

Dr. Anneke Bühler

IFT

IFT · Institut
für Therapieforschung
Parzivalstraße 25
80804 München

www.ift.de

Leiter des Instituts
Prof. Dr. Ludwig Kraus

Tel.: 089/360804-83
Fax: 089/360804-69
E-Mail: buehler@ift.de

2013-03-06

„Stellungnahme zum HTA Bericht *„Föderale Strukturen der Prävention von Alkoholmissbrauch bei Kindern und Jugendlichen“*, *Addendum 112a*, zur Beurteilung der Wirksamkeit der Alkoholprävention in Deutschland“

Träger: IFT · Institut für
Therapieforschung
Gemeinnützige Gesellschaft mbH
Registergericht München
HRB 46395

Geschäftsführung:
Prof. Dr. Gerhard Bühringer

Verwaltungsleiter:
Hartmut Behle

Commerzbank AG
Kto.-Nr. 03 276 806 00
BLZ 700 800 00

Stadtsparkasse München
Kto.-Nr. 23 168 370
BLZ 701 500 00

Ust.-IdNr.: DE 129521698

Steuer Nr. 143/237/40020

1. Hintergrund und Ziel der Stellungnahme

Der HTA-Bericht 112 „Föderale Strukturen der Prävention von Alkoholmissbrauch bei Kindern und Jugendlichen“ und insbesondere das Addendum 112a (Korczak, 2012, im folgenden „HTA_Bericht“ genannt und Autor genannt) haben mit ihren Ergebnissen hinsichtlich der Wirksamkeit deutscher Alkoholpräventionsprojekte in der Fachöffentlichkeit große Diskussionen ausgelöst.

Das Addendum, d.h. seine Methode, seine Ergebnisse und seine Schlussfolgerungen werden im Folgenden kommentiert. Zudem wird ein Gegenvorschlag zur Evaluation und Wirksamkeitsdefinition von suchtpreventiven Maßnahmen in Deutschland mit dem Ziel einer evidenz-basierten Praxis vorgestellt.

2. Methodische Vorgehensweise

Die Stellungnahme fasst zuerst die Methode, Ergebnisse und Schlussfolgerungen des Berichts zur Fragestellung der Evaluation und Wirksamkeit deutscher Alkoholpräventionsmaßnahmen (Fragestellung 3 und 4 von des HTA-Berichts, S. 5) zusammen. Sie beurteilt die Methode und allgemeinen Schlussfolgerungen aufgrund wissenschaftlicher Standards und internationaler Forschungsbefunde und diskutiert die Prämissen der Bewertung der Evaluationslage in Deutschland (S.7) unter Einbezug (inter-)nationaler Herangehensweisen und Perspektiven (z.B. DeGEval, 2008; Center for Substance Abuse Prevention, 2009; Society for Prevention Research, 2004, Elkeles & Broesskamp-Stone, 2010). Der Gegenvorschlag, wie „Wirksamkeit“ von Alkoholpräventionsmaßnahmen mit dem Ziel einer evidenz-basierten Praxis definiert werden kann, und welche Evaluationsanforderungen dafür notwendig sind, baut auf präventionsspezifischen Konzeptionen und Evaluationsstandards auf (EMCDDA, 2011;Walter et al., 2012; Broesskamp-Stone, 2012).

3. Zusammenfassung des Addendum des des HTA-Berichts 112a (Korczak, 2012)

3.1 Fragestellungen des HTA-Berichts

Ausgehend von der Einschätzung eines „strittigen“¹ Erfolgs der deutschen Alkoholpräventionsbemühungen setzt sich das Addendum u.a. die Aufgabe „die Wirkung deutscher Alkoholpräventionsprojekte zu erfassen“ und fragt „welche Alkoholpräventionsmaßnahmen wirksam sind“ (Korczak, 2012, S. 1). Konkret wird den Fragestellungen „Zu welchen

¹ Der strittige Erfolg der Alkoholpräventionsmaßnahmen wird im gleichbleibend hohen riskanten Alkoholkonsum und den ansteigenden Klinikeinweisungen Jugendlicher wegen Alkoholintoxikation begründet. Auf eine differenzierte Einordnung der Klinikstatistiken, die auch eine andere Interpretation zulässt (Kraus et al., 2012), wird verzichtet.

Alkoholpräventionsmaßnahmen liegen in Deutschland Evaluationen vor?“ (Fragestellung 3) und „Welche Alkoholpräventionsmaßnahmen in Deutschland sind effektiv (Ergebnisevaluation)?“ (Fragestellung 4) nachgegangen (Korczak, 2012, S. 5).

3.2 Methodik des HTA-Berichts

Um diese Fragen zu beantworten wurde eine schriftliche Befragung bei Bundeseinrichtungen, Ländereinrichtungen, Kommunen und Krankenversicherungen durchgeführt.

Rekrutierung Stichprobe. Durch eine Internetrecherche wurden Träger und Koordinatoren von Alkoholpräventionsmaßnahmen (APM) identifiziert. Zudem wurden Gesundheitsreferate von Großstädten (>500.000 Einwohner) und die für Suchtprävention zuständigen Referate von vier großen gesetzlichen Krankenkassen sowie des Verbandes der Privaten Krankenversicherung in die Datenbank aufgenommen. Dem BMG und der BZgA wurde die Datenbank übermittelt und um eventuelle Ergänzungen gebeten. Es erfolgten keine Ergänzungen.

Rücklauf. Auf Bundesebene wurden 17 Einrichtungen angeschrieben. Eine Einrichtung antwortete per Fragebogen, zwei Einrichtungen übermittelten relevante Informationen ohne den Fragebogen zu nutzen (5,8 bzw. 17,6% Rücklauf).

Auf Länderebene wurden insgesamt 69 Einrichtungen aus 16 Bundesländern angeschrieben (Gesundheitsministerien, Landesgesundheitsämter, Landes(koordinierungs)stellen, Landesarbeitsstellen der Aktion Jugendschutz und andere relevante Einrichtungen). Aus 10 Bundesländern wurden von 15 Einrichtungen Fragebogen ausgefüllt (21,7%).

Auf kommunaler Ebene wurde 165 Einrichtungen das Erhebungsinstrument zugesandt. Aus dem Land Bayern antworteten mehr kommunale Einrichtungen als von dem Projektteam angeschriebene (54 anstatt 28, 192,9%). Ansonsten konnte ein Rücklauf von 23 Einrichtungen aus einem bundesweit angeschriebenen Pool von 137 Einrichtungen festgestellt werden (16,8%).

Stichprobe. Die 208 Projekte, die von allen teilnehmenden Einrichtungen beschrieben wurden, stammen aus 15 Bundesländern. Aus Bayern werden dabei 130, also 62,5% der Projekte berichtet. Acht Projekte werden häufiger als einmal genannt.

Die Identifizierung der relevanten Einrichtungen scheint gelungen, wiewohl Kultusministerien und Stiftungen nicht befragt wurden. Der Rücklauf der Datenerhebung ist dagegen mangelhaft. Dass auf Erinnerungsschreiben aus „Budgetgründen“ verzichtet wurde (S.6), ist nicht verständlich. Die Gründe für die Rücklaufprobleme werden nicht berichtet. Zudem sind die berichteten Projekte regional sehr schief verteilt, ein Drittel stammt aus einem Bundesland. Damit ist die Aussagekraft

der erhobenen Daten stark beeinträchtigt. Der Autor selbst gibt an, „nicht den Anspruch repräsentativ für das Projektgeschehen in Deutschland zu sein“ zu erheben (S.17).

Instrument. Das Erhebungsinstrument wurde für die Befragung „eigens ... entwickelt“ (S. 6). Referenzen zu anderen Erhebungsinstrumenten oder Klassifikationen von APM werden nicht gemacht. Es umfasst 17 Fragen zu folgenden Inhalten:

- Titel/Name, Laufzeit der APM, Träger und Kooperationspartner
- Setting der APM
- Zielgruppe der APM

Die Zielgruppe der APM soll anhand ihres Alters beschrieben werden. Hinsichtlich des Risikoprofils der Zielgruppe, die eine Einteilung in universell, selektiv oder indiziert erlaubt, werden keine Daten erhoben. Auch ob geschlechtsspezifisch vorgegangen wird, wird nicht erfragt.

- Inhaltliche Ausgestaltung des Projekts
- Budget und Finanzierung der APM
- Formulierung der Ziele und Zielgrößen der APM Offen: Wenn ja, welche
- Evaluation der APM und Outcomeparameter der Evaluation
- Messzeitpunkte der Evaluation

Operationalisierung des Evaluationsgegenstands „Alkoholpräventionsmaßnahme (APM)“

Im Fragebogen wird offen nach Projekten gefragt, die der Prävention des Alkoholmissbrauchs dienen sollen. Damit beschränkt sich die Erhebung auf Projekte, die alkoholspezifisch sind, unspezifische Projekte zur Prävention von Problemverhalten (z.B. Elterntrainings, gesundheitsfördernde Maßnahmen) werden nicht abgebildet. APM auf Bundes-, Landes- oder kommunaler Ebene, die meist ein Maßnahmenbündel darstellen, werden genauso als eine APM-Einheit aufgefasst wie ein Elternabend, ein Klassenprogramm oder ein Arbeitskreis. In diesen heterogenen Einheiten werden die 208 APM in der Ergebnistabelle dargestellt und wird später die Auswertung vorgenommen.

Operationalisierung der Ergebnisvariablen des HTA-Berichts

Die Ziele des Addendums sind zu beurteilen, ob eine Maßnahme evaluiert und wirksam ist. Dies sind also die Ergebnisvariablen des HTA-Berichts. Im abschließenden 7. Kapitel des HTA-Berichts wird zudem implizit die Begründetheit der Inhalte der APM beurteilt, so dass diese Variable für die Stellungnahme auch als Ergebnisvariable aufgefasst und behandelt wird. Wie hat der Autor seine Ergebnisvariablen operationalisiert?

