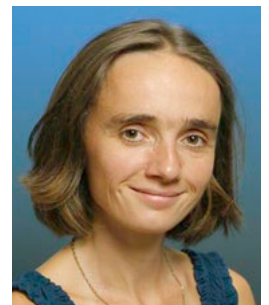




*Yvonne Blumenthal*



*Kathrin Mablau*

# Effektiv fördern – Wie wähle ich aus? Ein Plädoyer für die Evidenzbasierte Praxis in der schulischen Sonderpädagogik

## Zusammenfassung

Unterricht und Förderstunden sollten so gestaltet werden, dass die Schülerinnen und Schüler schnell, sicher und nachhaltig ein bestimmtes Lern- oder Entwicklungsziel erreichen. In der konkreten schulischen Umsetzung spielen dabei die professionelle Erfahrung der Lehrkräfte, vorhandenes Unterrichts- und Fördermaterial sowie Vorlieben der Schüler eine zentrale Rolle. Im Beitrag wird der Frage nachgegangen, ob dieses Vorgehen zugunsten einer effektiveren Förderung verbessert werden könnte. Dazu wird das Prinzip der Evidenzbasierten Praxis vorgestellt und auf sonderpädagogische Inhalte bezogen. Ein konkretes Vorgehen zur Auswahl evidenzbasierter Förder-, Trainings- und Therapieverfahren wird anhand eines Stufenmodells erläutert und durch ein Beispiel aus dem Förderbereich der emotional-sozialen Entwicklung verdeutlicht.

## Einleitung

Im Unterricht und in schulischen Förderstunden agieren verschiedene Personen wie Lehrkräfte an Allgemeinen Schulen oder Sonderpädagogen mit Schülern unterschiedlichster Lern- und Entwicklungsvoraussetzungen. Aufgabe schulischer Förderung ist es, alle Kinder zu individuellen Lern- und Entwicklungszielen zu führen. Doch wie erreicht ein Grundschul- oder Sonderpädagoge möglichst schnell, wirksam und nachhaltig bei einer Klasse oder einem einzelnen Schüler ein angestrebtes Lern- oder Entwicklungsziel? Im Unterricht und in der Förderung werden häufig auf persönliches Erfahrungswissen und auf in der Schule vorhandenes Unterrichts- und Fördermaterial zurückgegriffen und dabei die eigenen sowie die Vorlieben der Schüler berücksichtigt. Es stellt sich die Frage, ob dieses in der Praxis häufig anzutreffende Vorgehen mit dem Ziel einer effektiveren Förderung verbessert werden könnte.

In den letzten Jahren wird in der schulischen Förderung (Voss, Sikora & Hartke, 2015) und insbesondere in der schulischen Sonderpädagogik (Beushausen, 2014; Hartmann, 2013; Hillenbrand, 2014; Fingerle & Ellinger, 2008; Nussbeck, 2007; Hartke, 2005) das Konzept der sogenannten Evidenzbasierten Praxis thematisiert. Evidenzbasierte Praxis wird häufig so verstanden, dass nur die Methoden und Verfahren Anwendung finden sollten, bei denen in wissenschaftlichen Studien nachgewiesen wurde, dass sie tatsächlich das angestrebte Interventionsziel erreichen. Darüber hinaus sind sie im Vergleich zu anderen Methoden und Verfahren effektiver. Dieses Vorgehen ist aus der Medizin bekannt, in der es eine Selbstverständlichkeit ist, nur wissenschaftlich nachgewiesene wirksame Medikamente zu verschreiben.

Im Bereich der (Sonder-)Pädagogik greift jedoch zu kurz, nur den wissenschaftlichen Nachweis über die Wirksamkeit einer pädagogischen Maßnahme, die sogenannte externe Evidenz, zu berücksichtigen. Im pädagogischen Setting spielen auch das individuelle Wissen und die Erfahrung der Lehrkräfte sowie die Bedürfnisse der Schülerinnen und Schüler zu berücksichtigende Rollen.

Daher sollten evidenzbasierte Entscheidungen für Unterrichtsmethoden, Materialien und Fördermaßnahmen vor dem Hintergrund aktuell gültiger wissenschaftlicher Erkenntnisse (externe Evidenz), unter Einbezug des individuellen Wissens und der Erfahrung – der Expertise – der pädagogischen Fachkräfte (interne Evidenz) und der spezifischen Bedürfnisse der Schülerinnen und Schüler bzw. deren Erziehungsberechtigten (soziale Evidenz) getroffen werden. Ziel ist es, die bestmögliche und geeignetste Handlungsstrategie für die beteiligten Personen zu finden (s. Abbildung 1).

Wie könnte dies an einem praktischen Beispiel aussehen? Lehrkräfte stellen nicht selten fest, dass es einigen Kindern in ihrer Klasse erheblich an positiven sozialen Strategien und emotionalen Kompetenzen mangelt. Bei einer Meinungsverschiedenheit werden keine konstruktiven Lösungsstrategien eingesetzt, sondern es wird verbal ausfallend geantwortet oder schlimmstenfalls gleich körperlich reagiert. Diesem Verhalten entgegenzuwirken, ist beständiger Inhalt im Schulalltag eines jeden Pädagogen. Doch unterscheiden sich die eingesetzten Maßnahmen zur Entwicklung sozialer Kompetenzen und zum Unterbinden von Fehlverhalten bei Schülern zwischen den Lehrkräften erheblich. Während eine Lehrkraft Belohnungssysteme mit Gummibärchen und Smileys einsetzt, drohen der Kollege im Mathematikunterricht mit dem Ausfall der Klassenfahrt und die Sportlehrerin mit einhundert Sit-Ups. Zunehmend stehen Lehrkräfte vor der Aufgabe, die Förderung emotional-sozialer Kompetenzen explizit als übergreifendes Entwicklungsziel für einige oder auch für alle Kinder ihrer Klasse als Unterrichtsinhalt zu berücksichtigen. Dafür bietet sich der Einsatz von Trainingsverfahren an, deren Wirksamkeit in Studien nachgewiesen wurde.

In den letzten Jahren sind eine Reihe von Verfahren zur Förderung emotionaler und sozialer Kompetenzen entwickelt worden, wie z. B. Faustlos (Cierpka, 2003), Verhaltenstraining für Schulanfänger (Petermann, Natzke, Gerken & Walter, 2013a), Verhaltenstraining in der Grundschule (Petermann, Koglin, Natzke & von Marées, 2013b), Friedensstifter Training (Gasteiger-Klicpera & Klein, 2006), „Lubo aus dem All!“ (Hillenbrand, Hennemann & Hens, 2010) oder „Schwierige Schüler – 49 Handlungsmöglichkeiten bei Verhaltensauffälligkeiten“ (Hartke & Urban, 2014), die sich in den Unterrichtsalltag integrieren lassen und mehrheitlich sehr ansprechend und kindgerecht aufbereitet sind (Identifikationsfigur, Handpuppen), der Lehrkraft einen klaren Aufbau der Förderinhalte und -methoden vorgeben und auch gleich das notwendige Arbeitsmaterial (Lieder-CD, Arbeitsheft, Situationsbilder) mitliefern. Erklärte und wissenschaftlich nachgewiesene Ziele dieser Verfahren sind es, sozial-emotionale Kompetenzen aufzubauen und Verhaltensstörungen langfristig zu vermeiden. Mit anderen Worten: Setzt die Lehrkraft das Verfahren im Unterricht erfolgreich ein, erleichtert sie sich die pädagogische Arbeit in den nachfolgenden Schuljahren erheblich. Die Lehrkraft steht aber zunächst vor dem Problem zu entscheiden, welches Verfahren am effektivsten wirkt und was zu den Schülern der Klasse und zur ihr selbst am besten passt. Dabei kann das nachfolgend beschriebene Ablaufschema zum Treffen einer evidenzbasierten Entscheidung helfen.

In der Evidenzbasierten Praxis wird grundsätzlich ein stufenweises Vorgehen empfohlen, um die Entscheidung für ein geeignetes Interventionsverfahren oder Fördermaterial zu treffen. Die konkrete Anzahl der Stufen eines evidenzbasierten Entscheidungsprozesses variiert in der Regel zwischen fünf und acht, inhaltlich stimmen sie jedoch im Wesentlichen überein (Dollaghan, 2007; Mangold, 2013; Sackett et al., 2000; Nussbeck, 2013; Beushausen, 2014; Schlosser & Raghavendra, 2004). Abbildung 2 fasst den Ablauf des evidenzbasierten Entscheidungsprozesses zusammen.

