



# Sozio-emotionales Lernen

Pädagogik sozio-emotionaler  
Entwicklungsförderung

Dennis Christian Hövel, Claudia Schellenberg,  
Pierre-Carl Link, Olivia Gasser-Haas (Hrsg.)



Dennis Christian Hövel, Claudia Schellenberg,  
Pierre-Carl Link, Olivia Gasser-Haas (Hrsg.)

## **Sozio-emotionales Lernen**

Pädagogik sozio-emotionaler  
Entwicklungsförderung



Dennis Christian Hövel, Claudia Schellenberg,  
Pierre-Carl Link, Olivia Gasser-Haas (Hrsg.)

# **Sozio-emotionales Lernen**

Pädagogik sozio-emotionaler  
Entwicklungsförderung

© 2024

Edition SZH/CSPS

Stiftung Schweizer Zentrum für Heil- und Sonderpädagogik (SZH) Bern  
Fondation Centre suisse de pédagogie spécialisée (CSPS) Berne  
Fondazione Centro svizzero di pedagogia specializzata (CSPS) Berna  
Fundaziun Center svizzer da pedagogia speciala (CSPS) Berna



Diese Publikation wurde gefördert durch den SNF – Schweizerischer Nationalfonds zur Förderung der wissenschaftlichen Forschung.

Coverbild: Fatima Mohamed (Illustration) und Maja Davé (Grafik)

Layout: Weber Verlag AG

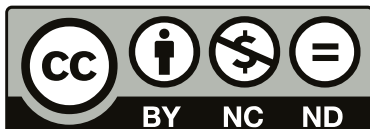
Lektorat: Edition SZH/CSPS

ISBN Print on demand: 978-3-905890-92-1

ISBN E-Book: 978-3-905890-91-4

Die Verantwortung für den Inhalt der Texte liegt bei der jeweiligen Autorin/beim jeweiligen Autor.

Dieses Werk ist lizenziert unter einer Creative Commons Lizenz 4.0 International (BY-NC-ND).



Edubook stellt alle Produkte ausschliesslich in der Schweiz her;  
dies vollumfänglich klimaneutral und auf FSC®-zertifizierten Papieren.



# Inhaltsverzeichnis

<b>I</b>	<b>Vorwort</b>	<b>9</b>
<b>II</b>	<b>Einführung und Aufbau des Buches</b>	<b>11</b>
<b>III</b>	<b>Fallvignetten</b>	<b>17</b>
	<b>Fallvignette 1: Lara</b>	<b>19</b>
	Christina Bär, Xenia Müller, Martin Jany und Max Schläfli Bieri	
	<b>Fallvignette 2: Louis</b>	<b>31</b>
	Xenia Müller, Christina Bär, Max Schläfli Bieri und Martin Jany	
	<b>Fallvignette 3: Martina</b>	<b>39</b>
	Martin Jany, Max Schläfli Bieri, Christina Bär und Xenia Müller	
<b>1</b>	<b>Grundlagen einer Pädagogik sozio-emotionaler Entwicklungsförderung</b>	<b>49</b>
	<b>1.1 Sozio-emotionales Lernen (SEL) für inklusive Bildung</b>	<b>51</b>
	Dennis C. Hövel	
	<b>1.2 Mehrstufige Förderung am Beispiel des «Schoolwide Positive Behaviour Support» SWPBS</b>	<b>63</b>
	Dennis C. Hövel, Anja Solenthaler, Annette Krauss, Pierre-Carl Link und Fabio Sticca	
	<b>1.3 Zur normativen Reichweite einer Pädagogik sozio-emotionaler Entwicklung</b>	<b>79</b>
	Pierre-Carl Link und Verena Muheim	

<b>1.4 Häufigkeit sozio-emotionaler Problemlagen</b>	<b>89</b>
Annette Krauss und Claudia Schellenberg	
<b>1.5 Ätiologie sozio-emotionaler Kompetenzen und deren Bedeutung für den Bildungsauftrag</b>	<b>99</b>
Olivia Gasser-Haas, Claudia Schellenberg, Pierre-Carl Link und Dennis C. Hövel	
<b>1.6 Stärkung sozio-emotionaler Kompetenzen im schulischen Unterricht</b>	<b>115</b>
Olivia Gasser-Haas und Claudia Schellenberg	
<b>1.7 Pädagogisch-therapeutische Auftrags- und Zielumsetzung</b>	<b>127</b>
Olivia Gasser-Haas, Anja Solenthaler, Lucia Maier Diatara und Pierre-Carl Link	
<b>1.8 Zusammenarbeit im multiprofessionellen Team</b>	<b>137</b>
Susan C. A. Burkhardt, Markus Matthys und Claudia Schellenberg	
<b>1.9 Transition – inklusive Bildung beim Übergang Schule–Beruf</b>	<b>147</b>
Claudia Schellenberg, Annette Krauss und Patrizia Rösli	
<b>2 Diagnostik im Bereich sozio-emotionaler Entwicklung</b>	<b>159</b>
<b>2.1 Klassifikationssysteme und deren Bedeutung für die Pädagogik</b>	<b>161</b>
Dennis C. Hövel, Noëlle Behringer und Pierre-Carl Link	
<b>2.2 Diagnose-Förderprozess</b>	<b>173</b>
Dennis C. Hövel	
<b>2.3 Statusdiagnostik</b>	<b>185</b>
Fabio Sticca, Anja Solenthaler, Patrizia Rösli, Markus Matthys und Dennis C. Hövel	
<b>2.4 Verlaufsdiagnostik</b>	<b>197</b>
Dennis C. Hövel, Jürgen Kohler und Fabio Sticca	



