

Differentielle Wirksamkeit eines Lerntrainings für Jugendliche mit Aufmerksamkeitsstörungen

Von der Carl von Ossietzky Universität Oldenburg – Fakultät I Bildungs-
und Sozialwissenschaften – zur Erlangung des Grades eines

Doktors der Philosophie (Dr. phil)

genehmigte Dissertation von

Timo Hennig

geboren am 26.09.1982 in Nordhorn

Referent: Prof. Dr. Friedrich Linderkamp

Korreferent: Prof. Dr. Gerhard W. Lauth

Tag der Disputation: 28.10.2013

Inhalt

ABBILDUNGEN	5
TABELLEN	6
1 EINLEITUNG	8
2 JUGENDALTER	11
2.1 DEFINITION	13
2.2 ENTWICKLUNGSPSYCHOPATHOLOGIE	15
2.3 ENTWICKLUNGSAUFGABEN	19
2.4 KOGNITIVE ENTWICKLUNG	21
2.4.1 GEHIRNENTWICKLUNG ALS BIOLOGISCHE GRUNDLAGE	22
2.4.2 QUALITATIVE VERÄNDERUNGEN	24
2.4.3 QUANTITATIVE VERÄNDERUNGEN	25
2.4.4 INFORMATIONSVERRARBEITUNG	27
2.4.5 GEDÄCHTNIS UND WISSEN	28
2.4.6 EXEKUTIVFUNKTIONEN	29
2.4.7 FAZIT	31
3 AUFMERKSAMKEITSTÖRUNGEN	33
3.1 KLASSIFIKATION	33
3.2 GESCHLECHTSSPEZIFIK	39
3.3 PRÄVALENZ	42
3.4 KOMORBIDITÄT	44
3.5 VERLAUF	51
3.6 ERKLÄRUNGSANSÄTZE	56
3.6.1 KONZEPTUELLE GRUNDLAGEN	57
3.6.2 NEUROPSYCHOLOGISCHE ERKLÄRUNGSMODELLE	60
3.6.3 GENETIK	64
3.6.4 PRÄ-, PERI- UND POSTNATALE KOMPLIKATIONEN	65
3.6.5 NEUROBIOLOGIE	67
3.6.6 NEUROPSYCHOLOGIE	69
3.6.7 SOZIALE INTERAKTIONEN	71
3.7 INTERVENTION	78
3.7.1 MEDIKATION	79
3.7.2 PSYCHOSOZIALE INTERVENTIONEN	80
3.7.3 DAS LERNTRAINING FÜR JUGENDLICHE MIT ADHS (LEJA)	81
4 EVALUATION PSYCHOLOGISCHER INTERVENTIONEN	88
4.1 PSYCHOTHERAPIEFORSCHUNG	88
4.2 OPERATIONALISIERUNG VON INTERVENTIONSERFOLG	91
4.3 VERÄNDERUNGSMESSUNG	94
4.3.1 VERZERRUNGSEFFEKTE BEI MESSWIEDERHOLUNGEN	95
4.3.2 RELIABILITÄT VON DIFFERENZWERTEN	96
4.3.3 REGRESSION ZUR MITTE	100
4.3.4 SIGNIFIKANZPRÜFUNG	102
4.3.5 EFFEKTSTÄRKEN IN GRUPPENVERGLEICHEN	104
4.3.6 ERFOLGSMASSE IM EINZELFALL	105

4.4	DIFFERENTIELLE WIRKSAMKEIT	108
4.4.1	WIRKSAME TEILKOMONENTEN EINER INTERVENTION	109
4.4.2	MODERATOREN UND MEDIATOREN	110
4.4.3	AUSWIRKUNG INITIALER WERTE	113
4.4.4	RESPONDER UND NON-RESPONDER	114
4.5	EMPIRISCHE ERGEBNISSE ZUR DIFFERENZIELLEN WIRKSAMKEIT	116
5	UNTERSUCHUNGSPLANUNG	134
5.1	ALLGEMEINE WIRKSAMKEITSPRÜFUNG	134
5.1.1	EVALUATIONSDESIGN	134
5.1.2	EINGESETZTE INSTRUMENTE	135
5.1.3	GESAMTPROBLEMMAB	152
5.1.4	VERÄNDERUNGSMESSUNG	155
5.1.5	VORLÄUFIGE ERGEBNISSE	156
5.2	DIFFERENTIELLE WIRKSAMKEIT	156
5.2.1	FRAGESTELLUNG	158
5.2.2	HYPOTHESEN	160
5.2.3	OPERATIONALISIERUNG	166
5.2.4	METHODISCHES VORGEHEN	168
5.2.5	STICHPROBE	176
5.2.6	DURCHFÜHRUNG	179
5.3	AUFBEREITUNG DER DATEN	181
5.3.1	PRÄ-POST-VERÄNDERUNGEN AUF VARIABLENEBENE	181
5.3.2	FEHLENDE WERTE	184
5.3.3	BESTIMMUNG EINES GESAMTPROBLEMMABES	187
5.3.4	BESTIMMUNG VON VERÄNDERUNGSMABEN	192
6	ERGEBNISSE	195
6.1	VARIABLENZENTRIERTE EINFACHE UNIVARIATE AUSWERTUNG	195
6.2	VARIABLENZENTRIERTE MULTIPLE UNIVARIATE AUSWERTUNG	210
6.3	VARIABLENZENTRIERTE MULTIVARIATE AUSWERTUNG	217
6.4	PERSONENZENTRIERTE UNIVARIATE AUSWERTUNG	218
6.5	PERSONENZENTRIERTE MULTIVARIATE AUSWERTUNG	221
6.6	WEITERFÜHRENDE AUSWERTUNG: ZUSAMMENHÄNGE DER AUSGANGSWERTE	227
7	DISKUSSION	229
7.1	INTERPRETATION DER BEFUNDE	229
7.2	FAZIT UND KLINISCHE IMPLIKATIONEN	244
LITERATUR		249
ANHANG		286
A.1	ERKLÄRUNGEN	287
A.2	VERGLEICH FEHLENDE LEHREREINSCHÄTZUNG	288
A.3	ÜBERPRÜFUNGEN AUF NORMALVERTEILUNG	290
A.4	WEITERE ERGEBNISSE DER UNIVARIATEN REGRESSIONSANALYSEN	292
A.5	WEITERE ERGEBNISSE DER MULTIVARIATEN REGRESSIONSANALYSE	297

Abbildungen

Abbildung 1: Modell der transaktionalen Entwicklung (aus Sameroff, 2000, S. 309).....	15
Abbildung 2: Mögliche kausale Zusammenhänge zwischen den von Werner (1993) untersuchten risikomildernden Bedingungen (aus Petermann, Niebank & Scheithauer, 2004, S. 350)	17
Abbildung 3: Strukturelle Gehirnentwicklung (aus Casey, Tottenham, Liston & Durston, 2005, S. 105)	24
Abbildung 4: Mechanics und Pragmatics in der kognitiven Entwicklung (aus Baltes, Staudinger & Lindenberger, 1999, S. 487).....	26
Abbildung 5: Körperliche pubertäre, soziale und emotionale Einflüsse auf das Entscheidungsverhalten im Lauf der Adoleszenz (aus Steinberg, 2005b, S. 70)	31
Abbildung 6: Diagnosen der hyperkinetischen bzw. Aufmerksamkeitsstörungen nach ICD-10 und DSM- IV (aus Döpfner, Banaschewski & Sonuga-Barke, 2008, S. 259).....	36
Abbildung 7: Zeitlicher Verlauf des Auftretens von ADHS und komorbiden Störungen (aus Taurines et al. 2010, S. 269)	46
Abbildung 8: Vermutete Prä-, simultane und Postkomorbidität der ADHS (aus Taurines et al., 2010, S. 268)	46
Abbildung 9: Entwicklungsmodell aggressiv-dissozialen Verhaltens, über den Lebenslauf stabiler Typus (aus Petermann, Niebank & Scheithauer, 2004, S. 384)	51
Abbildung 10: Vereinfachtes hypothetisches Kausalmodell (aus Coghill et al., 2005, S. 106)	58
Abbildung 11: V-förmiges Kausalmodell, bei dem unterschiedliche kausale Pfade zur phänotypischen Hyperaktivität auf der Verhaltensebene führen (aus Morton & Frith, 1995, S. 372)	58
Abbildung 12: Der Zusammenhang von genetischen und umweltbezogenen Risikofaktoren, Endophänotypen und Phänotypen bei ADHS (aus Rommelse et al., 2008, S. 2).....	60
Abbildung 13: Das kognitiv-energetische Modell (aus Sergeant, 2000, S. 8).....	61
Abbildung 14: Das Dual Pathway Modell der ADHS (aus Sonuga-Barke, 2002, S. 32).....	63
Abbildung 15: Kausalmodell der ADHS und Wirkweise von Interventionen (aus Rapport et al. , 2001, S. 49)	64
Abbildung 16: Grafische Veranschaulichung des Reliabilitäts-Validitäts-Dilemmas.....	98
Abbildung 17: Moderatormodell aus Baron und Kenny (1986).....	111
Abbildung 18: Mediatormodell aus Baron und Kenny (1986).....	112
Abbildung 19: Mediator- und Moderatoreffekte aus Holmbeck (1997).....	112
Abbildung 20: Veranschaulichung der stärker deduktiven und der stärker induktiven Vorgehensweise zur Beantwortung der Fragestellung nach der differentiellen Wirksamkeit des Lerntrainings	158
Abbildung 21: Altersverteilung der Stichprobe in Jahren	179
Abbildung 22: Verteilung der Häufigkeiten der Dauer des Trainings in Monaten	180
Abbildung 23: Verteilung der Häufigkeiten der Anzahl der durchgeführten Sitzungen pro Teilnehmer.	181
Abbildung 24: Screeplot der Outcomemaße vor Trainingsbeginn	188
Abbildung 25: Vergleich der Gesamtproblematik bei Mädchen und Jungen vor Trainingsbeginn	196
Abbildung 26: Vergleich der Problemmaße Symptomatik und Fehler vor Trainingsbeginn bei Mädchen und Jungen in Boxplots	197
Abbildung 27: Alter der Teilnehmer und Differenzwerte im Gesamtproblemmaß.....	203
Abbildung 28: Zusammenhang der Werte der Gesamtproblematik vor und nach dem Training	204
Abbildung 29: Gesamtproblematik vor und nach dem Training mit Reliable Change Index (RCI)	219
Abbildung 30: Darstellung der zu Clustern zusammengefassten Veränderungsmaße der Teilnehmer	225

Tabellen

Tabelle 1: Uneinheitliche Einteilung der Adoleszenz nach Lebensalter in verschiedenen Definitionen ...	14
Tabelle 2: Symptome der ADHS gemäß DSM-IV-TR (Saß et al., 2003)	33
Tabelle 3: Subtypen der ADHS gemäß DSM-IV-TR.....	35
Tabelle 4: Ablauf des Lerntrainings für Jugendliche mit ADHS (LeJA)	84
Tabelle 5: Versuchsdesign der allgemeinen Wirksamkeitsprüfung	135
Tabelle 6: Erfolgskriterien, Messinstrumente und Informationsquellen	135
Tabelle 7: Studien zur internen Konsistenz der deutschen Version des SDQ	142
Tabelle 8: Odd-even-Reliabilität der Testbatterie zur Aufmerksamkeitsprüfung (TAP) in für die Studie relevanten Subtests (Zimmermann & Fimm, 2009)	146
Tabelle 9: Studien zur Retest-Reliabilität relevanter Subtests der Testbatterie zur Aufmerksamkeitsprüfung (TAP).....	147
Tabelle 10: Stichprobenverteilung nach Geschlecht, Medikamenteneinnahme zum Prämesszeitpunkt und Schulform.....	178
Tabelle 11: Deskriptive Auswertung der Outcomemaße vor und nach dem Training	182
Tabelle 12: Fehlende Werte der Outcomemaße vor und nach dem Training nach der ersten einfachen Imputation	186
Tabelle 13: Hauptkomponentenanalyse (Promax- und Varimax-rotiert) der Outcome-Maße zum Zeitpunkt vor Trainingsbeginn	189
Tabelle 14: Verteilungsparameter der vier Faktoren bzw. Problemmaße und des Gesamtproblemmaßes	192
Tabelle 15: Verteilungsparameter der Differenzwerte der vier Faktoren bzw. Problemmaße und des Gesamtproblemmaßes	194
Tabelle 16: Vergleich der Problemmaße von Mädchen und Jungen zum Prämesszeitpunkt	196
Tabelle 17: Vergleich der Differenzwerte von Mädchen und Jungen	198
Tabelle 18: Vergleich der Werte von Teilnehmern mit und ohne Medikation zum Prämesszeitpunkt	199
Tabelle 19: Vergleich der Differenzwerte von Teilnehmern mit und ohne Medikation	200
Tabelle 20: Zusammenhang der Hyperaktivität und Impulsivität in der Elterneinschätzung mit den Problemmaßen vor dem Training	200
Tabelle 21: Zusammenhang der Hyperaktivität und Impulsivität in der Elterneinschätzung vor dem Training mit den Differenzwerten	201
Tabelle 22: Zusammenhang des Alters der Teilnehmer mit den Werten der Problemmaße zum Prämesszeitpunkt.....	202
Tabelle 23: Zusammenhang des Alters der Teilnehmer mit den Differenzwerten	203
Tabelle 24: Prä-post-Effektstärken der Gesamtproblematik nach Altersgruppen	203
Tabelle 25: Zusammenhang der initialen Störungsschwere mit den Differenzwerten	205
Tabelle 26: Zusammenhang der externalisierenden Komorbidität mit den Problemmaßen vor dem Training	206
Tabelle 27: Zusammenhang der internalisierenden Komorbidität mit den Problemmaßen vor dem Training	206
Tabelle 28: Zusammenhang der externalisierenden Komorbidität mit den Differenzwerten	207
Tabelle 29: Zusammenhang der internalisierenden Komorbidität mit den Differenzwerten	208
Tabelle 30: Zusammenhang der Trainingsdauer mit den Problemmaßen vor dem Training	209
Tabelle 31: Zusammenhang der Trainingsdauer mit den Differenzwerten	209
Tabelle 32: Zusammenhang der Anzahl der Sitzungen mit den Werten zum Prämesszeitpunkt	210

Tabelle 33: Zusammenhang der Anzahl der Sitzungen mit den Differenzwerten	210
Tabelle 34: Ergebnisse der schrittweisen linearen Regressionsanalyse. Abhängige Variable: Gesamtproblematik post	212
Tabelle 35: Ergebnisse der schrittweisen linearen Regressionsanalyse. Abhängige Variable: Differenz Gesamtproblematik	213
Tabelle 36: Schrittweise lineare Regressionsanalyse. Abhängige Variable: Symptomatik post.....	214
Tabelle 37: Schrittweise lineare Regressionsanalyse. Abhängige Variable: Differenz Symptomatik	214
Tabelle 38: Schrittweise lineare Regressionsanalyse. Abhängige Variable: Lehrereinschätzung post	214
Tabelle 39: Schrittweise lineare Regressionsanalyse. Abhängige Variable: Differenz Lehrereinschätzung	215
Tabelle 40: Schrittweise lineare Regressionsanalyse. Abhängige Variable: Reaktionszeit post.....	215
Tabelle 41: Schrittweise lineare Regressionsanalyse. Abhängige Variable: Differenz Reaktionszeit	216
Tabelle 42: Schrittweise lineare Regressionsanalyse. Abhängige Variable: Fehler post	217
Tabelle 43: Schrittweise lineare Regressionsanalyse. Abhängige Variable: Differenz Fehler.....	217
Tabelle 44: Ergebnisse der multivariaten Varianzanalyse (Pillai-Spur). Abhängige Variablen: Symptomatik und Lehrereinschätzung zum Postmesszeitpunkt.....	218
Tabelle 45: Deskriptive Ergebnisse für Responder und Non-Responder nach RCI (dichotome Variablen)	220
Tabelle 46: Deskriptive Ergebnisse für Responder und Non-Responder nach RCI (metrische Variablen)	220
Tabelle 47: Hierarchische Clusteranalyse nach Differenzwerten der Problemmaße (Linkage innerhalb der Gruppen)	222
Tabelle 48: Deskriptive Ergebnisse für die Cluster (dichotome Variablen).....	226
Tabelle 49: Deskriptive Ergebnisse für die Cluster (metrische Variablen).....	226
Tabelle 50: Vergleich der Zusammenhänge der Ausgangswerte (Prämesszeitpunkt) und der Differenzwerte.....	228

1 Einleitung

A fundamental question in child psychotherapy research is as follows: What type of treatment offered to what type of client who has what type of difficulty produces what type of effect? In other words: what are the factors that moderate treatment outcome? (Durlak, Wells, Cotten & Johnson, 1995, S. 144)

Die Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS) ist eine der häufigsten psychischen Störungen des Kindes- und Jugendalters. Die typischen Symptome sind Unaufmerksamkeit, Hyperaktivität und Impulsivität. Die Eltern betroffener Kinder bekommen oftmals schon im Vorschulalter die Rückmeldung, dass sich ihr Kind durch eine besonders starke Aktivität auszeichnet und sich somit von den anderen unterscheidet. Zumeist kurz nach dem Schuleintritt werden erste Probleme offenkundig, da die Anforderungen der Schule nicht zu den Besonderheiten der Kinder mit ADHS passen. Sie haben Schwierigkeiten, eigenständig still und konzentriert zu arbeiten, ruhig sitzen zu bleiben und sich zu gedulden (Gawrilow, 2012).

Entgegen der früheren Annahme, die Störung „wachse sich aus“, zeichnet sich immer stärker ab, dass die ADHS über die gesamte Lebensspanne bestehen bleibt (Turgay et al., 2012). Für viele der Betroffenen handelt es sich nicht um eine kindliche Besonderheit, die in der Adoleszenz langsam abgeflaut und spätestens im Erwachsenenalter ganz verschwunden ist. Viele Jugendliche, die als Kind eine ADHS-Diagnose bekommen haben, stellen fest, dass es ihnen auch in der Adoleszenz noch schwer fällt sich zu konzentrieren. Dies wirkt sich nicht nur auf das schulische Lernverhalten aus, sondern beeinträchtigt auch soziale Beziehungen wie Freundschaften und partnerschaftliche Beziehungen, die für Jugendliche besonders wichtig sind (Hoza et al., 2005).

Viele Lehrer verstehen nicht, warum es den Jugendlichen so schwer fällt, ihre durchaus vorhandenen Potenziale auszuschöpfen. In einigen Situationen zeigen sie, dass sie in der Lage sind, gut mitzuarbeiten und ihre Aufgaben zu erledigen, doch in anderen Situationen versagen sie. Die schulische Karriere ist daher für Kinder und Jugendliche mit ADHS oftmals steinig (Lauth & Naumann, 2009). Wichtige Entwicklungsaufgaben wie die Erlangung eines Schulabschlusses und das Finden eines passenden Berufes sind in Gefahr.

Die Diagnose ADHS ist zunehmend auch bei Erwachsenen vorhanden und aufgrund der steigenden Nachfrage gibt es neue Interventionskonzepte und Beratungsstellen. In den letzten Jahren wurden zahlreiche Untersuchungen zu ADHS veröffentlicht, viele im Erwachsenenbereich (Lauth & Minsel, 2009). Es zeigt sich, dass zu den Schwierigkeiten im beruflichen und sozialen Bereich bei vielen Betroffenen emotionale Problemlagen hinzukommen. Beispielsweise ist das Risiko, an einer Depression oder Angststö-

zung zu erkranken, erhöht. Für viele Menschen mit einer ADHS handelt es sich um eine Bedingung, die sie lebenslang begleitet, auch wenn sich das Erscheinungsbild ändert.

Das Vorliegen einer ADHS kann daher als eine ernstzunehmende Entwicklungsgefährdung verstanden werden (Ihle & Esser, 2002). Damit wird deutlich, wie wichtig die Entwicklung wirksamer Präventions- und Interventionsmaßnahmen ist. Diese Maßnahmen müssen entwicklungs sensitiv sein, also den jeweiligen Entwicklungsstand und die alterstypische Symptomatik der ADHS berücksichtigen. Auf Grundlage evidenzbasierter Behandlungstechniken und -strategien müssen angemessene Interventionen erprobt, evaluiert und weiterentwickelt werden. Ziel ist es, ein immer differenzierteres Wissen darüber zu erlangen, welche Intervention für welches Kind, welchen Jugendlichen und welchen Erwachsenen wirksam ist. Somit kann es nicht eine universell wirksame Intervention bei ADHS geben.

Das in der vorliegenden Untersuchung evaluierte Programm LeJA ist deshalb speziell auf die Entwicklungssituation und alterstypischen Probleme und Bedürfnisse von Jugendlichen mit ADHS zugeschnitten. Die vorliegende differenzielle Auswertung geht noch einen Schritt weiter und sucht in der Gruppe der Jugendlichen mit ADHS nach Unterschieden. Es handelt sich um keine homogene Gruppe, sondern eine Gruppe mit vielen Unterschieden, die auf das Alter, das Geschlecht und andere Aspekte zurückzuführen sind (Robin, 1998). Ziel dieser Untersuchung ist es daher, herauszufinden, für welche Jugendliche das LeJA-Training besonders wirksam ist und für welche weniger wirksam, um mit den gewonnenen Erkenntnissen eine Anpassung des Trainings an die jeweiligen Bedürfnisse zu entwickeln.

Die vorliegende Dissertation ist im Forschungsprojekt „Evaluation eines Lerntrainings für Jugendliche mit ADHS“ am Institut für Sonderpädagogik und Rehabilitation der Carl von Ossietzky Universität Oldenburg entstanden. Projektleiter war Prof. Dr. Friedrich Linderkamp, weiterer Projektmitarbeiter war Diplom-Psychologe Satyam A. Schramm. Aufbauend auf der Diplomarbeit des Autors (Hennig, 2008) wurde ein Manual für eine Interventionsmaßnahme für Jugendliche mit ADHS entwickelt und als „Lerntraining für Jugendliche mit ADHS (LeJA)“ publiziert (Linderkamp, Hennig, & Schramm, 2011). Von Januar 2009 bis Juli 2012 nahmen 115 Jugendliche mit ADHS an der Evaluationsstudie teil.

Die Evaluation der Interventionsmaßnahme im Rahmen des Forschungsprojektes ist zweiteilig. Im ersten Teil wird in einem Kontrollgruppendesign die allgemeine Wirksamkeit der Intervention überprüft. Im zweiten Teil werden differenzielle Effekte der Wirkung des Trainings untersucht. Die vorliegende Dissertation behandelt den zweiten Teil der Evaluation zur differenziellen Wirksamkeit.

Die Ausarbeitung beginnt mit der Darstellung theoretischer Hintergründe, die zum Verständnis und zur Beurteilung der weiterführenden Analysen notwendig sind. Den Anfang bildet ein Kapitel zur Lebensphase des Jugendalters, es folgen eine Vorstellung des Störungsbildes ADHS und ein Kapitel zur Evaluation von psychologischen Interventio-

nen. Es folgt die Darstellung der Planung der vorliegenden Untersuchung, an die sich die Ergebnisdarstellung und Diskussion der Ergebnisse anschließt. Den Abschluss bildet ein Kapitel, in dem ein Fazit gezogen und klinische Implikationen der Ergebnisse dargestellt werden.

Zugunsten der Lesbarkeit wurde darauf verzichtet, stets die männliche und die weibliche Form von Adjektiven und Substantiven zu nennen. Wenn nicht anders betont, sind alle Geschlechter mit einer Form angesprochen.

Zum erfolgreichen Abschluss dieser Arbeit haben einige Personen beigetragen, denen herzlich gedankt sein soll. Meinem Kollegen Satyam A. Schramm und unseren Hilfskräften Moritz Börnert, Helen Reurik und Antje Wagner für die gemeinsame Durchführung des Projektes und die Erhebung der Daten. Dem Betreuer der Arbeit, Prof. Dr. Friedrich Linderkamp, für die Überlassung des Themas und die Erstellung des Erstgutachtens. Den Jugendlichen, die an dem Training teilgenommen haben, und ihren Eltern und Lehrern für die gute Kooperation. Den Studierenden der Sonder- und Rehabilitationspädagogik, die das Training durchgeführt haben. Den Korrekturleserinnen Andrea Anschütz, Nanke Krieghoff, Conny Melzer, Carolin Reinck, Helen Reurik, Anna Schorer, Marie-Christine Vierbuchen und Antje Wagner für ihre hilfreichen Rückmeldungen. Prof. Dr. Gerhard W. Lauth für die Erstellung des Zweitgutachtens.

2 Jugendalter

Das Lerntraining LeJA richtet sich an Jugendliche im Alter von 12 bis 17 Jahren. In diesem Kapitel wird die Zielgruppe der Jugendlichen eingehend beschrieben, um die Entwicklungsspezifik dieser Altersphase transparent zu machen. Nach einer Einführung und Definition wird die Entwicklungspsychopathologie als ein theoretischer Rahmen eingeführt. Es folgt eine Vorstellung des Konzepts der Entwicklungsaufgaben, das zum Verständnis der alterstypischen Herausforderungen und Risiken beiträgt. Abschließend wird die kognitive Entwicklung im Jugendalter ausführlich dargestellt. Der Bereich Kognition ist im Zusammenhang mit einem Lerntraining als besonders wichtig zu bezeichnen. Eine der zentralen Hypothesen der späteren Auswertung leitet sich von diesem Aspekt ab.

Das Jugendalter befindet sich in der Lebensspanne zwischen Kindheit und Erwachsenenalter in einer Art „Sandwichposition“ (Grob & Jaschinski, 2003, S. 13). Mit diesem Bild wird betont, dass die Einordnung des Jugendalters im Wesentlichen über eine Abgrenzung von der vorangehenden und der nachfolgenden Lebensphase geschieht. Jugendliche legen kindliche Verhaltensweisen ab und es werden ihnen kindliche Privilegien aberkannt. Dafür fordern Jugendliche Erwachsenenrechte ein und es wird von ihnen die Übernahme von Erwachsenenpflichten erwartet. Das Jugendalter stellt damit eine Zeit des Übergangs, der Transition dar, die mit Veränderungen in vielen Bereichen verbunden ist. Diese Veränderungen geschehen in der Regel schrittweise mit unterschiedlicher Geschwindigkeit und vorübergehenden Inkonsistenzen.

Wie sich Jugendliche fühlen und verhalten und ob es eine kritische oder eher unproblematische Phase im Lebenslauf ist, wurde und wird kritisch diskutiert. Wenn antike Keilschriften zutreffen, so waren schon die Jugendlichen im Königreich Ur vor ca. 4000 Jahren heruntergekommen, zuchtlos und hörten nicht auf ihre Eltern, so dass der Weltuntergang als unmittelbar bevorstehend befürchtet wurde. Und auch ein Sokrates zugeschriebener Ausspruch lässt ihn über faule Heranwachsende mit schlechten Manieren klagen, die Eltern und Lehrer tyrannisieren. Als Begründung für den vermeintlich katastrophalen Zustand der immer neuen, immer wieder gleichen „Jugend von heute“ lässt sich deren chaotische Gefühls- und Erlebenswelt vermuten: „Sturm und Drang“ wie die gleichnamige Literaturepoche mit kraftvoller Dynamik und Veränderung der jugendlichen Gefühlswelt. Idealtypisch Goethes Werther, der ob der unerfüllbaren Liebe zu seiner angebeteten Lotte schier verzweifelt und sich letztlich selbst das Leben nimmt (1986, Original 1774). Nach Buchveröffentlichung bricht ein regelrechtes „Werther-Fieber“ aus, bei dem einige Jugendliche ihrem Romanvorbild sogar bis in den Selbstmord nachfolgen. Bis heute wird der „Werther-Effekt“ als ein kausaler Zusammenhang zwischen in Medien gezeigten Suiziden und einem Ansteigen der Suizidrate in der Bevölkerung thematisiert (Ziegler & Hegerl, 2002).

In die wissenschaftliche psychologische Literatur führt Hall den „Sturm und Drang“ als „storm and stress“ ein (1904). Er benennt Konflikte mit den Eltern, Stimmungsschwan-

kungen und Risikoverhalten als die drei zentralen Problembereiche dieser Phase. Die Darstellung des Jugendalters als kritischer Übergang findet sich auch später in der psychologischen Literatur, stark auch in der psychoanalytischen Tradition. Programmatisch betitelt ist Eriksons „Identity, Youth, and crisis (1968) bzw. „Jugend und Krise“ (1970), welches das psychologische Verständnis von Adoleszenz langfristig geprägt. Pointiert zusammengefasst z. B. als „Krach, Kummer und Krise“ (Fend, 2005, S. 28): Krach v. a. mit Eltern im schrittweisen Ablösungsprozess, Kummer mit dem eigenen sich verändernden Körper und Schwierigkeiten in wichtiger werdenden freundschaftlichen und beginnenden romantischen Beziehungen und schließlich Krise, weil steigende Anforderungen von außen nicht mit eigenen Bedürfnissen und der Selbstwahrnehmung vereinbar scheinen.

