Wirkungsforschung in der psychosozialen Praxis - Herausforderungen, Fallstricke und Lösungsmöglichkeiten. Methodische Reflexionen der Evaluation einer komplexen Intervention in den Settings Kindertageseinrichtung

Maike Rönnau-Böse, Claudia Grasy-Tinius, Rieke Hoffer & Klaus Fröhlich-Gildhoff
[Forum Gemeindepsychologie, Jg. 25 (2020), Ausgabe 1]

Zusammenfassung

In diesem Beitrag wird exemplarisch aufgezeigt, wie die Evaluation einer komplexen Intervention multimodal und multiperspektivisch designt werden kann. Es werden Konzept und Evaluation einer Weiterbildung für Pädagogische Fachkräfte in Kindertageseinrichtungen zum Thema "Herausforderndes Verhalten" vorgestellt; dabei handelte es sich um die Initiation von Organisations- und Teamentwicklungsprozessen nach dem Setting-Ansatz. Obwohl aus verschiedenen Gründen kein Vergleichsgruppendesign realisiert werden konnte, konnten durch das komplexe empirische Vorgehen plausible Wirkungszusammenhänge beschrieben werden.

Schlüsselworte: komplexe Interventionen, Evaluation, Wirkungsforschung, Setting-Ansatz, Organisationsentwicklung

Summary

Impact research in psychosocial practice - challenges, pitfalls and possible solutions. Methodological reflections on the evaluation of a complex intervention in the settings of daycare
This article shows the multimodal and multi-perspective evaluation design of a complex intervention. The concept and evaluation of a training for professionals in early childhood institutions on the subject of 'challenging behaviour' are presented; this involves the initiation of organisational and team development processes according to the setting approach. Although a comparison group design could not be realized for various reasons, the complex empirical procedure was able to describe plausible interdependencies.

Keywords: complex interventions, evaluation, outcome research, setting approach, organisational development

1 Einführung

Interventionen in Medizin, Psychotherapie, Gesundheitswissenschaften und - so die zunehmende Forderung - in der Pädagogik soll(t)en evidenzbasiert, also durch empirische Wirksamkeitsnachweise begründet sein (z.B. Fydrich, 2018). Allerdings stellt die wirkungsorientierte Evaluation komplexer Interventionen in der Praxis besondere Herausforderungen; die Erfassung von Wirkungen nach den Prinzipien des "Goldstandards" in Form von randomisierten Kontrollgruppen (RCT) Designs kann oft nicht umgesetzt werden und ist zumeist auch nicht gegenstandsangemessen (vgl. Kolip, 2006; Ziegler, 2010; Fröhlich-Gildhoff & Hoffer in dieser Ausgabe des "Forum Gemeindepsychologie"). So ist eine enggefasst-einheitliche Umsetzung der Intervention oftmals nicht "passgenau" genug für die jeweiligen Adressat*innen oder Zielgruppen, z.T. lassen sich aus ethischen oder praktikablen Gründen keine adäquaten Kontrollgruppen generieren (ebd.) und auch die Bildung homogener Gruppen, um Vergleiche zu ermöglichen, ist unter Praxisbedingungen in der Regel nicht gegeben (Kazdin, 2011; Wampold, 2001).

In diesem Beitrag werden anhand des Beispiels einer solchen komplexen Intervention Möglichkeiten eines umfangreichen Evaluationsdesigns aufgezeigt, das auch Rückschlüsse auf Zusammenhänge zwischen untersuchten Variablen und Wirkungen der Intervention zulässt. Hintergrund des Projekts "Herausforderungen: Für Dich? Für mich? Für alle? Herausforderungen durch Verhalten im pädagogischen Alltag professionell bewältigen" war die erlebte starke Belastung pädagogischer Fachkräfte in Kindertageseinrichtungen durch Verhaltensweisen von Kindern, die als besonders herausfordernd erlebt wurden, ein damit verbundenes geringes Kompetenzerleben und das Fehlen systematischer Konzepte zur Begegnung dieser Verhaltensweisen (GEW, 2007; Fröhlich-Gildhoff et al., 2013). Wenn die Kinder im pädagogischen Alltag durch die pädagogischen Fachkräfte keine angemessene Unterstützung erhalten und darüber hinaus bei weiterem Bedarf keine Diagnostik und Einleitung weiterer Maßnahmen durch externe Fachkräfte oder Institutionen erfolgt, besteht die Gefahr einer Chronifizierung des Verhaltens und der Entwicklung seelischer Störungen (z.B. Dougherty et al., 2015; Hoffer & Fröhlich-Gildhoff, 2019).

Die Fachkräfte in den Kindertageseinrichtungen geben an, dass sie Unterstützung im täglichen Umgang mit herausforderndem Verhalten benötigen; deutlich werden in diesem Zusammenhang auch spezifische Fortbildungswünsche. Die hier vorgestellte Intervention erfolgte als Setting-Intervention (Fröhlich-Gildhoff & Rönnau-Böse, 2018) unter Berücksichtigung aktueller Erkenntnisse der Präventions- (z.B. Christiansen, Ebert & Röhrle, 2018) und Implementationsforschung (z.B. Petermann, 2014; Hasselhorn et al., 2014).

2 Vorstellung des Projekts "Herausforderungen –", der Projektziele und der Fragestellungen

Das von der "Offensive Bildung" in Trägerschaft der BASF SE geförderte und von der Diakonie Pfalz koordinierte Forschungsprojekt wurde von Frühjahr 2016 bis Herbst 2018 durchgeführt. Dabei umfasste das Projekt *drei Phasen*, die von einer kontinuierlichen Evaluation begleitet wurden:

1. Von den Wissenschaftler*innen des Zentrums für Kinder- und Jugendforschung an der Evangelischen Hochschule Freiburg (ZfKJ) wurde ein Curriculum für die Qualifizierung von pädagogischen Fachkräften zur Begegnung mit herausforderndem Verhalten entwickelt (Fröhlich-Gildhoff, Rönnau-Böse & Tinius, 2017a).
2. Es wurden elf "Prozessbegleiter*innen" vom Diakonischen Werk (DW) Speyer rekrutiert und durch erfahrene wissenschaftliche Mitarbeiter*innen des ZfKJ (Coaches) fortgebildet, fachlich begleitet und beraten. Diese Prozessbegleiter*innen waren für die Team-/Organisationsentwicklungsprozesse in den zehn beteiligten Kitas zuständig, s.u.
3. Es wurden zehn Kindertageseinrichtungen aus der "Metropolregion Rhein/Neckar in ihrem jeweiligen Team-/Organisationsentwicklungsprozess zum professionelle(re)n Umgang mit herausforderndem Verhalten begleitet. Die Fachkräfte der beteiligten Kindertageseinrichtungen wurden dabei gemäß dem Curriculum über einen Zeitraum von etwa 18 Monaten von den geschulten Prozessbegleiter*innen fortgebildet (sechs Teamtage) und durch regelmäßige (i.d.R. zehn) kontinuierliche Prozessbegleitungs-/Coachingsitzungen fachlich begleitet. Die Fortbildungen durch die Prozessbegleiter*innen umfassten ein insgesamt sechstägiges Modulpaket auf der Grundlage eines dezidierten Curriculums (Fröhlich-Gildhoff, Rönnau-Böse & Tinius, 2017), an denen das gesamte Team einer Einrichtung teilnahm; am Standardangebot der Prozessbegleitungssitzungen nahmen z.T. das gesamte Team, z.T. nur einzelne Mitglieder teil.

