



Erfolgsrelevantes Verhalten in Coaching und
Führung -
Ein wissenschaftlich fundiertes
Coachingprogramm für Führungskräfte

Peter Behrendt, Freiburg Institut

In Kooperation mit:

Uwe Zimmermann, Stryker

Amel Karboul, Change Leadership & Partners

Prof.Dr. Anja Göritz, Albert-Ludwigs-Universität

Erfolgsrelevantes Verhalten in Coaching und Führung - Ein wissenschaftlich fundiertes Coachingprogramm für Führungskräfte

Coaching erfreut sich steigender Beliebtheit, insbesondere bei Fach- und Führungskräften (Stahl & Marlinghaus, 2000). Diese subjektive Beliebtheit wird durch wissenschaftliche Erkenntnisse untermauert: Coaching wirkt! Nach aktuellen Studien sind die tatsächlichen Effekte allerdings geringer als in anderen etablierten psychologischen Interventionen, wie beispielsweise der Psychotherapie. Die durchschnittlichen Verbesserungen im Coaching liegen lediglich bei Effektstärken von $g = .39$ im Vergleich zu Kontrollgruppen (Theeboom, Beersma, van Vianen, 2013). Diese mittleren Effektstärken von Coaching nehmen nicht wunder, wenn man sich vor Augen hält, dass Psychotherapeuten und deren Ausbilder auf Erkenntnissen jahrzehntelanger Forschung aufbauen können. Im Coaching hinkt die wissenschaftliche Erforschung der Wirkprozesse dem schnell wachsenden Markt noch deutlich hinterher (Künzli, 2005, 2009, Feldman und Lankau 2005). Die ersten Ansätze zur objektiven Analyse von erfolgsrelevantem Coachingverhalten stammen von Behrendt (2006) und Greif (2010). Diese Erkenntnisse der letzten Jahre haben aber noch kaum Eingang in die Praxis gefunden: Nach Greif sind in Deutschland nur 4,3% der Coachingtrainings voll wissenschaftlich fundiert, in den USA sogar nur 2,2% (2013).

Darüber hinaus gibt es in der Coachingforschung aktuell keine eigenständigen wissenschaftlichen Modelle, die die durch Coaching induzierten Verhaltensveränderungen beim Klienten und deren Zusammenhang mit objektiven Erfolgsmaßen untersuchen. In der Psychotherapie bestehen sehr differenzierte Erklärungsmodelle von Krankheit und Gesundheit, sowie den dazugehörigen sukzessiven Verhaltensveränderungen, die auf dem Weg von Krankheit zu Gesundheit beim Patienten induziert werden müssen (Grawe, 1998). Diese Erklärungsmodelle ermöglichen Therapeuten wissenschaftlich fundierte Verhaltensanalysen und Tests (1) zur Diagnose der Eingangsvoraussetzungen, (2) zur Verfolgung des Verbesserungsprozesses im Laufe der Intervention und (3) abschließend zur Feststellung eines erfolgreichen Endzustandes. Im Coaching liegt der Fokus nicht auf Krankheit und Gesundheit von Patienten. Über die Hälfte der Klienten sind Führungskräfte und eine der beiden Hauptgründe für Coaching ist der Wunsch nach Verhaltensänderungen (Jansen, Mäthner & Bachmann, 2003). Daher benötigt eine stärkere wissenschaftliche Fundierung von Coaching nicht nur ein wissenschaftliches Modell der Wirkprozesse im Coaching selbst, sondern auch ein wissenschaftliches Verhaltensmodell von erfolgreicher und weniger erfolgreicher Führung.

In dem vorgelegten Projektstudie haben sich dementsprechend die vier Hauptkooperationspartner die Aufgabe gestellt, ein Coachingprogramm für Führungskräfte umzusetzen, das (1) auf den aktuellen Erkenntnissen der Wirkungsforschung im Coaching beruht, (2) wissenschaftliche Modelle des Führungsverhaltens durch objektive Verhaltensanalysen für der Praxis nutzbar macht, (3) reale Verhaltensänderungen bei den Klienten im Laufe des Programms objektiv misst und (4) durch ein experimentelles Design den Effekt des Coachingprogramms auf objektive Zielkennzahlen prüft.

1 Theoretischer Hintergrund

1.1 Erfolgsrelevantes Verhalten des Coachs

Die Arbeiten von Greif (2010) und Behrendt (2006) sind die einzigen aktuell verfügbaren Studien, die anhand von objektiven Verhaltensanalysen des Coachs Coachingerfolg vorhersagen. Es existieren darüber hinaus zwar viele Fragebogenuntersuchungen (u.a. Bachmann, Jansen, & Mäthner, 2004, Jansen, Mäthner und Bachmann, 2003, Runde, 2003). Allerdings konnte Behrendt (2004) zeigen, dass das Verhalten des Coachs durch nachträgliche Befragungen von Coachs und Klienten nur sehr unreliabel erfasst wird. Die Übereinstimmung der Einschätzungen von Coach und Klient, mit dem objektiv gerateten Verhalten war sehr gering: Die Korrelationen lagen um $r = .2$ und $r = .3$ (Behrendt, 2004).

In der Studie von Behrendt (2004, 2006) wurde das therapeutische Wirkfaktorenmodell von Grawe (1998, s.a. Grawe, Donati, & Bernauer, 1994) auf den Coachingkontext übertragen (Behrendt, 2012). Das von Grawe entwickelte Videoanalyse-System (Grawe, Regli & Schmalbach, 1994) wurde durch Skalen- und Itemanalysen im Coaching weiter differenziert (Behrendt, 2004). Daraus ergaben sich sieben Verhaltensskalen, die über Videoratings sehr reliabel gemessen werden konnten (s. Abbildung 1).

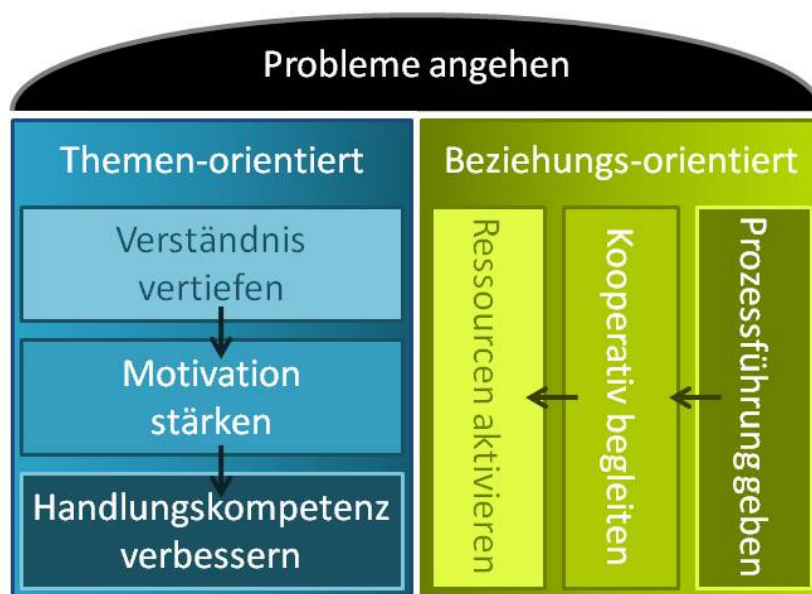


Abbildung 1: Sieben Verhaltensskalen für Coachs nach Behrendt – Die Benennung wurde grundsätzlich verhaltensnäher gewählt, um deutlich zu machen, dass es sich um Verhaltensskalen handelt. Die drei Beziehungs-orientierten Skalen resultieren aus Grawes wichtigstem Wirkfaktor, der Ressourcenaktivierung. Die drei Themen-orientierten Skalen basieren auf den beiden Wirkfaktoren, Bewältigung und motivationale Klärung. Da die beiden Aspekte der Motivation und Klärung nicht korrelierten, wurden diese in ‚Motivation stärken‘ und ‚Verständnis vertiefen‘ getrennt. Die Skala ‚Probleme angehen‘ korrespondiert mit dem Wirkfaktor Problemaktualisierung.

Der Coachingerfolg wurde in der Studie durch die direkte Bewertung des Klienten im Anschluss an das Coaching sowie die spätere Zielerreichung in der Praxis des Klienten gemessen (Behrendt, 2004, 2006). Dieser Coachingerfolg konnte durch drei der sieben Verhaltensskalen signifikant vorhergesagt werden¹:

¹ Die drei Skalen, die diese Verhaltensweisen messen, wurden in der Originalarbeit, (1) Ressourcenaktivierung, (2) (nicht-)distanzierendes Verhalten und (3) strukturierendes Verhalten genannt. Nachträgliche

- (1) **Prozessführung geben:** z.B. durch Strukturierung des Prozesses, kompetentes Auftreten und Vermittlung des Verfahrens
- (2) **Kooperativ begleiten:** z.B. durch Aufnehmen der Anliegen des Klienten und kein Hinterfragen und Belehren, sowie Fokus auf die eigenen Beiträge des Klienten
- (3) **Ressourcen aktivieren:** z.B. durch Wertschätzung des Klienten, Fokus auf seine Stärken, Mitgefühl und Zuversicht vermitteln

Die drei Beziehungs-orientierten Verhaltensweisen stellen für Coachs damit die entscheidenden Hebel für Erfolg dar (Regli, Bieber, Mathier, & Grawe, 2000). Die vier anderen Kategorien stellten im Sinne der Wirkung zwar wichtige Outputfaktoren von erfolgreichem Coaching dar, waren aber im Sinne des Coachverhaltens keine entscheidenden Inputfaktoren (Behrendt 2006). In anderen Worten war es für einen hohen Umsetzungserfolg wichtig, dass Klienten berichteten, dass Sie ihre Handlungskompetenz verbessern konnten. Dafür wiederum war es aber nicht bedeutsam, in wie weit der Coach selbst dafür konkrete Anregungen gab (Verhaltensskala Coach ‚Handlungskompetenz verbessern‘). Dies entspricht der Konzeption von Coaching als Hilfe zur Selbsthilfe (Rauen, 2008). Die Klienten konnten demnach insbesondere dann ihre Handlungskompetenz verbessern, wenn sie durch eine gute Beziehungsgestaltung des Coachs dafür die richtigen Rahmenbedingungen hatten (Behrendt 2006).

