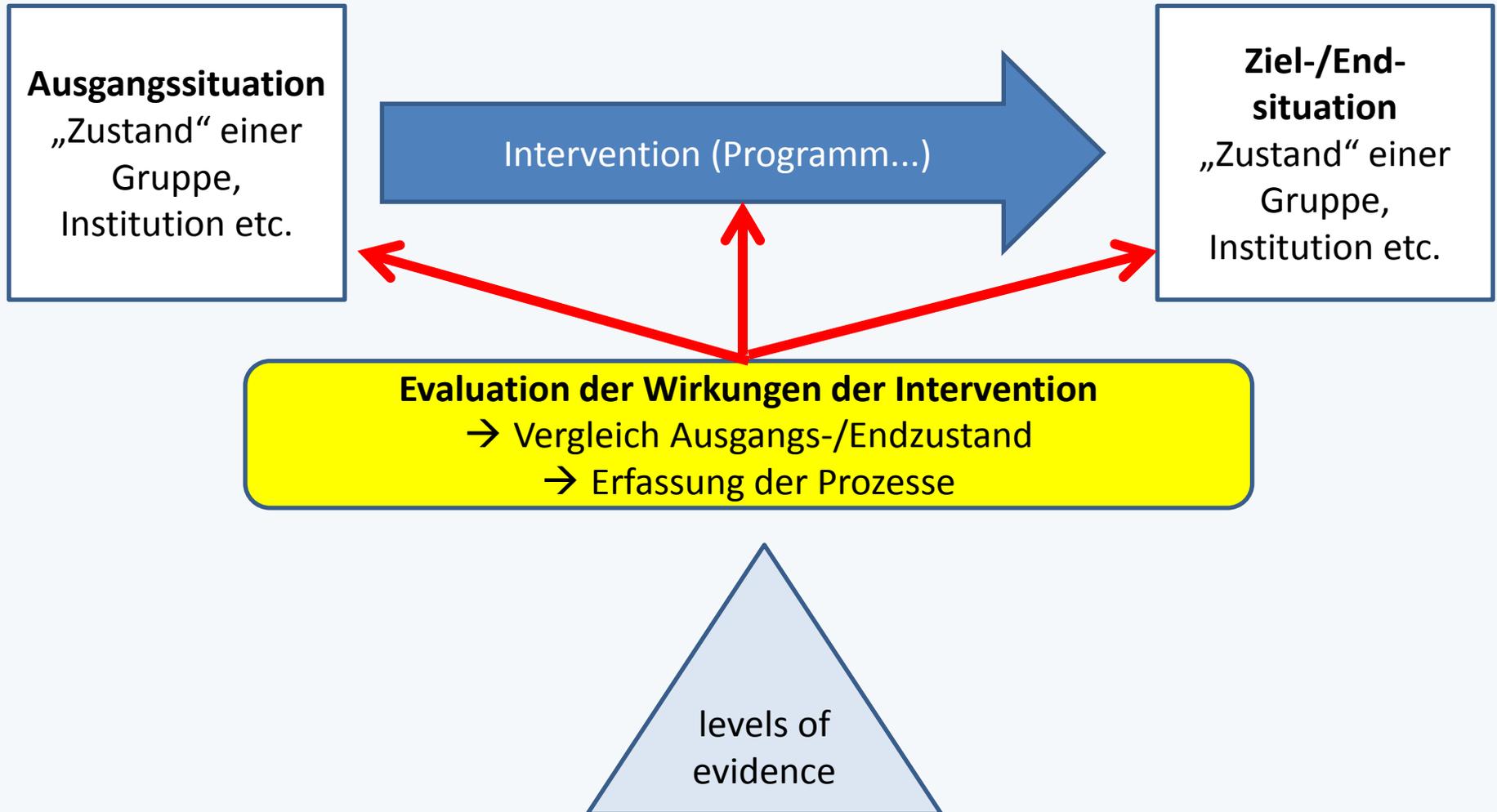


# **Methodologie bei der Evaluation komplexer Interventionen**

Prof. Dr. Klaus Fröhlich-Gildhoff  
Zentrum für Kinder- und Jugendforschung  
EH Freiburg

- (1) Klassisches Paradigma der Wirkungsforschung
- (2) Herausforderungen
- (3) Antworten

