

Prof. Dr. Johannes Mand

Bochum, 5.10.2017

Familienklassen funktionieren nicht

Im Internet verfügbar unter: www.johannes-mand.de/Fisch.html

Diese Homepage berichtet von der Evaluation zweier Familienklassenprojekte einer Stadt irgendwo in Deutschland. Eines der Projekte wurde im Schuljahr 2016/17 u.a. vom Autor dieses Beitrags wissenschaftlich begleitet. Das zweite Projekt ist Gegenstand einer empirischen Masterthesis, die dem Autor erfreulicherweise vorliegt. Die Ausrichtung der Projekte und auch das Evaluationskonzept beider Konzepte sind nicht identisch. Beide Projekte fühlen sich zwar der Multisystemischen Therapie verpflichtet. Das vom Autor evaluierte Modell ist ein Konzept, das als regionale Initiative entstand und in dem Sozialarbeiter und Lehrer als Elterncoaches arbeiten. Das studentisch evaluierte Konzept ist dagegen ein ausdrücklich von der Kommune unterstütztes Projekt. Hier arbeiten ausschließlich Sozialarbeiter als Elterncoaches. Sie wurden offenbar in einer 12-tägigen Fortbildung und in zwei Tagen Supervision zu "Multifamilientrainern" ausgebildet.

1. Was sind Familienklassen?

Familienklassen sind besondere Lerngruppen in Schulen, in denen Kinder, Eltern, besonders qualifizierte Elterncoaches und Lehrer gemeinsam an Verhaltensproblemen ihrer Kinder arbeiten. Dies geschieht nicht täglich, sondern nur während eines Teils der Schulzeit - in den beiden Familienklassenkonzepten, von denen hier berichtet werden soll, an einem Schulvormittag pro Woche. Die Idee, dass Elterncoaches gemeinsam mit Eltern, Schülern und Lehrern an Verhaltensproblemen arbeiten, ist eigentlich durchaus kompatibel mit unterschiedlichen therapeutischen Konzepten. In den beiden Familienklassenprojekten, von denen hier die Rede sein soll, geschieht dies aber in Ausrichtung an der multisystemischen Therapie.

2. Stand der Forschung

Das Konzept der Familienklassen ist bislang nur unzureichend erforscht. Es gibt zwar einige überzeugte Anhänger des Konzepts (z. B. : Castelleo et al 2016). Aber es wurden bislang nur wenige Studien veröffentlicht, die belastbare Aussagen über die Effektivität von Familienklassen erlauben.

Eine der wenigen Ausnahmen ist die Studie von Morris et al (2014), die von einem Familienklassenprojekt in London berichtet. Diese Studie orientiert sich an einigen wichtigen Standards von Evaluationsstudien. D.h.: Es handelt sich um eine Längsschnittstudie. Die Studie arbeitet mit Versuchsgruppe (50 Kinder aus einem Family Education Center Londons) und Kontrollgruppe (28 Kinder aus Nachbarbezirken). Es handelt sich damit zwar nicht um eine auf Basis von Zufallszahlen gewonnene, repräsentative Stichprobe, aber um eine anfallende Stichprobe immerhin einigermaßen angemessener Größe. Und es werden standardisierte Erhebungsinstrumente verwendet (u.a. der Strength and Difficulties Questionnaire). Bedeutsam ist weiter, dass auch die Einschätzungen von Personen erhoben wurden, die nicht direkt in den therapeutischen Prozess involviert waren (die Lehrer der Familienklassenschüler). Die Intervention war sehr intensiv. Sie wurde an bis zu vier Vormittagen pro Woche in Gruppen von bis zu 12 Kindern mit ihren jeweiligen Angehörigen durchgeführt und dauerte im Durchschnitt 23,4 Wochen.

Die Ergebnisse fallen bei nüchterner Betrachtung allerdings nicht sehr überzeugend aus. Bei den Eltern lassen sich zwar signifikante (also mehr als zufällige) Veränderungen in der Wahrnehmung

ihrer Kinder feststellen. Die Veränderungen sind aber nicht sehr stark, erreicht wurde gerade einmal moderater Effekt (Cohens Effektstärke liegt bei 0,56 – ein moderater Effekt beginnt bei 0,5). Und was von einiger Bedeutung ist: Die Lehrer erkennen keine bedeutsamen Veränderungen bei den Kindern der Familienklassen.