<b>2.5 Diagnostik als multiprofessionelle Zusammenarbeit</b>	<b>207</b>
Melanie Nideröst, Anja Solenthaler, Christina Bär und Xenia Müller	
<b>3 Pädagogische und therapeutische Handlungsansätze</b>	<b>219</b>
<b>3.1 Lerntheoretische und kognitionspsychologische Ansätze in der Förderung sozio-emotionaler Kompetenzen</b>	<b>221</b>
Annette Lütolf Belet, Susan C. A. Burkhardt und Dennis C. Hövel	
<b>3.2 Psychodynamische Handlungsansätze: Bindung und Mentalisierung</b>	<b>233</b>
Alex Neuhauser, Patrizia Rööslü, Verena Muheim, Iris Bräuninger und Pierre-Carl Link	
<b>3.3 Personenzentrierte, körperorientierte und systemische Ansätze</b>	<b>245</b>
Claudia Schellenberg, Margaretha Florin, Lucia Maier, Ilona Widmer, Thomas Lustig und Iris Bräuninger	
<b>IV Fazit einer Pädagogik sozio-emotionaler Entwicklungsförderung</b>	<b>257</b>
Claudia Schellenberg, Olivia Gasser-Haas, Dennis C. Hövel und Pierre-Carl Link	
<b>V Verzeichnis der Autorinnen und Autoren</b>	<b>263</b>

## 2.4 Verlaufsdagnostik

Dennis C. Hövel, Jürgen Kohler und Fabio Sticca

Ein zentrales Anliegen im Bereich der Professionalisierung in der Heil- und Sonderpädagogik ist die Evaluation der Wirksamkeit von durchgeführten Massnahmen. Besonders relevant ist die Möglichkeit, die Wirksamkeit einer Massnahme für einen Einzelfall zu evaluieren. Hierzu eignet sich die in diesem Kapitel vorgestellte Methode der *Verlaufsdagnostik*. Die Verlaufsdagnostik ist die Anwendung einer Methode aus den Sozialwissenschaften, nämlich die der kontrollierten Einzelfallstudie. Ziel des Kapitels ist es, der praktisch arbeitenden Fachperson eine Handlungsorientierung für die Durchführung einer kontrollierten Einzelfallstudie zu bieten. Der Bezug zur Praxis wird dabei fortlaufend hergestellt. Die Botschaft lautet: Die Wirksamkeit einer pädagogisch-therapeutischen Massnahme kann in der Praxis an einem Einzelfall aussagekräftig dargestellt werden.

Im Folgenden wird zuerst das Konzept einer Einzelfallstudie geschildert, gefolgt von einigen Variationen in der praktischen Umsetzung. Anschliessend wird die Methode der direkten Verhaltensbeurteilung als konkreter Anwendungsbereich vorgestellt. Abschliessend wird das Konzept der kontrollierten Einzelfallstudie in das Konstrukt der evidenzbasierten Praxis eingebettet, um die Bedeutung dieser Methode für die Professionalisierung der Heil- und Sonderpädagogik zu unterstreichen.

### **Einzelfallstudien**

Bei der Verlaufsdagnostik stehen Einzelfälle im Fokus. Einzelfallstudien können unterschiedlich angelegt sein (Petermann, 1996). Zunächst kann differenziert werden zwischen qualitativen Einzelfallstudien (Flick et al., 2012) und kontrollierten quantitativen Einzelfallstudien (Julius et al., 2000; Reicherts & Genoud, 2015). Letztere werden etwas vereinfachend auch als kontrollierte Einzelfallstudien bezeichnet. Während die qualitativen Einzelfallstudien der Hypothesenbildung dienen, eignen sich kontrollierte Einzelfallstudien für die Hypothesenprüfung. Im Rahmen dieses Kapitels konzentrieren wir uns auf die kontrollierte Einzelfallstudie.

### ***Die kontrollierte Einzelfallstudie***

Ein wichtiges Merkmal kontrollierter Einzelfallstudien ist die hochfrequente Messung (z. B. täglich oder mehrmals täglich) des Zielkonstrukts über einen längeren Zeitraum. Die Messung beginnt schon, bevor die Massnahme implementiert wird, nämlich in der sogenannten *A-Phase* oder *Baseline-Phase*. Die A-Phase ist somit eine Kontrollbedingung für den Einzelfall und dient dazu, den Verlauf

des Zielkonstrukts vor dem Beginn der Massnahme sichtbar zu machen. Mit dem Beginn der Massnahme wird die *B-Phase* der Messung initiiert. Forschungsmethodisch ist die Massnahme die *unabhängige Variable* (UV) und das Zielkonstrukt die *abhängige Variable* (AV). Im Vergleich der zeitlichen Verläufe der abhängigen Variablen lässt sich der Zusammenhang zwischen der Massnahme und dem Verlauf der Zielkonstrukte ermitteln.