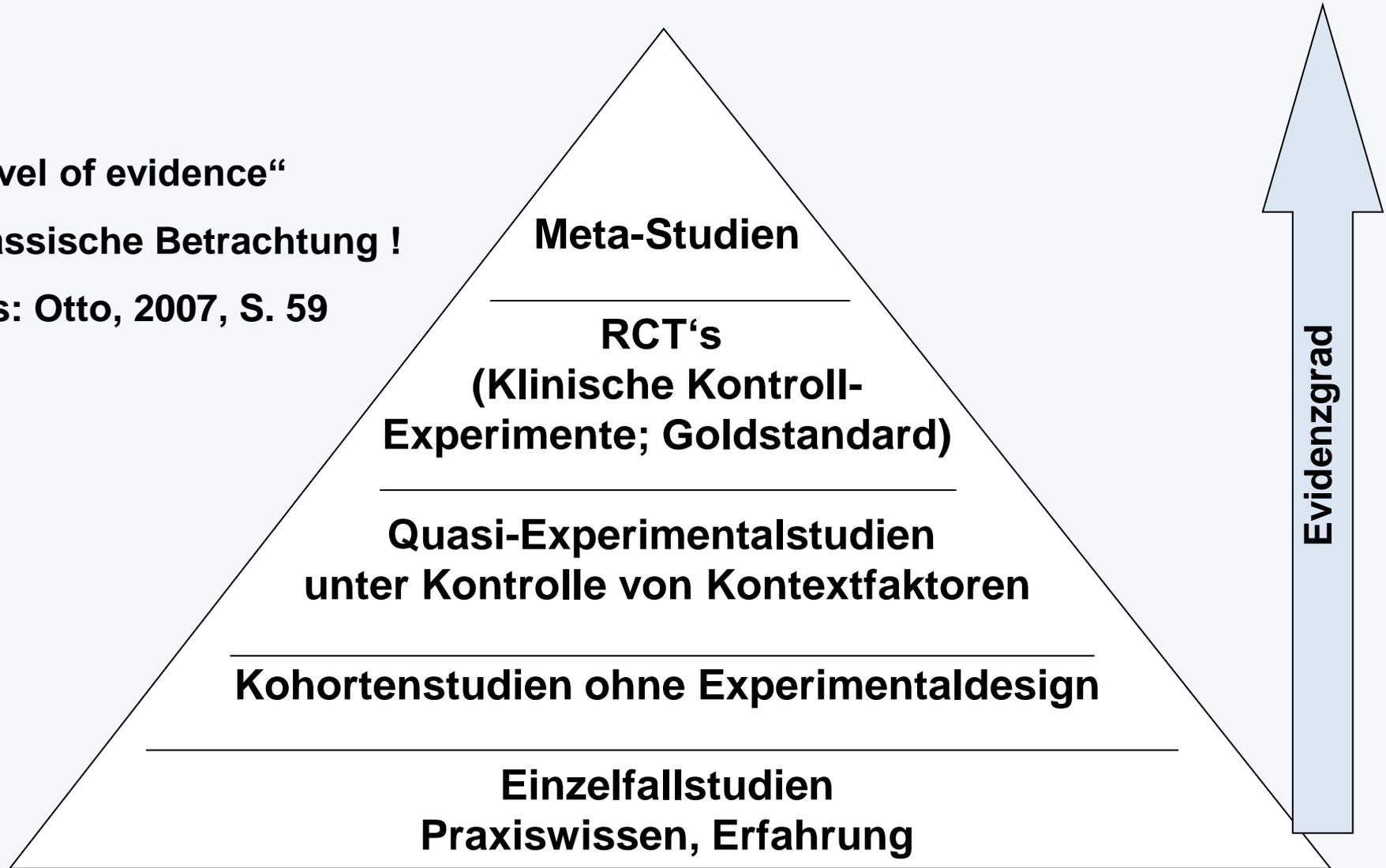
# (1) Klassisches Paradigma



„level of evidence“

Klassische Betrachtung !