Die formulierten *Ziele* des Projektes lassen sich auf mehreren Ebenen im Sinne einer Kette möglicher/zu erwartender Zusammenhänge beschreiben (Abbildung 1; Fröhlich-Gildhoff, Rönnau-Böse & Tinius, 2019b):

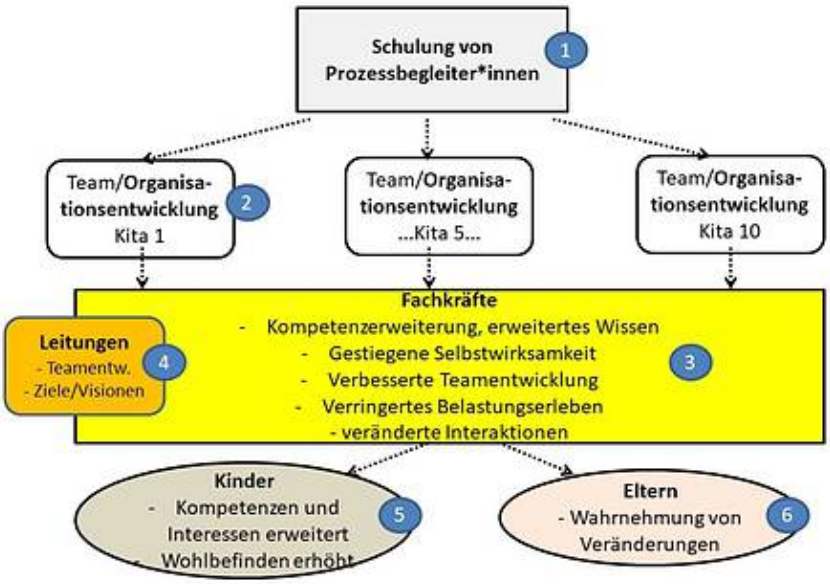

Abbildung 1: Projektablauf und Ziele

Aus den Zielen ergaben sich die verschiedenen *Fragestellungen:*

1. Wie verlaufen die Schulungen der Prozessbegleiter*innen? Wie werden diese von den Teilnehmer*innen bewertet? Wie fühlen sich die Prozessbegleitungen durch das ZfKJ auf ihre Rolle vorbereitet und begleitet?
2. Wie verläuft die Team-/Organisationsentwicklung? Wie bewerten die Teilnehmer*innen die jeweiligen Bausteine der Weiterbildungen?
3. Wie schätzen die Fachkräfte ihre Selbstwirksamkeit, die Zusammenarbeit im Team, ihre berufliche Belastung, sowie ihre eigenen Kompetenzen ein und was verändert sich im Vorher-Nachher-Vergleich? Wie verändern sich die fremdeingeschätzten Kompetenzen der Fachkräfte durch die Fortbildungen hinsichtlich der einzelnen Elemente des systematischen Vorgehens im Umgang mit herausforderndem Verhalten? Wie gestaltet sich (beispielhaft) die Fachkraft-Kind-Interaktion und wie verändert sich diese nach Abschluss der Qualifizierungsmaßnahmen?
4. Wie wird die Rolle der Leitungen von den Teams eingeschätzt? Verändern sich Maße der Teamkultur?
5. Wie schätzen die Fachkräfte die sozialen, emotionalen und motivationalen Kompetenzen sowie das Wohlbefinden und die sozialen Beziehungen der Kinder in ihrer Einrichtung ein und lassen sich Veränderungen im Vorher-Nachher-Vergleich erkennen?
6. Wie nehmen die Eltern das Klima in der Einrichtung, die Zusammenarbeit mit den Fachkräften und ihre Kinder in der Einrichtung wahr? Wie schätzen sie die Gestaltung der Fachkraft-Kind-Beziehung und -Interaktion ein? Welche Wünsche haben die Eltern an die Zusammenarbeit mit den Fachkräften? Nehmen sie nach Abschluss der Weiterbildungen Veränderungen wahr?
7. Wie schätzen die Fachkräfte und Leitungen den aktuellen Stand der Umsetzung in der jeweiligen Einrichtung ein? Wie bewerten die Leitungen und die Projektbeauftragten im Kita-Team den Prozess der Umsetzung und die Rolle der Prozessbegleitungen? Wie wird die Rolle der Leitung im Umsetzungsprozess aus Sicht der Prozessbegleitung, Projektbeauftragten und Coachs wahrgenommen? Gibt es Zusammenhänge zwischen Umsetzungsintensität und Outcome-Maßen?

3 Untersuchungsdesign

3.1 Vorüberlegungen

Es wird deutlich, dass zum einen die Intervention insgesamt sehr komplex war und verschiedene Wirkebenen intendiert waren. Dementsprechend wurde ein sehr komplexes Untersuchungsdesign geplant, das folgenden *Grundsätzen* folgte (s.a. Fröhlich-Gildhoff & Hoffer in dieser Ausgabe des "Forum Gemeindepsychologie"):

- Es sollte eine Kombination von Prozess- und Ergebnisevaluation umgesetzt werden.
- Um Veränderungen zu erfassen, sollte die Datenerhebung zu mehreren Zeitpunkten, zumindest vor und nach der Intervention (prä/post), erfolgen.
- Als Erhebungsverfahren sollten sowohl quantitative als auch - im Sinne einer Vertiefung - qualitative wissenschaftliche Erhebungsmethoden eingesetzt werden.
- Die verschiedenen beteiligten Gruppen (Prozessbegleitungen, Fachkräfte und Leitungen, Kinder, Eltern) sollten in die Evaluation einbezogen werden. Die pädagogischen Fachkräfte und die Einrichtungsleitungen aller Kindertageseinrichtungen (N=10) wurden dementsprechend in die Evaluation zu den Zeitpunkten t0 und t1 (Prä-Post-Erhebung) einbezogen.
- In der Hälfte der Kitas sollten über die Basisevaluation hinausgehend vertiefende Daten auf der Ebene der Kinder und Eltern erhoben werden.
- Die Umsetzungsintensität und -qualität sollte systematisch und mehrperspektivisch erfasst werden.