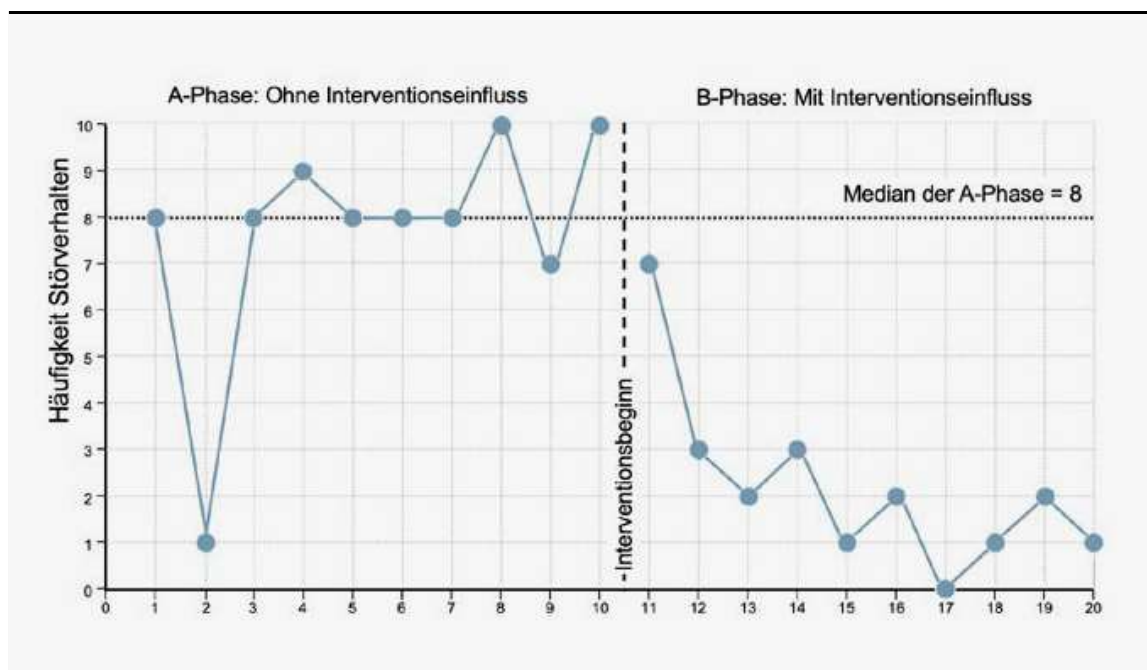
Die Frequenz der Messung hängt ab von der Variabilität des Zielkonstrukts selbst sowie von der Abwägung zwischen Präzision und Umsetzbarkeit der Messung. So kann beispielsweise die Aggression einer Schülerin mehrmals pro Tag nach jeder Lektion, jeweils täglich oder wöchentlich erfasst werden. Je höher die Frequenz angesetzt wird, desto höher ist die Präzision und desto niedriger die Machbarkeit (insbesondere auch über längere Zeiträume).

Das Ziel dieser Methode ist es, Unterschiede zwischen A-Phase und B-Phase festzustellen, welche wiederum auf die Massnahme zurückzuführen sind. Um dies zu bewerkstelligen, erfordert die kontrollierte Einzelfallstudie einen hohen Grad der Strukturierung (Julius et al., 2000). Entsprechend ist die Messung vorweg detailliert zu planen. Der Fokus liegt auf genau definierten Merkmalen oder Verhaltensweisen, die strukturiert gemessen werden (Döring & Bortz, 2016). Als Befragungsformen kommen zum Beispiel schriftliche Befragungen mit eher geschlossenen Antwortformaten oder stark strukturierte Interviews zum Einsatz. Auch Beobachtungen sind hochstrukturiert und mit möglichst klaren und eindeutigen Operationalisierungen des interessierenden Verhaltens versehen. Je nach Zielkonstrukt können auch Teile von Testverfahren und stark strukturierte Aufgaben eingesetzt werden. Wenn beispielsweise Aggression eingeschätzt wird, wird anhand von bestehenden Fragebögen und / oder Beobachtungsverfahren erarbeitet, was in Bezug auf den Einzelfall mit Aggression gemeint ist und wie diese durch die Fachperson mit einer oder einigen wenigen Fragen (z. B. einmalig am Ende des Tages) erfasst werden kann. Das Konstrukt «Aggression» kann hierbei sowohl im Sinne der Stärke ihrer Ausprägung (Qualität) als auch im Sinne der Häufigkeit ihres Auftretens (Quantität) erfasst werden. Als Beurteilungsskala bietet sich zum Beispiel eine Skalierung der Stärke der Aggression (von 0 *sehr schwache Aggression* bis 10 *sehr starke Aggression*) oder der Häufigkeit (von 0 *keine Aggression* bis 10 *sehr häufige Aggressionen*) an. Die Angaben werden in einem einfachen Protokoll festgehalten, dies kann sowohl analog als auch digital sein. Wichtig ist dabei, die geplanten Messungen darzustellen (z. B. tabellarisch) und die Angaben auch im geplanten Rhythmus zu tätigen. Sollte eine geplante Angabe nicht möglich sein, wird das dafür vorgesehene Feld leergelassen, sodass der fehlende Wert ersichtlich wird.