Ob die Jugendphase tatsächlich als Phase der normativen Krise verstanden werden sollte, erscheint im Lichte neuerer empirischer Forschung eher zweifelhaft (Arnett, 1999). Obwohl „Sturm und Drang“ im Jugendalter wahrscheinlicher auftritt als in anderen Lebensphasen, erleben ihn längst nicht alle Jugendlichen. Es gibt viele Beschreibungen der eigenen Jugendzeit als relativ problemlos und empirische Untersuchungen, die zeigen, dass die Adoleszenz zumeist ohne dramatische Krisen abläuft (Fend, 2005; Grob & Jaschinski, 2003).

Die Überschätzung krisenhaft verlaufender Jugendphasen mag unter anderem damit zusammenhängen, dass der Fokus bei der Untersuchung dieser Lebensphase oftmals auf problematisches Verhalten gelegt wird, nicht aber auf normale, unproblematische Verläufe. Steinberg und Morris (2001) formulieren drei allgemeine Prinzipien bei der Beurteilung adoleszenten Problemverhaltens: 1. Es ist wichtig zwischen gelegentlichem, experimentellen Verhalten und überdauernden Verhaltensmustern zu unterscheiden. Viele problematische Verhaltensweisen von Jugendlichen weisen einen eher singulären, ausprobierenden Charakter auf und verfestigen sich nicht. 2. Eine Unterscheidung zwischen Problemen, die im Jugendalter beginnen, und solchen, die schon zuvor begonnen haben, ist sinnvoll. Tatsächlich zeigen einige Jugendliche kriminelle oder delinquente Verhaltensmuster, bei genauerer Analyse stellt sich aber oftmals heraus, dass schon zuvor Probleme in verschiedenen Bereichen bestanden. Ein Problem mag sich zwar in der Adoleszenz besonders deutlich zeigen, das bedeutet aber nicht, dass es erst in der Adoleszenz begonnen hat. 3. Problematische Verhaltensweisen wie Drogenmissbrauch sind bei den meisten Jugendlichen vorübergehend und auf die Adoleszenz beschränkt. Moffitt (1993) unterscheidet daher zwischen „adolescence-limited“ und „life-course-persistent“ problematischen Verhaltensweisen. Beim Übergang zum Erwachsenenalter legen die meisten Jugendlichen das problematische Verhalten ab und übernehmen relativ konform die vorgesehenen Erwachsenenrollen ohne großes Maß an abweichendem Verhalten. Ob doch ein Risiko der Verfestigung problematischer Verhaltensmuster besteht, lässt sich nach Moffitt insbesondere darüber vorhersagen, ob schon in vorangegangenen Lebensphasen Probleme aufgetreten sind. Adoleszentes Problemverhalten sollte also immer im Kontext des gesamten Entwicklungsverlaufes beurteilt werden.

2.1 Definition

Die Phase zwischen der Kindheit und dem Erwachsenenalter wird mit unterschiedlichen Begriffen benannt, die kurz geklärt werden sollen (Fend, 2005):

- *Jugend* entstammt insbesondere dem soziologischen Sprachgebrauch und meint die soziale Gruppe der Jugendlichen, die sich von der Gruppe der Kinder, der Erwachsenen usw. unterscheidet.
- *Adoleszenz* ist ein Begriff, der in psychologischen Kontexten zumeist genutzt wird. Er betont das spezifische psychische Erleben in dieser Altersphase. Der Begriff wird als „adolescence“ im amerikanischen Sprachraum vorrangig benutzt.
- *Pubertät* entstammt der Biologie und bezieht sich vorrangig auf körperliche Veränderungen in dieser Altersphase.

In dieser Ausarbeitung werden die Begriffe Jugend und Adoleszenz weitgehend synonym verwendet, wobei an die Gruppe der Jugendlichen mit ihrer besonderen psychischen Entwicklungssituation gedacht ist. Der Begriff Pubertät wird nur benutzt, wenn die korrespondierenden biologischen Veränderungen betont werden sollen.

Der Beginn der Jugendphase wird zumeist einheitlich mit dem Eintreten der Geschlechtsreife definiert. Bei Mädchen ist dies durchschnittlich im Alter von 11 Jahren, bei Jungen im Durchschnitt mit 12 Jahren der Fall. Diese biologische Definition wird gewählt, da sich die vollziehenden Veränderungen nicht nur auf die körperliche Ebene beschränken, sondern sich weitreichend auch gravierend auf psychische und soziale Prozesse auswirken. Es kommt zu einer umfassenden Neuorientierung und Veränderung in vielen Lebensbereichen (Klaus Hurrelmann, 2007).

Weitaus weniger eindeutig lässt sich das Ende der Jugendphase abstecken. Definitionen, die lediglich das chronologische Alter als Maßstab einbeziehen sind wenig hilfreich, da sich die Entwicklung mit einer großen interindividuellen Variabilität und Inkonsistenz auf verschiedenen intraindividuellen Ebenen vollzieht. Biologische Definitionen beziehen sich vor allem auf das Ende der Pubertät und den damit abgeschlossenen körperlichen Veränderungen im Alter von 14 bis 16 Jahren. Jugendliche in diesem Alter werden jedoch kaum als erwachsen bezeichnet, da sie wesentliche Rollen und Erwartungen noch nicht erfüllen. Der Übergang vom Jugendalter ins Erwachsenenalter wird daher oft soziologisch gefärbt über die Übernahme bestimmter Funktionen und Rollen definiert, so der Berufsrolle, der Partner- und Familienrolle, der Konsumentenrolle und der Rolle als mündiger Bürger. In dieser Sichtweise ist der Übergang in das Erwachsenenalter vollzogen, wenn alle das Erwachsenenalter konstituierenden Teilrollen erfüllt werden. Die Rollenübernahme läuft dabei schrittweise ab und wird nicht notwendigerweise vollständig abgeschlossen; der Übergang ins Erwachsenenalter ist fließend (Hurrelmann, 2007). Die Teilrollen korrespondieren mit den Entwicklungsaufgaben des Jugendalters,

die noch beschrieben werden. Kurz und bündig zusammengefasst: „Die Adoleszenz beginnt in der Biologie und endet in der Kultur“ (Herbert, 1989, S. 15)

Im historischen Wandel zeigt sich eine Ausweitung der Lebensphase Jugend, die sich dadurch erklärt, dass sich einerseits die Geschlechtsreife nach vorn verlagert und andererseits die Übernahme von Erwachsenenrollen nach hinten verschiebt, v. a. in den Bereichen Beruf und Familiengründung (Hurrelmann, 2007, S. 39). Das Jugendalter erstreckt sich über ca. 10 bis 15 Jahre, wobei es starke Variabilität zwischen Jugendlichen gibt, die maßgeblich mit der Dauer des Verbleibes in der schulischen und universitären Ausbildung zusammenhängt. Jugendliche, die nach der Sekundarstufe 1 eine Ausbildung absolvieren, treten viel eher ins Berufsleben ein und werden finanziell unabhängig als solche, die die gymnasiale Oberstufe besuchen und dann ein Studium absolvieren und damit viel länger in finanzieller Abhängigkeit bleiben. Erstere Jugendliche werden im Sinne der obigen Definition früher erwachsen als letztere.

Die zeitliche Strukturierung der Jugendphase wird in der Literatur nicht einheitlich vorgenommen. Psychologische Definitionen unterscheiden meist die frühe, mittlere und späte Adoleszenz mit unterschiedlichen Altersangaben, die auf eine große Variabilität der individuellen Entwicklung zurückgeführt wird (siehe Tabelle 1).

Tabelle 1: Uneinheitliche Einteilung der Adoleszenz nach Lebensalter in verschiedenen Definitionen (Auswahl)

	Frühe Adoleszenz	Mittlere Adoleszenz	Späte Adoleszenz
Steinberg, 2005a	10 – 13	14 – 17	18 – 22
Smetana, Campione-Barr, & Metzger, 2006	11 – 13	14 – 17	18 – Anfang 20
Berk, 2011	11/12 – 14 rasche pubertäre Veränderungen	14 – 16 pubertäre Verände- rungen beinahe ab- geschlossen	16 – 18 Erwachsenes Aus- sehen, Beginn Übernahme Er- wachsenenrollen

Soziologische Definitionen sind weiter gesteckt und unterteilen die Lebensphase Jugend in die frühe Jugendphase (12 – 17 Jahre, pubertäre Phase), mittlere Jugendphase (18 – 21 Jahre, nachpubertäre Phase) und späte Jugendphase (22 – 27 Jahre, Übergangszeit zur Erwachsenenrolle) (Hurrelmann, 2007, S. 41). Um die Instabilität der Lebensphase zwischen 18 und 25 Jahren herauszustellen, wurde von Arnett (2000) der Begriff der „emerging adulthood“ als eigenständiger Entwicklungsphase geprägt. Damit wird die Veränderbarkeit, Unsicherheit und auch die voranschreitende Identitätsentwicklung in diesem Lebensabschnitt betont.

2.2 Entwicklungspsychopathologie

Menschliche Entwicklung vollzieht sich in kontinuierlichen Wechselwirkungen zwischen einem aktiven Individuum und einer aktiven Umwelt vor dem zeitlichen Verlauf der Lebensspanne. Dabei spielen genetische Grundlagen (Genotypen), die in Wechselwirkung mit Umweltfaktoren („enviromtype“ nach Sameroff) die Phänotypen, also verschiedene Merkmale eines Menschen, bestimmen. Die Diskussion um die Vorherrschaft von Anlage und Umwelt, nature vs. nurture, wird immer mehr zu einer integrativen Sicht geführt, bei der beide nicht mehr trennbar miteinander verflochten sind. Es geht nicht mehr um die Frage, wie viel Anteile jeder Faktor hat, sondern um Mechanismen, wie Anlage und Umwelt miteinander interagieren mit dem Ziel effektive Umweltinterventionen zu entwickeln und durchzuführen (Petermann, Niebank, & Scheithauer, 2004).

Sameroff (2000) spricht in diesem Zusammenhang von einem „transaktionalen“ Entwicklungsmodell. Dabei beeinflussen sowohl Genotyp als auch „enviromtype“ den Phänotypen und werden wiederum von diesem reguliert (siehe Abbildung 1). Dass Genotypen auch von Phänotypen verändert werden können, widerspricht „Crick's zentralem Dogma“ der unidirektionalen Wirkungsweise von Genen auf Entstehung von Proteinen und Zellen, die wiederum Verhalten beeinflussen. Mittlerweile lässt sich allerdings zeigen, dass auch Gene durch Umwelteinwirkungen verändert werden können und die Entdeckung der Plastizität des Gehirns, bei der sich das Gewebe des Kortex flexibel an Umwelteinflüsse anpasst, stellen das Dogma in Frage (Petermann et al., 2004).

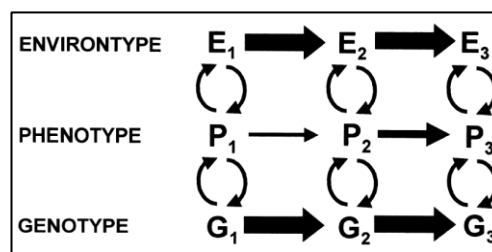


Abbildung 1: Modell der transaktionalen Entwicklung (aus Sameroff, 2000, S. 309)

Die Annahme einer wechselwirksamen bzw. transaktionalen Entwicklung beeinflusst die Interpretation psychischer Störungen und Abweichungen. So werden psychische Abweichungen als das Ergebnis einer Vielzahl verwobener adaptiver Entwicklungsschritte gesehen. Eine psychische Störung ist nicht etwas, das eine Person „hat“ oder eine der Person innewohnende Pathologie, die sich im Verhalten äußert, sondern eine Anpassung einer Person an ihre Umgebung (Sroufe, 1997). Forschungsfragen in diesem Bereich thematisieren die Identifikation von Risiko- und Schutzfaktoren, Entwicklungspfaden normativer und abweichender Entwicklung, Resilienz und Komorbidität. Bei der Entwicklungspsychopathologie handelt sich um keine singuläre Theorie, sondern um eine integrative Disziplin, die versucht in einem entwicklungsbezogenen Lebensspannenrahmen Erkenntnisse verschiedener Forschungslinien zu vereinen. Dabei fungiert sie

als eine Art „Makroparadigma“ (Achenbach 1990, zitiert nach Cicchetti & Rogosch, 2002) oder wird als eine „Sichtweise“ bezeichnet (Resch, 1999).

Sroufe (1997) grenzt das „Entwicklungsmodell“ vom klinischen Modell ab, welches kindliche und jugendliche Verhaltens- und emotionale Probleme mit Krankheiten in Analogie bringt. Dabei werde oft nach einem abgrenzbaren einzigen Krankheitsverursacher wie einer bestimmten genetischen Komponente gesucht, welches die Störung erkläre. In einem Entwicklungsmodell aber sind Organismus und Kontext untrennbar miteinander verbunden; Abweichung entstehe durch eine einzigartige Kombination des Wirkens von Risiko- und Schutzfaktoren. Dem Individuum kommt dabei eine aktive Rolle zu, insbesondere die Geschichte vorangegangener Adaptation sei zentral. Zur Veranschaulichung des Entwicklungsmodells eigne sich das Konzept der Entwicklungspfade. Pathologie werde dabei gedacht als eine Abfolge von Verzweigungen, die das Kind von einem Pfad zu kompetentem Funktionieren wegführe. Sroufe (1997) benennt fünf wichtige Implikationen des Modells:

1. Das Verständnis von Störung setzt ein Verständnis der normativen Entwicklung mit den verschiedenen Arten positiver Entwicklung voraus. Eine bedeutsame Abweichung stelle eine erhöhte Wahrscheinlichkeit dar, wichtige Entwicklungsaufgaben nicht bewältigen zu können. Pathologie stelle damit ein Ergebnis wiederholter misslungener Anpassung dar. Maladaptive Muster werden nicht per se als pathologisch gesehen, sondern als ein Entwicklungsrisiko.
2. Verschiedene Entwicklungspfade können zu ähnlichen Ergebnissen führen, was als Äquifinalität bezeichnet wird. Individuen sind phänotypisch ähnlich, haben aber eine unterschiedliche Prognose und benötigen verschiedene Interventionen. Moffitt (1993) konnte das in Bezug auf antisoziales Verhalten zeigen.
3. Individuen, die auf dem gleichen Pfad starten, können einen anderen Ausgang nehmen (Multifinalität).
4. Veränderung ist vielfach möglich. Pathologie wird nicht als etwas verstanden, das das Kind „hat“ im Sinne einer permanenten Kondition. Die Rückkehr zu positiver Adaptation besteht vielfach.
5. Je länger ein Kind allerdings auf einem maladaptiven Entwicklungspfad unterwegs ist, desto weniger wahrscheinlich wird es, dass eine Rückkehr zu positiver Adaptation möglich ist.

Zentrale Bedeutung haben frühe Adaptationsmuster, die als Prototypen wirken. Daher ist die Erfassung der Entwicklungspfade schon ab einem sehr frühen Zeitpunkt sehr wichtig. Früher Beginn, Dauer und Anzahl der Probleme komme eine große prognostische Aussagekraft zu.

Ein wichtiges Konzept im Entwicklungsmodell ist das der Resilienz. Damit ist gemeint, dass manche Kinder trotz widriger Bedingungen einen guten Entwicklungsverlauf nehmen bzw. nach einer Fehlanpassung wieder zu einem positiven Weg zurückkehren. Im medizinischen Modell werde Resilienz oft als ein dem Kind innewohnende Eigenschaft („trait“) verstanden und nicht als sich verändernder Zustand („state“) wie im Entwicklungsmodell. Die Erklärung, warum manche Kinder trotz widriger Umstände keine Störung entwickeln, werde schlicht damit begründet, dass sie resilient seien. Sroufe (1997) vergleicht dies damit, wenn die Frage warum ein Kind Aufmerksamkeitsprobleme habe, damit beantwortet werde, dass es ADHS habe. Gleichzeitig werde die Diagnose aber gestellt, weil ein Kind Aufmerksamkeitsprobleme habe, was somit tautologisch sei. Im Entwicklungsmodell wird Resilienz als eine Entwicklung der Person in Interaktion mit seiner Umwelt über vielfache Anpassungsschritte konzeptualisiert. Die vorhergehende Anpassungsgeschichte des Individuums kann dazu wichtige Hinweise liefern.

Eine klassische Untersuchung im Zusammenhang mit Resilienz ist die Kauai-Längsschnittstudie, bei der 72 resiliente Kinder über 35 Jahre begleitet wurden (Werner, 1993). Trotz schwieriger Bedingungen wie Komplikationen bei der Geburt oder Armut entwickelten sich diese Kinder im Vergleich zu den anderen Kindern der Risikogruppe sehr positiv. Werner untersuchte Resilienz- bzw. risikomildernde Faktoren in unterschiedlichen Lebensphasen, z. B. umgängliches Temperament als Baby, positive Eltern-Kind-Interaktionen oder größerer schulischer Erfolg. Solche Merkmale, Kompetenzen und Schutzbedingungen lassen sich auch in mögliche kausale Zusammenhänge überführen (siehe Abbildung 2).

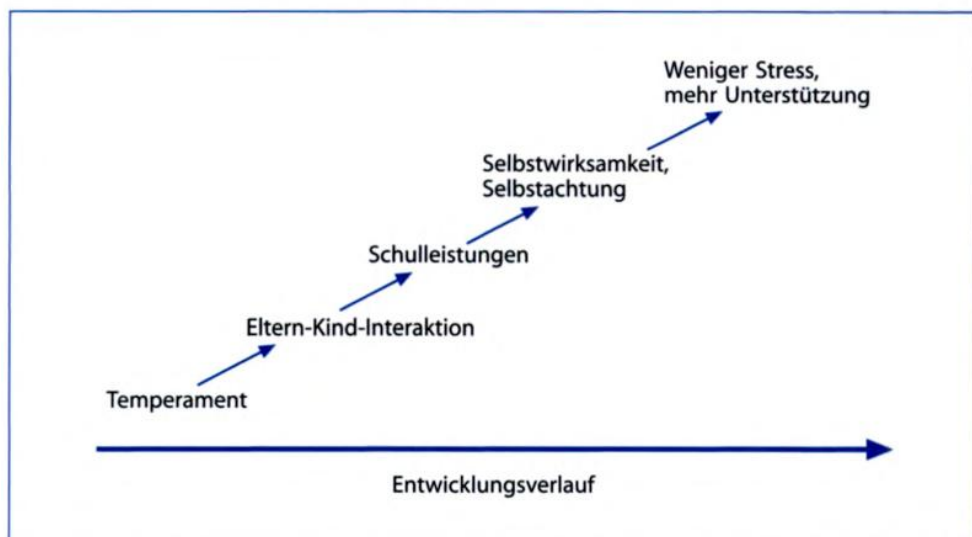


Abbildung 2: Mögliche kausale Zusammenhänge zwischen den von Werner (1993) untersuchten risikomildernden Bedingungen (aus Petermann, Niebank & Scheithauer, 2004, S. 350)

Im Licht des Entwicklungsmodells erscheint auch das Konzept der Komorbidität in einem anderen Licht. Komorbidität bezeichnet den Umstand, dass viele Patienten mehr als eine Diagnose erfüllen und gleichzeitig mehrere abgrenzbare Störungsbilder und

Probleme vorliegen (Feinstein, 1970). Im Kindesalter tritt beispielsweise häufig eine Störung des Sozialverhaltens und ADHS gemeinsam auf (Petermann, Niebank & Scheithauer, 2004). Die hohen Angaben zu Komorbiditäten bei emotionalen und Verhaltensstörungen zeigen, dass Komorbidität eher die Regel als die Ausnahme ist. Daher erscheint es fraglich, ob die Abgrenzung der Störungen und damit die Integrität der Kategorien zutreffend sind. Nach Sroufe (1997) könnten die Erkenntnisse zur Komorbidität die Klassifikation kindlicher psychischer Probleme revolutionär verändern und zu Klassifikationsschemata führen, die nicht allein an der aktuellen Symptomatik orientiert seien, sondern an Adaptationsmustern und entwicklungsgemäßen Verlaufsmustern.

Es wird deutlich, dass sich das Entwicklungsmodell als nicht vorherbestimmt, nicht deterministisch versteht. Es geht von einer Vielzahl denkbarer Entwicklungspfade aus, die entweder normativ und damit adaptiv sind oder davon abweichend. Der deterministischen, vermeintlich medizinischen Sichtweise, wird eine probabilistische entgegengestellt, bei der Risiko- und Schutzfaktoren die Wahrscheinlichkeiten für das Auftreten von Pathologien beeinflussen im Sinne einer Fuzzylogik (Rutter, 2005). Forschung im Entwicklungsmodell fokussiert daher auf Kontinuitäten und Veränderung, Prozesse und Mechanismen, die Individuen auf bestimmten Entwicklungspfaden halten oder sie von anderen wegleiten. Pathologie ist nicht gegeben, sondern entwickelt als ein Muster der Anpassung, das den gesamten Entwicklungskontext zu einem bestimmten Zeitpunkt abbildet.

Sroufe (1997) sieht bei Kindern und Jugendlichen mit emotionalen und Verhaltensproblemen eher einen quantitativen Unterschied zu anderen Kindern und Jugendlichen. Dagegen grenzt er klar Störungen wie Autismus ab, die als tiefgreifende Entwicklungsstörungen bezeichnet werden, was den qualitativen Unterschied im Denken und Handeln dieser Kinder und Jugendlichen betonen soll. Werde beispielsweise ADHS als eine genetisch bedingte, im Kind verankerte Störung wahrgenommen, so werde der Aspekt der gegenseitigen Beeinflussung mit anderen Faktoren unterschätzt. Behandlung der Wahl sei damit die medikamentöse Intervention auf der biologischen Ebene, während Veränderungen der Umgebung im Wesentlichen als Symptommanagement gesehen werden.

Der Verlauf der Störungen vom Kindesalter über die Adoleszenz bis ins Erwachsenenalter zeichnet sich zumeist homotypisch ab. D. h. dass eine Störung bestehen bleibt bzw. später dieselbe Störung erneut auftritt. Dem Vorliegen einer bestimmten Störung komme somit prädiktive Aussagekraft für das Vorliegen genau dieser Störung zu einem späteren Zeitpunkt zu. Diesem Verlauf liegt zugrunde, dass eine Störung sich über den Entwicklungsverlauf erstreckt. Es gibt aber auch heterotypische Verlaufsmuster, bei denen eine Störung eine andere Störung vorhersagt. Dabei wird vermutet, dass eine zugrunde liegende Problematik sich je nach Entwicklungsstadium in unterschiedlichen phänotypischen Ausprägungen zeigt. Empirische Evidenz finden insbesondere drei Muster heterotypischer Prädiktion: Angst und Depression können sich über Kreuz von Kindheit bzw. Adoleszenz zum Erwachsenenalter vorhersagen; Störungen des Sozialverhaltens und oppositionelles Verhalten erhöhen die Wahrscheinlichkeit für das Vor-

liegen von Angst und Depression im Erwachsenenalter und Störungen des Sozialverhaltens können Substanzmissbrauch im Erwachsenenalter vorhersagen (Costello, Copeland & Angold, 2011).

Die Anfänge dieses zusammengesetzten entwicklungswissenschaftlichen Paradigmas gehen auf die Forschung Sameroffs (Elkind & Sameroff, 1970; Sameroff, 2000) zurück. In der Debatte um „nature versus nurture“, also der Frage, ob Erbe oder Umwelt unser Verhalten bestimmen, ging Sameroff davon aus, dass beide Komponenten nicht zu trennen seien. Beide bedingen sich gegenseitig und über die Zeit kontinuierlich wechselseitig. Bezogen auf Störungen oder problematische Verhaltensweisen von Kindern und Jugendlichen bedeutet dies, dass Gründe nicht ausschließlich in einer biologischen Disposition liegen oder in unangemessenem Erziehungsverhalten, sondern dass sich beides gegenseitig bedingt und schließlich nicht mehr voneinander getrennt beurteilbar ist. Untersucht werden dabei sowohl proximale als auch distale Einflussgrößen. Sameroff nutzt angelehnt an die Biologie und Genetik die Begriffe Genotyp und Phänotyp und stellt diesen eine weitere Kategorie, den „environment“ zur Seite, der die Einflüsse der Umwelt darstellt.

2.3 Entwicklungsaufgaben

Das Konzept der Entwicklungsaufgaben wurde von Havighurst (1948) eingeführt und findet seitdem Anwendung in der Entwicklungspsychologie. Es verknüpft fortlaufende gesellschaftliche, biologische und subjektiver Aspekte der menschlichen Entwicklung (Flammer & Alsaker, 2002). Daher ist es gut geeignet, die menschliche Entwicklung im Sinne einer „Entwicklungspsychologie der Lebensspanne“ (Brandstädter & Lindenberger, 2007) abzubilden. Havighurst ging davon aus, dass jeder Mensch im Lauf seines Lebens vor bestimmte für die jeweilige Entwicklungsphase typische Aufgaben gestellt wird. Dies beginnt damit, krabbeln, laufen und sprechen zu lernen, führt über die Bewältigung schulischer Herausforderungen über die Ablösung von den Eltern und Gründung einer eigenen Familie bis zur Auseinandersetzung mit dem eigenen Ruhestand und gesundheitlichen Problemen im höheren Alter. Einige Entwicklungsaufgaben sind dabei in einer festgelegten Reihenfolge zu bewältigen. So geht in der Regel der Gründung einer Familie mit eigenen Kindern das Eingehen einer Partnerschaft voraus. Die Bewältigung von Entwicklungsaufgaben früherer Entwicklungsphasen führt zur Erlangung von Kompetenzen, die für spätere Aufgaben benötigt werden. Tendenziell werden die Entwicklungsaufgaben als universell gültig gedacht, aber im Zuge gesellschaftlicher Veränderungsprozesse verändern sich auch die Entwicklungsaufgaben. So hat sich beispielsweise der Stellenwert der Ehe seit den 1950er Jahren stark verändert (Nave-Herz, 2007). Soziologen stellen heute eine starke Individualisierung von Lebensläufen und eine Pluralisierung von Lebensstilen fest.

Als Ursprung von Entwicklungsaufgaben nennen Grob & Jaschinski (2003) drei Quellen: 1. physische Reifungsprozesse, 2. gesellschaftliche Erwartungen, die auf kulturellen

Normen basieren, und 3. individuelle Zielsetzungen und Werthaltungen. Oftmals ergeben sich konkrete Entwicklungsaufgaben aus einem Zusammenspiel mehrerer Quellen. Im Jugendalter finden sowohl starke körperliche Veränderungen im Rahmen der Pubertät statt und gleichzeitig werden an Jugendliche von ihrer Umgebung höhere Anforderungen gestellt, so sollen sie viel stärker als Kinder Verantwortung für ihr eigenes Handeln übernehmen. Schließlich entwickeln Jugendliche eigene Ideen und Wünsche für die Gestaltung ihrer Zukunft. Einige Entwicklungsaufgaben sind zeitlich beschränkt wie das Erlernen des Laufens, andere sind zeitlich unbegrenzt wie die Beziehungsgestaltung zu Gleichaltrigen, die über die gesamte Lebensspanne relevant ist. Manche Aufgaben bauen aufeinander auf und müssen deshalb nacheinander bearbeitet werden, andere können oder müssen parallel bearbeitet werden. Jugendliche müssen in der Regel erst einen Schulabschluss erwerben, um dann einen Ausbildungs- oder Studienplatz zu erlangen. Hingegen müssen sie gleichzeitig die Ablösung von den Eltern bewerkstelligen und gleichzeitig eine Annäherung an die Gleichaltrigen erreichen. Weiterhin lassen sich normative und nicht-normative Entwicklungsaufgaben unterscheiden. Die meisten Entwicklungsaufgaben haben normativen Charakter und gelten für die meisten Menschen in einem bestimmten Kulturkreis. Diese sind einfacher zu bewältigen, da sie vorherzusehen sind. Nicht-normative Entwicklungsaufgaben stellen sich dagegen nicht allen Menschen und ergeben sich z. B. nach kritischen Lebensereignissen wie dem Tod von Familienmitgliedern. Diese Entwicklungsaufgaben sind oftmals schwerer zu bewältigen, da sie unerwartet auftreten und da die Erfahrung mit den meisten Gleichaltrigen nicht geteilt werden kann (Flammer & Alsaker, 2002).

Nach Fend (2005) lassen sich die Entwicklungsaufgaben des Jugendalters in drei Bereiche einordnen: 1. intrapersonal oder persönlich, 2. interpersonal oder zwischenmenschlich und 3. kulturell-sachlich oder gesellschaftlich. Ergänzt durch Hurrelmann (2007) fällt in den Bereich der intrapersonalen Entwicklungsaufgaben die Akzeptanz des sich verändernden eigenen Körpers, der erwachsene Umgang mit Sexualität und die Übernahme einer stabilen Geschlechtsidentität. Weiterhin finden sich hier leistungsbezogene Entwicklungsaufgaben, die mit der Entwicklung intellektueller Kompetenzen und der Bewältigung schulischer Anforderungen in Verbindung stehen. Das Ziel in diesem Bereich ist der Erwerb eines schulischen Abschlusses als Voraussetzung für die Aufnahme eines Berufes. Die Aufnahme einer Erwerbsarbeit fungiert als Grundlage der finanziellen Unabhängigkeit und einer selbstständigen Existenz. In den Bereich der interpersonalen Entwicklungsaufgaben fallen die Entwicklung sozialer Kompetenzen zum Aufbau von Freundschaften und einer partnerschaftlichen Beziehung und auch die Umstrukturierung der Beziehung zur Herkunftsfamilie, die sich sowohl durch Ablösungs- als auch durch Wiederannäherungsprozesse auszeichnet. Als Ziel in diesem Bereich lässt sich die Einrichtung sozialer Unterstützungssysteme formulieren, klassischerweise die Gründung einer eigenen Familie mit eigenen Kindern. In den kulturell-sachlichen bzw. gesellschaftlichen Bereich fallen der Bereich Konsum und Freizeit und die Herausbildung einer ethischen und politischen Orientierung. Ziele sind ein kontrollierter Umgang mit Freizeitangeboten, Geld, Drogen und Medien. Zusätzlich soll ein eigenes Werte- und Normsystem gebildet werden und das eigene Handeln in Übereinstimmung mit

diesem gebracht werden. Das Ziel ist die verantwortliche Übernahme gesellschaftlicher Partizipation. Aktuell formulieren Hurrelmann und Quenzel (2012) diese vier Ziele als „Binden“, „Qualifizieren“, „Konsumieren“ und „Partizipieren“.