Bei der Detailplanung und Realisierung der Untersuchung stellten sich einige besondere *Herausforderungen:*

Es war zunächst geplant, ein (Wartelisten-)Kontrollgruppendesign umzusetzen, um Vergleiche zwischen Kitas/Fachkräften, die die Intervention erhielten und solchen ohne Intervention, ziehen zu können und so den Kriterien des klassischen, sog. "Goldstandards" der Wirksamkeitsforschung im Sinne eines quasi-experimentellen Designs genügen zu können (s. Fröhlich-Gildhoff & Hoffer, 2017, und in diesem Band). Allerdings fand sich kein Kita Team, das bereit war, an der Evaluation teilzunehmen, ohne zugleich an den Fortbildungen teilzunehmen. So konnte dieser Ansatz nicht verwirklicht werden und die Wirkungszusammenhänge mussten aus den vorhandenen Daten hergestellt werden.

Eine Grundschwierigkeit bestand und besteht darin, eine - in diesem Fall curricular gestützte - Rahmenkonzeption auf die Situation, Bedürfnisse und Entwicklungsstände der je einzelnen Kindertageseinrichtungen und Teams zu adaptieren. Dieses Prinzip der Passung muss im Sinne einer gelingenden Implementierung vor der manualgetreuen Umsetzung der Intervention Vorrang haben (Petermann, 2014; Haselhorn et al., 2014; Durlak & DuPree, 2008) - zugleich ergibt sich das Problem, dass bei zehn teilnehmenden Einrichtungen zehn zumindest partiell unterschiedliche Varianten der Intervention realisiert werden (müssen). Diese Varianz muss durch eine genaue Prozessdokumentation präzise und möglichst umfassend erfasst werden.

3.2 Erhebungsinstrumente[1]

Entsprechend der o.g. Prinzipien wurden für die verschiedenen Untersuchungsebenen und -zielgruppen verschiedene quantitative und qualitative Erhebungsinstrumente ausgewählt und z.T. entwickelt. Diese werden im Folgenden in ihrem Bezug zu den Fragestellungen kurz vorgestellt; ergänzend sind die Stichprobengrößen aufgeführt. Die Auswahl der Instrumente und die z.T. erfolgte Begrenzung auf Teilstichproben der gesamten Teilnehmer*innen am Projekt musste auch unter Berücksichtigung forschungsökonomischer Grenzen vorgenommen werden; so war z.B. eine Vollerhebung unter allen Eltern der beteiligten Kitas nicht umzusetzen.

Die Erhebungen erfolgten zu zwei Messzeitpunkten (MZP), jeweils vor (t0) und nach (t1) der Intervention (Abstand ca.18 Monate). Die Fragebögen zu den Fortbildungssitzungen wurden bei jeder Sitzung bearbeitet.

(1) Schulung der Prozessbegleiter*innen	[A] Die Prozessbegleitungen bearbeiteten einen Fragebogen (FB) zur Qualität der eigenen Vorbereitung/Schulung.	Sechs von zehn FB vollständig über mehrere Messzeitpunkte (MZP) ausgefüllt
(2) Qualifizierungsmaßnahmen in den Teams (Prozessevaluation und Teamentwicklung)	[B] Kontinuierliche schriftliche Befragung der Prozessbegleitungen, Fachkräfte und Leitungen nach jeder Fortbildungseinheit zur Zufriedenheit mit den Fortbildungen.	
	Evaluationsbogen für die Teilnehmer*innen Die Evaluationsbögen wurden von den Teilnehmer*innen jeweils nach jedem durchgeführten Baustein ausgefüllt. Die Bestandteile des Bogens waren einerseits Fragen zur Didaktik auf Dozent*innen-Ebene sowie andererseits die Einschätzung von Motivation und Akzeptanz der Teilnehmer*innen.	503 vollst. ausgefüllte FB
	Protokolle der Prozessbegleitungen Die Prozessbegleitungen bearbeiteten nach jedem abgeschlossenen Baustein ebenfalls einen Bogen. Dieser umfasste die gleichen Teile wie der Bogen der Fachkräfte, fokussierte jedoch die Perspektive der Prozessbegleitungen auf die Fachkräfte sowie die Eignung und Praktikabilität des Curriculums.	24 vollst. ausgefüllte FB
	[C] Fragebögen zur Teamentwicklung	
	Zusammenarbeit im Team (t_0/t_1) Zur Erfassung der Zusammenarbeit im Team wurden zwölf Items ausgewählt, die aus dem FAT (Fragebogen zur Arbeit im Team; Kauffeld, 2004) stammen.	t0 143 FB t1 120 FB
(3) Entwicklung der pädagogischen Fachkräfte	Basisevaluation [D] Schriftliche Befragung der Fachkräfte und Leitungen durch Fragebögen (t_0/t_1): Erfassung des Kompetenzerlebens, der Selbstwirksamkeit sowie der beruflichen Belastung. = SELBSTeinschätzung	t0 143 FB t1 120 FB
	[D1] Kompetenzerleben Die Selbsteinschätzung der Kompetenzen erfolgt durch einen standardisierten Fragebogen mit 20 Items und drei Subskalen (Cronbachs von .730 bis .912).	