Die Daten werden anschliessend mit statistischen Methoden analysiert (Kazdin, 2011). Dabei wird vor allem geprüft, ob sich bedeutsame Unterschiede in den Messungen zwischen A-Phase und B-Phase ergeben. Die Ergebnisse der kontrollierten Einzelfallstudie gelten zunächst nur für den einen Fall. Jedoch hat die kontrollierte Form der Einzelfallanalyse grosses Potenzial, um Ursache-Wirkung-Zusammenhänge darzustellen. Das liegt zum einen daran, dass sich kontrollierte Einzelfallstudien auf wenige Zielkonstrukte konzentrieren, die besser kontrollierbar sind als viele miteinander interagierende Variablen. Es ist aber auch möglich, dass mehrere Variablen nebeneinander beziehungsweise nacheinander kontrolliert werden (kontrollierte Einzelfallstudie im *Multiple-Baseline-Design*). Zum andern können Ergebnisse mehrerer kontrollierter Einzelfallstudien zusammengetragen und gesamthaft ausgewertet werden. Dies ermöglicht es, Aussagen zu treffen, die über den Einzelfall hinausgehen.

Das oben vorgestellte Vorgehen wird im Folgenden am Beispiel des Störverhaltens im Unterricht veranschaulicht. Die Fachperson hat festgelegt, welche Verhaltensweisen im Unterricht als störend zu verstehen sind (z. B. «Steht auf und läuft während des Unterrichts herum»). Sie beobachtet das Störverhalten der Schülerin im Unterricht vor einer Massnahme zwei Wochen lang und nach Beginn der Massnahmen weitere zwei Wochen lang. Die Fachperson hält das Störverhalten auf einer Skala von 0 bis 10 am Ende des Tages fest. Die Daten sind in Abbildung 1 dargestellt.

Abbildung 1: Prototypischer Aufbau einer einfachen quantitativ-kontrollierten Einzelfallstudie (Jain & Spiess, 2012, S. 214)



Im Beispiel zielt die Intervention der Fachperson auf eine Verbesserung des Störverhaltens einer Schülerin ab. Die Abbildung zeigt auf, dass das Störverhalten vor Beginn der Massnahme meist sehr häufig vorkam und dass es mit Beginn der Massnahme sichtbar abgenommen hat. Da es sich hierbei um dieselbe Person handelt und sowohl die A-Phase als auch die B-Phase eine hohe Menge an Messungen enthalten, sind erste Hinweise dafür gegeben, dass die Massnahme erfolgreich war. Weitere Berechnungen zur Quantifizierung werden weiter unten erläutert.

### ***Variationen der kontrollierten Einzelfallstudie***

Es gibt viele Variationen des in Abbildung 1 veranschaulichten einfachen A-B-Versuchsplanes der quantitativen Einzelfallstudie. Davon werden im Folgenden drei genauer ausgeführt:

1. Der *A-B-A-Versuchsplan* misst die abhängige Variable (z. B. das Störverhalten) auch eine Zeit lang nach Beenden der Massnahme. Damit lässt sich ein Langzeit- oder Transfereffekt ermitteln. Im Beispiel liesse sich die Wirkung der Massnahme in Form einer Linie darstellen, die auch dann noch wenig Störverhalten anzeigt, wenn die Massnahme bereits beendet wurde.
2. Die oben schon erwähnte *Einzelfallstudie im Multiple-Baseline-Design* versucht nicht nur eine, sondern mehrere Variablen einzubeziehen. Das ist sehr praxisnah, denn oftmals interessiert nicht nur eine Variable, sondern mehrere. So könnte im angeführten Beispiel die Hoffnung bestehen, dass die schulischen Leistungen sich im Zuge der Massnahme ebenfalls verbessern. Dann hätte man eine Variablenvielfalt aufseiten der abhängigen Variablen. Oder man bezieht neben der implementierten Massnahme auch noch weitere Massnahmen ein. Beispielsweise könnte neben einem Training der Emotionsregulation auch ein Aufmerksamkeitstraining durchgeführt werden. Dann hätte man eine Variablenvielfalt aufseiten der unabhängigen Variablen. Die Eindeutigkeit des Ursache-Wirkung-Zusammenhangs wird allerdings bei zunehmender Variablenvielfalt immer weniger gut identifizierbar.
3. Bei der *Einzelfallstudie im psychometrischen Design* wird das Zielkonstrukt durch ein Testverfahren je einmal vor und einmal nach der Massnahme erfasst (anstelle vieler Messzeitpunkte). Damit lässt sich der Ursache-Wirkung-Zusammenhang möglichst sicher nachvollziehen. Dieses Vorgehen hat den Vorteil, dass nur zwei Messzeitpunkte erforderlich sind. Gleichzeitig hat es aber auch den Nachteil, dass die Testverfahren an sich meist aufwendiger sind als die beiden oben beschriebenen Prozeduren. Der Vergleich geschieht zwischen zwei Testergebnissen, wobei die Testergebnisse (Testscores) auch einen Vergleich mit einer Normstichprobe umfassen. Zudem beinhalten die

