

Bewertungs- und Klassifikationssysteme in der Suchtprävention

Jens Kalke, Hermann Schlömer



Erklärung Interessenkonflikt

- Die folgende Präsentation basiert nicht auf einem finanziell geförderten Projekt.
- Das Bewertungssystem des Qualitätszirkels Nordverbund (BQN) wurde in Eigeninitiative ohne finanzielle Förderung Dritter entwickelt.
- An der Entwicklung des BQN war auch der Referent beteiligt.

Inhalt

- 1. Ausgangssituation**
- 2. Bewertungssysteme**
- 3. Empfehlungen**

Ausgangssituation

Projektvielfalt im „Programm-Dschungel“



**60 evaluierte Projekte
in der DHS Übersicht**

**470 Projekte
2014 bei
PrevNet**

**208 Projekte im HTA
Bericht von Korczak**

**745 eingereichte Projekte bei 7 Bundeswettbewerben
„Strategien kommunaler Suchtprävention“ 2001 bis 2016**

Projekttypologie des „Programm-Dschungels“

**Weltrettungsideen
Einzelner**

**Leuchtturmprojekte
von Politikern**

**innovative,
kreative Projekte**

**evidenzbasierte
Maßnahmen**

**kommerzielle
Angebote**

**Duftmarken
von Fachstellen**

Vorteile von Bewertungssystemen im „Programm - Dschungel“

- **Orientierungshilfe für die Praxis**
- **Verbreitung evidenzbasierter Suchtprävention**
- **verantwortungsvoller Einsatz von Ressourcen**

Empfehlungen aus dem Kölner Memorandum 2014

- Die Praxis sollte in der Zukunft verstärkt taugliche Instrumente zur Recherche von evidenzbasierten Maßnahmen, Projekten und Programmen nutzen.
- Dafür stehen Projekt-, Methodendatenbanken und Bewertungsinstrumente wie EDDRA oder die Grüne Liste Prävention zur Verfügung.
- Es ist anzustreben, dass zukünftig ein einheitliches, qualitativ hochwertiges Bewertungsinstrument zur Verfügung stehen könnte.

Bewertungssysteme



Bewertungssysteme

- EDDRA: Exchange on Drug Demand Reduction Action
- BQN: Bewertungssystem des Qualitätszirkels Nordverbund
- GLP: Grüne Liste Prävention

EDDRA

About EDDRA quality levels

The purpose of EDDRA is to compile evaluated practices in drug demand reduction implemented in Member States of the European Union. In order to take into account differences in the quality of evaluated practices, a set of criteria were developed in 2007. These criteria focus on the quality assessment of intervention evaluation. They are used to categorise evaluated projects according to three quality levels:

- Level 1 projects
- Level 2 projects: promising projects
- Level 3: top level projects

Level 2 and 3 are determined via a points system.

The criteria focus primarily on implementation and evaluation design based on the [logic model](#) and data collection. Criteria that help to judge the strength of the available evidence on the effectiveness of the intervention will be introduced once the majority of submitted projects are rated as promising or top level projects.

Criteria for Level 1 projects

All projects that meet the following entry criteria will be added to EDDRA:

- The project has been evaluated
- It has a theory base/assumptions that are clearly related to its objectives
- It has clear evaluation indicators in place that relate to its objectives and initial situation
- It contains a clear description of the evaluation design
- The project must be at least one year old

EDDRA: Bewertung und Kategorisierung von Projekten

1. Zielsetzung

Zusammenstellung evaluierter, in den EU Staaten implementierter, suchtpreventiver Projekte

2. Kriterien zur Bewertung evaluierter Projekte in Bezug auf:

- logische Stimmigkeit des Maßnahme- und Evaluationskonzepts
- Qualität des Evaluationsdesigns und der genutzten Instrumente
- Qualität der evaluierten Maßnahme

3. Kriteriengestützte Kategorisierung evaluierter Projekte durch Punktvergabe

- Level 1 Projekte
- Level 2 erfolgsversprechende Projekte
- Level 3: "Top level" Projekte