Auch die von ihren Anhängern vorgebrachten allgemeinen Wirksamkeitsbelege multisystemischer Therapie bei dissozialem Verhalten wollen bei etwas genauerer Analyse nicht wirklich überzeugen. Castello et al (2014) verweisen in diesem Zusammenhang zwar auf eine Forschungsübersicht in einer Art Lehrbuch der multisystemischen Therapie (Henggeler et al 2012). Und diese verweist in der Tat auf 33 Arbeiten, die von effektiven Interventionen berichten. Die Überzeugungskraft dieser Forschungsübersicht leidet allerdings ein wenig darunter, dass ein Drittel dieser Arbeiten von den Autoren des Lehrbuchs verfasst wurde. Eine Datenbankrecherche vom 7.1.2017 ermittelt zudem, dass die meisten dieser Arbeiten nicht in den sicher einschlägigen Datenbanken ERIC und Psyn dex auftauchen (jeweils vier Treffer bei den Suchbegriffen Verhaltensstörung - Effekte - systemisch / Multisystemic Therapie). Dies verweist auf nicht erstklassige, möglicherweise sogar auf zweifelhafte Publikationsorte.

Die Autoren versäumen zudem, auf die verwendeten Datenbanken und Auswahlkriterien zu verweisen. Die hier durchgeführte Datenbankrecherche ermittelt z. B. allein 38 bzw. 452 Treffer bei den Suchbegriffen Verhaltensstörung - Effekte – Verhaltensmodifikation (Psyn dex und ERIC). Es ist also zu befürchten, dass diese Forschungsübersicht Methoden anderer Provenienz systematisch ausblendet. Zusammenfassend kann man also sagen, dass Methoden der Multisystemischen Therapie keineswegs als Methoden der Wahl in der Arbeit an Verhaltensstörungen gelten können. Auch das Konzept der Familienklassen muss man als in seiner Wirkung nicht angemessen erforschte Interventionsmethode einstufen.

3. Methode

Das von dem Autor dieses Beitrags evaluierte Familienklassenprojekt verwendet vor allem Methoden der quantitativen Forschung. Die Studie hat eine Versuchsgruppe (n=24 aus fünf Grundschulen) und eine kleine Kontrollgruppe (n=9) in einer von der Sozialstruktur her vergleichbaren Stadt der Region. Die Partnerschüler wurden von den Schulleitern der Kontrollgruppenschulen auf Basis der Befunde aus dem Pretest ausgewählt. Befragt wurden Eltern (nur in der Versuchsgruppe) und Klassenlehrer (in Versuchs- und Kontrollgruppe; das sind mit einer Ausnahme nicht die Coaches). Verwendet wird hier mit der Child Behaviour Checklist ein standardisierter Fragebogen, der in der nationalen und internationalen Forschung zum Thema Verhaltensstörung weit verbreitet ist. Die Klassenlehrer in Versuchs- und Kontrollgruppe haben zudem ein Soziogramm durchgeführt und berichten von den Ergebnissen dieser Befragung in einem gesonderten Fragebogen über Schüler ihrer Klassen (n= 140).

Die Evaluation wird leider nicht von allen Lehrern freundlich unterstützt. Die Vorbehalte gegen die Evaluation finden sich allerdings ausschließlich in der Versuchsgruppe. Wertet man Fälle, in denen mindestens ein Fragebogen im Pretest und Posttest vorliegt, als hinreichenden Beleg zur Bereitschaft der Teilnahme an der Studie, so kommt man insgesamt auf einen Rücklauf von knapp 73 %. Ursprünglich war vorgesehen, dass nach 12 Wochen Interventionsdauer neue Kinder in die Familienklassen aufgenommen werden.

Die studentische Evaluation des zweiten Familienklassenprojekts ist rein qualitativ. Die Arbeit macht keine Angaben darüber, wie viele Kinder insgesamt an dem Projekt teilnahmen. Erhebungsinstrument ist ein Leitfadeninterview mit insgesamt sechs Elternteilen. Es gibt auch in diesem Projekt Rücklaufprobleme. Drei der sechs teilnehmenden Schulen verweigern sich der Hospitation und Befragung. Auch in diesem zweiten Familienklassenprojekt ist vorgesehen, dass in die Familienklassen nach einem Zeitraum von etwa drei Monaten neue Schüler aufgenommen werden.