Testwerte durch den hohen Grad an theoretischer Fundierung und Strukturiertheit von Tests auch ein Versprechen zu deren Güte (zu den Gütekriterien Validität, Reliabilität und Objektivität vgl. Kap. 2.3).

Allen Variationen experimenteller Designs (auch den hier angeführten kontrollierten Einzelfallstudien) ist gemein, dass sie einen Ursache-Wirkung-Zusammenhang immer nur mit einer gewissen Irrtumswahrscheinlichkeit abbilden können. In einem Laborexperiment ist die Irrtumswahrscheinlichkeit in der Regel kleiner als in einem Feldexperiment. Trotzdem gibt es auch in der laborähnlichen Situation (wie bei standardisierten Tests) Fehlerquellen.

Im Allgemeinen zeichnet sich die kontrollierte Einzelfallstudie durch eine sorgfältige Datenerhebung aus, bei der geschlossene und hochstrukturierte Methoden angewendet werden. In diesem Kontext werden Befragungen und Beobachtungen als zentrale Instrumente eingesetzt, wobei die Auswahl zwischen einem Fragebogen und standardisierten Aufgaben oder Tests erfolgt. Die Datenauswertung geschieht auf mehreren Ebenen. Es wird eine visuelle Analyse angewendet, die es ermöglicht, Muster und Trends in den gesammelten Daten zu identifizieren. Gleichzeitig wird eine quantitative Auswertung durchgeführt, um präzise und objektive Ergebnisse zu erhalten. Die Dateninterpretation in einer kontrollierten Einzelfallstudie konzentriert sich auf idiografische Ansätze und einzelne Zielkonstrukte. Die Ergebnisse sind dabei nur begrenzt auf andere Fälle verallgemeinerbar. Die Methode zielt darauf ab, Ursache-Wirkung-Zusammenhänge zu beschreiben und zu erklären, wobei ein besonderes Augenmerk auf die hypothesenprüfende Methode durch ein quasi-experimentelles Design gelegt wird. Ein weiteres wichtiges Merkmal ist die hohe Reliabilität der Ergebnisse, was auf die Genauigkeit und Zuverlässigkeit der erhobenen Daten hinweist.

### **Direkte Verhaltensbeurteilung**

Die direkte Verhaltensbeurteilung (DVB) ist eine Kombination aus systematischer Verhaltensbeobachtung und Verhaltensbeurteilung (Huber & Rietz, 2015). Eine Fachperson beobachtet ein bestimmtes Verhalten einer Person in einem klar definierten Zeitfenster und beurteilt direkt im Anschluss dieses Verhalten. Zur Bestimmung von Items für eine direkte Verhaltensbeurteilung können etablierte Beurteilungsskalen genutzt werden. Ein Beispiel hierfür wäre der Fragebogen *Schulische Einschätzung des Verhaltens und der Entwicklung* (SEVE) von Hartke et al. (2020). Der Bogen umfasst 91 Items (u. a. zum Bereich Arbeitsverhalten; Beispielitem «Arbeitet durchgehend konzentriert»; ebd., S. 13). Die Aussagen können von der Fachperson beispielsweise täglich (über mehrere Wochen hinweg) auf einer Skala zwischen 0 und 10 oder zwischen 0 und 100 Prozent beurteilt werden.

Die Beurteilung des Zielverhaltens erfolgt über ein Item (*Single-Item-Scale*, SIS) oder über bis zu fünf Items (*Multi-Item-Scale*, MIS). Erstere sind sehr ökonomisch, Letztere erweisen sich entlang der Studie von Volpe und Briesch (2012) dafür als zuverlässiger (reliabler). Der zeitliche Aufwand ist auch bei fünf zu beurteilenden Items noch als ökonomisch zu beurteilen, sodass aufgrund ihrer höheren Zuverlässigkeit die *Multi-Item-Scale* in der Praxis bevorzugt werden sollte. Zudem fanden Volpe und Briesch (2012) Hinweise dafür, dass eine *Multi-Item-Scale* weniger Beobachtungszeitpunkte erfordert als eine *Single-Item-Scale*, sodass erstere in einem kürzeren Zeitraum umsetzbar ist.