	[D2] Selbstwirksamkeit Fragebogen von Schwarzer & Jerusalem (1999).	
	[D3] Berufliche Belastung Erfasst mittels Hamburger Burnout Inventar in der Version HBI 40 (Burnout-Institut Norddeutschland, 2017).	
	[E] Schriftliche Befragung der Fachkräfte und Leitungen durch Fallvignetten (t_0/t_1) im Sinne einer FREMDeinschätzung: Kompetenzerfassung durch Dilemmasituationen anhand einer Fallvignette (Methode nach Fröhlich-Gildhoff et al., 2014). Die pädagogischen Fachkräfte bearbeiteten hierfür eine Fallvignette bzw. Dilemma-Situation, die sie durch fünf offen gestellte Fragen beantworteten. Diese erfassen die Bereiche Situationswahrnehmung, Grad der Analyse, Handlungsplanung und -perspektiven sowie Evaluation.	t0: 133 t1: 110
	[F] Beobachtung und Reflexion der Fachkraft-Kind-Interaktion mittels Videografie (t_0/t_1): Durch Videoaufnahmen und Video-Fall-Analysen sollte eine systematische Betrachtung der Interaktion zwischen Fachkraft und Kind ermöglicht werden (Weltzien et al., 2017). Analyse mittels Evaluationsinstrument ‚GInA-E' (Gestaltung von Interaktionsgelegenheiten im Alltag-Evaluation; Weltzien et al., 2017).	t0: 57 Videosequenzen von n= 15 Fachkräften t1: 30 Videosequenzen von n = 10 Fachkräften
(4) Leitungsqualität	[G] Fragebogen zur Leitungsqualität (t_0/t_1) Multifactor Leadership Questionnaire (MLQ, Felfe und Goihl, 2003) (Testkennwerte von Cronbachs .744 bis .880). 18 Items wurden in die hier vorliegende Untersuchung einbezogen; der Bogen wurde ebenfalls aus Sicht der Leitungen bearbeitet.	
(5) Entwicklung des psychischen Wohlbefindens der Kinder (nur im Rahmen der vertieften Evaluation, t_0/t_1 in fünf Kitas)	[H] Einschätzung der Kinder mittels schriftlicher Befragung der Fachkräfte (t_0/t_1): Hierbei wurden soziale, emotionale und motivationale Kompetenzen sowie das Wohlbefinden und die sozialen Beziehungen der Kinder erfasst. Als Instrument wurden Einzelskalen des KOMPIK (Kompetenzen und Interessen von Kindern) eingesetzt. (Testkennwerte von Cronbachs =.893 bis .955; Staatsinstitut für Frühpädagogik, 2014).	t0 353 Kinder t1 307 Kinder
(6) Erfahrungen der Eltern (nur i.R. der vertieften Evaluation in fünf Kitas)	[I] Gruppendiskussionen mit zusätzlichem Kurzfragebogen für Eltern (t_0/t_1): Erfassung der Perspektive der Eltern zum Klima in den	t0: 5 Gruppendiskussionen,

	Einrichtungen und der Zusammenarbeit mit den Fachkräften sowie der Wahrnehmung des Kindes und der Fachkraft-Kind-Beziehung und der Wünsche der Eltern an die Zusammenarbeit mit den Fachkräften und an das Projekt. Zum Zeitpunkt t_1 sollte zudem das Projekt reflektiert werden. Auswertung mittels Inhaltsanalyse (Mayring, 2015).	30 Kurzfragebögen t1: 4 Gruppendiskussionen, 15 Kurzfragebögen
(7) Betrachtung der Implementationsqualität Um die Implementationsqualität zu erfassen wurden einerseits schriftliche Befragungen der Teilnehmer*innen bezüglich der Fortbildungen und deren Umsetzung sowie andererseits Befragungen der Leitungen, Projektbeauftragten und Prozessbegleitungen eingesetzt. Zu t_1 wurden die Umsetzungsintensität des Gesamt-Projektprozesses sowie die Rolle der Leitung und Prozessbegleitungen nochmals erfasst.	[J] Implementierungsprozess der Fortbildungsinhalte Zum Zeitpunkt der Post-Erhebung wurden die Teilnehmer*innen bezüglich der Fortbildungen und der Umsetzung der Inhalte befragt. Hierfür wurde ein Fragebogen konstruiert, der auf den „Kriterien für gelingende Implementation“ nach Fröhlich-Gildhoff & Hoffer (2017) basiert (Testkennwerte von Cronbachs =.747 bis .832; 20 Items).	t1: 120 Fragebögen
	[K] Beurteilung des Umsetzungsindexes (UI) Zur Erfassung des Umsetzungsindexes wurden verschiedene Indikatoren, erfasst, die Hinweise auf eine „gelingende“ Implementierung geben können und erlauben, „die empirisch erfassten Ergebnisse einer Intervention – z.B. die Unterschiede in standardisierten Tests bei der Vorher-Nachher-Messung – in ein Verhältnis zu relevanten Faktoren im Umsetzungs-/Implementationsprozess zu setzen. Der UI ist damit von der Grundintention eine Möglichkeit, relevante Prinzipien der Implementationsforschung zu operationalisieren.“ (Fröhlich-Gildhoff & Hoffer, 2017, S. 197). Dabei wurden jeweils die Perspektiven der Leitungen, der Projektbeauftragten im Kita-Team sowie der Prozessbegleitungen erfasst, indem diese den Bogen bearbeiteten.	t1: 6 Prozessbegleitungen 8 Leitungen, 7 Projektverantwortliche auf Teamebene

4 Ergebnisse und Zusammenhänge

Die Ergebnisse der Untersuchungen sind andernorts ausführlich dargestellt (Fröhlich-Gildhoff et al., 2019 a, b); die wichtigsten lassen sich wie folgt zusammenfassen:

- Die *Organisationsentwicklung* in den jeweiligen Einrichtungen erreichte hohe Zufriedenheiten in den Skalen der Prozessevaluation (MW=5,02, SD=.86; sechsstufiges Antwortformat von 1="trifft gar nicht zu", 6="trifft voll zu"). Die Fachkräfte zeigten eine hohe Zufriedenheit mit dem Aufbau des Curriculums und den dabei eingesetzten Methoden und Vorgehensweisen. Besonders wurde in den offenen Antworten der hohe/gute Praxisbezug hervorgehoben.
- Die *Fachkräfte* beschrieben selbst ein bedeutsam gestiegenes Kompetenzerleben mit einem starken Effekt ($p<.001$; $d=.83$). Zugleich zeigte sich, dass die Fachkräfte nach der Intervention ein signifikant

sinkendes Bedürfnis an Weiterbildung (p<.001; d=.67) haben. In der Fremdeinschätzung (Auswertung der Fallvignetten) ergaben sich im Verlauf von t0 zu t1 in den Kompetenzbereichen "Handlungsplanung" und "Fachliche Begründung der Handlungsoptionen" positive Veränderungen; diese war im Bereich "Fachliche Begründung" hoch signifikant (p<.001; Effektstärke von d=.66). Das bedeutet, die Fachkräfte können ihr Handeln besser begründen, sie haben klarere Handlungsoptionen und erleben sich kompetenter. In den weiteren Kompetenzbereichen (Situationswahrnehmung, Situationsanalyse) ließen sich keine signifikanten Unterschiede erkennen.