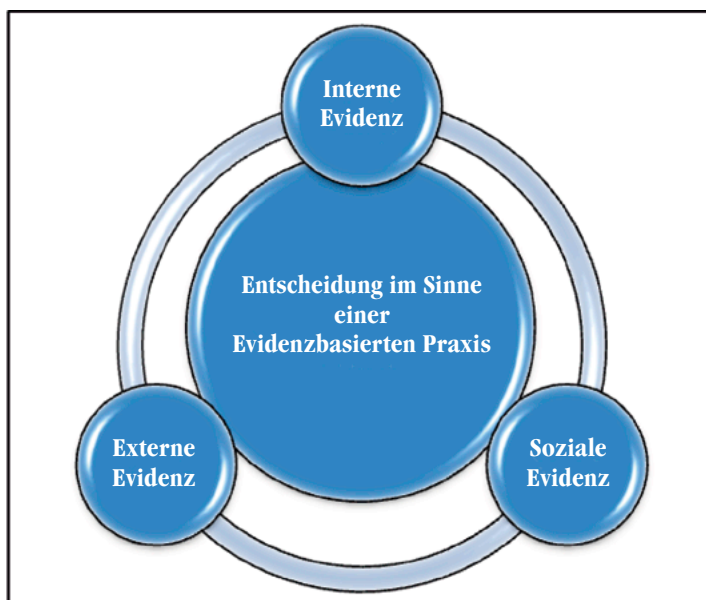
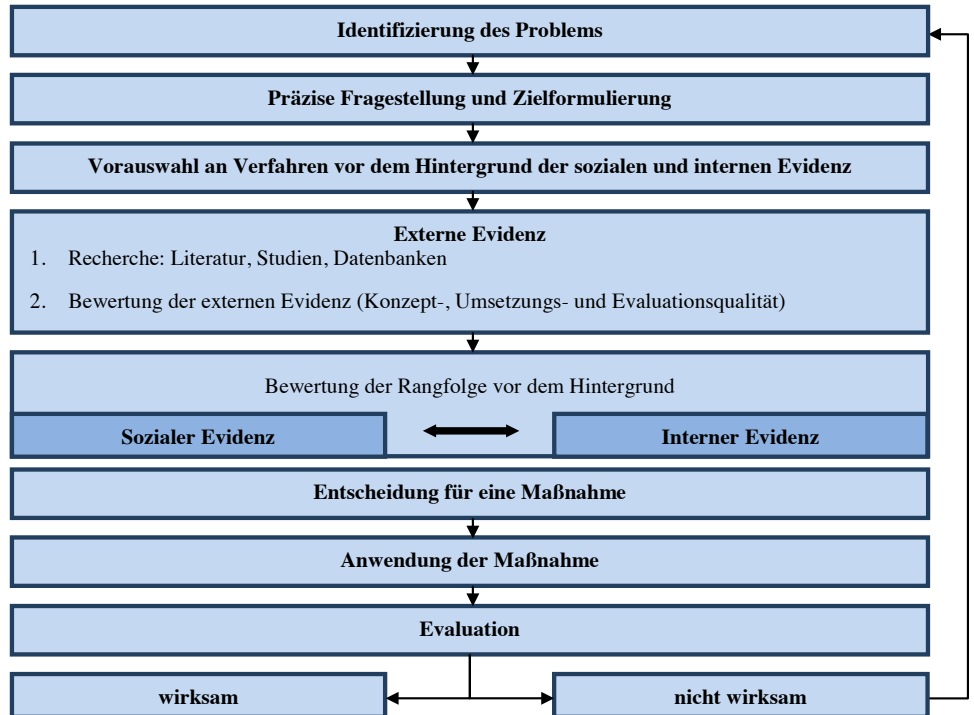


Abbildung 1:  
Evidenzaspekte im  
Entscheidungsfindungsprozess

## Ablauf einer evidenzbasierten Entscheidung

Abbildung 2:  
Ablauf eines evidenzbasierten  
Entscheidungsprozesses



Übernimmt beispielsweise eine Lehrkraft eine jahrgangsübergreifende Klasse der Klassenstufen eins bis drei an einer Grundschule und stellt fest, dass mehrere Kinder erhebliche Schwierigkeiten im sozial-emotionalen Bereich haben, dann sollte sie in einem ersten Schritt das Problem differenzierter beschreiben. Dazu eignen sich Fragebögen zur Erfassung des Sozialverhaltens durch die Lehrkraft oder die Eltern (z. B. The Strengths and Difficulties Questionnaire - SDQ, Goodman, 2005), Fragebögen zur Erfassung der emotionalen und sozialen Schulerfahrungen durch die Schüler (FEESS 1-2/3-4, Rauer & Schuck, 2004; 2003) sowie Fragebögen zur Erfassung der sozialen Stellung (Blumenthal & Marten, 2014). Für einzelne Kinder könnte der Einsatz spezifischerer Diagnostikverfahren, wie der Teachers Report Form (TRF, Arbeitsgruppe Deutsche Child Behavior Checklist, 1993) oder der Fragebogen Schulische Einschätzung des Verhaltens und der Entwicklung (SEVE, Hartke & Vrban, 2014) sinnvoll sein.

In unserem Beispiel stellt die Lehrkraft anhand der Ergebnisse des SDQ-Eltern sowie des SDQ-Lehrerfragebogens (Goodman, 2005) fest, dass mehrere Schüler ihrer Klasse sowohl zu Hause als auch in der Schule ein sozial unangepasstes Verhalten zeigen. Die Informationen aus dem FEESS 1-2 (Rauer & Schuck, 2004) bzw. FEESS 3-4 (Rauer & Schuck, 2003) verdeutlichen, dass ca. zwei Drittel der Kinder das Klassenklima als angespannt empfinden und über die Hälfte eine geringe Lernfreude hat. Als besonders problematisch erweist sich in der Auswertung des soziometrischen Schülerfragebogens, dass einzelne Kinder von mehreren anderen Schülern der Klasse sozial abgelehnt werden. Anhand informeller Elterngespräche erfährt die Lehrkraft darüber hinaus, dass der familiäre Hintergrund einer größeren Anzahl der Kinder ihrer Klasse stark belastet ist, so dass von Seiten der Eltern lediglich geringe Unterstützung zur Lösung des Problems zu erwarten ist.

Zusammengefasst sieht die Lehrkraft ihre Annahme eines ungünstigen sozialen Miteinanders der Schüler durch die standardisierte Diagnostik bestätigt und hat dieses nun für jeden Schüler individuell und für die gesamte Klasse differenziert als Problem identifiziert.

### Präzise Fragestellung und Zielformulierung

Die sich aus dem vorliegenden Problem entwickelnde präzise Fragestellung, von Richardson, Scott, Wilson, Nishikawa und Hayward (1995, A 12) auch als „well-built clinical question“ bezeichnet, bildet die Grundlage für das Formulieren einer Zielstellung. Darüber hinaus konkretisiert sie auch die betroffene Person bzw. den Personenkreis (z. B. Altersspanne, Geschlecht), die Situation sowie das vorliegende Problem (z. B. aggressives Verhalten seit Trennung der Eltern) und präzisiert damit wiederum die Informationsbasis und Fragestellung für die sich anschließende Literatur- und Studienrecherche (Schlosser & Raghavendra, 2004).

Nachdem im obigen Beispiel die Lehrkraft das Problem differenziert überblickt, stellt sie sich zielorientiert Fragen: Mit welchem Verfahren kann es gelingen, dass die Schüler meiner Klasse schnell und nachhaltig ein altersgerechtes positives Sozialverhalten auf- und ein emotional-soziales Problemverhalten abbauen? Mit welchen Strategien und Maßnahmen können die sozial ausgeschlossenen Kinder in die Klassengemeinschaft integriert werden? Wie können das Klassenklima nachhaltig verbessert und die Lernfreude erhöht werden? Ziel ist es, die Klassensituation insgesamt zu verbessern und eskalierenden Entwicklungsverläufen einzelner Kinder vorzubeugen.

### Vorauswahl an Verfahren vor dem Hintergrund der sozialen und internen Evidenz

Im nächsten Schritt muss die Lehrkraft in einer Vorauswahl unter Berücksichtigung der internen und sozialen Evidenz entscheiden, welche Verfahren überhaupt in Frage kommen, um das anvisierte Ziel zu erreichen.

