die ausgewählten Subskalen repräsentieren diejenigen vier Kompetenzbereiche, auf deren Verbesserung FKTs im Allgemeinen abzielen (Fach-, Methoden-, Kommunikations- und Selbstkompetenz; vgl. Sonntag & Schäfer-Rauser, 1993; Sonntag & Schaper, 2006). Die Items der ausgewählten Subskalen wurden in einem mehrstufigen Prozess modifiziert. Items mit geringer Augenscheinvalidität für den angestrebten Einsatzbereich wurden eliminiert und die Formulierung der beibehaltenen Items durch Experteneinschätzung mit Konsensfindung modifiziert. Parallel wurden neue Items entwickelt. Bei allen Items wurde der Begriff „Lehrveranstaltung“ durch „Seminar“ ersetzt. Es resultierte ein Pool mit 17 Items. Für jede der modifizierten Subskalen ist im Folgenden ein Itembeispiel genannt. Fachkompetenz (besteht aus 4 Items): „Ich kann einen Überblick über die Themen des Seminars geben.“; Methoden- (3 Items) „Ich habe meine Arbeitstechniken durch den Besuch des Seminars verbessert.“; Kommunikations- (5 Items): „In Folge des Seminars kann ich besser Gespräche führen.“; Selbst- (5 Items): „Durch das Seminar ist mein Interesse für das Thema gewachsen“.

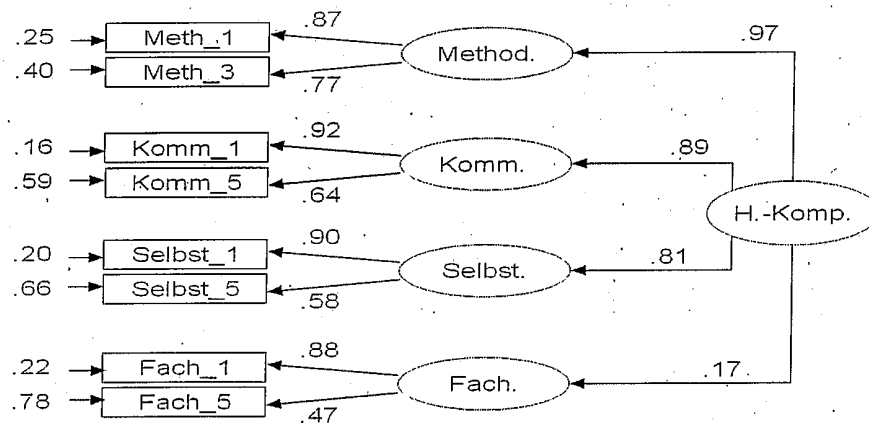
Die Gütekriterien des adaptierten Instrumentes wurden anhand der klassischen Testtheorie überprüft und Item- sowie Skalenparameter berechnet. Die Konstruktvalidität wurde mit einer konfirmatorischen Faktorenanalyse untersucht, die diskriminante Validität wurde über Interkorrelationen zwischen den vier Kompetenzbereichen, Zufriedenheit und Nutzen geschätzt.

Ergebnisse

Faktorielle Struktur

Anhand der Daten erfolgte eine konfirmatorische Überprüfung der postulierten Struktur des adaptierten BEvaKomp. Zunächst wurde auf der Ebene von Primärfaktoren überprüft, ob die Items als Indikatoren von vier latenten Variablen (Fach-, Methoden-, Kommunikations- und Selbstkompetenz) betrachtet werden können. Im Einklang mit bisheriger Forschung (z.B. Braun & Hannover, im Druck; Braun & Leidner, 2009; Kauffeld, 2006) wurde angenommen, dass die vier latenten Variablen miteinander korrelieren. Im Anschluss wurde ein hierarchisches Modell getestet, das neben den vier latenten Variablen auf der Ebene von Primärfaktoren einen Sekundär- bzw. Generalfaktor der Handlungskompetenz postuliert. Letzteres Modell stellt das restriktivere Modell dar und sollte bei ähnlicher Anpassung beider Modelle bevorzugt werden (Parsimony).

Beide Modelle wurden mittels konfirmatorischer Faktorenanalysen mit LISREL (Version 8.8; Jöreskog & Sörbom, 2006) überprüft, wobei Kovarianzen analysiert wurden und die Maximum-Likelihood-Methode zur Schätzung der Parameter zum Einsatz kam. Um die Modellpassung zu überprüfen, wurden sowohl die χ^2 -Prüfgröße als auch verschiedene Fit-Indizes herangezogen, wobei wir uns bei der Interpretation an Schermelleh-Engel, Moosbrugger und Müller (2003) orientierten.