- Operationalisierung der Ergebnisvariable „evaluiert“

Zur Definition von Evaluation wird auf Balzer zurückgegriffen: „Evaluation (ist) ein Prozess, „in dem nach zuvor festgelegten Zielen und explizit auf den Sachverhalt bezogenen und begründeten Kriterien ein Evaluationsgegenstand bewertet wird (...) um den Evaluationsgegenstand zu optimieren und zukünftiges Handeln zu unterstützen.“ (Balzer, S.16, zit. nach Korczak, 2012). Zudem weisen die Autoren auf eine eigene Konzeption des Evaluationsprozesses hin, die letztendlich zur Einteilung in Struktur-, Prozess- und Ergebnisevaluation führt. Damit werden wesentliche Tätigkeitsfelder der Evaluation angesprochen, allerdings fehlt die Konzeptevaluation oder auch präformative Evaluation, die u.a. ein evidenzinformiertes Vorgehen der Praxis bei der Entwicklung einer Maßnahme überprüft (s. z.B. Mittag, 2006; Uhl, 1998).

Der HTA-Bericht versteht in seinen Schlussfolgerungen somit unter „evaluiert“ nur ergebnisevaluiert oder „wirkungsevaluiert“ (Korczak, 2012, S. 67). Ergebnisevaluationen seien „systematische Untersuchungen, die empirisch gewonnene Information bereitstellen, so dass die Wirkung einer Aktivität, einer Aktion oder eines Programm(ablaufs) nachvollziehbar beurteilt werden kann“ (Korczak, 2012, S.8).

- Operationalisierung der Ergebnisvariable „wirksam“

Als eigentliche Aufgabe des Addendums wird die Erfassung und Messung der Resultate und Wirksamkeit von präventiven Maßnahmen und ihre Bewertung hinsichtlich der vorher festgelegten Zielparameter bezeichnet. Als „aussagekräftige“ (Korczak, 2012, S. 68) Zielparameter werden - ohne Referenz und ohne Differenzierung nach Altersgruppe - genannt die Reduktion

- der Alkoholtrinkhäufigkeit
- der Alkoholtrinkmenge
- der Rauschhäufigkeit
- der Alkoholintoxikationen
- des riskanten Alkoholkonsums.

Maßnahmen, die Einstellungsänderungen feststellen, die nicht im Rahmen des Projekts zur Verhaltensänderung führen, werden als in ihrer Wirksamkeit als fragwürdig eingeschätzt. Eine genauere Definition der „Trinkeinstellungen“ wie sie im Fragebogen genannt werden und die sehr Verschiedenes (z.B. Wertungen, Wirkerwartungen, Trinkmotive) beinhalten können wird dem Anschein nach weder dem Befragten noch dem Leser des HTA-Berichts gegeben (Frage 16, s. Kasten 1).

Kasten 1: Item zur Messung der Evaluationsform und -parameter

16.1 Wenn „ja“ oder „geplant“: In welcher Form ist die Evaluation vorgenommen worden/ geplant?

<input type="checkbox"/> Erfassung der Teilnehmerzahlen	<input type="checkbox"/> nach Alter	<input type="checkbox"/> nach Geschlecht
<input type="checkbox"/> Erfassung der Abbrecherquoten	<input type="checkbox"/> zu Beginn	<input type="checkbox"/> am Ende der Maßnahme
<input type="checkbox"/> Erfassung der Trinkhäufigkeit	<input type="checkbox"/> zu Beginn	<input type="checkbox"/> am Ende der Maßnahme
<input type="checkbox"/> Erfassung der Trinkmenge	<input type="checkbox"/> zu Beginn	<input type="checkbox"/> am Ende der Maßnahme
<input type="checkbox"/> Erfassung des Risikokonsums	<input type="checkbox"/> zu Beginn	<input type="checkbox"/> am Ende der Maßnahme
<input type="checkbox"/> Erfassung von Trinkeinstellungen	<input type="checkbox"/> zu Beginn	<input type="checkbox"/> am Ende der Maßnahme
<input type="checkbox"/> Sonstige Kriterien, und zwar: _____		

Damit wird die Beurteilung von Präventionsmaßnahmen auf ihre Fähigkeit beschränkt, jedwedes Alkoholkonsumverhalten zu beeinflussen. Die Veränderung von Risiko- und Schutzfaktoren des (riskanten) Konsums oder der Abhängigkeitsentwicklung gelten nicht als Wirksamkeitsparameter. Diese Parameter werden ohne Referenz oder Begründung gewählt.

Die von den Einrichtungen genannten sehr „umfassend(en)“ (Korczak, 2012, S.62) Ziele der berichteten APM werden trotz der zuvor verwendeten, sehr breiten Definition von Evaluation ebenfalls nicht als Wirksamkeitsparameter aufgefasst.

- Operationalisierung der Ergebnisvariable „inhaltlich evidenzorientiert begründet“

Es wird zwar im Fragebogen mit einem Item der inhaltliche Ansatzpunkt der APM abgefragt, das zwar nicht dem Stand der Forschung (z.B. Ennett et al., 2003) entspricht, mit Hilfe dessen man aber versuchen könnte einzuschätzen, ob sich die Inhalte der APM an Ergebnissen der internationalen Wirksamkeitsforschung orientieren (Frage 11, s. Kasten 2). Die Auswertung dieses Items wird aber nicht berichtet.

Kasten 2: Item zur Messung der Inhalte einer APM

11. Wie ist das Projekt ausgestaltet?

Selbstwert- stärkung	Wettbewerb/ Gewinnspiel	Kompetenz- förderung	Alternative Erlebnisse	Informations- vermittlung	Bildung kritischer Einstellungen/ Normenbildung	Sonstiges
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Aus den Ausführungen in den letztendlichen Schlussfolgerungen (Kapitel 7, 6. Absatz) ist keine stimmige Definition und Operationalisierung zu diesem Aspekt zu finden (s.a. Punkt 4.4 der Stellungnahme).

3.3 Ergebnisse des HTA-Berichts

Mittels der bisher beschriebenen Methode und ihren Schwächen kommt der HTA-Bericht zu folgenden Ergebnissen:

- Anteil evaluierter Projekte in der Stichprobe: 98 der berichteten 208 APM (47%) werden nach Auskunft der Befragten in irgendeiner Form evaluiert, bei weiteren 28 APM (13%) ist die Evaluation in Planung.
- Verteilung der Outcomeparameter bei Projekten mit Evaluationsinformation: Von 125 ausgewerteten APM geben alle an, sie erfassten die Teilnehmerzahlen, nur 13,6% erfassten die Abbrecherquote. Der Anteil der APM mit Evaluationsinformation, der über eine Erfassung des Alkoholkonsums und der Trinkeinstellungen berichtet, liegt zwischen 9 und 18%.
- Anteil Prozess- und Ergebnisevaluation: Aus den „eingereichten bzw. beigefügten Unterlagen“ ersehen die Autoren bei 17 APM, ob es sich um eine Prozess- oder Ergebnisevaluation handelt und bestimmen einen Anteil von 2,9% prozessevaluierten und 5,3% ergebnisevaluierten APM in der Stichprobe.

Als weiteres Ergebnis wird eine Übersicht über die 208 berichteten Projekte zur Verfügung gestellt. Jede APM wird beschrieben hinsichtlich Träger, Projekttitle, Laufzeit, Finanzierung, Kooperationspartner, Alter der Zielgruppe, Setting und Evaluation. Es ist nicht dargestellt, welche spezifischen Ziele die jeweiligen APM verfolgen. Diese werden in der Zusammenfassung für die gesamte Stichprobe genannt: Förderung von Gesundheits- und Lebenskompetenzen, Vorbeugung von Sucht und Gewalt, Vermittlung von Wissen, Reduzierung der psychischen Belastung, Stärkung der Handlungsfähigkeit und verantwortungsbewusstem Konsum- und Genussverhalten, Förderung der Veränderungsbereitschaft, Änderung der Einstellung, Verzögerung des Einsteigalters, Reduzierung des exzessiven Alkoholkonsums sowie Vorbildverhalten im Umgang mit Alkohol“. (Korczak, 2012, S. 62). In der Übersicht fehlt die Information, ob es sich jeweils um eine universelle, selektive oder indizierte Maßnahme handelt. Der Autor weist im Text darauf hin, dass es sich „überwiegend um universelle Präventionsprojekte“ handelt (Korczak, 2012, S.18). Zudem stellen sie fest, dass die „Heterogenität der Träger, der Zielstellung und der Evaluation der Präventionsmaßnahmen [...] deutlich sichtbar“ wird (Korczak, 2012, S. 18).

Schließlich werden die 17 APM ausführlicher beschrieben, zu denen prozess- oder ergebnisevaluative Angaben extrahiert werden konnten, oder die eine geplante Evaluation berichten (n=3). Für diese Projekte liegt auch jeweils die Einschätzung vor, ob es sich um eine universelle, selektive oder indizierte Maßnahme handelt und es werden die anscheinend selbst formulierten Ziele der Anbieter aufgelistet.

Die Autoren beantworten die Forschungsfrage „Zu welchen APM liegen in Deutschland Evaluationen vor?“ zusammenfassend: „Die weit überwiegende Anzahl der gemeldeten APP ist somit nicht evaluiert oder zumindest nicht wirkungsevaluiert“ (Korczak, 2012, S.67). Zur

Forschungsfrage „Welche APM in Deutschland sind wirkungsvoll?“ lässt sich aus Sicht der Autoren „festhalten, dass eine wirksame und nachhaltige Verhaltensänderung des Risikoalkoholkonsums durch die Präventionsprogramme anhand der vorgelegten Evaluationsstudien nicht zu beobachten ist“ (Korczak, 2012, S.67).