Die interne Evidenz einer Maßnahme wird durch die fachlich-pädagogische Expertise, d. h. die Berufserfahrung sowie durch das daraus resultierende Professionswissen, gewährleistet. Es ist zu berücksichtigen, mit welchen Konzepten, Verfahren und Materialien die jeweilig handelnde Lehrkraft positive (Selbst-)Wirksamkeitserfahrungen verbindet. Vor diesem Hintergrund spielen auch Aspekte wie die Zugänglichkeit von Verfahren eine Rolle. Ist das Verfahren in der Schule vorhanden bzw. einfach und kostengünstig zu erwerben? Welche Erfahrung hat die Lehrkraft mit einzelnen Verfahren bereits sammeln können? Was haben Kollegen über Durchführungsaspekte und Wirksamkeit berichtet? Erscheint der Vor- und Nachbereitungsaufwand den persönlichen Möglichkeiten angemessen? Mit dem Aspekt der internen Evidenz sind auch die Festlegung der Messverfahren, der Messzeitpunkte sowie die Form der Datensammlung und -auswertung verbunden (Beushausen, 2014). Welche diagnostischen Verfahren sind sinnvoll, um den Verlauf und den Erfolg der Maßnahme zu kontrollieren?

**Die interne Evidenz einer Maßnahme wird durch die fachlich-pädagogische Expertise, d. h. die Berufserfahrung sowie durch das daraus resultierende Professionswissen, gewährleistet. Es ist zu berücksichtigen, mit welchen Konzepten, Verfahren und Materialien die jeweilig handelnde Lehrkraft positive (Selbst-)Wirksamkeitserfahrungen verbindet.**

Weiterhin müssen Kriterien der sozialen Evidenz wie Vorlieben, Werte und Erfahrungen der betroffenen Schüler und ggf. seiner Angehörigen berücksichtigt werden. Durch narrative Methoden (z. B. Gespräche, Interviews) zwischen den Schülern, seinen Erziehungsberechtigten und dem Pädagogen wird gemeinsam das (messbare) Therapieziel möglichst so festgelegt, dass es evaluiert werden kann (Beushausen, 2014; Hartmann, 2013). Wichtig ist, dass die Lehrkraft die Bedürfnisse der Schüler genau kennt und sie in den Entscheidungsprozess einbezieht.

Übertragen auf den pädagogischen Kontext ist bei der Auswahl von Materialien zu prüfen, inwieweit diese kindgemäß gestaltet und für die Arbeit in einer konkreten Lerngruppe geeignet sind. Ist das Verfahren für das Alter der Kinder geeignet? Müssen weitere Personen, z. B. die Erziehungsberechtigten, involviert werden? Lassen sich die Kinder der Klasse der Erfahrung nach mit Freude auf das Material, wie z. B. auf eine Identifikationsfigur, ein oder finden sie das kindisch oder albern? Kann das Verfahren in der sozialen Situation einer Klasse durch-

geführt werden oder müssen auch Kleingruppen oder Einzelförderungen berücksichtigt werden, was sich im Schulalltag nicht ohne erheblichen Aufwand umsetzen lässt? Bei individualisierten Fördermaßnahmen gilt es, zumindest die motivationale Lernausgangslage und weitere individuelle Merkmale des Kinds und ggf. seine Ansichten bzw. die seiner Eltern zu erfassen und zu berücksichtigen.

Nachdem die Lehrkraft die Aspekte der internen und sozialen Evidenz geprüft hat, kann sie eine Vorauswahl zu den in Frage kommenden Verfahren treffen. Diese wird in der Regel nicht sehr umfangreich sein.

Der Lehrkraft im obigen Beispiel, die sich das Ziel gesetzt hat, die sozial-emotionalen Kompetenzen der Schüler ihrer jahrgangübergreifenden Klasse zu fördern, liegen zwei mögliche Verfahren vor, die sowohl versprechen, das angestrebte Ziel zu erreichen, als auch in der Schule vorhanden sind, vom Vorbereitungsaufwand angemessen erscheinen (interne Evidenz), als auch dem Alter und der motivationalen Lage der Klasse entsprechen (soziale Evidenz): „Lubo aus dem All!“ (Hillenbrand et al., 2010, geeignet für Kinder der 1. und 2. Klasse) und das „Verhaltenstraining in der Grundschule“ (Petermann, Koglin, Natzke & von Marées, 2013b; geeignet für Kinder der Klassenstufen 3 und 4). Dabei ist zu beachten, dass in der jahrgangübergreifenden Klasse Kinder der ersten drei Jahrgänge lernen. Eigene Erfahrungen hat die Lehrkraft noch mit keinem der beiden Verfahren sammeln können, weiß aber von Kollegen, dass beide Verfahren von Schülern gerne angenommen werden. Um zu entscheiden, ob „Lubo aus dem All!“ (Hillenbrand et al., 2010) oder das „Verhaltenstraining in der Grundschule“ (Petermann et al., 2013b) das höchstwahrscheinlich wirksamere Verfahren für ihre Klasse ist, schaut sie sich in einem nächsten Schritt die Nachweise zur externen Evidenz beider Verfahren an.

Die Bedeutung einer besonderen Beachtung der externen Evidenz liegt zum einen bei der Auswahl des geeignetsten Verfahrens, wenn mehrere zur Auswahl stehen, die nach der internen und sozialen Evidenz gleichwertig sind. Zum anderen muss beachtet werden, dass die interne und die soziale Evidenz sehr fehleranfällig sein können. Das Erfahrungswissen einzelner, am Entscheidungs- und Interventionsprozess beteiligter Pädagogen reicht als Beleg für die Wirksamkeit einer Maßnahme nicht aus. So können unsystematische Alltagsbeobachtungen, nachträgliche Einschätzungen und erfolgreiche Einzelfälle leicht zu Übergeneralisierungen hinsichtlich der Evidenz einer Interventionsmaßnahme führen, die sich in Studien eventuell nicht zeigen. Daher spielt die externe Evidenz, also die Einschätzung der Wirksamkeitsstudien, eine herausragende Rolle bei der Entscheidung für ein Förderverfahren.

### Externe Evidenz

Die Recherche einzelner Studien, die prüfen, wie wirksam ein Verfahren ist, kann sehr aufwändig sein. Sofern überhaupt vorhanden, müssen mehrere relevante Einzelbeiträge aus unterschiedlichen Quellen (z. B. im Manual des jeweiligen Materials, aus Beiträgen in Fachzeitschriften, Datenbanken und Herausgeberwerken) zusammengetragen werden. Zur Bewertung der einzelnen Studien benötigt die Lehrkraft umfassendes Hintergrundwissen über die Qualität von Studiendesigns und die Bedeutung der statistischen Kennwerte. Sinnvoll ist es, eine Rangfolge der in Frage kommenden Materialien oder Trainingsverfahren nach ihrer externen Evidenz aufzustellen. Diese Rangfolge muss dann vor dem Hintergrund der internen und sozialen Evidenz kritisch geprüft werden, um eine abschließende Entscheidung treffen zu können.

An dieser Stelle ist es wichtig darauf hinzuweisen, dass für eine große Anzahl von Verfahren und Materialien keine Wirksamkeitsstudien vorhanden sind. Dies ist besonders problematisch, da sich Lehrpersonen einer beinahe unüberschaubaren Menge an erhältlichen Materialien, Ansätzen und Programmen gegenübersehen. Während die Forderung nach Evidenzbasierung von Unterricht und Förderung in einigen Ländern gesetzlich verankert ist (z. B. USA; Simpson, LaCava & Graner, 2004), ist der Anspruch nach empirisch belegbarer Wirksamkeit deutscher Unterrichtslehrwerke und schulischer Förderkonzepte derzeit noch keine Realität, sondern eine bloße Forderung (Voss et al., 2015). Das Fehlen eines empirischen Nachweises lässt jedoch

nicht automatisch den Schluss zu, dass ein Förderverfahren, Unterrichtswerk bzw. -konzept nicht wirksam ist. Oftmals lassen sich in einem Lehrwerk oder Förderverfahren Ansätze und Elemente finden, die als nachweislich wirksam gelten oder deren Wirksamkeit empirisch geprüft und belegt wurde. Solange noch keine evidenzbasierten Materialien verfügbar sind, ist jeweils im Einzelfall zu prüfen, inwieweit einzelne empirisch bewährte Elemente in den Materialien enthalten oder mithilfe der Materialien zu realisieren sind.

Liegen für die in Frage kommenden Materialien bereits Ergebnisse aus wissenschaftlichen Studien vor, erfolgt die Einschätzung der externen Evidenz. Da dieser Teilschritt besonders komplex und nicht ganz unstrittig ist, soll er im Folgenden ausführlich erläutert werden.