Anmerkungen. Method. = Methodenkompetenz, Komm. = Kommunikationskompetenz, Selbst. = Selbstkompetenz, Fach. = Fachkompetenz, H.-Komp. = übergeordneter Faktor der Handlungskompetenz

Abb. 1: Standardisierte Ladungen für das Sekundärfaktormodell

Die Kennwerte des Primärfaktorenmodells wichen nur geringfügig von denen des Sekundärfaktormodells ab und werden daher aus Platzgründen im Folgenden nicht berichtet. Das Sekundärfaktormodell mit standardisierten Lösungen ist in Abbildung 1 dargestellt, wobei aus Gründen der Übersichtlichkeit pro Primärfaktor nur die Items mit den höchsten und den niedrigsten Ladungen abgebildet sind.

Die Items der drei Kompetenzbereiche Methoden-, Kommunikations- und Selbstkompetenz zeigten durchweg hohe Ladungen auf den Primärfaktoren. Letztere wiederum wiesen hohe bis sehr hohe Ladungen auf dem Sekundärfaktor auf. Für den Faktor Fachkompetenz lagen die Ladungswerte der Items auf dem Primärfaktor in einem guten bis befriedigenden Spektrum. Die zugehörige Ladung des Primär- auf dem Sekundärfaktor hingegen war zwar signifikant, der Koeffizient erreichte aber nur eine mäßige Höhe (vgl. Schermelleh-Engel et al., 2003). Während die χ^2 -Statistik auf eine befriedigende Passung zwischen Modell und Datenmatrix hinwies ($\chi^2 = 280$, $df = 115$, $\chi^2/df = 2,43$, $p = .00$), indizierten die übrigen Kriterien eine akzeptable ($RMSEA = .060$, $SRMR = .080$) bis gute Anpassung des Modells ($CFI = .98$, $NFI = .97$; vgl. Schermelleh-Engel et al., 2003). Insgesamt weist das hierarchische Modell Methoden-, Kommunikations- und Selbstkompetenz als gut differenzierbare und gleichwertig wichtige Komponenten einer übergeordneten Handlungskompetenz aus. Fachkompetenz bildet nach dem Modell ebenfalls einen gut definierbaren Kompetenzbereich, der allerdings eine schwächere Facette von Handlungskompetenz als die drei o. g. Bereiche darstellt.

Nach der konfirmatorischen Überprüfung der Skalenstruktur wurden die Mittelwerte, Standardabweichungen, Trennschärfen und internen Konsistenzen für die Gesamtstichprobe berechnet. Die entsprechenden Werte für die Subskalen und die Gesamtskala sind in Tabelle 1 aufgeführt.

Tab. 1: Mittelwerte, Standardabweichungen und interne Konsistenzen der Subskalen und der Gesamtskala

Subskalen	<i>M</i>	<i>SD</i>	α	$r_{i(i-1)}^a$
Methoden. (3 Items)	9.68	2.40	.84	.66 - .72
Komm. (5 Items)	17.90	3.52	.87	.56 - .79
Selbst. (5 Items)	19.41	3.82	.84	.49 - .77
Fach. (4 Items)	16.50	1.95	.74	.38 - .61
Gesamtskala (17 Items)	63.50	9.21	.89	.20 - .72

Anmerkungen. Methoden = Methodenkompetenz, Komm = Kommunikationskompetenz, Selbst = Selbstkompetenz, Fach = Fachkompetenz.

^a Korrigierte Trennschärfe (part-whole-Korrektur), Werte bezogen auf die jeweilige Subskala, angegeben sind der höchste und der geringste Wert pro Subskala.

Wie die Ergebnisse zeigen, lagen die Mittelwerte der Subskalen im mittleren bis hohen Bereich. Den niedrigsten Mittelwert wies die Subskala Methodenkompetenz auf, den höchsten die Subskala Fachkompetenz. Auf Subskalenniveau waren die Trennschärfekoeffizienten relativ hoch und erreichten Werte zwischen $r_{i(t-i)} = .38$ und $r_{i(t-i)} = .79$. Für die Gesamtskala lagen die Trennschärfen erwartungsgemäß niedriger, hier verteilten sich die Koeffizienten in einem Bereich zwischen $r_{i(t-i)} = .20$ und $r_{i(t-i)} = .72$. Die Reliabilitäten der Subskalen (nach Cronbach's Alpha) lagen in einem guten bis sehr guten Bereich. Eine durchweg hohe Homogenität erreichten die Subskalen Methoden-, Kommunikations- und Selbstkompetenz. Etwas geringer fiel der Wert für Fachkompetenz aus. Für die Gesamtskala lag der Koeffizient nur knapp unter einem sehr guten Wert von .90.

