

Bernd Röhrle

Die Forschungslage zur Prävention psychischer Störungen und Förderung psychischer Gesundheit

Der Ausbau der Prävention psychischer Störungen und die Förderung psychischer Gesundheit hängen, wie alle präventiven Maßnahmen, von Maßnahmen der Qualitätssicherung und damit auch vom Ausmaß und der Qualität der Evaluation einschlägiger Interventionen ab. Das findet sich nicht erst im zweiten Entwurf eines Gesetzes zur Stärkung der Gesundheitsförderung und gesundheitlicher Prävention sowie zur Änderung anderer Gesetze der Bundesregierung vom 23.11.2007, sondern schon in zahlreichen vorher erschienenen Verlautbarungen (vgl. z. B. *Christiansen, 1999*). Diese Vorgaben lenken den Blick auf die vorhandenen Effektivitätsnachweise auch bei Interventionen zur Prävention psychischer Störungen. Dabei werden gewohnte Analytika, wie die Kriterien einer evidenzbasierten Medizin (EBM), genutzt, die ohne Einschränkungen zu einer fragwürdigen Bewertung vorhandener Präventionsprogramme führen. Vor der Darstellung der entsprechenden Ergebnislage soll deshalb eine kritische Würdigung dieser Kriterien gemacht werden, damit die Effektivität und Effizienz dieser Programme angemessener eingeschätzt werden kann, als dies in ersten Gutachten, die gerade auch für weitere politische Maßnahmen verheerende Folgen haben könnten, der Fall ist.

Notwendige methodische Vorbemerkungen

Obgleich erkannt wird, dass die Methoden der evidenzbasierten Medizin (vgl. *Kunz, et al., 2007*) oder auch die Kriterien der Psychotherapieforschung nicht ohne weiteres auf die Bewertung von Maßnahmen der Prävention, insbe-

sondere psychischer Störungen und Förderung der (psychischen) Gesundheit, übertragen werden können, wird dies in einem neuen, von der Hans-Böckler-Stiftung in Auftrag gegebenen Gutachten mit dem Ergebnis gemacht, dass die bislang entwickelten Programme in ihrer Wirksamkeit eher als fraglich einzustufen sind (*Klever-Deichert, et al., 2007*).¹ Die zentralen Probleme bei der Übertragung der Kriterien der evidenzbasierten Medizin und auch der Psychotherapieforschung auf die Bewertung von Programmen zur Prävention psychischer Störungen und Förderung psychischer Gesundheit sind neben vielen weiteren Argumenten, wie die mangelnde Individualisierung der Aussagen, folgende:

1. Die Maßstäbe dieser Kriterien sind fraglich: Meta-Analysen zu randomisierten Studien sind in ihrer Bedeutung relativierbar². So berichten Meta-Analysen nur über auswertbare Studien. Dass schwer zugängliche Studien oder solche mit fehlenden statistischen Angaben nicht genutzt werden können, ist durch die Angabe sog. „Fail Safe Number“ (Hochrechnungen, wie viele nicht signifikante Studienergebnisse eine identifizierte Effektgröße zu Fall bringen können) einigermaßen kompensierbar. Auch die Tatsache, dass in Meta-Analysen Äpfel mit Birnen verglichen werden, ist durch die Berechnung sog. Zufallseffektmodelle (Annahmen, dass die untersuchten Studien nicht zu einer Grundgesamtheit gehören) auszugleichen. Gravierender sind wenigstens drei Aspekte: Zum Einen sind Meta-Analysen selbst von sehr unterschiedlicher methodischer Wertigkeit, zum Zweiten berichten sie nicht über komplexe Studien. Die-

Zusammenfassung

Es wird ein Überblick zu den vorliegenden evidenzbasierten Ergebnissen zu Studien der Prävention psychischer Störungen und zur Förderung psychischer Gesundheit geboten. Dies geschieht auf dem Hintergrund der in der evidenzbasierten Medizin üblichen Kriterien und deren Problematik für die Präventionsforschung. Der Beitrag schließt ab mit erkennbaren Trends in der Evaluation von entsprechenden Präventionsprogrammen.