EDDRA Einschlusskriterien für “Level 1”- Projekte

- ein evaluiertes Projekt
- abgeleitete Projektziele aus theoretischem Konzept und Annahmen
- Evaluationsindikatoren mit Bezug auf Ziele und Ausgangssituation
- ein beschriebenes Evaluationsdesign
- ein mindestens ein Jahr altes Projekt

Kriterien und Punktgewichtung für Level 2 und 3 EDDRA - Projekte

logisches Konzept	Punkte	Evaluation	Punkte	zusätzliche Informationen	Punkte
spezifische Ziele	1	Prozess-Ev. *	2	Koordination mit anderen Angeboten/Programmen *	2
Indikatoren für alle Ziele	1				
ein oder mehr quantifizierbare Indikator- Dimensionen pro Ziel	1				
Bezug der Ziele zur Ausgangssituation	1	Outcome-Ev.:			
zu den Zielen passende dargestellte Ergebnisse *	1	Follow up	2	Outcome-Ev.- Instrumente verfügbar u. übermittelt an EMCDDA	2
Outcome Ergebnisse	1	naturalistisches Prä-/Postdesign ohne Kontrollgruppe	4	neue oder noch nicht in der EIB vorhandene Outcome-Ev.- Instrumente	1
Bezug der Arbeitshypothese zur Ausgangssituation *	1	quasi experimentelles Prä-/ Postdesign mit Kontrollgruppe	8 (12) ¹	validierte neue Outcome-Ev.- Instrumente	1

Kriterien und Punktgewichtung für Level 2 und 3 EDDRA - Projekte

logisches Konzept	Punkte	Evaluation	Punkte	zusätzliche Informationen	Punkte
evidenzbasierte (mind. per Kontrollgruppenstudie) Arbeitshypothese *	2	Prä-/Post-design und RCT	12	ein Projektmanual ist verfügbar und an EMCDDA übermittelt	2
Verbindung der Arbeitshypothese mit Zielen und deren Indikatoren *	2	Outcome Ev. mit modifiziertem validiertem Instrument	2		
Maßnahmen/Programminhalte passen zu den Zielen *	1	Outcome Ev. mit validiertem Instrument	4		
Maßnahmen/Programminhalte passen zu den Zielen und der Arbeitshypothese *	1				
maximale Punktzahl	13	maximale Punktzahl	18	maximale Punktzahl	8

* = Kriterien auch für Prozess-Evaluation; Projekte mit nur Prozessevaluation können Level 2 erreichen und ¹ = für das setting das passendste Design

Cut-off-points für Level 2 (erfolgsversprechende Projekte): > 12

Cut-off-points für Level 3 („Top Level“ Projekte): > 28

Die neue Fassung in der Pilotphase: About Xchange ratings

nützlich: Interventionen, für die in zwei oder mehr europäischen Studien von excellenter Qualität überzeugende, eindeutige and nachhaltige Effekte für relevante Outcomes nachgewiesen wurden.

wahrscheinlich nützlich: Interventionen, für die in mindestens einer europäischen Evaluationsstudie von excellenter Qualität überzeugende und eindeutige Effekte für relevante Outcomes nachgewiesen wurden.

möglicherweise nützlich: Interventionen, für die in mindestens einer europäischen Evaluationsstudie von akzeptabler Qualität einige Effekte für relevante Outcomes nachgewiesen wurden. Eine als “möglicherweise nützlich” bewertete Intervention, eignet sich für die Anwendung im Kontext einer strengeren Evaluation.

zusätzliche Studien empfohlen: Interventionen, bei denen wegen Bedenken hinsichtlich der Qualität der Evaluation oder der Eindeutigkeit der Outcomes in Europa es schwierig ist zu beurteilen ob sie wirksam sind oder nicht, auch wenn Outcomes für die Intervention zu sprechen scheinen.

unwahrscheinlich nützlich: Interventionen für die mindestens eine europäische Evaluationsstudie von excellenter Qualität überzeugend belegt, dass sie keine oder schädliche Effekte hinsichtlich relevanter Outcomes hat.