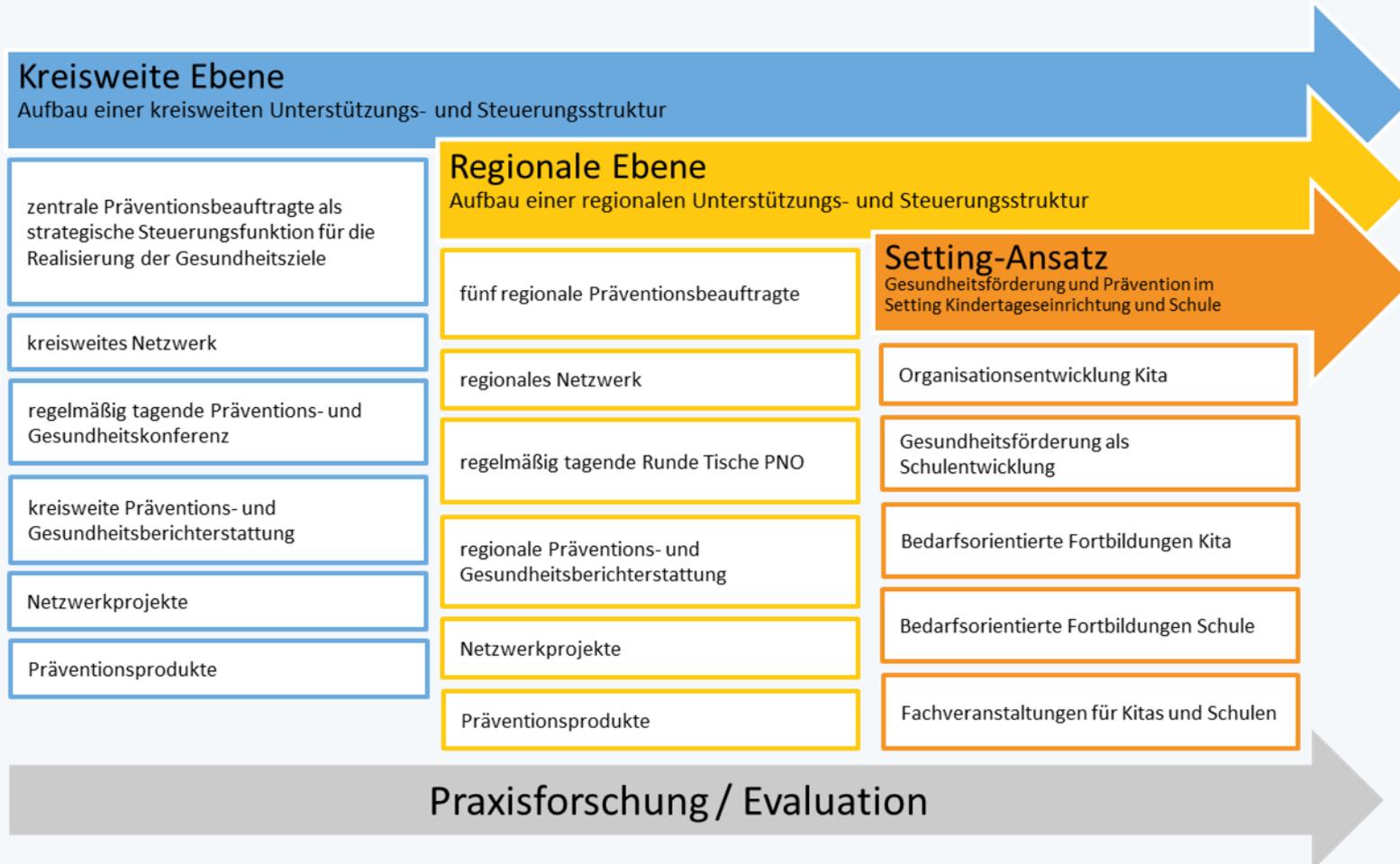
aus: Otto, 2007, S. 59



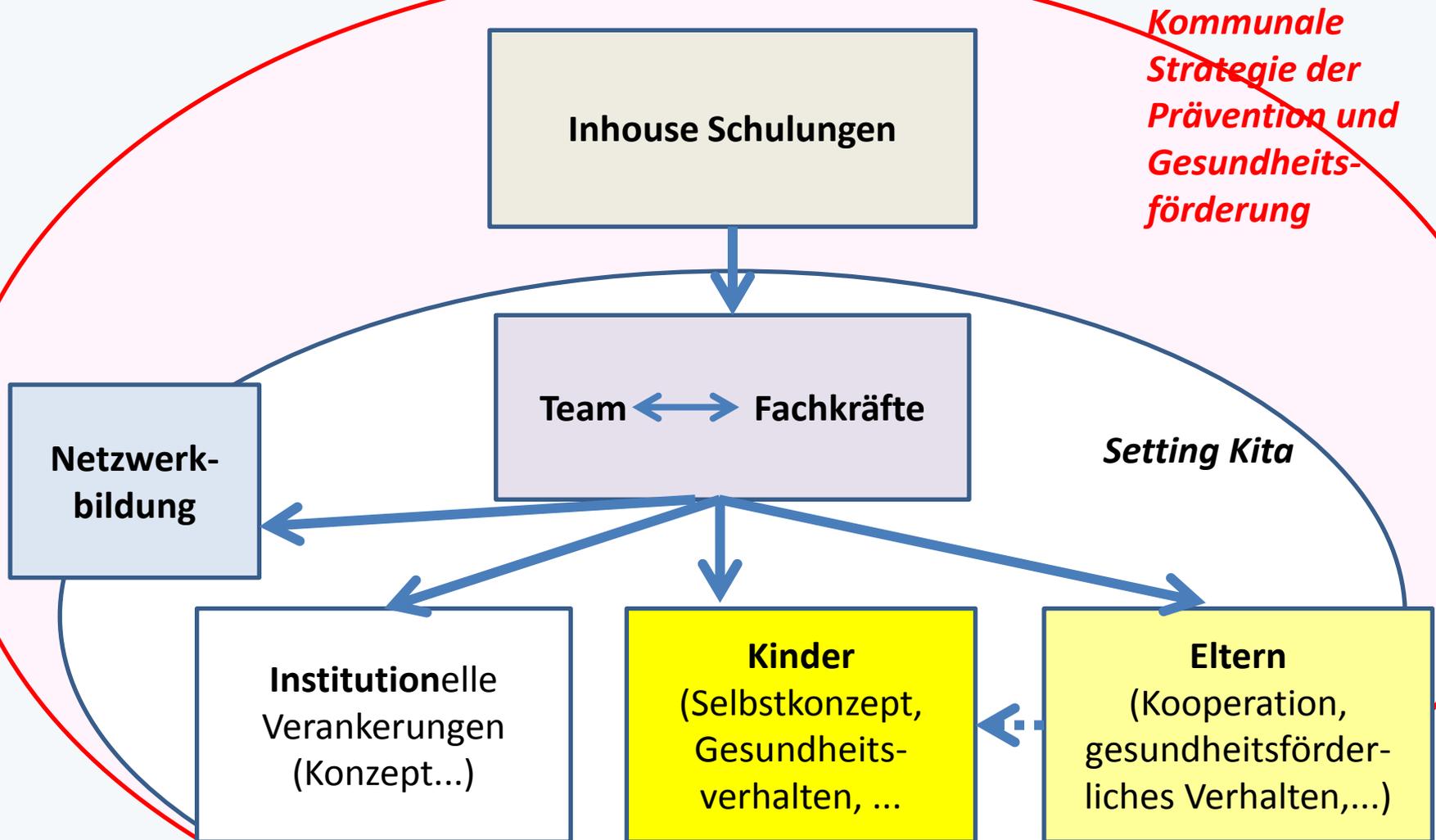
## **(2) Herausforderungen**

# Beispiel Präventionsnetzwerk Ortenaukreis

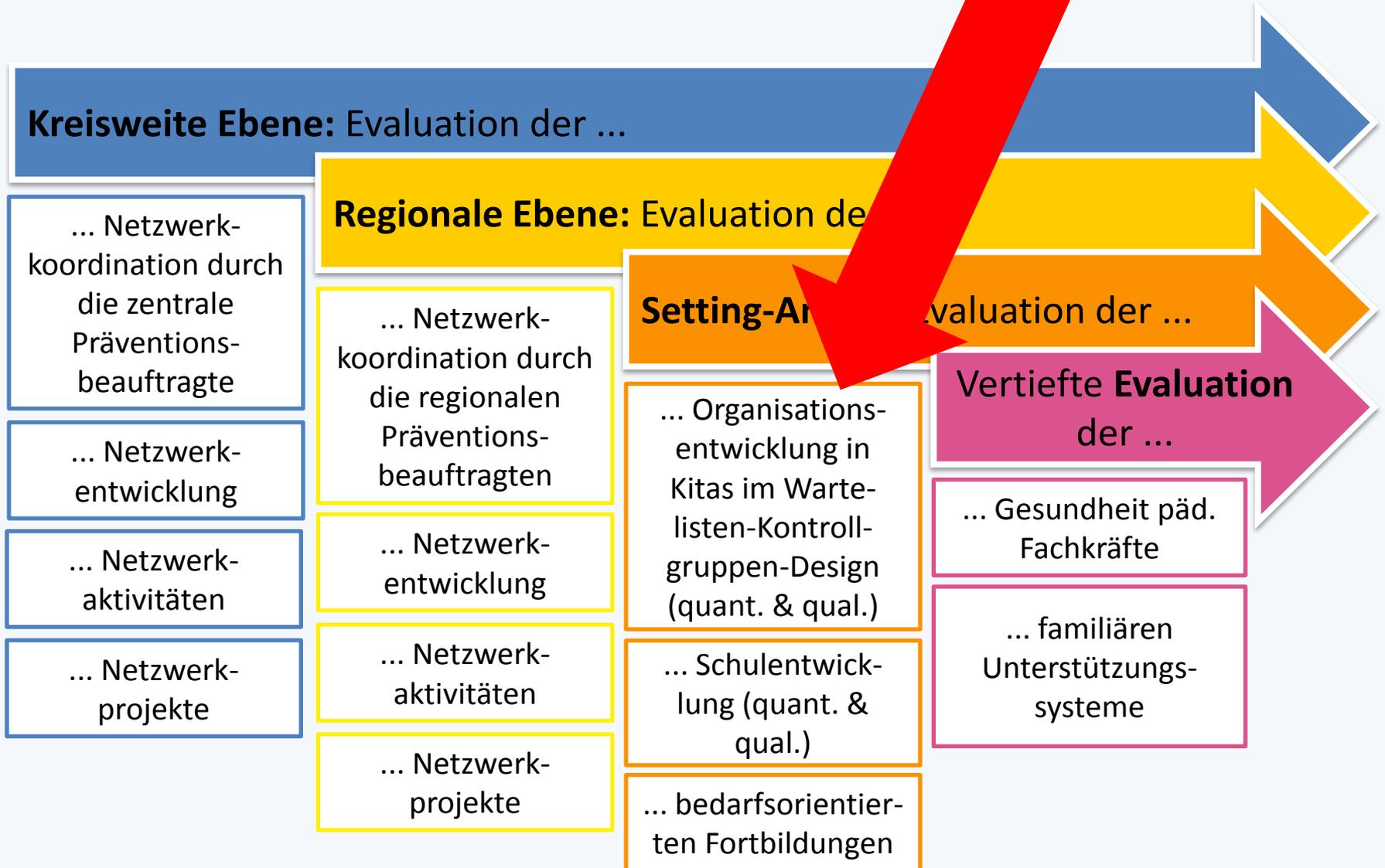
## PNO-Gesamtstruktur



# Das Beispiel „Präventionsnetzwerk Ortenaukreis“ Angenommene Wirkungskette



# PNO -Untersuchungsdesign



| Vertiefte Evaluation im gewählten Schwerpunkt<br>Wartelisten-Kontrollgruppendesign |   |  |   | Vertiefte Evaluation im Bereich Fachkräftegesundheit   |  |
|--|---|--|---|--|--|
| 5 DG Kitas bzw. 16 Grundschulen  |   | 5 KG Kitas   |   | 20 Kitas   |  |
| 1. MZP   | Basisevaluation t <sub>0</sub>  | 1. MZP   | Basisevaluation t <sub>0</sub>  | 1. MZP   | Basisevaluation t <sub>0</sub>   |
| Organisations-/ Schulentwicklungsprozess Bausteine 1-3                             |   | Prozessevaluation                                      | 12-18 Monate ohne Intervention  | Organisations-/ Schulentwicklungsprozess Bausteine 1-3 |  |
| 2. MZP   | Vertiefte Evaluation im gewählten Schwerpunkt t <sub>0</sub>                                    |  |   | 2. MZP   | Vertiefte Evaluation im Bereich Fachkräftegesundheit t <sub>0</sub>                                    |
| Organisations-/ Schulentwicklungsprozess Bausteine 4-6                             |   |  |   | Organisations-/ Schulentwicklungsprozess Bausteine 4-6 |  |
| 3. MZP   | Basisevaluation t <sub>1</sub> und Vertiefte Evaluation im gewählten Schwerpunkt t <sub>1</sub> | 2. MZP   | Basisevaluation t <sub>1</sub>  | 3. MZP   | Basisevaluation t <sub>1</sub> und Vertiefte Evaluation im Bereich Fachkräftegesundheit t <sub>1</sub> |
| 12 Monate ohne weitere Intervention  |   | Organisations-/ Schulentwicklungsprozess Bausteine 1-3 |   | 12 Monate ohne weitere Intervention                    |  |
| 4. MZP follow-up   | Basisevaluation t <sub>2</sub> und Vertiefte Evaluation im gewählten Schwerpunkt t <sub>2</sub> | 3. MZP   | Vertiefte Evaluation im gewählten Schwerpunkt t <sub>0</sub>                                    | 4. MZP follow-up                                       | Basisevaluation t <sub>2</sub> und Vertiefte Evaluation im Bereich Fachkräftegesundheit t <sub>2</sub> |
|  |   | Organisations-/ Schulentwicklungsprozess Bausteine 4-6 |   |  |  |
|  |   | 4. MZP   | Basisevaluation t <sub>2</sub> und Vertiefte Evaluation im gewählten Schwerpunkt t <sub>1</sub> |  |  |

Erläuterungen: DG=Durchführungsgruppe; KG=Kontrollgruppe; MZP=Messzeitpunkt



# → Herausforderungen

- In der Praxis ist die Vielzahl der Bedingungen nicht zu kontrollieren (keine Gleichartigkeit von Einrichtungen, Gruppen, Personen...) – oder es braucht sehr große „Zahlen“/Stichproben
- Standardisierung/Manualisierung der Intervention: Variablenkontrolle erfordert genau gleiches Vorgehen („treatment fidelity“) – Fachlichkeit besteht darin, eine Intervention auf die jeweilige Zielgruppe zu adaptieren („Passgenauigkeit“)
- Randomisierung ist oft ethisch nicht zu vertreten (Bsp. Frühe Hilfen)
- „Sleeper Effekte“ sind schwer zu erfassen
- Relativ hohe drop out Raten (Wegzug von Familien...)

→ Ausbalancieren des Verhältnisses von interner und externer Validität

„...die Kausalbeschreibungen von RCTs [gelten] nur für die tatsächlich geprüften Kontexte und Programmgestaltungen. Eine Übertragung ihrer Ergebnisse auf andere Kontexte und Programmgestaltungen ist nicht möglich. In Struktur und Ausprägung veränderte Rahmenbedingungen machen die Aussagen der Kausalbeschreibung als ‚objektive‘ Entscheidungsgrundlage wertlos“ (Ziegler, 2010, S. 164)

## (3) Antworten

- Kombination von Prozess- und Ergebnisevaluation