Die Erfüllung der Gütekriterien Reliabilität und Validität variiert beim Einsatz der direkten Verhaltensbeurteilung in Abhängigkeit des zu beurteilenden Verhaltens (Huber & Rietz, 2015). In ihrem systematischen Review kamen Huber und Rietz (ebd.) zum Schluss, dass Beobachtungen von Verhaltensweisen in den Bereichen Lern- und Arbeitsverhalten von Schüler:innen sowie störendes Sozialverhalten besonders zuverlässig sind. Für die Anwendung sollte eine Verlaufsmessung immer durch die gleiche Person erfolgen. Zudem ist es wünschenswert, dass mehrere Personen unabhängig voneinander Verlaufsmessungen durchführen. Huber und Rietz (ebd.) konnten nachweisen, dass die Übereinstimmung der Beurteilungen zwischen Beurteilenden (Interrater-Reliabilität) zwischen  $r = .67$  und  $r = .78$  variiert. Casale et al. (2015) untersuchten die Generalisierbarkeit und Zuverlässigkeit der direkten Verhaltensbeurteilung anhand der Beurteilung des Lern- und Arbeitsverhaltens an einer inklusiven Primarschule. Die Ergebnisse zeigen, dass etwa 50 Prozent der Varianz durch Verhaltensunterschiede zwischen den Kindern erklärt werden können (Casale et al., 2015), die restlichen 50 Prozent sind auf Unterschiede zwischen den beurteilenden Personen zurückzuführen (diese beurteilten z. B. strenger oder weniger streng). Die Untersuchung der Interrater-Reliabilität zeigte wie bereits erwähnt, dass die Einschätzungen der Rater hinsichtlich der Höhe der eingeschätzten Werte nicht übereinstimmen. Dennoch können sie die strukturelle Varianz der Verhaltensänderung über die Zeit abbilden. Insgesamt kann festgehalten werden, dass die direkte Verhaltensbeurteilung ein geeignetes Instrument zur Verlaufsdiagnostik des Verhaltens von Lernenden ist.

### **Kurze Differenzierung der Auswertungsmethodik**

Nebst der Beurteilung der Wirkung einer Massnahme durch die visuelle Inspektion der Verlaufsdaten ist es wichtig, die Grösse des Effekts numerisch zu bestimmen. So kann er besser eingeordnet werden. Für die statistische Auswertung von Einzelfallverläufen bieten sich sogenannte *Non-Overlap-Indices* an (u. a. Parker et al., 2011). Hierzu wird die Wertemenge aus der A-Phase mit jener aus

der B-Phase oder der nachfolgenden A-Phase in Beziehung gesetzt. Die entsprechende Berechnungsformel gibt dem *Non-Overlap-Index* seinen Namen. Bei *Non-Overlap-Indices* handelt es sich um nicht-parametrische Effektstärkemasse.<sup>1</sup> Sie basieren auf den individuell erhobenen Daten der Baseline- und der Interventionsphase. Für die Interpretation der Werte gibt es (wie auch für andere Effektmasse) Konventionen. In ihrer Übersichtsarbeit stellen Parker et al. (ebd.) neun verschiedene *Non-Overlap-Indices* vor, mit deren Hilfe ein statistischer Vergleich zwischen den unterschiedlichen Phasen von Verlaufsmessungen vorgenommen werden kann. Einige der Analysestrategien eignen sich eher für schulleistungsbezogene Daten wie zum Beispiel für Veränderungen der Lesefertigkeiten oder für das Schreiben oder Rechnen. Andere Methoden sind eher für Verhaltensdaten geeignet. Da das Verhalten in der Regel stark situations- und kontextabhängig ist, und da somit auch starke Schwankungen und extrem hohe oder extrem tiefe Werte vorkommen können, ist es wichtig, dass die Methoden bei der Analyse von Verhaltensbeurteilungen robust gegen Ausreisser sind.

Im Folgenden werden drei Verfahren vorgestellt, die aufgrund ihrer Berechnungsmethoden diese Eigenschaft aufweisen und die somit auch bei Verhaltensdaten benutzt werden können.

(1) *Percentage of data points exceeding the median (PEM)*: Der PEM-Wert bezieht sich auf die Datenpunkte der B-Phase, die den Median der A-Phase übersteigen beziehungsweise unterschreiten je nach gewünschter / erwarteter Veränderungsrichtung (Parker et al., 2011). Zur Prozentberechnung wird die Anzahl dieser Datenpunkte durch die Gesamtzahl aller Datenpunkte der B-Phase geteilt. Durch die Berücksichtigung des Medians ist der PEM relativ robust gegenüber Ausreißern. Der PEM kann Werte zwischen 0 Prozent und 100 Prozent erreichen. Der Median der A-Phase beträgt im vorherigen Beispiel 8 (vgl. Abb. 1). Alle zehn Datenpunkte der B-Phase sind unter diesem Median der A-Phase und somit nicht-überlappend. Hierdurch ergibt sich die Rechnung:  $10 / 10 = 100 \%$ . Liegt der Koeffizient zwischen 70 Prozent und 90 Prozent, wird von einem moderaten Effekt, ab 90 Prozent von einem starken Effekt ausgegangen (Ma, 2006).