Interkorrelationen der Skalen

Um die Zusammenhänge der Subskalen zu untersuchen, wurden Interkorrelationen berechnet (siehe Tabelle 2). Für Methoden-, Kommunikations- und Selbstkompetenz lagen die Koeffizienten in einem hohen Bereich und erreichten durchweg Werte $> .60$. Alle drei Facetten können demnach klar als Komponenten einer generellen Handlungskompetenz interpretiert werden, die zwar deutlich wechselseitig verbunden sind, aber inhaltlich separierbare Facetten bilden. Alle drei Kompetenzfacetten waren zudem positiv mit Fachkompetenz korreliert, die Höhe der Koeffizienten lag in einem niedrigen bis mittleren Bereich.

Tab. 2: Interkorrelationen der Subskalen

	Methoden.	Komm.	Selbst.	Fach.
Methoden.		.73	.64	.18
Komm.			.62	.19
Selbst.				.27
Fach.				

Anmerkungen. Methoden = Methodenkompetenz, Komm = Kommunikationskompetenz, Selbst = Selbstkompetenz, Fach = Fachkompetenz. Bei allen Korrelationen $p < .01$

Kriteriumsvalidität

Um Hinweise auf die diskriminante Validität des adaptierten BEvaKomp zu erhalten, wurden Interkorrelationen zwischen den vier Kompetenzbereichen (Fach-, Methoden-, Kommunikations- und Selbstkompetenz), dem wahrgenommenen Nutzen und der Zufriedenheit

mit dem Seminar berechnet (siehe Tabelle 3). Im Einklang mit bisherigen Studien zu diesem Variablenbereich (z.B. Oeffner & Preiser, 2007; Braun et al., 2008; Kauffeld, 2006) zeigten sich mittlere bis hohe Interkorrelationen. Die geringsten Koeffizienten wies der wahrgenommene Nutzen bzw. die Zufriedenheit mit Fachkompetenz auf, die höchsten Koeffizienten hingegen mit Selbstkompetenz. Offensichtlich nehmen Führungskräfte Kompetenzzuwachs in verschiedenen Bereichen, Nutzen und Zufriedenheit als unterschiedlich wahr und schätzen die Konstrukte somit als zusammenhängend aber gut differenzierbar ein.

Tab. 3: Korrelationen der vier Subskalen mit Nutzen und Zufriedenheit

	Methoden.	Komm.	Selbst.	Fach.
Wahrgenommener Nutzen	.62	.60	.68	.22
Zufriedenheit mit dem FKT	.49	.51	.62	.28

Anmerkungen. Methoden = Methodenkompetenz, Komm = Kommunikationskompetenz, Selbst = Selbstkompetenz, Fach = Fachkompetenz. Bei allen Korrelationen $p < .01$

Fazit und Ausblick

Die Ergebnisse der aktuellen Studie legen nahe, dass das adaptierte BEvaKomp über ausreichende Validität und Reliabilität verfügt, um als Selbstberichtsinstrument zur Evaluation von FKT eingesetzt zu werden. Item- und Skalenparameter weisen gute Werte auf und die Ergebnisse der konfirmatorischen Modellprüfung bestätigen im Hinblick auf Faktorkennwerte, Ladungsmuster und Interkorrelationen die angenommene Struktur der vier Subskalen. Die Zusammenhänge der vier Subskalen mit Zufriedenheit und dem wahrgenommenen Nutzen indizieren eine befriedigende diskriminante Validität. Insgesamt decken sich die Ergebnisse mit den bisherigen Befunden zum BEvaKomp (Braun et al., 2008, Braun & Hannover, im Druck; Braun & Leidner, 2009), was als zusätzlicher Hinweis auf die psychometrische Güte des adaptierten Instruments gewertet werden kann.