Schlüsselworte

Prävention, psychische Krankheit, psychische Gesundheit, Gesundheitsförderung, Evaluation, Meta-Analysen

Abstract

Evidenced-based results of existing studies in mental health prevention and promotion are offered in a review. The review is based on the usual criteria of evidenced based medicine and their methodological restrictions in prevention research. The article is completed with the presentation of perceptible trends in the evaluation of prevention programs.

Keywords

prevention, health promotion, mental disorder, mental health, evaluation, meta-analysis

¹ Kriterien der evidenzbasierten Medizin werten Forschungsergebnisse am höchsten, wenn Nachweise aus systematischen Überblicksarbeiten (Meta-Analysen) über zahlreiche randomisiert-kontrollierte Studien gegeben werden. An zweiter Stelle stehen Nachweise auch von einer oder wenigen randomisierten Studien.

² Meta-Analysen stellen Verfahren dar, die objektiv alle Studien zu einem Thema, hier zur Prävention psychischer Störungen, sammeln, und die Effektstärken der einzelnen Studien in der Regel unter Berücksichtigung der Zahl der Studienteilnehmer mitteln (mittlere Effektstärke). Zum Begriff der Effektstärke siehe Fußnote 4

se geben z. B. über die vermittelnde Wirkung von bestimmten Faktoren Auskunft; entsprechende Statistiken aber sind in Meta-Analysen nicht verwertbar. Vor allem aber, und dies zum Dritten, werden die Vergleichsmaßstäbe oder auch normative Vorgaben für Effektstärken völlig unkritisch verwendet³. Sehr geringe Effektstärken können bei sehr bedeutsamen Erfolgskriterien, wie der Letalität, in absoluten Zahlen ausgedrückt (binominale Darstellungen von Effektgrößen) besonders dann überzeugend sein, wenn es zum Erhalt des einzelnen Lebens keine Alternative gibt als ein Interventionsprogramm mit eben dieser geringen Effektstärke (das gilt z. B. für die Programme zur Vermeidung von Suizidalität, die über geringe Effektstärken verfügen). Außerdem sind die realen empirischen Vergleichsmaßstäbe oft weit weg von den normativen Vorgaben zu dem, was eine bedeutsame Effektstärke ausmacht. So wurde in einer extrem weit angelegten Analyse psychologisch begründeter klinischer, pädagogischer und organisationspsychologischer Interventionen nur eine schwache bis mittlere Effektstärke von .42 errechnet (Lipsey/Wilson, 1993).⁴ Hinzu kommt die Tatsache, dass auch in vielen medizinisch kurativen Bereichen keine hinreichenden evidenzbasierten Maßnahmen vorliegen, wie z. B. in der Pädiatrie. Dennoch wird nicht auf entsprechende Maßnahmen verzichtet. So gilt z. B. die Behandlung depressiver Kinder mit Antidepressiva eher als fragwürdig (vgl. Hetrick, et al. 2007).

2. Auch die Angemessenheit des Einsatzes solcher Kriterien kann – zumindest in der Präventionsforschung – aus folgenden Gründen bezweifelt werden: Nutzt man pathogene Erfolgskriterien, so können Maßnahmen allenfalls bei Risikopersonen erfolgreich sein (sekundäre oder selektive Prävention), jedoch sind bei gesunden Personen im Sinne der universellen Prävention keine Verbesserungen möglich (Dach- oder Bodeneffekte). Dies gilt umso mehr, als die Auftretenswahrscheinlichkeit be-

stimmter psychischer Störungen nicht hoch und allenfalls über sehr lange Zeiträume zunimmt (Grundratenproblem). Nutzt man salutogene Erfolgskriterien (z. B. soziale Kompetenz) so muss man damit rechnen, dass diese bei der gesunden Bevölkerung schon hinreichend ausgebildet sind und zudem mäßige Vorhersagekraft in Hinsicht auf die Inzidenz, z. B. depressiver Störungen, besitzen (Segrin, 2000).