3.4 Schlussfolgerungen des HTA-Berichts

Die Schlussfolgerungen des HTA-Berichts lassen sich anhand der Ausführungen und Zitate im Kapitel 6 „Diskussion und Beantwortung der Forschungsfragen“ und Kapitel 7 „Schlussfolgerung/Empfehlung“ wie folgt pointieren:

1. APM in Deutschland sind unzureichend/nicht evaluiert

„Die Primärerhebung zur Alkoholpräventionslandschaft in Deutschland zeigt, dass nur 5,3% der Alkoholpräventionsprojekte evaluiert sind.“ und „Es werden „in Deutschland weitestgehend nicht evaluierte Alkoholpräventionsprojekte eingesetzt“ (Korczak, 2012, S.68)

2. APM in Deutschland sind schlecht evaluiert

„Ergebnisevaluationen werden nur in sehr geringem Umfang durchgeführt.“ (Korczak, 2012, S.66)

„Wirksamkeitsprüfungen sind das entscheidende Instrument für die Bewertung von Präventionsmaßnahmen“ (Korczak, 2012, S. 66)

„Wirksamkeitsprüfungen haben eine ungleich höhere Bedeutung als Struktur- und Prozessevaluationen“ (Korczak, 2012, S.67)

3. APM in Deutschland sind nicht wirksam

„Insgesamt lässt sich festhalten, dass eine wirksame und nachhaltige Verhaltensänderung des Risikoalkoholkonsums durch die Präventionsprojekte anhand der vorgelegten Evaluationsstudien nicht zu beobachten ist.“ (Korczak, 2012, S.67)

„Derzeit kann nur für zwei Präventionsprojekte (1%) eine gesicherte Evidenz der Wirkung belegt werden.“ (Korczak, 2012, S.68)

4. APM in Deutschland sind inhaltlich nicht ausreichend/evidenzorientiert begründet

„Die Begründung für die Durchführung von APP ... stützt sich in der Regel auf eine einzige Studie“ (Korczak, 2012, S.68)

„Alle Projekte der universellen Prävention unterliegen der Limitation, dass mit wenigen Kontakten eine Sozialisation korrigiert werden soll, die am Entstehen und Aufrechterhalten gesundheitsriskanten Verhaltens beteiligt ist. Babor et al weisen nach einer umfangreichen Analyse der internationalen Studienlage darauf hin, dass schulische Erziehungspräventionsprogramme allein zur Korrektur eines riskanten Alkoholkonsumverhaltens nicht ausreichen“ (Korczak, 2012, S.68)

Darauf aufbauend ergibt sich für die Autoren die „Notwendigkeit einer grundsätzlich neuen Ausrichtung und Umsetzung der Präventionsforschung und des Einsatzes von Präventionsmaßnahmen zum Alkoholmissbrauch von Kindern, Jugendlichen und jungen Erwachsenen“ (Korczak, 2012, S. 68).

4. Stellungnahme

Die Stellungnahme erfolgt zuerst einzeln zu den vier Schlussfolgerungen und dann zusammenfassend.

4.1 APM in Deutschland sind unzureichend/nicht evaluiert

Grundsätzlich sind anhand der Daten, auf denen der HTA-Bericht aufbaut, keine Generalisierungen auf die Lage in Deutschland und keine Quantifizierungen zulässig. Dies liegt vor allem an dem geringen Rücklauf der ausgesendeten Fragebogen, der schiefen Verteilung der berichteten APM über Deutschland mit einer Überrepräsentation Bayerns und der Definition des Evaluationsgegenstands „Alkoholpräventionsprojekt“, die ein „Äpfel-Birnen-Problem“ mit sich bringt. Der Autor selbst beurteilt zu Beginn seine Stichprobe kritisch und gibt an, „nicht den Anspruch repräsentativ für das Projektgeschehen in Deutschland zu sein“ zu erheben (Korczak, 2012, S.17). Es ist also methodisch nicht korrekt, quantifizierte und deutschlandweite Schlussfolgerungen zu ziehen. Die Bemerkung „Rücklaufquoten von mehr als 15% bei schriftlichen Befragungen gelten als bemerkenswert hoch“ (Korczak, 2012, S.65) kann dabei deutlich machen, wie schwierig eine solche Befragung umzusetzen ist, kann aber nicht zur Aufwertung der Datenbasis beitragen.

Um die Unangemessenheit der Quantifizierung zu verdeutlichen: Zum Beispiel wurde das Programm Klasse2000 im Schuljahr 2011/12 bundesweit in 17.376 Klassen mit 400.000 Kindern durchgeführt (Auskunft des Trägers auf www.klasse2000.de/seiten/das_programm.php). Das Programm kommt aber nur fünf Mal in der HTA-Stichprobe vor. Laut Schulstatistik gab es 2011/2012 in Deutschland 2.790.138 Grundschülerinnen. Demnach hätten etwa 14% der deutschen Grundschüler an einem evaluierten, evidenz-basierten Programm teilgenommen.

Unabhängig von den für diese Frage nicht sehr hilfreichen Erhebungen des HTA-Berichts ist die Einschätzung, in Deutschland werden (Alkohol-) Präventionsmaßnahmen zu wenig evaluiert, weit verbreitet und wird –wie auch von Korczak ausgeführt - von Praxis, Politik und Forschung gleichermaßen geteilt. Dies ist auch meine Einschätzung, allerdings bin ich wie noch deutlich

werden wird hinsichtlich der Art der Evaluation, die für eine wirkungsvollere Praxis nötig wäre, anderer Meinung als der Autor des HTA-Berichts.

2. APM in Deutschland sind schlecht evaluiert

Der HTA-Bericht stellt fest, dass sich zumindest für seine Stichprobe an APM belegen lässt, dass es sich bei den erfolgten oder geplanten „Evaluationen zu 80% um Struktur- und Prozessevaluationen handelt. Ergebnisevaluationen werden nur in sehr geringem Umfang durchgeführt.“ (Korczak, 2012, S. 66). Dabei hätten Wirksamkeitsprüfungen „eine ungleich höhere Bedeutung als Struktur- und Prozessevaluationen“ (Korczak, 2012, S.66), die „in der Regel nur die Frage beantworten wie das Präventionsprojekt unter Einbeziehung von Kosten hinsichtlich der Implementation, seiner Abläufe und der Durchführung optimiert werden kann“ (Korczak, 2012, S.66) Wirksamkeitsprüfungen seien „das entscheidende Instrument für die Bewertung von Präventionsmaßnahmen“, da durch sie die „grundsätzliche Eignung der Präventionsmaßnahme, die gewünschten Ergebnisse der Inzidenzreduktion auch erfüllen zu können, ermittelt“ werde (Korczak, 2012, S.66f).

Mit dieser Einschätzung steht der Autor nicht alleine da (es wird eine Arbeit renommierter deutscher Suchtforscher zitiert) und redet dem „Gespenst der Evidenzbasierung“ (Bödeker, 2007) das Wort. Evidenz-basierung wird in den Gesundheitswissenschaften sehr unterschiedlich definiert.² Im HTA-Bericht wird Evidenzbasierung nur den Programmen zugestanden, die eine randomisierte Evaluationsstudie mit Konsumvariablen als Ergebnisparameter aufweisen können (Korczak, 2012, S. 67). Arbeiten, die sich explizit mit der Evaluation von Prävention und Gesundheitsförderung beschäftigen, werden nicht diskutiert. In diesen Arbeiten wird deutlich, dass Ergebnisevaluation eine wichtige aber nur eine Form der Evaluation bei der Entwicklung und Umsetzung von Präventionsmaßnahmen (Mittag, 2006; Uhl, 1998; EMCDDA, 2011; Elkeles & Broeskamp-Stone, 2010; Walter et al., 2012) neben Konzept-, Struktur- und Prozessevaluation ist. „Als grundlegendes Qualitätskriterium ist von allen Gesundheits- und Präventionsprogrammen zu fordern dass sie sich hinsichtlich ihrer Ziele als wirksam und effektiv erweisen. Entsprechend gilt der empirische Wirksamkeitsnachweis als zentrale Aufgabe der Evaluationsforschung, auch wenn dies nur eine unter vielen Aufgaben im Rahmen von Evaluationen darstellt“ (Mittag, 2006, S.95).

² Auch ist Evidenzbasierung nicht gleich evidenz-basierte Praxis. So definiert die EMCDDA (2011): Drug prevention work is ‚evidence-based‘ if it is based upon a systematic analysis of relevant professional literature (e.g. scientific journals), makes use of the evidence reported in the literature, and ensures correspondence with this evidence. ... When reviewing evidence of effectiveness, priority should be given to studies that represent the highest available level of evidence. ... The evidence from scientific research must inform practice, but it cannot replace the professional experience of practitioners“ (S.127)

Das sogenannte Präventionsdilemma (Präventionsangebote sind zwar effektiv, werden aber von der Zielgruppe nicht in Anspruch genommen; z.B. Bauer, 2005) führt vor Augen, dass es nicht zielführend ist, nur auf Ergebnisevaluation zu setzen und dass die Evaluation von Inanspruchnahme und Zielgruppenerreichung essentiell für die Breitenwirkung ist. Ein Programm muss nicht nur effektiv sein, sondern auch in den Strukturen umsetzbar und seine Zielgruppe erreichen.