(2) *Non-overlap of all pairs (NAP)*: Zur Berechnung des NAP-Wertes werden die Daten aller einzelnen Messzeitpunkte der A-Phase mit allen Messzeitpunkten der B-Phase paarweise verglichen. Er bildet den prozentualen Wert aller sich in die gewünschte Richtung verändernden Datenpaare (Parker et al., 2011). Zur

<sup>1</sup> Effektstärken sind statistische Kennwerte. Sie werden dafür genutzt, um die praktische Bedeutsamkeit pädagogisch-therapeutischer Massnahmen anzugeben.



Berechnung wird eine Summe gebildet aus (a) der Anzahl jener Datenpaare, bei denen sich die Werte der B-Phase in die gewünschte Richtung verändern, und (b) der Anzahl jener Datenpaare, die identische Werte besitzen, multipliziert mit 0,5. Diese Summe wird geteilt durch die Anzahl aller möglichen Datenpaare. Der erhaltene NAP-Wert liegt zwischen 0,00 und 1,00. Er ist jedoch erst ab einer Grösse von über 0,50 relevant, da sonst von einem Zufallsereignis ausgegangen wird. Am Beispiel der Abbildung 1 ergibt sich folgende Rechnung: Es gibt zehn Werte in der A-Phase und zehn Werte in der B-Phase, wodurch es zu 100 Paarvergleichen kommt. 92 Paare überlappen sich nicht und vier Paare sind identisch  $(92 + 4 \cdot 0,5) / 100 = 0,94 = 94 \%$ .

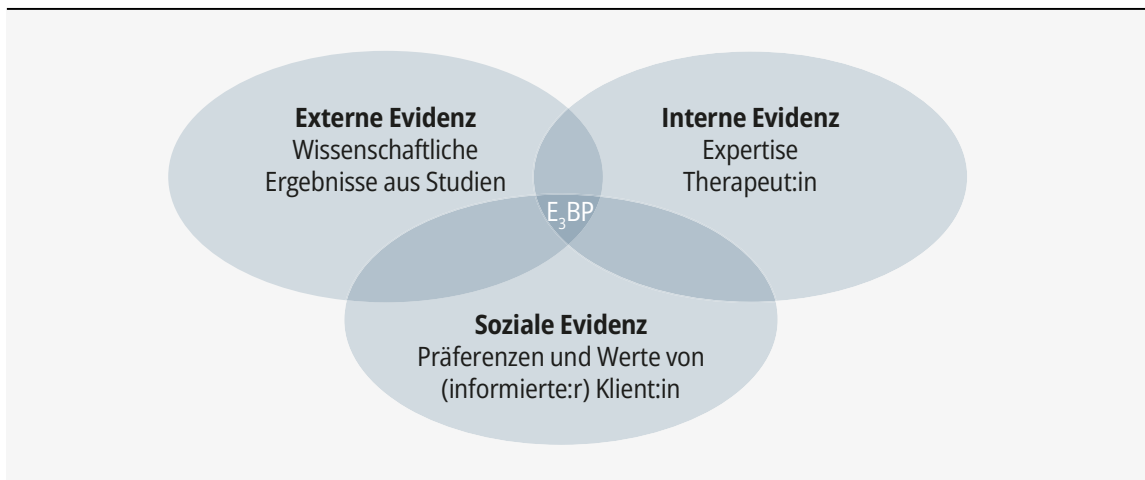
(3) *Percentage of all non-overlapping data (PAND)*: Der PAND-Wert ist der Prozentsatz der überlappungsfreien Datenpunkte im Vergleich zur Gesamtzahl an Datenpunkten (Parker et al., 2011). Um diesen zu bestimmen, werden sowohl aus der Baseline- als auch aus der Interventionsphase schrittweise einzelne Werte eliminiert, die als Ausreisser angesehen werden. Diese auffällig hohen beziehungsweise niedrigen Werte werden gestrichen, um ein Überlappen mit dem Wertebereich der jeweils anderen Phase zu vermeiden. Das Ziel ist, die geringste Anzahl an Datenpunkten zu eliminieren, um eine Überlappungsfreiheit der Baseline- und der Interventionsphase zu erreichen (Brunstein & Julius, 2014). Die Anzahl der verbleibenden Datenpunkte wird dann durch die Anzahl aller Datenpunkte geteilt. Der PAND-Wert kann zwischen 0 und 100 variieren, er ist jedoch erst ab einem Wert von 50 von Bedeutung. Nur wenn sich mindestens 50 Prozent der Datenpunkte für eine überschneidungsfreie Darstellung der Baseline- und Interventionsphase nutzen lassen, kann davon ausgegangen werden, dass die Unterschiede zwischen den beiden Phasen nicht rein zufällig zustande gekommen sind. Damit die Daten von A- und B-Phase überlappungsfrei sind, müssen im Beispiel der Abbildung 1 zwei Datenpunkte (Messungen der Tage 2 und 9) gestrichen werden. Hierdurch ergibt sich die Rechnung:  $18 / 20 = 0,90 = 90 \%$ . Die Faustregel zur Interpretation des PAND-Wertes lautet: 70 Prozent bis 90 Prozent deuten auf einen angemessenen Effekt der Intervention hin, ein PAND-Wert von über 90 Prozent wird als starker Effekt interpretiert (Alresheed et al., 2013).