Von diesen Entwicklungszielen kann auch eine Definition des Übergangs ins Erwachsenenalter abgeleitet werden. Jemand ist demnach erwachsen, wenn er bestimmte Rollen übernommen hat, die für das Erwachsenenalter konstitutiv sind. Korrespondierend mit den genannten Entwicklungsaufgaben sind dies die Berufsrolle als finanziell selbstständiger Mensch, die Partner- und Familienrolle, die Konsumentenrolle und die Rolle als politischer Bürger mit einer eigenen Wertorientierung. Charakteristisch ist, dass die Rollenübernahme unterschiedlich schnell abläuft. Hurrelmann (2007) kennzeichnet diese Statusinkonsistenz als ein wesentliches strukturelles Merkmal des Jugendalters. Damit ist gemeint, dass Jugendliche in einigen Bereichen schon sehr früh selbstständig sind, z. B. in der Mediennutzung und im Warenkonsum. Im Vergleich mit vorangehenden Generationen zeigt sich auch eine frühere Selbstständigkeit im Bereich Sexualität und Finanzen. Hingegen zeigt sich eine spätere Übernahme der Rolle als Familiengründer mit einem höheren Durchschnittsalter, in dem Kinder geboren werden.

Zum Umgang mit Entwicklungsaufgaben unterscheiden Flammer und Alsaker (2002) analog zu Seiffge-Krenke (1995) drei Arten von Strategien: 1. aktive, aufgabenorientierte Problemlösung, 2. kognitive Umwertung und 3. Vermeidung und Ablenkung. Welche Strategie sinnvoll und Erfolg versprechend ist, hängt von dem jeweiligen Kontext ab. Prinzipiell lösbare Aufgaben sind am besten aktiv anzugehen. Sind Entwicklungsaufgaben momentan nicht lösbar, kann eine Umbewertung sinnvoll sein und manchmal auch eine Vermeidung der Auseinandersetzung mit dieser Aufgabe. So kann der Selbstwert geschützt werden, das Problem bleibt allerdings bestehen.

2.4 Kognitive Entwicklung

Die kognitiven Fähigkeiten von Jugendlichen sind denen von Kindern deutlich überlegen. Im Vergleich fallen sowohl qualitative als auch quantitative Unterschiede in den Blick. Nach Oerter und Dreher (2008) ist der Denkstil von Jugendlichen als mehr formal und abstrakt zu bezeichnen, so sind sie ähnlich einem Wissenschaftler zu hypothetisch-deduktivem Denken in der Lage. Weniger stark als Kinder benötigen Sie für Ihre Überlegungen materielle Grundlagen, beziehen kontrafaktische Erwägungen mit ein und sind in der Lage zu Transfer und Übertragung auf andere Inhalte. Ein zentraler Fortschritt ist auch die größere Fähigkeit zu relativistischem Denken. Jugendliche berücksichtigen beim Denken und Entscheiden unterschiedliche Dimensionen und Perspektiven und wägen Werte und Regeln nach ihrer jeweiligen situativen und kontextuellen Angemessenheit gegeneinander ab. Jugendliche erlernen es, kritisch zu denken, zu diskutieren und zu argumentieren. Auch entwickeln sich metakognitive Kompetenzen bei der Planung, Überwachung und Steuerung des eigenen Denkens und Verhaltens. Neben diesen qualitativen Unterschieden zeigen sich auch deutliche quantitative Verän-

derungen. Jugendliche können sich Dinge besser merken, haben eine größere Aufmerksamkeitskapazität und können Informationen schneller verarbeiten.

Vergleicht man also kontextunabhängig ein Kind mit einem Jugendlichen hinsichtlich ihrer kognitiven Fertigkeiten, werden gravierende Unterschiede in der Qualität ihres Denkens offenbar. Diese qualitativen Unterschiede haben Entwicklungspsychologen dazu geführt, von unterschiedlichen Stadien oder Stufen der Entwicklung des Denkens auszugehen. Einer der prominentesten und ersten Vertreter dieser Forschungstradition ist Piaget, der aus Beobachtungen von Kindern und den von ihnen gemachten Fehler ein vierstufiges Entwicklungsmodell aufgestellt hat (Piaget, 1947). Dieses einflussreiche Modell und seine (quasi-)experimentellen Versuchsaufgaben sind vielfach repliziert und zunehmend auch kritisch betrachtet worden. Mittlerweile belegen viele Überprüfungen, dass das Stufenmodell in seiner strengen Auslegung nicht in allen Annahmen zutrifft. Insbesondere konnte gezeigt werden, dass Kinder niedriger Stufen teilweise auch Lösungen höherer Stufen benutzen, so dass von einer großen Variabilität auszugehen ist, die mit dem Piaget'schen Modell nur schwer in Einklang gebracht werden kann (Siegler, 2007).

Ob die Unterschiede besser als qualitative oder quantitative Fortschritte zu beschreiben sind, wurde und wird in der Entwicklungspsychologie viel diskutiert. Strukturgenetische Ansätze betonen die qualitativen Veränderungen, während informationsverarbeitende Ansätze stärker die quantitativen Veränderungen in den Blick nehmen (Sodian, 1998). Im Folgenden werden nach einer Vorstellung der biologischen Grundlagen ausgewählte Theorien der kognitiven Entwicklung kurz dargestellt. Es wird begonnen mit dem Piaget'schen Stufenmodell als Urgedanken, insbesondere mit Piagets Blick auf die Jugendphase, nachfolgend ein quantitativer Ansatz nach Baltes bzw. Cattell, anschließend eine Integration im informationsverarbeitenden Ansatz nach Case und eine Beschreibung der Entwicklung zentraler Einzelaspekte der Kognition.

2.4.1 Gehirnentwicklung als biologische Grundlage

Der Beginn der Adoleszenz ist gemäß gängigen Definitionen durch das Einsetzen der körperlichen Veränderungen im Zuge der Pubertät festgelegt. Ausgelöst durch verschiedene hormonelle Vorgänge, bei denen genetisch beeinflusst Wachstumshormone ausgeschüttet werden, beginnt eine Phase der körperlichen Umstrukturierung. Erstes sichtbares Anzeichen ist daher ein Wachstumsschub (Berk, 2011). Auch das Gehirn entwickelt sich kontinuierlich von vor der Geburt bis in die späte Adoleszenz. Das zerebrale Gesamtvolumen verändert sich zwar nach dem fünften Lebensjahr nicht mehr signifikant, es finden aber bis in die Adoleszenz gleichzeitige regionale Volumenvergrößerungen und –verminderungen statt (Konrad, 2007). Daran beteiligt ist einer der wichtigsten gehirnanatomischen Entwicklungsprozesse, der als Myelinisierung bezeichnet wird. Die Axonen der Nerven bekommen dabei eine Ummantelung durch eine lipidreiche Schicht, die als Myelinscheide oder Markscheide bezeichnet wird. Durch diesen Vorgang erhöhen sich die Leitfähigkeit der Nervenzellen und der Schutz vor Erregung durch andere

Nervenzellen, die das Axon kreuzen. Vermutlich durch voranschreitende Myelinisierung der Axone nimmt der Anteil weißer Substanz des Gehirns bis ins Jugendalter hinein zu. Die graue Substanz nimmt ebenfalls über das Kindesalter hinweg zu, bis in der Adoleszenz ein Zelltod von Neuronen und Gliazellen einsetzt, der zu einer Abnahme des Gehirnvolumens bis zum Erwachsenenalter führt. Bildgebende Verfahren zeigen, dass sich in der Adoleszenz ein großer Verlust an grauer Substanz in verschiedenen Hirnarealen vollzieht, der mit dem Absterben von Synapsen in Verbindung gebracht wird. Der Verlust grauer Substanz vollzieht sich im präfrontalen Kortex über einen längeren Zeitraum als in anderen Teilen, z. B. dem parietalen Kortex. Gleichzeitig entwickelt sich die weiße Substanz, insbesondere durch Myelinisierung, besonders ausgeprägt im Frontallappen (Luciana, Conklin, Hooper & Yarger, 2005).

Der Anteil grauer Substanz im Frontalhirn scheint bei gesunden Kindern im Alter von 12 Jahren am höchsten zu sein. Das Volumen der Basalganglien und damit verbunden das Volumen des frontalen Kortex nimmt im Jugendalter ab. Diese Veränderungen werden als eine kritische Periode der Reifung der fronto-strialen Kreisläufe angesehen (Konrad, 2007). Der Grad der Myelinisierung und die Synapsendichte in den Frontallappen weisen auf eine sehr lange Reifungsphase hin. Die Konnektivität des medialen Temporallappens mit dem präfrontalen Kortex nimmt in der Adoleszenz zu. Das Volumen der grauen Substanz verändert sich bis ins junge Erwachsenenalter nicht-linear. In frontalen und parietalen Bereichen ist das Volumenmaximum bei ca. 12 Jahren, im Temporallappen erst mit etwa 16 Jahren und im Okzipitallappen erst mit ca. 20 Jahren (Gleissner, 2007).

Hirnareale, die mit sensorischen und motorischen Prozessen assoziiert sind, reifen eher. Hirnareale wie der präfrontale Kortex, die mit höheren kognitiven Funktionen in Verbindung gebracht werden, reifen bis in die späte Adoleszenz hinein (siehe Abbildung 3).

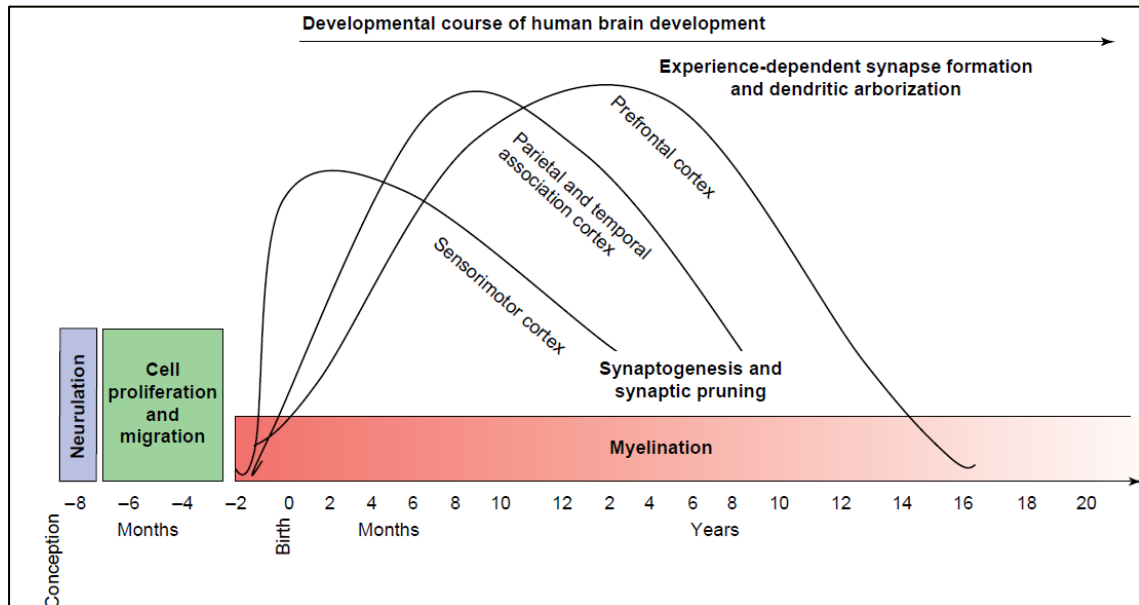


Abbildung 3: Strukturelle Gehirnentwicklung (aus Casey, Tottenham, Liston & Durston, 2005, S. 105)

Zusammenfassen lässt sich, dass die Gehirnentwicklung je nach Areal unterschiedlich verläuft, im präfrontalen Kortex am langsamsten. Eine vollständige Reifung ist oft erst in der mittleren Adoleszenz erreicht. Durch voranschreitende Myelinisierung erhöhen sich die Weiterleitungsgeschwindigkeit der Nerven und der Schutz vor Aktivierung durch benachbarte Nervenzellen. Wenig genutzte synaptische Verbindungen werden gekappt und neue synaptische Verbindungen entstehen. Die Aktivierungsmuster des Gehirns von Erwachsenen sind eher fokal als diffus und sind damit scheinbar effizienter (Konrad, 2007).

2.4.2 Qualitative Veränderungen

Qualitative Modelle gehen im Gegensatz zu quantitativen Ansätzen davon aus, dass sich die kognitive Entwicklung nicht kontinuierlich vollzieht, sondern dass es bestimmte Übergänge gibt, an denen ältere Denkmuster von neueren abgelöst werden. Piaget beobachtete dazu, welche Denkfehler Kinder verschiedenen Alters machen, und leitete daraus ein vierstufiges Modell der intellektuellen Entwicklung ab. Wird eine neue Stufe erlangt, stehen dem Kind neue Denk- und Handlungsmöglichkeiten zur Verfügung, wobei Erkenntnisse vorheriger Stufen integriert werden. Piaget ging davon aus, dass die Abfolge der zu durchlaufenden Stufen bei jedem Kind gleich sei und machte auch Angaben zu Alter. Diese sind allerdings eher als Orientierung zu sehen, da die interindividuelle Variabilität sehr groß ist.

Piaget (1947) unterscheidet vier Stufen der intellektuellen Entwicklung:

- (1) Sensorimotorisch (0 – 2 Jahre)

- (2) Präoperational (2 – 7 Jahre)
- (3) Konkrete Operationen (7 – 11 Jahre)
- (4) Formale Operationen (ab 11 Jahre)

Besondere Relevanz für das Jugendalter hat demnach die letzte Stufe. Sie zeichnet sich durch deutliche Veränderungen und Fortschritte im Denken aus. Es gelingt erstmalig kognitive Operationen durchzuführen, die nicht auf eine materielle Grundlage angewiesen sind. Dieses seien z. B. die gedankliche Umkehrung oder Negation von Handlungen. Jugendliche benötigen zum Denken weniger stark tatsächlich vorliegende, greifbare Objekte, sondern können mit deren verbalen oder symbolischen Repräsentationen operieren. Das Denken ist zudem abstrakter und auch auf die Zukunft bezogen. Der Jugendliche wird mit einem Wissenschaftler verglichen, der Hypothesen aufstellt, diese prüft und anhand der Fakten annimmt oder verwirft (Piaget, 1969).

Die Vorstellung von Entwicklungsstufen, die zwingendermaßen nacheinander erreicht werden müssen, und getroffene Altersangaben sind vielfach kritisiert worden und vermutlich überholt. Trotzdem bilden die Beobachtungen von Piaget eine treffende Beschreibung des jugendlichen Denkens.

2.4.3 Quantitative Veränderungen

Quantitative Modelle gehen von einer kontinuierlichen Entwicklung der kognitiven Leistungsfähigkeiten über die Lebensspanne aus. Einige Modelle unterscheiden zwei Komponenten mit jeweils spezifischem Verlaufsprofil. Die erste Komponente bezieht sich dabei auf grundlegende Prozesse der Informationsverarbeitung und wird von Cattell (1963) als „fluide Intelligenz“ bezeichnet bzw. als „mechanics“ bei Baltes und Kollegen (Baltes, Staudinger & Lindenberger, 1999). Es wird angenommen, dass diese „mechanischen“, inhaltlich nicht gebundenen Prozesse eng mit neurobiologischen Hirnfunktionen zusammenhängen, die genetisch prädispositioniert sind. Sie bestimmen wie schnell und akkurat Informationen verarbeitet werden können und koordinieren elementare Verarbeitungsschritte. Die zweite Komponente wird von Cattell als „kristalline Intelligenz“ bzw. als „pragmatics“ bei Baltes et al. bezeichnet. Sie ist wissensbasiert, stark kulturell geprägt und beinhaltet deklaratives und prozedurales Wissen, das über Generationen tradiert wird. Die intellektuelle Entwicklung hat damit sowohl eine biologische als auch eine kulturell-historischen Basis.

In Kindheit und Jugend entwickeln sich beide Komponenten sehr schnell. Diese Entwicklung zeigt sich in einem großen Zugewinn der möglichen Leistung wie in Abbildung 4 verbildlicht wird. Ausschlaggebend scheint dabei die Verbesserung der Verarbeitungsgeschwindigkeit zu sein. So konnten Coyle, Pillow, Snyder und Kochunov (2011) zeigen, dass die adoleszenztypische große Verbesserung des allgemeinen Intelligenzfaktors g nach Spearman fast vollständig durch schnellere Verarbeitungsgeschwindigkeit erklärt werden konnte. Der Zusammenhang von kristalliner und fluider Intelli-

genz bis ungefähr zum 25. Lebensjahr wurde als synergetisch bezeichnet: Die enorme Entwicklungsgeschwindigkeit der mechanics durch biologische Reifung führt in Kombination mit einer Förderung der pragmatics durch kulturellen Input zu einer sehr schnellen Entwicklung der allgemeinen Leistungsfähigkeit. Im weiteren Entwicklungsverlauf kann man von einer kompensatorischen Beziehung der Komponenten sprechen, da die mechanics langsam abnehmen, während die pragmatics stabil bleiben und ausgleichend wirken.

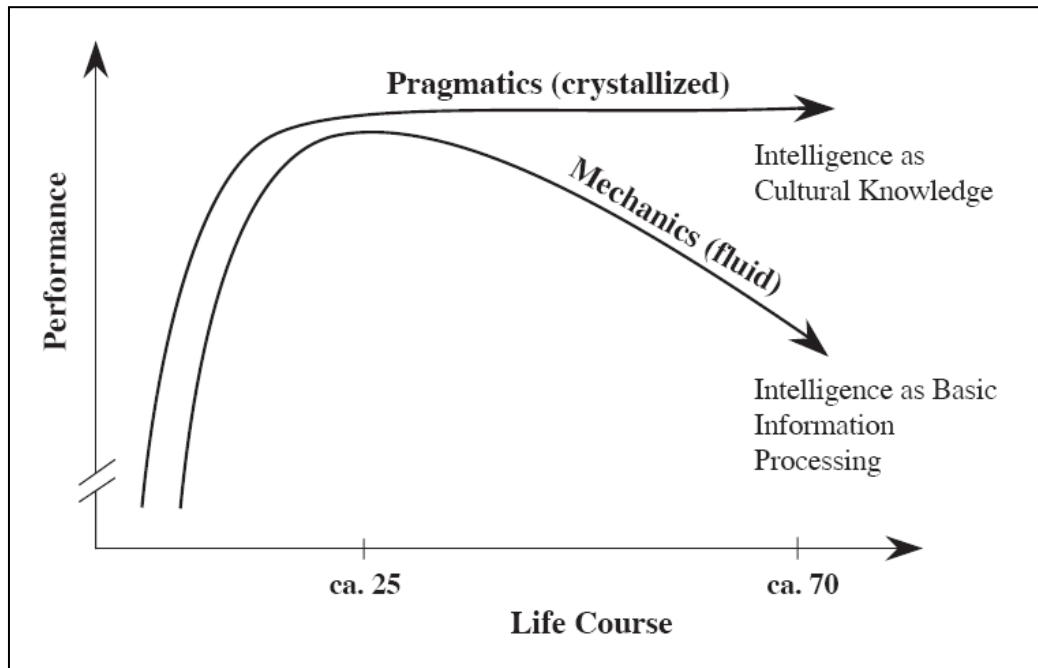


Abbildung 4: Mechanics und Pragmatics in der kognitiven Entwicklung (aus Baltes, Staudinger & Lindenberger, 1999, S. 487)

Auf Grundlage ausgewerteter Normtabellen von Intelligenztests und eigener empirischer Studien kommt Rindermann (2011) zu folgenden Ergebnissen: Der Intelligenzquotient steigt bei Kindern und Jugendlichen im Jahr durchschnittlich um 5.62 IQ-Punkte an. Dabei ist der Anstieg der kristallinen Intelligenz mit 6.12 Punkten höher als der Anstieg der fluiden Intelligenz mit im Schnitt 3.58 Punkten. Der Zuwachs bei jüngeren Kindern ist größer als bei älteren, ist also negativ beschleunigt. Zudem zeigte sich, dass der Zugewinn bei höherem Intelligenzniveau größer ist. Es liegt damit ein Schereffekt vor, bei dem Kinder und Jugendliche mit höherem Intelligenzquotienten größere Zuwächse aufweisen als solche mit niedrigerem. Ebenso stellt Rindermann fest, dass sich die kristalline Intelligenz ab Beginn der Schulzeit stärker verbessert als die fluide Intelligenz. Daher fordert Rindermann eine Korrektur von Cattells Modell dahingehend, dass ab Beginn der Schulzeit kein gleichförmiger Anstieg beider Intelligenzkomponenten anzunehmen ist, sondern ein stärkerer Zuwachs der kristallinen Intelligenz.

2.4.4 Informationsverarbeitung

Piagets Stufenmodell hat die Arbeit vieler nachfolgender Entwicklungspsychologen geprägt. Im Zuge der Verbreitung der Computermetapher als Analogie zum menschlichen Denken lehnten sich viele Theoretiker an den Ansatz der Informationsverarbeitung an. Die sogenannten Neo-Piagetianer behielten teilweise die Stufenidee bei, interpretierten sie aber unter der Perspektive des Informationsverarbeitungsansatzes neu. Beispielsweise hielt Case (1992) an vier Stufen der Entwicklung fest, nahm aber andere Übergangsmechanismen an, so die Automatisierung durch gesteigerte Verarbeitungskapazität, biologische Reifung und die Veränderung zentraler begrifflicher Strukturen, die in Form semantischer Netzwerke als bereichsspezifische Wissenskerne repräsentiert sind.

Die Stufen sind dabei nicht ausschließlich zu sehen, da Strategien unterschiedlicher Stufen weiterhin nebeneinander genutzt werden. Siegler (1996, 2007) konnte zeigen, dass Kinder sogar bei dem gleichen Aufgabentyp in kurz aufeinanderfolgenden Versuchsdurchläufen Strategien unterschiedlicher Stufe nutzen. Er stellte eine sehr langsame Generalisierung der neuen Strategien fest, die nur langsam ältere Strategien ablösen, was aus seinen kurzen Längsschnittstudien für intraindividuelle Entwicklungsverläufe deutlich wird. Beobachtet man nur ein Kind in kurz aufeinander folgenden Intervallen, wird dies viel stärker deutlich, als bei einem Querschnittsvergleich oder auch bei einem Längsschnitt mit den üblicherweise großen Pausen zwischen einzelnen Beobachtungen. Dadurch wird laut Siegler der Übergang von alten zu neuen Strategien überschätzt, so dass er statt des Treppen-Modells der kognitiven Entwicklung zu einem Modell „überlappender Wellen“ kommt. Dabei verfügen Kinder zu jedem Zeitpunkt über mehrere Strategien, die sie nutzen können. Mit steigendem Alter werden langsam eher effektivere Strategien angewendet, aber auch noch die älteren. Laut Siegler ziehen die Kinder Vorteile aus der Variabilität in ihren Strategien, da sie so auf Besonderheiten des jeweiligen Problems, dem Kontext und ihren eigenen Kompetenzen reagieren können. Ältere Strategien können mit größerer Wahrscheinlichkeit sicher und richtig umgesetzt werden und haben damit ein geringeres Risiko zu misslingen (Siegler, DeLoache & Eisenberg, 2008). Variabilität wird in dieser Theorie als Norm verstanden und es steht die Frage im Vordergrund, wie und nach welchen Kriterien welche Strategie eingesetzt wird, also die Selektion zwischen alternativen Problemlösewegen.

Viele Entwicklungstheorien, die sich am Informationsverarbeitungsansatz orientieren, haben ihren Fokus stärker auf die quantitativen Veränderungen und Zuwächse gelegt, die im intraindividuellen Entwicklungsverlauf deutlich werden. Einzelne Komponenten der Kognition werden getrennt betrachtet. Es zeigt sich, dass sich kognitive Mechanismen von der einfachen Enkodierung von Informationen bis zur Anwendung komplexerer Strategien bis zum Einsetzen der Pubertät deutlich weiterentwickeln. Insbesondere ist eine starke Beschleunigung der Verarbeitungsgeschwindigkeit in der Kindheit auffällig, die zum Einsetzen der Adoleszenz im Wesentlichen abgeschlossen ist. Dafür werden vor allem die biologische Reifung durch zunehmende Myelinisierung der Neuronen

und die steigende Erfahrung verantwortlich gemacht (Siegler, DeLoache & Eisenberg, 2008).

Zusammenfassend zeigen sich folgende Fortschritte des adoleszenten Denkens aus Sicht der Informationsverarbeitung (Berk, 2005, S. 505):

- *Zielgerichteter und flexibler Einsatz von Aufmerksamkeit*
- *Effektivere Strategienutzung*
- *Ausweitung von Wissensbeständen*
- *Weiterentwicklung der Fähigkeit zur Metakognition*
- *Verbesserung der kognitiven Selbstregulation*
- *Steigende Kapazität zur Informationsverarbeitung*

2.4.5 Gedächtnis und Wissen

Im Vergleich mit Kindern sind Jugendliche und Erwachsene praktisch in allen Gedächtnisleistungen überlegen. Die Verbesserung des Gedächtnisses ist auf mehrere Entwicklungsschritte zurückzuführen (Gleisser, 2007, S. 183):

- *Zunahme von Gedächtnisstrategien*
- *Vergrößerung der Wissensbasis*
- *Metagedächtnis*
- *Größeres kognitives Verarbeitungstempo*

Anteile des nicht-deklaratives Gedächtnisses scheinen sich am schnellsten und über die Lebensspanne hinweg konstant zu entwickeln. Die Entwicklung des deklaratives Gedächtnisses vollzieht sich über die Lebensspanne hinweg in zunehmender Komplexität und Umfang (Gleissner, 2007).

Die unmittelbare Gedächtnisspanne verbessert sich bis ins junge Erwachsenenalter kontinuierlich. Bis in die später Adoleszenz entwickelt sich insbesondere auch die Kontroll- und Regulationsfunktion der zentralen Exekutive im Arbeitsgedächtnismodell nach Baddeley und Hitch (Baddeley, 2012). In einer empirischen Studie mit insgesamt 133 Teilnehmern im Alter von 9 bis 20 Jahren fanden Luciana et al. (2005) für das nonverbale Arbeitsgedächtnis folgende Ergebnisse: Die Fähigkeiten zur Lösung einfacher, erinnerungsbasierte Aufgaben waren bereits im Alter von 11 bis 12 Jahren vorhanden, während sich die Möglichkeiten multiple Informationen aufrechtzuerhalten und zu manipulieren erst zwischen 13 und 15 Jahren entwickelten und erst mit 16 bis 17 Jahren

entwickelte sich eine strategische Selbstorganisation zur Bewältigung komplexerer Anforderungen.

Es konnte empirisch gezeigt werden, dass die Reihenfolge der von Piaget beschriebenen Entwicklungsstufen nicht immer eingehalten wird (Siegler, 2007). Piaget untermauerte sein Modell mit Beobachtungen von Kindern und den Fehlern die Kinder verschiedenen Alters typischerweise beim Lösen von Aufgaben machten. Piagets experimentelle Settings sind vielfach repliziert worden, auch unter Bedingungen, die eine Abweichung vom Modell deutlich machten. Insbesondere die Striktheit des Stufenmodells wurde kritisiert, indem gezeigt wurde, dass Kinder niedriger Entwicklungsstufen teilweise auch Problemlösestrategien höherer Stufen nutzen. Die Problemlösung hängt dabei eng mit der jeweiligen Domäne zusammen, in welcher sich das Problem erstreckt, und nicht nur mit grundlegenden, allgemeinen kognitiven Funktionen. Es gibt eine große Variabilität in der Problemlösung von Kindern, die sogar auch bei der gleichen Aufgabe vorliegen kann, wenn sie zum zweiten Mal vorgelegt wird; die intrapsychische Variabilität ist also sehr hoch (Siegler, 2007).