Wichtige Ergebnisse

Zentraler Evaluationsbefund des Autors dieses Beitrags ist: In der Versuchsgruppe konnten nach 12 Wochen keine Hinweise auf bedeutsame Veränderungen zwischen Pre- und Posttest gefunden werden, und dies gilt für alle Erhebungsinstrumente.

Welche Befunde lassen sich in der Lehrerbefragung der Begleitung der multiprofessionellen Familienklassen ermitteln? Am häufigsten werden die Grenzen der Auffälligkeit beim Index Gesamtwert überschritten (Tabelle 2). Auffällige Werte erreichen insgesamt zwei Schüler der Versuchsgruppe und drei Schüler der Kontrollgruppe. In den Grenzbereich der Auffälligkeit werden drei Schüler der Versuchsgruppe und ein Schüler der Kontrollgruppe eingestuft. Unauffällig sind zwei Schüler der Versuchsgruppe und zwei Schüler der Kontrollgruppe.

Betrachtet man die Veränderungen individuell von Kind zu Kind, so lässt sich zunächst festhalten, dass die beobachteten Veränderungen nicht sehr beeindruckend sind. Und sie erfolgen auch nicht in einer einheitlichen Richtung (Tabelle 1). Es gibt zwei Schüler der Versuchsgruppe und drei Schüler der Kontrollgruppe, deren Verhalten von den Klassenlehrern weitgehend als unverändert eingestuft wird (t-Wert-Differenz unter fünf Punkten). Es gibt einige Kinder, die im Posttest etwas weniger häufig problematische Verhaltensweisen zeigen (vier Kinder der Versuchsgruppe, zwei Schüler der Kontrollgruppe). Und es gibt ein Kind, dem im Posttest mehr problematische Verhaltensweisen attestiert werden (in der Kontrollgruppe). Berechnet man - zu heuristischen Zwecken - Cohens Effektstärke, so kommt man auf einen Wert von $d = 0,05$. also weit von der Grenze zu einem schwacher Effekt entfernt. Läge eine größere Stichprobe vor und würden die Ergebnisse so ausfallen wie hier, müsste man also sagen: Die Intervention ist nicht einmal schwach effektiv.

In den Indices Internationalisierende Störungen fallen Verbesserungen insbesondere in der Kontrollgruppe ins Auge. Die Klassenlehrer berichten in diesem Bereich im Posttest weniger häufig von problematischer Verhaltensweisen. Das Bild in den Familienklassen ist dagegen uneinheitlich: Ein Schüler verbessert sich, einer verschlechtert sich, der Rest bleibt weitgehend unverändert. Hier kann man sogar Werte errechnen, die auf einen starken Effekt verweisen würden. Dieser muss allerdings der Kontrollgruppe zugerechnet werden ($d = -0,83$). Das Vorzeichen ist negativ. Und dies bedeutet: Die Versuchsgruppe schneidet schlechter ab als die Kontrollgruppe. Im Index Externalisierende Störungen gibt es bei drei Schülern der Kontrollgruppe und zwei Schülern der Versuchsgruppe substanzielle Verbesserungen. Aber auch bezogen auf diesen Index würden die Befunde wiederum nicht für einen schwachen Effekt ausreichen ($d = 0,13$).

Tabelle 1: Veränderungen in den Einschätzungen der Klassenlehrerinnen (cbcl/trf Lehrerfragebogen)

	Internalisierende Störungen		Externalisierende Störungen		Gesamtwert	
	(t-werte)		(t-Werte)		(t-wert)	
	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
VG	Schüler 101: 62	Schüler 101: 52	Schüler 101: 44	Schüler 101: 41	Schüler 101: 61	Schüler 101: 50
	Schüler 103: 53	Schüler 103: 55	Schüler 103: 47	Schüler 103: 43	Schüler 103: 49	Schüler 103: 57