Die Effektstärken waren mit Korrelationen bis zu $r > .6$ sehr groß. Die Validität dieser Ergebnisse wird etwas eingeschränkt durch die relativ kleine Stichprobe ($n = 35$ Coachingsitzungen), sowie das korrelative Design der Verhaltensanalysen. Allerdings replizieren die Ergebnisse von Behrendt (2006) die früheren Ergebnisse im therapeutischen Kontext aus der Wirkfaktorenforschung um Grawe (Flückiger, Frischknecht, Wüsten, & Lutz, 2008, Flückiger & Grosse Holforth, 2008, Regli, Flückiger, & Grawe, 2004, Smith & Grawe, 2003, 2005, Smith, Grawe, & Regli, 1999), was für eine hohe externe Validität spricht.

Die Erkenntnisse von Grawe (1998) und Behrendt (2006) wurden 2010 von Greif weiterentwickelt. Basierend auf der Videoanalyse von 44 Coachingsitzungen zwischen Studierenden analysierten Greif, Schmidt und Thamm (2010) insgesamt acht Verhaltensweisen und deren Zusammenhang zu verschiedenen Erfolgsmaßen. Nicht alle Verhaltensweisen konnten mit zufriedenstellender Reliabilität gemessen werden. Dennoch fanden auch Greif et al. starke positive Effekte auf Coachingerfolgsmaße (2010). Zusätzlich zu den von Behrendt (2006) etablierten Verhaltensweisen wurden zwei weitere als reliabel und erfolgsrelevant etabliert:

- (4) **Praxistransfer unterstützen:** Z.B. durch Fokus auf Planung, Umsetzung und das Ausräumen von Umsetzungshindernissen
- (5) **Feedback einholen:** z.B. durch verbales Feedback des Klienten oder quantitative Evaluationen des Coachingprozesses²

Faktorenanalysen haben eine leichte Umgruppierung der Verhaltensweisen in die drei hier berichteten Verhaltensskalen ergeben.

² Diese Ratingskala war jedoch keine reine Verhaltensskala bezogen auf den Coach. Da ein hohes Rating nur vergeben wurde, wenn das Feedback des Klienten positiv war, beinhaltete das Rating zudem eine mit dem Verhalten konfundierte Erfolgsmessung,

Zusammenfassend sollte ein wissenschaftlich fundiert arbeitender Coach insbesondere fünf Verhaltensweisen einsetzen: Er sollte (1) Prozessführung geben, (2) kooperativ begleiten, (3) Ressourcen aktivieren, (4) Praxistransfer unterstützen und (5) Feedback einholen.

1.2 Erfolgsrelevantes Verhalten der Führungskraft

Neben den erfolgsrelevanten Verhaltensweisen für Coaches benötigt wissenschaftlich fundiertes Coaching auch ein Erklärungsmodell erfolgreicher Führung. Wenn ein Coach beispielsweise Ressourcen aktivieren möchte, braucht er eine klare Indikation dafür, was eine relevante Verhaltensweise des Klienten ist, die er als Stärke fokussieren kann. Wenn er den Praxistransfer unterstützen möchte, dann braucht er ein klares Bild, welche Verhaltensweisen die Führungskraft stärker in der Praxis umsetzen sollte, um erfolgreicher zu sein, und welche nicht.

In der Führungsforschung der vergangenen Jahre dominiert das Führungsmodell der Transformationalen Führung (Bass & Riggio, 2006). Dieses wurde inzwischen jedoch stark kritisiert, da es fast ausschließlich auf methodisch relativ schwachen Fragebogenstudien beruht, die trennscharfe theoretische Fundierung in Frage gestellt ist und einige zentrale Konstrukte sehr verhaltensfern sind – insbesondere Charisma (van Knippenberg & Sitkin, 2013). Yukl hat 2012 auf Basis einer umfassenden Literaturanalyse eine neue Taxonomie von erfolgsrelevantem Führungsverhalten entwickelt (s.a. Yukl, Gordon, & Taber, 2002). Diese Taxonomie ist verhaltensnah und berücksichtigt im Gegensatz zu transformationaler Führung alle Verhaltensweisen, die Erfolg von Führung in wissenschaftlichen Studien vorhersagen konnten.

Yukl's hierarchische Taxonomie von erfolgsrelevantem Führungsverhalten (2012) unterscheidet vier Metakategorien: Themen-orientierte, Beziehungs-orientierte, Veränderungs-orientierte und Extern-orientierte (Burke et al., 2006) Verhaltensweisen. Diesen vier Metakategorien ordnet Yukl 15 spezifische Verhaltenscluster und diesen wiederum jeweils mehrere ganz konkrete Verhaltensweisen zu. Durch diese Systematik ist die Taxonomie sehr umfassend und konkret. Allerdings steht dem ein Nachteil gegenüber: Die Taxonomie ist eine Zusammenstellung von erfolgsrelevanten Verhaltensweisen, aber kein kohärentes, theoretisch fundiertes Erklärungsmodell.

Dementsprechend kann die Taxonomie von Yukl die wesentlichen Inhalte eines Modells liefern, nicht aber dessen ordnenden Erklärungsrahmen. In der vorliegenden Studie wurde daher auf das Modell von Grawe und dessen Spezifizierung durch Behrendt zurückgegriffen, um die verschiedenen Verhaltensweisen aus Yukl's Taxonomie in ein kohärentes Erklärungsmodell von Erfolg einzuordnen (s. Abbildung 2):

Die drei Beziehungs-orientierten Verhaltensweisen beruhen auf der neuropsychologischen Erkenntnis, dass es wesentlich einfacher ist, funktionale Muster zu reaktivieren oder neu zu verknüpfen, als etablierte dysfunktionale Muster vollständig zu inhibieren oder ganz neue Muster aufzubauen (Grawe, 1998, Behrendt, 2012). Diese neuropsychologische Erkenntnis gilt auch im Kontext von Führung. In einer erfolgreichen Führungsbeziehung gelingt es demnach, die vorhandenen Ressourcen (bzw. funktionalen neurologischen Muster) effektiv zu nutzen. Die Beziehungs-orientierten und extern-orientierten Verhaltensweisen von Yukl (2012) lassen sich hier gut einordnen. Beide beziehen sich auf eine effektive Beziehungsgestaltung, in einem Fall auf Beziehungen im Team (Beziehungs-orientiert), im anderen Fall auf Beziehungen außerhalb des Teams (extern-orientiert). In beiden Fällen müssen Führungskräfte die vorhandenen Ressourcen (1) durch eine klare Prozessführung auf ein gemeinsames Ziel ausrichten (Drath et al., 2008), (2) durch

Kooperation sicherstellen, dass alle Ressourcen effektiv eingebracht werden können (Spreitzer, 1995) und (3) die hilfreichsten Ressourcen weiter selektiv verstärkt. Die Meta-Analyse von Marrone zu extern-orientierten Verhaltensweisen hat eine vergleichbare drei-stufige Systematik entwickelt (2010).

Die drei Themen-orientierten Verhaltensweisen sind aus dem allgemeinspsychologischen Rubikonmodell von Heckhausen abgeleitet (Heckhausen & Gollwitzer, 1987). Es beschreibt die menschlichen Phänomene der Motivation und Volition (Achtziger & Gollwitzer, 2008), die auch im Kontext von Führung relevant sind. Yukl's Verhaltensweisen aus den beiden Metakategorien der Themen- und Veränderungs-Orientierung (2012) lassen sich hier einordnen. Beide beziehen sich auf fachliche Prozesse, im einen Fall in Routine (Themen-orientiert) im anderen Fall in Veränderungsprozessen (Veränderung-orientiert). In Routine- und Veränderungsprozessen müssen Individuen (4) ihr Verständnis der Situation vertiefen, für daraus abgeleitete Ziele (5) ihre Motivation stärken und dann zur Umsetzung (6) ihre Handlungskompetenz verbessern. Dabei ist die Fokussierung in Routineprozessen mehr auf den Ist-Zustand, in Veränderungsprozessen auf einen neuen Soll-Zustand. In den Routineprozessen ist das vertiefte Verständnis z.B. auf Daten und Fakten konzentriert, bei Veränderungsprozessen dagegen auf die Entwicklung ganz neuer Perspektiven und die Überwindung überholter Annahmen.

Die siebte Verhaltensweise ‚Probleme angehen‘ ist ebenfalls neuropsychologisch fundiert (Behrendt, 2012; Caspar, Rothenfluh, & Segal, 1992). Vorhandene dysfunktionale neuronale Muster können nur verändert werden, wenn sie auch aktiviert sind. Neuronale Netzwerke werden generell nur neu verknüpft oder auch inhibiert, wenn sie vorher aktiviert waren (Grawe, 1998). In anderen Worten, Probleme können nicht gelöst werden, solange sie von den beteiligten Individuen vermieden und nicht bewusst angegangen werden. Dementsprechend identifiziert auch Yukl im Führungskontext einige problem-orientierte Verhaltensweisen als erfolgsrelevant (2012).

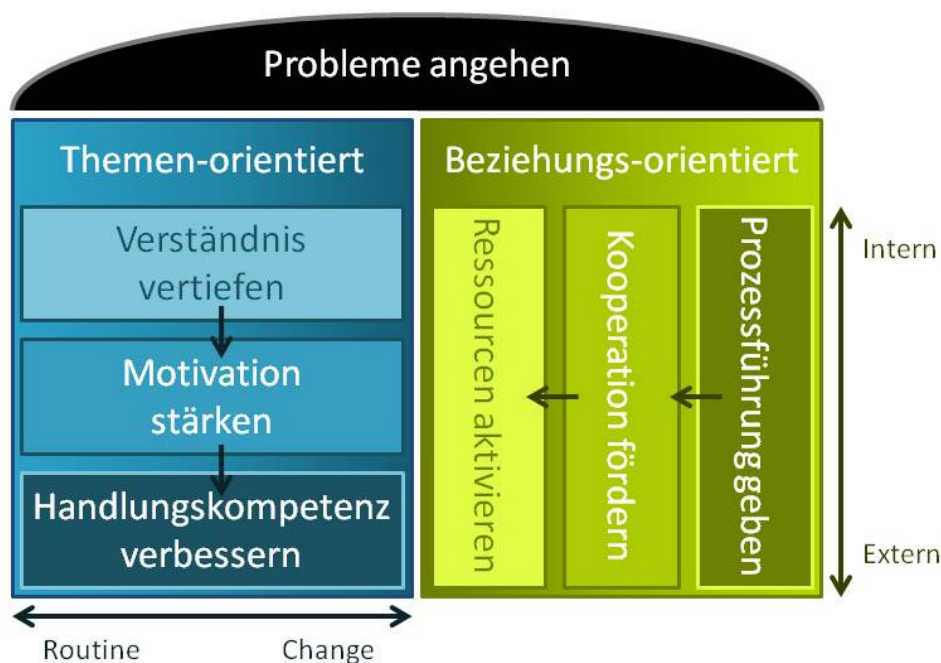


Abbildung 2: Sieben Verhaltensskalen für Führungskräfte nach Behrendt

Zusammenfassend sollten Führungskräfte also sieben Verhaltensweisen nutzen, um ihren Führungserfolg zu erhöhen. Auf der Beziehungsebene sollten sie gegenüber Ihren Mitarbeitern (1) eine klare Prozessführung geben, (2) die Kooperation fördern und (3) deren Ressourcen aktivieren. Auf der Themenebene sollten sie in ihren Teams (4) das Verständnis vertiefen, (5) die Motivation stärken und (6) die Handlungskompetenz verbessern. Zudem sollten Führungskräfte (7) Probleme angehen, um diese zu lösen.