Die einzige Ausnahme stellt die geringe Interkorrelation zwischen Fachkompetenz und den drei anderen Kompetenzfacetten dar. Der verringerte Zusammenhang könnte durch die unterschiedliche Formulierung der Items in den vier Subskalen bedingt sein. Während die Items der Subskala Fachkompetenz die aktuelle Kompetenzausprägung erheben, erfassen die Items der drei restlichen Skalen den Kompetenzzuwachs. Beispielsweise konnten Kauffeld und Kollegen (z.B. Kauffeld, 2006; Kauffeld & Grote, 2000) zeigen, dass variiierende Itemformulierungen den Zusammenhang der interessierenden Kompetenzfacetten beeinflussen. Sogar Braun et al. (2008, S. 38) räumen in der Diskussion der Originalstudie ein, dass „zwei verschiedene Arten von Itemformulierungen [...] zu einer Unterschätzung der empiri-

schen Zusammenhänge zwischen den auf unterschiedliche Weise gemessenen Kompetenzbereichen führen [können].

Alternative Gründe könnten ebenfalls für den geringen Zusammenhang verantwortlich sein. So erscheint es aus methodischer Sicht plausibel, dass die Ergebnisse der Konstrukt- und Kriteriumsvalidierung je nach Inhalt eines FKT variieren (vgl. Kaufhold, 2006; Sonntag & Schäfer-Rauser, 1993). Um einen solchen Effekt zu kontrollieren, haben Braun et al. (2008) das BEvaKomp zur Evaluation in mehr als 10 Studiengängen an 9 Universitäten eingesetzt und die Daten auf Gruppen- und Studierendenebene ausgewertet. Zukünftige Studien sollten das adaptierte BEvaKomp daher in FKTs mit unterschiedlichen Inhalten einsetzen und die Ergebnisse auf verschiedenen Aggregationsebenen kontrastieren.

Weiterhin ist denkbar, dass sich die vier Kompetenzfacetten unterschiedlich schnell aufbauen und über die Zeit wechselseitig beeinflussen. So hat sich im beispielsweise gezeigt, dass Methoden-, Kommunikations- und Selbstkompetenz eine Voraussetzung bilden, um Fachkompetenz aufzubauen (Staudt, Kottmann & Merker, 1997). Die Zusammenhänge der vier Facetten verstärken sich also, wenn die dahinter liegenden Verhaltensweisen im Führungsalltag angewendet werden. In der aktuellen Studie wurde das BEvaKomp direkt nach Ende eines FKT ausgefüllt. Die Teilnehmer des FKT konnten somit noch keine neue Verhaltensweisen in der Praxis testen, weshalb eine tatsächliche Verbesserung der einzelnen Kompetenzfacetten – vor allem der Fachkompetenz – nur eingeschränkt eingeschätzt werden konnte. In Zukunft sollte das BEvaKomp somit erst mehrere Wochen bis Monate nach Abschluss eines FKT vorgelegt werden (vgl. Kauffeld, 2006; Rowold et al., 2008).

Schließlich fehlt trotz der ermutigenden Ergebnisse noch ein umfassender Nachweis der Kriteriumsvalidität im Sinne einer Multitrait-Multimethod-Analyse (Campbell & Fiske, 1959). Anstelle von Einzelitems sollten in zukünftigen Studien bestehende Skalen oder Checklisten als Selbstbericht, Fremdrating und/oder für Verhaltensbeobachtungen eingesetzt werden, um die diskriminante und die konvergente Validität differenzierter einschätzen zu können. Parallel bietet sich mittel- bis langfristige an, „harte“ Kriterien wie Krankenstand oder Fluktuationsrate hinzuzunehmen.

Trotz der genannten Einschränkungen liegt mit dem adaptierten BEvaKomp ein kurzes und einfach einsetzbares Instrumentes vor, das sich als Selbstbericht zu eignen scheint, um den Zuwachs in verschiedenen Kompetenzfacetten durch ein FKT für Führungskräfte abzubilden.