	Schüler 106: 48	Schüler 106: 48	Schüler 106: 58	Schüler 106: 57	Schüler 106: 56	Schüler 106: 48
	Schüler 406: 36	Schüler 406: 49	Schüler 406: 52	Schüler 406: 49	Schüler 406: 62	Schüler 406: 63
	Schüler 414: 43	Schüler 414: 46	Schüler 414: 50	Schüler 414: 43	Schüler 414: 65	Schüler 414: 59
	Schüler 418: 39	Schüler 418: 39	Schüler 418: 63	Schüler 418: 58	Schüler 418: 75	Schüler 418: 63
	Schüler 421: 51	Schüler 421: 51	Schüler 421: 56	Schüler 421: 58	Schüler 421: 61	Schüler 421: 61
KG	Schüler 09: 58	Schüler 09: 49	Schüler 09: 43	Schüler 09: 62	Schüler 09: 62	Schüler 09: 72
	Schüler 12: 54	Schüler 12: 51	Schüler 12: 65	Schüler 12: 61	Schüler 12: 74	Schüler 12: 70
	Schüler 16: 43	Schüler 16: 36	Schüler 16: 53	Schüler 16: 31	Schüler 16: 56	Schüler 16: 40
	Schüler 17: 51	Schüler 17: 36	Schüler 17: 43	Schüler 17: 31	Schüler 17: 55	Schüler 17: 38
	Schüler 21: 54	Schüler 21: 37	Schüler 21: 41	Schüler 21: 58	Schüler 21: 64	Schüler 21: 66
	Schüler 26: 51	Schüler 26: 51	Schüler 26: 55	Schüler 26: 53	Schüler 26: 66	Schüler 26: 62

Bei den Elternfragebogen (Tabelle 2) sieht die Situation nicht viel anders aus. Hier liegen zwar keine Werte aus der Kontrollgruppe vor. Aber wieder ist die Befundlage bestenfalls uneinheitlich. Beim Gesamtwert gibt es im Pretest zwei Schüler im Grenzbereich der Auffälligkeit und einen weiteren auffälligen Schüler. Der Posttest vermerkt gleich vier auffällige Schüler. Das Schülerverhalten wird also hier nach 12 Wochen zumindest für diese Schüler ernster eingeschätzt als zu Projektbeginn. Die Einschätzungen von sechs Schülern verändern sich dagegen nur geringfügig. Ähnliches gilt auch für die Indices Internalisierende Störungen und externalisierende Störungen. Im Index internalisierende Störungen verbessern sich drei Schüler, einer verschlechtert sich, der Rest bleibt stabil. Im Index externalisierende Störungen verschlechtert sich ein Schüler substantiell. Der Rest bleibt stabil.

Tabelle 2: Veränderungen in den Einschätzungen der Eltern (cbcl Elternfragebogen)						
	Internalisierende Störungen (t-werte)		Externalisierende Störungen (t-Werte)		Gesamtwert (t-wert)	
	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
VG	Schüler 101: 51	Schüler 101: 45	Schüler 101: 32	Schüler 101: 32	Schüler 101: 47	Schüler 101:46
	Schüler 103: 59	Schüler 103: 63	Schüler 103: 45	Schüler 103: 41	Schüler 103: 56	Schüler 103: 58
	Schüler 104: 70	Schüler 104: 76	Schüler 104: 40	Schüler 104: 46	Schüler 104: 59	Schüler 104: 71
	Schüler 105: 67	Schüler 105: 70	Schüler 105: 45	Schüler 105: 47	Schüler 105: 62	Schüler 105: 65
	Schüler 106: 38	Schüler 106: 38	Schüler 106: 48	Schüler 106: 49	Schüler 106: 54	Schüler 106: 65
	Schüler 217: 38	Schüler 217: 38	Schüler 217: 45	Schüler 217: 44	Schüler 217: 50	Schüler 217: 50

Schüler 406: 44	Schüler 406: 38	Schüler 406: 41	Schüler 406: 30	Schüler 406: 46	Schüler 406: 41
Schüler 414: 59	Schüler 414: 63	Schüler 414: 47	Schüler 414: 44	Schüler 414: 61	Schüler 414: 59
Schüler 418: 59	Schüler 418: 38	Schüler 418: 58	Schüler 418: 59	Schüler 418: 71	Schüler 418: 66