Übersieht man diese Probleme, dann kann man zu sehr diskussionswürdigen Schlussfolgerungen kommen, wie sie Klever-Deichert et al. (2007) für die universelle Prävention depressiver Störungen gezogen haben: „Eine flächendeckende Umsetzung kognitiv verhaltenstherapeutischer Interventionen zur universalen Prävention der Depression erscheint gegenwärtig aufgrund der offenen Fragen zur optimalen Ausgestaltung verfrüht.“ (S. 11, 166). Solche Schlussfolgerungen sind umso irreführender, je mehr an fundierten Belegen übersehen werden, die auf unterschiedlichen Evidenzebenen vorliegen. Dies gilt auch für den Bereich der depressiven Störungen. Auf der zweiten Ebene der Evidenzkriterien liegen zu diesem Störungsbereich methodisch gute, randomisierte Studien, aber auch weitere im Gutachten von Klever-Deichert et al. (2007) nicht erwähnte Meta-Analysen vor.⁵ Sehr bedauerlich ist u. a., dass nicht über die Meta-Analyse von Kühner (2003) berichtet wird, die für ein bestimmtes verhaltenstheoretisch fundiertes präventives Vorgehen über Effektstärken von 1.13 berichtet, einem Ergebnis, von dem viele Forscher träumen. An diesem Beispiel ist zu erkennen, dass die übersichtsartige Bewertung von Meta-Analysen die gleiche Gefahr der selektiven Wahrnehmung in sich birgt, wie die den Meta-Analysen vorausgehenden Reviews. Aus diesem Grunde ist auch an die Möglichkeit von Meta-Analysen zu denken (vgl. Röhrle, 1999). Zumindest sind vollständige Reviews über Meta-Analysen notwendig, was im Folgenden versucht wird zu leisten.

Ergebnisse zur Effektivität

Eine objektive Sammlung aller Meta-Analysen zur Prävention psychischer Störungen und Förderung psychischer Gesundheit führt bis einschließlich zum Jahre 2007 zu einer Zahl von insgesamt 113 einschlägigen Publikationen. Um zu einer konservativen Schätzung der Gesamtzahl der empirischen Studien zu gelangen, welche in diesen Meta-Analysen genutzt wurden, muss man die jeweils umfangreichsten Meta-Analysen selektieren, da in den einzelnen Meta-Analysen teilweise die gleichen Studien genutzt wurden. Auf diese Weise kommt man dann auf eine Zahl von insgesamt 1799 Studien, welche die minimale empirische Basis dieses Forschungsbereichs ergibt. Wenn man davon ausgeht, dass durch die methodischen Restriktionen von Meta-Analysen eine Vielzahl von Studien dabei unberücksichtigt bleibt und damit in der Regel auch die zugleich methodisch besseren Studien aufgenommen wurden, so kann man davon sprechen, dass dieser Forschungsbereich eine solide empirische Grundlage hat. Fasst man die Ergebnisse dieser Meta-Analysen zusammen, so weisen sie mit einer durchschnittlichen Effektstärke von .37 auf eine schwache bis allenfalls mittlere Effektivität der für psychische Störungen bzw. gesundheitsrelevanten präventiven Interventionen hin.

Gruppiert man die Studien nach verschiedenen Anwendungsfeldern, so verändert sich das Bild: Störungsspezifische Interventionen (Depression, Suizidalität, Verhaltensstörungen bzw. Delinquenz und Gewalt; Posttraumatische Stresssyndrome, Essstörungen, Drogenmissbrauch) erreichen eine durchschnittliche Effektstärke von .29 (-0.02 - 1.04). Am effektivsten erweisen sich Programme zur Vermeidung von depressiven Störungen, wobei hier insbesondere solche erfolgreich sind, die sich an Risikopersonen richten (sekundäre oder selektive Prävention; vgl. Abb. 1). Andere Störungsbereiche wie Angst-

³ Siehe Fußnote 4

⁴ Effektstärken beschreiben in der Regel das Ausmaß der Unterschiede von Interventions- zu Kontrollgruppe oder auch von Vorher-Nachher-Vergleichen (meist als Differenz der Mittelwerte in Relation zur gemeinsamen Streuung ausgedrückt; manchmal noch durch die Stichprobengröße gewichtet). Effektstärken von $d = .2$ gelten als klein, $d = .5$ als mittel und solche mit einem Wert ab $d = .8$ als groß.