Zur Einschätzung der externen Evidenz wurden international und national unterschiedliche Klassifikationssysteme und Kriterien für sogenannte Evidenzhierarchien, -level bzw. -stufen erarbeitet. Wegweisend nicht nur für die Medizin ist zweifelsohne die von Alejandro Jadad in Oxford entwickelte Jadad Skala, auch bekannt unter dem Namen Oxford-Skala. Eine direkte Übertragung dieser Evidenzhierarchie oder ihrer Weiterentwicklungen (Oxford Centre for Evidence-Based Medicine, 2001) auf die Pädagogik ist jedoch nur eingeschränkt möglich, da für die pädagogische Praxis relevante Komponenten wie zugrundeliegende wissenschaftliche Theorien und Bedingungen der Umsetzbarkeit keine Berücksichtigung finden. Auf nationaler Ebene haben Hartke (2005) sowie Fingerle und Ellinger (2008) Kriterien zur Einschätzung von Förderprogrammen oder präventiver Maßnahmen erarbeitet. Aus diesen Kriterien heraus ergibt sich folgender Mindeststandard an Evidenz: Materialien, Ansätze oder Maßnahmen sind erst dann „bedingt empfehlenswert“ (Hartke, 2005, S. 17) bzw. „potenziell effektiv“ (Fingerle & Ellinger, 2008, S. 9), wenn sie wissenschaftlich begründet sind und darüber hinaus erste Hinweise auf die Wirksamkeit im Rahmen empirisch angelegter, methodisch sorgfältig geplanter Studien bestehen. Die ermittelten Effekte sollten in der Fachliteratur publiziert sein.

Für die Auswahl von Fördermaterialien gilt somit generell, dass diese im Hinblick auf ihre Wirksamkeit empirisch geprüft wären. Evidenzbasiert ist ein Verfahren letztlich erst dann, wenn es durch ein wissenschaftliches Modell begründet ist und sich in zumindest einer methodisch gut durchgeführten Studie als nachweislich wirksam erwiesen hat.

Noch konkreter und sehr umfassend sind die Auswahl- und Bewertungskriterien für die CTC (Communities that care)-Programm-Datenbank (Groeger-Roth & Hasenpusch, 2011; Veermann & van Yperen, 2007) des Landespräventionsrats Niedersachsen. Die CTC-Programm-Datenbank bewertet Verfahren zur Prävention von Gewalt, Sucht, Kriminalität und Problemverhalten hinsichtlich ihrer Wirksamkeit und der Beweiskraft der für die Wirksamkeitsprüfung durchgeführten Studien. Danach sollten Interventionen nicht punktuell in evidenzbasiert (d. h. nachweislich wirksam) und nicht evidenzbasiert (d. h. nachweislich nicht wirksam) unterschieden werden. Vielmehr wird bei der Beurteilung von Verfahren und Materialien berücksichtigt, dass sich diese oftmals in einer fortlaufenden Entwicklung befinden. So können Verfahren, wenn sie durch zunehmend qualitativ höherwertigere, gut durchgeführte Studien ihre Wirksamkeit nachweisen, in der Evidenzhierarchie aufsteigen. Die meisten Verfahren durchlaufen damit von der Erstveröffentlichung bis zu ihrem Abschluss bzw. zu ihrer Vervollkommnung verschiedene Stufen der Beweiskraft.

***Für die Auswahl von Fördermaterialien gilt somit generell, dass diese im Hinblick auf ihre Wirksamkeit empirisch geprüft wären. Evidenzbasiert ist ein Verfahren letztlich erst dann, wenn es durch ein wissenschaftliches Modell begründet ist und sich in zumindest einer methodisch gut durchgeführten Studie als nachweislich wirksam erwiesen hat.***

Nachfolgend wird ein auf Grundlage der Kriterien der CTC-Programm-Datenbank (Groeger-Roth & Hasenpusch, 2011) weiterentwickeltes Vorgehen zur Beurteilung der externen Evidenz von (sonder-)pädagogisch-psychologischen Materialien, Methoden und Programmen für den

## Evidenzbasierte Auswahl von Fördermaterialien

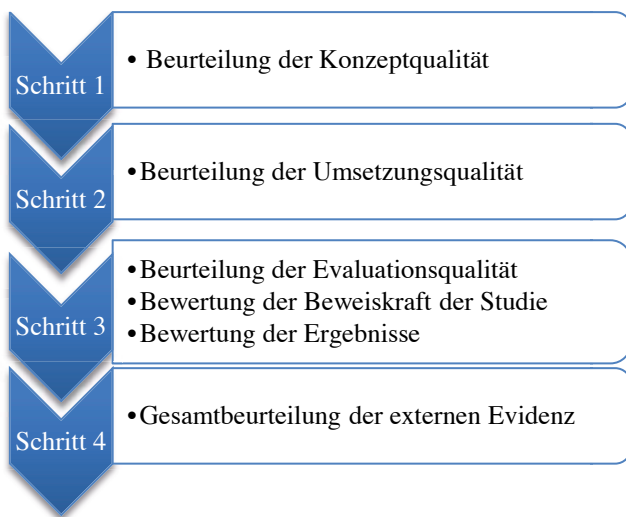


Abbildung 3: Beurteilung der externen Evidenz in vier Schritten

schulischen Kontext differenziert erläutert. Grundlegend folgt die Beurteilung der externen Evidenz vier Schritten: die Beurteilung der Konzeptqualität, die Umsetzungsqualität, die Evaluationsqualität und eine abschließende Gesamtbeurteilung (s. Abbildung 3).

Auf der Grundlage dieses Vorgehens wurde an der Universität Rostock ein Kriterienkatalog zur Beurteilung der externen Evidenz von Trainings-, Förder- und Therapieverfahren entwickelt. In diesem Kriterienkatalog wird auf einer Viererskala die Konzept- und Umsetzungsqualität bewertet.

**Schritt 1: Beurteilung der Konzeptqualität**

Der erste Schritt besteht darin, das dem Material, dem Verfahren bzw. dem Training zugrunde liegende Konzept bzw. die theoretischen Modelle zu beurteilen. Für die schulische Sonderpädagogik

sind folgende Indikatoren für eine angemessene Konzeption bedeutsam (s. Tabelle 1).

1. Konzeptqualität	ist erfüllt	<input type="checkbox"/>	Trifft nicht zu	Trifft eher nicht zu	Trifft eher zu	Trifft zu
	ist nicht erfüllt	<input type="checkbox"/>				
Es liegt ein theoretisch begründetes Modell über die voraussichtliche Wirksamkeit des Verfahrens vor.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die angenommenen Wirkfaktoren sind eindeutig benannt.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die Theorien sind in der Fachwissenschaft bekannt.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die eingesetzten Methoden sind theoretisch gut begründet.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die eingesetzten Instrumente sind theoretisch gut begründet.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Es ist ein logischer Zusammenhang von „Problemanalyse – Einflussfaktoren – Zielen – Zielgruppen – Methoden“ erkennbar.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Das Verfahren bezieht sich auf (aus der Forschung bekannte) Bedingungs-, Risiko- und Schutzfaktoren.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die Zielgruppe(n) ist (sind) präzise beschrieben.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Handlungsanleitungen und Manuale sind nachvollziehbar aus dem Modell abgeleitet.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die Ziele sind explizit formuliert.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die Ziele sind messbar/überprüfbar.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Bei nicht in Deutschland entwickelten Verfahren werden der ursprüngliche Kontext und die Anpassungen beschrieben.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Tabelle 1: Beurteilung der Konzeptqualität einfügen

Bezogen auf das obige Beispiel, in dem die Lehrperson die Förderverfahren „Lubo aus dem All!“ (Hillenbrand et al., 2010) und „Verhaltenstraining in der Grundschule“ (Petermann et al., 2013b) bewertet, zeigt sich in der Konzeptqualität bei beiden Verfahren ein durchgängig positives Bild. Beide Verfahren vermögen allen Kriterien gerecht zu werden und weisen damit eine sehr gute Konzeptqualität auf.

**Schritt 2: Beurteilung der Umsetzungsqualität**

Ein wesentlicher Bestandteil in der Beurteilung eines Materials ist die Qualität der Umsetzung der zugrundeliegenden Konzepte und Ansätze sowie Kriterien, die besonders für die Schulpraxis relevant sind (s. Tabelle 2).

Alle Urheberrechte liegen beim Verband Sonderpädagogik e. V. – Veröffentlichung und Wiedergabe sind nur mit Genehmigung des Rechteinhabers gestattet.