Blieben noch die Einschätzungen der Schüler. Zu sagen: Diesen Schüler mag ich am wenigsten, sagt sicher einiges auch über die Verhaltensprobleme dieser Schüler aus. Ein Projekt, das an schulischen Verhaltensproblemen arbeitet, kann also auch Veränderungen soziometrischer Daten verwenden, um Aussagen über die Effekte der Familienklassen zu machen. In dieser Studie ist das so, wie in vielen anderen Studien (z. B. bei Mand 2007): Schüler mit Verhaltensproblemen werden häufig abgelehnt (Tabelle 3). Gerade einmal fünf von 106 Schülern, denen von ihren Lehrern keine Verhaltensprobleme zugeschrieben werden, erreichen einen „Abgelehnt“- Status. Bei den Schülern mit Verhaltensproblemen sind es sechs von 34. Dies reicht für hoch signifikante Befunde ($p = 0.1$).

	Lehrereinschätzung im Klassenfragebogen: keine gravierende Verhaltensprobleme	Lehrereinschätzung im Klassenfragebogen: gravierende Verhaltensprobleme
Soziometrischer Status: nicht abgelehnt	95,3 %	82,4 %
Soziometrischer Status abgelehnt	4,7 %	17,6 %
$p = .01$	$n=106$	$n = 34$

Nun wird man den einmal erworbenen Abgelehnt-Status nur schwer los, denn es müssen ja nicht nur ein oder zwei Schüler ihre Meinung ändern, sondern es müssen mehr Kinder diesen Schüler als „nett“ kategorisieren als unter der Kategorie „mag ich am wenigsten“ einstufen. Dies gelingt in der vorliegenden Studie gerade einmal zwei Schülern. Beide stammen aus den Kontrollgruppenklassen. Versucht man die Messlatte etwas niedriger zu legen, und untersucht, wer immerhin seine Nicht-Nett-Nennungen reduzieren kann, so zeigt sich, dass diese Veränderungen nur sechs von 34 Schülern mit Verhaltensproblemen gelingen Und wieder kommt keiner dieser Schüler in der Familienklasse. Der pädagogischen Arbeit in den Familienklassen gelingt also nichts, was die Kinder der Herkunftsklassen dazu bewegt, etwas an dem Beliebtheitsstatus der Familienklassenkinder mit Verhaltensproblemen zu ändern.

Im studentisch evaluierten Projekt sehen die Befunde auf den ersten Blick etwas freundlicher aus. Die befragten Eltern berichten positive Dinge über das Projekt. Sie erzählen z. B. von Veränderungen bei ihren Kindern. Diese hätten sich in ihren Schul- und Lernleistungen verbessert, würden Strukturen und Regeln besser akzeptieren, und hätten neue Handlungsmuster erlernt. Sich selbst attestieren die Eltern mehr Sicherheit im Umgang mit den Kindern, ein ruhigeres Verhalten, ein höheres Bewusstsein für die Bedürfnisse ihrer Kinder. Vier von sechs Eltern wollen zudem das Projekt anderen Eltern weiterempfehlen (Schmidt 2016, 74 f.).

Die Stichprobenauswahl scheint allerdings selbst für die Maßstäbe qualitativer Forschung etwas unglücklich verlaufen zu sein. Die Autorin vermutet, dass ggf. nur Elternteile interviewt wurden, die sehr zufrieden mit der Gestaltung des Projekts waren. In diesem Zusammenhang verweist sie

zusätzlich auf das Problem, dass es in Familienklassenprojekten fast unmöglich erscheine, alle Familien zu erreichen. Sie vermutet, dass vor allem Familien das Projekt in Anspruch nehmen, die hohes Engagement für ihr Kind zeigen (Schmidt 2016, 81). Das Problem: Man kann nicht gerade sagen, dass Familien mit hohem Engagement für ihre Kinder unbedingt typisch sind für die eigentliche Zielgruppe des Projekts: Familien in denen Kindeswohlgefährdung befürchtet werden muss. Möglicherweise werden also die Befunde nicht allein dadurch beschädigt, dass ausschließlich begeisterte Anhängerinnen des Projekts befragt wurden. Sondern es besteht zusätzlich das Problem, dass die befragten Eltern möglicherweise überhaupt keine Hilfen benötigen.