- Bezüglich des Selbstwirksamkeitserlebens zeigte sich im Vorher/Nachher-Vergleich eine bedeutsame Steigerung in Bezug auf den Umgang mit neuen und unbekannten Situationen (p<.001; mittlerer Effekt d=.438). Die berufliche Belastung (Hamburger Burnout Inventar, Burnout Institut Norddeutschland, 2017) lag prä/post im mittleren, "unauffälligen" Bereich; die Leistungszufriedenheit veränderte sich über die Zeit positiv (Effektstärke d=.26; Signifikanzniveau p<.05).
- Auf der Ebene der *Kinder* konnten zum zweiten Erhebungszeitpunkt die Werte in den Skalen des KOMPIK signifikant mit deutlichen Effekten gesteigert werden. Die positive Entwicklung in der Skala "Wohlbefinden und soziale Beziehungen" - die nach den Testautor*innen als Indikator für seelische Gesundheit gesehen werden kann (Staatsinstitut für Frühpädagogik, 2018) - ist hochsignifikant (p<.001) mit einem mittleren Effekt (d=.48). Beschreibungen der Fachkräfte verdeutlichten, dass sie durch die Fortbildungen die Kinder in ihrer Entwicklung entsprechend ihrer Bedürfnisse individuell deutlicher unterstützen und stärken.
- Bei der *Analyse der Umsetzungsqualität* bzw. des Zusammenhangs zwischen Umsetzungsintensität und Ergebnissen auf Fachkraft- und Kinder-Ebene fand sich ein Zusammenhang zwischen dem Ausmaß der praktischen Anwendung/Umsetzung von Weiterbildungselementen und einem verbesserten Selbstwirksamkeitserleben der Fachkräfte: Je höher der Grad der konkreten Umsetzung in den Weiterbildungen war bzw. je konkreter die vermittelten Inhalte, Handlungsformen und Maßnahmen umgesetzt wurden, desto höher war die Steigerung des Selbstwirksamkeitserlebens im Bereich des Umgangs mit neuen und überraschenden Situationen (r=.28*; p=.02; N=71). Bei der Steigerung der Wahrnehmung emotionaler kindlicher Kompetenzen zeigten sich ähnliche Effekte: Je höher der Grad der konkreten Umsetzung in den Weiterbildungen war, umso höhere Werte ließen sich bei den Fragen zur Wahrnehmung emotionaler kindlicher Kompetenzen erkennen (p=.03). Diese zeigten eine mittlere bis hohe Effektstärke (Eta-Quadrat von .09).

Die intendierten Veränderungen über die Zeit und bezüglich der Projektziele lassen sich wie folgt skizzieren:

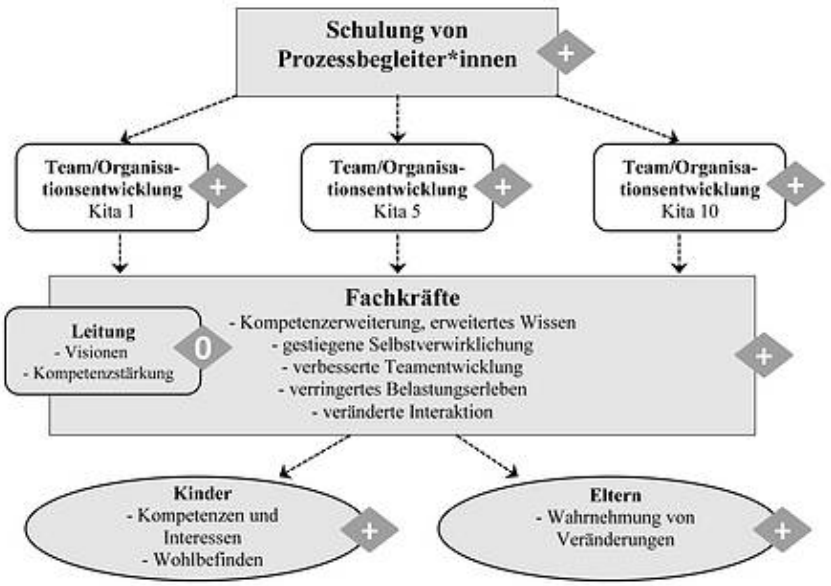

Abbildung 2: Zusammenfassung der Veränderungen über die Zeit und bzgl. der Ziele

Die Qualifizierung der Prozessbegleiter*innen war sinnvoll; sie konnten die verschiedenen Kita-Teams so in den Fortbildungen schulen und in den Umsetzungsprozessen begleiten, dass die Fachkräfte sich als kompetenter erlebten - dies zeigte sich dann analog in den Fremdeinschätzungen der Kompetenzen und im Selbstwirksamkeitserleben. Parallel zeigen sich signifikant positive Veränderungen des Wohlbefindens und der sozialen Beziehungen der Kinder. Die (wenigen) befragten Eltern können positive Entwicklungen bei den Fachkräften beschreiben. Hinsichtlich der Teamentwicklungsprozesse insgesamt und dem Handeln der Leitungen konnten keine messbaren Veränderungen erfasst werden; ein Grund dafür war ein schon hohes Ausgangsniveau zu Projektbeginn/ vor der Intervention.

5 Reflexion des methodischen Vorgehens

Ziel der Untersuchung war es, mögliche Veränderungen auf den verschiedenen Ebenen der Intervention abzubilden und mögliche Zusammenhänge zwischen den durchgeführten Qualifizierungsmaßnahmen und erfassten Veränderungen herzustellen.

Dazu wurde ein komplexes Untersuchungsdesign mit folgenden Vorgaben entwickelt, das die aus der Literatur zur (Wirkungs-)Evaluation (z.B. Bortz & Döring, 2006; Stockmann, 2006; Fröhlich-Gildhoff, 2012) zusammengefassten zentralen Vorgaben erfüllte:

- eine Intervention oder Handlungspraxis wurde über *mehrere Messzeitpunkte* beobachtet
- es wurde eine *Vielfalt von Methoden* eingesetzt, um verschiedene Zugänge zum Untersuchungsgegenstand zu ermöglichen; es wurde ein mixed methods Design eingesetzt (Cresswell, 2003)
- es wurden *verschiedene Perspektiven* (Selbstbeobachtung, Fremdbeobachtung) berücksichtigt
- es wurde eine *Verbindung von Prozessdaten und Ergebnissen* durch den Umsetzungsindex hergestellt

Dieses Design konnte grundsätzlich umgesetzt werden; durch die hohe Objektivität und Zuverlässigkeit, schnelle und einfache Auswertung wie auch die kurzen Bearbeitungszeiten war der Einsatz der Fragebögen auf verschiedenen Ebenen sehr geeignet. Auch für den Prä-Post-Vergleich waren alle Fragebögen verwendbar.