Die Evaluation der Stimmigkeit des Maßnahmenkonzepts (Konzept- oder präformative Evaluation), wird als wichtiger Teil bei der Entwicklung von Präventionsmaßnahmen bzw. als Voraussetzung der Ergebnisevaluation beschrieben (Mittag, 2006; Uhl, 1998). Bei der Konzeptevaluation geht es um die Überprüfung des Wirkmodells. Das Wirkmodell ist eine anschauliche Darstellung der Frage: Was kann ich begründeterweise (begründet in Theorie, Empirie, Präventionsforschung, Praxiswissen) von meiner Maßnahme erwarten? Oder anders herum: Um realistischerweise einen Effekt auf den riskanten Konsum zu erwarten, was muss meine Maßnahme an Elementen beinhalten? So ein Wirkmodell wird auch vom Autor selbst eingefordert (Korczak, 2012, S.67). In seiner Operationalisierung von „Evaluation“ ist die Konzeptevaluation allerdings nicht inbegriffen und wird demnach auch nicht erhoben (s. Kasten 1). Somit ist auch nicht zu beurteilen, inwieweit diese grundlegende Art der Evaluation von den Trägern der APM geleistet wird. Die Wichtigkeit von Ergebnisevaluation ist damit nicht angezweifelt, nur ihre Überlegenheit für die Bestimmung der „grundsätzlichen Eignung einer Präventionsmaßnahme“.

Der Autor setzt eine *gute* Evaluation mit Studien gleich, die „Qualitätsstandards entsprechen (z.B. Randomisierung, Kontrollgruppen, Baseline-Messungen, Festlegung der Wirkungsparameter, Kontrolle von Confounder, Follow-Up-Erhebungen etc.)“ (Korczak, 2012, S. 66). Wessen Qualitätsstandards hier gemeint sind ist aus Ermangelung einer Referenz nicht nachvollziehbar. Die Standards der DGEval, auf die der Autor in der Einleitung Bezug nimmt, legen sich nicht auf diesen Studientypus fest. Im Gegenteil, um valide Ergebnisse und Schlussfolgerungen bei Evaluationen zu erhalten, „sollen mehrere – quantitative und/oder qualitative – Verfahren zur Datengewinnung verwendet werden“ (DGEval, 2008, S. 35).³

Die Rolle von RCTs wird meiner Ansicht nach – trotz ihrer unbestrittenen methodischen Vorzüge – überschätzt, wenn es darum geht, sie als alleinige/entscheidende Entscheidungsgrundlage für den Einsatz psychosozialer Präventions- und Gesundheitsförderungsmaßnahmen heranzuziehen. Die

³ Selbst die US-amerikanische Gesellschaft für Präventionsforschung stellt in ihren Standards of Evidence voran: „we believe that (a) no single method can be used to address all interesting and important questions about prevention and (b) even when causal relations are of primary interest, other types of research tools and designs are often needed to yield important information about when, why, and how interventions work, and for whom.“ (SPR, 2004)

kritischen Punkte der quantitativen Forschung, die „Grenzen des Wirksamkeitsnachweises“ (Uhl, 2012) allgemein und im Rahmen von Suchtprävention werden ausgeblendet (Stichworte: Komplexität und Komplexitätsreduktion, Heterogenität der Wirkung, Probleme der Messung der Intervention, des Zielkriteriums und der Mediatoren, Papageieffekt, Vagheit verbaler Ausdrücke, implizite Kommunikation, Selbstangaben sowie Langzeiteffekte, (Uhl, 2012)).

Auch sollten die Ansprüche an Evaluation differenziert werden. In den Europäischen Standards Drogenprävention (EMCDDA, 2011) werden „basic standards“ und „expert standards“ unterschieden, um der Vielfalt an Drogenprävention in Hinblick auf Zweck, Tätigkeitsbereich, Zielgruppe, Setting, Methode, Dauer, und/oder vorhandene oder nötige Ressourcen Rechnung zu tragen. Basic standards gelten für alle Aktivitäten. Expert standards, die eine qualitativ hochwertigere Evaluation beschreiben, sind zusätzliche Standards, die für ressourcenstärkere Akteure gelten sollen, wie wohl sie für alle wünschenswert sind. Unter den Verfassern der Standards wurde die Ansicht geäußert: „standards should be applied more rigorously to large-scale projects, while small-scale projects could not be expected to adhere to all standards“ (S.40). Diese Unterscheidung trifft der Autor nicht und hält damit den von ihm selbst zitierten Nützlichkeits-Standard der DeGEval nicht ein: „Der Aufwand für Evaluation soll in einem angemessenen Verhältnis zum Nutzen der Evaluation stehen“ (DeGEval, 2008, D3, S.29).

Die bisherige Diskussion betrifft die „Top-Down“-Präventionsstrategie (Lavereck, 2012). Präventionsverantwortliche entwickeln eine möglichst standardisierte Maßnahme, die dann in die Praxis transferiert und angewendet werden soll. Sie steht in Gegensatz zur „Bottom-Up“-Strategie, die eine partizipative Entwicklung einer Präventionsmaßnahme mit der konkreten Zielgruppe vor Ort verfolgt. Für diesen Ansatz wird das Top-down-Evaluationsinstrumentarium und sein Evidenzbegriff als mindestens „fragwürdig“ eingeschätzt (Elkeles & Broeskamp-Stone, 2010). Da aber auch für die partizipative Gesundheitsförderung Ziele und Endpunkte aufgestellt werden, sollte Ergebnisevaluation auch für diesen Ansatz möglich und nötig sein (Bödecker, 2007). RCTs spielen hier auch eine Rolle für die „wissensbasierte Praxis“ (Broeskamp-Stone, 2012), allerdings hängt ihr Einsatz von der Fragestellung ab und sie sind auf gleicher Höhe mit anderen Methoden im Evidenzprisma und nicht an oberster Stelle der Evidenzhierarchie angesiedelt (Elkeles & Broeskamp-Stone, 2010).

3. APM in Deutschland sind nicht wirksam

Gepaart mit der Forderung nach RCTs ist die Festlegung, dass als Wirksamkeitsparameter nur Konsumvariablen zählen. Damit folgt der Autor der Definition von „evidence-based prevention“ der US-amerikanischen Suchtpräventionsforscherin Pentz (2003) und den Evidenz-Standards der US-

amerikanischen Society for Prevention Research (SPR, 2004). So schlussfolgert der Autor, „dass eine wirksame und nachhaltige Verhaltensänderung des Risikoalkoholkonsums durch die Präventionsprojekte anhand der vorgelegten Evaluationsstudien nicht zu beobachten ist.“ (Korczak, 2012, S.67) und „Derzeit kann nur für zwei Präventionsprojekte (1%) eine gesicherte Evidenz der Wirkung belegt werden.“ (Korczak, 2012, S.68). Abgesehen von der bereits diskutierten Unangemessenheit der Quantifizierung und Generalisierung, möchte ich bei dieser Schlussfolgerung zu Bedenken geben, wie sinnvoll die Operationalisierung dieses Wirksamkeitsparameters ist.

Als „aussagefähige Parameter“ für die Wirksamkeit von APM gelten für den Autor „eine signifikante und deutliche Reduzierung des Alkoholkonsums, des riskanten Trinkens und des Rauschtrinkens“ (S.68). Allgemein gesagt, macht es natürlich Sinn, dass man mit Alkoholprävention das Konsumverhalten im Sinne der Vorbeugung eines Missbrauchs oder Abhängigkeit beeinflussen möchte und dieses zum letztendlichen Ziel aller Bemühungen festlegt. Dies aber als ausschließlichen Maßstab für die Wirksamkeit von einzelnen Präventionsmaßnahmen heranzuziehen, finde ich aus methodischen Gründen nicht gerechtfertigt und nicht sinnvoll, wenn Evaluation dazu da ist, „den Evaluationsgegenstand zu optimieren und zukünftiges Handeln zu unterstützen.“ (Balzer, S.16, zit. nach Korczak, 2012, S.7).

Konsum als alleinigen Wirksamkeitsparameter in der Suchtprävention anzusetzen ist aufgrund folgender Spezifika der Präventionspraxis und -forschung nicht sinnvoll:

- Präventionsstichproben: Konsum noch sehr instabil, Indikatoren (30-Tage) fragwürdig
In den Stichproben der Präventionsforschung liegen entweder noch keine Konsumerfahrungen oder sehr inkonsistente Konsummuster vor und damit sind die notwendigerweise groben Indikatoren für Alkoholmissbrauch fragwürdig (z.B. Lebenszeitprävalenz oder 30-Tage-Prävalenz: Hier heißt dann wirksam, dass ein Kind oder Jugendlicher noch nie oder kein Alkohol im letzten Monat getrunken hat, und unwirksam, wenn das Kind oder der Jugendliche es einmal getan hat)
- Messgenauigkeit von Alkoholkonsum fragwürdig (Schmutterer et al., 2007)
Auch in einer alkoholerfahreneren Altersgruppe gelingt die korrekte Beantwortung sorgfältig entwickelter, etablierter Alkoholkonsumitems in vielen Fällen nicht (Schmutterer et al., 2007). Dies mag für Studien, in denen die Prävalenz hoch ist als „Rauschen“ toleriert werden, in Präventionsstudien mit geringer Prävalenz aber entscheidend sein.
- Unspezifische Prävention im Kindesalter, die Risiko- und Schutzfaktoren beeinflusst, wäre dann „unwirksam“
Frühzeitige, alkoholunspezifische Präventionsansätze (Elterntrainings, Familienprogramme), die ausgewiesene Schutz- und Risikofaktoren des späteren Alkoholmissbrauchs beeinflussen (s.

Spoth et al., 2008) würden bei dieser Vorgehensweise als „nicht wirksam“ klassifiziert werden müssen⁴, wenn sie nicht in Studien mit einem Nachbefragungszeitraum von über 10 Jahren evaluiert werden, deren Validität dann aber auch wieder fraglich ist (s.u.).