Neuere Informationsverarbeitungsansätze beziehen daher die Entwicklung bereichsspezifischen Wissens unter dem Begriff der „Theorie-Theorie“ (Sodian, 1998) bzw. der „Theorie des Kernwissens“ (R. Siegler, et al., 2008) mit ein. Die Entwicklung vollzieht sich demnach nicht für alle Domänen und Wissensbereiche gleichsam. Die Entwicklung von Expertise vollzieht sich nur in Bereichen, die von besonderer Relevanz sind. Das erklärt auch, warum nicht alle Menschen alle Aufgaben auf formal-operationalem Niveau sensu Piaget erreichen. Dieses Niveau wird nicht universell erreicht, sondern nur in Bereichen, in denen der Mensch eine gewisse Expertise besitzt (Berk, 2005). Darüber hinaus gibt es konnektionistische Theorien, die von einer parallelen Verarbeitungsart ausgehen, z. B. in Form neuronaler Netzwerke.

2.4.6 Exekutivfunktionen

Eng verbunden mit den Gedächtnisfunktionen sind die exekutiven Funktionen, zu denen die Planung, Kontrolle und Steuerung des eigenen Verhaltens gezählt wird. Mit neuropsychologischen Methoden konnte gezeigt werden, dass bei Aktivitäten, die mit exekutiven Funktionen in Verbindung gebracht werden, insbesondere frontale Areale des Gehirns aktiviert sind. Früher wurden daher die Begriffe exekutive Funktionen und Frontalhirnfunktionen synonym verwandt. Es hat sich aber gezeigt, dass ein ausgedehntes Netzwerk kortikaler und subkortikaler Komponenten für die exekutiven Prozesse zuständig ist (Konrad, 2007).

In der Leistung in unterschiedlichen kognitiven Aufgaben und Paradigmen zeigen sich frühe Verbesserungen ab dem Kindesalter, die bis in die späte Adoleszenz anhalten. Dabei haben insbesondere die exekutiven Funktionen einen langen Entwicklungsprozess. Es wird vermutet, dass Kinder bei der Bearbeitung von kognitiven Aufgaben diffusere Netzwerke aktivieren als Erwachsene. Die kortikale Aktivität scheint sich mit

voranschreitender Entwicklung in mit der Leistung korrelierten Hirnarealen stärker zu fokalisieren und in irrelevanten Bereichen abzunehmen. Diese veränderten Hirnaktivierungsmuster hängen vermutlich mit Veränderungen der funktionalen und strukturellen Konnektivität der Bereiche zusammen (Konrad, 2007).

In einer Entwicklungsperspektive fassen Best und Miller (2010) altersgemäße Veränderungen für drei zentrale exekutive Funktionen zusammen: Bei der Inhibition zeigten sich schon im vorschulischen Kindesalter große Verbesserungen und nur noch geringere Zugewinne während der Adoleszenz. In der Leistung bei komplexen Arbeitsgedächtnisaufgaben, die ein höheres Maß an Aufrechterhaltung und Manipulation von Informationen erfordern, zeigten sich dagegen kontinuierliche Verbesserungen vom Vorschulalter bis in die Adoleszenz. Ebenso bei der Fähigkeit des „shifting“, also dem Wechsel von einer gültigen Regel zu der nachfolgend aktuellen oder von einem abgeschlossenen relevanten Aufgabenaspekt zum nächst folgenden. Vorschulkinder konnten Wechsel bei einfachen Aufgaben bewältigen und später auch unerwartete Änderungen bei komplexeren Aufgaben. Aber erst in der Adoleszenz wurde eine Leistung auf Erwachsenenniveau beim komplexen Wechseln erreicht, vermutlich mitbedingt durch zunehmendes Monitoring eigener Fehler.

In einer Studie mit 890 Teilnehmern zwischen 10 und 30 Jahren untersuchten Albert und Steinberg (2011a) die Leistungen in der Aufgabe „Tower of London“ zum strategischen Planen. Es zeigten sich Leistungszuwächse im Jugendalter, bis im Alter von 17 Jahren ein stabiles Leistungsniveau erreicht wurde. Aber auch darüber hinaus zeigten sich für besonders schwierige Probleme weitere Verbesserungen bis in die frühen zwanziger Jahre. Während die Leistungszuwächse in der Adoleszenz teilweise über Reifungsprozesse sowohl des Arbeitsgedächtnisses als auch der Impulskontrolle erklärt werden können, wurden die späteren Zugewinne nur noch durch die bessere Impulskontrolle allein erklärt. Die Autoren kommen zu dem Schluss, dass die späte Adoleszenz eine Zeit der weiteren Verbesserung des geplanten Ziel gerichteten Verhaltens ist.

Kognitive Verbesserungen bis in die später Adoleszenz zeigen sich auch in Studien zur Beurteilung des Entscheidungsverhaltens von Jugendlichen. Auch wenn die kognitiven Möglichkeiten schon in der frühen Adoleszenz gegeben sind, so zeigt sich doch, dass Jugendliche Entscheidungen anders treffen als Erwachsene. Verschiedene entwicklungsbedingte Einflussfaktoren auf das Entscheidungsverhalten von Jugendlichen sind u. a. emotionale Faktoren wie das Bedürfnis nach Belohnung und Aufregung („sensation-seeking“) und soziale Faktoren, zum Beispiel Akzeptanz und Zugehörigkeit durch riskantes Verhalten in der Gruppe (siehe Abbildung 5). Wenn Jugendliche beispielsweise in experimentellen Setting hypothetischer Dilemmas lösen sollen, so zeigen sie ein anderes Entscheidungsverhalten als in realen Situationen, insbesondere wenn Gleichaltrige involviert sind. Zum Beispiel konsumieren Jugendliche Drogen obwohl ihnen die Risiken bewusst sind (Albert & Steinberg, 2011b).

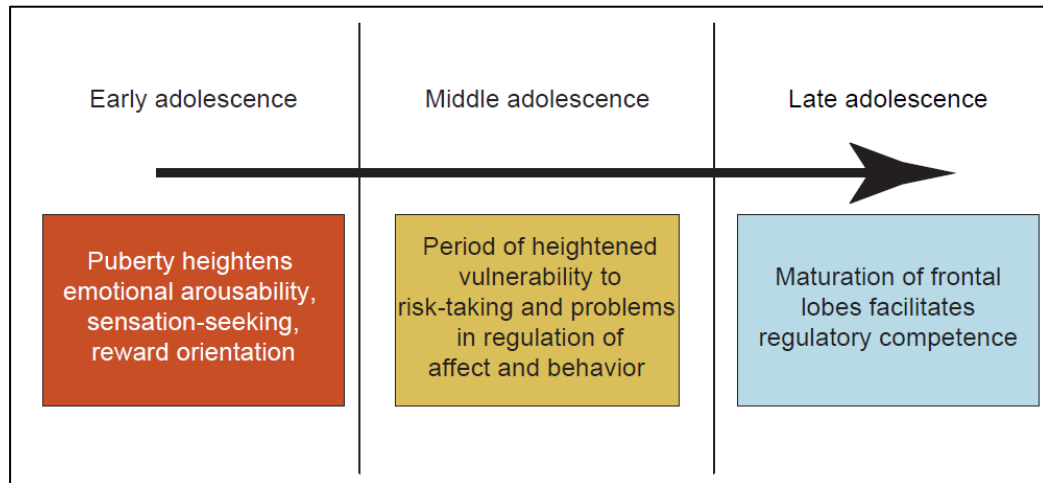


Abbildung 5: Körperliche pubertäre, soziale und emotionale Einflüsse auf das Entscheidungsverhalten im Lauf der Adoleszenz (aus Steinberg, 2005b, S. 70)

Im Kontext der Frage nach der Schuldfähigkeit von Jugendlichen betonen Cauffman und Steinberg (2000) wie wichtig es ist, neben der im Wesentlichen abgeschlossenen kognitiven Reifung die langsamer verlaufende Entwicklung psychosozialer Komponenten zu berücksichtigen. In einer Studie mit 1000 Jugendlichen und Erwachsenen zwischen 12 und 48 fanden sie eine voranschreitende psychosoziale Reifung (Verantwortlichkeit, flexible Perspektivübernahme und Mäßigung, z. B. Impulskontrolle und Selbstbeherrschung) und weniger antisoziale Entscheidungsfindung mit zunehmendem Alter. Generell hatten die Erwachsenen eine höhere psychosoziale Reife und trafen weniger antisoziale Entscheidungen. Allerdings scheint es nach dem Alter von 19 Jahren keinen weiteren Anstieg mehr zu geben. Die Altersspanne zwischen 16 und 19 scheint eine kritische Übergangsphase der psychosozialen Entwicklung darzustellen, die sich damit später vollzieht als die kognitive Reifung. Die Entwicklung von Verantwortungsübernahme, flexiblere Perspektivübernahme und Selbstkontrolle scheint also nicht lediglich von den kognitiven Voraussetzungen dieser Kompetenzen abzuhängen. Allerdings fanden die Autoren auch große Schwankungen, so hatten einige Jugendliche schon sehr früh hohe Ergebnisse, teilweise höher als einige der Erwachsenen. Die psychosoziale Reife wird somit erst in der späten Adoleszenz erreicht und diese Entwicklung hat eine gravierende Auswirkung auf die Fähigkeit reife Entscheidungen zu treffen.

2.4.7 Fazit

Es wird deutlich, dass sich im Jugendalter ein großer Zuwachs an intellektuellen Fertigkeiten einstellt, wodurch viele neue Möglichkeiten eröffnet werden. Diese bilden die Grundlage dafür, Probleme kompetent lösen zu können und mit den Anforderungen des Erwachsenenlebens zurechtzukommen.

Es zeigt sich, dass sich vor allem im Kindesalter, aber auch in der Adoleszenz wesentliche Fortschritte in den kognitiven Fähigkeiten vollziehen. Quantitativ zeigt sich dies in einer Verbesserung der Gedächtnisfunktionen durch bessere Enkodierung, Speicherung

und Abruf von Inhalten und vor allem auch im Arbeitsgedächtnis. Die „mechanics“ bzw. die „fluide Intelligenz“ verbessern sich deutlich. Damit einher gehend verbessert sich die Effektivität der Informationsverarbeitung, indem mehr Informationen mit höherer Verarbeitungsgeschwindigkeit verarbeitet werden. Diese Verbesserungen scheinen sich zum großen Teil aus biologischen Veränderungen, z. B. der steigenden Myelinisierung der Neuronen zu entstammen, aber wesentlich auch durch den Zuwachs an Wissen („Theorie-Theorie“ bzw. „pragmatics“), das untereinander stärker Domäne bezogen verknüpft wird (semantische Netzwerke) und einer größeren Adaptivität in der Strategieanwendung. Im Längsschnitt erscheint so ein qualitativ anderes Bild der kognitiven Mechanismen, das insgesamt gut über den Begriff der formalen Operationen nach Piaget beschrieben werden kann. Piaget eignet sich demnach sehr gut, das charakteristische Denken eines Jugendlichen zu beschreiben, nicht aber dafür, den Übergang vom kindlichen ins adoleszente Denken zu charakterisieren. Die Veränderungsprozesse sind besser durch Theorien der Informationsverarbeitung zu erklären. Die Adoleszenz kann als eine Hochphase der kognitiven Entwicklung angesehen werden. Allerdings ist die Entwicklung reifer Entscheidungsverhaltensweisen noch nicht abgeschlossen und auch der Zugewinn an Expertise und Wissen erstreckt sich über die gesamte Lebensspanne.

3 Aufmerksamkeitsstörungen

In diesem Kapitel wird die Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS) vorgestellt. Nach der einführenden Klassifikation werden geschlechtsbedingte Unterschiede aufgezeigt, da diese eine Grundlage für die spätere Analyse der differenziellen Wirksamkeit bilden. Es folgen Angaben zu Prävalenz und Komorbidität. Letztere wird ebenfalls im Bezug zu differenzieren Wirksamkeit genauer untersucht. Es folgt eine Zusammenstellung aktueller Erkenntnisse zum Verlauf der Störung. Im Anschluss werden Erklärungen und Modelle zur Entstehung, Auslösung und Aufrechterhaltung der Störung vorgestellt. Den Abschluss des Kapitels bildet ein Einblick in die Behandlung der ADHS. Dabei bildet das Lerntraining für Jugendliche mit ADHS (LeJA), das im Weiteren evaluiert wird, einen Schwerpunkt.

3.1 Klassifikation

Zur Stellung der Diagnose einer Aufmerksamkeitsstörung ist es erforderlich, dass starke Beeinträchtigungen im Bereich Unaufmerksamkeit und bzw. oder Hyperaktivität und Impulsivität vorliegen. Kinder mit ADHS haben große Schwierigkeiten, ihre Aufmerksamkeit selektiv zu steuern und dauerhaft aufrecht zu erhalten. Viele leiden zusätzlich unter einer starken motorischen Unruhe und Rastlosigkeit. Weiterhin ist häufig ein unbedachtes, vorschnelles Verhalten vorhanden (Gawrilow, 2012). Die hyperkinetischen bzw. Aufmerksamkeitsstörungen werden anhand von Verhaltensmerkmalen sowohl im DSM-IV-TR (Saß, Wittchen, Zaudig & Houben, 2003) als auch in der ICD-10 (Dilling, Mombour & Schmidt, 2004) sehr ähnlich definiert. Neben der Erfüllung einer bestimmten Anzahl von Symptomen ist es zur Diagnosestellung wichtig, dass zumindest ein Teil der Symptome bereits vor dem siebten Lebensjahr begonnen hat, dass sich die Beeinträchtigungen in mindestens zwei wichtigen Lebensbereichen (z. B. Zuhause und in der Schule) zeigen, dass eine gravierende Entwicklungsgefährdung bzw. Leidensdruck besteht und dass ausgeschlossen werden kann, dass eine andere Störung die Symptome besser erklären kann. Tabelle 2 zeigt die Definition gemäß DSM-IV-TR.

Tabelle 2: Symptome der ADHS gemäß DSM-IV-TR (Saß et al., 2003)

1. Unaufmerksamkeit

- (a) beachtet häufig Einzelheiten nicht oder macht Flüchtigkeitsfehler bei den Schularbeiten, bei der Arbeit oder bei anderen Tätigkeiten
- (b) hat oft Schwierigkeiten, längere Zeit die Aufmerksamkeit bei Aufgaben oder beim Spielen aufrechtzuerhalten
- (c) scheint häufig nicht zuzuhören, wenn andere ihn/sie ansprechen
- (d) führt häufig Anweisungen anderer nicht vollständig durch und kann Schularbeiten, andere Arbeiten oder Pflichten am Arbeitsplatz nicht zu Ende bringen (nicht aufgrund oppositionellen Verhaltens oder Verständigungsschwierigkeiten)

- (e) hat häufig Schwierigkeiten, Aufgaben und Aktivitäten zu organisieren
- (f) vermeidet häufig oder hat eine Abneigung gegen oder beschäftigt sich häufig nur widerwillig mit Aufgaben, die längerandauernde geistige Anstrengungen erfordern (wie Mitarbeit im Unterricht oder Hausaufgaben)
- (g) verliert häufig Gegenstände, die für Aufgaben oder Aktivitäten benötigt werden (z. B. Spielsachen, Hausaufgabenhefte, Stifte, Bücher oder Werkzeug)
- (h) lässt sich oft durch äußere Reize leicht ablenken
- (i) ist bei Alltagstätigkeiten häufig vergesslich

2. Hyperaktivität

- (a) zappelt häufig mit Händen oder Füßen oder rutscht auf dem Stuhl herum
- (b) steht in der Klasse oder in Situationen, in denen Sitzen bleiben erwartet wird, häufig auf
- (c) läuft herum oder klettert exzessiv in Situationen, in denen dies unpassend ist (bei Jugendlichen oder Erwachsenen kann dies auf ein subjektives Unruhegefühl beschränkt bleiben)
- (d) hat häufig Schwierigkeiten, ruhig zu spielen oder sich mit Freizeitaktivitäten ruhig zu beschäftigen
- (e) ist häufig „auf Achse“ oder handelt oftmals, als wäre er/sie „getrieben“
- (f) redet häufig übermäßig viel

3. Impulsivität

- (g) platzt häufig mit Antworten heraus, bevor die Frage zu Ende gestellt ist
- (h) kann nur schwer warten, bis er/sie an der Reihe ist
- (i) unterbricht und stört andere häufig (platzt z. B. in Gespräche oder Spiele anderer hinein)

Im DSM-IV-TR werden Aufmerksamkeitsstörungen als Aufmerksamkeitsdefizit/Hyperaktivitätsstörung (ADHS) bezeichnet. Das DSM-IV-TR kennt drei Subtypen der ADHS: 1. den vorwiegend unaufmerksamen Subtypus, bei dem mindestens sechs Symptome der Unaufmerksamkeit vorliegen, aber weniger als sechs Symptome aus dem Bereich Hyperaktivität und Impulsivität, 2. den vorwiegend hyperaktiv-impulsiven Subtypus, bei dem mindestens sechs Symptome der Hyperaktivität und Impulsivität vorliegen, aber weniger als sechs Symptome der Unaufmerksamkeit und 3. den Mischtypus, bei dem mindestens sechs Symptome der Unaufmerksamkeit erfüllt sind und auch der Hyperaktivität und Impulsivität (siehe Tabelle 3).

Tabelle 3: Subtypen der ADHS gemäß DSM-IV-TR

Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung	Anzahl Symptome Unaufmerksamkeit	Anzahl Symptome Hyperaktivität und Impulsivität
Mischtypus	≥ 6	≥ 6
Vorwiegend Unaufmerksam Typus	≥ 6	< 6
Vorwiegend Hyperaktiv-Impulsiver Typus	< 6	≥ 6

Im ICD-10 finden sich die Aufmerksamkeitsstörungen unter der Bezeichnung Hyperkinetische Störungen (F9). Als weiteres Unterscheidungsmerkmal und zur Definition von Subtypen wird im ICD-10 das gleichzeitige Vorliegen einer Störung des Sozialverhaltens herangezogen. Bei Vorliegen von sechs oder mehr Symptomen der Unaufmerksamkeit und zusammengenommen mindestens sechs Symptomen der Hyperaktivität und Impulsivität kann eine Einfache Aufmerksamkeit- und Hyperaktivitätsstörung (F90.0) diagnostiziert werden. Liegt zusätzlich eine Störung des Sozialverhaltens vor, so wird eine Hyperkinetische Störung des Sozialverhaltens (F90.1) festgestellt. Es gibt auch die Diagnose einer Aufmerksamkeitsstörung ohne Hyperaktivität, diese wird aber abweichend vom DSM-IV-TR nicht unter den Hyperkinetischen Störungen aufgeführt, sondern als Sonstige näher bezeichnete Verhaltens- und emotionale Störungen mit Beginn in der Kindheit und Jugend (F98.8) eingeordnet. Abbildung 6 zeigt einen Überblick der Diagnosen gemäß beider Diagnostiksysteme aus Döpfner, Banaschewski und Sonuga-Barke (2008).

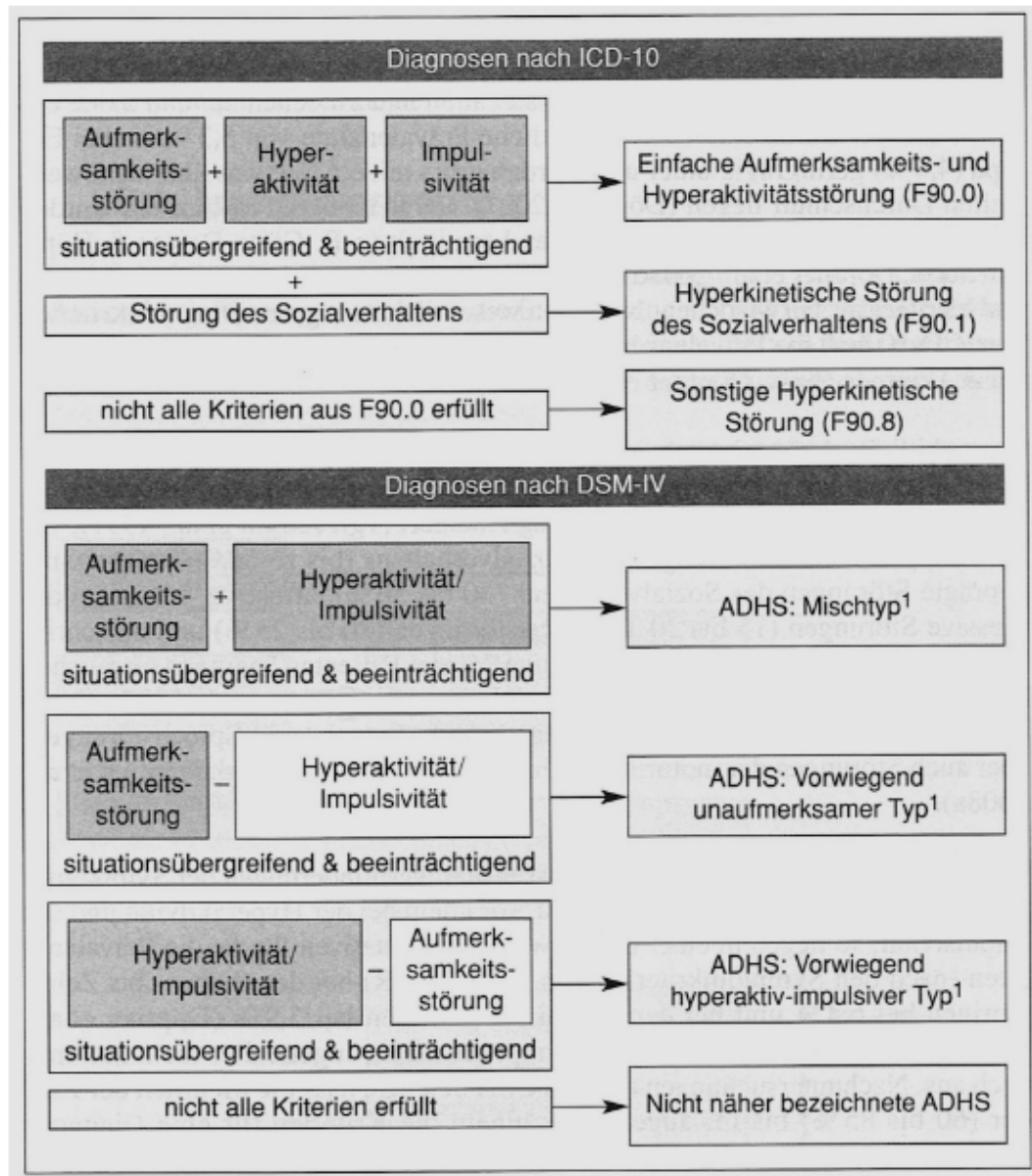


Abbildung 6: Diagnosen der hyperkinetischen bzw. Aufmerksamkeitsstörungen nach ICD-10 und DSM-IV (aus Döpfner, Banaschewski & Sonuga-Barke, 2008, S. 259)

Die vorgestellten Verhaltenskriterien wurden für das Kindesalter spezifiziert und sind für Jugendliche und Erwachsene teilweise nicht angemessen. Eine Diagnostik anhand dieser Kriterien ist daher insbesondere im Erwachsenenalter schwierig, auch deshalb weil die retrospektive Schätzung, ob die Symptome im Kindesalter durchgehend erfüllt waren, nicht unproblematisch möglich ist (McGough & Barkley, 2004). Neben den Definitionen der gängigen Diagnostiksysteme wurden von Wender (2002) die sog. Wender-Utah-Kriterien der adulten ADHS vorgeschlagen, die sich in sieben Symptomcluster aufteilen:

- (1) Aufmerksamkeitsstörung
- (2) Motorische Hyperaktivität

- (3) Affektlabilität
- (4) Desorganisiertes Verhalten
- (5) Affektkontrolle
- (6) Impulsivität
- (7) Emotionale Überreagibilität

Damit die Diagnose erfüllt gemäß Wender-Utah-Kriterien als erfüllt gilt, muss sowohl eine motorische Hyperaktivität vorliegen und darüber hinaus mindestens zwei der weiteren Symptomcluster. McGough und Barkley (2004) weisen darauf hin, dass so Erwachsene mit einer vorrangig unaufmerksamen Symptomatik nicht erfasst werden. Zudem ist die Zusammenstellung der Cluster nicht ausreichend empirisch evaluiert. Veränderungen der Klassifikation der ADHS im Erwachsenenalter werden mit dem DSM-V eingeführt (Coghill & Seth, 2011).

Dass die im DSM-IV-TR vorgeschlagene Aufteilung der Symptomatik nach Unaufmerksamkeit und den zusammengekommenen Bereichen Hyperaktivität und Impulsivität sinnvoll ist, bestätigt sich auch in empirischen Studien. Toplak und Kollegen (2012) untersuchten eine internationale klinische Stichprobe mit 1.373 Kindern und Jugendlichen mit ADHS und 1.772 Geschwistern dieser Kinder. Es wurden die Einschätzungen von Eltern und Lehrern genutzt. Ein hierarchisches Modell mit einem allgemeinen ADHS-Faktor und zwei darunter liegenden spezifischen Faktoren (Unaufmerksamkeit vs. Hyperaktivität/Impulsivität) konnte die in den Daten gefundene Struktur am besten abbilden. Dies galt unabhängig von der Quelle der Einschätzung und für die gesamte Altersspanne.

Im Folgenden werden die Bezeichnungen des DSM-IV-TR verwendet, da sie zur empirischen Evidenz passen und konzeptuell sinnvoller und besser gegliedert erscheinen. Sie haben sich auch im internationalen Praxis- und Forschungskontext stärker durchgesetzt. Insbesondere wird im Folgenden die Abkürzung ADHS verwendet.

Es gibt aber auch begründete Vorschläge, den vorwiegend unaufmerksamen ADHS-Typ und den Mischtypus als zwei unterschiedliche Störungen zu konzeptualisieren anstatt als zwei Ausprägungen derselben Störung (Milich, Balentine & Lynam, 2001). Als ein zentrales Problem des vorrangig unaufmerksamen Störungstypes werden Beeinträchtigungen der Arbeitsgedächtnisleistung gesehen, die auch zur Diagnostik genutzt werden können. Kinder mit einem „wahren“ unaufmerksamen ADHS-Typus zeichneten sich v. a. dadurch aus, dass sie schnell gelangweilt seien. Ihre Defizite lägen weniger als bei Kindern des Mischtypus in der inhibitorischen Kontrolle, sondern vielmehr in der Motivation, die insbesondere durch eine Untererregung („underarousal“) entstehen (Diamond, 2005).

Im Zusammenhang mit dem vorwiegend unaufmerksamen Subtypus ist auch das Konzept des „sluggish cognitive tempo“ (SCT) relevant. Es beschreibt Symptome wie Tagträumereien und einen schwerfälligen, schleppenden Arbeitsstil (McBurnett, 2008). McBurnett, Pfiffner und Frick (2001) stellten anhand einer Stichprobe mit 692 Kindern einer auf ADHS spezialisierten Klinik fest, dass sich diese Symptomgruppe von der Unaufmerksamkeit abgrenzen ließ. In Faktorenanalysen mit den Daten der Kinder ohne Hyperaktivität und Impulsivität ließen sich zwei Faktoren identifizieren, einer für das SCT und einer für die Unaufmerksamkeit. Die Autoren schlussfolgern, dass beide Symptomgruppen für den unaufmerksamen Störungssubtypus typisch sind und dass das SCT den unaufmerksamen Typus in zwei weitere Subtypen unterteilt. Dieses Ergebnis konnte durch eine aktuellere Studie mit 140 puertoricanischen Kindern mit ADHS bestätigt werden. Die Autoren fanden in Analysen der Einschätzungen von Müttern und Lehrern drei Faktoren: Unaufmerksamkeit, Hyperaktivität/Impulsivität und SCT (Bauermeister, Barkley, Bauermeister, Martinez & McBurnett, 2012). Ebenso wurden die drei Faktoren anhand der Eltern- und Lehrereinschätzungen in einer Studie mit 322 Kindern und Jugendlichen gefunden, die wegen Lern-, Verhaltens- und emotionalen Störungen an einer Universitätsklinik untersucht wurden (Garner, Marceaux, Mrug, Patterson & Hodgins, 2010). Die Faktoren sind dabei abgrenzbar, aber korrelieren miteinander. Analoge Ergebnisse brachte eine Untersuchung der Eltern- und Lehrerurteile einer allgemeinen Stichprobe mit 296 8- bis 18jährigen Zwillingen, bei denen ADHS und Lernstörungen überrepräsentiert waren (Hartman, Willcutt, Rhee & Pennington, 2004).

Carlson und Mann (2002) stellten in einem Vergleich fest, dass sich Schulkinder mit Symptomen des SCT in der Einschätzung ihrer Lehrer von ihren nicht betroffenen Mitschülern zwar nicht in ihren Aufmerksamkeits- und Lernleistungen unterschieden, aber als unzufriedener eingeschätzt wurden und ihnen eher affektive und soziale Schwierigkeiten wie Rückzugsverhalten zugeschrieben wurden. Dagegen zeigten sie seltener externalisierendes Verhalten.

Für das Erwachsenenalter untersuchte Barkley (2012) eine repräsentative Stichprobe mit 1.249 Erwachsenen im Alter von 18 bis 96 Jahren. Ebenso wie im Kindesalter konnte eine abgrenzbare Dimension SCT gefunden werden, die nicht mit Alter, Geschlecht oder Ethnizität in Verbindung stand. Die Erwachsenen mit SCT berichteten mehr Schwierigkeiten sich selbst zu organisieren und Probleme zu lösen als die Erwachsenen, die nur ADHS hatten, und als die Kontrollgruppe ohne ADHS. Sie wiesen große Defizite in ihren Exekutivfunktionen auf. Barkley schlussfolgert, dass SCT eine von ADHS unabhängige Störung konzeptualisiert werden könne, die bei schätzungsweise der Hälfte der Betroffenen mit einer Komorbidität einhergehe.