Da durch die Fragebögen nur ein Teilaspekt von Kompetenz erfasst werden kann und durch die Selbsteinschätzung von Kompetenz nicht vorhergesagt werden kann, ob und wie sich diese im Handlungsprozess umsetzen, waren bezüglich der Kompetenzerfassung auf Ebene der Fachkräfte mehrere spezifische Methoden und Verfahren notwendig. Um die Kompetenzentwicklung der Fachkräfte mehrperspektivisch zu erfassen, kamen entsprechend quantitative sowie qualitative Verfahren zum Einsatz. Die schriftliche Befragung bezüglich der Kompetenz konnte dabei die subjektive Wahrnehmung der Fachkräfte erfassen, die Fallvignetten und Videografie konnten hingegen eine Fremdeinschätzung generieren. Es wurde also versucht, die verschiedenen Facetten von Kompetenz mehrperspektivisch messbar zu machen.

Bei der Realisierung dieses Designs ergaben sich allerdings auch einige Probleme:

- Es lag eine relativ hohe Drop-Out Rate vor; zum Teil konnte nur die Hälfte der zu t0 befragten Fachkräfte auch zur Abschlusserhebung erreicht werden. Diese lässt sich neben der hohen Personalfluktuation - die allgemein in Kindertageseinrichtungen vorliegt (z.B. Weltzien et al., 2016) - auf Krankheitsausfälle, falsch verwendete IDs etc. zurückführen. Auch wenn die Drop-out-Analyse keine Hinweise auf systematischen Datenverlust ergab, so ist der - in der Praxis nicht unübliche - "Ausfall" an Daten bedauerlich.

- Bei der Fremdeinschätzung der Kompetenzen mittels Fallvignetten ergab sich die Schwierigkeit, dass die Bearbeitung für die teilnehmenden Fachkräfte viel Zeit erforderte und sich dadurch möglicherweise die Konzentration verringerte und "Motivationseinbrüche" vorkamen. Dies erklärt die z.T. nicht vollständigen oder weniger substantiellen Antworten. Die zeitliche Begrenzung der Bearbeitung des Falles auf 35 Minuten wurde zudem mehrfach als zu knapp angemerkt. Diese Faktoren müssen bei der Ergebnisinterpretation berücksichtigt werden.
- Bezüglich der Kompetenzerfassung der Fachkraft-Kind-Interaktion mittels Videografie ergab sich zudem das Problem, dass die Aussagen nur begrenzt aussagekräftig sind, da hier ein Drittel (t0: N=15; t1: N=10) der Fachkräfte am zweiten Untersuchungszeitpunkt nicht teilnahm (Personalfluktuation, Krankheit). In weiteren Projekten sollten hier größere Stichproben generiert werden.

Zusammenfassend konnten die verfassten Fragestellungen mit dem verwendeten Design größtenteils beantwortet werden. Die Ergebnis- und Prozessdaten konnten systematisch in Beziehung gesetzt werden. Die eingesetzten Verfahren konnten sich gegenseitig ergänzen und so auch methodische Schwächen abmildern.

6 Methodologische Reflexion im Bezug zur Wirkungsforschung bzw. Evidenzbasierung

Auch wenn kein klassisch-experimentelles Design zur Erfassung von Wirkungen umgesetzt wurde, muss zunächst festgehalten werden, dass die vorliegende Untersuchung auf einer methodisch sorgfältig kontrollierten Kombination von summativer Evaluation und Prozessevaluation mit einer Kombination aus quantitativen und qualitativen Methoden basiert (z.B. Cresswell, 2003; s.a. Fröhlich-Gildhoff & Nentwig-Gesemann, 2008). Zudem wurden alle Ebenen der komplexen Intervention empirisch erfasst und auch mögliche sekundäre Veränderungen (auf Ebene der Eltern und der Kinder) untersucht.

Die Gesamtbetrachtung der vielfältigen Daten lässt folgenden Schluss zu: Da die beobachteten Veränderungen bei allen eingesetzten Instrumenten und untersuchten Gruppen mit mittleren oder starken Effekten konsistent die Richtung eines positiven Zuwachses (an Kompetenzen der Fachkräfte, im Wohlbefinden der Kinder–) aufweisen, kann von einer hohen Wirkwahrscheinlichkeit der Intervention (Weiterbildungen) ausgegangen werden. Dies wird unterstützt durch die hoch positive Beurteilung der Intervention durch die Zielgruppe der pädagogischen Fachkräfte.

Auch wenn grundsätzlich nicht auszuschließen ist, dass nicht (allein) die Intervention zu den beobachtbaren Veränderungen führte - denkbar wären Einflüsse wie eine besonders hohe Motivation der Teilnehmenden, die alleinige Wirkung der "Zuwendung" durch die externe Begleitung oder weitere, nicht erfasste Einflüsse - finden sich aus der Prozessevaluation und über den Umsetzungsindex, der eine Aussage über den Zusammenhang zwischen Umsetzungsintensität und prä/post-Veränderungen macht (Fröhlich-Gildhoff & Hoffer, 2017), deutliche Hinweise über Zusammenhänge zwischen Intervention und den referierten Veränderungen, sodass von einer Wirksamkeit der Intervention ausgegangen werden kann (Fröhlich-Gildhoff et al., 2018, 2019b). Möglicherweise existierten auch hier nicht erfasste Zusammenhänge zwischen der Motivation der Teilnehmer*innen und der Umsetzungsintensität - dies sollte in Folgestudien berücksichtigt werden.

Wichtige Erkenntnisse der Implementationsforschung (Durlak & DuPree, 2008; Hasselhorn et al., Fröhlich-Gildhoff & Hoffer in dieser Zeitschrift) lassen sich auch in der hier referierten Studie wieder erkennen:

Angemessenheit und Passung: Die generelle Intervention, also die curriculumgestützte Weiterbildung der Fachkräfteteams, musste adaptiert werden auf die jeweilige Situation der Zielgruppe - die wurde von den

Adressat*innen als bedeutsam für die Akzeptanz herausgestellt. Damit ist auch *Verhältnis von Wiedergabetreue und Passgenauigkeit* angesprochen (Petermann, 2014; Beelmann & Karing, 2014). Die im Sinne der Implementation nötige "compatibility (contextual appropriateness, fit, congruence, match)"(Durlak & DuPree, 2008, S. 337) führt, wie dargelegt, empirisch zu dem Problem, dass zehn zumindest teilweise abweichende Umsetzungswege beschritten wurden. Diese Herausforderung ist nur durch eine sehr sorgfältige Dokumentation aller Prozessschritte zu bewältigen, so dass ex post Zusammenhänge zwischen den "Umsetzungspfaden" und den Ergebnissen auf den verschiedenen Ebenen hergestellt werden können.