2 Methodik und Durchführung

2.1 Fragestellung

In der vorliegenden Studie wurde ein Coachingprogramm entwickelt, durchgeführt und evaluiert, das auf den erfolgsrelevanten Verhaltensweisen im Coaching beruht, (1) gezielt erfolgsrelevantes Führungsverhalten fördert, (2) die realen Verhaltensänderungen der Führungskräfte im Laufe des Programms objektiv misst und durch ein experimentelles Design den Effekt des Programms auf objektive Zielkennzahlen erfasst.

Darauf basierend untersucht die Studie die folgenden zwei Hauptfragestellungen:

1. **Erfolgsrelevantes Führungsverhalten:** Welche der erfolgsrelevanten Führungsverhaltensweisen sagen Zielerreichung vorher, wenn beides durch objektive Forschungsmethoden gemessen wird?
2. **Effekt des Coachingprogramms:** Erhöht ein wissenschaftlich fundiertes Coachingprogramm objektiv messbar die erfolgsrelevanten Verhaltensweisen? Und erhöht dieses wissenschaftlich fundierte Coachingprogramm objektiv messbar die Zielerreichung?

Neben diesen beiden Hauptfragestellungen wurden in der Studie zwei Methoden entwickelt und auf ihren Nutzen untersucht:

- A) **Videorating:** Können die sieben theoretisch abgeleiteten Führungsverhaltensskalen mit einem ökonomischen Ratingverfahren objektiv und reliabel gemessen werden?
- B) **Ein praktisch implementierbares Coachingprogramm:** Wird das wissenschaftlich fundierte Coachingprogramm in der Praxis von den Teilnehmern angenommen und als hilfreich erlebt?

2.2 Ausgangslage und unternehmerische Rahmenbedingungen

In heutigen Unternehmen gibt es meist eine Vielzahl von etablierten Führungsprogrammen für Führungskräfte. Gleichzeitig hat in viele Unternehmen eine projektorientierte Matrixorganisation Einzug gehalten, in denen viele Mitarbeiter für ihre tagtäglichen Aufgaben in Projekten arbeiten und dort von Projektleitern geführt werden. Obwohl diese Projektführungskräfte daher einen großen Teil der täglichen Mitarbeiterführung vollziehen, sind sie in fast allen Fällen weniger in Führung ausgebildet als die Linienführungskräfte. Die meisten Entwicklungsprogramme für Projektleiter und Projektmanagement-Zertifizierungen fokussieren zu großen Teilen auf fachliche Inhalte und Tools.

Auch bei Stryker erhielten Projektleiter bisher zwar fachliche Projektmanagement-Schulungen, aber keine systematischen Programme zur Entwicklung der persönlichen Führungskompetenz. Seit 2010 wurden bei Stryker Organisationsentwicklungsprogramme angestoßen, um die Anerkennung von Projektleitern weiter zu steigern und damit auch die Attraktivität von Projektleiterkarrieren für talentierte Führungskräfte zu erhöhen. Im Kontext dieser organisatorischen Rahmenbedingungen

wurde ein wissenschaftlich fundiertes Coachingprogramm zur Entwicklung der persönlichen Führungskompetenz von Projektleitern gebraucht.

Das hier vorgestellte Coachingprogramm wurde dafür in 2011 vom Freiburg Institut entwickelt und dann in Zusammenarbeit mit Stryker sowie dem Trainingspartner Change, Leadership & Partners für eine internationale Durchführung weiter spezifiziert. Als externes, uninahes Institut bot das Freiburg Institut das Coachingprogramm neben Stryker auch anderen Firmen auf dem Coachingmarkt an.

2.3 Setting und Stichprobe

Das Coachingprogramm startete im Mai 2012 mit 46 teilnehmenden Projektleitern. 19 Projektleiter kamen aus Deutschland, neun aus der Schweiz und 18 aus den USA. 37 Teilnehmer waren Stryker-interne Projektleiter, die anderen neun Projektleiter kamen aus vier anderen deutschen Organisationen. Stryker ist ein amerikanischer Fortune 500 Medizintechnikkonzern mit Produktions- und Entwicklungsstandorten in USA und Europa. Die neun Stryker-externen Projektleiter kamen aus einer deutschen Tochter eines anderen großen amerikanischen Industriekonzerns, zwei mittelständischen Serviceunternehmen sowie einer Nicht-Regierungs-Organisation. 78% der Projektleiter waren Männer, das durchschnittliche Alter lag bei 37.0 Jahren ($SD = 7.9$) und die durchschnittliche Berufserfahrung lag bei 10.3 Jahren ($SD = 8.3$ Jahre). Alle Projektleiter leiteten während des Coachingprogramms ein aktives Projekt mit mindestens 3 weiteren Mitarbeitern.

2.4 Design der Studie

Die Studie bestand aus drei Phasen: der Phase der Vorstudie, der Phase der Durchführung und begleitenden Evaluation des Coachingprogramms für die Experimentalgruppe sowie der Follow-Up-Phase, in der auch die Wartekontrollgruppe das Coachingprogramm durchlief (s. Abbildung 3).

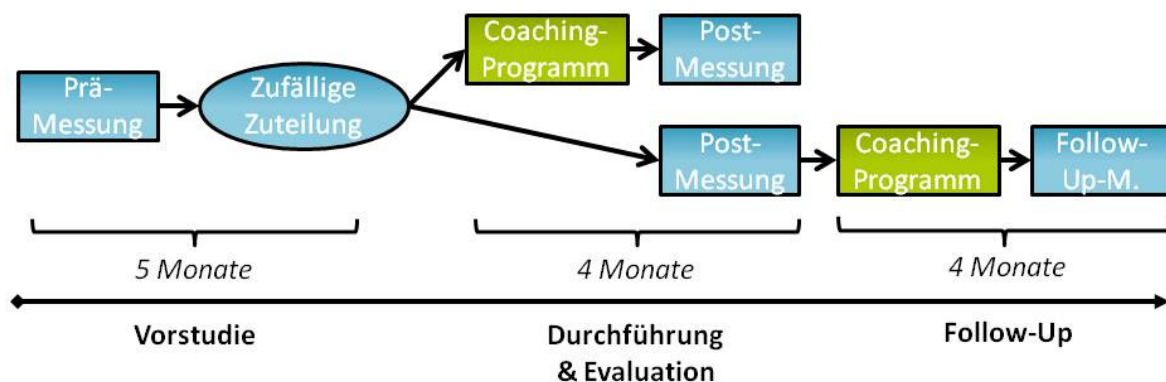


Abbildung 3: Experimentelles Studiendesign in drei Phasen: Vorstudie, Durchführungs-Phase mit & Evaluation sowie Follow-Up-Phase

Die Vorstudie wurde dazu genutzt, das theoretische Modell der erfolgsrelevanten Verhaltensweisen wissenschaftlich zu prüfen und objektiv messbare Verhaltensweisen zu identifizieren, die objektiv messbare Zielerreichung im Projekt vorhersagen. Diese Verhaltensweisen wurden dann durch das Coachingprogramm bei allen beteiligten Führungskräften verstärkt –in der Durchführungs-Phase zunächst bei der Experimentalgruppe (Abbildung 3 oben) und dann in der Follow-Up-Phase bei der Wartekontrollgruppe (Abbildung 3 unten). Die Zuteilung in Experimental- (EG) und Kontrollgruppe (KG) geschah randomisiert: Auf Basis der Prämessung wurden alle 46 Projektleiter durch Parallelisierung in maximal ähnliche Paare eingeteilt und daraufhin je ein Projektleiter des Paares zufällig für die EG und der andere für die KG ausgewählt.

2.5 Forschungsmethoden

Alle drei Messungen zum Prä-, Post- und Follow-Up-Zeitpunkt bestanden aus der unabhängigen Messung von zwei aufeinanderfolgenden ca. 4-wöchigen Projektperioden. Jede der sechs Projektperioden wurde anhand folgender Instrumente untersucht:

- A) Vor Start der Projektperiode:
 - a. **Zielsetzung** anhand des Goal Attainment Scalings (Bergin & Garfield, 1994; Kiresuk, Smith, & Cardillo, 1994) durch den Projektleiter (PL): Definition von 1-3 Projektperiodenzielen inklusive der quantitativen Definition der möglichen Zielerreichungsgüte auf einer 5-stufigen Skala (ungünstig bis bestmöglich)
- B) Am Start der Projektperiode:
 - a. **Rating des Führungsverhaltens** der Projektleiter anhand der Videoaufnahme des ersten Projektteammeetings durch einen unabhängigen trainierten Videoanalysten mittels des dafür entwickelten Freiburger Leadership Behavior Ratings (FLBR).
- C) Am Ende der Projektperiode:
 - a. Bewertung der **Zielerreichung** durch die Führungskraft des Projektleiters (FK) anhand der vorher vom Projektleiter definierten 5-stufigen Skala für die Zielerreichungsgüte

Das Goal Attainment Scaling (GAS) kombiniert zur Messung der abhängigen Variable Zielerreichung im Projekt zwei erfolgreiche Messstrategien und deren Vorteile. Einerseits basiert es auf einer subjektiven Zielsetzung, die garantiert, dass die Ziele für das jeweilige Projekt tatsächlich relevant sind. Andererseits erhebt das GAS wesentlich objektivere Daten als anderen subjektiven Zielerreichungsskalen, da vor Beginn der Projektperiode quantitativ definiert werden muss, anhand welcher messbaren Kriterien die Zielerreichung bewertet werden wird (z.B. durch spezifische Zeit-, Kosten- oder Qualitätskriterien). Aufgrund dieser Kombination von subjektiver Zielsetzung und objektiver Bewertung der Zielerreichung ist das GAS ein etabliertes Forschungsinstrument zur Evaluation von individueller Veränderung in Therapie und Coaching. Um darüber hinaus die Objektivität der Zielerreichungsmessung weiter zu erhöhen, wurde die Bewertung der Zielerreichung nicht vom Projektleiter selbst sondern von dessen Führungskraft eingeholt.