Auch in dem studentisch evaluierten Familienklassenprojekt fällt eine Entscheidung gegen eine befristete Interventionsdauer. Der Bericht der Studentin verweist darauf, dass Lehrerinnen und Sozialarbeiterinnen eine Beendigung der Maßnahme (und damit auch eine Neuaufnahme von weiteren Kindern in die Familienklassen) nach 10 Wochen als zu früh empfunden hätten und deshalb den Wunsch formulierten, ohne festen Zeitplan zu arbeiten (Schmidt 2016, 49 f.).

Diskussion

Wie muss man die Ergebnisse einordnen?

Zunächst muss man wohl festhalten, dass immerhin 11 Schulen einer Stadt irgendwo in Deutschland das Konzept der Familienklassen in zwei leicht unterschiedlichen Varianten ausprobiert haben. Zwei unterschiedliche Evaluationsansätze haben sich mit der Arbeit der Schulen befasst. Beide Ansätze sind auf Widerstände seitens der durchführenden Schulen gestoßen – in einem ersten Fall weigerten sich einige Lehrer aus zwei Schulen standardisierte Fragebogen auszufüllen. Im zweiten Fall verweigerten sich drei von sechs Schulen Hospitation und Befragungswünschen. Dies ist ziemlich bedenklich und verweist möglicherweise auf einen beachtlichen Anteil von Projektbeteiligten, die keine empirische Untersuchungen benötigen, um zu entscheiden, was wirksam ist.

Die quantitative Evaluation kann keinerlei Hinweise auf nennenswerte Verhaltensänderungen bei den Kindern ermitteln. Die qualitative Evaluation findet sechs, möglicherweise handverlesenen Eltern, die bereit sind, von Erfolgen zu sprechen. Und diese Eltern gehören offenbar nicht zur eigentlichen Zielgruppe des Projekts.

Bei beiden Konzepten entscheiden sich die Beteiligten, die Interventionsdauer drastisch auszudehnen. Dabei lassen die verwendeten Formulierungen bei der Begründung den Schluss zu, dass auch die beteiligten Coaches zumindest keine ausreichenden Verhaltensänderungen bei ihren Klienten wahrnehmen können.

10 bis 12 Wochen Intervention sind also in diesen Projekten ganz offensichtlich nach allen verfügbaren Informationen nicht ausreichend, um spürbare Verhaltensänderungen in den Familien zu erreichen. Besondere Tragweite erhält dieser Befund vor dem Hintergrund der Befunde von Morris et al (2014). Wenn man festhält, dass eine Intervention, die fast doppelt so lange Zeit in Anspruch nimmt, wie die hier vorgestellten Projekte und mit bis zu vierfach längeren wöchentlichen Familienklassenzeiten arbeitet, und ermittelt dann als Ergebnis, dass lediglich festzuhalten ist, dass die befragten Eltern ihre gestörten Kinder in einem etwas milderem Licht sehen, dann ist es wohl berechtigt, Zweifel an der Wirksamkeit des Konzepts der Familienklassen zu formulieren. In Kurzfassung lässt sich demnach festhalten: Familienklassen funktionieren nicht.

Nun könnte man argumentieren, 12 Wochen, vielleicht auch 24 Wochen seien eine kurze Zeit für die Arbeit an Verhaltenproblemen, die Familienverhältnisse seien bekanntlich schwierig, es bedürfe eben mehr Zeit, um wirksam zu intervenieren. Diese Argumentation ist zwar nachvollziehbar. Sie übersieht aber, dass anderen Konzepten eben dies – eine wirksame Veränderung von problematischem Verhalten bei Kindern aus schwierigen Familien in vergleichbaren Zeiträumen

durchaus gelingt. Dies gilt insbesondere für Trainingsprogramme verhaltenstherapeutischer Provenienz. Am meisten Bekanntheit hat Triple P erreicht, ein Programm, das ausführlich und mit guten Befunden evaluiert ist (eine Übersicht bei Halweg & Heinrichs 2007).