- Einzelmaßnahmen mit fundiertem Wirkmodell, die realistischerweise „nur“ auf Beeinflussung einiger Risiko- oder Schutzfaktoren abzielen, wären dann „unwirksam“

Angenommen, eine konzeptuell stimmige APM, wie z.B. eine 15-minütige Peer education Maßnahme, setzt sich zum Ziel, die kritische Diskussion über Alkohol im Freundeskreis anzuregen und damit die sozialen Normen zu beeinflussen und erreicht dieses auch nachweislich im Vergleich zu einer Kontrollgruppe. Von der Maßnahme wird begründeterweise nicht erwartet, sie alleine könne missbräuchliches Konsumverhalten verhindern. Solch eine Maßnahme würde bei dem vom HTA-Bericht vorgeschlagenen Bewertungslinie als „nicht wirksam“ klassifiziert.

Warum finde ich Studien mit einem Nachbefragungszeitraum von 10 bis 20 Jahren, wie sie der Autor für verhaltensbezogene Präventionsmaßnahmen am Ende des HTA-Berichts als Vision für die Präventionsforschung empfiehlt, nur bei sehr differenzierter Studienanlage als sinnvoll und valide? Warum sollen meiner Meinung nach Präventionsprogramme, die einen Effekt auf begründete vermittelnde Schutz- und Risikofaktoren (Surrogatparameter) nachweisen können, nicht als „unwirksam“ eingestuft werden?

Dies liegt an der Komplexität der Entwicklung von Alkoholmissbrauch (Brown et al., 2008; Zucker et al., 2011), dessen psychosozialer Präventionsansatz sich meiner Ansicht nach nicht analog zum medizinisch physiologischen Präventionsansatz begreifen lässt. Gemeinhin stellt man sich vor, dass psychosoziale Prävention funktioniert wie die Schutzimpfung zur Vorbeugung von Infektionskrankheiten. Es bedarf *nur einer* richtigen Maßnahme, um in späteren Risikosituationen und -phasen für die Entwicklung eines Missbrauchs gewappnet zu sein. Als gäbe es eine ganz genau spezifizierbare Auslösesituation, eine Ursache, der man dann – durch die Präventionsmaßnahme vorbereitet - begegnen kann. Wie der Körper, der durch die Impfung Antikörper gebildet hat, die dann auf einen spezifischen Krankheitserreger mit einer Immunantwort reagieren. Gegen die Übertragung dieser Vorstellung auf psychosoziale Phänomene spricht zum einen die

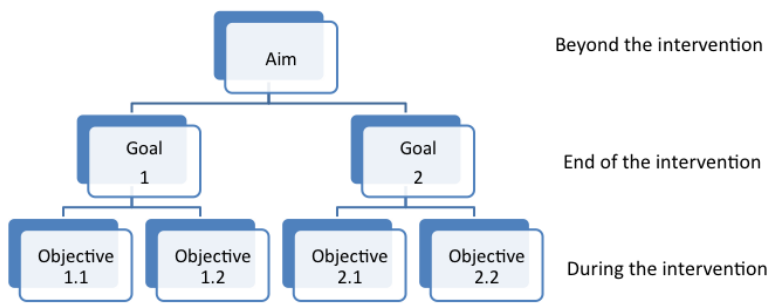
⁴ Hierzu auch die Europäischen Standards Drogenprävention (EMCDDA, 2011, S. 121): Aims, goals, and objectives need not relate directly to drug use. For example, goals and objectives may not relate to drug use if the programme hopes to prevent drug use that would potentially occur many years after the intervention (e.g. drug prevention activities with young children). Conversely, aims may address wider health or social issues (e.g. social inclusion projects), meaning that drug prevention might form only one aspect of the programme, expressed in a drug-related goal among other general goals. It also depends on the scope and duration of the individual drug prevention programme whether a certain target constitutes an aim, a goal or an objective. For example, behaviour change may be an overall aim for a short-term programme, but could represent a feasible objective within a long-term programme.

Multikausalität, von der man bei der Entwicklung des Substanzkonsums ausgehen muss. Zum anderen das Wissen, dass sich im Gegensatz zu den physiologischen die psychosozialen „Antikörper“ oder die „Immunantwort“, also die gebildeten Schutzfaktoren, sich im weiteren Lebensverlauf wieder verändern. Beispielsweise kann die Selbstwirksamkeit, Alkohol widerstehen zu können, durch nicht präventive Erfahrungen wieder verändert werden. Insofern erscheint es fraglich, die Wirkung von psychosozialen Maßnahmen in der gleichen Weise wie die Wirkung von physiologischen Vorsorgeinstrumenten zu überprüfen.

Um nicht missverstanden zu werden: Studien mit langfristigeren Nachbefragungszeiträumen wären unbestritten für die Wissensgenerierung zu möglichen Wirkmechanismen einer APM von erheblichem Vorteil. Als Beweis für die Effektivität einer APM, die in ihrer Intensität stark umschrieben ist (z.B. 4 Unterrichtsstunden), überzeugen sie allerdings nur, wenn über diese lange Zeit, in der viele andere Einflussfaktoren in viel stärkerer Intensität wirken, die nachhaltige Veränderung von begründeten, den Präventionserfolg vermittelnde Variablen auch überprüft wird (wie z.B. Spoth et al., 2009).

Die Komplexität der Entwicklung von Alkoholmissbrauch bringt es auch mit sich, dass man mit einer umschriebenen Maßnahme in einem Setting Risiko- und Schutzfaktoren beeinflussen kann, dass aber (viele) andere Faktoren in anderen Settings nicht angegangen werden können. Damit ist die Wahrscheinlichkeit, letztendlich das Alkoholmissbrauchsverhalten direkt zu verändern, geringer als wenn man viele unterschiedliche Settings und Maßnahmen umsetzt (Policy-Mix). Es stellt sich also die Frage, was realistischerweise von einer APM erwartet werden kann. Die Europäischen Qualitätsstandards Drogenprävention fordern, dass die unmittelbaren (objectives) und mittelbaren Ziele (goals), die während oder zum Ende einer Maßnahme angestrebt und überprüft werden sollten, u.a. spezifisch und realistisch sein sollen (s. Abbildung 1, EMCDDA, 2011, Standard 3.3.5, S.122). Anscheinend haben zumindest einige Projekte in der Stichprobe des HTA-Berichts sich Gedanken über realistische Ziele gemacht. Die Träger der APM berichten von Zielen ihrer Maßnahmen wie der Förderung von Gesundheits- und Lebenskompetenzen, der Reduzierung von psychischer Belastung, der Förderung der Veränderungsbereitschaft etc.

Abbildung 1: Connection between aims, goals and objectives (aus EMCDDA, 2011)



Es lässt sich darüber streiten und kann nur eine Konsensusentscheidung sein, welche Nachweise erbracht werden müssen, um eine APM als „wirksam“ oder „unwirksam“ zu klassifizieren. Ich möchte mit diesen Überlegungen dem Trend zur Deutungshoheit der quantitativen Forschung, des RCT-Designs und der Konsumvariablen als Wirksamkeitsparameter in der Suchtprävention begegnen und die Diskussion fortführen, die bereits u.a. von der Sucht- und Drogenkommission beim Bundesministerium für Gesundheit und dem Kooperationsnetzwerk nachhaltige Präventionsforschung (KNP) (Walter et al., 2012) angestoßen wurde.⁵

4.4 APM in Deutschland sind nicht ausreichend/evidenzorientiert begründet

Ohne im Ergebnisteil Daten zur inhaltlichen Begründung und Ausrichtung der 208 berichteten Projekte darzustellen, kritisiert der Autor, dass sich die „Begründung für die Durchführung von APP ... in der Regel auf eine einzige Studie“ stütze. Dem ist entgegenzuhalten, dass es sich bei dieser Studie (Bühler und Kröger, 2006) um eine Zusammenfassung von 49 Übersichtsartikeln handelt, die wiederum die Information von einem Vielfachen an Einzelstudien zusammentragen.

Der Autor führt weiter an, dass andere Studien nicht berücksichtigt würden und die Ergebnisse der internationalen Studienlage (hier Babor et al., 2010) keine Berücksichtigung bei der Konzeption von APM fänden. Im Einzelnen wird argumentiert, dass fast alle APM universelle Projekte seien.

„Alle Projekte der universellen Prävention unterliegen der Limitation, dass mit wenigen Kontakten eine Sozialisation korrigiert werden soll, die am Entstehen und Aufrechterhalten gesundheitsriskanten Verhaltens beteiligt ist. Babor et al weisen nach einer umfangreichen Analyse der internationalen Studienlage darauf hin, dass schulische Erziehungspräventionsprogramme allein zur Korrektur eines riskanten Alkoholkonsumverhaltens nicht ausreichen: [...] Diese Ausgangssituation erfordert ein umfassendes Präventionskonzept, das nachhaltig

⁵ Auch in der eigenen Arbeit (Bühler & Kröger, 2006) wurde auf den Konsumparameter zurückgegriffen, weil er derzeit den kleinsten gemeinsamen Nenner der Wirksamkeitsforschung darstellt. Die Verwendung haben wir kritisch diskutiert und vorgeschlagen, den frühzeitigen Konsum als einen unter mehreren wichtigen Prädiktoren späteren Substanzmissbrauchs als Wirksamkeitskriterium für Präventionsmaßnahmen zu verstehen.

auf das Verhalten wirkt als auch die Verhältnisse im Blick hat (z.B. Durchsetzung des Jugendschutzgesetzes, regulatorische Eingriffe“ (S.68).