Zur objektiven Messung des Führungsverhaltens anhand von Verhaltensdaten existiert kein den Autoren bekanntes, veröffentlichtes Forschungsinstrument. Daher wurde im Rahmen der Projektstudie das Ratingmanual FLBR entwickelt. Anhand dieses Manuals werten unabhängige Videoanalysten die erfolgsrelevanten Führungsverhaltensweisen (Yukl, 2012) aus, die in einem Projektteammeeting gezeigt und beobachtet werden können. Die wichtigste Einschränkung beruht darauf, dass die extern-orientierten Verhaltensweisen der Projektleiter nicht beobachtet werden können, da die externen Stakeholder in projekt-internen Meetings nicht anwesend sind. Die anderen, insgesamt 54 einzelnen Verhaltensweisen wurden anhand der sieben theoretisch abgeleiteten Verhaltensskalen eingeordnet (s. Abbildung 2). Das grundsätzliche methodische Vorgehen orientiert sich an dem Videoanalysesystem für Verhaltensrating im Coaching (Behrendt 2004, Behrendt 2006; s.a. Grawe, Regli, & Schmalbach, 1994; Hartwanger, 1996).

Vor der Auswertung wurden alle fünf Videoanalysten in einem 5-tägigen Ratertraining auf das 24-seitige Ratingmanual durch den Studienleiter trainiert. Das Training wurde mit einer Prüfung der Interraterreliabilität abgeschlossen, zu deren Bestehen zwei aufeinanderfolgende Videos mit einer Interraterreliabilität von $r_{\text{inter}} > .80$ ausgewertet werden mussten. Für die Auswertung eines Meetings sieht ein Rater das Video zwei Mal komplett durch. Bei der ersten Durchsicht transkribiert der Rater

stichwortartig das verbale und nonverbale Verhalten des Projektleiters. Bei der zweiten Durchsicht bewertet er alle zehn Minuten die 54 Verhaltensweisen auf den dazugehörigen 54 5-stufigen Itemskalen. Für ein maximales Rating auf der Itemskala muss der Projektleiter das jeweilige Führungsverhalten in hoher Intensität und hoher Quantität zeigen.

2.6 Design des Coachingprogramms

Die experimentelle Intervention, die in der Studie als unabhängige Variable untersucht werden sollte, bestand aus einem umfassenden Coachingprogramm. Das Coachingprogramm war darauf ausgerichtet, die erfolgsrelevanten Verhaltensweisen von Führung zu verstärken.

Das entwickelte Coachingprogramm setzte sich aus drei Modulen zusammen: einer zwei-tägigen Gruppensitzung zum Start und danach zwei ein-stündigen Einzelsitzungen jeweils im Abstand von zwei Monaten. In diesen zwei Monaten wurde für die Post- bzw. Follow-Up-Messung jeweils eine Projektperiode ausgewertet. Die Auswertungen der Projektperioden dienten zudem als Feedback zu Fortschritten des Projektleiters in den Einzelsitzungen des Coachingprogramms. Die Inhalte der drei Module kombinierten vier zentrale Methoden der Verhaltensänderung zur Verstärkung der erfolgsrelevanten Verhaltensweisen der Führung: persönliche Zielsetzung, individuelles Feedback, Training und Coaching.

- 2-tägige Gruppensitzung zum Start des Programms:
 - a. Trainingseinheit zu den erfolgsrelevanten Verhaltensweisen
 - b. Video-Bestpractices
 - c. Persönliches Feedback: Ergebnisreport 1 auf Basis der Prä-Messung
 - d. Peer-Coaching in Kleingruppen
 - e. Praktischen Übungen mit persönlichem Feedback
 - f. Persönliche Zielsetzung zur Veränderung in der Praxis für die nächsten zwei Monate
- Zwei 1-stündige Einzelsitzungen in den vier Monaten nach dem Start des Programms:
 - a. Persönliches Feedback: Ergebnisreport 2 und 3 auf Basis der Post- bzw. Follow-Up-Messungen
 - b. Persönliches Video-Feedback mit Fokus auf persönliche Best-Practices
 - c. Coaching zu persönlichen Themen und Fragen
 - d. Persönliche Zielsetzung zur Veränderung in der Praxis für die nächsten zwei Monate (erste Einzelsitzung)
 - e. Persönliches Abschluss-Feedback des Klienten zum Programm (letzte Einzelsitzung)

Um eine hohe Wirksamkeit des Coachingprogramms zu gewährleisten, wurden in dem Design des Programms die aktuellen Erkenntnisse aus der Coachingforschung umgesetzt. Daher orientierten sich alle Methoden an den fünf erfolgsrelevanten Verhaltensweisen im Coaching. Für die fünf erfolgsrelevanten Verhaltensweisen im Coaching (s. Abschnitt 1.1) hieß das im Einzelnen:

- (1) **Ressourcen aktivieren:** In allen Feedbacks wurde besonderer Wert auf eine Stärkenorientierung gelegt. Die quantitative Aufbereitung der Feedback-Reports beinhaltete für jeden Projektleiter persönliche Stärken und persönliche Entwicklungspotenziale ohne Vergleich zu einer objektiven Norm. Für das Video-Feedback wurden die besten Momente des Projektleiters ausgewählt. Insbesondere bei unterentwickelten Verhaltensweisen wurden Sequenzen gezeigt, in denen das Verhalten bereits in guten Ansätzen sichtbar war, um den Projektleiter ein positives Leitbild zu geben, welches Verhalten sie mehr zeigen

sollten. Der Vorteil dieses Vorgehens war, dass diese Verhaltensansätze nicht neu entwickelt, sondern nur verstärkt und ausgebaut werden mussten. Für die Peercoachings bekamen die Teilnehmer eine klare ressourcen-orientierte Anleitung und Einführung.

- (2) **Kooperativ begleiten:** Die gesamte Programmkonzeption war sehr interaktiv ausgerichtet und eröffnete den Teilnehmer viele Gelegenheiten zum Mitgestalten. Die rein frontale Wissensvermittlung beschränkte sich auf einen 1-stündigen ersten Input für die theoretischen Grundlagen. Die Inputs zu einzelnen Verhaltensweisen wurden den Teilnehmern übertragen, die in der Prä-Messung dieses Verhalten am erfolgreichsten gezeigt hatten. In den Kleingruppen-Übungen und Peercoaching konnten die Teilnehmer aktuelle Anliegen einbringen bzw. ihre Peers in verantwortungsvoller Rolle als Peer-Coach aktiv begleiten. In den Einzelsitzungen des Coachingprogramms wurde die inhaltliche Schwerpunktsetzung vollständig auf die persönlichen Anliegen, Fragen und Themen des Klienten ausgerichtet.
- (3) **Prozessführung geben:** Auch wenn die konkreten Inhalte stark auf die Anliegen des Teilnehmer ausgerichtet wurden, war das methodische Vorgehen stets klar strukturiert. Für die Peercoachings und Übungen wurden detaillierte Anleitungen ausgegeben, ebenso für die persönliche Auswertung des Feedbacks. Auch die Einzelsitzungen im Coaching folgten einer klaren transparenten Struktur: (A) Besprechung der Anliegen für die Sitzung, (B) Besprechung der bisherigen Fortschritte, (C) Video-Feedback mit Auswertung, (D) nächste Schritte vereinbaren. Innerhalb dieser klaren Struktur konnten die Teilnehmer ihre eigenen inhaltlichen Schwerpunkte setzen (s. (2) Kooperativ begleiten).
- (4) **Praxistransfer unterstützen:** Um den Praxistransfer zu unterstützen wurden zu allen erfolgsrelevanten Verhaltensweisen konkrete Tools an die Hand gegeben, anhand derer die Teilnehmer die Umsetzung in Ihren Projektteams planen konnten. Zudem wurden durch die Video-Best-Practices den Projektleitern viele realitätsnahe Rollenmodelle an die Hand gegeben, an denen sie sich orientieren konnten. Am Ende jeder der drei Sitzungen wurden von allen Teilnehmern konkrete persönliche Umsetzungsschritte festgehalten und öffentlich gemacht. Als zentrales Instrument der Transferunterstützung ist darüber hinaus die Messung nach dem Coachingprogramm zu werten: Alle Projektleiter wussten, dass sie im Rahmen der Post- bzw. Follow-Up-Messung noch zweimal in ihrer realen Praxis gefilmt und darauf basierend ein persönliches Feedback zu ihren Fortschritten erhalten würden. Auf Basis dieser Auswertungen wurde der persönliche Praxisfortschritt in beiden Einzelsitzungen besprochen und weiter unterstützt.
- (5) **Feedback einholen:** Nach dem ersten Tag und am Ende der Gruppensitzung, sowie in der letzten Einzelsitzung des gesamten Programms wurden ausführliche Rückmeldungen aller Teilnehmer eingeholt. Am Ende der Gruppensitzung und am Ende des gesamten Programms wurden diese persönlichen Rückmeldungen durch eine schriftliche und quantitative Evaluation ergänzt.

3 Ergebnisse

3.1 Die Raterreliabilität der Messung des Führungsverhaltens mittels FLBR

Zur Berechnung der Raterreliabilitäten für intervall-skalierte Daten ist nach Wirtz & Caspar (2002) der Intraklassenkoeffizient ICC heranzuziehen. Der ICC_{unjust}-Koeffizient kontrolliert neben der Konsistenz

der Rating-Streuung auch die Konsistenz der Mittelwerte. Der ICC(3,1)unjust wurde gewählt, da die Rater weder zufällig ausgewählt noch gepooled waren.

Zur Messung der Intraraterreliabilität musste jeder der fünf Rater ein zufällig ausgewähltes Video aus seiner ersten Auswertungsphase mindestens drei Monate später erneut auswerten. Alle fünf Intrarater-Koeffizienten lagen über $ICC_{unjust.intra} = .88$ mit einem Mittelwert von $ICC_{unjust.intra} = .96$. Damit war die Konsistenz der wiederholten Ratings desselben Raters nahezu perfekt.

Die Interrater-Reliabilität wurde durch ein Rating derselben Videoaufnahme durch alle zu dem Studienzeitpunkt aktiven Rater gemessen. Aus jeder Studienphase wurden dafür zufällig Videos ausgewählt. Insgesamt wurden dadurch 7 Videos von durchschnittlich 3,7 Ratern ausgewertet. Die durchschnittliche Interrater-Reliabilität lag bei $ICC_{unjust.inter} = .71$. Damit ist die Konsistenz der Ratings unterschiedlicher Rater als hoch einzustufen.