### **Evidenzbasierung**

Die darstellbare Kontrolle der heilpädagogischen Intervention und ihrer Wirkung (z. B. durch ein Liniendiagramm oder ein Effektstärkemass) ist ein wesentliches Mittel der Evidenzbasierung in der (Heil-)Pädagogik (Hillenbrand, 2015).

Im Modell der evidenzbasierten Praxis (vgl. Abb. 2) betrifft das die Modellkomponenten der internen und sozialen Evidenz (vgl. Kap. 2.2).

Abbildung 2: E<sub>3</sub>BP-Modell der evidenzbasierten Praxis nach Dolloghan (2007; leicht modifiziert aus Beushausen, 2014, S. 97)



Wenn Schulische Heilpädagog:innen ihr praktisches Handeln mit den Strategien einer kontrollierten Einzelfallstudie nachvollziehbar gestalten, erfüllen sie den Anspruch der Evidenzbasierung für ihre einzelfallorientierte Praxis in hohem Masse. Das gilt selbst dann, wenn die sogenannte externe Evidenz nicht in einer befriedigenden Art und Weise umgesetzt werden kann, weil es beispielsweise kein passgenaues Interventionsmodell für den Einzelfall gibt. Es besteht trotzdem durch systematisch angelegte Messreihen oder den Einsatz von Tests die Chance, das individuell-angepasste (heil-)pädagogische Handeln in seiner Wirksamkeit darzustellen. So gesehen kann die kontrollierte Einzelfallstudie mit ihrer Denkweise aus der quantitativen Sozialwissenschaft zur Professionalisierung und Legitimation der Heil- und Sonderpädagogik beitragen.

## Literatur

- Alresheed, F., Hott, B. L. & Bano, C. (2013). Single Subject Research: A Synthesis of Analytic Methods. *The Journal of Special Education Apprenticeship*, 2 (1), 1–18.
- Beushausen, U. (2014). Chancen und Risiken einer evidenzbasierten Praxis. *Logos*, 22, 961–04.
- Brunstein, J. C. & Julius, H. (2014). Evaluation von Interventionen durch Einzelfallstudien. In G. W. Lauth, M. Grünke & J. C. Brunstein, *Interventionen bei Lernstörungen. Förderung, Training und Therapie in der Praxis*. (2., überarb. und erw. Aufl.) (S. 119–138). Hogrefe.
- Casale, G., Hennemann, T., Huber, C. & Grosche, M. (2015). Testgütekriterien der Verlaufsdiagnostik von Schülerverhalten im Förderschwerpunkt Emotionale und soziale Entwicklung. *Heilpädagogische Forschung*, 41, 37–54.

- Döring, N. & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5. Aufl.). Springer. <https://doi.org/10.1007/978-3-642-41089-5>
- Flick, U., von Kardoff, E., Keupp, H., von Rosenstiel, L. & Wolff, S. (2012). *Handbuch Qualitative Sozialforschung*. Beltz.
- Hartke, B., Blumenthal, Y., Carnein, O. & Vrban, R. (2020). *Schwierige Schüler – Förderschule: 84 Handlungsmöglichkeiten bei Verhaltensauffälligkeiten und sonderpädagogischem Förderbedarf* (2. Aufl.). Persen.
- Hillenbrand, C. (2015). Evidenzbasierung sonderpädagogischer Praxis: Widerspruch oder Gelingensbedingung? *Zeitschrift für Heilpädagogik*, 66, 312–324.
- Huber, C. & Rietz, C. (2015). Direct Behavior Rating (DBR) als Methode zur Verhaltensverlaufsdiagnostik in der Schule: Ein systematisches Review von Methodenstudien. *Empirische Sonderpädagogik*, 5, 75–98.
- Jain, A., & Spiess, R. (2012). Versuchspläne der experimentellen Einzelfallforschung. *Empirische Sonderpädagogik*, 4, 211–245. <https://doi.org/10.25656/01:9300>
- Julius, H., Schlosser, R. & Goetze, H. (2000). Kontrollierte Einzelfallstudie. Hogrefe.
- Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings*. Oxford University Press.
- Ma, H.-H. (2006). An alternative method for quantitative synthesis of single-subject researches: Percentage of data points exceeding the median. *Behavior Modification*, 30, 598–617. <https://doi.org/10.1177/0145445504272974>
- Parker, R. I., Vannest, K. J. & Davis, J. L. (2011). Effect Size in Single-Case Research: A Review of Nine Nonoverlap Techniques. *Behavior Modification*, 35, 303–322. <https://doi.org/10.1177/0145445511399147>
- Petermann, F. (1996). *Einzelfallanalyse*. De Gruyter.
- Reichert, M. & Genoud, P. A. (2015). *Einzelfallanalysen in der psychosozialen Forschung und Praxis*. ZKS-Verlag.
- Volpe, R. J. & Briesch, A. M. (2012). Generalizability and Dependability of Single-Item and Multiple-Item Direct Behavior Rating Scales for Engagement and Disruptive Behavior. *School Psychology Review*, 41, 246261.