- Erfassung der Zusammenhänge zwischen Intervention(Intensität) und Ergebnis über den Umsetzungsindex
- ... aus der Implementationsforschung
- Pragmatic controlled trials (Porszolt et al., 2013; Bergmann et al., 2010) “studies under ideal or real world conditions use different methods, generate different information and cannot replace each other” (Porszolt et al., 2013, S. 307)

# Grundsätze der Evaluation

- Mehrere Messzeitpunkte (MZP), um Entwicklungen abbilden zu können
- Mehrere Perspektiven/Untersuchungsebenen (Bsp: Präv.beauftragte selbst *und* standardis. Tätigkeitsdoku; Selbsteinschätzung Kinder *und* Fremdeinschätzung durch Fachkräfte)
- Mehrere Methoden (Kombinationsdesign): quantitativ (Fragebögen, standardis. Protokolle) *und* qualitativ (Interviews, Dokumentenanalyse)
- Prozess- *und* Ergebnisevaluation

| Interventionsebene  | Stichprobe/Befragte/Datenbasis  | Messzeit-punkte (MZP)      | Instrumente  |
|---|---|----------------------------|--|
| <b>A1 Netzwerkentwicklung auf Kreisebene (kommunale Strategie)</b>                        | Zentrale Präventionsbeauftragte   | 4 MZP (ca. alle 12 Monate) | Interview; Tätigkeitsanalyse mit standardis. Instrument  |
|   | Stakeholder/Entscheidungsträger/Leitungskräfte und Multiplikator*innen in den drei Systemen (Gesundheit, Bildung, Jugendhilfe) auf Kreisebene | 2 Befragungszeitpunkte     | Bestands- & Bedarfserhebung (Interviews & Fragebögen), Netzwerkabfrage   |
|   | Prozessdokumentation aller Aktivitäten auf Kreisebene (Konferenzen, zentrale Veranstaltungen...)  | kontinuierlich             | Netzwerkdatenbank-analyse<br>Protokoll-/ Dokumenten-Analysen   |
| <b>A2 Netzwerkentwicklung auf regionaler Ebene („Raumschaften“) (regionale Strategie)</b> | regionale Präventionsbeauftragte  | 4 MZP (ca. alle 12 Monate) | Interview; Tätigkeitsanalyse mit standardis. Instrument  |
|   | Netzwerkakteure/ regionale Entscheidungsträger in den drei Systemen (Gesundheit, Bildung, Jugendhilfe) auf Raumschaftsebene                   | 2 Befragungszeitpunkte     | Bestands- & Bedarfserhebung ((Interviews & Fragebögen), Netzwerkabfrage  |
|   | Prozessdokumentation aller Aktivitäten auf Raumschaftsebene (regionale Runde Tische, regionale Veranstaltungen, lokale Netzwerke...)          | Kontinuierlich             | Netzwerkdatenbankanalyse<br>Protokoll/ Dokumenten-Analysen   |
| <b>B Organisationsentwicklung in Kitas und Schulen</b>                                    | Basisevaluation aller teilnehmenden Kitas und Schulen   | Prä/post                   | „Checkliste“ zum Stand der PräV und GF und zu Entwicklungsbedarfen; Erfassung der Umsetzungsintensität („Umsetzungsindex“) aus drei Perspektiven |
|   |   | post                       | s.u.   |
|   | Vertiefte Evaluation im Wartelisten-Vergleichsgruppendesign   | Prä/post/follow up         |  |
| <b>C Ebene Institution</b>  | Konzeption der Einrichtung, Netzwerk der Einrichtung  | Prä/post                   | Dokumentenanalyse, Netzwerkkarten  |
|   | Einrichtungsleitungen   | Prä/post                   | Interviews   |
| <b>D Ebene Fachkräfte</b>   | <b>B</b> Fachkräfte und Referent*innen bei den Inhouse-Schulungen   | Prä/post rig               | Fragebögen zur Zufriedenheit und selbsteingesch. Erfolg der Schulungen   |
|   |   | regelmäßig                 | Standardis. Protokollbögen zum Verlauf der Schulung  |
|   | <b>C</b> Fachkräfte   | Prä/post                   | Standardis. Fragebögen zur Kompetenzentwicklung, Arbeitsbelastung/-zufriedenheit, Einstellung zu PräV und GF                                     |
|   |   | Prä/post                   | leitfadengestützte Interviews zur Einstellung zu PräV/GF und zum Projekt   |