3.2 Rücklaufquoten und Drop-Out

Die Rücklaufquote für die GAS-Zielsetzung lag über die gesamte Studie hinweg bei 64%, für die GAS-Zielbewertung bei 56%. Die aufgenommenen Projektmeeting-Videos wurden zu 100% ausgewertet.

Aufgrund der langen Laufzeit des Projekts und der Anforderung, dass alle Projektleiter ein aktives Projekt benötigen, gab es über die 13 Monaten hinweg immer wieder Missings und Drop-outs. Diese waren im Einzelnen begründet in ad-hoc gecancelten Meetings, Terminkonflikten z.B. durch Urlaube, vorzeitigen Projektabschlüssen während der Studienlaufzeit und individuellen Prioritätskonflikten. Dadurch wurden letztendlich $n_{ges} = 170$ Projektperioden ausgewertet. Diese verteilen sich auf die Studienphasen und Messzeitpunkte wie folgt: $n_{prä1} = 46$, $n_{prä2} = 43$, $n_{postEG1} = 19$, $n_{postKG1} = 15$, $n_{postEG2} = 16$, $n_{follow-upKG1} = 16$, $n_{follow-upKG2} = 15$. Daraus wird deutlich, dass das Drop-Out v.a. nach der Vorstudie vor dem Start des Coachingprogramms entstand und sich zahlenmäßig gleich auf die EG und KG verteilte.

3.3 Vorstudie: 1. Fragestellung - Analyse des erfolgsrelevanten Führungsverhaltens

Für die Vorstudie lagen $n_{prä} = 89$ Projektperioden vor. Da die experimentelle Intervention noch nicht eingeführt war, wurden die Verhaltens- und Zielerreichungsdaten in einer korrelativen Analyse ausgewertet. Die Testung erfolgte anhand von Pearson's Korrelationskoeffizient, zwei-seitig und mit einem α -Signifikanz-Niveau von 5%. Die Produkt-Moment-Korrelationen zwischen der Zielerreichung und 13 der Verhaltensweisen wurden signifikant (s. Abbildung 4). Zehn der 13 signifikanten Verhaltensweisen ordnen sich den drei Verhaltensskalen zu, die auch im Coaching zu den drei erfolgsrelevanten Verhaltensskalen zählten: (1) Prozessführung geben, (2) Kooperation fördern und (3) Ressourcen aktivieren. Darüber hinaus wurden drei weitere Verhaltensweisen signifikant: persönliche Beiträge der Mitarbeiter klären, externe Unterstützung sicherstellen und Risiken identifizieren. Diese drei ordnen sich den drei folgenden Verhaltensskalen der Führung zu: (4) Verständnis vertiefen, (6) Handlungskompetenz verbessern und (7) Probleme angehen.

Zusammenfassend kann daher gesagt werden, dass 13 erfolgsrelevante Verhaltensweisen identifiziert werden konnten, die auch anhand objektiver Messungen Zielerreichung im Projekt vorhersagen.

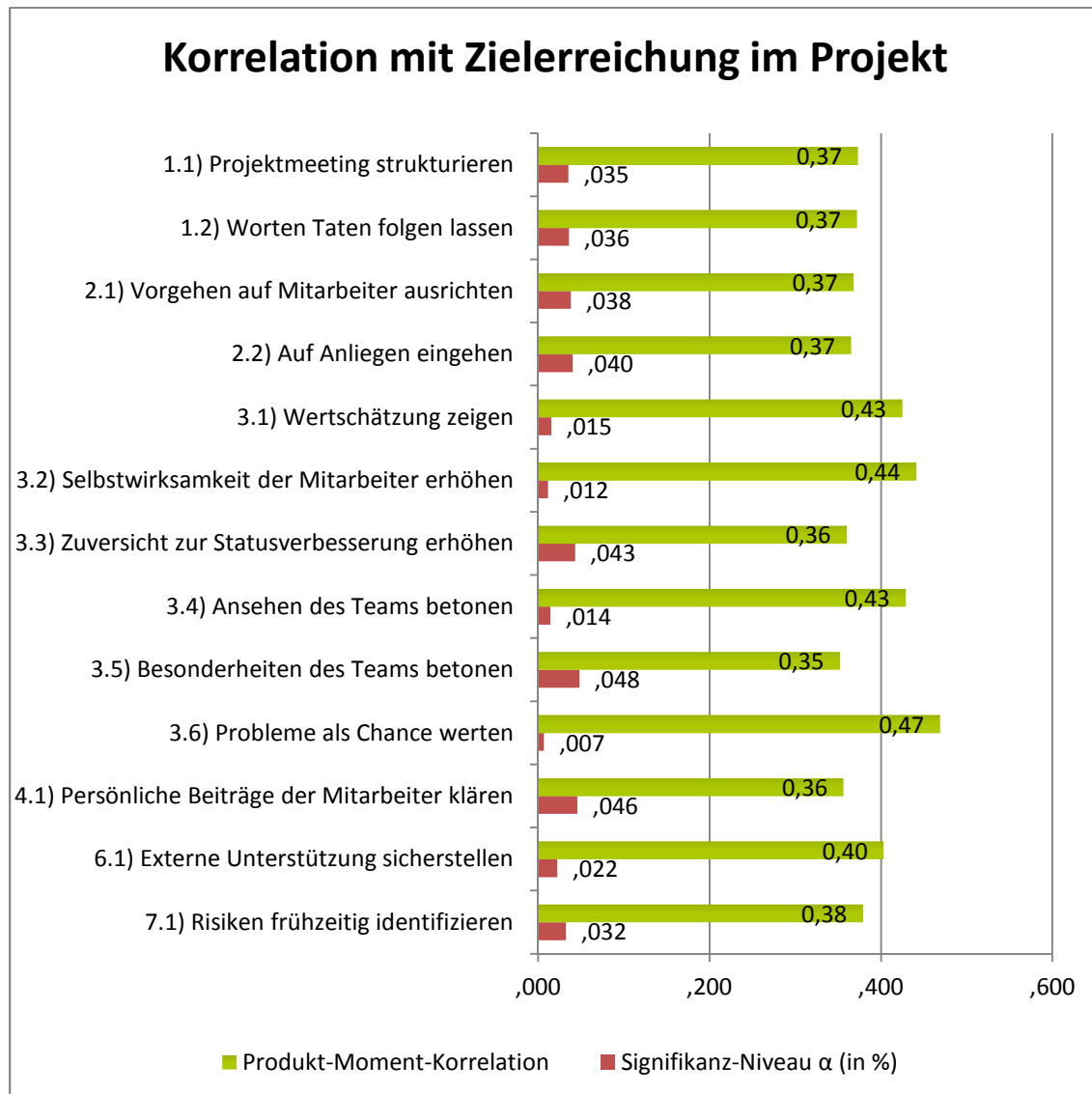


Abbildung 4: Ergebnisse der Vorstudie –13 signifikante Produkt-Moment-Korrelationen zwischen Verhaltensweisen aus der Videoanalysen FLBR sowie der Zielerreichung im Projekt (GAS bewertet durch die Führungskraft des Projektleiters); ausgewertet wurden $n_{\text{prä}}=89$ Projektperioden; vollständige Datensätze mit Zielsetzung und Bewertung der Zielerreichung lagen für $n=32$ Projektperioden vor. Die Verhaltensweisen 1.1-1.2 ordnen sich der Verhaltensskala (1) Prozessführung geben zu, die Verhaltensweisen 2.1-2.2 der Skala (2) Kooperation fördern und die Verhaltensweisen 3.1-3.6 der Verhaltensskala (1) Ressourcen aktivieren.

3.4 2. Fragestellung - Analyse des Effekts des Coachingprogramms

Entsprechend der Analysen der Vorstudie wurde das folgende Coachingprogramm darauf ausgerichtet, die 13 objektiv messbaren, erfolgsrelevanten Verhaltensweisen bei den teilnehmenden Projektleitern zu fördern. Dementsprechend fokussierten sich die Trainingseinheit, Best-Practices, Feedbackreports, Übungen und Coachings auf diese 13 Verhaltensweisen und die Projektleiter wurden angehalten, sich konkrete Ziele zu setzen, um diese 13 Verhaltensweisen auszubauen.

Der Effekt des Coachingprogramms auf die Verhaltensweisen der Projektleiter wurden anhand von Mehrebenenanalysen (Raudenbush & Bryk, 2002) überprüft (s. Tabelle 1). Die abhängige Variable der Modellrechnungen war jeweils eine der 13 Verhaltensweisen. Dabei wurden für jede der 13 Verhaltensweisen sechs Modelle gerechnet und verglichen: zunächst die drei Basismodelle der Mehrebenenanalyse, das ‚Unconditional Means‘- das ‚Random Intercept‘-, sowie das ‚Random

Intercept & Random Slope'-Modell, und dann die drei gleichen Modelle mit zusätzlichem fixen Effekt durch das Coachingprogramm. Die sechs Modelle wurden jeweils mittels der beiden Informationskriterien AIC und BIC bzgl. ihrer Güte verglichen. Das AIC testet sensitiver auf die erklärte Varianz, das BIC sensitiver auf die Sparsamkeit des Modells. Von den drei Modellen mit Prüfung des fixen Programmeffekts wurde das 'Unconditional Means'-Modell in allen 13 Fällen am besten bewertet, so dass im Folgenden nur dieses berichtet wird (Tabelle 1).

Tabelle 1 zeigt, dass die experimentelle Intervention des Coachingprogramms dazu führte, dass die Projektleiter tatsächlich alle 13 Verhaltensweisen in der Tendenz ausbauen konnten. Die Unterschiede sind je nach Verhaltensweisen jedoch sehr unterschiedlich groß. Bei fünf der Verhaltensweisen ist der Anstieg bei einem α -Niveau von 5% signifikant, bei zwei weiteren bei einem Signifikanz-Niveau von 10%. Die frühzeitige Identifikation von Risiken (Verhalten 7) nimmt besonders deutlich zu: Hier steigt der Einsatz von geschätzten 18.3% vor dem Training um geschätzte 13.4% auf insgesamt 31.7% nach dem Training.

Zusammenfassen kann festgehalten werden, dass das Coachingprogramm die 13 Verhaltensweisen verstärkt, fünf von diesen in signifikantem Umfang.