Literaturverzeichnis

- Bogner, J. (2007). *Vorhersage und Förderung einer erfolgreichen Führungskarriere: Längsschnittliche Befunde zur Vorhersage des Karriereerfolgs und Meta-Analyse zur Wirksamkeit von Führungskräfte-Training*. Unveröffentlichte Dissertation: Universität Nürnberg-Erlangen.
- Braun, E. & Gussy, B. (2006). Perspektiven der Lehrevaluation. In G. Krampen & H. Zayer (Hrsg.), *Didaktik und Evaluation in der Psychologie* (S. 152-166). Göttingen: Hogrefe.
- Braun, E., Gussy, B., Leidner, B. & Hannover, B. (2008). Kompetenzorientierte Lehrevaluation - Das Berliner Evaluationsinstrument für selbsteingeschätzte, studentische Kompetenzen (BEvaKomp). *Diagnostica*, 54, 30-42.
- Braun, E. & Hannover, B. (im Druck). Gelegenheiten zum Kompetenzerwerb in der universitären Lehre. Zusammenhänge zwischen den Einschätzungen Studierender und unabhängigen Beobachtungen relevanter Merkmale universitärer Lehrveranstaltungen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*.
- Braun, E. & Leidner, B. (2009). Academic course evaluation. Theoretical and empirical distinctions between self-rated gain in competences and satisfaction with teaching behavior. *European Psychologist*, 14, 297-306.
- Campbell, D. T. & Fiske, D. W. (1959). Convergent and discriminant validation by the multi-trait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
- Erpenbeck, J. & von Rosenstiel, L. (2003) (Hrsg.). *Handbuch Kompetenzmessung. Erkennen, Verstehen und Bewerten von Kompetenzen in der betrieblichen, pädagogischen und psychologischen Praxis*. Stuttgart: Schaeffer-Poeschel.
- Gosling, S. D., John, O. P., Craik, K. H., & Robins, R. W. (1998). Do People know how they behave? Self-reported act frequencies compared with on-line codings by observers. *Journal of Personality and Social Psychology*, 74, 1337-1349.
- Jöreskog, K. G. & Sörbom, D. (2006). *LISREL 8.8*. Chicago: SSI Scientific Software International.
- Kauffeld, S. (2006). *Kompetenzen messen, bewerten, entwickeln*. Stuttgart: Schäffer-Poeschel.
- Kauffeld, S. & Grote, S. (2000). Sozialkompetenz als der Schlüssel zur erfolgreichen Bewältigung von Optimierungsaufgaben? Zur Bedeutung der Kompetenzfacetten. In E. Frieling, S. Kauffeld, S. Grote & H. Bernard (Hrsg.), *Flexibilität und Kompetenz: Schaffen flexible Unternehmen kompetente und flexible Mitarbeiter?* (S. 49-74). Münster: Waxmann.

- Kaufhold, M. (2006). *Kompetenz und Kompetenzerfassung: Analyse und Beurteilung von Verfahren der Kompetenzerfassung*. Wiesbaden: VS-Verlag.
- Klieme, E., Artelt, C. & Stanat, P. (2002). Fächerübergreifende Kompetenzen. In F. E. Weinert (Hrsg.), *Leistungsmessungen in der Schule* (S. 203-218). Weinheim: Beltz.
- Lucas, R. E. & Baird, B. M. (2006). Global self-assessment. In M. Eid & E. Diener (Hrsg.), *Handbook of multimethod measurement in psychology* (S. 29-42). Washington, DC: American Psychological Association.
- Oeffner, S. & Preiser, S. (2007). Supervision - ein Instrument zur berufsbezogenen Kompetenzerweiterung: Eine Evaluationsstudie in einer Non-Profit-Organisation. In M. Krämer, S. Preiser & K. Brusdeylins (Hrsg.), *Psychologiedidaktik und Evaluation in Aus-, Fort- und Weiterbildung* (S. 125-133). Göttingen: Vandenhoeck & Ruprecht.
- Rowold, J., Hochholding, S. & Schaper, N. (Hrsg.) (2008). *Evaluation und Transfersicherung betrieblicher Trainings - Modelle, Methoden, Befunde*. Göttingen: Hogrefe.
- Schermelleh-Engel, K., Moosbrugger, H. & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8, 23-74.
- Sonntag, K.-H. & Schäfer-Rauser, U. (1993). Selbsteinschätzung beruflicher Kompetenzen bei der Evaluation von Bildungsmaßnahmen. *Zeitschrift für Arbeits- und Organisationspsychologie*, 37, 163-171.
- Sonntag, K.-H. & Schaper, N. (2006). Förderung beruflicher Handlungskompetenz. In K. H. Sonntag (Hrsg.), *Personalentwicklung in Organisationen* (S. 230-311). Göttingen: Hogrefe.
- Staudt, E., Kottmann, M. & Merker, R. (1997). *Chemiker: Hochqualifiziert aber inkompetent?* Institut für angewandte Innovationsforschung: Bochum.