2. Umsetzungsqualität	ist erfüllt	<input type="checkbox"/>	Trifft nicht zu	Trifft eher nicht zu	Trifft eher zu	Trifft zu
	ist nicht erfüllt	<input type="checkbox"/>				
Die eingesetzten Methoden und Instrumente sind didaktisch gut aufbereitet.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die eingesetzten Methoden und Instrumente sind nachvollziehbar beschrieben.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die benötigten Materialien und Manuale sind verfügbar.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die benötigten Materialien und Manuale sind auf einem aktuellen Stand.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Instrumente zur Qualitätskontrolle innerhalb der Umsetzung sind verfügbar.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Einführende Schulungen sind vorhanden.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Begleitende Unterstützungsmaßnahmen, die nach Bedarf bei der Umsetzung angefordert werden können, sind möglich.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Transparenz der anfallenden Gesamtkosten, z. B. durch zusätzliche Kostenaufstellungen, ist gegeben.		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Tabelle 2:  
Beurteilung der Umsetzungsqualität

Ein Vergleich der beiden Verfahren „Lubo aus dem All!“ (Hillenbrand et al., 2010) und „Verhaltenstraining in der Grundschule“ (Petermann et al., 2013b) zeigt leichte Unterschiede. So erscheinen die didaktischen Aufbereitungen bei „Lubo aus dem All!“ (Hillenbrand et al., 2010) tendenziell strukturierter, es braucht weniger Vorbereitungsaufwand, da das meiste Material (Arbeitshefte, Bilder) bereits vorhanden ist. Beim „Verhaltenstraining in der Grundschule“ (Petermann et al., 2013b) müssen die Materialien erst ausgedruckt werden. Positiver ist dort, dass die Materialien insgesamt deutlich preiswerter sind. Trotz der leichten Unterschiede in der Umsetzungsqualität ist diese bei beiden Verfahren im schulischen Kontext erfüllt.

### Schritt 3: Beurteilung der Evaluationsqualität

In Schritt 1 und 2 wurde zunächst geprüft, inwieweit ein Material, Verfahren oder Training die Kriterien der Konzeptqualität vollständig und die Kriterien der Umsetzungsqualität zum größten Teil erfüllt. Ist dies der Fall, gelten die Materialien als wissenschaftsbasiert, da es begründete Hinweise auf die Wirksamkeit der Verfahren gibt. Evidenzbasiert sind Materialien, Verfahren oder Trainingsprogramme jedoch erst dann, wenn sich über die Konzept- und Umsetzungsqualität hinaus Hinweise auf die Wirksamkeit des Verfahrens in Evaluationsstudien zeigen. Liegen Evaluationsstudien für das betreffende Verfahren vor, dann erfolgt die Bewertung der Beweiskraft dieser Evaluationsstudien. Die Beurteilung der Evaluationsqualität ist der wohl schwierigste Schritt bei der Einschätzung der externen Evidenz, bei dem zwei Teilschritte durchlaufen werden müssen.

#### Schritt 3.1: Bewertung der Beweiskraft der Evaluationsstudien

Die Beweiskraft der vorliegenden Evaluationsstudien wird mit Hilfe eines Pfaddiagramms (Blumenthal, Mahlau, Voss & Hartke, 2015), herunterzuladen unter [www.rim.uni-rostock.de](http://www.rim.uni-rostock.de), beurteilt. Hiernach erfolgt die Einschätzung der Studien anhand zentraler Kriterien, wie dem Untersuchungsdesign, der Durchführung im schulischen Kontext, der Stichprobengröße und der ermittelten Effektstärke. Das Pfadmodell führt zu einer Angabe der Beweiskraft der Evaluationsstudien von 0 bis 6 Punkten.

Beim Untersuchungsdesign wird eingeschätzt, ob bereits Metaanalysen vorliegen, deren Befundlage überwiegend homogene Effekte in die gewünschte Richtung belegen. Damit wäre die höchste Beweiskraft gegeben, die mit sechs Punkten bewertet wird. Bei Einzelstudien wird eingeschätzt, ob es sich um ein Kontrollgruppendesign mit mindestens einer Vorher(Prä)- und Nachher(Post)-Messung handelt. Wird diese Frage positiv beantwortet, ist die Randomisierung,



also die zufällige Zuteilung zu den Untersuchungsgruppen von Interesse. Bei einem randomisierten Kontrollgruppendesign mit Prä-Postmessung liegt eine Randomisierte Kontrollgruppenstudie (randomized controlled trial, RCT) vor. Randomized controlled trial gelten als Goldstandard in der empirischen Forschung und werden hinsichtlich ihrer Beweiskraft hoch bewertet. Zudem spielt hier eine wesentliche Rolle, ob eine Nachuntersuchung (ein sogenanntes Follow-Up) stattgefunden hat. Dies gilt auch bei einem nicht randomisierten Kontrollgruppendesign (= quasi-experimentelle Kontrollgruppenstudie) mit einer Prä-Post-Messung. Das Kriterium einer positiven Follow-Up-Erhebung ist sehr bedeutsam, um zu zeigen, ob ein Verfahren auch langfristig wirksam ist.

Neben den Randomisierten Kontrollgruppenstudien (RCT-Studien) und den quasi-experimentellen Studien finden in dem hier erarbeiteten Vorgehen explizit kontrollierte Einzelfallstudien Berücksichtigung. Die Aufnahme begründet sich durch deren Bedeutung in der empirischen sonderpädagogischen Forschung (Baker & McLeod, 2011; Grünke, 2012; Horner, Carr, Halle, Mc Gee, Odom & Wolery, 2005; Julius, Schlosser & Goetze, 1999). Die Sonderpädagogik beschäftigt sich mit sehr vielfältigen Phänomenen, die teilweise sehr spezifische Merkmalskonfigurationen bei Probanden mit sich bringen und nur eine sehr begrenzte Anzahl an Personen betreffen, wie z. B. das Angelman-Syndrom. Zu Studienzwecken können folglich nicht immer hinreichend große Gruppen gebildet, geschweige denn Personen zufällig unterschiedlichen Gruppen zugewiesen, werden. Horner et al. (2005) sind der Ansicht, dass Einzelfallstudien in ihrer Güte mit randomisierten Kontrollgruppenstudien vergleichbar sind, da sie eine experimentelle Kontrollgruppensituation beinhalten (baseline).

Liegen weder Randomisierte Kontrollgruppenstudien, quasi-experimentelle Studien noch Einzelfallstudien vor, wird im Pfadmodell der Oberbegriff anderes Design verwendet. Hierunter sind u. a. folgende Studien zu subsumieren: Zufriedenheitsmessungen, Vorher-Nachher-Messungen ohne Kontrollgruppe oder Ziel-Erreichungsmessungen. Diese werden mit null Punkten bewertet, da die Beweiskraft nicht ausreicht, um die Wirksamkeit einer Maßnahme zu überprüfen. Die niedrigste Bewertung von null Punkten wird auch dann vergeben, wenn keine positiven Ergebnisse ermittelt werden konnten (s. Schritt 3.2).

Im Kontext der Auswahl schulisch einsetzbarer Materialien, Verfahren und Trainingsprogrammen sind Studien dann beweiskräftiger, wenn sie unter schulischen Rahmen- bzw. Alltagsbedingungen stattfinden und relativ große Untersuchungsgruppen berücksichtigen. So kann es sein, dass bei gleichem Stichprobenumfang eine Randomisierte Kontrollgruppenstudie, die weder eine Follow-Up Erhebung ausweisen kann, noch im schulischen Kontext durchgeführt wurde, in der Beweiskraft geringer eingestuft wird (z. B. vier Punkte) als eine quasi-experimentelle Kontrollgruppenstudie mit Follow-Up, die in einem schulischen Setting stattfand (fünf Punkte).

Bedeutsam für die Beurteilung der Beweiskraft der Studie ist eine Einschätzung der berichteten Effektstärken. Hierzu finden sich in der Literatur verschiedene Klassifikationen. Nach Cohen (1992) sind Effektstärken unter .20 nicht relevant, Effektstärken zwischen .20 und .50 deuten auf einen kleinen, zwischen .50 und .80 auf einen mittleren und eine Effektstärke  $\geq$  .80 auf einen großen Effekt hin. Diese Bewertung hat sich in der Human- und Sozialwissenschaft als sinnvolle Einschätzung bei Interventions- und Experimentalstudien bewährt (Bortz & Döring, 2006). In der Bildungsforschung hat Hattie (2013) eine Neuinterpretation der Effektstärken angeregt, indem er den Nullpunkt bei .40 setzt. Demnach sind Effektstärken zwischen 0 und .20 sogenannte „Entwicklungseffekte“, zwischen .20 und .40 „Schulbesuchseffekte“ und erst ab .40 kann von erwünschten Effekten gesprochen werden. Im Bereich der sozialen und emotionalen Entwicklung ist grundlegend von niedrigeren Effektstärken zwischen .20 und .50 auszugehen. Dieser Fakt muss in der Bewertung der Beweiskraft berücksichtigt werden. Die Interpretation der in den einzelnen Studien verwendeten Effektstärkemaßen wird in aller Regel von den Autoren in der Ergebnisdarstellung benannt.