Norddeutsches Bewertungssystem für Unterrichtsprogramme und Projekte schulischer Suchtprävention



Teil I: Unterrichtsprogramme

Norddeutsches Bewertungssystem für Unterrichtsprogramme (BQN I)

	Primar	Sek I	Sek II
I. Effektkriterien (signifikante Veränderungsnachweise durch Kontrollgruppenstudien)	ja/nein	ja/nein	ja/nein
1. Nachweis von Effekten für Life skills (WHO)			
2. Nachweis von Konsumreduktion/-stabilität bei Risikokonsum	-		
3. Nachweis von Konsumreduktion/-stabilität generell bzw. Abstinenzerrhöhung/-stabilität			
4. Nachweis von anderen Gesundheitseffekten			
5. Nachweis von Einstellungsveränderungen			
6. Nachweis von Wissenssteigerung			
7. Nachweis eines positiven Einflusses auf Schul- / Klassenklima			

Es muss mindestens ein Kriterium der Bereiche 1. bis 3. oder zwei der Bereiche 4. bis 7. erfüllt sein, um die beste Einstufungsklasse erreichen zu können.

BQN (II)

	Punkte	Punkte	Punkte
II. Extra: Hinweise auf Effekte durch systematische Erhebungen ohne Kontrollgruppen	6	6	6
III. Akzeptanz	4	4	4
1a. hohe Akzeptanz bei mind. 2/3 der Adressaten lt. Adressatenrückmeldung			
1b. hohe Akzeptanz bei ausschließlicher Fremdbewertung	2	2	2
2. hohe Akzeptanz bei mind. ¾ der Durchführenden lt. eigener Rückmeldung	4	4	4
IV. Vorhandensein nachgewiesener inhaltlicher Erfolgsmerkmale (evidenzbasierte/-orientierte Merkmale nach Bühler/Kröger):	6	6	6
1. life-skill-fördernd ausgerichtet (siehe Operationalisierung)			
2. interaktiv (siehe Operationalisierung)	6	6	6
3. sozial beeinflussend (Indikatoren: Standfestigkeitstrainings, Normen setzend)	6	6	6
4. Verpflichtung/Übungen zum Konsumverzicht/-zur Konsumreduktion	4	4	4
5. Peers als Vermittler einbeziehend		2	2
6. mittlere Dauer (d. h. 11 bis 30 Std. pro Schuljahr)	2	2	2
V. didaktische - und Durchführungsqualität (Merkmale nach PISA etc.)	6	6	6
1. enthält binnendifferenzierende Unterrichtselemente (für verschiedene Lernniveaus)			
2. Berücksichtigung von Genderaspekten (differenzierte Angebote für Mädchen und Jungen)	4	4	4
3. Eltern einbeziehend	3	3	

BQN (III)

Qualitätsklasse AA: Unterrichtsprogramme mit Wirkungsnachweis durch Kontrollgruppenstudie

(siehe Tabelle auf S. 10: mindestens ein Kriterium von 1.-3. oder zwei von 4.-7. im Bereich I. zutreffend)

	Primarstufenprogramme (Wegfall des Kriteriums „Peers“ mit 2 Punkten)	Programme für die Sekundarstufe I	Programme für die Sekundarstufe II (Wegfall des Kriteriums „Einbeziehung der Eltern“ mit 3 Punkten)
erreichbare Maximalpunktzahl	47 Punkte aus III. bis V.	49 Punkte aus III. bis V.	46 Punkte aus III. bis V.
empfehlenswert	24 Punkte (über 50%)	25 Punkte (über 50%)	24 Punkte (über 50%)
sehr empfehlenswert	36 Punkte (über 75%)	37 Punkte (über 75%)	35 Punkte (über 75%)

BQN (IV)