Die *Durchdringung*, also die "Implementationstiefe" (Hasselhorn et al, 2014, S. 142) bzw. breite "Integration einer Maßnahme in die betreffende Institution" (Petermann, 2014, S. 125) ließ sich mittels des Instruments des Umsetzungsindex erfassen. Dabei konnten Zusammenhänge zwischen der Umsetzungsintensität und outcome Maßen beobachtet werden.

Bei der Umsetzung hatten *"Merkmale der einzelnen Einrichtung"* (Hasselhorn et al., 2014, S. 144) wie Institutionsklima, Kooperation und Unterstützung im Team, die Unterstützung und enge Begleitung durch die Leitung oder Transferunterstützung durch den Träger ebenso eine Bedeutung wie *"Merkmale der pädagogischen Fachkräfte"* bzw. Anwender/Zielgruppe (Hasselhorn et al., 2014, S. 144). Diese Merkmale wurden allerdings - abgesehen von der Variable der Leitungsqualität sowie dem Selbstwirksamkeitserleben und der Belastung der Fachkräfte - nicht systematisch erhoben. Aus den Protokollen der Weiterbildungs- und Prozessbegleitungssitzungen fanden sich jedoch Hinweise auf die Relevanz dieser Faktoren. Sie sollten in zukünftigen Implementationen noch sorgfältiger erhoben werden.

Ein Problem stellte, wie beschrieben, die Fluktuation der pädagogischen Fachkräfte, mithin der unmittelbaren Zielgruppe der Intervention, dar. Dies hatte nicht nur Auswirkungen auf die Evaluation, sondern stellt Anforderungen an die Sicherung der Nachhaltigkeit: Es muss strukturell gesichert werden, dass das im Team erworbene Wissen und die entsprechenden Kompetenzen an neue Mitarbeiter*innen weitergegeben werden und die konzeptionellen Veränderungen auch weiter "gelebt" werden. Hier ist es wichtig, schon im Projektverlauf Verantwortliche neben der Leitung festzulegen, die für das "Wissensmanagement" und die Sicherung der erzielten Fortschritte verantwortlich sind.

Der im Fachdiskurs nur selten (Ausnahme: Beelmann & Karing, 2014) diskutierte Aspekt der *Rolle der Person und wahrgenommenen Kompetenz der Umsetzer*innen/Referent*innen* (hier: "Prozessbegleiter*innen) konnte hingegen aus der qualitativen Analyse der Protokolle der Weiterbildungssitzungen und entsprechenden Evaluationsbögen der Teilnehmer*innen sehr deutlich extrahiert werden. Die wahrgenommene Kompetenz und auch die Motivation bzw. "Begeisterung" der Referent*innen war ebenso wie die didaktische Gestaltung bedeutsam für die Akzeptanz des Angebots und der Thematik.

Die ausführliche Darstellung der Evaluation, der genutzten Instrumente, der Ergebnisse und der beobachteten Zusammenhänge macht deutlich, dass Wirkungen einer Intervention auch ohne das klassische RCT-Design beschrieben werden können. Dabei müssen komplexe Interventionen durch ein komplexes Design, v.a. ein multiperspektivisches und multimethodisches Vorgehen, untersucht werden.

Diese Erkenntnisse sollten die Diskussion um die Evidenzbasierung des Handelns in den Humanwissenschaften befruchten: Gezielte und systematische Interventionen sollen hinsichtlich ihrer Wirksamkeit betrachtet werden, die gleichfalls systematisch untersucht werden sollte - allerdings müssen dazu auch Methoden bzw. Designs genutzt werden, die über die eng gefassten und für die Human- bzw. Sozialwissenschaften oft nicht gegenstandsangemessenen empirischen Ansätze des medizinischen Modells hinausgehen. Die hier referierte Untersuchung ist ein Beispiel dafür, wie entsprechende Wirkungszusammenhänge durch ein methodisch umfassendes und anspruchsvolles Vorgehen hergestellt werden und die Evidenzbasierung des gewählten Vorgehens zeigen können.

Endnote

1. Die Erhebungsinstrumente sind ausführlich in Fröhlich-Gildhoff et al. (2019b) dargestellt.

Literatur

Beelmann, A. & Karing, C. (2014). Implementationsfaktoren und -prozesse in der Präventionsforschung: Ergebnisse, Probleme, Strategien, Perspektiven. Psychologische Rundschau, 65, 129 - 139.

Bortz, J. & Döring, N. (2006). Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. Heidelberg: Springer.

Burnout Institut Norddeutschland (2017). Burnout-Test: Über das Hamburger Burnout Inventar (HBI 40). Hg. v. Matthias Burisch. Verfügbar unter: http://www.burnout-institut.eu/Burnout-Test.8.0.html [18.10.2018].

Christiansen, H., Ebert, D. & Röhrle, B. (Hrsg.) (2018). Prävention und Gesundheitsförderung Bd. VI. Entwicklungen und Perspektiven. Tübingen: DGVT.

Cresswell, J. W. (2003). Research Design. Qualitative and Mixed Methods Approaches. Thousand Oaks: Sage.

Dougherty, L. R., Smith, V. C., Bufferd, S. J., Kessel, E., Carlson, G. A. & Klein, D. N. (2015). Preschool irritability predicts child psychopathology, functional impairment, and service use at age nine. Journal of child psychology and psychiatry, and allied disciplines, 56, 999-1007. DOI 10.1111/jcpp.12403

Durlak, J. A. & DuPre, E. (2008). Implementation Matters: A Review of Research on the Influence of Implementation on Program Outcomes and the Factors Affecting Implementation. American Journal of Community Psychology, 41(3-4), 327-350.

Fröhlich-Gildhoff, K., Strohmer, J., Rönnau-Böse, M., Braner, K. & Grasy-Tinius, C. (2019a). Herausforderndes Verhalten in Kindertageseinrichtungen - Evaluation eines Qualifizierungsprogramms für Fachkräfte-Teams. In D. Weltzien, H. Wadepohl, C. Schmude, H. Wedekind & A. Jegodtka (Hrsg.), Forschung in der Frühpädagogik XII (S. 141 - 172). Freiburg: FEL.

Fröhlich-Gildhoff, K., Strohmer, S., Rönnau-Böse, M., Braner, K. & Grasy-Tinius, C. (2019b). Herausforderungen: Für Dich? Für mich? Für alle? Herausforderungen durch Verhalten im pädagogischen Alltag professionell bewältigen. Wissenschaftlicher Abschlussbericht. Freiburg: FEL.