Die Anzahl der Punkte zur Studiengüte ist zudem abhängig von der Stichprobengröße. Da bei kleinen Stichproben ( $N < 20$ ) in der Regel nur große Effekte sichtbar werden, wird für solche Stichprobenumfänge auf eine differenzierte Abfrage zur Effektstärkengröße verzichtet. Im Pfaddiagramm führt der Pfeil von der Stichprobe gleich zur entsprechenden Punktzahl. Anders verhält es sich bei mittleren oder großen Stichprobenumfängen. Bei denen wird zwischen einem kleinen und einem mindestens mittleren Effekt unterschieden. Der Pfeil führt bei mindestens mittleren Stichprobengrößen zum jeweils höheren Punktwert.

Auf der Basis dieses Vorgehens gestaltet sich der Vergleich von Studien, die zur Wirksamkeit der Verfahren „Lubo aus dem All!“ (Hillenbrand et al., 2010) und „Verhaltenstraining in der Grundschule“ (Petermann et al., 2013b) durchgeführt wurden, relativ einfach. Die Autoren beschreiben die durchgeführten Evaluationsstudien detailliert in den entsprechenden Handbüchern bzw. verweisen auf entsprechende Quellen in der Fachliteratur. So kann festgestellt werden, dass die Wirksamkeitsstudie zu „Lubo aus dem All!“ (Hillenbrand et al., 2010) hinsichtlich der Evaluationsqualität eine Beweiskraft von fünf Punkten erhält (quasi-experimentelles Kontrollgruppendesign mit Prä-, Post-, Follow-Up-Messung, durchgeführt unter schulischen Alltagsbedingungen mit mehr als 50 Probanden und überwiegend mittleren Effektstärken). Die Evaluationsstudie(n) zur Wirksamkeit des „Verhaltenstrainings in der Grundschule“ (Petermann et al., 2013b) beinhalten eine Beweiskraft von ebenfalls fünf Punkten (quasi-experimentelles Kontrollgruppendesign mit Prä-, Post-, Follow-Up-Messung, durchgeführt unter schulischen Alltagsbedingungen mit mehr als 50 Probanden, allerdings überwiegend kleinen Effektstärken).

### Schritt 3.2: Beurteilung der Ergebnisse

Nach der Einschätzung der Beweiskraft der Evaluationsstudien, erfolgt nun die Bewertung der Evaluationsergebnisse nach folgenden Richtlinien:

Das Ergebnis der Evaluationsstudie(n)

- ist eindeutig positiv, wenn bei dem/den erfassten Leistungs- bzw. Verhaltensmerkmal/en signifikante Effekte in die gewünschte Richtung auftreten.
- ist überwiegend positiv, wenn bei der Mehrzahl der erfassten Leistungs- bzw. Verhaltensmerkmale signifikante Effekte in die gewünschte Richtung vorliegen.
- enthält keinen Effekt, wenn keine signifikanten Effekte auftreten.
- ist überwiegend negativ, wenn bei der Mehrzahl der erfassten Leistungs- bzw. Verhaltensmerkmale signifikante negative Effekte in die nicht erwünschte Richtung weisen.
- ist negativ, wenn bei dem/den erfassten Leistungs- bzw. Verhaltensmerkmal/en signifikante Effekte in die nicht erwünschte Richtung auftreten.

Die Evaluationsstudien zu „Lubo aus dem All!“ (Hillenbrand et al., 2010) und zum „Verhaltenstraining in der Grundschule“ (Petermann et al., 2013b) zeigen durchgängig eindeutig positive Ergebnisse, die sich jedoch zwischen den Effektstärken unterscheiden.

So konnte in der Evaluationsstudie zu „Lubo aus dem All!“ (Hillenbrand et al., 2010) nachgewiesen werden, dass das Trainingsprogramm den geförderten Kindern hilft, signifikant mehr sozial angemessene Verhaltensstrategien zur Problem- und Konfliktlösung zu entwickeln (großer Effekt). Insbesondere Risikokinder mit einer Disposition zu Verhaltensstörungen sowie Kinder mit leichten intellektuellen Einschränkungen profitierten von der Förderung (ebenfalls großer Effekt). Auch der soziale Status, als Maß der sozialen Integration, verbesserte sich bei den geförderten Kindern im Vergleich zu nichtgeforderten Kindern signifikant (kleiner Effekt). Zudem entwickelte sich das soziale Klima in den geförderten Klassen signifikant positiv. Der Aggressivitätsindex konnte sich bei Risikokindern in der geförderten Gruppe ebenfalls signifikant bei mittlerer Effektstärke verbessern. Nach Einschätzung der Lehrkräfte entwickelte sich das prosoziale Verhalten signifikant positiv (mittlerer Effekt).

Im „Verhaltenstraining in der Grundschule“ (Petermann et al., 2013b) zeigen sich überwiegend kleine Effektstärken zwischen .10 und .43. Die Kinder der Trainingsgruppe zeigten positive Ver-

änderungen in den sozialen Kompetenzen, im prosozialem Verhalten, der Emotionsregulation und auch in den schulischen Fertigkeiten. Etwas weniger effektiv war das Trainingsprogramm in Bezug auf die Abnahme des aggressiven Verhaltens.

**Schritt 4: Gesamtbeurteilung der externen Evidenz**

Die abschließende Beurteilung der Evaluationsergebnisse resultiert aus der Konzept- und Umsetzungsqualität, der Beweiskraft der Evaluationsstudie(n) und den berichteten Effekten.

Wichtig zu erwähnen ist, dass in der sonderpädagogischen Praxis in Deutschland zunehmend Methoden und Maßnahmen Anwendung finden, die im Ausland entwickelt und erfolgreich erprobt wurden (z. B. das Good Behavior Game). Nach Groeger-Roth und Hasenpusch (2011) ist die Beweiskraft für im Ausland entwickelte Verfahren auch für Deutschland gegeben, wenn mehrere Studien mit mindestens hinreichender Beweiskraft (hier mindestens drei Punkte) im deutschsprachigen Raum umgesetzt wurden und vergleichbare Ergebnisse liefern. Liegt keine Studie im deutschsprachigen Raum vor, aber eine sachlogisch nachvollziehbare Begründung über vergleichbare Zielgruppen und Umstände, ist das Material/Verfahren bzw. Training der Evidenzstufe Wissenschaftsbasierte Effektivität zuzuordnen.

Insgesamt ergibt sich aus diesem Vorgehen eine vierstufige Skala zur Beurteilung der externen Evidenz eines Fördermaterials oder Trainingsverfahrens (s. Tabelle 3).

Gesamtbeurteilung der externen Evidenz				Einschätzung: Das Verfahren ist	
Konzeptqualität	Umsetzungsqualität	Beweiskraft (Punkte)	Ergebnisse		
nicht erfüllt	nicht erfüllt	keine Studien	keine Ergebnisse	nicht wissenschaftsbasiert	nicht empfehlenswert
nicht erfüllt	erfüllt	keine Studien	keine Ergebnisse	nicht wissenschaftsbasiert	nicht empfehlenswert
erfüllt	nicht erfüllt	keine Studien	keine Ergebnisse	nicht praktikabel	nicht empfehlenswert
erfüllt	erfüllt	keine Studien	keine Ergebnisse	wissenschaftsbasiert	bedingt empfehlenswert
erfüllt	erfüllt	1 bis 6	negativ	wissenschaftsbasiert, aber eher schädlich	nicht empfehlenswert
erfüllt	erfüllt	1 bis 6	kein Effekt	wissenschaftsbasiert, aber unwirksam	nicht empfehlenswert
erfüllt	erfüllt	1	(überwiegend) positiv	wissenschaftsbasiert und potentiell wirksam	
erfüllt	erfüllt	2	(überwiegend) positiv	= Evidenzstufe 1	empfehlenswert
erfüllt	erfüllt	3	(überwiegend) positiv	wissenschaftsbasiert und wahrscheinlich	
erfüllt	erfüllt	4	(überwiegend) positiv	wirksam = Evidenzstufe 2	empfehlenswert
erfüllt	erfüllt	5	(überwiegend) positiv	höchstwahrscheinlich wirksam = Evidenzstufe 3	empfehlenswert
erfüllt	erfüllt	6	(überwiegend) positiv	nachweislich wirksam = Evidenzstufe 4	sehr empfehlenswert

Tabelle 3:  
Beurteilung der  
Evaluationsergebnisse

Nach diesem Vorgehen ergibt sich für beide Verfahren „Lubo aus dem All!“ (Hillenbrand et al., 2010) und für das „Verhaltenstraining in der Grundschule“ (Petermann et al., 2013b) die Evidenzstufe 3. Beide Verfahren sind also nachgewiesen wirksam und empfehlenswert, wobei hinsichtlich der Effektstärken leichte Vorteile für „Lubo aus dem All!“ (Hillenbrand et al., 2010) bestehen.