3.2 Geschlechtsspezifik

ADHS wird häufiger bei Jungen als bei Mädchen diagnostiziert. Das Verhältnis der Prävalenzraten reicht von 2:1 bis 6:1 (Gawrilow, 2012). Im für Deutschland repräsentativen Kinder- und Jugendgesundheitssurvey (KiGGS) mit 7569 Jungen und 7267 Mädchen im Alter von 3 bis 17 Jahren lag die Diagnosehäufigkeit von ADHS für Jungen bei 7.7 % und für Mädchen bei 1.8 % (Huss, Hölling, Kurth & Schlack, 2008). Ähnliche Ergebnisse wurden in einer repräsentativen Studie aus den USA mit 1420 Kindern gefunden (Costello, Mustillo, Erkanli, Keeler & Angold, 2003). Für die Jungen wurde eine Prävalenzrate von 7.0 % gefunden und für die Mädchen eine Prävalenzrate von 1.1 %.

In einer weltweiten Metaanalyse über die Jahre 1978 bis 2005 mit insgesamt 171.756 Kindern und Jugendlichen wurde für die Jungen eine gepoolte Prävalenz von ca. 10 % gefunden, für die Mädchen nur ca. 4 % (Polanczyk, de Lima, Horta, Biederman & Rohde, 2007). Größere Diskrepanzen werden dabei in klinischen Stichproben gefunden als in repräsentativen Studien (Gaub & Carlson, 1997). Mädchen sind häufiger dem vorrangig unaufmerksamen ADHS-Subtypus zuzuordnen, sowohl in klinischen als auch in nichtklinischen Stichproben (Mikami & Hinshaw, 2008).

Als Ursachen für die Gründe der höheren Diagnosehäufigkeit bei Jungen wird u. a. diskutiert, dass bei Jungen eine stärkere Disposition zur Entwicklung expansiven Verhaltens als bei Mädchen vorhanden zu sein scheint (Saß, et al., 2003). Weiterhin daran, dass sich generell mehr Jungen in klinischen Stichproben befinden und Jungen eher klinisch vorgestellt werden, dass die Formulierung der diagnostischen Kriterien vorrangig an männlichen Stichproben orientiert ist und dass sich die Symptomatik bei Mädchen anders äußert, u. a. da sie stärker vom vorrangig unaufmerksamen Subtypus betroffen sind (Retz-Junginger, Sobanski, Alm, Retz & Rösler, 2008).

Bei der Bewertung von klinischen Fallvignetten konnten Bruchmüller, Margraf und Schneider (2012) zeigen, dass ADHS häufiger diagnostiziert wurde, wenn es sich um eine Falldarstellung eines Jungen handelte. Die Autoren schickten je eine Fallvignette zu 1.000 Kinderpsychologen, Psychiatern und Sozialarbeitern mit der Bitte eine Diagnose abzugeben. Bei den männlichen Versionen der Fälle wurde ADHS ca. doppelt so häufig diagnostiziert, auch wenn sich die Vignetten ansonsten nicht unterschieden. Sie schlussfolgern, dass das Geschlecht einen Einfluss in der klinischen Diagnostik hat. Ein ähnlicher Bias konnte auch für Grundschullehrer festgestellt werden. Scitutto, Nolfi und Bluhm (2004) baten 199 Lehrer ein fiktionales Profil über schulische Leistungen eines Kindes zu bewerten und anzugeben, ob sie das Kind für eine Überprüfung anmelden würden. Jungen wurden dabei häufiger für eine Überprüfung ausgewählt, insbesondere bei stark ausgeprägter Hyperaktivität. Derk, Hudziak und Boomsma (2007) fanden den gleichen Effekt in der Einschätzung von Lehrern. In einem nichtklinischen Setting erhoben sie Einschätzungen für 283 Jungen und 291 Mädchen mit und ohne ADHS, sowohl von den Müttern der Kinder als auch von den Lehrern. In der Einschätzung der Mütter unterschieden sich die Jungen und Mädchen nicht in der Ausprägung psychi-

scher und schulischer Probleme. Es wurden ähnliche Schweregrade für Aggression und Aufmerksamkeitsprobleme angegeben. Die Lehrer beurteilten die Problematik bei den Jungen hingegen als schwerwiegender. Es wird also berichtet, dass sich Mädchen und Jungen zuhause nicht unterschiedlich verhalten, in der Schule aber schon. Auch in einer Metaanalyse von Gershon (2002), in die 38 Artikel einbezogen wurden, beurteilten die Lehrer die Problematik der Jungen als deutlich schwerwiegender als die der Mädchen. Abweichende Ergebnisse wurden in einer aktuelleren Studie gefunden (Coles, Slavec, Bernstein & Baroni, 2012). Dabei sollten 50 Lehrer anhand von acht Fallvignetten von Jungen und Mädchen mit ADHS einschätzen, welchen Grad der Beeinträchtigung und Bedarf an Förderung sie jeweils sahen. Dabei wurden die Beeinträchtigung und Behandlungsbedürftigkeit der Mädchen als signifikant höher eingeschätzt. Vignetten, die einen vorrangig unaufmerksamen Subtypus beschrieben, wurden als am wenigsten beeinträchtigt bewertet. Als am stärksten beeinträchtigt wurden Mädchen mit hyperaktiv-impulsiven Verhalten eingestuft. Es mag sich ein zeitgemäßer Wandel vollziehen, bei dem sich die Einschätzung durchsetzt, dass ADHS keine Störung ist, die fast ausschließlich bei Jungen auftritt bzw. dass die Störung bei Mädchen eher als unproblematisch zu betrachten ist.

Zur Frage, ob es Unterschiede in der Ausprägung der Störung bei Jungen und Mädchen gibt, finden sich widersprüchliche Befunde. Eine Untersuchung, in der genug Mädchen vertreten waren, um generalisierbare Effekte festzustellen, ist die „Multimodal Treatment Study of Children With ADHD“ (MTA; Jensen, Arnold, Richters, Severe, Vereen, Vitiello, Schiller, Hinshaw, Elliott, Conners, Wells, March, Swanson, Wigal, Cantwell, Abikoff, Hechtman, Greenhill, Newcorn, Pelham, Hoza & Kraemer, 1999). An dieser multizentrischen Studie nahmen 395 Jungen und 103 Mädchen im Alter von 7 bis 10 Jahren mit einer ADHS des kombinierten Subtypus teil. Sie wurden randomisiert einer der vier experimentellen Gruppen zugeteilt. Die Gruppen bestanden aus 1. reiner Medikation, 2. reiner Verhaltenstherapie, 3. einer Kombinationsbehandlung aus Medikation und Verhaltenstherapie und 4. der standardmäßigen Behandlung („community care“). Im Vergleich der Werte vor Beginn der Interventionen zeigten die Mädchen eine signifikant geringere Störungsschwere als die Jungen (Newcorn et al., 2001).

Auch in der Metaanalyse von Gershon (2002) zeigten sich im Vergleich mit den Jungen für Mädchen geringere Werte in der Hyperaktivität ($d = .29$), Unaufmerksamkeit ($d = .23$), Impulsivität ($d = .22$) und weiteren externalisierenden Probleme ($d = .21$). Mädchen hatten weiterhin größere intellektuelle Schwierigkeiten ($d = -.27$) und mehr internalisierende Symptome ($d = -.12$). Keine Unterschiede zeigten sich für die schulischen Leistungen, neuropsychologische Maße und das Sozialverhalten. Auch in einer Studie von Greene et al. (2001) zeigte sich, dass die 267 untersuchten Kinder mit ADHS im Vergleich zu 234 Kindern ohne ADHS größere Probleme im Sozialverhalten aufwiesen. Diese Probleme waren allerdings im Vergleich der 127 Mädchen mit den 140 Jungen mit ADHS gleichsam ausgeprägt. Bezogen auf das schulische Verhalten fanden Abikoff et al. (2002) in einer Studie mit 403 Jungen und 99 Mädchen mit ADHS, dass die Mädchen mit ADHS häufiger verbal aggressives Verhalten zeigten als Mädchen ohne

ADHS. Jungen mit ADHS zeigten dagegen häufiger Regelverstöße und externalisierende Verhaltensweisen als Mädchen mit ADHS. Biederman, Faraone und Monuteaux (2002) verglichen 280 Kinder und Jugendliche mit ADHS mit 242 Probanden ohne ADHS. Sie konnten zeigen, dass es unabhängig vom Geschlecht einen Zusammenhang zwischen psychosozialen Risikofaktoren und dem Ausbilden einer ADHS gab. Aber die Jungen hatten die schlechteren Ergebnisse bezogen auf Lernstörungen und andere funktionale Maße. Die Autoren schlussfolgern, dass die Jungen eine höhere Vulnerabilität aufwiesen als die Mädchen.

Andere Studien fanden geringere bzw. gar keine Geschlechtsunterschiede. Novik et al. (2006) stellen zwar fest, dass Mädchen seltener als Jungen eine Behandlung ihrer ADHS erfuhren, dafür aber gleiche Muster in ihrer Beeinträchtigung aufwiesen. Die Autoren analysierten die Daten von 231 Mädchen und 1.222 Jungen mit ADHS aus 10 europäischen Staaten. Es zeigten sich keine signifikanten Unterschiede in der Schwere der ADHS-Kernsymptomatik und komorbiden Störungen. Lediglich das Ausmaß der emotionalen Probleme wurde von den Eltern für die Mädchen als gravierender eingeschätzt. Weiterhin zeigten die Mädchen mehr prosoziales Verhalten und waren häufiger Opfer und seltener Täter von Bullying. Auch in einer australischen Studie unterschieden sich 225 Jungen und 99 Mädchen mit ADHS nicht in Bezug auf ihre Kernsymptomatik, Komorbidität und Grad der Beeinträchtigung (Graetz, Sawyer & Baghurst, 2005). Die Mädchen gaben allerdings mehr körperlichen Beschwerden an und die Jungen zeigten die schlechteren schulischen Leistungen. Unterschiede ließen sich aber feststellen, wenn der ADHS-Subtypus berücksichtigt wurde. Jungen des kombinierten und hyperaktiv-impulsiven Subtypus wurden im sozialen und schulischen Bereich und im Selbstwertgefühl als stärker beeinträchtigt eingestuft, aber als gleich oder weniger beeinträchtigt im vorrangig unaufmerksamen Subtypus. Die Unterschiede scheinen insbesondere bei nichtklinischen Stichproben gering zu sein. So fanden Biederman und Kollegen (2005) in einer eher kleineren Stichprobe mit 73 männlichen und 25 weiblichen Probanden mit ADHS keine Unterschiede in der Zugehörigkeit zu Subtypen, Komorbidität und Behandlungsgeschichte. Rucklidge und Tannock (2002) fanden keine Geschlechtsunterschiede in kognitiven Leistungstests im Vergleich von 71 jugendlichen Probanden mit ADHS und 37 Probanden ohne ADHS bei gleicher Geschlechtsverteilung in den Gruppen.

Im Jugendalter werden die Unterschiede in der Diagnosehäufigkeit geringer oder sind nicht mehr festzustellen. So erfüllten 5,4 % von 10.987 isländischen Schülern im Alter von 14 bis 16 Jahren die Kriterien für das Vorliegen einer ADHS. Dabei waren es 5,1 % bei den Jungen und bei den Mädchen sogar 5,7 % (Gudjonsson, Sigurdsson, Sigfusdottir & Young, 2012). Für das Erwachsenenalter scheinen die Unterschiede ebenfalls geringer zu sein. Retz-Junginger und Kollegen (2007) verglichen 166 Frauen und 643 Männer mit ADHS. Unter Berücksichtigung der Subtypen fanden sie keine signifikanten Geschlechtsunterschiede bei einer persistenten ADHS des vorwiegend unaufmerksamen Subtypus. Grevet et al. (2006) untersuchten eine klinische Stichprobe mit 219 Erwachsenen mit ADHS mit einem Geschlechterverhältnis von 1,1 : 1. Männer und Frauen

hatten die gleiche Häufigkeit in den ADHS-Subtypen und es gab keine signifikante Interaktion zwischen Geschlecht und Subtyp in den erhobenen Variablen. Auch Babinski et al. (2011) fanden in einer gematchten Stichprobe mit insgesamt $N = 87$, dass die Ausprägungen relevanter klinischer Maße der Männer und Frauen mit ADHS sehr ähnlich waren.

Robison und Kollegen (2008) fanden in einer klinischen Stichprobe mit 188 Frauen und 348 Männern mit ADHS sogar mehr Symptome bei den Frauen als bei den Männern. Die Frauen zeigen darüber hinaus mehr Symptome emotionaler Dysregulation. Zudem waren mehr Frauen dem kombinierten Subtypus als dem vorrangig unaufmerksamen zuzuteilen. Von einer ADHS bezogenen Behandlung konnten Frauen und Männer gleichsam profitieren. Dieses Ergebnis zeigte sich auch in einer Studie an drei US-amerikanischen Universitäten mit 164 Studierenden mit ADHS und 710 Studierenden ohne ADHS (Fedele, Lefler, Hartung, & Canu, 2012). In einer online ausgefüllten Einschätzung ihrer ADHS-Symptomatik berichteten die Studentinnen mit ADHS mehr ADHS-Symptome als die Studentinnen ohne ADHS und auch mehr ADHS-Symptome als die Studenten mit ADHS. Problembereiche waren dabei sowohl das häusliche Leben und alltägliche Aktivitäten, der Sozialbereich, Schwierigkeiten beim Lernen und der Umgang mit Geld.

Zusammenfassend kann gesagt werden, dass im Kindesalter mehr Jungen als Mädchen die Diagnose ADHS bekommen. Dies hängt unter anderem auch mit einem Beurteilerfehler zusammen, nach dem Jungen eher die Diagnose ADHS bekommen als Mädchen bei gleicher Symptombeschreibung. Das Verhalten der Jungen scheint insbesondere in der Schule auffälliger zu sein als das der Mädchen. Unter anderem deswegen werden Jungen in der Kindheit eher zu einer Überprüfung auf ADHS vorgestellt. Sie bekommen im Kindesalter auch häufiger eine ADHS bezogene Behandlung. Während einige ältere Studien eine gravierendere Problematik bei den Jungen mit ADHS feststellten, zeigen neuere Studien, dass die Unterschiede relativ gering sind. Die Unterschiede sind insbesondere klein, wenn es sich um nichtklinische Stichproben handelt. Weiterhin scheinen die Unterschiede mit steigendem Alter abzunehmen. Im Jugend- bzw. Erwachsenenalter zeigen sich kaum noch Unterschiede in der Ausprägung der Symptomatik und Zuordnung zu Subtypen. Möglicherweise sind aber emotionale Probleme bei jugendlichen Mädchen und erwachsenen Frauen mit ADHS stärker ausgeprägt.

3.3 Prävalenz

Im Schulalter sind nach DSM-IV-TR 3 bis 7 % der Kinder von einer ADHS betroffen (Saß et al., 2003). Übereinstimmend fanden Ihle und Esser (2002) in einem Überblick über 19 internationale Studien mit Stichproben zwischen $N = 150$ und 3021 eine Periodenprävalenz von 4.4 %. Eine der größten internationalen Studien zur Prävalenz der ADHS umfasste 102 Studien aus den Jahren 1978 bis 2005 von beinahe allen Kontinenten mit einer Stichprobengröße von insgesamt $N = 171.756$ (Polanczyk et al., 2007). Die

gepoolte Punktprävalenz für eine ADHS lag bei 5.29 % bei einem Konfidenzintervall von 5.01 – 5.56 (95 %). Es wurde eine große Variabilität in den Ergebnissen der zugrunde liegenden Studie festgestellt, die nach den Autoren am ehesten dem unterschiedlichen methodischen Vorgehen der Studien zuzuschreiben sei. Eine geringere Rolle wird der geografischen Herkunft der Studien zugeschrieben. Zwischen den Ergebnissen der nordamerikanischen Studien und den Studien aus Europa, Südamerika, Asien und Ozeanien wurden keine signifikanten Unterschiede gefunden. Die Ergebnisse aus Afrika und dem Nahen Osten waren signifikant geringer als die Ergebnisse aus Nordamerika. Im Vergleich zu Studien mit Kindern wurde eine geringere Punktprävalenz für Jugendliche in Höhe von 2.75 % festgestellt bei einem Konfidenzintervall von 2.04 – 3.45 (95 %) (Polanczyk, 2009).

Eine weitere umfangreiche Metaanalyse mit einer Gesamtgröße von $N = 163.688$ aus 86 Studien für das Kindes- und Jugendalter hat Willcutt (2012) vorgelegt. Gemäß DSM-IV-TR-Kriterien ergab sich im Elternurteil eine Prävalenz von 6.1 % bei einem Konfidenzintervall von 5.2 – 7.1 (95 %), im Lehrerurteil von 7.1 % bei einem Konfidenzintervall von 6.6 – 7.5 (95 %) und nach einer kombinierten Einschätzung („best estimate“) von 5.9 % bei einem Konfidenzintervall von 4.6 – 7.5 (95 %). Die Studie gibt zudem Aufschluss über die altersgemäße Verteilung auf die ADHS-Subtypen. Es zeigte sich, dass im Alter von 3 bis 5 Jahren der hyperaktiv-impulsive Subtypus am häufigsten vertreten war. Für alle anderen Altersspannen war der unaufmerksame Subtypus am häufigsten vertreten, besonders ausgeprägt im Jugendalter, in dem der hyperaktiv-impulsive und der kombinierte Subtypus nur noch selten auftraten. Auch in dieser Metaanalyse wurden keine signifikanten Unterschiede in den Prävalenzen nach der geografischen Herkunft der Originalstudien gefunden. Allerdings konnten aus manchen Regionen nur wenige Studien gefunden werden.

Im Kinder- und Jugendgesundheitsurvey (KiGGS) wurden repräsentative Daten zur Diagnosehäufigkeit von ADHS in Deutschland von 7569 Jungen und 7267 Mädchen im Alter von 3 bis 17 Jahren erhoben (Huss et al., 2008). Für 4.8 % der Stichprobe war jemals die Diagnose ADHS gestellt worden. Zumeist wurden die Diagnosen beim Schuleintritt bzw. in den ersten Klassenstufen gestellt. Dies zeigt sich in einem Zuwachs der Diagnosehäufigkeit mit zunehmendem Alter. Während von den 3- bis 6jährigen nur 1.5 % jemals die Diagnose ADHS bekommen hatten, waren es bei den 7- bis 10jährigen 5.3 %, bei den 11- bis 13jährigen 7.1 % und bei den 14- bis 17jährigen 5.6 %. Die größte Diagnosehäufigkeit bzw. Lebenszeitprävalenz lag damit bei den Jugendlichen im Alter von 11 bis 13 Jahren.

Für das Erwachsenenalter schätzten Kessler und Kollegen (2006) basierend auf einer repräsentativen Umfrage aus den USA eine Prävalenz von 4.4 % für das aktuelle Vorliegen einer ADHS. An der Studie nahmen 3199 Teilnehmer zwischen 18 und 44 Jahren teil. Dies deckt sich mit Ergebnissen aus einer deutschen Stichprobe mit Erwachsenen im Alter von 18 bis 64 Jahren für die eine Prävalenzrate von 4.7 % festgestellt werden konnte (de Zwaan et al., 2012). In der Metaanalyse von Willcutt (2012) wurden 11 Stu-

dien zur Prävalenz von ADHS im Erwachsenenalter mit einer Größe von insgesamt $N = 14.112$ zusammengefasst. Es ergab sich eine Prävalenz von 5.0 % bei einem Konfidenzintervall von 4.1 – 6.2 (95 %). Am häufigsten war der unaufmerksame Störungssubtypus vertreten, gefolgt vom hyperaktiv-impulsiven Subtypus und letztlich dem kombinierten Subtypus. Eine etwas geringere Prävalenzrate von 3.4 % fanden Fayyad et al. (2007) in einem weltweiten ADHS-Screening. Insgesamt nahmen 11.422 Teilnehmer aus zehn Ländern in Amerika, Europa und dem Mittleren Osten teil. Die geringere mittlere Prävalenzrate schien mit der geografischen Herkunft der Teilnehmer zusammenzuhängen. So zeigte sich, dass die Prävalenzschätzung geringer ausfiel für Länder mit geringem Einkommen (1.9 %) als für Länder mit höherem Einkommen (4.2 %). Eine geringere Prävalenzrate von 2.5 % (Konfidenzintervall 2.1 – 3.1; 95 %) fanden Simon, Czobor, Balint, Meszaros und Bitter (2009) in einer Metaanalyse mit einer Stichprobengröße von $N = 5307$.

Für ältere Erwachsene liegt eine epidemiologische Studie aus den Niederlanden vor (Michielsen et al., 2012). Von 1494 Teilnehmern einer längsschnittlichen Altersstudie wurde ein ADHS-Screeningbogen ausgefüllt. Mit 231 Personen, die geantwortet hatten, wurde ein zusätzliches diagnostisches Gespräch durchgeführt. Es zeigte sich, dass bei 2.8 % der Stichprobe das klinische Vollbild der ADHS vorlag und bei insgesamt 4.2 % eine Teilsymptomatik. Die jüngeren Teilnehmer der Stichprobe im Alter von 60 bis 70 Jahren wiesen signifikant mehr ADHS-Symptomatik auf als die älteren Teilnehmer im Alter von 71 bis 90 Jahren.

Kessler et al. (2006) weisen darauf hin, dass es sich bei den Untersuchungen zum Erwachsenenalter eher um konservative Schätzungen handelt, da die aktuellen Diagnosekriterien, die für das Kindesalter formuliert wurden, für Erwachsene eher ungeeignet sind. Diese Einschätzung teilen auch McGough und Barkley (2004).

3.4 Komorbidität

Dass neben der ADHS noch eine weitere Störung vorliegt, ist sehr häufig und trifft auf ca. 2/3 aller Kinder mit ADHS zu (Gawrilow, 2012). Mittlerweile liegen internationale Studien mit großen Stichproben vor. Unter den 61.779 Kindern und Jugendlichen des US National Survey of Children's Health im Alter von 6 bis 17 Jahren befanden sich 5028 mit ADHS. Davon hatten 33 % eine zusätzliche, komorbide Erkrankung, 16 % zwei komorbide Erkrankungen und 18 % drei oder mehr Komorbiditäten. Im Elternurteil betrug die Prävalenz für ADHS in der Stichprobe 8.2 %. Die Wahrscheinlichkeit, dass eine weitere Störung vorlag, war für die Kinder und Jugendlichen mit ADHS signifikant höher. Im Elternurteil zeigte sich, dass 46 % der Kinder mit ADHS zusätzlich eine Lernstörung aufwiesen im Vergleich zu 5 % der Kinder ohne ADHS. 27 % der Kinder mit ADHS hatten eine Störung des Sozialverhaltens und nur 2 % der Kinder ohne ADHS. Angststörungen waren bei 18 % der Kinder mit ADHS komorbid vorhanden und nur bei 2 % der Kinder ohne ADHS. Eine Depression wurde bei 14 % der Kin-

der mit ADHS festgestellt und nur bei einem Prozent der Vergleichsfelder. Sprachprobleme lagen bei 12 % der Kinder mit ADHS vor und nur bei 3 % der Kinder ohne ADHS (Larson, Russ, Kahn & Halfon, 2011). Für eine puertoricanische Stichprobe konnten Bauermeister und Kollegen (2007) zeigen, dass die Komorbiditätsraten sowohl in einer klinischen Stichprobe als auch in einer nichtklinischen Stichprobe ähnlich hoch ausgeprägt waren. In der untersuchten Stichprobe von insgesamt 2660 Kindern und Jugendlichen im Alter von 4 bis 17 Jahren befanden sich 344 mit einer ADHS. Bei 13.18 % der Kinder und Jugendlichen mit ADHS aus der nichtklinischen Stichprobe bzw. 10.22 % aus der klinischen Stichprobe lag eine komorbide Störung des Sozialverhaltens vor, bei 38.89 bzw. 61.19 % eine Störung mit oppositionellem Trotzverhalten, bei 24.43 bzw. 33.51 % eine Angststörung und bei 9.27 bzw. 22.73 % eine depressive Störung. Ähnliche Ergebnisse lieferte eine Geburtskohortenstudie, in der sich 379 Kinder mit ADHS befanden, die mit 758 gematcht zugeordneten Kontrollkindern bis zu einem Alter von 19 Jahren untersucht wurden (Yoshimasu et al., 2012). Das Risiko einer komorbiden Erkrankung war in der ADHS-Gruppe stark erhöht. Am höchsten war das Risiko für Störungen des Sozialverhaltens und oppositionellem Trotzverhalten, gefolgt von Ticstörungen, Persönlichkeitsstörungen, Essstörungen, Substanzabhängigkeiten, Anpassungsstörungen, Affektiven Störungen und Angststörungen. In einer 10-jährigen Längsschnittstudie mit anfangs 260 Kindern im Alter von 6 bis 18 Jahren wurde festgestellt, dass das kumulative Morbiditätsrisiko im Alter von 21 Jahren höher war als für eine Kontrollgruppe (Biederman et al., 2006). Das Risiko war signifikant höher für Depression und Psychosen (Faktor 6.1), Angststörungen (Faktor 2.2), antisoziale Störungen (Faktor 5.9), Entwicklungsstörungen (Faktor 2.5) und Substanzmissbrauch (Faktor 2.0). In einer Studie über 13 Jahre wurden 149 hyperaktive Kinder und 72 Kindern ohne ADHS untersucht (Barkley, Fischer, Smallish & Fletcher, 2006). Neben Problemen in Ausbildung, Arbeit, Sozialkontakten und Finanzen wurde festgestellt, dass die jungen Erwachsenen früher Eltern wurden und häufiger aufgrund von sexuell übertragbaren Krankheiten in Behandlung waren. Neben komplett erfüllten weiteren Störungsbildern liegen oftmals weitere Symptome vor, die keinen Syndromcharakter erreichen. Connor und Ford (2012) konnten dies bei 245 Kindern und Jugendlichen mit ADHS im Vergleich zu einer Kontrollgruppe mit $n = 72$ zeigen. Es lagen in der ADHS-Gruppe sowohl mehr externalisierende als auch mehr internalisierende Symptome vor.

Insbesondere bei Mädchen mit ADHS wurde ein erhöhtes Risiko für das Vorliegen von Essstörungen gefunden. In einem längsschnittlichen Vergleich von 123 Mädchen mit ADHD und 122 Mädchen ohne ADHS stellten Biederman und Kollegen (2007) fest, dass von den Mädchen mit ADHS 16 % jemals unter einer Essstörung gelitten hatten. Bei 30 % dieser Gruppe lag eine anorektische Störung vor, bei 50 % eine bulimische und bei 20 % beide Störungen zu jeweils unterschiedlichen Zeitpunkten. Beginn der Störungen war durchschnittlich im Alter von 12.3 Jahren. Hingegen waren in der Gruppe der Mädchen ohne ADHS nur 5 % von einer Essstörung betroffen. Davon lag in 60 % der Fälle eine Anorexie vor, in 20 % eine Bulimie und in 20 % sowohl eine anorektische als auch bulimische Störung. Der Störungsbeginn lag in der Kontrollgruppe im Durchschnitt bei 14 Jahren. Die Mädchen mit ADHS und zusätzlicher Essstörung hatten

zudem ein signifikant höheres Risiko, dass zusätzlich affektive Störungen und Störungen des Sozialverhaltens bzw. Störungen mit Trotzverhalten vorlagen.

Auf einer konzeptuellen Ebene ordnen Taurines und Kollegen (2010) die Komorbidität nach ihrem zeitlichen Auftreten im Sinne einer „developmental comorbidity“. Sie unterscheiden Präkomorbidität, die auftritt, bevor sich ADHS-Symptome erstmalig manifestieren, simultane Komorbidität, die gleichzeitig mit der ADHS auftritt, und Postkomorbidität, die nach der Manifestation der ADHS auftritt, zumeist mit Beginn in der Adoleszenz (siehe Abbildung 7).

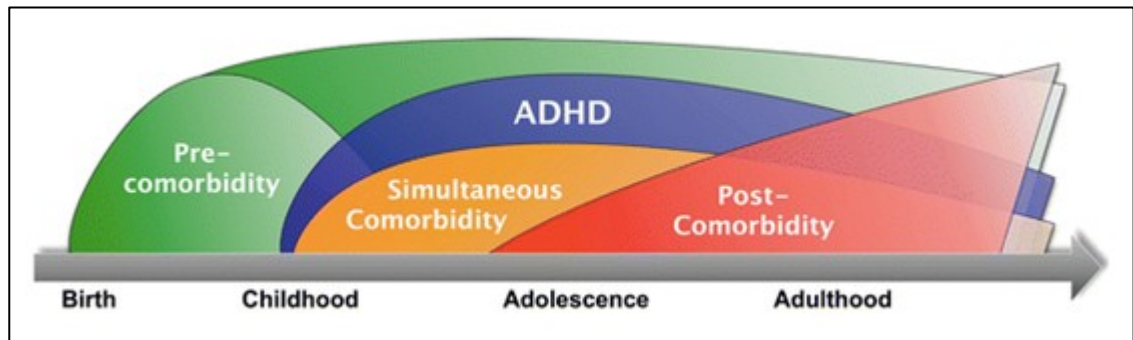


Abbildung 7: Zeitlicher Verlauf des Auftretens von ADHS und komorbiden Störungen (aus Taurines et al., 2010, S. 269)

Zusätzlich zu psychiatrischen Störungsbildern bringen die Autoren auch nichtpsychiatrische Erkrankungen wie das Atopische Ekzem („Neurodermitis“) und damit verbundene Schlafstörungen und Einflüsse des kindlichen Temperaments in Zusammenhang mit ADHS. Abbildung 8 zeigt die von den Autoren auf Grundlage empirischer Studien zusammengestellten komorbiden Konditionen geordnet nach ihrem zeitlichen Auftreten.