Fröhlich-Gildhoff, K. & Rönnau-Böse, M. (2018). Prävention und Gesundheitsförderung im Setting Kindertageseinrichtung - welche Strategien sind erfolgreich? In H. Christiansen, D. Ebert & B. Röhrle (Hrsg.), Prävention und Gesundheitsförderung Bd. VI. Entwicklungen und Perspektiven (S. 287 - 308). Tübingen: DGVT.

Fröhlich-Gildhoff, K., Braner, K., Strohmer, J.,Rönnau-Böse, M. & Tinius, C. (2018).

Herausforderndes Verhalten in Kindertageseinrichtungen - Evaluation eines Qualifizierungsprogramms für Fachkräfte-Teams. Perspektiven der empirischen Kinder- und Jugendforschung, 4(2), 26-52.

Fröhlich-Gildhoff, K., Rönnau-Böse, M. & Tinius, C. (2017). Herausforderndes Verhalten von Kindern professionell bewältigen. Ein Curriculum für die Weiterbildung pädagogischer Fachkräfte in Kindertageseinrichtungen. Freiburg: FEL.

Fröhlich-Gildhoff, K. & Hoffer, R. (2017). Methodische und methodologische Herausforderungen der Wirkungsforschung unter Praxisbedingungen in der Frühpädagogik - Lösungen jenseits des "Goldstandards". In I. Nentwig-Gesemann & K. Fröhlich-Gildhoff (Hrsg.), Forschung in der Frühpädagogik X. Zehn Jahre frühpädagogische Forschung - Bilanzierungen und Reflexionen (S. 209-228). Freiburg: FEL.

Fröhlich-Gildhoff, K., Rönnau-Böse, M. & Tinius, C. (2017a). Herausforderndes Verhalten in Kita und Grundschule. Erkennen, Verstehen, Begegnen. 1. Auflage. Stuttgart: Kohlhammer Verlag.

Fröhlich-Gildhoff, K., Lorenz, F. L., Tinius, C. & Sippel, M. (2013). Überblickstudie zur pädagogischen Arbeit mit Kindern mit herausforderndem Verhalten in Kindertageseinrichtungen. Frühe Bildung (2), 59-71.

Fröhlich-Gildhoff, K. (2012). Forschung in der Frühpädagogik in Deutschland - Geschichte, Stand und Standards - oder: Von den Methoden zu empiriebasiertem und profiliertem Handeln. In S. Kägi & U. Stenger (Hrsg.), Forschung in Feldern der Frühpädagogik (S. 26-58). Hohengehren: Schneider.

Fröhlich-Gildhoff, K. & Nentwig-Gesemann, I. (2008). Forschung in der Frühpädagogik. Sinn - Standards - Herausforderungen. In K. Fröhlich-Gildhoff, I. Nentwig-Gesemann. & R. Haderlein (Hrsg.), Forschung in der Frühpädagogik I (S. 13 - 36). Freiburg: FEL.

GEW (2007). Wie geht's im Job? Kita-Studie der GEW. Frankfurt am Main: GEW (Jugendhilfe und Sozialarbeit).

Fydrich, T. (2018). Evidenzbasierung in der Psychotherapie. Familiendynamik, 43 (4), 268-279.

Hasselhorn, M., Köller, O., Maaz, K. & Zimmer, K. (2014). Implementation wirksamer Handlungskonzepte im Bildungsbereich als Forschungsaufgabe. Psychologische Rundschau, 65, 140-149.

Hoffer, R. & Fröhlich-Gildhoff, K. (2019). Inanspruchnahme von Hilfe und Versorgung bei psychischen Auffälligkeiten im Kindergartenalter. Kindheit und Entwicklung, 28(1), 33-45. DOI 10.1026/0942-5403/a000269

Kazdin, A.E. (2011). Single-case research designs. Methods for clinical and applied settings (2nd ed.). New York: Oxford University Press.

Kolip, P. (2006). Evaluation, Evidenzbasierung und Qualitätsentwicklung. Zentrale Herausforderungen für Prävention und Gesundheitsförderung. Prävention und Gesundheitsförderung, 1 (4), 234-239.

Petermann, F. (2014). Implementationsforschung: Grundbegriffe und Konzepte. Psychologische Rundschau 3 (65), 122-128.

Staatsinstitut für Frühpädagogik (2018). KOMPIK. Kompetenzen und Interessen von Kindern. Verfügbar unter: www.kompik.de [13.08.2018].

Stockmann, R. (2006). Evaluationsforschung: Grundlagen und ausgewählte Forschungsfelder (Sozialwissenschaftliche Evaluationsforschung). Münster: Waxmann.

Wampold, B. E. (2001). The great psychotherapy debate. Models, methods and findings. Mahwah, NJ: Erlbaum.

Weltzien, D., Fröhlich-Gildhoff, K., Strohmer, J., Reutter, A. & Tinius, C. (2016). Multiprofessionelle Teams in Kindertageseinrichtungen. Evaluation der Arbeitsprozesse und Arbeitszufriedenheit von multiprofessionell besetzten Teams in Baden-Württemberg. Weinheim: Beltz Juventa.

Ziegler, H. (2010). Ist der experimentelle Goldstandard wirklich Gold wert für eine Evidenzbasierung der Praxis Früher Hilfen? Bundesgesundheitsblatt, 53(10), 1060-1066.

Autor*innen

Prof. Dr. Maike Rönnau-Böse
Zentrum für Kinder- und Jugendforschung
an der Evangelischen Hochschule Freiburg
Buggingerstr. 38
79114 Freiburg
roennau-boese@bitte-keinen-spam-eh-freiburg.de

Professorin für Kindheitspädagogik an der Evangelischen Hochschule Freiburg

Dr. Claudia Grasy-Tinius
Claudiatinius@bitte-keinen-spam-gmail.com

MA Kindheitspädagogin, ehem. Wiss. Mitarbeiterin im Zentrum für Kinder- und Jugendforschung an der Evangelischen Hochschule Freiburg

Rieke Hoffer
rieke.hoffer@bitte-keinen-spam-eh-freiburg.de

Dipl. Psych., Kinder- und Jugendlichenpsychotherapeutin, hauptamtl. Dozentin an der Evangelischen Hochschule Freiburg

Prof. Dr. Klaus Fröhlich-Gildhoff
froehlich-gildhoff@bitte-keinen-spam-eh-freiburg.de

Dipl.Psych., Psychologischer Psychotherapeut und Kinder- und Jugendlichenpsychotherapeut, Co-Leiter des Zentrums für Kinder- und Jugendforschung an der Evangelischen Hochschule Freiburg