| Interventionsebene  | Stichprobe/Befragte/Datenbasis  | Messzeit-punkte (MZP)                         | Instrumente  |
|---|---|---|--|
| <b>Organisationsentwicklung in Kitas und Schulen</b>                                |   |   |  |
| <b>E Ebene Eltern</b>   | Alle Eltern aus den 11 Wartelisten- Vergleichsgruppen- Einrichtungen      | Prä/post                                      | Fragebögen: Einstellung zu PräV/GF; Kooperation mit der Kita   |
|   |   | Prä/post                                      | Leitfadengestützte Interviews  |
| <b>F Ebene Kinder</b>   | Alle Kinder aus den 11 Wartelisten- Vergleichsgruppen- Einrichtungen      | Prä/post/follow up (nur DG, nicht Vgl gruppe) | Standardis. Fragebögen zur Selbsteinschätzung (Resilienz, Selbstkonzept) und Fremdeinschätzung durch Fachkräfte (psych. Wohlbefinden, soziale Beziehungen) |
| <b>Bedarfsorientierte Fortbildungen</b>   | Fachkräfte und Referent*innen bei den jeweiligen Fortbildungen            | Prä/post                                      | Fragebögen zur Zufriedenheit und selbsteingesch. Erfolg der Fortbildungen  |
|   |   | regelmäßig                                    | Standardis. Protokollbögen zum Verlauf der Schulung  |
| <b>Zusätzliche Fragestellungen</b>  |   |   |  |
| <b>Gesundheit der pädagogischen Fachkräfte (i.R. der Organisations-entwicklung)</b> | Fachkräfte aus 19 Einrichtungen (Auswahl nach Kriterien der Repräsentanz) | 3 MZP im Projektverlauf                       | Standardis. Fragebogen zum Gesundheitsverhalten und zur Einschätzung der Gesundheitsstrategie der Institution  |
|   | Leitungen der 19 ausgew. Einrichtungen                                    | 3 MZP im Projektverlauf                       | Interviews zur Gesundheitsstrategie der Institution  |
| <b>Netzwerkentwicklung von Familien mit bes. Unterstützungsbedarf</b>               | 17 Familien mit bes. Unterstützungsbedarf                                 | 5 MZP im Projektverlauf                       | Leitfadengest. Interviews zur Netzwerkentwicklung (privat und institutionell); Netzwerkkarten  |

# Zusammenhänge Intervention - Effekte

- Erfassung der Umsetzungsintensität
  - Zusammenhangsanalysen
  - Systematische Analyse des Implementationsprozesses anhand der Erkenntnisse der Implementationsforschung
- (Petermann, 2014; Hasselhorn et al., 2014; Durlak, 2013; Durlak & DuPree, 2008; Fröhlich-Gildhof & Hoffer, 2017)

# Umsetzungsindex

**Frage: Hängt die Intensität der Umsetzung der Intervention [OE/SE] mit den prä/post-Veränderungen auf Kinderebene zusammen?**

## **Schritte:**

- 1) Einschätzung der Umsetzungsqualität/-intensität anhand von 13 Kriterien aus drei Perspektiven (Leitung, Beteiligte, Prozessbegl.) → Umsetzungsindex (UI) pro Einrichtung
- 2) Rangreihe der Einrichtungen (hohe vs. niedrige Umsetzungsintensität)
- 3) Vergleich dieser Rangreihe mit der Rangreihe der gemessenen Veränderungen in den einzelnen Tests
- 4) Extremgruppenvergleiche

## **Ergebnisse:**

- a) Keine signifikante Unterscheidung der Rangreihen → hoher Rangplatz Umsetzungsintensität unterscheidet sich nicht von hohem Rangplatz Veränderungsmaße
- b) Korrelation der Rangreihen
- c) Einrichtungen mit hohem UI zeigen bessere Entwicklungen bei einzelnen Tests (nicht bei allen!)

| Indikator |  | Erläuterungen   | Umsetzungsgrad |   |   |   |
|-----------|--|---|----------------|---|---|---|
|           |  |   | 0              | 1 | 2 | 3 |
| 1         | Begleitung des Umsetzungsprozesses durch die Einrichtungsleitung           | <p>0 Leitung hat nicht am Prozess teilgenommen</p> <p>1 Leitung war am Anfang und/oder sporadisch am Prozess beteiligt</p> <p>2 Leitung war an Sitzungen/Aktivitäten beteiligt, hat kontinuierlich für das Projekt geworben und einen verbindlichen Auftrag an die Steuerungsverantwortlichen erteilt</p> <p>3 Leitung war kontinuierlich am Projekt beteiligt, hat es aktiv kontinuierlich unterstützt, hat für verbindliche Verankerung und die Einbettung in andere Prozesse der Organisationsentwicklung gesorgt</p>                      |                |   |   |   |
| 2         | Funktionsfähigkeit und Bedeutung der Steuerungsverantwortlichen (STV)      | <p>0 STV hat/haben keinen Einfluss auf den Umsetzungsprozess</p> <p>1 STV hat/haben unregelmäßig den Umsetzungsprozess begleitet</p> <p>2 STV hat mehrfach getagt und den Umsetzungsprozess begleitet</p> <p>3 STV war/waren wesentlicher Motor für den Umsetzungsprozess und hat Steuerung übernommen</p>  |                |   |   |   |
| 3         | Bedeutung der Gesundheitsförderung (GF) allgemein im Kitaleben             | <p>0 GF wurde cursorisch erwähnt und hatte keinen wirklichen Platz im Kitaleben/in der Organisationsentwicklung</p> <p>1 GF wurde oft angesprochen, hatte jedoch keinen festen Platz im Kitaleben/in der Organisationsentwicklung (Umsetzung lediglich durch einzelne)</p> <p>2 GF wurde regelmäßig angesprochen und der Umsetzungsprozess reflektiert</p> <p>3 GF war regelmäßig Thema bei Teamsitzungen etc.; GF hat festen Platz (z.B. durch verbindliche „Resilienzstunde“, regelhafte Reflexion der Essensituation etc.) in der Kita</p> |                |   |   |   |
| 4         | Teilnahme der päd. Fachkräfte an den Fortbildungen (FB)                    | <p>0 Päd. Fachkräfte nahmen nur als Teilgruppe und nicht kontinuierlich an den FB teil</p> <p>1 eine feste, kleine Gruppe von päd. Fachkräften nahm an FB teil</p> <p>2 FB waren regelmäßig gut besucht und es gab offene, produktive Diskussionen</p> <p>3 im Prinzip war immer das gesamte Team bei den FB anwesend und es gab offene, produktive Diskussionen</p>  |                |   |   |   |
| 5         | Umsetzung eines konkreten Projekts zur Gesundheitsförderung in den Gruppen | <p>0 es wurde kein Projekt umgesetzt</p> <p>1 Es wurde ein Einzelprojekt in einzelnen Gruppen umgesetzt</p> <p>2 Ein konkretes Projekt wurde regelmäßig in einer Mehrzahl der Gruppen umgesetzt</p> <p>3 Ein konkretes Projekt wurde regelmäßig in allen Gruppen umgesetzt</p>  |                |   |   |   |