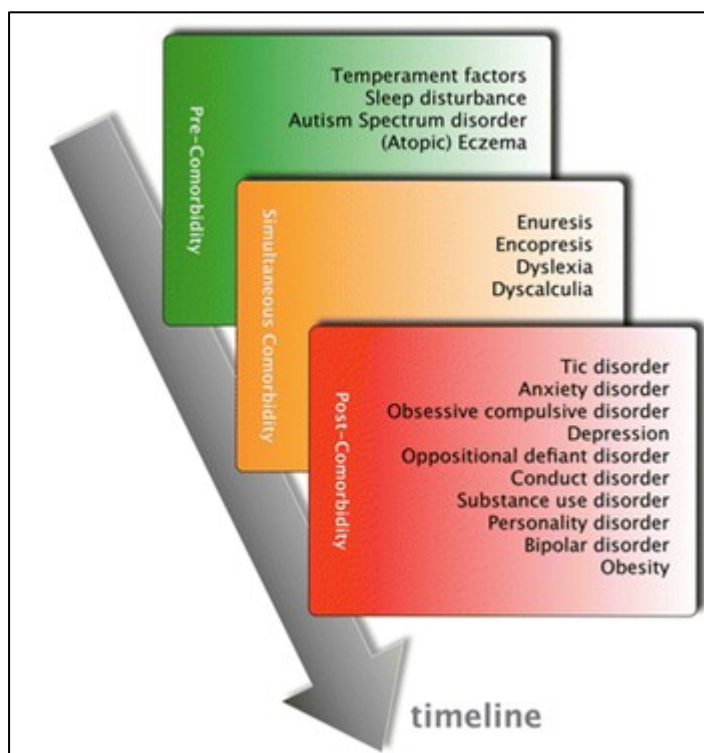


Abbildung 8: Vermutete Prä-, simultane und Postkomorbidität der ADHS (aus Taurines et al., 2010, S. 268)

In einer für Deutschland repräsentativen Stichprobe mit 327 Teilnehmern im Alter von 14 bis 25 Jahren zeigte sich, dass Betroffene von ADHS eine deutlich negative psychische Befindlichkeit angaben. Diese blieb über die Lebensspanne stabil. Die ADHS Betroffenen gaben an, durchweg höher belastet zu sein Vergleichspersonen ohne ADHS (Schmidt, Brähler, Petermann & Koglin, 2012). In einer Studie mit 211 Erwachsenen mit ADHS konnte gezeigt werden, dass eine erhöhte ADHS-Problematik mit einer höheren Rate an aktuellen negativen Lebensereignissen verbunden war. Die Autoren führen weiter aus, dass die negativen Lebensereignisse der Patienten mit ADHS nur mit der Schwere der ADHS zusammenhängen, unabhängig von der Komorbidität (Garcia et al., 2012).

Bei vielen Kindern und Jugendlichen mit ADHS sind Probleme im schulischen Lern-Leistungsbereich feststellbar. Im Vergleich zu 740 Kindern ohne ADHS hatten 370 Kinder mit ADHS signifikant schlechtere Ergebnisse in den Lesefähigkeiten, waren häufiger schulabstinent, wiederholten dreimal häufiger eine Klasse und verließen die Schule 2,7-mal häufiger ohne Abschluss (Barbatesi, Katusic, Colligan, Weaver & Jacobsen, 2007). Auch im US-amerikanischen „National Survey of Children’s Health“ mit insgesamt 61.779 hatten die 5.028 Kinder mit ADHS ca. 5-mal häufiger schulische Probleme als die Kinder ohne ADHS (Larson et al., 2011). In einer Regressionsanalyse in einer allgemeinen Stichprobe von 324 Jugendlichen war die ADHS-Symptomatik ein signifikanter Prädiktor für schulische Leistungen (Birchwood & Daley, 2012). Auch in einer Studie mit 2.463 Schülern aus Taiwan waren ADHS-Symptome negativ mit schulischer Leistungen und darüber hinaus positiv mit sozialen Problemen verbunden (Kawabata, Tseng & Gau, 2012).

Die meisten Studien finden ein höheres Risiko für Suchterkrankungen bei Jugendlichen mit ADHS. Im Vergleich zu 38 College-Studenten ohne ADHS zeigten 53 College-Studenten mit ADHS mehr Konsum legaler und illegaler Drogen (Rooney, Chronis-Tuscano & Yoon, 2012). Sie rauchten mehr Tabak, zeigten häufiger gefährlichen Alkoholkonsum und hatten größere Einschränkungen aufgrund des Konsums von Marihuana und anderen illegalen Drogen. Die Probleme waren unabhängig von einer Störung des Sozialverhaltens in der Vorgeschichte. In einer Studie von Glass und Flory (2012) nahmen 889 Bachelorstudierende an einem Online-Survey teil. Korrelationen von ADHS-Symptomatik mit allgemeinem Alkoholkonsum und mit illegalem Drogengebrauch waren nicht mehr signifikant, wenn für Symptome der Störung des Sozialverhaltens kontrolliert wurde. Allerdings blieben die Zusammenhänge der ADHS-Symptome, insbesondere der Unaufmerksamkeit, mit Nikotinkonsum und problematischem Alkoholkonsum auch nach der Kontrolle bestehen. In einer isländischen Studie wurden 10.987 Schüler im Alter von 14 bis 16 Jahren untersucht, von denen 594 von einer ADHS betroffen waren (Gudjonsson et al., 2012). Es zeigte sich ein signifikanter Zusammenhang zwischen ADHS-Symptomen und dem Konsum von Nikotin, Alkohol und illegalen Drogen. Die Anzahl der konsumierten illegalen Drogen war in der ADHS-Gruppe signifikant höher. Jugendliche mit ADHS zeigten insbesondere einen stärkeren Konsum von LSD, Kokain, halluzinogenen Pilzen und Amphetaminen. Deshalb vermuten die Auto-

ren, dass der Konsum der ADHS-Gruppe teilweise auch der Selbstmedikation diene. Nach Kontrolle von Geschlecht und Schulklasse konnten ADHS-Symptome in logistischen Regressionen Nikotin- und Alkoholkonsum und den Konsum illegaler Drogen unabhängig von Angst- und depressiver Symptomatik und dem Establishment entgegengesetzten Einstellungen vorhersagen.

ADHS-Betroffene sind von einem höheren Risiko für suizidale Gedanken, Handlungen und versuchte und erfolgreiche Suizide betroffen (Impey & Heun, 2012). In einer Studie von Chronis-Tuscano und Kollegen (2010) wurden 125 Kinder mit ADHS und 123 gematchten Kontrollkindern über einen Zeitraum von 18 Jahren fortlaufend untersucht. Das Risiko für einen versuchten Selbstmord war 3,6-mal höher in der ADHS-Gruppe. Weiterhin war das Risiko für eine Depression und Dysthymia mehr als viermal so hoch. In diesem Zusammenhang wird die Rolle der Emotionsregulation als Mediator diskutiert. So hatten in einem Vergleich zu 32 Jugendlichen ohne ADHS im Alter von 10 bis 14 37 Jugendliche mit ADHS ein höheres Maß an depressiver Symptomatik. Gleichzeitig zeigten sie aber auch geringere Fähigkeiten in der Emotionsregulation, die als vollständiger Mediator des Zusammenhangs der ADHS-Symptome mit der Depressivität fungierten (Seymour et al., 2012). In einer finnländischen Studie wurden 104 Jugendliche mit ADHS und 169 Jugendliche ohne ADHS verglichen (Hurtig, Taanila, Moilanen, Nordstrom, & Ebeling, 2012). Die Jugendlichen mit ADHS hatten signifikant häufiger suizidale Gedanken (57 % vs. 28 %) und zeigten zudem signifikant häufiger selbstverletzende Verhaltensweisen (69 % vs. 32 %). In beiden der zuvor vorgestellten Studien war das Risiko höher für Mädchen als für Jungen. In der deutschen BELLA-Studie, an der 2.863 Familien teilnahmen, wurden starke Zusammenhänge zwischen suizidalen Verhaltensweisen und Hyperaktivität festgestellt (Resch, Parzer & Brunner, 2008). So waren die Skalenwerte der Hyperaktivität im Selbsturteil des SDQ bei Jugendlichen, die sich selbst verletzten oder Selbstmordversuche unternommen hatten, fast eine halbe Standardabweichung höher als die der unauffälligen Jugendlichen.

Auch für Erwachsene mit ADHS werden hohe Komorbiditätsraten festgestellt. Zwischen 65 und 89 % der erwachsenen Patienten mit ADHS weisen zumindest eine zusätzliche psychische Störung auf (Sobanski, 2006). Die hohe Komorbiditätsrate ist nicht allein auf einen klinischen Bias zurückzuführen. So untersuchten McGough und Kollegen (2005) 435 Eltern von Kindern mit ADHS in einem nichtklinischen Setting. Von den Eltern mit ADHS hatten 87 % eine weitere und 56 % mindestens zwei weitere psychiatrische Störungen, während die Eltern ohne ADHS geringere nur Raten von 64 % bzw. 27 % aufwiesen.

Empirische Studien zeigen eine signifikant höhere Lebenszeitprävalenz für Erwachsene mit ADHS im Vergleich zu Erwachsenen ohne ADHS. Im Bereich affektiver Störung zeigt sich, dass 35 bis 50 % aller Erwachsenen mit ADHS von zumindest einer depressiven Episode im jeweiligen Untersuchungszeitraum betroffen waren. In der Normalpopulation sind es hingegen nur 15 %. Weiterhin sind 40 bis 60 % von einer oder mehreren Angststörungen betroffen, insbesondere von sozialen Phobien, die bei 20 bis 34 %

der Betroffenen auftreten. Bis zu 50 % der Erwachsenen mit ADHS weisen einen Substanzmissbrauch auf. Weiterhin beklagen viele Erwachsene mit ADHS einen gestörten Schlaf und es liegen häufiger antisoziale Persönlichkeitsstörungen vor (Esther Sobanski, 2006).

Einen Zusammenhang gibt es auch von ADHS mit der Borderline-Persönlichkeitsstörung, die große Überschneidungen in der Symptomatik aufweist, insbesondere in der Affektregulation und Selbstorganisation (Schmidt & Petermann, 2008). In einer klinischen Studie mit 181 Erwachsenen mit einer Borderline-Persönlichkeitsstörung wiesen 38 % eine komorbide ADHS auf. Diese Gruppe zeigte ein stärkeres Ausmaß an Substanzmissbrauch, Antisozialer Persönlichkeitsstörung und Zwangsstörung (Ferrer et al., 2010). Der Störungsbeginn bei Substanzmissbrauch und delinquenter Verhaltensweisen bei der Antisozialen Persönlichkeitsstörung ist für ADHS-Betroffene früher (Sobanski, 2006). In einer Stichprobe mit 266 männlichen und 181 weiblichen Erwachsenen mit ADHS zeigten sich ebenso die beschriebenen höheren Prävalenzraten. Darüber hinaus zeichnete sich ein Zusammenhang mit dem Geschlecht der Patienten ab. Männer waren stärker von der Antisozialen Persönlichkeitsstörung betroffen und zeigten stärkeren Substanzmissbrauch, während die Frauen stärker von Panikstörung, Borderline-Persönlichkeitsstörung und von Anorexie und Bulimie in der Vergangenheit betroffen waren (Cumyn, French & Hechtman, 2009). In einer Langzeitstudie über 34 Jahre wurde eine Stichprobe von 551 Kindern begleitet, von denen 72 von ADHS betroffen waren (Brook, Brook, Zhang, Seltzer & Finch, 2013). Dabei wurden sechs Messzeitpunkte umgesetzt. Die ADHS-Gruppe wies einen höheren Stresslevel auf, hatte im Gruppenvergleich eine beeinträchtigte physische und psychische Gesundheit, eine beeinträchtigte Arbeitsleistung und mehr finanziellen Stress. In einer deutschen Stichprobe im Alter zwischen 18 und 64 Jahren mit einer Größe von $N = 1.655$ zeigte sich für die in der Stichprobe befindlichen 78 Erwachsenen mit ADHS ein geringeres Bildungsniveau, mehr Arbeitslosigkeit, weniger Eheschließungen und mehr Scheidungen (de Zwaan et al., 2012). Weiterhin wurden in Screenings für Depression und Angststörungen in der ADHS-Gruppe auffälliger Werte erzielt.

Cherkasova, Sulla, Dalena, Pondé und Hechtman (2013) tragen die Ergebnisse von elf kontrollierten prospektiven Follow-up-Studien zusammen. Es wurden Effektstärken zum Vergleich der ADHS-Gruppen mit den Kontrollgruppen bestimmt, wobei nicht zu jedem Bereich Ergebnisse aus allen Studien vorlagen. Im Bereich Bildung, der unter anderem Klassenwiederholung und erreichte Schulabschlüsse beinhaltet, finden sie einen Effekt von $d = .63$, für den Bereich Arbeit und Beschäftigung einen Effekt von $d = .41$, im Bereich Komorbidität von $d = .56$, für den Bereich Kriminalität von $d = .77$, für Auffälligkeiten im Straßenverkehr von $d = 1.04$, für Autounfälle $d = 1.46$, für partnerschaftliche bzw. eheliche Probleme von $d = .58$, für risikoreiche sexuelle Praktiken von $d = .45$ und für frühe Schwangerschaften von $d = 1.18$. Die Ergebnisse der drei letztgenannten Bereiche basieren allerdings jeweils lediglich auf einer bzw. auf zwei Studien.

Besonderes Augenmerk ist auf die Komorbidität der ADHS mit den anderen externalisierenden Verhaltensstörungen zu richten. Externalisierende Verhaltensweisen umfassen neben hyperkinetischen Symptomen aggressives und oppositionelles bzw. Trotzverhalten (Petermann et al., 2004). Das DSM-IV-TR unterscheidet zwischen der Störung des Sozialverhaltens (SSV), der Störung mit Oppositionellem Trotzverhalten (SOT) und der Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS). In klinischen Stichproben zeigen sich Übereinstimmungen zwischen 84 und 96 % zwischen SSV und SOT, so dass es fraglich erscheint, ob es sich wirklich um zwei unabhängige Störungen handelt (Petermann et al., 2004). Vielmehr scheinen sich zwei Dimensionen unterscheiden zu lassen, eine, die aggressives Verhalten und eine, die dissoziales und delinquentes Verhalten abbildet.

Kinder mit SSV und SOT haben ein erhöhtes Risiko für familiäre Probleme und weiterer klinisch relevanter Probleme, auch Angststörungen und Depression. Kinder mit gleichzeitigem Vorliegen von SSV und ADHS haben stärkere funktionale Einschränkungen bei der Impulskontrolle und der sozial-kognitiven Informationsverarbeitung. Sie zeigen mehr aggressive Verhaltensweisen und stabilere aggressive und dissoziale Verhaltensweisen. Zudem werden sie von Gleichaltrigen stärker zurückgewiesen als Kinder mit nur einer der beiden Störungen. Es wird daher diskutiert, ob es sich um einen eigenständigen Störungstyp handelt, dem gemeinsame dysfunktionale Verhaltensweisen zugrunde liegen, z. B. eine mangelhafte Impulskontrolle und somit soziale Interaktionen gestört werden, die aggressives Verhalten begünstigen (Petermann et al., 2004). Wie von Sroufe (1997) herausgestellt, weist die große Überlappung von Störungsbildern und Verhaltensweisen darauf hin, dass es sich nicht um abgrenzbare Störungsbilder handelt und die Integrität vorherrschender klinischer Kategorien unklar ist.

Zum ätiologischen Zusammenhang zwischen SSV bzw. SOT und ADHS ist zu sagen, dass einer SSV in der Entwicklung oftmals eine ADHS vorhergeht (Petermann et al., 2004). Dies gilt insbesondere für den hyperaktiv-impulsiven ADHS-Subtypus und nicht für den unaufmerksamen Störungssubtypus. In Follow-up-Studien konnte gezeigt werden, dass eine ADHS-Symptomatik späteres antisoziales Verhalten vorhersagt, aber nicht andersherum. Dieser Zusammenhang ist allerdings vorsichtig kausal zu interpretieren, da es auch sein könnte, dass sich beide Störungen lediglich zu einem unterschiedlichen Alter manifestieren (Rutter, Kim-Cohen & Maughan, 2006). Ein möglicher Entwicklungsverlauf ist in Abbildung 9 dargestellt.

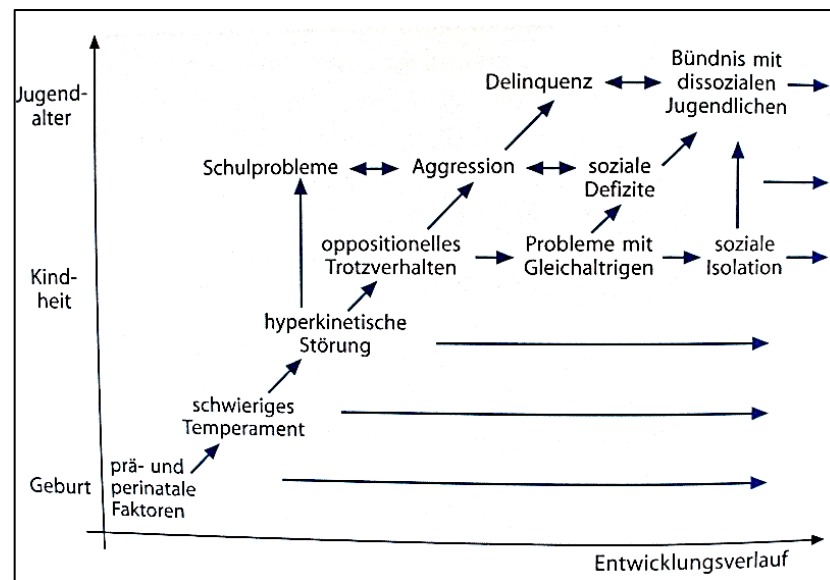


Abbildung 9: Entwicklungsmodell aggressiv-dissozialen Verhaltens, über den Lebenslauf stabiler Typus (aus Petermann, Niebank & Scheithauer, 2004, S. 384)

Zusammenfassen lässt sich, dass Jugendliche mit Aufmerksamkeitsstörungen insbesondere Schwierigkeiten im sozialen oder zwischenmenschlichen und im schulischen oder Leistungsbereich aufweisen. Es wurde bereits thematisiert, dass diese Bereiche für das Jugendalter allgemein von großer Bedeutung sind und sich aus ihnen herausfordernde Entwicklungsaufgaben ableiten, deren Bewältigung nicht immer reibungslos abläuft. Das Vorliegen einer Aufmerksamkeitsstörung lässt sich als ein Risikofaktor verstehen, der die Bewältigung dieser Entwicklungsaufgaben zusätzlich erschweren kann. Insbesondere die gesteigerten Anforderungen an Autonomie bereiten Jugendlichen mit Aufmerksamkeitsstörungen häufig Schwierigkeiten, da ihnen selbständiges Organisieren und planvolles Agieren oft besonders schwer fallen. So verschärfen sich die spezifische Entwicklungssituation und die störungsspezifische Problematik gegenseitig.

Abschließend ist zu sagen, dass die Konzeptualisierung psychischer Störungen und Entwicklungsabweichungen im Kindes- und Jugendalter noch viel Bedarf an Weiterentwicklung aufweist. Die vorherrschenden medizinischen Kategoriensysteme, an denen sich die Klassifikation der ADHS orientiert, betonen den aktuellen Entwicklungsstand des Individuums, der sich in der akuten Symptomatik äußert, sehr stark. Dabei werden die vorausgegangene Adaptationsgeschichte und die Entwicklungsperspektive vernachlässigt.

3.5 Verlauf

Bis zur vierten Auflage des DSM und auch in der ICD-10 werden Aufmerksamkeitsstörungen als Störungen des Kindes- und Jugendalters konzeptualisiert. Eine Diagnose für Erwachsene ist nicht möglich, da lange davon ausgegangen wurde, dass sich die Störung über das Jugendalter auswächst und spätestens im Erwachsenenalter nicht mehr vorliegt. Dieser Konzeptualisierung wird mittlerweile ein Verständnis von der ADHS

als einer lebenslangen Bedingung bzw. Störung vorgezogen, die sich alters- und entwicklungsentsprechend über den Lebensverlauf unterschiedlich äußert (Schmidt & Petermann, 2008; Turgay et al., 2012).

In einer Metaanalyse fassten Faraone, Biederman und Mick (2006) 32 kontrollierte Follow-up-Studien zur Persistenz der ADHS zusammen. Es zeigte sich, dass der volle diagnostische Status im Alter von 25 Jahren nur noch bei ca. 15 % der ADHS-Betroffenen erreicht wird. Dies wird von den Autoren als syndromatische Persistenz bezeichnet. Weitere Analysen zeigten aber, dass noch ca. 65 % der Stichprobe eine Teilsymptomatik behielt, was die Autoren als symptomatische Persistenz bezeichnen. Ramsay und Rostain (2007) fassen in einem Review zusammen, dass in den meisten Studien Persistenzraten zwischen 50 und 70 % berichtet werden.

Zum Entwicklungsverlauf der ADHS-Kernsymptomatik existieren einige Längsschnittstudien:

Hart, Lahey, Loeber, Applegate und Frick (1995) untersuchten eine Stichprobe von 106 Jungen mit ADHS, die anfangs im Durchschnitt 9,4 Jahre alt waren. In einem Zeitraum über vier Jahre mit jährlichen Messzeitpunkten zeigte sich, dass alle ADHS-Symptome rückläufig waren. Die Symptome der Unaufmerksamkeit reduzierten sich allerdings nur vom ersten zum zweiten Messzeitpunkt, während die Symptome der Hyperaktivität und Impulsivität über alle Messzeitpunkte zurückgingen. Da keine Kontrollgruppe vorlag, schließen die Autoren auf einen Regressionseffekt in der Unaufmerksamkeit. Der Rückgang der Symptomatik war stärker für die Elterneinschätzung als für die Lehrereinschätzung.

Lahey, Pelham, Loney, Lee and Willcutt (2005) erfassten die Symptomatik von 118 4- bis 6-jährigen Kindern mit ADHS. Verteilt über acht Jahre wurden sieben Messzeitpunkte realisiert. Die Symptome der Hyperaktivität und Impulsivität gingen zurück, nicht aber die Unaufmerksamkeit.

Biederman, Mick and Faraone (2000) untersuchten eine Stichprobe von 128 Jungen mit ADHS über vier Jahre, in denen fünf Messzeitpunkte durchgeführt wurden. Alle Symptome waren rückläufig, insbesondere die Hyperaktivität und Impulsivität. Bei einem Follow-up nach zehn Jahren mit 110 verbleibenden Teilnehmern wurde ein starker Abfall der Aufmerksamkeitsproblematik zwischen der Altersgruppe 19 bis 21 und 22 bis 24 ersichtlich (J. Biederman, Petty, Evans, Small, & Faraone, 2010). Für die über 22-jährigen waren die Persistenzraten für alle Symptomgruppen sehr ähnlich. Die Autoren diskutieren, ob der Übergang ins Berufsleben zur Reduktion der Probleme der Unaufmerksamkeit beigetragen haben könnte. Die jungen Erwachsenen könnten solche Berufe gewählt haben, in denen ihre Probleme weniger Relevanz haben und nicht mehr als problematisch eingeschätzt werden.

In einem Längsschnitt wurden 140 Mädchen mit ADHS im Alter von 11 bis 18 Jahren begleitet (Hinshaw, Owens, Sami & Fargeon, 2006). Im Verlauf über fünf Jahre zeigte

sich eine Reduktion der Symptomatik, stärker für die Hyperaktivität und Impulsivität als für die Unaufmerksamkeit. Im Vergleich mit 88 Kontrollpersonen wiesen die Mädchen mit ADHS kontinuierlich stärkere Beeinträchtigungen auf. Eine Studie mit 96 Mädchen mit ADHS und 91 Mädchen ohne ADHS, die im Alter zwischen 6 und 17 Jahren zum ersten Mal und im Alter von 15 bis 30 Jahren erneut untersucht wurden, kommt zu ähnlichen Ergebnissen (Biederman, Petty, O'Connor, Hyder & Faraone, 2012). Beim Follow-up nach elf Jahren erfüllten noch 33.3 % der Stichprobe die ADHS-Symptomkriterien komplett, 29.2 % zeigten noch eine Teilsymptomatik, 10.4 % hatten funktionale Einschränkungen und bei 4.2 % war die Störung zwar remittiert, es fand aber trotzdem eine Behandlung statt. Insgesamt waren also bei 77.1 % der Stichprobe auch elf Jahre später noch Probleme vorhanden.

In einer schwedischen Stichprobe wurden längsschnittliche Daten von 1480 Zwillingspaaren erhoben (Larsson, Larsson & Lichtenstein, 2004). Die Jungen in der Stichprobe hatten signifikant mehr Symptome als die Mädchen. Die ADHS-Symptome gingen zwischen dem Alter 8 bis 9 und 13 bis 14 sowohl für Jungen als auch für Mädchen zurück. Der Rückgang war am stärksten für Hyperaktivität und Impulsivität. Die Unaufmerksamkeit ging nur bei den Jungen zurück. In einem Follow-up im Alter von 16 bis 17 Jahren waren die Hyperaktivität und Impulsivität bei den Mädchen nicht mehr signifikant geringer als bei den Jungen (Larsson, Lichtenstein & Larsson, 2006). Der Rückgang in der Symptomatik wurde nur für die Hyperaktivität und Impulsivität über alle drei Messzeitpunkte signifikant, nicht aber für die Unaufmerksamkeit. Die Autoren formulieren zwei differenzierte Entwicklungsmuster („trajectories“). Für die Hyperaktivität und Impulsivität wurde ein Muster mit konstant geringer Ausprägung gefunden und ein Muster mit einer anfangs hohen Ausprägung, die dann geringer wird. Für die Unaufmerksamkeit ein Muster mit konstant geringer Ausprägung und eins mit konstant hoher bzw. ansteigender Ausprägung.

Übereinstimmend werden also ein Rückgang der Hyperaktivität und Impulsivität berichtet und ein weniger starker Rückgang bzw. Persistenz der Unaufmerksamkeit. Dies mag damit zusammenhängen, dass Hyperaktivität und Impulsivität mit steigendem Alter dafür sorgen, dass sich Probleme der Unaufmerksamkeit überhaupt erst ausbilden oder verfestigen. Zu diesem Schluss kommen die Autoren einer repräsentativen Stichprobe mit ungefähr 7000 Zwillingspaaren aus Großbritannien (Greven, Rijdsdijk, Asherson & Plomin, 2012). Dabei konnten Symptome der Hyperaktivität und Impulsivität in der mittleren Kindheit Symptome der Unaufmerksamkeit in der frühen Adoleszenz vorher-sagen, aber nicht umgekehrt. Dementgegen steht eine Studie, bei der die Hyperaktivität mit einem Bewegungserfassungssystem gemessen wurde anstatt mit Ratingfragebögen (Teicher, Polcari, Furligas, Vitaliano & Navalta, 2012). Im Vergleich von 40 Erwachsenen mit ADHS und 60 Kontrollpersonen zeigte sich, dass die Personen mit ADHS bei der Bearbeitung einer kognitiven Kontrollaufgabe signifikant mehr körperliche Aktivität zeigten. Die Autoren schlussfolgern, dass Hyperaktivität auch im Erwachsenenalter vorhanden ist und wie sie in ROC-Analysen darlegen auch als valider Indikator für die Diagnostik genutzt werden kann. Ein anderer Grund für die gefundenen Veränderungen

kann darin vermutet werden, dass die zur Erfassung der Symptomatik benutzten Fragebögen, die zumeist am DSM-IV orientiert sind, nicht entwicklungsensibel sind (Brassett-Harknett & Butler, 2007). Die Formulierung der Verhaltenskriterien orientiert sich an der Symptomatik der Störung im Kindesalter und ist wenig angemessen für Jugendliche und noch weniger für Erwachsene. So fanden Todd und Kollegen (2008), dass die Stabilität der ADHS-Symptome über fünf Jahre sehr gering war. In einer Stichprobe von 708 Zwillingen im Alter von 7 bis 19 Jahren war der Prozentsatz von zum ersten Messzeitpunkt erfüllten Symptomen gemäß DSM-IV, die nach fünf Jahren immer noch vorlagen, unter 50 % für jedes der Kriterien. Besonders gering waren die Prozentzahlen für die Symptome der Hyperaktivität und Impulsivität. Am geringsten war die Stabilität für das Herausplatzen mit Antworten (24 %), Schwierigkeiten sich ruhig zu beschäftigen (26 %), ständig auf Achse sein (28 %) und häufig den Platz verlassen (28 %). Diese Formulierungen scheinen nicht angemessen für die Ausprägung der ADHS im Jugendalter zu sein. Ob die gefundenen Veränderungen der Kernsymptomatik auf reifungsbedingte Veränderungen zurückzuführen sind oder mit der Abhängigkeit von der Erfassungsmethode zusammenhängen oder beides, ist bislang nicht endgültig geklärt.