Qualitätsklasse A: Evidenzorientierte Unterrichtsprogramme ohne Wirkungsnachweis durch Kontrollgruppenstudie (bei vorliegenden Effektnachweisen durch Evaluationen inkl. 6 Bonuspunkten)

	Primarstufenprogramme (Wegfall des Kriteriums „Peers“ mit 2 Punkten)	Programme für die Sekundarstufe I	Programme für die Sekundarstufe II (Wegfall des Kriteriums „Einbeziehung der Eltern“ mit 3 Punkten)
erreichbare Maximalpunktzahl	53 Punkte aus II. bis V.	55 Punkte aus II. bis V.	52 Punkte aus II. bis V.
empfehlenswert	27 Punkte (über 50%)	28 Punkte (über 50%)	27 Punkte (über 50%)
sehr empfehlenswert	40 Punkte (über 75%)	42 Punkte (über 75%)	39 Punkte (über 75%)

Grüne Liste Prävention – CTC - Datenbank empfohlener Präventionsprogramme

In Deutschland existiert eine Fülle von Programmen, die zum Ziel haben, der Gewalt, der Kriminalität, dem Suchtverhalten und anderen Problemverhaltensweisen von Kindern und Jugendlichen vorzubeugen. Diese Programme sind in Bereichen wie der Eltern- und Familienbildung, der Kompetenzförderung bei Kindern und Jugendlichen, in der Kindertagesstätte und der Schule oder im sozialen Umfeld der Nachbarschaft angesiedelt. Im internationalen Vergleich liegen in Deutschland aber nur wenige hochwertige Evaluationsstudien vor, die zeigen, welche Programme tatsächlich wirksam sind.

Trotz dieses Defizites an Evaluationsforschung können die in Deutschland angebotenen Präventionsprogramme nach der Güte ihrer Wirkungsüberprüfung und ihrer Konzeptqualität unterschieden werden. Die vorliegende Datenbank "Grüne Liste Prävention" bietet auf der Basis nachvollziehbarer [Kriterien](#) einen Überblick über empfehlenswerte Präventionsansätze in den Bereichen Familie, Schule, Kinder/Jugendliche und Nachbarschaft: Welche Programme können mit Aussicht auf Erfolg wo, wann und wie eingesetzt werden, um die Entwicklung von Kindern und Jugendlichen zu unterstützen? Auf welche Weise beugen diese Programme der Entstehung oder der Verfestigung von Problemverhaltensweisen vor? Und schließlich: Was weiß man darüber, ob diese Programme funktionieren und zu welchen Ergebnissen sie führen?

Die "Grüne Liste Prävention" will den Stand der Entwicklung von ausgewählten Programmansätzen möglichst genau abbilden. Daher werden die empfohlenen Präventionsprogramme noch einmal in drei Stufen bezüglich des Nachweises ihrer Wirksamkeit eingeteilt:

Für weitere Informationen bitte anklicken.

- 1 Effektivität theoretisch gut begründet
- 2 Effektivität wahrscheinlich
- 3 Effektivität nachgewiesen

Die Bewertung „Auf der Schwelle“ kommt zustande, wenn nicht alle [Kriterien](#) auf der Konzeptebene für ein theoretisch überzeugendes Modell erfüllt sind; wenn nötige Informationen zur Einstufung noch nicht vorliegen oder wenn noch überhaupt keine Evaluation vorgenommen wurde.

Datenbank

Information

[Kriterien \(.pdf\)](#)

[Programm vorschlagen](#)

[Neue Programme](#)

[Alle empfohlenen Programme](#)

[Programme auf der Schwelle](#)

[In Bearbeitung](#)

Programmsuche

[erweiterte Suche](#)

[Risikofaktoren](#)

[Schutzfaktoren](#)

[FAQ - Häufig gestellte Fragen](#)

Grüne Liste Prävention (GLP) – CTC - Datenbank empfohlener Präventionsprogramme

... bietet auf Basis nachvollziehbarer Kriterien einen Überblick über empfehlenswerte Präventionsprojekte in den Bereichen Familie, Kindertagesstätte, Schule und Nachbarschaft für die Vorbeugung von Gewalt, Kriminalität, Suchtverhalten und anderen Problemverhaltensweisen von Kindern und Jugendlichen.