# Umsetzungsindex

**Frage: Hängt die Intensität der Umsetzung der Intervention [OE/SE] mit den prä/post-Veränderungen auf Kinderebene zusammen?**

## **Schritte:**

- 1) Einschätzung der Umsetzungsqualität/-intensität anhand von 13 Kriterien aus drei Perspektiven (Leitung, Beteiligte, Prozessbegl.) → Umsetzungsindex (UI) pro Einrichtung → Bogen
- 2) Rangreihe der Einrichtungen (hohe vs. niedrige Umsetzungsintensität)
- 3) Vergleich dieser Rangreihe mit der Rangreihe der gemessenen Veränderungen in den einzelnen Tests
- 4) Extremgruppenvergleiche

## **Ergebnisse:**

- a) Keine signifikante Unterscheidung der Rangreihen → hoher Rangplatz Umsetzungsintensität unterscheidet sich nicht von hohem Rangplatz Veränderungsmaße
- b) Korrelation der Rangreihen
- c) Einrichtungen mit hohem UI zeigen bessere Entwicklungen bei einzelnen Tests (nicht bei allen!)

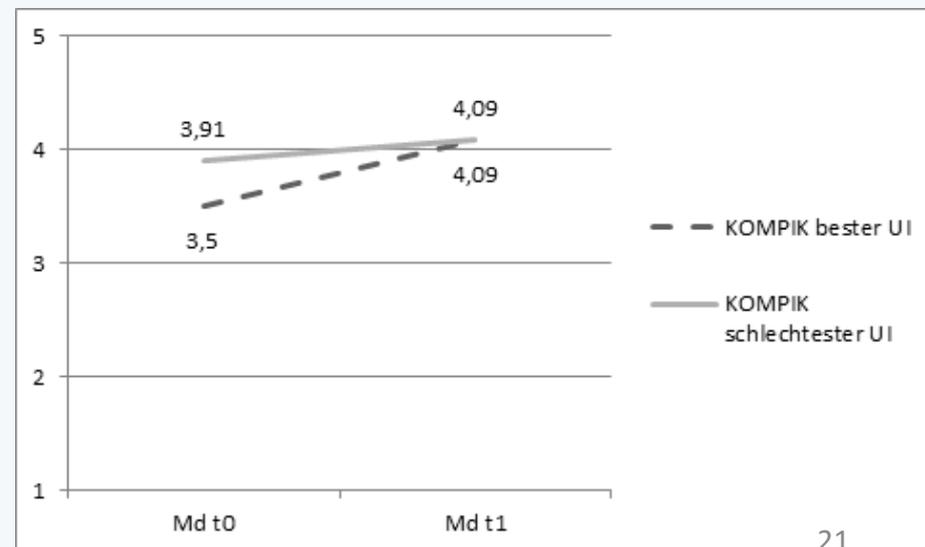
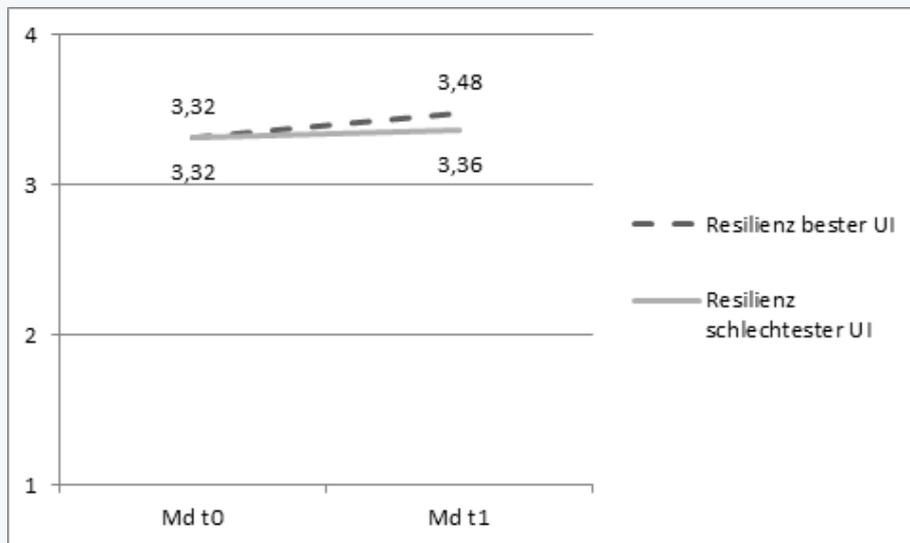
Zu b) Korrelation UI – Tests → keine signifikante Korrelation, deutl. Unterschiede zwischen den Tests

|  |                             | Resilien<br>z | SKF_F | SKF_K | SKF_S | PsWb        | SozBez | KOMPI<br>K  |
|--|-----------------------------|---------------|-------|-------|-------|-------------|--------|-------------|
| Gesamtscore<br>Umsetzungsindex<br>(UI) | Korrelationsk<br>oeffizient | <b>.427</b>   | .164  | .373  | -.082 | <b>.573</b> | .327   | <b>.473</b> |
|  | Sig. (2-seitig)             | .190          | .631  | .259  | .811  | .066        | .326   | .142        |
|  | N                           | 11            | 11    | 11    | 11    | 11          | 11     | 11          |