Nicht nur die ADHS-Kernsymptomatik persistiert, sondern auch damit verbundene Problematik, wie eine prospektive Follow-up-Studie mit 140 Jungen mit ADHS und 120 Jungen ohne ADHS zeigt (Biederman et al., 2012). Auch noch nach 16 Jahren hatten die Jungen mit ADHS signifikant schlechtere Werte bezogen auf antisoziales Verhalten, affektive Symptome und Suchterkrankungen und hatten mehr Probleme im psychosozialen, bildungsbezogenen und neuropsychologischen Bereich. Die gefundenen Gruppenunterschiede konnten nicht durch andere psychopathologische Symptome erklärt werden. Die Autoren schließen, dass ADHS ein starkes Risiko für psychische Probleme im Erwachsenenalter mit sich bringt. Dies gilt ebenso für gesundheitsbezogene und finanzielle Probleme im Erwachsenenalter, wie die Langzeitstudie von Brook und Kollegen (2013) über 34 Jahre belegen konnte. In einer weiteren Langzeitstudie mit einer Geburtskohorte von 5718 Kindern, davon 367 mit ADHS, konnte gezeigt werden, dass die Kinder mit ADHS ein höheres Sterblichkeitsrisiko hatten als die Kontrollpersonen.

Die entwicklungsgemäße Ausgestaltung der ADHS-Problematik äußert sich über die Lebensspanne unterschiedlich. Auffälligkeiten zeigen sich oftmals bereits im frühen Kleinkindalter. So wurden Zusammenhänge von frühen Regulationsstörungen, die sich in anhaltendem Schreien und Schlaf- und Fütterstörungen äußern, mit späterer Hyperaktivität gefunden (Wolke, Rizzo & Woods, 2002). Es wird ein Zusammenhang mit Polymorphismen der mit dem Transport und der Aufnahme von Dopamin verbundenen Gene diskutiert, bislang sind aber keine kausalen Schlüsse möglich (Becker, El-Faddagh, Schmidt & Laucht, 2007). Im Kindesalter fällt insbesondere die gesteigerte Aktivität der Kinder auf. Hinzu kommen häufig aggressive Verhaltensweisen, Schwierigkeiten sich an Regeln zu halten und eine geringe Frustrationstoleranz (Newcorn, 2008). Beim Übergang in die Grundschule kommt es in Deutschland zum stärksten An-

stieg der gestellten ADHS-Diagnosen (Huss et al., 2008), da oftmals zu diesem Zeitpunkt die Probleme im Vergleich zu den Gleichaltrigen auffällig werden und die schulischen Anforderungen an die Aufmerksamkeit der Kinder nicht zufriedenstellend erfüllt werden können. Aber auch schon im Vorschulalter werden ADHS-Diagnosen gestellt. So wurde in der für Deutschland repräsentativen KiGGS-Studie festgestellt, dass bis zu 2 % der Kinder bereits im Vorschulalter eine ADHS-Diagnose bekommen hatten (Schlack, Hölling, Kurth & Huss, 2007).

Im Jugendalter rücken Probleme der Unaufmerksamkeit stärker in den Vordergrund. In der Schule kommt es zu Problemen, da die Jugendlichen leicht ablenkbar sind und schlecht im Unterricht folgen können, da sie schnell gelangweilt sind (Newcorn, 2008). Viele Jugendliche werden von ihren Lehrern als „Underachiever“ beschrieben, die ihre Fähigkeiten und ihr Potenzial nicht ausschöpfen können. Jugendliche mit ADHS haben Schwierigkeiten sich selbst beim Lernen zu organisieren, insbesondere längerfristige komplexe Projekte bereiten ihnen Schwierigkeiten. Sie schieben zu erledigende Aufgaben häufiger auf. Dieses prokrastinative Verhalten führt insbesondere auch dann zu Schwierigkeiten, wenn sie nach der Schule ein Studium aufnehmen (Wilens, Prince & Biederman, 2008). Weitere Probleme entstehen im Sozialbereich, so ist es für viele Jugendliche mit ADHS schwierig, feste Freundschaften und längerfristige partnerschaftliche Beziehungen zu knüpfen (Hoza et al., 2005).

Erwachsene mit ADHS sind oftmals ungeduldig und ruhelos (Newcorn, 2008). Viele Erwachsene mit ADHS leiden unter funktionalen Einschränkungen in unterschiedlichen Bereichen. Es ergeben sich Probleme in der Ausbildung und im Beruf, da ihr Arbeitsstil oftmals chaotisch ist und sie Schwierigkeiten haben zu planen und sich zu organisieren. Im sozialen Bereich führen Schwierigkeiten in der Impulskontrolle und Neigung zu starken emotionalen Reaktionen wie Wutausbrüchen oftmals zu Konflikten. Auch im Straßenverkehr fallen Erwachsene mit ADHS durch riskantes Verhalten auf und sind häufiger in Unfälle verwickelt. (Sobanski & Alm, 2004). Zudem spielen emotionale Probleme eine große Rolle, da sich die mangelhafte Selbstregulation auch im affektiven Bereich auswirkt (Retz, Stieglitz, Corbisiero, Retz-Junginger & Roesler, 2012). Das Störungsbild der ADHS weist bei einigen Erwachsenen Ähnlichkeiten mit einer Borderline-Persönlichkeitsstörung auf, da bei beiden Störungsbildern impulsives Verhalten im Vordergrund steht (Schmidt & Petermann, 2008).

Cherkasova, Sulla, Dalena, Pondé und Hechtman (2013) fassen auf Grundlage prospektiver Follow-up-Studien Risikofaktoren für einen ungünstigen Verlauf der ADHS zusammen. Im Vorschulalter sei die Schwere der Symptomatik und das Ausmaß der Beeinträchtigungen in den Exekutivfunktionen signifikante Prädiktoren für die Persistenz und die Schwere der ADHS. Dies gelte ebenso für das Schulalter, in dem zudem Komorbiditäten, insbesondere SOT und SSV, als zusätzliche Prädiktoren der ADHS-Persistenz und Funktionsbeeinträchtigung im Jugendalter und Erwachsenenalter dazukämen. In einer prospektiven Langzeitstudie über elf Jahre verglichen Biedermann, Petty, O'Connor, Hyder und Faraone (2012) 96 Mädchen mit ADHS und 91 Mädchen ohne

ADHS. Als Prädiktoren für die Persistenz der ADHS wurde das Vorliegen komorbider Erkrankungen gefunden, weiterhin Psychopathologie in der Familie und das Ausmaß der familiären und schulischen Funktionalität bei der ersten Erhebung.

Aufmerksamkeitsstörungen können damit zusammenfassend als eine schwerwiegende Entwicklungsgefährdung charakterisiert werden. Von einer spontanen Remission ist in den meisten Fällen nicht auszugehen. Es findet sowohl ein homo- als auch heterotypischer Verlauf statt, d.h. dass bei vielen Betroffenen auch in späteren Lebensphasen weiterhin ADHS-Symptome vorliegen, aber auch andere komorbide Erkrankungen und Probleme dazukommen können. Es besteht ein Risiko, dass spätere entwicklungsgemäße Anforderungen nicht bewältigt werden können, wenn Vorläuferfertigkeiten in der normgemäßen Altersphase nicht erlernt werden, z. B. im sozialen oder schulischen Bereich. Im Einklang mit Ihle und Esser (2002) sollte der Verlauf der hyperkinetischen Störung also nicht als ein entwicklungsbezogenes Übergangsphänomen verstanden werden, sondern als eine gravierende und gesundheitspolitisch sehr relevante Störung.

3.6 Erklärungsansätze

Es scheint keinen Einzelfaktor zu geben, der ADHS verursacht und aufrecht erhält, aber es konnten einige Risikofaktoren zur Ausbildung und Manifestation der Störung gefunden werden (Brassett-Harknett & Butler 2007; Millichap, 2008). Es wird allerdings darauf hingewiesen, dass zumeist nicht nachgewiesen werden kann, dass die Faktoren tatsächlich kausale Wirkung haben (Thapar, Cooper, Eyre & Langley, 2013). Es handelt sich vor allem um korrelative Zusammenhänge aus Vergleichen von Gruppen mit und ohne ADHS.

ADHS kann als eine Störung mit multifaktoriellen Verursachungs- und Aufrechterhaltungskomponenten verstanden werden, die sowohl biologisch, als auch psychologisch und sozial sind (Döpfner et al., 2008). Dabei kommen genetischen Grundlagen eine wichtige Rolle in der Entstehung der Störung bei. Sie werden in einem Diathese-Stress-Modell bzw. Vulnerabilitäts-Stress-Modell (Petermann et al., 2004) als genetische Prädispositionen verstanden, die das Risiko erhöhen, eine ADHS auszubilden. Prä-, peri- und postnatale Schädigungen des zentralen Nervensystems und ungünstige psychosoziale Stressoren können im Zusammenspiel mit den genetischen Prädispositionen zu einer Auslösung der Störung führen (Petermann & Toussaint, 2009). Die ADHS äußert sich auf der Verhaltensebene in Form der zentralen Merkmale Unaufmerksamkeit, Hyperaktivität und Impulsivität. Neurobiologische und neuropsychologische Strukturen und Prozesse spielen dabei eine vermittelnde Rolle. Ungünstige Reaktionen aus der Umgebung und die eigene reaktive Verarbeitung wirken sich auf den Verlauf und die Persistenz der Störung aus (Linderkamp, Hennig & Schramm, 2011).

Dieser Abschnitt beginnt mit einer Darstellung der konzeptuellen Grundlagen aktueller Kausalmodelle der ADHS. Dabei werden insbesondere die Zusammenhänge zwischen biologischen, kognitiven und verhaltensbezogenen Erklärungsebenen thematisiert. Wei-

terhin werden das Konzept der Endophänotypen und die wichtigsten neuropsychologischen Erklärungsmodelle vorgestellt, welche die Grundlage für das Verständnis und die Behandlung der Störung bilden. Es folgt die Vorstellung wichtiger die Störung verursachender und aufrecht erhaltender Bedingungen. Als kausale Einflussgrößen werden genetische Auffälligkeiten, prä-, peri- und postnatale Komplikationen und neurobiologische Aspekte diskutiert. Den Abschluss bildet die Diskussion wichtiger Faktoren, welche die Störung aufrechterhalten bzw. verschärfen. Dies sind soziale Interaktionen innerhalb der Familie, mit Lehrern und mit Gleichaltrigen.

3.6.1 Konzeptuelle Grundlagen

Viele aktuelle Kausalmodelle zur Erklärung von ADHS betonen neuropsychologische Defizite der Betroffenen, die auf neurobiologischen Auffälligkeiten beruhen, die wiederum teilweise genetisch prädispositioniert sind. Wie die Zusammenhänge zwischen genetischer, neurobiologischer und neuropsychologischer Ebene konkret aussehen, ist nicht abschließend geklärt. Ob, wann und in welcher Form die Störung auftritt, hängt von auslösenden und aufrechterhaltenden Einflüssen ab, die zumeist in der Umwelt bzw. Umgebung verortet werden (Banaschewski, Roessner, Uebel & Rothenberger, 2004). Ein allgemein gültiges Erklärungsmodell, das für alle Personen mit ADHS zutrifft, liegt allerdings bislang nicht vor. Es ist von einer kausalen Heterogenität der Störung auszugehen, bei der unterschiedliche Entwicklungspfade vorliegen (Nigg, Willcutt, Doyle & Sonuga-Barke, 2005). Es werden im Weiteren einige aktuelle Erklärungsansätze vorgestellt.

Die meisten Kausalmodelle der ADHS gehen von einer grundlegenden genetischen Ursache bzw. Prädisposition aus, die sich vermittelt über kognitive Prozesse, auf die auch Einflüsse der Umgebung wirken, auf das Verhalten auswirken. Abbildung 10 zeigt ein vereinfachtes hypothetisches Kausalmodell (Coghill, Nigg, Rothenberger, Sonuga-Barke & Tannock, 2005). Es besteht aus den drei intrapersonalen Analyseebenen biologisch, kognitiv und verhaltensbezogen und der Ebene umweltbezogener Einflüsse, die sich auf jede der anderen Ebenen auswirken kann. Der ursächliche genetische Faktor A_1 bewirkt auf der biologischen Ebene eine Abweichung der biologischen Struktur Br_1 , die wiederum dazu führt, dass sich ein kognitiver Prozess verändert wird bzw. normativ abweichend entwickelt C_1 . Diese kognitive Veränderung wird wiederum als ursächlich für das abweichende Verhalten S_1 gesehen. Umweltfaktoren wie hier E_1 können sich auf alle der anderen drei Ebenen auswirken

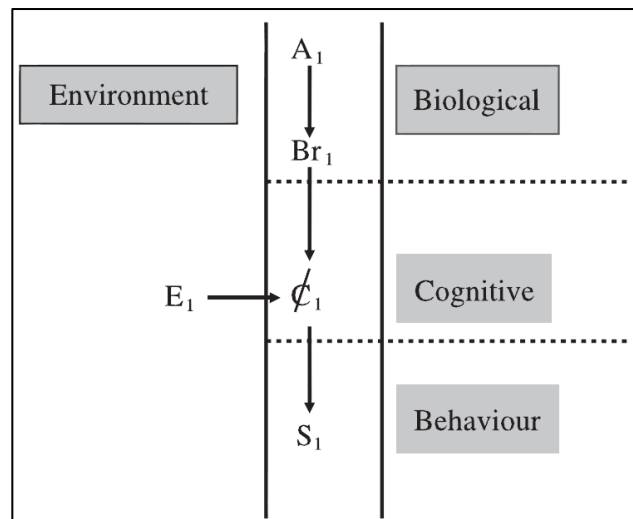


Abbildung 10: Vereinfachtes hypothetisches Kausalmodell (aus Coghill et al., 2005, S. 106)

Morton und Frith (1995) führten dieses Schema als „developmental causal modelling framework“ ein. In solchen Kausalmodellen werden Elemente derselben oder anderer Ebenen zu Kausalketten zusammengestellt. Morton und Frith (1995) liefern weiterhin eine Anwendung ihres Modells zur Erklärung hyperkinetischer Symptome. Sie schlagen ein v-förmiges Modell vor, bei dem unterschiedliche biologische, kognitive und Umweltfaktoren in der hyperaktiven Symptomatik münden (Abbildung 11).

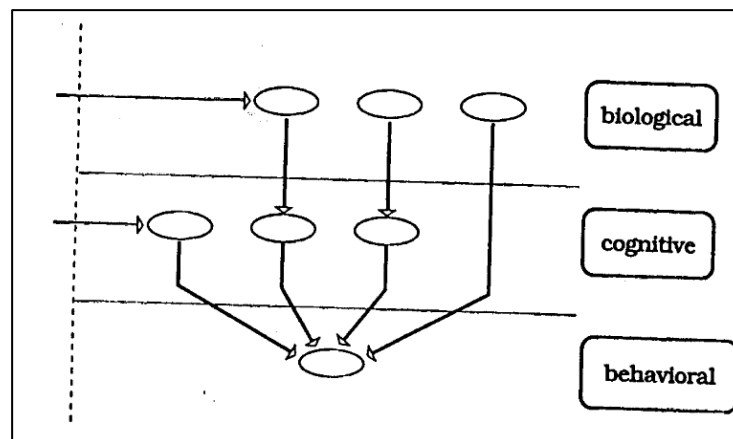


Abbildung 11: V-förmiges Kausalmodell, bei dem unterschiedliche kausale Pfade zur phänotypischen Hyperaktivität auf der Verhaltensebene führen (aus Morton & Frith, 1995, S. 372)

Der Pfad ganz rechts deutet eine direkte Verursachung der Hyperaktivität durch biologische Faktoren an. So werden beispielsweise strukturelle und funktionale Auffälligkeiten des Kleinhirns bei ADHS-Betroffenen als neurobiologisches Korrelat der Hyperaktivität diskutiert (O'Halloran, Kinsella & Storey, 2011). Dagegen fungieren bei den mittleren Pfaden kognitive Faktoren als Mediatoren. Diese können beispielsweise beeinträchtigte Arbeitsgedächtnisfunktionen sein, die von gehirnstrukturellen Abweichungen verursacht werden und sich auf die sichtbare Verhaltensebene auswirken (Burgess et al., 2010). Die von links kommenden Pfeile in der Abbildung kennzeichnen Umweltfaktoren.

ren, die sich sowohl auf die biologische als auch auf die kognitive Ebene auswirken können. Ein Umwelteinfluss auf die biologische Ebene ist beispielsweise die Wirkung von Nahrungsmitteln auf stoffwechselbezogene Prozesse wie sie auch im Zuge der diätetischen und komplementärmedizinischen Behandlung von ADHS diskutiert werden (Sarris, Kean, Schweitzer & Lake, 2011). Ein Umwelteinfluss auf die kognitive Ebene ohne Beteiligung der biologischen Ebene wäre die Veränderung einer schulbezogenen Einstellung (Wu & Gau, 2013). Beispielsweise könnten Rückmeldungen des Lehrers die Erwartung des Schülers dahingehend prägen, dass er selbst davon ausgeht, dass Bemühungen die eigene Hyperaktivität zu unterdrücken keine Aussicht auf Erfolg haben, da er vom Charakter her ein „Zappelphilipp“ sei. Alle Pfade resultieren in diesem Modell in demselben Verhaltenskomplex, der Hyperaktivität, so dass sich eine V-Form ergibt. Zur Erklärung von Variabilität im Störungsbild werden weitere Pfade im Modell eingeführt und auch protektive Faktoren, welche die Auswirkung pathogener Faktoren verhindern können (Morton & Frith, 1995). Die Autoren verweisen in diesem Zusammenhang auf die Nähe ihres kausalen Modells zu den psychosozialen Pfadmodellen von Rutter und Kollegen (1989). Der Unterschied bestehe darin, dass die Pfadmodelle im Wesentlichen Trends über die Entwicklung von Gruppen von Individuen über die Zeit ausdrücken. Diese beruhen auf Korrelationen und machen daher keine Aussagen über die Kausalität der Zusammenhänge. Im Kausalmodell nach Morton und Frith (1995) liegen die ursächlichen Faktoren der Hyperaktivität also auf der biologischen Ebene, die zumeist vermittelt über kognitive Prozesse auf die Verhaltensebene wirken. Ebenfalls ursächlichen Charakter haben Umwelteinflüsse, die auf biologische und kognitive Prozesse wirken. Es handelt sich aber um Modellvorstellungen – der Beweis der Kausalität ist empirisch nicht abschließend überprüft bzw. überprüfbar.

In Einklang mit den vorgestellten Kausalmodellen kann auch die Modellvorstellung sogenannter Endophänotypen gebracht werden (Castellanos & Tannock, 2002; Schachar & Crosbie, 2008). Diese stellen einen neueren Ansatz in der psychiatrischen Forschung dar, die sich als ein Ziel die Erforschung genetischer Ursachen psychiatrischer Erkrankungen gesetzt hat (Zobel & Maier, 2004). Psychische Störungen werden demnach als extreme Ausprägungen von zugrunde liegenden kontinuierlichen Dimensionen psychischer Konditionen verstanden. Die sichtbaren Aspekte der Störungen werden als Phänotyp bezeichnet. Ein solcher wäre die über Verhaltensmerkmale definierte ADHS. Es wird angenommen, dass den Phänotypen bestimmte genetische Auffälligkeiten zugrunde liegen, die als Genotypen bezeichnet werden. Diese sind nur schwer exakt zu bestimmen bzw. direkt mit Phänotypen in Verbindung zu bringen, da Phänotypen oftmals eine heterogene Ätiologie aufweisen, also unterschiedlich verursacht werden können. Ein vermittelndes Konstrukt stellen sogenannte Endophänotypen dar, die eine Verbindung von Geno- und Phänotypen bilden. Die Bandbreite der Endophänotypen ist da dabei sehr breit, so werden neurophysiologische, biochemische, endokrinologische, neuroanatomische, kognitive und neuropsychologische Auffälligkeiten herangezogen (Gottesman & Gould, 2003). Im Zusammenhang mit ADHS werden erklärend insbesondere neurobiologische und neuropsychologische Endophänotypen herangezogen, so

dass sich eine Ähnlichkeit mit dem Modell nach Morten und Frith (1995) zeigt (siehe Abbildung 12).

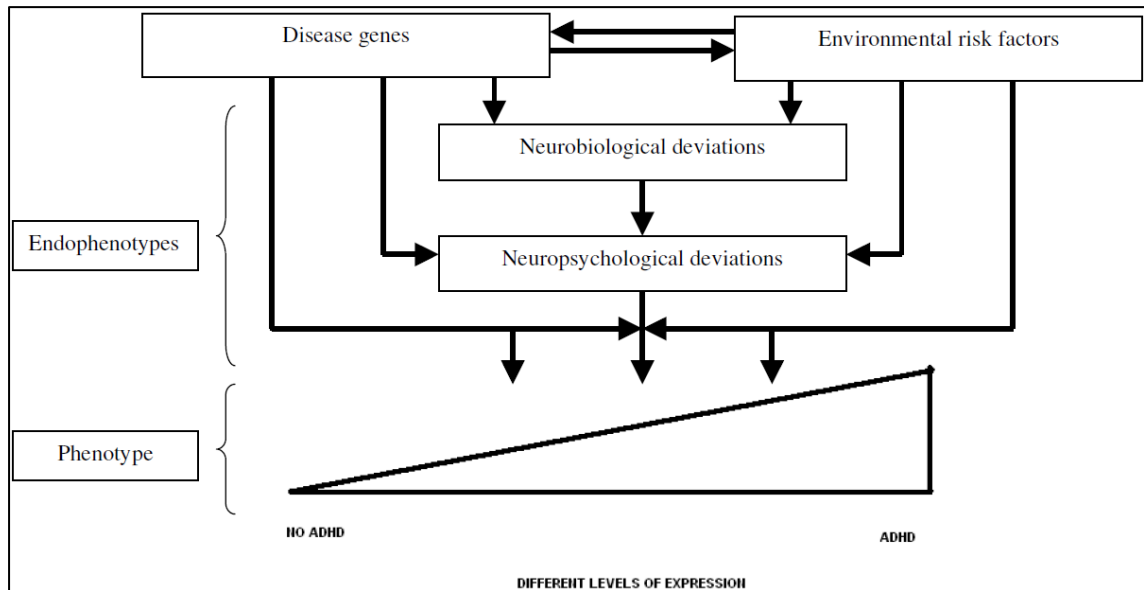


Abbildung 12: Der Zusammenhang von genetischen und umweltbezogenen Risikofaktoren, Endophänotypen und Phänotypen bei ADHS (aus Rommelse et al., 2008, S. 2)

Endophänotypen werden hier als neurokognitive Abweichungen verstanden, auf deren Grundlage das Risiko, eine Störung im Phänotyp auszubilden, erhöht ist. Neurokognitive Auffälligkeiten beeinflussen mentale Prozesse, die nicht direkt beobachtbar sind, aber experimentell untersucht werden können. Erst die Interaktion mehrerer Endophänotypen bestimmt das sichtbare Verhalten im Phänotyp. Der Vorteil der Erforschung von Endophänotypen im Gegensatz zu Phänotypen liegt darin, dass erstere genetisch weniger komplex seien. Damit könnten solche Gene, die an der Entstehung einer Störung beteiligt sind, besser identifiziert werden. Ferner seien auch unterschiedliche Pfade in der Entwicklung derselben Störung untersuchbar (Rommelse et al., 2008).

3.6.2 Neuropsychologische Erklärungsmodelle

Ein einflussreiches neuropsychologisches Modell zur Erklärung der ADHS, das auch mit den Endophänotypen in Übereinstimmung gebracht werden kann, stammt von Barkley (1997). Es postuliert eine beeinträchtigte Verhaltensinhibition als das Kernelement der Störung. Solche und verwandte neuropsychologische Defizite der exekutiven Selbststeuerung werden als potentielle Endophänotypen bei der ADHS bezeichnet (Doyle et al., 2005). Barkley subsumiert unter der Verhaltensinhibition die Fähigkeit zur Unterdrückung eines dominanten Handlungsimpulses, das Unterbrechen einer aktuell ausgeführten Handlung und das Ausblenden konkurrierender Reize. Zur Verhaltensinhibition gehören nach Barkley (1997) vier Komponenten, die bei ADHS beeinträchtigt seien: das Arbeitsgedächtnis, die Selbstregulation von Gefühlen, Motivation und Aktivierung, die Internalisierung von Sprache und die Rekonstitution, womit die Analy-

se, Synthese und Wiederherstellung von Handlungssequenzen gemeint ist. Auf Grundlage dieser Beeinträchtigungen sei die ADHS-Symptomatik erklärbar.

Auch im kognitiv-energetischen Modell nach Sanders (1983), das von Sergeant (2000, 2005) auf die ADHS bezogen und konkretisiert wurde, wird eine defizitäre Selbstregulation als zentrales Element der Störung postuliert. Wichtig zur angemessenen Verhaltensreaktion auf einen Reiz müsse nach den Modellannahmen eine Grundaktivierung („arousal“) vorliegen, damit Reize erkannt und kodiert werden können, weiterhin müssen Energien zur Verarbeitung der Reize aufgewendet werden („effort“) und schließlich sind Energien zur Aktivierung einer motorischen Reaktion erforderlich (siehe Abbildung 13). Die zentrale Steuerung dieser Funktionen sei bei Menschen mit ADHS nicht optimal.

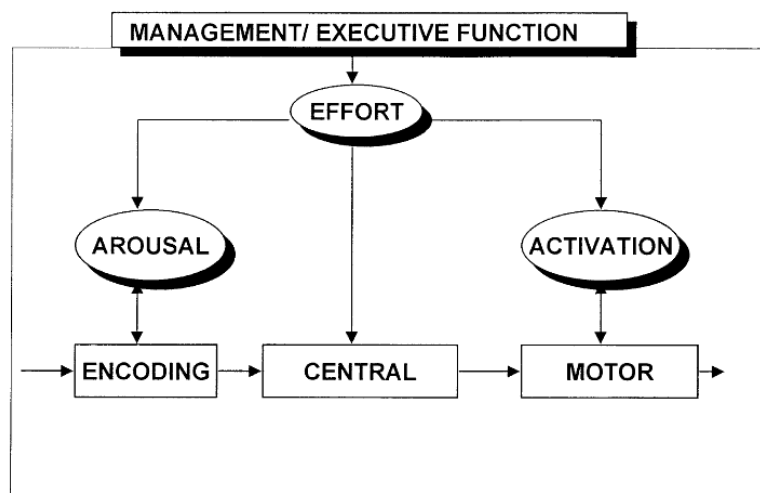


Abbildung 13: Das kognitiv-energetische Modell (aus Sergeant, 2000, S. 8)

Kinder und Jugendliche mit ADHS machen in kognitiven Aufmerksamkeitstests zu meist mehr Fehler und haben längere Reaktionszeiten (Banaschewski et al., 2004). Schwache Leistungen treten insbesondere bei wenig stimulierenden Aufgaben auf. Diese Kontextabhängigkeit wird auf eine geringe kortikale Stimulierung und bzw. oder motivationale Defizite zurückgeführt (Banaschewski et al., 2004). Zentall (1975) geht in ihrer Theorie der optimalen Stimulation davon aus, dass hyperaktive Kinder ein geringeres Arousalniveau aufweisen und sie versuchen, dieses zu erhöhen. Das hyperaktive Verhalten kann damit als eine Form der Selbststimulation gesehen werden (Zentall & Zentall, 1983). Diese Vorstellung ist gut mit dem kognitiv-energetischen Modell in Einklang zu bringen. Auch in diesem Modell wird ein geringeres Arousal bei ADHS vermutet, welches über eine Anstrengungsfunktion reguliert wird, die ebenfalls die zentrale Verarbeitung und die motorische Aktivierung steuert. Diese Regulierung, an denen die exekutiven Funktionen beteiligt sind, ist bei ADHS beeinträchtigt. Dies kann auch als Selbstregulationsstörung bezeichnet werden (Gerjets, Graw, Heise, Westermann & Rothenberger, 2002). Der Ansatz ADHS als Selbstregulationsdefizit zu verstehen wurde u. a. von Douglas (1972) eingeführt, die in diesem Zusammenhang insbesondere Probleme der Daueraufmerksamkeit und Reaktionsinhibition betont. Mangelnde Inhibitions-

kontrolle, die im Modell von Barkley (1997) als zentral angesehen wird, ist eine der exekutiven Funktionen, in denen die Beeinträchtigungen der Menschen mit ADHS im Vergleich zu Kontrollpersonen am größten sind. Trotzdem sind sie nicht der einzige Faktor, mit dem alle mit ADHS verbundenen Auffälligkeiten erklärt werden können und der für alle Betroffenen zutrifft (Banaschewski et al., 2004).