Projekte werden nach Güte ihrer Wirkungsüberprüfung und ihrer Konzeptqualität unterschieden und in drei bzw. fünf Stufen eingeteilt.

GLP Einstufung der Programme und Kriterien I

Stufe1: Effektivität theoretisch gut begründet1

1 a) Konzept

- Konzept, Methoden und Instrumente theoretisch gut begründet
- logischer Zusammenhang von: „Problemanalyse – Zielen – Zielgruppen – Methoden“
- Programmbezug auf wissenschaftlich belegte Risiko- und Schutzfaktoren
- Zielgruppe(n) präzise beschrieben
- Handlungsanleitungen und Manuale nachvollziehbar und aus dem Modell abgeleitet
- Ziele explizit formuliert und überprüfbar
- für nicht in Deutschland entwickelte Programme ursprünglicher Kontext und erfolgte Anpassungen beschrieben

1b) Umsetzung

- Methoden und Instrumente didaktisch gut aufbereitet und nachvollziehbar
- benötigte Materialien/ Manuale verfügbar und auf aktuellem Stand
- notwendige Trainings und Schulungen vorhanden
- entstehende Kosten transparent
- Support für die Umsetzung verfügbar
- Instrumente für Qualitätskontrolle bei der Umsetzung verfügbar

1c) Evaluation

- mindestens eine Evaluation mit Niveau 0 Sterne und (überwiegend) positivem Ergebnis

GLP Einstufung der Programme und Kriterien II

Stufe 2: Effektivität wahrscheinlich

- 1 a) und b) wie oben
- mindestens eine Evaluationsstudie * -*** Sterne mit (überwiegend) positivem Ergebnis

Stufe 3: Effektivität nachgewiesen

- 1 a) und b) wie oben
- mindestens eine Evaluationsstudie **** oder ***** Sterne mit (überwiegend) positivem Ergebnis und mindestens hinreichender Beweiskraft

Sobald bei der Stufe 1 ein Punkt (1 a, b oder c) nicht erfüllt ist, führt dies zu der Einordnung **“Auf der Schwelle”**. Sind die ersten drei Punkte der Konzeptqualität (1 a) von Stufe 1 nicht erfüllt, erfolgt die Einordnung **“keine Aufnahme”**

GLP (II)

Konzept- und Umsetzungs-qualität	Evaluations-niveau und Beweiskraft	Evaluations-ergebnisse	Gesamtbewertung
1 a – c sind erfüllt	4 bis 5 Sterne: hinreichende Beweiskraft oder besser	(überwiegend) positiv	Effektivität nachgewiesen (Stufe 3)
1 a – c sind erfüllt	1 bis 3 Sterne 4 – 5 Sterne mit nur schwacher Beweiskraft	(überwiegend) positiv	Effektivität wahrscheinlich (Stufe 2)
1 a – c sind erfüllt	0 Sterne	(überwiegend) positiv	Effektivität theoretisch gut begründet (Stufe 1)
1 a – c sind (teilweise) erfüllt	keine Evaluation oder unterhalb 0 Sterne	teils positiv und negativ oder unklar	auf der Schwelle
1 a – c sind (nicht) erfüllt	Alle	keine oder negative Effekte	Keine Aufnahme

GLP (III)