**Bewertung der Rangfolge vor dem Hintergrund der sozialen und internen Evidenz und Entscheidung für eine Maßnahme und deren Umsetzung**

Unter Berücksichtigung der sozialen und der internen Evidenz entscheidet die Lehrkraft nun, welches der eingeschätzten Verfahren sie einsetzt. Nicht zwangsläufig muss dies das Verfahren auf dem ersten Rang sein. So könnte es beispielsweise sein, dass das Verfahren mit der höchsten Evidenz einen hohen Anteil an Elternmitarbeit voraussetzt. Die Lehrkraft als Quelle der internen Evidenz hat bei dem jeweiligen Kind jedoch die Erfahrung gemacht, dass die Eltern wenig Bereitschaft zeigen, mit der Schule zu kooperieren und auch das Kind selbst die Zusammenarbeit mit den Eltern nicht stressfrei erleben kann. In diesem Fall würde die Wahl auf das zweitbeste Training in der Rangfolge der externen Evidenz fallen, bei dem der Anteil der Elternarbeit weniger erforderlich ist, um gute Erfolge zu erzielen.



Die Lehrkraft in unserem Beispiel entscheidet sich wegen der besseren Effektstärken (externe Evidenz) und des Umstands, dass sich etwas mehr Schüler ihrer jahrgangübergreifenden Klasse in der ersten und zweiten Klassenstufe befinden (soziale Evidenz) für „Lubo aus dem All!“ (Hillenbrand et al., 2010) und führt das Förderverfahren in der Art und Weise und in dem Zeitrahmen durch, wie es im Manual angegeben wird.

### Evaluation

Grundsätzlich muss sich zum Ende einer Therapie- bzw. Förderphase eine Evaluation der Förderung anschließen. War die Vorgehensweise bei diesem Kind bzw. in dieser Klasse wirksam? Konnte das Verhalten in die gewünschte Richtung verändert werden? Wurde bei den ehemals abgelehnten Schülern ein höherer Grad an sozialem Status erreicht? Um nicht allein auf subjektive Einschätzungen oder Meinungen anderer angewiesen zu sein, ist hierbei der Einsatz von standardisierten Diagnostikverfahren, wie schon unter Identifizierung des Problems benannt, angezeigt. Auf der Basis der Dokumentation der Lern- und Verhaltensentwicklung kann entschieden werden, ob die Maßnahmen sich als wirksam oder als nicht wirksam erwiesen haben. Wenn möglich sollte nach einiger Zeit, z. B. drei Monate nach dem Therapieende, ebenfalls mit standardisierten Testverfahren überprüft werden, ob sich die durch das Verfahren, das Training oder das Material erlernten Verhaltensweisen auch ohne unmittelbare therapeutische Intervention zeigen. Letzteres gibt wertvolle Hinweise auf die Nachhaltigkeit der Wirksamkeit des entsprechenden Verfahrens.

Um die Wirksamkeit des ausgewählten Verfahrens „Lubo aus dem All!“ (Hillenbrand et al., 2010) in ihrer jahrgangübergreifenden Klasse zu prüfen, kann die Lehrkraft beispielsweise wiederum die o. g. Fragebögen bzw. differenzierende Diagnostikverfahren bei einzelnen Schülern einsetzen. Dieses Vorgehen hätte den Vorteil, dass sich die Ergebnisse der Verhaltensänderungen – gemessen mit demselben Diagnostikverfahren vor und nach der Durchführung des Interventionsprogramms – sehr gut miteinander vergleichen ließen. Grundsätzlich eignen sich auch die in einem Testverfahren implementierten Diagnostikverfahren.

Auf der Grundlage der Evaluationsergebnisse entscheidet die Lehrkraft, ob mit dem eingesetzten Verfahren das Ziel erreicht werden konnte. Sollten sich noch keine zufriedenstellenden Ergebnisse zeigen, beginnt der Prozess einer evidenzbasierten Entscheidung für ein angemessenes pädagogisches Vorgehen von neuem.

### Schlussfolgerungen für die Materialauswahl in der schulischen Sonderpädagogik

Grundsätzlich ist es mehr als wünschenswert, wenn sich die Auswahl an Förder-, Therapie- und Trainingsverfahren in der schulischen (Sonder-)Pädagogik nicht am Preis oder an der Werbewirksamkeit der Verlage orientiert, sondern an der Effektivität der einzelnen Verfahren. Die nachgewiesene Effektivität der Förder-, Trainings- oder Therapieverfahren erhöht die Wahrscheinlichkeit, Kinder mit besonderen Risikokonstellationen schnell, wirksam und nachhaltig zu fördern.

Abschließend soll darauf verwiesen werden, dass es im deutschsprachigen Raum nur sehr wenige im unterrichtlichen Kontext erprobte und durch hochwertige wissenschaftliche Studien evaluierte Therapie- und Förderverfahren gibt. Die hier beispielhaft betrachteten Verfahren „Lubo aus dem All!“ (Hillenbrand et al., 2010) und „Verhaltenstraining in der Grundschule“ (Petermann et al., 2013b) stellen hinsichtlich ihrer Güte eine positive Ausnahme dar. Betrachtet man verschiedene, im Handel erhältliche Verfahren und deren Evidenz aus den einzelnen sonderpädagogischen Förderbereichen, wird deutlich, dass aktuell eine geringe Auswahl an Trainings-, Förder- und Therapie-

### Schlüsselbegriffe

Evidenzbasierte Praxis, externe Evidenz, interne Evidenz, soziale Evidenz, Trainings- und Fördermaterialien

### Abstract

Education and special education should be organised in a way that enables students to achieve a specific learning or development goal quickly, certainly and sustainably. For the practical implementation at school, experiences of teachers, teaching materials and preferences of the students play a central role. The following article focuses on the question whether this approach could be modified in favour of a more effective support. Thus, the principle of ‘evidence-based practice’ relates to the content of special education. A specific procedure for selecting evidence-based teaching materials is explained by using a stage model and an example of the behavioural development as an illustration.

### Keywords

evidence-based practice, external evidence, internal evidence, social evidence, support materials

## Literatur

verfahren mit einer hohen externen Evidenz (Stufe 3 und Stufe 4) für den schulischen Kontext besteht. Evaluationsstudien an hinreichend großen Stichproben, im Kontext eines schulischen Alltagssettings sind daher unbedingt erforderlich, auch wenn sich damit die grundlegenden Limitationen einer evidenzbasierten Forschung im Kontext der pädagogischen Wissenschaft (Schad, 2015) nicht vollständig überwinden lassen.