Universelle Prävention heißt nicht, mit wenigen Kontakten eine Sozialisation korrigieren zu wollen, die gesundheitsriskantes Verhalten begünstigt. Universelle Prävention heißt, seine Maßnahme an eine unausgelesene Zielgruppe mit Individuen von unterschiedlichem Risikoprofil zu richten (Mrzazek & Haggerty, 1994; National Academy of Sciences, 2009). Die Sozialisation positiv zu beeinflussen, das haben psychosoziale Ansätze im Sinn. Psychosoziale Maßnahmen sind vom Ansatz her grundsätzlich anders als kognitive Aufklärungsmaßnahmen, wie man den von Babor et al. verwendeten Begriff „Education“ eigentlich übersetzen müsste (s. hierzu auch Babor et al., 2012, S. 210). Dieser Abschnitt zeigt, wie eine ausbleibende oder unsaubere Operationalisierung von Präventionsinhalten und damit Klassifikation von Präventionsmaßnahmen zu un schlüssigen Folgerungen führt. Gleichzeitig stellt nämlich der Autor die Wirksamkeit von „aktion glasklar“ heraus, das dem Aufklärungsansatz zuzurechnen ist, sowie von „Klasse2000“, das aus aufklärenden und psychosozialen Elementen besteht.

Tatsächlich muss davon ausgegangen werden, dass rein kognitive, aufklärende Maßnahmen nicht ausreichen (Babor et al., 2010; Bühler & Kröger, 2006; Foxcroft & Tsertsvadze, 2011). Der neueste Cochrane Review zur schulbasierten Alkoholprävention (Foxcroft & Tsertsvadze, 2011) kommt zu dem Ergebnis, dass bestimmte psychosoziale Maßnahmen auf den Alkoholkonsum Effekte zeigen (zwei Lebenskompetenzprogramme, ein Verhaltensmanagementprogramm) und diese entwicklungsunterstützenden Programme im Vergleich zu sehr unterschiedlichen, alkoholspezifisch ausgerichteten Maßnahmen ein größeres Potenzial hätten („Overall, we conclude that the evidence supports certain generic prevention programs over alcohol-specific prevention programs.“ S. 14).

Inwiefern APM in Deutschland in ihren Inhalten nicht oder nicht ausreichend begründet oder gar entgegen wissenschaftlichen Empfehlungen ausgerichtet sind, lässt sich nicht einschätzen. Deutlich wird hier, dass für eine valide Einschätzung erst einmal bestehende Klassifikationen gleichsam Eingang in die Präventionsforschung, -politik- und -praxis finden müssen und bei Bedarf eine einheitliche Begrifflichkeit entwickelt werden müsste.

5. Zusammenfassende Kommentierung der Methode und Stellungnahme zu den Schlussfolgerungen

Der HTA- Bericht hat zu große methodische Mängel, um Schlussfolgerungen von der Tragweite einer grundsätzlichen Neuausrichtung der alkoholpräventiven Praxis und Forschung zu

rechtfertigen. Die Anlage der Studie, nämlich die Akteure im Feld zu befragen ist positiv hervorzuheben und auch die Identifizierung der Akteure scheint gelungen. Mängel sind dagegen (a) insbesondere der zu geringe Rücklauf (ohne dass Maßnahmen zur Erhöhung des Rücklaufs getroffen wurden), (b) das Äpfel und Birnen Problem bei der Definition einer APM, (c) teilweise nicht forschungsstandgemäße Items im Erhebungsinstrument, (d) die lückenhafte Operationalisierung der Ergebnisparameter „evaluiert“ und „wirksam“, (e) triviale Rechenfehler und Inkorrektheiten bei der Verwendung von etablierten Begriffen der Präventionsforschung, (f) mangelnde Differenzierung von Evaluationskontexten (Praxis, Forschung), (g) sich widersprechende Argumentation.

Vor diesem Hintergrund die „Notwendigkeit einer grundsätzlich neuen Ausrichtung und Umsetzung der Präventionsforschung und des Einsatzes von Präventionsmaßnahmen zum Alkoholmissbrauch von Kindern, Jugendlichen und jungen Erwachsenen“ zu konstatieren ist meiner Meinung nach nicht angemessen.

Trotz ihrer Unbegründetheit in den erhobenen Daten, geben die Schlussfolgerungen Anlass, wichtige Fragen der Evaluation in der Suchtprävention zu diskutieren und den derzeitigen Trend zu hinterfragen, darauf nur mit dem Zauberwort „Evidenzbasierung“ zu antworten.

Wie der Autor und viele andere vor ihm nehme ich auch wahr, dass in der Suchtprävention in Deutschland zu wenig evaluiert wird. Auch finde ich, dass mehr sinnvoll evaluiert werden sollte, lehne aber die Priorisierung von Ergebnisevaluation und Evidenzbasierung im engeren/medizinischen Sinne ab und finde deren Überhöhung kombiniert mit der Abwertung von Konzept-, Struktur- und Prozessevaluation für die Informierung von Politik und die Weiterentwicklung der Praxis nicht zielführend.

In den letzten Jahren beobachte ich (repräsentative Daten habe ich dafür keine)

- eine verstärkte Evidenzorientierung bei der Entwicklung von Präventionsmaßnahmen im Sinne von expliziteren Zielsetzungen, realistischeren Zielsetzungen, mehr begründeten/evidenzinformierten Wirkmodellen
- eine verstärkte Thematisierung von „evidenzinformierter Arbeit in der Suchtprävention“ in Fortbildungen und Tagungen (z.B. bundesweite Fachtagungen zur Qualität der Suchtprävention, DHS Konferenzen)
- und weise auf die hochwertigen Evaluationsstudien (zu Trampolin, Familien stärken, HaLT) hin, die für flächendeckend einzuführende verhaltensbezogene Maßnahmen in den letzten Jahren begonnen wurden, die ich als EINE wichtige Informationsquelle für evidenzorientierte Praxis und die Überprüfung der Präventionsmodelle einordne.

Somit kann ich die Einschätzung der „Notwendigkeit für eine grundsätzlich neue Ausrichtung und Umsetzung der Präventionsforschung und des Einsatzes von Präventionsmaßnahmen“ im Sinne der hier vertretenen Richtung der Evidenzbasierung nicht unterstützen und sehe in dieser Haltung in seiner Ausschließlichkeit eher ein Bremsklotz als eine Hilfe für die Weiterentwicklung der Alkoholprävention in Deutschland.

6. Vorschlag zur Evaluation von Suchtprävention mit dem Ziel einer evidenz-basierten Praxis in Deutschland

Was wir über die Fehlerbehaftetheit der Forschung zu Substanzkonsum bei Jugendlichen wissen (Uhl, 2012), sollte nicht dazu führen, zu resignieren und die Hände in den Schoß zu legen. Aber eine selbstkritische Haltung in Hinblick auf die Aussagekräftigkeit auch qualitativ hochwertiger Forschung ist angebracht. Ihr Gewissheitsversprechen sollte einen nicht dazu verleiten, die Evidenzbasierung einer Maßnahme (nachweisbare Effekte auf Konsum in einer RCT-Studie) als *alleiniges* Kriterium für die politische Entscheidung der Förderung oder Nichtförderung von Maßnahmen heranzuziehen⁶. **Dies ist kein Freibrief für keine Evaluation, weil "man ja sowieso nichts sagen kann" oder qualitativ minderwertige Evaluation, die dann - mit noch weniger aussagekräftigeren Ergebnissen - den gleichen Anspruch erfüllen will.**

Das Kooperationsnetzwerk nachhaltige Präventionsforschung (KNP) und die Drogen- und Suchtkommission beim Bundesministerium für Gesundheit erkennen die „methodischen Schwierigkeiten und prinzipiellen Grenzen der Wirkevaluation von Prävention“ an und sehen einen „dringenden Bedarf an praktikablen Studiendesigns, aussagekräftigen Instrumentarien sowie an der Identifikation angemessener Zielgrößen auf der Prozess- und Ergebnisebene“ (Walter et al.,

⁶ Siehe auch Broeskamp-Stone (2012): Eine einseitige Fokussierung auf Evidenzbasierung als dem allgemeinen Qualitäts- oder „Muss“-Kriterium für „akzeptables“ Entscheiden und Handeln, für optimale Praxis und Politik im Handlungsfeld Gesundheitsförderung und darüber hinaus, wird seit Jahren durch die zunehmende Verwendung des Begriffs „evidenzinformiert“ statt „evidenzbasiert“ relativiert (anfänglich im Bereich [Healthy Public] Policy, später auch in anderen Handlungsbereichen von Gesundheitsförderung und Public Health, z.B. bei wichtigen Entscheidungsprozessen). Das bedeutet nicht die Ablehnung der Forderung nach Evidenzbasierung, sondern die Anerkennung dessen, dass Evidenzbasierung zumeist nicht das alleinige Maß ist, auf dem optimale Entscheidungen und Maßnahmen im Bereich Gesundheitsförderung beruhen. Wie neuere Best-Practice-Ansätze reflektieren, ist „evidenzbasiert“ nur eines von mehreren wichtigen Kriterien für Qualität und nachhaltige Wirksamkeit von Interventionen.

Interessanterweise hat die amerikanische Suchtpräventionsbehörde SAMHSA, die früher einzelne Maßnahmen als model, effective oder promising programs eingestuft hat, diese Einstufung 2004 aufgegeben. Im National Registry of Evidence-based Programs and Practices werden die Maßnahmen jetzt auf mehreren methodischen Dimensionen mit einer kontinuierlichen Einschätzung (1-4) beurteilt, eine kategoriale Empfehlung gibt es nicht mehr.

2012). Abschließend soll ein Vorschlag für diese Diskussion formuliert werden, der darauf eingeht, wo das wissenschaftliche Wissen in der Evidenz-basierten Praxis zu verorten ist und wie Evaluation ausgestaltet sein sollte, um tatsächlich „...den Evaluationsgegenstand zu optimieren und zukünftiges Handeln zu unterstützen.“ (Balzer, S.16, zit. nach Korczak, 2012).