	Intercept		Program		AIC/BIC
	Est.	(SE)	Est.	(SE)	
1.1) Projektmeeting strukturieren	0.429***	(0.018)	0.006	(0.025)	-
1.2) Worten Taten folgen lassen	0.528***	(0.007)	0.012	(0.012)	-
2.1) Vorgehen auf Mitarbeiter ausrichten	0.708***	(0.019)	0.051*	(0.023)	-
2.2) Auf Anliegen eingehen	0.938***	(0.010)	0.002	(0.015)	-
3.1) Wertschätzung zeigen	0.232***	(0.022)	0.081**	(0.024)	+
3.2) Selbstwirksamkeit der Mitarbeiter erhöhen	0.081***	(0.014)	0.046*	(0.019)	+
3.3) Zuversicht zur Statusverbesserung erhöhen	0.527***	(0.008)	0.021°	(0.012)	-
3.4) Ansehen des Teams betonen	0.542***	(0.011)	0.007	(0.012)	-
3.5) Besonderheiten des Teams betonen	0.554***	(0.009)	0.004	(0.013)	-
3.6) Probleme als Chance werten	0.543***	(0.011)	0.028°	(0.015)	-
4.1) Persönliche Beiträge der Mitarbeiter klären	0.200***	(0.018)	0.097***	(0.026)	+
6.1) Externe Unterstützung sicherstellen	0.541***	(0.007)	0.011	(0.013)	-
7.1) Risiken frühzeitig identifizieren	0.183***	(0.017)	0.134***	(0.028)	+

Tabelle 1: Mehrebenenanalyse des Trainingseffekts auf die 13 erfolgsrelevanten Führungsverhaltensweisen der Projektleiter, n = 144 Projektperioden; Projektleiter, die das Projekt zwischenzeitlich gewechselt hatten oder nicht an dem zufällig zugewiesenen Coachingprogramm teilnahmen, wurden von dieser Analyse ausgeschlossen. Die Bewertung der Zielerreichung wurde umkodiert, um einen standardisierten Range von 0 (= minimales Verhaltensrating) bis 1 (maximales Verhaltensrating) zu erreichen. Est. = Schätzung des Parameters; SE = Standardabweichung; ° = $p < .1$, * = $p < .05$, ** = $p < .01$, *** = $p < .001$. AIC/BIC-Spalte: += das berichtete Modell

wurden von allen sechs getesteten Modelle anhand der Informationskriterien am besten bewertet; -- eines der drei Basismodelle ohne fixen Effekt wurden anhand der Informationskriterien besser bewertet, als das berichtete Modell mit fixem Effekt für das Coachingprogramm

Der Programmeffekt auf die abhängige Variable, Zielerreichung im Projekt, wurde ebenfalls anhand einer Mehrebenenanalyse (Raudenbush & Bryk, 2002) ermittelt (s. Tabelle 2). Zunächst wurden die drei Basismodelle der Mehrebenenanalyse anhand der Informationskriterien AIC und BIC geprüft. Danach wurde die unabhängige Variable Coachingprogramm sowie die mögliche Kovariable nationaler Standort getestet in wiederholten Modellrechnungen getestet. Der nationale Standort wurde für die USA und die Schweiz jeweils dummy-kodiert. Die beiden Kovariablen wurden in keinem Modell signifikant, so dass diese verworfen wurden. Von allen getesteten Modellen wurde das Random Intercept Modell mit der unabhängigen Variable Coachingprogramm am besten bewertet. Das reine Unconditional Means Modell erwies sich zwar anhand des BIC als etwas sparsamer, war aber anhand des AIC deutlich weniger Varianzaufklärend. In dem finalen Mehrebenenanalysemodell (s. Tabelle 2 ganz rechts) wurde die unabhängige Variable Coachingprogramm bei einem α – Signifikanz-Niveau von 5% signifikant. Der fixe Parameter wurde dabei auf 0.25 geschätzt, was einer sehr großen Effektstärke von $d = 1.30$ entspricht (Bortz & Schuster, 2010). Das bedeutet, dass ein Projektleiter durch das Coachingprogramm von seiner Führungskraft eine 25% höhere Zielerreichung im Projekt attestiert bekommt (auf einer Skala von 0% bis 100%).

Damit kann festgehalten werden, dass die experimentelle Intervention des Coachingprogramms die objektiv messbare Zielerreichung signifikant um 25% erhöht.

	Unconditional Means Model		Random Intercept Model		Random Intercept & Random Slope		Random Intercept Model & Program	
	Est.	(SE)	Est.	(SE)	Est.	(SE)	Est.	(SE)
Intercept	0.70***	(0.028)	0.70***	(0.039)	0.70***	(0.039)	0.71***	(0.037)
Repeated Measures			0.00	(0.021)	0.00	(0.021)	-0.07*	(0.034)
Random Repeated Measures					0.00	(0.000)		
Program							0.25*	(0.097)
AIC	-9.5		-7.5		-5.5		-11.8	
BIC	-3.3		0.7		4.7		-1.6	

Tabelle 2: Mehrebenenanalyse des Trainingseffekts auf die Zielerreichung im Projekt; $n = 56$; Projektleiter, die das Projekt zwischenzeitlich gewechselt hatten oder nicht an dem zufällig zugewiesenen Coachingprogramm teilnahmen, wurden von dieser Analyse ausgeschlossen. Die Bewertung der Zielerreichung wurde umkodiert, um einen standardisierten Range von 0 (= ungünstiges Ergebnis) bis 1 (bestmögliches Ergebnis) zu erreichen. Niedrigere AIC- und BIC-Werte zeigen eine höhere Modellgüte an. Est. = Schätzung des Parameters; SE = Standardabweichung; * = $p < .05$, ** = $p < .01$, *** = $p < .001$.

3.5 Zusatzfragestellung A - Prüfung der Verhaltensskalen und Entwicklung einer ökonomischen Kurzversion des FLBR

Zur Prüfung und Optimierung des neu entwickelten Videoratings FLBR wurden die sieben theoretisch abgeleiteten Skalen des erfolgsrelevanten Führungsverhaltens (s. Abbildung 2) faktorenanalytisch überprüft. Aufgrund der erstmaligen Prüfung wurden die Skalen zunächst durch eine explorative

Faktorenanalyse³ anhand der Daten der Vorstudie untersucht. Die daraufhin optimierten Skalen wurden dann anhand der Daten der Post- und Follow-Up-Messungen durch eine konfirmatorische Faktorenanalyse⁴ (Byrne, 2010) cross-validiert.

Um eine ökonomische Version des FLBR zu erhalten, wurde die Skalenlänge im ersten Schritt nach der explorativen Faktorenanalyse auf drei Items pro Verhaltensskala begrenzt. Neben den Ergebnissen der Faktorenanalyse wurden auch die Ergebnisse einer umfassenden Itemanalysen zur Auswahl der drei finalen Items berücksichtigt. Angestrebt wurde eine hohe Kriteriumsvalidität zur Vorhersage der Zielerreichung im Projekt, eine mittlere Itemschwierigkeit und hohe Standardabweichung, sowie eine hohe interne Konsistenz der Skalen.

Die konfirmatorische Faktorenanalyse zur Cross-Validierung der sieben Verhaltensskalen bestätigte diese mit zufriedenstellenden bis sehr guten Model-Fit-Werten (Hu & Bentler, 1999; Jöreskog, 1993). Die Skalenkonsistenz der resultierenden Skalen war mit einem Cronbach's Alpha von $\alpha = .44$ bis $\alpha = .65$ jedoch nur mittlerer Höhe.

Damit kann festgehalten werden, dass die sieben theoretisch abgeleiteten Verhaltensskalen der Führung ökonomisch, objektiv und von Ratern konsistent gemessen werden können. Die interne Konsistenz der Skalen ist im Vergleich zu typischen Fragebogenskalen jedoch kleiner. Daher sollten für wissenschaftliche Verhaltensanalysen immer auch die einzelnen Verhaltensweisen berichtet werden, wie dies bei allen vorangegangenen Auswertungen der Ergebnisse erfolgte.

3.6 Zusatzfragestellung B – praktischer Nutzen und Evaluation des wissenschaftlich fundierten Coachingprogramms

Am Ende des Coachingprogramms wurde von den teilnehmenden Projektleitern, sowie deren Teams und Führungskräften ein quantitatives und qualitatives Feedback eingeholt (s. Abbildung 5 und Tabelle 3).

Die quantitative Bewertung der Projektleiter war sehr gut. Sie bewerteten das Programm als wertvolles Investment in ihre Karriere (durchschnittliche Schulnote 1,44) und empfahlen ihrem Unternehmen mehr vergleichbare Angebote einzuführen (Schulnote 1,75). Beachtenswert ist außerdem, dass von den Führungskräften und Mitarbeitern der Projektleiter ebenfalls ein guter Fortschritt in der Führung des Projektleiters bemerkt wurde, obwohl diese nicht in dem Coaching involviert waren (Schulnoten 2,04 bzw. 2,27).

³ Schätzung mittels Hauptachsen-Methode; Rotation: Oblimin direkt; Selektionskriterium: Eigenwert > 2

⁴ Parameter-Schätzung anhand der Maximum-Likelihood-Methode

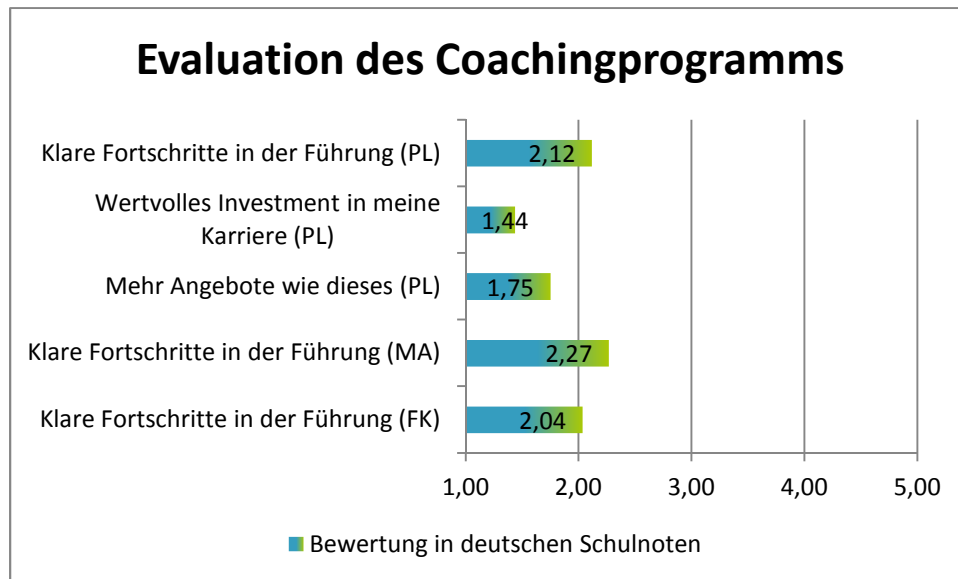


Abbildung 5: Evaluation des Coachingprogramms nach der letzten Coachingsitzung, PL = Feedback des teilnehmenden Projektleiters, MA = Feedback der Mitarbeiter des Projektleiters, FK = Feedback der Führungskraft des Projektleiters

Die qualitativen Rückmeldungen der teilnehmenden Projektleiter, sowie der Führungskräfte und Mitarbeiter unterstützen und untermauern die positive Bewertung (s. Tabelle 3). In einigen Kommentaren wurde der herausragende Erfolg des Coachingprogramms hervorgehoben. Auch in den qualitativen Kommentaren der Mitarbeiter und Führungskräfte lässt sich der spezifische Effekt des Coachingprogramms ablesen. Obwohl den nicht beteiligten Personen die Inhalte des Coachings nicht kommuniziert wurden, finden sich die angestrebten Verhaltensweisen in den positiven Feedbacks wieder (z.B. mehr Zuversicht, klarere Struktur, höhere Kooperation). Besonders deutlich wurde das positive Feedback in einem zufällig aufgenommenen Feedback einer Projektleiterin, deren letztes Coaching für eine andere Studie gefilmt wurde. Dieses ist für Sie über den folgenden Link verfügbar: [Lifefeedback online](https://lifecyclefeedback.com/).