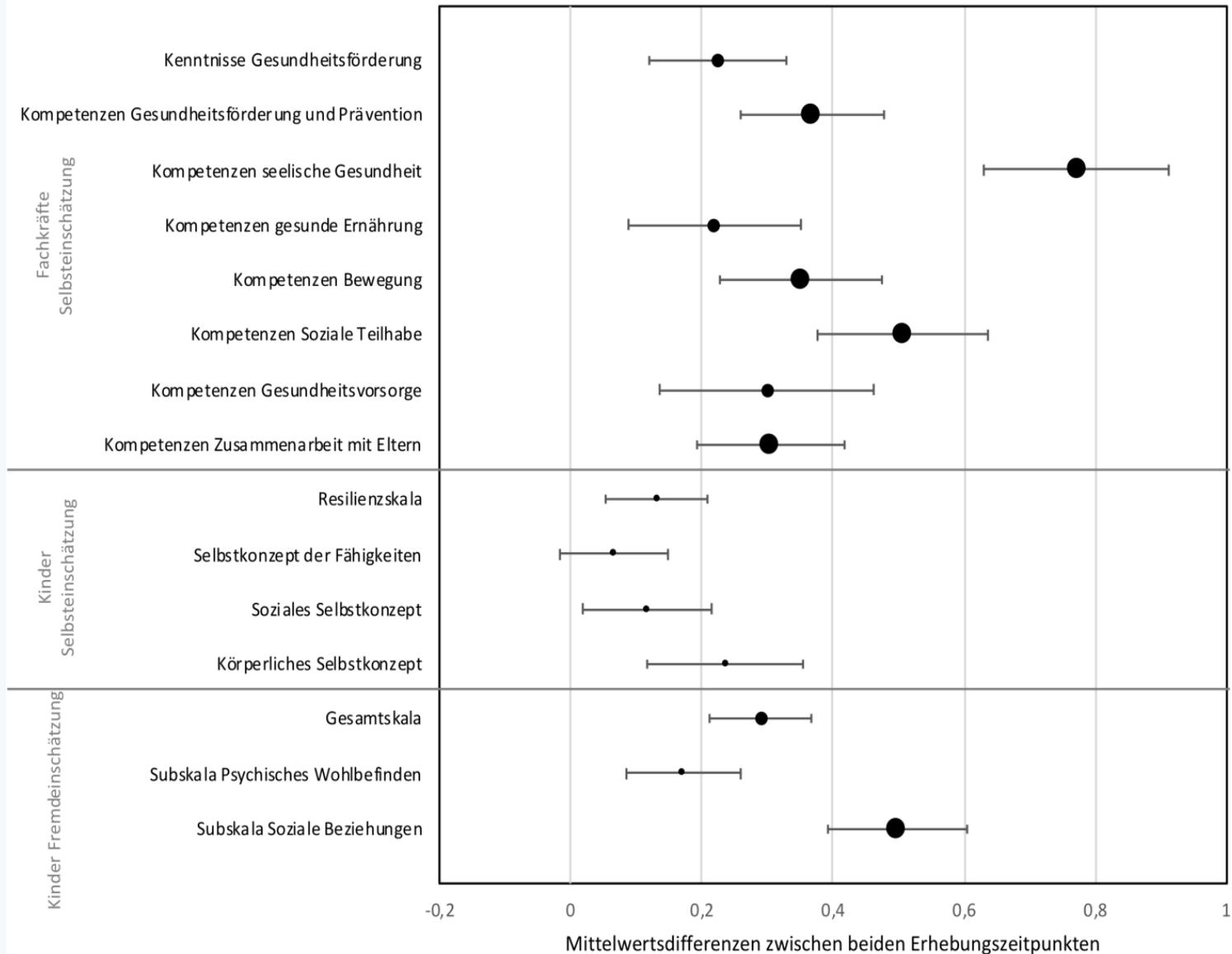
Anmerkungen: \*Korrelation ist auf dem .05 Niveau signifikant.

\*\* Korrelation ist auf dem .001 Niveau signifikant.

Zu c) Extremgruppenvergleiche: Vergleich der besten beiden Einrichtungen im UI mit denen, die am Ende der Rangreihe stehen → tendenzielle Veränderungen



# Zusammenhangsanalysen „Forest Plot“ (Lewis, 2001)



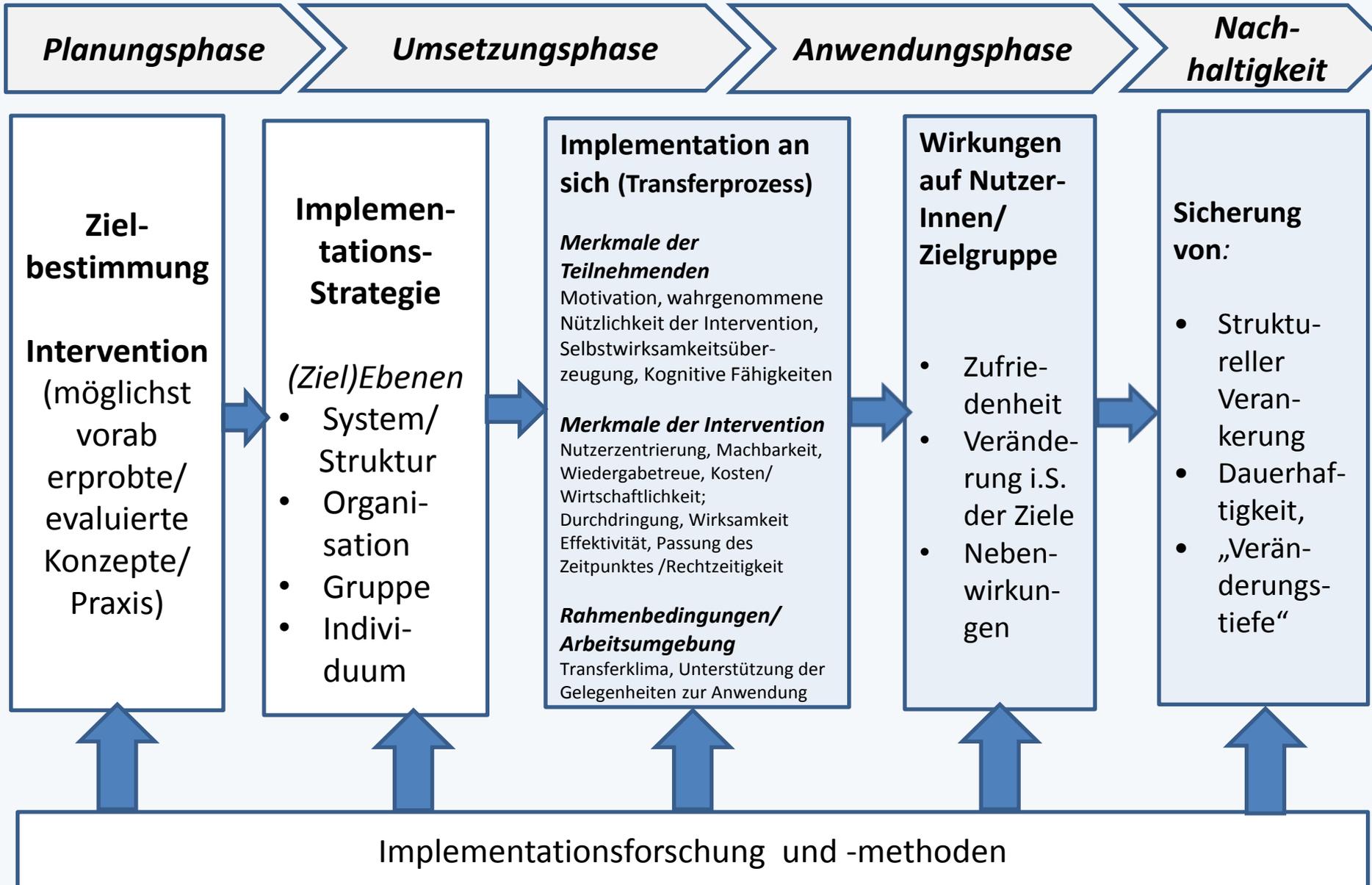
# Untersuchung des Zusammenhangs zwischen dem Kompetenzzuwachs der Fachkräfte und dem fremdeingeschätzten Wohlbefinden der Kinder