Wissenschaftliche Ergebnisse können im Rahmen dessen genutzt werden, was als evidenzinformierte, evidenzbasierte Praxis auch für den Bereich Prävention und Gesundheitsförderung auf Basis internationaler Arbeiten und Diskurse von der Gesundheitsförderung Schweiz beschrieben worden ist (Elkeles & Broesskamp-Stone, 2010). Gleich einem Radarstrahl müssen bei der Planung und Umsetzung von gesundheitsfördernden und präventiven Aktivitäten drei Dimensionen wiederholt reflektiert werden (s. Abbildung 2):

- Werte: (ethische) Grundwerte und Prinzipien der Gesundheitsförderung und Public Health wie z.B. Schadensvermeidung, Respekt vor Autonomie, gesundheitliche Chancengleichheit, Nachhaltigkeit, Empowerment
- Wissen: systematisch recherchierte aktuelle wissenschaftliche Erkenntnisse, eigenes Evaluationswissen und Reflektion des Nichtwissens, ergänzend Erfahrungs-, Praxis und Expertenwissen
- und Kontext: vorhandene Kapazitäten für Maßnahmen, Gesetze, politische/kulturelle Faktoren und ihre Bedeutung für die Übertragbarkeit von wissenschaftlichen Erkenntnissen und Interventionen auf den eigenen Kontext

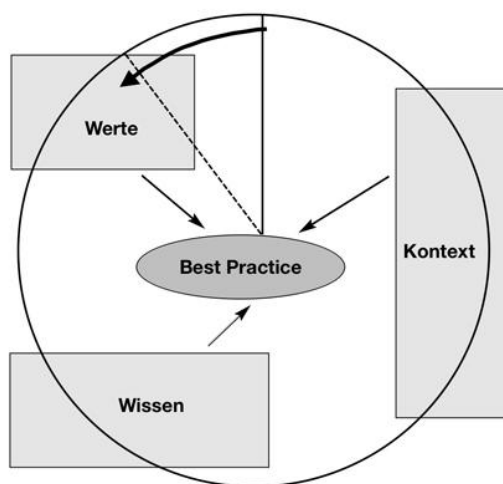


Abbildung 2: Radarschirm-Modell Best Practice (aus Broesskamp-Stone, 2012)

In dieser Auffassung von evidenz-basierter, oder besser „evidenzinformierter“ Gesundheitsförderung und Prävention ist das wissenschaftliche Wissen ein zentrales, aber nicht einziges Kriterium, das bei der Planung und Umsetzung von Suchtprävention beachtet werden muss.

Wie kommen wir an wissenschaftliche Ergebnisse? Wissenschaftliche Ergebnisse setzen sich aus Resultaten verschiedener methodischer Herangehensweisen an das Phänomen (hier den Substanzmissbrauch und seine Vorbeugung) zusammen und münden letztlich in einer Modellvorstellung von dem, was man tun sollte, um Substanzmissbrauch vorzubeugen (Wirkmodell)⁷. Empirische Daten können - dem wissenschaftstheoretischen Falsifikationsprinzip von Karl Popper nach - dieses Wirkmodell nur falsifizieren. Sprechen die Ergebnisse einer Wirksamkeitsstudie gegen das Wirkmodell, muss es verworfen oder angepasst werden. Stehen die Ergebnisse im Einklang mit dem Wirkmodell darf man an der Wirktheorie und seiner Umsetzung in einer APM festhalten. Nach diesem Prinzip können empirisch gewonnene Daten keinesfalls als „die Wahrheit“ interpretiert werden, weil es in der Empirie immer methodische Schwächen gibt. So muss auch in der Wirksamkeitsforschung das Wirkmodell im Mittelpunkt stehen. Wirksamkeitsstudien sind dann als Studien zu verstehen, die das Wirkmodell überprüfen und nicht positivistisch „die Wirksamkeit“ bestätigen.

Konkret sehe ich in den Europäischen Qualitätsstandards Drogenprävention (EMCDDA, 2011) ein sinnvolles Konzept für den nächsten Schritt in der Evaluationsarbeit der Alkoholprävention in Deutschland. Es macht Sinn grundlegende Evaluationsstandards für alle Präventionsaktivitäten und zusätzliche Standards für groß angelegte Programme vorzugeben.

- Entscheidend für jede Maßnahme ist aus meiner Sicht das evidenzinformierte, begründete Wirkmodell (Konzeptevaluation). Eine Dokumentation sollte Aufschluss darüber geben, ob die anvisierte Zielgruppe erreicht wird und ob das Projekt wie geplant umgesetzt werden kann (Prozess- und Strukturrevaluation).
- Als Ergebnisevaluation würde ich bei einem Projekt mit knappen finanziellen Mitteln und wenigen Evaluationskompetenzen eine initiale Studie akzeptieren, die eine Veränderung der Risiko- und Schutzfaktoren überprüft, welche sich in bisheriger Forschung als relevant für ein späteres Missbrauchsverhalten erwiesen haben und realistischerweise von dem Projekt beeinflusst werden können.
- Flächendeckend einzuführende oder eingeführte, verhaltensbezogene Maßnahmen und Programme sollten in Anlehnung an das 6-Phasen-Modell der Entwicklung einer Präventionsprogramms (Uhl, 1998) entwickelt und evaluiert werden (Phase 3 bis Phase 6, s. Tabelle 1 im Anhang). In dessen „Testungsphase“ sind (möglichst randomisiert)

⁷ Ähnlich auch Elkeles & Broesskamp-Stone (2010): „Insgesamt kommt es auf allen Gebieten der Gesundheitsförderung darauf an, die Evidenzbasis mittels wissenschaftlicher Evaluation einerseits und Qualitätsentwicklung andererseits zu verbessern. Auf der Grundlage des gegebenen wissenschaftlichen Kenntnisstandes sollte eine plausible Abschätzung vorgenommen werden, ob mit den vorgeschlagenen Interventionen positive Effekte in einem grob abschätzbaren Umfang erreichbar erscheinen. Dies umfasst in der Regel die Skizzierung eines Wirkungsmodells mit zu testenden Hypothesen.“

kontrollierte Interventionsstudien mit ausreichender statistischer Power vorgesehen. Ergebnisparameter sind hier begründete, relevante, realistische Risiko- und Schutzfaktoren (darunter auch früher Konsum) und im Jugendalter auch riskantes Konsumverhalten unter Berücksichtigung potenzieller Störgrößen, die dem komplexen Entstehungsgeflecht von Substanzmissbrauch Rechnung tragen.

- Auch oder gerade für die Evaluation von sogenannten „komplexen Interventionen“ im Sinne von einer „Vielzahl von Interventionskomponenten, Akteuren oder Kontexten“ (Bödeker, 2012, S. 37) meist auf kommunaler oder Bundesebene ist das Wirkmodell essentiell. Ergebnisevaluationsstudien der Teilkomponenten können sich auf die mittelbaren Schutz- und Risikofaktoren als Wirkparameter stützen, Endpunkte müssen über Surveys auf kommunaler oder Bundesebene evaluiert werden. Die Interaktionen zwischen den Teilkomponenten mit vertretbarem Aufwand (DGEval, 2008) anhand quantitativer Methoden überprüfen zu können, sehe ich nicht. Hier wären sicher qualitative Verfahren hilfreich, die hypothesengeleitet aus Narrativen der Zielgruppe die Wechselwirkungen zwischen den Wirkungen der Teilkomponenten abbilden und überprüfen.

Fazit

Die Prävention bringt eine Neuausrichtung der Forschung und Praxis, die neben einer sehr begrenzt verstandenen Evidenzbasierung nichts gelten lässt, nicht weiter. Das menschliche Bedürfnis nach Gewissheit und simplen Entscheidungsgrundlagen sollte zwar antreiben aber nicht blind treiben, der empirische Blick kann auch in Bezug auf die Wirksamkeit von Interventionen nur unscharf sein. Immer weniger unscharf kann er werden, indem mit verschiedenen Sehhilfen, also Methoden der Erkenntnisgewinnung darauf geblickt wird. Dank des bereits angesammelten internationalen Forschungs- und nationalen Praxiswissens muss auf der Suche nach wirkungsvollen Ansätzen nicht mehr im Dunkeln gestochert werden, können Wissensstücke zusammengesetzt werden und dieses Ergebnis der Politik und den Entscheidern als Arbeitshypothesen zur Verfügung gestellt werden. Diese Arbeitshypothesen koordiniert zu verfeinern mit Hilfe von Evaluation in Forschung *und* Praxis sollte die weitere Ausrichtung der Präventionsforschung sein.