⁵ Auf eine solche randomisierte Studie wird verwiesen und somit wäre Evidenzstufe 2 erreicht, dennoch bleibt man einer skeptischen Haltung Depressionspräventionsprogrammen gegenüber (vgl. Merry et al., 2004). Eine zweite wird erfasst, aber ihre Ergebnisse nicht positiv dargestellt (Clarke et al., 1993). Nicht erwähnt werden die Meta-Analysen von Cujpers et al. (2007), Horowitz/Garber (2006), Kühner (2003), Larun et al. (2006) und Merry et al. (2004).

störungen und schizophrene Störungen sind bislang noch nicht meta-analysiert, verfügen aber über randomisierte Studien mit entsprechenden Erfolgen. Wenig untersucht ist die Prävention von Hyperaktivitäts-Aufmerksamkeitsstörungen, Persönlichkeitsstörungen und auch die der häufig auftretenden somatoformen Störungen. Belastungsbezogene Präventionsformen (allgemeine Stressbewältigung; Hilfen zur Vermeidung von kritischen Lebensereignissen bzw. chronischen Belastungen und ihren Folgen als auch Hilfen zu deren Bewältigung [Ehekrise; Scheidung; sexueller, emotionaler oder körperlicher Missbrauch, Verlust von Angehörigen; Traumata, Schuleintritt, Belastung bei der Pflege von Demenzkranken]) erreichen im Durchschnitt eine Effektstärke von .40 (-.08 - 1.07). Für folgende Belastungen liegen zudem randomisierte oder kontrollierte Präventionsstudien mit günstigen Ergebnissen vor: Einsamkeit, Arbeitslosigkeit und psychisch kranke Eltern.

Programme zur Stärkung der psychischen Gesundheit im Sinne der Förderung von Kompetenzen und Ressourcen (Problemlösen, soziale Kompetenz, Wohlbefinden älterer Menschen, Entwicklungspotenziale von Kindern, familiäre und erzieherische Qualitäten) erzielen eine Effektstärke von durchschnittlich .44 (.19 - .62). Mit einiger Sicherheit sind diese Unterschiede auf die Art der in den verschiedenen Anwendungsfeldern genutzten Erfolgskriterien zurückzuführen (s. o.). Insgesamt kann man also davon ausgehen, dass unter Vernachlässigung der spezifischen methodischen Probleme der Präventionsforschung die Effektivität der entsprechenden Interventionsprogramme denen anderer psychologisch fundierter Programme gleicht (nur psychotherapeutische Interventionen erzielen davon abweichend im Schnitt eine Effektstärke von .76). Zugleich sind nicht auf Störungen bezogene präventive Programme mit einer in der Tendenz mittleren Effektstärke ausgestattet.

Ergebnisse dieser Art beschreiben eher eine konservative Einschätzung zur Effektivität von Programmen zur Prävention psychischer Störungen und Förderung psychischer Gesundheit. Bezogen auf die zweite Stufe der EBM sind randomisierte Studien und auch

Meta-Analysen vorhanden, welche sich wenigstens in den Erfolgsbereichen der Psychotherapieforschung bewegen (s. o. das Beispiel von Kühner, 2003). Zudem kann man über Studien berichten, die über Ergebnisse in Nachuntersuchungen verfügen, die teilweise 40 Jahre nach Beendigung der Intervention zu sozialpolitisch bedeutsamen Ergebnissen, wie geringere Störungs-raten, besserem sozialen Status und erheblichen Kosteneinsparungen, geführt haben (vgl. z. B. Belfield et al., 2006).

Ergebnisse zur Effizienz

Neben Antworten zur Frage, ob die Programme zur Prävention psychischer Störungen effektiv sind, sind auch schon Antworten auf die Frage gegeben worden, ob und unter welchen Umstän-

den ihre Implementation in die Praxis erfolgreich war. Drei qualitative Studien haben die erfolgreichsten Präventionsprogramme ausgesucht und inhaltsanalytisch in Hinsicht auf ihre gemeinsamen Merkmale ausgewertet (Catalano et al., 2002; Greenberg et al., 2001; Nation et al., 2003). Es zeigte sich, dass erfolgreiche Programme über eine Reihe von Merkmalen verfügten, die den Erfolg begünstigten. Diese waren:

- theoretisch gut begründet
- individuen- und umweltzentriert
- Domain übergreifend (intersektoral)
- reich an Komponenten
- multimodal gemessen
- Ressourcen-, fertigkeiten- und defizit-orientiert
- kulturell angepasst
- strukturiert (manualisiert, Einsatz von Curricula)

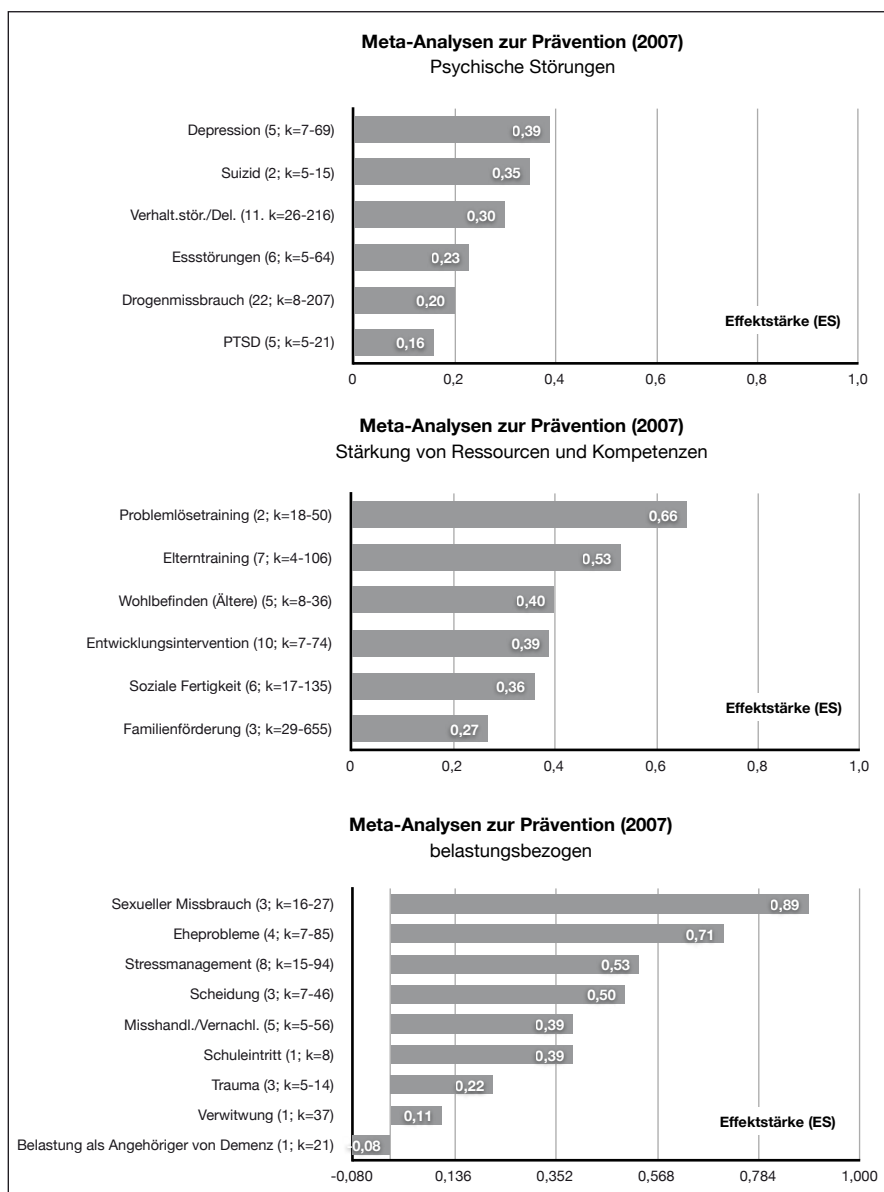


Abbildung 1: Meta-Analysen zur Prävention psychischer Störungen bis 2007

- mit variabler Didaktik ausgestattet
- durch ausgebildete TrainerInnen angeleitet
- lang anhaltend (mindestens 9 Monate von Dauer; kürzere waren nur bei Risikogruppen erfolgreich)
- alle Interaktanten berücksichtigend (partizipativ)
- in die Versorgung implementiert

Längst haben noch nicht alle Präventionsprogramme diese Erfahrungswerte genutzt, um die Effizienz ihres Vorgehens zu erhöhen. Das gilt insbesondere auch für die deutschsprachigen Programme, von denen es, gemessen an den internationalen Studien, noch wenige gut überprüfte gibt. Zu den am meisten im deutschsprachigen Bereich entwickelten und evaluierten Präventionsprogrammen gehören solche, die dem Missbrauch von Drogen und depressiven Erkrankungen vorbeugen und Hilfen bei allgemeinen sowie sozialen Belastungen (z. B. Scheidung) anbieten (vgl. *Kaluza/Lohaus, 2006*).