|   | Gesamtskala              | Psychisches Wohlbefinden | Soziale Beziehungen    |
|---|--------------------------|--------------------------|------------------------|
| Kenntnisse Gesundheitsförderung                 | $\tau (221) = .061$      | $\tau (221) = .094$      | $\tau (221) = .015$    |
| Kompetenzen Gesundheitsförderung und Prävention | $\tau (221) = .140^{**}$ | $\tau (221) = .191^{**}$ | $\tau (221) = .033$    |
| Kompetenzen seelische Gesundheit                | $\tau (221) = .035$      | $\tau (221) = .115^*$    | $\tau (221) = -.084$   |
| Kompetenzen gesunde Ernährung                   | $\tau (221) = -.039$     | $\tau (221) = .030$      | $\tau (221) = -.123^*$ |
| Kompetenzen Bewegung                            | $\tau (221) = .052$      | $\tau (221) = .131^{**}$ | $\tau (221) = -.084$   |
| Kompetenzen Soziale Teilhabe                    | $\tau (221) = .080$      | $\tau (221) = .160^{**}$ | $\tau (221) = -.053$   |
| Kompetenzen Gesundheitsvorsorge                 | $\tau (221) = .059$      | $\tau (221) = .134^{**}$ | $\tau (221) = -.060$   |
| Kompetenzen Zusammenarbeit mit Eltern           | $\tau (221) = .077$      | $\tau (221) = .119^*$    | $\tau (221) = -.005$   |

Anmerkung: Kendall's tau  $\tau$ : \* $p < .05$ ; \*\* $p < .01$ ; \*\*\*  $p < .001$ . Signifikanz (2-seitig).

→ schwächerer, aber statistisch bedeutsamer Zusammenhang zwischen dem selbsteingeschätzten Anstieg der Kompetenzen zur GF und Der positiven Veränderung des Psychischen Wohlbefindens der Kinder

# Bezüge zu Erkenntnissen der Implementationsforschung



# Bezüge zu relevanten Implementationsfaktoren I

- Die **Akzeptanz der Intervention bei den Beteiligten** : Dort wo die Ziele des Projektes und das Vorgehen transparent dargestellt und mit den Betroffenen diskutiert worden sind, gelang die Umsetzung besser.
- Die Inhalte der Intervention mussten für die Zielgruppe *angemessen und passend* sein.
- Die **Übernahme der Intervention** in die Praxis der Einrichtungen gelang dann, wenn zum einen die Methoden (zum Beispiel Inhalte der Resilienzförderung) für die Praktiker\*innen gut (und schnell) nutzbar waren, zum anderen, wenn entsprechende Rahmenbedingungen (zum Beispiel Zeiträume für "Resilienzstunden") gegeben waren.
- Wichtig: **Sicherung der Kosten**, geringer/kein finanzieller Aufwand für die Beteiligten - die eigens bereit zu stellenden Ressourcen (Personal und Zeit) waren an sich, zumindest am Anfang, eine Hürde.
- Für die **Nachhaltigkeit** war es wichtig, sehr konkret Maßnahmen zu planen und auch verbindlich zu vereinbaren – hierzu waren klar überschaubare Projekte ein wichtiger Anker, ebenso aber auch konzeptionelle Verankerung.

## Bezüge zu relevanten Implementationsfaktoren II

- Bedeutung der "**Merkmale der einzelnen Einrichtung**" (Hasselhorn et al., 2014; Durlack und DuPre, 2008) auch im Projekt PNO bestätigt. Das betraf besonders die Offenheit für den Wandel, die Kooperation und Unterstützung im Team, die Unterstützung durch die Leitung, aber auch relevant unterstützende Akteur\*innen im Team.
- Ebenso wichtig war der **fachpolitische Rückhalt**, also die Einbettung in das Gesamtprojekt und damit der grundlegende Einbezug der Träger der Institution.
- Als bedeutsam stellen sich auch die „**Merkmale der pädagogischen Fachkräfte**“ heraus. Besonders wichtig war hier, außer Kompetenz und persönlicher Überzeugung, die Motivation der Fachkräfte.
- Eine wichtige Rolle spielte der in der Literatur oft vernachlässigte Faktor der **Kompetenz der Umsetzer\*innen**, hier der Referent\*innen beziehungsweise Prozessbegleiter\*innen. Deren Bewertung durch die Teilnehmer\*innen war in der Regel sehr hoch. Wenn hier die Passung zwischen Teilnehmer\*innengruppe und deren Bedürfnissen einerseits und den Kompetenzen, aber auch der Adaptabilität der Referent\*innen gelang, fanden sich deutliche Hinweise für positive Umsetzungsprozesse.

## (4) Fazit, Perspektiven

*„Eine komplexe Realität erfordert einen komplexen Zugang“;  
Engelke 1992*

- Kombination aus quantitativen und qualitativen Methoden
- Multiperspektivität
- Kombination aus Prozess- und Ergebnisevaluation
- Wirksamkeitsforschung in der Praxis benötigt auch Vergleichsgruppendedesigns – aber nicht nach dem RCT Prinzip

Vielen Dank für Ihre Aufmerksamkeit



[www.zfkj.de](http://www.zfkj.de)

[froehlich-gildhoff@eh-freiburg.de](mailto:froehlich-gildhoff@eh-freiburg.de)

Prof. Dr. Klaus Fröhlich-Gildhoff, ZfKJ EH  
Freiburg