Literatur

- Babor, T. et al. (2010). *Alcohol: No ordinary Commodity – Research and Public policy*. 2nd edition. Oxford:
- Bauer, U. (2005). *Das Präventionsdilemma. Schulische Kompetenzförderung im Spiegel sozialer Polarisierung*. Wiesbaden: VS
- Bödeker, W. (2007) Evidenzbasierung in Gesundheitsförderung und Prävention. *Prävention extra*. www.dnbgf.de/fileadmin/texte/.../Praevention_extra_0703.pdf
- Bödeker, W. (2012). Wirkungen und Wirkungsnachweis bei komplexen Interventionen. In RKI & BayLGL (Hrsg.). *Evaluation komplexer Interventionsprogramme in der Prävention: Lernende Systeme, lehrreiche Systeme?* (S.33-39). http://www.rki.de/DE/Content/Gesundheitsmonitoring/Gesundheitsberichterstattung/GBEDownloadsB/evaluation_praevention.pdf?__blob=publicationFile
- Broeskamp-Stone, U. (2012). Evidenz- und was noch? Ein Qualitätsrahmen für „Best Practice“ in der Gesundheitsförderung. In BZgA (Hrsg.): *Gesund aufwachsen in Kita, Schule, Familie und Quartier. Nutzen und Praxis verhaltens- und verhältnisbezogener Prävention – KNP-Tagung am 18. Und 19. Mai 2011 in Bonn* (S. 33-46). Bd 41 Forschung und Praxis der Gesundheitsförderung. Köln: BZgA.
- Brown, S.A., McGue, M., Maggs, J., Schulenberg, J., Hingson, R., Swartzwelder, S., Martin, C., Chung, T., Tapert, S.F., Sher, K., Winters, K.C., Lowman, C., & Murphy, S. (2008). A developmental perspective on alcohol and youths 16 to 20 years of age. *Pediatrics*, 121 (4), S290-S310.
- Bühler, A. & Kröger, C. (2006). *Expertise zur Prävention des Substanzmissbrauchs* (Forschung und Praxis der Gesundheitsförderung Band 29). Köln: Bundeszentrale für gesundheitliche Aufklärung.
- Center for Substance Abuse Prevention (CSAP) (2009). Identifying and Selecting Evidence-Based Interventions. <http://store.samhsa.gov/shin/content/SMA09-4205/SMA09-4205.pdf>
- DeGEval-Gesellschaft für Evaluation e.V. (2008). *Standards für Evaluation*. <http://www.degeval.de/degeval-standards>.
- Elkeles, T. & Broeskamp-Stone, U. (2010). *Evidenzbasierte Gesundheitsförderung . Leitbegriffe der Gesundheitsförderung BZgA*. <http://www.leitbegriffe.bzga.de/?uid=3404d2fc77f4eadcc4a9ddce020ebf4e&id=angebote&idx=170>
- EMCDDA (2011). *European drug prevention quality standards. A manual for prevention professionals*. EMCDDA Manuals 7. Lisbon: EMCDDA.
- Ennett, S., Ringwalt, C., Thorne, J., Rohrbach, L., Vincus, A., Simons-Rudolph, A. & Jones, S. (2003). A comparison of current practice in school-based substance use prevention programs with meta-analysis findings. *Prevention Science*, 4 (1), 1-14.
- Foxcroft, D. & Tsertsvadze, A. (2011). Universal school-based prevention programs for alcohol misuse in young people. *Cochrane Database Syst Rev.*, 11;(5):CD009113. doi: 10.1002/14651858.CD009113
- Kraus, L., Hannemann, T., Pabst, A., Müller, S., Kronthaler, F., Grübl, A., Stürmer, M. & Wolstein, J. (2012). Stationäre Behandlung von Jugendlichen mit akuter Alkoholintoxikation: Die Spitze des Eisbergs? *Das Gesundheitswesen*. Aug 29. [Epub ahead of print]
- Laverack, G. (2012). Bottom-up und Top-down-Programme: Plädoyer für eine Strategie der parallelen Zielverfolgung (parallel tracking). In BZgA (Hrsg.): *Gesund aufwachsen in Kita, Schule, Familie und Quartier. Nutzen und Praxis verhaltens- und verhältnisbezogener Prävention – KNP-Tagung am 18. Und 19. Mai 2011 in Bonn* (S. 33-46). Bd 41 Forschung und Praxis der Gesundheitsförderung. Köln: BZgA.
- Mrazek, P. J. & Haggerty, R. J. (1994). *Reducing risks for mental disorders*. Washington DC: National Academy Press.
- Mittag, W. (2006). Evaluation von Gesundheitsförderungsmaßnahmen. In A. Lohaus, M- Jerusalem & J. Klein-Heßling (Hrsg.), *Gesundheitsförderung im Kindes- und Jugendalter*. S. 87-112. Göttingen: Hogrefe.

Pentz, M. (2003). Evidence-based prevention: characteristics, impact, and future direction. *Journal of Psychoactive Drugs*, 35 Suppl 1,143-52.

Schmutterer, I. et al., 2007. *ESPAD – Austria 2007, Band2: Validierungsstudie*. Wien: BMG.

Society für Prevention Research (2004). *Standards of Evidence. Criteria for efficacy, effectiveness and dissemination*. www.preventionsresearch.org/StandardsOfEvidencebook.pdf

Spoth, R., Greenberg, M., & Turrisi, R. (2008). Preventive interventions addressing underage drinking: State of the evidence and steps toward public health impact. *Pediatrics*, 121, S311-S336.

Spoth, R., Trudeau, L., Gyll, M., Shin, C. & Redmond, C. (2009). Universal intervention effects on substance use among young adults mediated by delayed adolescent substance initiation. *J Consult Clin Psychol.*, 77(4), 620-32.

Uhl, A. (1998). 6 Phase Model In A. Springer & A. Uhl (Hrsg.), *Evaluation research in regard to primary prevention of drug abuse*. Brussels: European Commission Social Sciences. Downloadable at <http://www.api.or.at/sp/download.htm#1995>

Uhl, A. (2012). Methodenprobleme bei der Evaluation komplexer Sachverhalte: Das Beispiel Suchtprävention. In RKI & BayLGL (Hrsg.). *Evaluation komplexer Interventionsprogramme in der Prävention: Lernende Systeme, lehrreiche Systeme?* (S.57-79). http://www.rki.de/DE/Content/Gesundheitsmonitoring/Gesundheitsberichterstattung/GBEDownloadsB/evaluation_praevention.pdf?__blob=publicationFile

Walter et al., 2012. Kooperationsnetzwerk nachhaltige Präventionsforschung (KNP) Memorandum zur Präventionsforschung- Themenfelder und Methoden (Langfassung). *Gesundheitswesen*, 74, e99-e113.

Zucker, R.A., Donovan, J.E., Masten, A.S., Mattson, M.E., & Moss, H.B. (2011). Early developmental processes and the continuity of risk for underage drinking and problem drinking. *Pediatrics*, 121 (4), S252-272.

Tabelle 1: 6-Phasen-Modell der Entwicklung einer Präventionsprogramms (Uhl, 1998)

Phase 1: Grundlagenforschung	<p>Hilfreich für Präventionsarbeit, aber nicht direkt für Präventionsfragen unternommen</p> <ul style="list-style-type: none"> • Entwicklung valider und zuverlässiger Instrumente zur Messung von Konsum und relevanten Faktoren • Epidemiologische Studien zur Bestimmung von Problembereichen und Zielgruppen • Forschung zum Phänomen Substanzmissbrauch
Phase 2: Präventionsforschung	<p>Direkt mit Präventionsfragestellungen befasst</p> <ul style="list-style-type: none"> • Entwicklung und Überprüfung von Theorien und Modellen zur Entstehung, Veränderung, Reduktion oder Beendigung des Substanzkonsums • Bestimmung von (veränderbaren)* Risiko- und Schutzfaktoren • (Maßnahmenübergreifende Wirksamkeitsforschung)
Phase 3: Konzeptphase	<p>Entwicklung eines Maßnahmenkonzepts per Reflexion, am Ende vorläufiger Maßnahmenentwurf (mit kausalem oder Interventionsmodell)</p> <ul style="list-style-type: none"> • Ethische Überlegungen • (Kontextuelle Informationen) • Analyse bereits existierender Forschung: Sichtung und Beurteilung existierender Studien zu Grundlagen (Phase 1), Präventionsfragestellungen (Phase 2) und Evaluationen existierender Maßnahmen aus dem Bereich
Phase 4: Entwicklungsphase	<p>Bearbeitung des Konzepts und ersten Maßnahmenentwurfs, um eine endgültige Version zu erhalten</p> <ul style="list-style-type: none"> • Mängel identifizieren, Maßnahme kontinuierlich verbessern • Pilotstudie zur Machbarkeit, Wirksamkeit und unerwünschten Nebenwirkungen (sowie Zugang zur Zielgruppe) • Explorative Herangehensweise • Sollte ausreichend lange Phase sein: Flexibel, schnell und kostengünstige Informationen für Verbesserung • Erfolg ist Voraussetzung für das Eintreten in die Testungsphase
Phase 5: Testungsphase	<p>Bestätigung des Nutzens der Maßnahme : Überprüfung im Rahmen einer groß angelegten Studie unter kontrollierten Bedingungen</p> <ul style="list-style-type: none"> • Randomisierte oder nicht-randomisierte, kontrollierte Studie <ul style="list-style-type: none"> ○ Überprüfung in realer Welt ○ Machbarkeit, Umsetzbarkeit ○ (Zielgruppenerreichung) ○ Wirksamkeit, Effektivität ○ Nebenwirkungen • Bei finanziellen oder methodologischen Einschränkungen zumindest teilweise Überprüfung des kausalen Modells/Interventionsmodells und Bezugnahme auf existierende Evaluationsstudien
Phase 6: Routinephase	<p>Routineanwendung. Wirksamkeit demonstriert, deswegen Fokus auf</p> <ul style="list-style-type: none"> • Qualitätssicherung bei der Durchführung • Verwendung und Inanspruchnahme der Maßnahme • Verallgemeinerung der Befunde auf Zielpopulation • Eventuell sich verändernde Voraussetzungen und Rahmenbedingungen • Beobachtung unerwarteter Effekte • Kosten-Nutzen-Beurteilung • Anpassung/Optimierung der Maßnahme mit erneuter Testungsphase unabdingbar, wenn <ul style="list-style-type: none"> ○ Durchführung nicht adäquat, gar keine Durchführung mehr ○ Kontext sich fundamental geändert hat ○ Negative, langfristige Effekte

* In Klammern: Punkte zu Uhl (1998) hinzugefügt