- Arbeitsgruppe Deutsche Child Behavior Checklist (1993). *Lehrerfragebogen über das Verhalten von Kindern und Jugendlichen*. Deutsche Bearbeitung Teacher's Report Form (TRF) der Child Behavior Checklist. Einführung und Anleitung zur Handanweisung, (bearbeitet von Döpfner, M. & Melchers, P.). Köln: Arbeitsgruppe Kinder-, Jugend- und Familiendiagnostik (KJFD).
- Baker, E. & McLeod, S. (2001). Evidence-Based Practice for Children With Speech Sound Disorders: Part 1 Narrative Review. *Language, Speech, and Hearing Services in Schools*, 42, S. 102–139.
- Beushausen, U. (2014). Chancen und Risiken einer evidenz-basierten Sprachtherapie. *Logos*, 22, 96–194.
- Blumenthal, Y., Mahlau, K., Voss, S. & Hartke, B. (2015). *Pfadmodell zur Einschätzung der externen Evidenz von Trainings-, Förder- und Therapieverfahren*. Verfügbar unter: [http://www.rim.uni-rostock.de/uploads/media/PfadmodellExterneEvidenz\\_BlumenthalMahlauVossHartke.pdf](http://www.rim.uni-rostock.de/uploads/media/PfadmodellExterneEvidenz_BlumenthalMahlauVossHartke.pdf). [28.05.2015]
- Blumenthal, Y. & Marten, K. (2014). *Soziale Integration von abgelehnten Kindern in einer Schulklasse – Diagnostik und Interventionen*. In Y. Blumenthal, K. Mahlau & B. Hartke (Hrsg.), *Lernen nachhaltig fördern Klasse 4*. (S. 141–173). Rostock: Universität Rostock.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4. Aufl.). Berlin: Springer.
- Cierpka, M. (2001). *FAUSTLOS. Ein Curriculum zur Prävention von aggressivem und gewaltbereitem Verhalten bei Kindern der Klassen 1 bis 3*. Göttingen: Hogrefe.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, S. 155–159.
- Dollaghan, C. (2007). *The handbook for evidence-based practice in communication disorders*. Baltimore: Paul H. Books.
- Fingerle, M. & Ellinger, S. (2008). *Sonderpädagogische Förderprogramme im Vergleich*. Orientierungshilfen für die Praxis (S. 9-10). Stuttgart: Kohlhammer.
- Gasteiger-Klicpera, B. & Klein, G. (2006). *Das Friedensstifter Training. Grundschulprogramm zur Gewaltprävention*. München: Reinhardt.
- Goodman, R. (2005). *Fragebogen zu Stärken und Schwächen (SDQ-Deu)*. Verfügbar unter <http://www.sdqinfo.com/> [16.03.2015]
- Groeger-Roth, F. & Hasenpusch, B. (2011). *Grüne Liste Prävention. Auswahl- und Bewertungskriterien für die CTC-Programm-Datenbank*. Verfügbar unter [http://www.gruene-liste-praevention.de/communities-that-care/Media/\\_Grne\\_Liste\\_Kriterien.pdf](http://www.gruene-liste-praevention.de/communities-that-care/Media/_Grne_Liste_Kriterien.pdf) [16.03.2015]
- Grünke, M. (2012). Auswertung von Daten aus kontrollierten Einzelfallstudien mit Hilfe von Randomisierungstests. *Empirische Sonderpädagogik*, 3/4, S. 247–264.
- Hartke, B. (2005). *Schulische Prävention - welche Maßnahmen haben sich bewährt?* In S. Ellinger & M. Wittrock (Hrsg.), *Sonderpädagogik in der Regelschule*. Konzepte, Forschung, Praxis (S. 11-37). Stuttgart: Kohlhammer.
- Hartke, B. & Vrbán, R. (2014). *Schwierige Schüler: 49 Handlungsmöglichkeiten bei Verhaltensauffälligkeiten* (9. Aufl.). Buxtehude: Persen.
- Hartmann, E. (2013). Evidenzbasiertes Denken und Handeln in der Logopädie/Sprachheilpädagogik. State of the Art und Perspektiven. *Vierteljahreszeitschrift für Heilpädagogik*, 4, S. 339–343.
- Hattie, J. (2013). *Lernen sichtbar machen*. Überarbeitete deutschsprachige Ausgabe von „Visible Learning“. Hohengehren: Schneider.
- Hillenbrand, C. (2014). *Evidenzbasierte Praxis im Förderschwerpunkt emotional-soziale Entwicklung*. In R. Stein & T. Müller (Hrsg.), *Inklusion im Förderschwerpunkt emotional-soziale Entwicklung* (S. 170–215). Stuttgart: Kohlhammer.
- Hillenbrand, C., Hennemann, T. & Hens, S. (2010). *Lubo aus dem All! Programm zur Förderung emotional-sozialer Kompetenzen in der Schuleingangsphase*. München: Reinhardt.

- Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odo, S. & Wolery, M. (2005). The Use of Single-Subject Research to Identify Evidence-Based Practice in Special Education. *Council for Exceptional Children*, 71 (2), S. 165–179.
- Julius, H., Schlosser, R.W. & Goetze, H. (1999). *Kontrollierte Einzelfallforschung*. Göttingen: Hogrefe.
- Mangold, S. (2013). *Evidenzbasiertes Arbeiten in der Physio- und Ergotherapie*. Berlin: Springer.
- Nussbeck, S. (2007). Evidenzbasierte Praxis – ein Konzept für sonderpädagogisches Handeln? *Sonderpädagogik*, 37 (2/3), S. 146-155.
- Nussbeck, S. (2013). *Evidenzbasierte Praxis*. In U. Heimlich, R. Stein & F. B. Wember (Hrsg.), *Handlexikon Lernschwierigkeiten und Verhaltensstörungen* (S. 247–249). Stuttgart: Kohlhammer.
- Oxford Centre for Evidence-Based Medicine (2001). *Evidenzhierarchie des Oxford Centre for Evidence-Based Medicine*. Verfügbar unter <http://www.cebm.net/> [15.03.2015]
- Petermann, F., Koglin, U., Natzke, H. & Marées, N. von (2013b). *Verhaltenstraining in der Grundschule. Ein Präventionsprogramm zur Förderung emotionaler und sozialer Kompetenzen* (2. überarbeitete Aufl.). Göttingen: Hogrefe.
- Petermann, F., Natzke, H., Gerken, N. & Walter, H. J. (2013a). *Verhaltenstraining für Schulanfänger. Ein Programm zur Primärprävention von aggressivem und unaufmerksamem Verhalten bei Kindern in den ersten beiden Grundschulklassen* (3. Aufl.). Paderborn: Schöningh UTB.
- Rauer, W. & Schuck, K. D. (2003). *Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern dritter und vierter Klassen (FEES 3–4)*. Göttingen: Hogrefe.
- Rauer, W. & Schuck, K. D. (2004). *Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern erster und zweiter Klassen (FEES 1–2)*. Göttingen: Hogrefe.
- Richardson, W., Scott, M. D., Wilson, M. C., Nishikawa, J. & Hayward, R. S. A. (1995). *The well built clinical question: A key to evidence based decisions (Editorial)*. *American College of Physicians Journal Club*, 123, November–December, A-12.
- Sackett, D. L., Straus, S. E., Richardson, W. S., Rosenberg, W. & Haynes, R. B. (2000). *Evidence-based medicine: how to practice and teach EBM*. London: Churchill-Livingstone.
- Schad, G. (2015). Evidenzbasierte Erziehung? *Zeitschrift für Heilpädagogik*, 66, S. 335–344.
- Simpson, R. L., LaCava, P. G., & Graner, P. S. (2004). The no child left behind act: Challenges and implications for educators. *Intervention in School & Clinic*, 40 (2), S. 67–75.
- Schlosser, R. W. & Raghavendra, P. (2004). Evidence-Based Practice in Augmentative and Alternative Communication. *Augmentative and Alternative Communication*, 20 (1), S. 1–21.
- Veermaann, J. W. & van Yperen, T. A. (2007). Degrees of freedom and degrees of certainty: A developmental model for the establishment of evidence-based youth care. *Evaluation and Program Planning*, 30, S. 212–221.
- Voss, S., Sikora, S. & Hartke, B. (2014). *Kriterienliste zur wissenschaftlich begründeten Materialauswahl im Mathematikunterricht der Grundschule*. Verfügbar unter [http://www.rim.uni-rostock.de/uploads/media/KriterienkatalogZurMaterialauswahl-Mathematik\\_VossSikoraHartke.pdf](http://www.rim.uni-rostock.de/uploads/media/KriterienkatalogZurMaterialauswahl-Mathematik_VossSikoraHartke.pdf) [12.03.2015]
- Voss, S., Sikora, S. & Hartke, B. (2015). Was heißt hier Evidenzbasiert? – Kriterien zur wissenschaftlich begründeten Auswahl von Materialien für den Mathematikunterricht in der Grundschule. *Zeitschrift für Heilpädagogik*, 66, S. 85–101.

Dr. Yvonne Blumenthal  
 Institut für Sonderpädagogische Entwicklungsförderung und Rehabilitation  
 Universität Rostock  
 August-Bebel-Straße 28, 18051 Rostock  
 E-Mail: [yvonne.blumenthal@uni-rostock.de](mailto:yvonne.blumenthal@uni-rostock.de)

Dr. Kathrin Mahlau  
 Institut für Sonderpädagogische Entwicklungsförderung und Rehabilitation  
 Universität Rostock  
 August-Bebel-Straße 28, 18051 Rostock  
 E-Mail: [kathrin.mahlau@uni-rostock.de](mailto:kathrin.mahlau@uni-rostock.de)