Dass Elterntrainingsprogramme verhaltenstherapeutischer Provenienz auch innerhalb eines vergleichbar kurzen Zeitraums in der Lage sind, das Verhalten von Kindern zu beeinflussen, belegt z. B. auch Otte (2011). Nur zum Vergleich: Otte untersucht die Effekte des Kompetenztrainings für sozial auffällige Kinder (Lauth & Heibek 2006) im Einzel- und Gruppentraining. Dieses Konzept umfasst sechs Trainingseinheiten und eine Auffrischungssitzung. Die Studie von Otte (2011, Seite 86, Seite 93) ist als randomisierte Längsschnittstudie konzipiert (N = 258, Warte-Kontrollgruppe, zwei Messzeitpunkte, Follow up nach einem halben Jahr). Die Studie verwendet mehrere standardisierte Fragebogen, u.a. auch die in der Evaluation der Familienklassen eingesetzte Child Behaviour Checklist und kommt auf Effektstärken von 0,7 bis 0,88. Die mittlere Effektstärke liegt bei $d = 0,65$. Dem Elterntraining gelingt also eine Verringerung der Intensität der Verhaltensstörungen bei den untersuchten hyperkinetischen Kindern (Otte 2011, 122), und dies bei wesentlich geringerer Interventionsdauer als in den hier untersuchten Familienklassenprojekten oder gar im Vergleich zur Londoner Studie. Im eigentlichen Schwerpunkt der Studie – dem Vergleich von Einzel- und Gruppentraining konnten in zusammenfassender Wertung der Befunde aller Instrumente übrigens keinerlei Hinweise auf bedeutsame Unterschiede ermittelt werden (Otte 2011, 122). Das Einzeltraining von Eltern bietet also keinerlei belastbaren Vorteil.

Man muss also festhalten: Elterntrainingsprogramme verhaltenstherapeutischer Provenienz funktionieren. Vermutlich muss man sogar davon ausgehen, dass dies auch gelingen würde, wenn das Training in Schulen, vielleicht sogar unter Anwesenheit der Schüler durchgeführt würde. Dass aber multisystemisches Elterncoaching in Familienklassen hilfreich ist, das konnte bislang nicht überzeugend nachgewiesen werden.

Literatur

Arbeitsgruppe Deutsche Child Behaviour Checklist (1993). Lehrerfragebogen über das Verhalten von Kindern und Jugendlichen. Köln.

Castello, A., Bierkandt S. Suchy, J. (2016). Familienklassen: Schulische Interventionen im Multifamiliensetting. Zeitschrift für Heilpädagogik, 67, 227-233.

Hahlweg, U., Heinrichs, N. (2007). Prävention von kindlichen Verhaltensstörungen mit dem Triple P-Elterntraining. In: Von Suchodoletz W. (Ed.), Prävention von Entwicklungsstörungen (pp. 183-201). Göttingen: Hogrefe.

Henggeler, S. W., Schoenwald S., Borduin, C. M, Rowland, M. D. & Cunningham, P. B. (2012). Multisystemische Therapie bei dissozialem Verhalten von Kindern und Jugendlichen. Berlin: Springer.

Lauth, G. W., Heubeck, B. G. (2006). Kompetenztraining für Eltern sozial auffälliger Kinder – KES. Göttingen: Hogrefe.

Morris, E., Le Huray C., Skagerberg F., Comes, R. & Ninteman, A. (2014). Families changing Families: The Protective function of multi-family therapy for children in education. In: Clinical Child Psychology and Psychiatry 19 (4), 617-632.

Otte, T. A. (2011). Klinische Studie zur Wirksamkeit kognitiv behavioraler Elterntrainings bei hyperkinetischen Störungen. Ein Vergleich von Einzel- und Gruppenintervention. Dissertation Universität zu Köln. [kups...ub.uni-koeln.de/4521/1/Dissertation_Th_A-Otte_Juni_2011.pdf](https://kups.ub.uni-koeln.de/4521/1/Dissertation_Th_A-Otte_Juni_2011.pdf); Zugriff am 29.1.2017.

Schmidt, S. (2016): Familienklassen als Instrument der sozialen Inklusion in der Schule. Bochum (unveröffentlichte Masterthesis).