Was die Implementationsbreite einschlägiger Programme anbelangt, so besitzen wir aufgrund der rudimentären Versorgungsforschung in Deutschland nur ein sehr unzulängliches und auf bestimmte Anwendungsbereiche reduziertes Bild. So kann man mit einiger Sicherheit sagen, dass folgende Programme relativ breit angeboten werden: Hilfen zur Stressverarbeitung, Hilfen zur Minderung von Scheidungsfolgen, Prävention von sexuellem Missbrauch und Drogenmissbrauch, Prävention von Gewalt im schulischen Bereich und Förderung elterlicher Kompetenzen (vgl.

z. B. *Burger/Reiter, 1997; Medizinischer Dienst der Spitzenverbände der Krankenkassen, 2007; Statistisches Bundesamt, 2007*).

Abschließende Bemerkungen

Insgesamt stellt sich, je nach Sichtweise, ein ernüchterndes bzw. ermutigendes Bild zur Frage nach der Effektivität und Effizienz der Prävention psychischer Störungen und Förderung psychischer Gesundheit. Noch ist man zu weit weg von deutlicheren Wirksamkeitsnachweisen. Dies wird auch durch wenig objektive Analysen der Befundlage erschwert (vgl. *Klever-Deichert et al., 2007*).

Mehrere Wege deuten sich an, diese Situation zu verbessern. Ein erster besteht darin, durch entsprechende Veränderungen der Kriterien der EBM zu anderen Ergebnissen zu kommen und darüber hinaus entsprechend bessere zu nutzen. Ein zweiter Weg besteht darin, Forschungsarbeiten voranzutreiben, welche sowohl die spezifischen methodischen Probleme der Präventionsevaluation umgehen, als auch ihre Fragestellungen präzisieren. Dazu gehören die in letzter Zeit vermehrt auftauchenden differentiellen Fragestellungen, also solche, welche die Wirksamkeit von bestimmten Programmen für spezifische Populationen überprüfen. Hinzu kommen vermehrt Studien, welche die Wirksamkeit einzelner Programmkomponenten oder auch die Bedeutung vermittelnder Variablen untersuchen. Weil sich die Risikofaktoren für verschiedene psychische Stö-

rungen recht ähnlich sind, ist auch ein Trend erkennbar, Programme zu entwickeln, die gewissermaßen komorbid bzw. kopräventiv wirksam sind. Und einen letzten Trend mag man darin erkennen, dass zunehmend verhältnis- und individuenorientierte Programme kombiniert und bedarfsgerecht angeboten werden.

So bewegt sich die derzeitige Forschung immer mehr im Bereich von Forschungsstrategien, die der Psychotherapieforschung schon seit geraumer Zeit bekannt sind. Ein deutlicher Nachholbedarf ist darin zu erkennen, dass demgegenüber keine Prozessanalysen durchgeführt wurden, die neue Wirkfaktoren zu entdecken helfen. Zumindest für den bundesdeutschen Bereich gilt, dass Präventionsforschung insgesamt quantitativ und in diesem Sinne auch qualitativ zu stärken ist, wobei für sozialpolitische Maßnahmen sich auch eine hinreichende Versorgungsforschung als notwendig erweist.

Literatur:

Bei dem Verlag, dem Verfasser und vorübergehend im Internet unter www.zeitschrift-praevention.de bei diesem Artikel nach der Zusammenfassung.

Anschrift des Verfassers:

Prof. Dr. Bernd Röhrle
 Fachbereich Psychologie
 Philipps-Universität Marburg
 Gutenbergstr. 18
 35032 Marburg
roehrle@staff-uni-marburg.de
<http://www.uni-marburg.de/fb04/ag-klin/mitarbeiter/broehrle>