	Qualitative Feedbacks zu den Fortschritten durch das Programm
Projektleiter	- The program was, by far, the best project management training I have ever attended. The meeting recordings, interaction with other PMs and applicability to my every day work were all phenomenal. The personal coaching sessions were just as helpful. “Overall I felt the training was exceptional and was at the same level with the Instruments Leadership Academy in the excitement I feel leaving with tangible tools to affect the team, project, and outcomes.”
Mitarbeiter	“My manager is more confident and can redirect the team when we get off task!” “Better collaboration and organization of meetings. Driving a clearer plan for the team.”
Führungskräfte	„Der Projektleiter bindet die Teammitglieder mehr ein; führt das Team zielorientierter und macht, wenn es sein muss, klare Vorgaben.“ „The Project Manager made great progress in leading the team and aligning the broader group across functions and departments to achieve the project goals. He also established a mentality of integrity trust and commitment in the team that helped the team when it was facing some significant pushbacks.“

Tabelle 3: Qualitative Feedbacks zu den Fortschritten durch das Coachingprogramm von den teilnehmenden Projektleitern, sowie deren Mitarbeitern und Führungskräften

Zusammenfassend kann festgehalten werden, dass das wissenschaftlich fundierte Coachingprogramm von den beteiligten Projektleitern als auch den indirekt betroffenen Mitarbeitern und Führungskräften sehr gut angenommen und als sehr förderlich für ihre Praxis bewertet wurde.

4 Diskussion

Durch die vorliegende Studie konnten 13 Verhaltensweisen der Führung identifiziert werden, die objektiv messbare Zielerreichung vorhersagen. Zudem konnte ein Coachingprogramm etabliert werden, das sich an den wissenschaftlich fundierten Verhaltensweisen orientiert, die im Coaching Erfolg vorhersagen. Es konnte gezeigt werden, dass dieses Coachingprogramm fünf erfolgsrelevante Verhaltensweisen der Führung signifikant verstärkt, die objektiv messbare Zielerreichung um 25% steigert und auch in der Praxis sehr gut angenommen und bewertet wird. Darüber hinaus wurde mit dem FLBR ein Videorating etabliert, das die theoretisch abgeleiteten Verhaltensskalen der Führung ökonomisch, objektiv und mit hoher Raterreliabilität misst.

Die methodischen Stärken der wissenschaftlichen Studie sind:

- **Hohe interne Validität durch das experimentelle Design** mit vollständiger Randomisierung und Wartekontrollgruppe
- **Hohe Objektivität der Messverfahren** durch die Messung von Verhalten mittels Videorating und Bewertung der Zielerreichung anhand objektiver Kennzahlen, bewertet durch einen unabhängigen Dritten (die Führungskraft des Projektleiters)
- **Hohe Objektivität der Verhaltensmessung** durch eine nahezu perfekte Intra- und hohe Interraterreliabilität
- **Hohe Bedeutsamkeit** durch die Etablierung von großen Effekten auf praktisch relevante Variablen: z.B. 25% höhere Zielerreichung durch das Coachingprogramm (Kirkpatrick, 1976)
- **Hohe Nachhaltigkeit der untersuchten Effekte** durch Post- bzw. Follow-Up-Messungen zwei und vier Monate nach der experimentellen Hauptintervention
- **Hohe statistische Validität** durch die Verwendung von Mehrebenenanalysen
- **Hohe externe Validität** durch eine Studie im Feld mit Teilnehmern aus drei verschiedenen Nationen und zwei Kontinenten

Die methodischen Verbesserungspotenziale der wissenschaftlichen Studie sind:

- **Eingeschränkte externe Validität für Führung außerhalb von Projekten und in anderen Unternehmen** durch die Untersuchung von Projektleitern, die zu einem Großteil aus einem Unternehmen stammten
- **Eingeschränkte externe Validität für die Führung außerhalb von Teammeetings** durch die Beschränkung der Verhaltensanalysen auf interne Projektteammeetings
- **Eingeschränkte Bedeutsamkeit** dadurch, dass das vorgestellte Coachingprogramm nur fünf der 13 erfolgsrelevanten Verhaltensweisen signifikant erhöhen konnte
- **Mittlere interne Konsistenz der theoretisch abgeleiteten Verhaltensskalen** der Führung durch mittlere Cronbach's Alpha der Verhaltensskalen im FLBR
- **Eingeschränkte statistische Validität** durch eine relativ hohe Anzahl von Missings, begründet durch die aufwendige, lang-dauernde Untersuchung
- **Eingeschränkte statistische Validität** durch ein niedriges n, insbesondere in der Vorstudie zur Identifizierung der erfolgsrelevanten Verhaltensweisen

4.1 Ausblick für die Wissenschaft

Da es sich bei der vorliegenden Studie um einen ganz neuen Forschungs- und Methodenansatz im Coaching handelt, bedürfen die Ergebnisse einer Replikation in weiteren Studien. Diese sollten an unterschiedlichen Stichproben in unterschiedlichen Settings durchgeführt werden, um die Einschränkungen der externen Validität zu adressieren. Besonders interessant sind Studien mit Linienführungskräften in anderen Unternehmen, die das Verhalten außerhalb von Teammeetings erfassen. Darüber hinaus gilt es die neu entwickelten Methoden weiter zu optimieren. Hierbei gilt es insbesondere zu untersuchen, ob die niedrigeren Konsistenzen der Verhaltensskalen in der grundsätzlichen Methode begründet sind oder ob sich auch für Videoratings Verhaltensskalen konstruieren lassen, die nicht nur valide, objektiv und rater-reliabel, sondern darüber hinaus auch intern konsistent sind.

Um das Verständnis der erfolgreichen Prozesse in Coaching und Führung weiter zu differenzieren, erscheint es außerdem vielversprechend, Coachingprogramme zu untersuchen, die spezifischer auf einzelne Verhaltensweisen der Führungskräfte ausgerichtet sind. Das würde ermöglichen, Verbesserungen in der Zielerreichung nicht nur auf ein erfolgreiches Coachingprogramm, sondern auch eindeutig auf einzelne Verhaltensweisen der Führung zurückzuführen.

Der generelle Forschungsansatz erscheint im Lichte der positiven Ergebnisse sehr fruchtbar und könnte daher auf andere Bereiche ausgeweitet werden. Neben dem weit verbreiteten Führungcoaching könnten auch andere Coachingformen profitieren, wenn detaillierte Verhaltensmodelle für die Klienten entwickelt und durch fundierte Methoden messbar gemacht würden – als ein interessanter Kontext sei z.B. Coaching im Bereich Stress- oder Burnout genannt.

4.2 Ausblick für die Coachingpraxis

Für die Coachingpraxis wurden mit der vorliegenden Studie zwei sehr vielversprechende Methoden etabliert: das wissenschaftlich fundierte Coachingprogramm sowie eine fundierte Verhaltensanalyse zu erfolgsrelevantem Führungsverhalten.

Für Unternehmen, denen eine hohe Zielerreichung ihrer Projektleiter und Führungskräfte am Herzen liegt, bietet die vorliegende Studie ein umfassendes Coachingprogramm mit erprobtem Programmdesign. Dieses kann so übernommen oder als Anregung für die Gestaltung eigener zielgenauer Programme genutzt werden. Dabei sollten insbesondere die systematische Ausrichtung an den erfolgsrelevanten Verhaltensweisen im Coaching sowie die fundierte Verhaltensanalyse der Klienten bedacht werden.

Für Coaches bietet die vorliegende Studie zwei neue Möglichkeiten. Zum einen liefert die Studie weitere Belege dafür, dass es erfolgsversprechend ist, Coachingangebote systematisch auf die erfolgsrelevanten Verhaltensweisen von Coaching auszurichten. Die Ansätze von Behrendt (2006) und Greif (2010) liefern dazu fundierte Ansatzpunkte. Es existieren dazu etablierte Verhaltensanalysen, die für Coaches als wertvolles Feedback genutzt werden können. Zum anderen liegt mit der FLBR erstmals eine fundierte Verhaltensanalyse vor, die Führungcoaches einsetzen können, um ihr Coaching noch gezielter auf die erfolgsrelevanten Aspekte des Führungsverhaltens auszurichten und ihren Klienten fundiertes, wirksames Feedback zu geben.

4.3 Zusammenfassung

Die vorliegende Studie folgt dem Ruf nach verstärkter Forschung im Coaching, nach höheren wissenschaftlichen Standards in der Coachingforschung (Künzli, 2005) und nach einer verstärkten wissenschaftlichen Fundierung der Coachingpraxis (Anders, 2002). Die Studie bricht in mehrerer Hinsicht mit existierenden Traditionen und eröffnet so der Coachingforschung und der Coachingpraxis neue Möglichkeiten:

Die Studie verbindet in innovativer Weise kontrollierte Wissenschaft mit zielorientierter Praxis. Künzli konstatiert, dass „Forschung von Praktizierenden für ihre Arbeit mit Klienten als wenig hilfreich wahrgenommen wird“ (Künzli 2009, S. 11). Die vorliegende Arbeit belegt das Gegenteil. Die wissenschaftliche Fundierung und die anspruchsvollen methodischen Kontrollen haben die Wirksamkeit des Coachingprogramms nicht untergraben. Im Gegenteil: Die Ausrichtung an wissenschaftlichen Erkenntnissen in der Gestaltung des Coachingprogramms und insbesondere die wissenschaftliche Verhaltensanalyse des Führungsverhaltens haben dazu geführt, dass das Coachingprogramm signifikante Effekte erzeugte. Wichtige Teile des Coachingprogramms wären ohne die Videoanalyse nicht möglich gewesen: die detaillierten Feedbackreports, die persönlichen Video-Feedback und auch die vielen Video-Best-Practices von anderen erfolgreichen Projektleitern und ganz besonders die systematische Ausrichtung an den 13 erfolgsrelevanten Verhaltensweisen.