Niveau Evaluations- design	schwache Beweiskraft	hinreichende Beweiskraft	starke Beweiskraft	sehr starke Beweiskraft
4 – 5 Sterne	<p>Interventions- und Kontrollgruppen n < 20</p> <p>oder</p> <p>keine Studie im deutschsprachigen Raum, aber eine oder mehrere ausländische Studien mit mindestens hinreichender Beweiskraft*</p>	<p>Interventions- und Kontrollgruppen n = 20 - 50</p> <p>oder</p> <p>nur schwache Effektstärken</p> <p>und</p> <p>mindestens eine Studie im deutschsprachigen Raum</p>	<p>Interventions- und Kontrollgruppen n > 50</p> <p>oder</p> <p>mittlere und größere Effektstärken</p> <p>oder</p> <p>2 oder mehr Studien mit hinreichender Beweiskraft im deutschsprachigen Raum und mehrere ausländische Studien*</p>	<p>Kriterien wie bei starker Beweiskraft plus mindestens ein groß angelegtes Feld-Experiment im deutschsprachigen Raum</p>
1 - 3 Sterne	vorläufige (* - **) oder schwache (***) Beweiskraft			
0 Sterne	keine Beweiskraft			

Vergleich der Bewertungssysteme

	BQN	EDDRA	GLP
Konzept- und Durchführungsqualität des Programms/Projekts	differenziert nach Konzept und Durchführung	differenziert nach Konzept	differenziert nach Konzept und Durchführung
Forschungsdesign der Evaluation	differenziert nach Studien mit und ohne Kontrollgruppe	differenziert nach Studien mit und ohne Kontrollgruppe und Randomisierung	differenziert nach Studien mit und ohne Kontrollgruppe und Randomisierung
Evaluationsinstrumente	nicht enthalten	differenziert nach Validität	nicht enthalten
Evaluationsergebnisse: Akzeptanz	differenziert nach Zielgruppe und Durchführenden	nicht enthalten	nicht enthalten
Evaluationsergebnisse: Effekte	differenziert nach einzelnen Zielsetzungen	im Pilot enthalten	nach Zielsetzungen insgesamt

Bewertung von schulischen Unterrichtsprogrammen

	BQN	EDDRA	GLP
Klasse 2000	Qualitätsklasse AA: „empfehlenswert“	Qualitätsstufe 3: „Top-Level-Projekt“	Gesamtbewertung Stufe 2: „Effektivität wahrscheinlich“
ALF	Qualitätsklasse AA: „empfehlenswert“	Qualitätsstufe 1: „in Datenbank aufgenommen“ im neuen Pilot: wahrscheinlich nützlich	Gesamtbewertung Stufe 3: „Effektivität nachgewiesen“
Erwachsen werden	Qualitätsklasse AA: „sehr empfehlenswert“	Qualitätsstufe 1: „in Datenbank aufgenommen“	Gesamtbewertung Stufe 2: „Effektivität wahrscheinlich“
Eigenständig werden	Qualitätsklasse AA: „empfehlenswert“	im neuen Pilot: nützlich	Gesamtbewertung Stufe 2: „Effektivität wahrscheinlich“

Empfehlungen

Empfehlungen für die Praxis

- intensivere Nutzung der Bewertungssysteme in der Praxis
- verstärkte Schulung der suchtpreventiven Fachkräfte im Themenbereich „Evaluation und Qualitätssicherung“

Empfehlungen für die Forschung

- Ein einheitliches Bewertungssystem (unter Beteiligung von Praktikern/innen) entwickeln.
- Dabei die Schwächen der vorhandenen Bewertungssysteme überwinden.
- Zwischen notwendiger Ausdifferenzierung der Qualitätsparameter und der Gefahr der Überkomplexität die richtige Balance finden.

Links & Literatur

- [EDDRA: www.emcdda.europa.eu](http://www.emcdda.europa.eu) (EDDRA)
- Nordverbund suchtpräventiver Fachstellen:
Norddeutsches Bewertungssystem für
Unterrichtsprogramme, Hamburg 2012
- [GLP: www.gruene-liste-praevention.de](http://www.gruene-liste-praevention.de)
- Schlömer H. & Kalke J. (2013). Bewertungs- und Klassifikationsinstrumente für die Suchtprävention: Aktueller Stand und zukünftige Herausforderungen. Suchttherapie 14 (3): 121-127.

**Vielen Dank
für Ihre
Aufmerksamkeit !**

Institut für interdisziplinäre Sucht- und Drogenforschung (ISD-Hamburg)