Die stärkste Veränderung liegt dabei sicherlich in der Öffnung der Black Box Führung. Das Coachingprogramm bricht mit der Tradition von vertraulichem Coaching hinter verschlossenen Türen, in denen Coaches Führungskräften Hilfe zur Selbsthilfe geben. Der vorliegende Ansatz erfordert, dass Führungskräfte das Coaching Ihrem Team gegenüber transparent machen, sich mittels Kamera sehr detailliert in die Karten schauen lassen und dabei einer objektiven Verhaltensanalyse stellen (s.a. Van der Weide & Wilderom, 2004).

Die wichtigste Erkenntnis für die Praxis dabei ist: Coaching wirkt – Führungsverhalten kann durch wissenschaftlich fundierte Coachingprogramme signifikant verbessert werden und es zahlt sich direkt, nachhaltig und bedeutsam in Verbesserungen der Zielkennzahlen des Unternehmens aus!

5 Literatur

- Achtziger, A. & Gollwitzer, R.M. (2008). Motivation and volition in the course of action. In: J. Heckhausen & H. Heckhausen (Eds.), *Motivation and action* (pp. 272-295). New York, NY, US: Cambridge University Press.
- Anders, S. (2002). Qualitätsmanagement und Qualitätssicherung im Coaching: konzeptionelle und inhaltliche Überlegungen zur Qualität der Einzelberatung von Führungskräften. *Organisationsberat Superv Coach*, 9(2), 101–118.
- Bass, B. M., & Riggio, R. E. (2006). *Transformational leadership*. New York: Psychology Press.
- Bachmann, T., Jansen, A. & Mäthner, E. (2004). *Check-the-Coach. Standardisierter Fragebogen zur Evaluation von Einzel-Coaching-Prozessen*. Berlin: artop-Institut an der Humboldt Universität zu Berlin.
- Behrendt, P. (2012). Freiburger Erfolgsfaktoren-Coaching – Vier Erfolgsfaktoren zur Etablierung von Konsistenz bei Coachees. *Organisationsberat Superv Coach*, 19(4), 391-404. doi: 10.1007/s11613-012-0296-7
- Behrendt, P. (2006). Wirkung und Wirkfaktoren von psychodramatischem Coaching - Eine experimentelle Evaluationsstudie. *Zeitschrift für Psychodrama und Soziometrie*, 5(1), 59-87. doi: 10.1007/s11620-006-0006-6
- Behrendt, P. (2004). *Wirkfaktoren und Psychodrama im Transfercoaching*. Unpublished diploma thesis, Albert-Ludwigs-Universität of Freiburg.
- Bergin, A. E., & Garfield, S. L. (1994). *Handbook of Psychotherapy and behavior change*. New York: Wiley.
- Bortz, J., & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler*. Berlin: Springer.

- Burke, C. S., Stagl, K. C., Klein, C., Goodwin, G. F., Salas, E., Halpin, S. M. (2006). What type of leadership behaviors are functional in teams? A meta-analysis. *The Leadership Quarterly*, 17, 288-307.
- Byrne, B. M. (2010). *Structural equation modeling with Amos: Basic concepts, applications, and programming* (2nd ed.). New York, NY: Taylor and Francis Group.
- Caspar, F., Rothenfluh, T., & Segal, Z. (1992). The appeal of connectionism for clinical psychology. *Clinical Psychology Review*, 12(7), 719-762.
- Drath, W. H., McCauley, C. D., Palus, C. J., Van Velsor, E., O'Connor, P. M. G., & McGuire, J. B. (2008). Direction, alignment, commitment: Toward a more integrative ontology of leadership. *The Leadership Quarterly*, 19, 635-653.
- Feldman, D., & Lankau, M. (2005). Executive coaching: a review an agenda for future research. *Journal of Management*, 31(6), 829-848.
- Flückiger, C., Frischknecht, E., Wüsten, G., & Lutz, W. (2008). Ressourcenpriming – Veränderung der Aufmerksamkeitsfokussierung bei Novizen und erfahrenen Therapeuten zu Therapiebeginn. *Zeitschrift für Psychiatrie, Psychologie und Psychotherapie*, 56(1), 61-68. doi: 10.1024/1661-4747.56.1.61
- Flückiger, C. & Grosse Holforth, M. (2008). Focusing the therapist's attention on the patient's strengths – A preliminary study to foster a mechanism of change in outpatient psychotherapy. *Journal of Clinical Psychology*, 56, 61-67.
- Grawe, K. (1998). *Psychologische Therapie*. Göttingen: Hogrefe.
- Grawe, K., Regli, D., & Schmalbach, S. (1994). *Cubus-Analyse*. Unpublished ratinginstrument, Universität Bern.
- Grawe, K., Donati, R., & Bernauer, F. (1994). *Psychotherapie im Wandel: Von der Konfession zur Profession*. Göttingen: Hogrefe.
- Greif, S. (2013). *Coaching – Practice or science based?* Presentation at the EAWOP conference.
- Greif, S. (2010). A new frontier of research and practice: Observation of coaching behavior. *The Coaching Psychologist*, 6 (2), 97-105.
- Greif, S., Schmidt, F. & Thamm, A. (2012). Warum und wodurch Coaching wirkt: Ein Überblick zum Stand der Theorieentwicklung und Forschung über Wirkfaktoren. *Organisationsberat Superv Coach*, 19, 375-390.
- Hartwanger, E. (1996). *Cubus-Analyse*. Unpublished diploma thesis, University of Bern.
- Heckhausen, H. & Gollwitzer, P. M. (1987). Thought contents and cognitive functioning in motivational versus volitional states of mind. *Motivation and Emotion*, 11(2), 101-120.
- Hu & Bentler (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55.
- Jansen, A., Mäthner, E. & Bachmann, T. (2003). Evaluation von Coaching. Eine Befragung von Coaches und Klienten zur Wirksamkeit von Coaching. *Organisationsberatung, Supervision, Coaching*, 10 (3), 245-254.
- Jöreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 294-316). Newbury Park, CA: Sage.
- Kiresuk, T., Smith, A., & Cardillo, J. (1994). *Goal Attainment Scaling: Applications, theory, and measurement*. New Jersey: Lawrence Erlbaum Associates.
- Kirkpatrick, D. L. (1976). Evaluation of training. In R. I. Craig (Hrsg.), *Training and evaluation handbook* (S. 87-112). New York: McGraw-Hill.
- Künzli, H. (2005). Wirksamkeitsforschung im Führungskräfte-Coaching. *Organisationsberat Superv Coach*, 12(3), 231-243.
- Künzli, H. (2009). Wirksamkeitsforschung im Führungskräfte-Coaching. *Organisationsberat Superv Coach*, 16(1), 4-18.
- Marrone, J. A. (2010). Team Boundary Spanning: A Multilevel Review of Past Research and Proposals for the Future. *Journal Of Management*, 36(4), 911-940.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models*. Thousand Oaks: Sage.
- Rauen, C. (2008). *Coaching*. Göttingen: Hogrefe.
- Regli, D., Bieber, K., Mathier, F., & Grawe, K. (2000). Beziehungsgestaltung und Aktivierung von Ressourcen in der Anfangsphase von Therapien. *Verhaltenstherapie und Verhaltensmedizin*, 21(4), 399-420.
- Regli, D., Flückiger, C., & Grawe, K. (2004). Die Realisierung von Wirkfaktoren im Therapieprozess von erfolgreichen Gruppen- und Einzeltherapien. *Verhaltenstherapie & psychosoziale Praxis*, 36(1), 41-50.
- Runde, B. (2003). *S-C-Eval- ein Instrument zur summativen Evaluation von Coaching-Prozessen*. Bissendorf: Methodos.

- Smith, E. & Grawe, K. (2003). What makes psychotherapy sessions productive? A new approach to bridging the gap between process research and practice. *Clinical psychology and psychotherapy*, 10, 275-285.
- Smith, E. & Grawe, K. (2005). Which therapeutic mechanisms work when? A step towards the formulation of empirically validated guidelines for therapists' session-to-session decisions. *Clinical Psychology & Psychotherapy*, 12(2), 112-123.
- Smith, E., Regli, D., & Grawe, K. (1999). Wenn Therapie weh tut: Wie können Therapeuten zu fruchtbaren Problemaktualisierungen beitragen? *Verhaltenstherapie & Psychosoziale Praxis*, 31(2), 227–251.
- Spreitzer, G. M. (1995). Psychological empowerment in the workplace: Dimensions, measurement, and validation. *Academy of Management Journal*, 38(5), 1442-1465.
- Stahl, G., & Marlinghaus, R. (2000). Coaching von Führungskräften: Anlässe, Methoden, Erfolg. *Zeitschrift für Führung und Organisation*, 69(4), 199–207.
- Theeboom, T., Beersma, B., & van Vianen, A.E.M. (2013). Does coaching work? A meta-analysis on the effects of coaching on individual level outcomes in an organizational context. *The Journal of Positive Psychology*. DOI: 10.1080/17439760.2013.837499
- van Knippenberg, D., & Sitkin, S. B. (2013). A Critical Assessment of Charismatic—Transformational Leadership Research: Back to the Drawing Board?. *Academy Of Management Annals*, 7(1), 1-60. doi:10.1080/19416520.2013.759433
- Yukl, G. (2012). Effective Leadership Behavior: What We Know and What Questions Need More Attention. *Academy Of Management Perspectives*, 26(4), 66-85. doi:10.5465/amp.2012.0088
- Yukl, G., Gordon, A. & Taber, T. (2002). A Hierarchical Taxonomy of Leadership Behaviour: Integrating a Half Century of Behavior Research. *Journal of Leadership and Organizational Studies*, 9(1), 15-32.
- Van der Weide, J. G., & Wilderom, C. P.M. (2004). Deromancing leadership: what are the behaviours of highly effective middle managers? *International Journal of Management Practic*, 1(1), 3-20.
- Wirtz, M., & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität*. Göttingen: Hogrefe.