In anderen Modellen werden zusätzlich motivationale Faktoren zur Erklärung der ADHS herangezogen. Verhaltensbezogene Auffälligkeiten in der Unterdrückung von Reizen werden darin nicht ausschließlich als Kompetenzdefizite verstanden, sondern auch als Mangel an Motivation eine Verhaltensweise auszuführen. Viel zitiert ist das „Dual-pathway model“ von Sonuga-Barke (2002, 2003). Es konkretisiert insbesondere die ablaufenden kognitiven Prozesse und vereint zwei Erklärungspfade in einem Modell. Ein Pfad konzeptualisiert ADHS als eine Störung der Exekutivfunktionen, die sich insbesondere in einer gestörten Reizinhibition äußert. Dem zugrunde liegend vermutet Sonuga-Barke neurobiologische Abweichungen in meso-kortikalen Kreisläufen. Die gestörte Reizinhibition führe gemäß Modell zu einer Verhaltensdysregulation und trage darüber vermittelt zur ADHS-Symptomatik bei. Gleichsam zeige sich eine kognitive Dysregulation, die dazu führe, dass ADHS-Betroffene weniger Anstrengung bei der Bewältigung von Aufgaben aufwenden. So könnten sie sich nicht lange auf eine Arbeit konzentrieren und erzielen oftmals auch qualitativ schlechtere Ergebnisse. Ein zweiter Pfad, in dem ADHS als ein motivationaler Stil konzeptualisiert wird, ergänzt das Modell. Auf Grundlage von Abweichungen in meso- limbischen Prozessen, die mit der Verarbeitung von Belohnungen in Verbindung gebracht werden, entstünden Probleme, wenn Belohnungen verzögert dargeboten werden. Demgegenüber stünden basierend auf kulturellen Normen und Wertvorstellungen erhöhte Anforderungen, wenn Belohnungen erst verzögert erlangt werden können. Es resultiere eine generalisierte Verzögerungsaersion der Betroffenen. So hätten Kinder mit ADHS eine starke Präferenz für unmittelbare Belohnungen, auch wenn eine noch größere Belohnung nach Verzögerung in Aussicht gestellt werde. Diese wiederum erkläre einen Teil der ADHS-Symptomatik und vermittelt darüber auch einen Teil der geringeren Anstrengungsaufwendung bei der Aufgabenbearbeitung (siehe Abbildung 14).

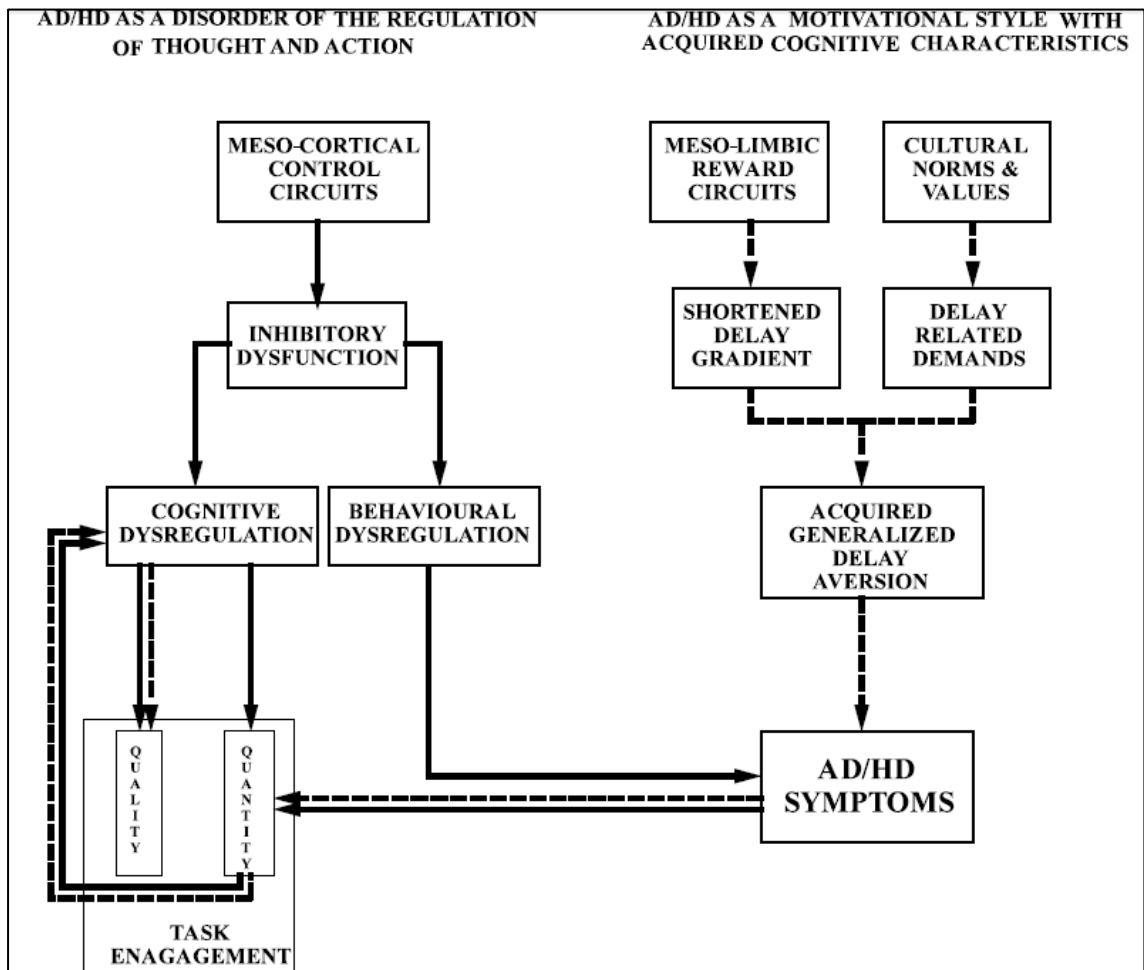


Abbildung 14: Das Dual Pathway Modell der ADHS (aus Sonuga-Barke, 2002, S. 32).

Ein ähnliches konzeptuelles Modell der ADHS, welches die Wirkungen von Interventionen auf den unterschiedlichen Ebenen kennzeichnet, wurde von Rapport, Chung, Shore und Isaacs (2001) vorgeschlagen (Abbildung 15).

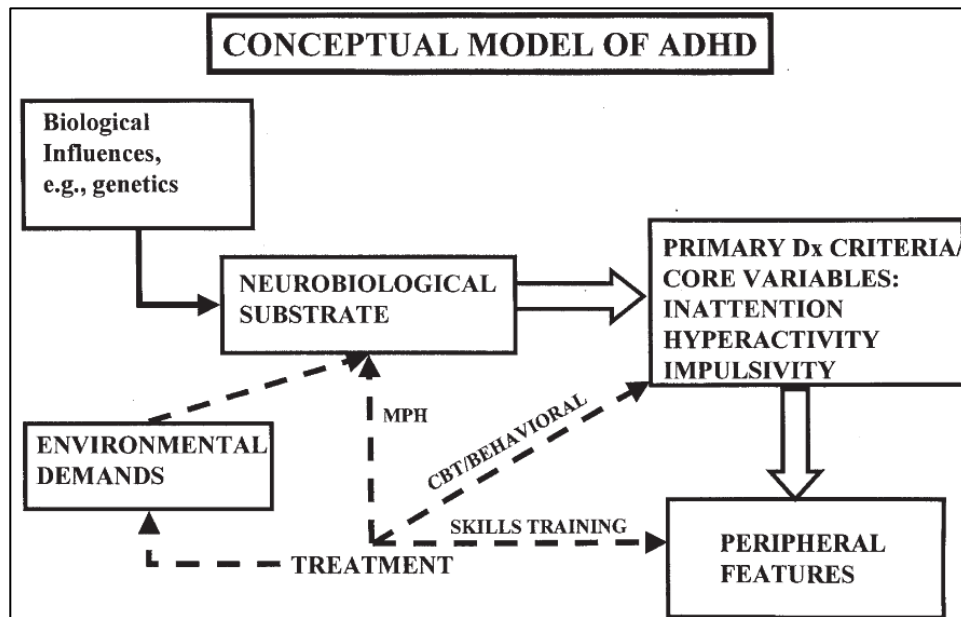


Abbildung 15: Kausalmodell der ADHS und Wirkweise von Interventionen (aus Rapport et al., 2001, S. 49)

Auch hier wird eine grundlegende biologische Ursache der Störung angenommen, die sich insbesondere als genetischer Einfluss und pränatalen Schädigungen äußern. Diese wirken sich auf ein individuell unterschiedlich ausgeprägtes neurobiologisches „Substrat“ aus, das sich u. a. in funktionalen Besonderheiten des Neurotransmittersystems (Dopamin und Noradrenalin) äußert. Diese Besonderheiten führen zur Herausbildung der zentralen psychologischen (kognitiven und verhaltensbezogenen) Problematiken, Unaufmerksamkeit, Hyperaktivität und Impulsivität. Auf dieses Substrat wirken sich auch die Anforderungen aus der Umgebung aus. Weitere Auffälligkeiten wie Underachievement werden als Begleit- bzw. Folgeerscheinungen konzeptualisiert; ebenso mangelhafte soziale Fähigkeiten, geringe Frustrationstoleranz und angespannte familiäre Beziehungen. Diese werden angenommen als verursacht von zentralen Verhaltens- und kognitiven Einflüssen. Rapport und Kollegen schlussfolgern, dass also vier Ziele von Interventionen anvisiert werden können, je nachdem auf welcher Ebene eine Veränderung erreicht werden soll: Jeweils bezogen auf das neurobiologische „Substrat“, die zentralen bzw. Kernprobleme, periphere Schwierigkeiten oder umweltbezogene.

3.6.3 Genetik

Es findet sich weitgehende Zustimmung, dass der ADHS eine starke genetische Grundlage unterliegt, wobei bislang keine störungsspezifischen Chromosomenstrukturen eindeutig identifiziert werden konnten (Barkley, 2006; Thapar, O'Donovan & Owen, 2005). Voeller (2004) schätzt, dass in ca. 80 % der Fälle eine genetische Ursache für die Entstehung der ADHS vorliegt, wobei unterschiedliche Gene beteiligt seien. In den übrigen 20 % der Fälle entstehe ADHS als Resultat einer erworbenen Hirnverletzung. In einigen Fällen liege auch eine Kombination aus genetisch verursachter und erworbener ADHS vor. Die Ausprägung beider Störungshintergründe sei auf der Symptomebene

identisch und bislang nicht unterscheidbar. Beide reagierten gleichsam gut auf eine Behandlung mit Stimulantien. Auf der Grundlage von Zwillingsstudien wird die Heritabilität von ADHS auf 60 bis 90 % geschätzt (Waldman & Gizer, 2006). In einer Metaanalyse von Faraone und Kollegen (2005) wurden sieben Gene identifiziert, die in zumindest drei Studien einen Zusammenhang mit dem Auftreten einer ADHS aufwiesen. Diese waren zumeist Dopaminregulatorgene, so wie auch in einer Metaanalyse von Li, Sham, Owen und He (2006), die alle bis dato publizierten europäischen und asiatischen Studien einschloss. Aber es wird auch die Relevanz von Genen diskutiert, die mit dem noradrenergen und serotonergen Neurotransmittersystem in Verbindung gebracht werden (Waldman & Gizer, 2006). Wie bereits dargelegt werden stärkere Zusammenhänge gefunden, wenn die Zusammenhänge der genetischen Grundlage mit Endophänotypen untersucht werden, die wiederum mit dem bzw. den Phänotypen der ADHS in Verbindung gebracht werden (Kebir & Joober, 2011).

3.6.4 Prä-, peri- und postnatale Komplikationen

In einem systematischen Review über den Zeitraum 1973 bis 2002 wurden 24 Studien zum Zusammenhang von mütterlichem Nikotinkonsum während der Schwangerschaft und ADHS bei den Kindern gefunden (Linnet et al., 2003). Alle Studien berichteten einen signifikanten Zusammenhang. Es wird geschätzt, dass das Verhältnis der Wahrscheinlichkeiten (Odds Ratio) vom Auftreten der ADHS bei mütterlichem Nikotinkonsum vs. kein Konsum 2.39 beträgt und es einen Zusammenhang von Dosis und Wirkung gibt (Thapar, Cooper, Jefferies & Stergiakouli, 2012). So verglichen Thakur und Kollegen (2013) 436 Kinder mit ADHS mit und ohne mütterlichem Nikotinkonsum während der Schwangerschaft. Dabei wurden die Variablen sozioökonomischer Status, Ethnizität, Alter der Mutter bei Geburt des Kindes und mütterlicher Alkoholkonsum während der Schwangerschaft kontrolliert. Kinder, deren Mütter während der Schwangerschaft geraucht hatten, zeigten schwerwiegendere Verhaltensprobleme mit mehr externalisierenden Verhaltensweisen, einen geringeren verbalen IQ und ein Sluggish Cognitive Tempo im Continuous Performance Test (CPT). In linearen Regressionsanalysen zeigte sich ein Dosis-Wirkungs-Zusammenhang zwischen durchschnittlicher Anzahl gerauchter Zigaretten pro Tag und dem verbalen IQ, Auslassungsfehlern im CPT und anderen Variablen. Je mehr Zigaretten in der Schwangerschaft geraucht worden waren, desto ungünstiger waren die Ergebnisse des Kindes. Die Autoren schlussfolgern, dass mütterlicher Nikotinkonsum in der Schwangerschaft zu einer schwereren Form der ADHS führe.

Ob es sich allerdings um eine direkte Wirkung des Nikotins auf den Embryo handelt, konnte nicht bewiesen werden. Langley, Heron, Smith und Thapar (2012) untersuchten die Effekte mütterlichen und väterlichen Nikotinkonsums. Wenn ein reiner intrauteriner Effekt vorläge, dürfte das väterliche Rauchen nicht mit dem Risiko für ADHS verbunden sein. Anhand einer Kohorte von 8.324 Kindern konnte gezeigt werden, dass das Vorliegen von ADHS-Symptomen sowohl mit mütterlichem als auch väterlichem Nikotinkonsum während der Schwangerschaft in Verbindung stand. Auch in Abwesenheit

von mütterlichem Rauchen blieben die Verbindungen bestehen, auch ohne den Einfluss von passivem Mitrauchen. Dieser Befund legt nahe, dass der Zusammenhang von Nikotinkonsum in der Schwangerschaft und kindlichem ADHS nicht auf einen kausalen intrauterinen Effekt zurückgeht, sondern vermutlich eher mit genetischen oder häuslichen Faktoren zusammenhängt. Auch Knopik weist darauf hin, dass aktuelle Studien, die Zusammenhänge finden, oft Konfundierungen mit Umwelt- und genetischen Faktoren vorliegen, so dass die Kausalität des Zusammenhangs nicht eindeutig nachgewiesen werden kann (Knopik, 2009). Einige neuere Studien finden keinen Zusammenhang zwischen Nikotinkonsum während der Schwangerschaft und ADHS, z. B. Ball und Kollegen (2010) in einer prospektiven Studie mit einer Größe von $N = 2024$.

Weniger konsistent sind die Ergebnisse für den Zusammenhang von mütterlichem Alkoholkonsum während der Schwangerschaft und ADHS. Linnet et al. (2003) fanden für den von ihnen untersuchten Zeitraum acht Studien mit widersprüchlichen Ergebnissen. Mick und Kollegen (2002) verglichen 280 Kinder mit ADHS mit 242 Kindern ohne ADHS. Die Kinder mit ADHS waren 2.5mal häufiger einem Alkoholkonsum der Mutter während der Schwangerschaft ausgesetzt gewesen. Dieser Zusammenhang blieb auch dann bestehen, wenn die elterliche Psychopathologie und andere Variablen kontrolliert wurden. Schwerer Alkoholmissbrauch in der Schwangerschaft kann zum Auftreten einer Alkohol-Embryopathie führen, die ebenfalls Symptome der Unaufmerksamkeit und Hyperaktivität mit sich bringt (Thapar et al., 2012). Die Effekte weniger extremen Konsums und ADHS sind hingegen weniger klar. Alkoholkonsum scheint ein Risikofaktor beim Ausbilden einer ADHS zu sein, dieser Zusammenhang ist allerdings weniger deutlich als bei Nikotin und insgesamt nur korrelativ und nicht bewiesenermaßen kausal (Thapar et al., 2013). Es wird auch vermutet, dass andere illegale Drogen Risikofaktoren für das Entstehen einer ADHS sind. Die Effekte konnten allerdings nicht trennscharf untersucht werden, da zum Konsum der illegalen Drogen oftmals zusätzlich Nikotin- oder Alkoholkonsum vorlag (Linnet et al., 2003).

Auch niedriges Geburtsgewicht wird oft in Zusammenhang mit ADHS festgestellt (Thapar et al., 2013). Mick, Biederman, Prince, Fischer und Faraone (2002) verglichen 252 Kinder mit ADHD mit 231 Kindern ohne ADHD. Die Kinder mit ADHD waren dreimal häufiger mit einem niedrigen Gewicht geboren worden als die Kinder ohne ADHD. Die Autoren schließen, dass niedriges Geburtsgewicht ein eigenständiger Risikofaktor für ADHD sei. Halmoy, Klungsoyr, Skjaerven und Haavik (2012) untersuchten 2323 Erwachsene, die sich wegen ADHD in Behandlung begeben hatten, anhand des norwegischen medizinischen Geburtsregisters. Sie verglichen die Daten der ADHD-Gruppe mit den übrigen Erwachsenen ohne ADHD, die in denselben Jahren geboren worden waren. Es zeigte sich, dass in der Gruppe mit ADHD sowohl ein niedrigeres Geburtsgewicht als auch frühe Geburt und niedrigere Apgar-Scores, die den Gesundheitszustand von Neugeborenen bewerten, vorlagen als in der Vergleichsgruppe. Die Autoren schlussfolgern, dass diese Faktoren Risikofaktoren für das Persistieren der ADHD bis zum Alter von 40 Jahren nach der Geburt darstellen. Strang-Karlsson et al. (2008) vermuten, dass nicht das geringe Geburtsgewicht an sich, sondern eine intraute-

rine Wachstumsverzögerung mit ADHS bezogenen Problemen in Verbindung stehe. Die Wachstumsverzögerung bestimmten die Autoren über die relative Größe des Kindes bei Geburt in Bezug zum Gestationsalter, also der Dauer der Schwangerschaft seit dem ersten Tag der letzten normalen Regelblutung der Mutter. In ihrer Stichprobe mit insgesamt 162 Erwachsenen mit sehr geringem Geburtsgewicht und 172 Kontrollpersonen hatten nur diejenigen Erwachsenen, die zu klein für ihr Gestationsalter gewesen waren, schlechtere Ergebnisse bei Tests der Exekutivfunktionen und berichteten mehr Probleme in ihrer Emotionsregulation. Die Erwachsenen mit sehr geringem Geburtsgewicht, die aber normal groß für ihr Gestationsalter gewesen waren, erreichten vergleichbare Werte wie die Kontrollpersonen.

Als weiterer Risikofaktor für das Ausbilden einer ADHS wird Stress der Mutter während der Schwangerschaft diskutiert (Thapar et al., 2012). Durch den Stress der Mutter erhöht sich nicht nur der Cortisol- und Adrenocorticotropinspiegel im Blutplasma der Mutter, sondern auch im Plasma des Fötus (Weinstock, 2008). Dies kann zu dauerhaften Veränderungen des Stoffwechsels führen, die zur Ausbildung einer ADHS-Symptomatik beitragen. Es ist aber auch für diesen Risikofaktor nicht nachgewiesen, dass es sich um eine kausale Verursachung der Symptome durch den mütterlichen Stress handelt. In diesem Zusammenhang werden ebenso Interaktionen von genetischen und umweltbezogenen Faktoren diskutiert (Rice et al., 2010).

Auch die Belastung mit Toxinen scheint ein Risikofaktor für ADHS zu sein. Es wurden unter anderem die Effekte von Bleivergiftungen und Pestiziden untersucht und als Risikofaktoren identifiziert (Thapar et al., 2012). Aber auch hier liegt kein Nachweis der kausalen Verursachung vor. Es werden im Zusammenhang mit ADHS einige weitere Umgebungsfaktoren gefunden, die besonders häufig gemeinsam auftreten, momentan sind diese aber nur als Korrelate anzusehen, deren ursächlicher Zusammenhang nicht nachgewiesen ist (Thapar et al., 2013).

3.6.5 Neurobiologie

Da es sich bei ADHS um eine Störung der Selbstregulation handelt, werden Beeinträchtigungen des frontalen subkortikalen Systems vermutet (Voeller, 2004). Bezogen auf die Gehirnanatomie zeigt sich, dass das Volumen des zerebralen Kortex bei Personen mit ADHS, insbesondere das Volumen des Frontallappens, geringer zu sein scheint (Castellanos & Tannock, 2002). Weiterhin sind einige Bereiche der Basalganglien kleiner, vor allem der Nucleus caudatus. Die Größenunterschiede werden mit steigendem Alter geringer. Areale der rechten Gehirnhälfte sind dabei mehr betroffen als solche der linken Hemisphäre. Auch das Kleinhirn scheint kleiner zu sein (Krain & Castellanos, 2006). Seidman, Valera und Makris (2005) kommen zu dem Schluss, dass mittlerweile viele Hinweise darauf bestünden, dass das Gehirnvolumen insgesamt und spezifischer Areale bei Gruppen von Personen mit ADHS signifikant geringer ist als bei Gruppen von Personen ohne ADHS. Die beeinträchtigten Areale seien insbesondere mit der Modulation von Aufmerksamkeit und Impulsivität assoziiert. Der Frontallappen sei mit den

exekutiven Funktionen verbunden, ebenso stärker die rechte als die linke Gehirnhälfte und auch das Kleinhirn über seine reziproken Verbindungen zum präfrontalen Kortex. Neben den anatomischen Volumenunterschieden scheint ein bestimmtes System beeinträchtigt zu sein, das mit den ADHS-Symptomen in Verbindung gebracht werden kann. Nach Durston (2003) sind dabei vermutlich die frontalen Hirnregionen, die Basalganglien und das Kleinhirn beteiligt. Es scheint aber keine völlig abweichende Entwicklung vorzuliegen, sondern eine Verlangsamung der Reifungsprozesse. So verglichen Shaw und Kollegen (2007) 223 Kinder mit ADHS und 223 sich typisch entwickelnde Kinder. Sie stellten eine gleichförmige Entwicklung des Gehirns in beiden Gruppen fest mit einer Verlangsamung der Reifungsprozess in der ADHS-Gruppe, insbesondere in präfrontalen Arealen des Cortex.

Auch Methoden der funktionalen Bildgebung zeigen Unterschiede im Vergleich von Personen mit und ohne ADHS (Voeller, 2004). Insbesondere wurde eine Unteraktivierung eines Netzwerkes festgestellt, das mit Aufmerksamkeitsfunktionen in Verbindung gebracht wird (Bush, 2011). An diesem Netzwerk seien sowohl frontale als auch parietale und cinguläre Areale beteiligt, so wurden beispielsweise bei Kindern, die eine Inhibitionsaufgabe bearbeiteten, eine veränderte Gehirnaktivität im Frontalbereich, im Nucleus caudatus und dem anterioren Cingulus festgestellt (Bush et al., 1999). Die Erregungsmuster bei Kindern seien diffuser als die von Erwachsenen (Durston, 2003), beide zeigten aber eine geringere Aktivierung von Bereichen, die mit Emotion und Gedächtnis in Verbindung gebracht werden. Im EEG zeigten Kinder mit ADHS häufiger Theta-Wellen-Aktivität als Kontrollkinder, die mit einer geringeren Altermtness und Untererregung in Verbindung gebracht werden. Dieses Muster war weniger stark für Kinder des vorwiegend unaufmerksamen Subtyps (Voeller, 2004).

Valera et al. (2007) kommen allerdings zu dem Schluss, dass noch kein Marker gefunden werden konnte, der zuverlässig zwischen Probanden mit und ohne ADHS unterscheiden könne. Die Autoren stellten in einer Metaanalyse 21 Einzelstudien über bildgebende Verfahren zusammen und kommen zu dem Ergebnis, dass die hirnstrukturellen Unterschiede in den Gruppen mit und ohne ADHS nur mittelgroß seien und zudem nicht auf einzelne Regionen beschränkt.

Auch im Neurotransmitterstoffwechsel finden sich Besonderheiten bei Personen mit ADHS. Viel spricht für einen durch mehrere Gene verursachten Dopaminmangel im synaptischen Spalt, der durch eine Behandlung mit Stimulantien ausgeglichen werden kann (Petermann & Toussaint, 2009). Dopamin wird eine zentrale Rolle bei der Steuerung motorischer Aktivitäten und der Verarbeitung von Belohnungen zugeschrieben, so dass Zusammenhänge mit ADHS plausibel erscheinen (Brassett-Harknett & Butler, 2007). Es konnte gezeigt werden, dass Dopamin eine Rolle bei der Regulation von Lernprozessen zukommt, sowohl in der Aufrechterhaltung gelernter Verhaltensweisen und auch in der Motivation (Voeller, 2004). Weiterhin ist Dopamin an Gedächtnisprozessen beteiligt, insbesondere dem Arbeitsgedächtnis (Voeller, 2004). Es wird aber auch die Beteiligung von Noradrenalin diskutiert, das an Prozessen der Aufrechterhaltung

von Aufmerksamkeit („alertness“) beteiligt ist. Noradrenerge Neurone werden von neuen und wichtigen Reizen aktiviert und sind inaktiv im Schlaf. Die Behandlung mit Stimulanzien bei ADHS ist vermutlich deshalb so wirksam, da die Medikation die Wiederaufnahme von Dopamin und Noradrenalin aufhalten, so dass die Konzentration der beiden Transmitter im synaptischen Spalt höher ist (Voeller, 2004). Eine Beteiligung bzw. Abweichung des katecholaminen Systems, an dem insbesondere Dopamin, Noradrenalin und Adrenalin beteiligt sind, ist daher bei Menschen mit ADHS wahrscheinlich.

Als strukturellen Beleg des Dopaminregulationsdefizits stellen Max und Kollegen (2002) fest, dass Läsionen im dopaminreichen ventralen Putamen zu einer Zunahme von ADHS bezogenen Verhaltensweisen führten. In einer Metaanalyse analysierten Ellison-Wright, Ellison-Wright und Bullmore (2008) sieben Studien mit insgesamt 114 Patienten mit ADHS und ähnlichen Störungen und 143 Kontrollpersonen. Sie fanden Hinweise, dass bei den Patienten mit ADHS das Volumen der grauen Substanz im rechten Putamen und beim Globus pallidus geringer war. Die Autoren diskutieren, ob dieser Befund ein anatomischer Marker für eine abweichende Funktion des frontostrialen Kreislaufes und damit der ADHS sein könnte.

3.6.6 Neuropsychologie

In neuropsychologischen Tests zeigen Kinder und Jugendliche mit ADHS zumeist schlechtere Ergebnisse, aber es scheint kein ADHS spezifisches Profil zu geben, das auf alle Betroffenen gleichsam zutrifft. Probleme in der Informationsverarbeitung ergeben sich insbesondere dann, wenn die Menge und Komplexität der zu verarbeitenden Informationen gesteigert wird (Banaschewski et al., 2004).

Viele empirische Untersuchungen fanden eine Beeinträchtigung der Exekutivfunktionen bei Menschen mit einer ADHS. In einer Metaanalyse fassten Willcutt, Doyle, Nigg, Faraone und Pennington (2005) die Ergebnisse von 83 Studien zusammen, insgesamt mit einer ADHS-Gruppe von $n = 3734$ und einer Kontrollgruppe von $n = 2969$. Die ADHS-Gruppe erreichte signifikant schlechtere Leistungen in allen Tests der Exekutivfunktionen. Es wurden Effektstärken zur Größe des Unterschiedes im mittleren Bereich gefunden ($d = 0.46 - 0.69$). Die größten Unterschiede wurden bei Aufgaben zur Reizinhibition, Vigilanz, Arbeitsgedächtnis und Vorausplanung festgestellt. Die schlechteren Leistungen der ADHS-Gruppe konnten nicht durch Unterschiede der Gruppen in der Intelligenz, schulischen Leistungen oder Komorbidität erklärt werden. Es wurde allerdings kein ADHS spezifisches Muster oder Profil festgestellt. Die Autoren vermuten, dass defizitäre Exekutivfunktionen allein nicht alle mit der ADHS verbundenen Probleme erklären können und auch nicht notwendigerweise zur Entstehung der Störung vorliegen müssen. Auch wenn einige Studien keine signifikante Beeinträchtigung der Exekutivfunktionen feststellen, so zeigen die meisten Studien, dass Menschen mit ADHS im Gruppenvergleich unterdurchschnittliche Leistungen in ihren Exekutivfunktionen zeigen (Seidman, 2006).

Durch Modifikation von Kontextfaktoren können Defizite in den Exekutivfunktionen zumindest teilweise ausgeglichen werden, was für eine Beeinflussbarkeit durch motivationale Aspekte spricht. Dies kann z. B. geschehen, wenn durch externe Hilfestellung die Anforderung an die Informationsverarbeitung gesenkt wird oder wenn Belohnungen zur Erhöhung der Motivation in Aussicht gestellt werden (Banaschewski et al., 2004). In einer Studie von Dovis, Van der Oord, Wiers und Prins (2012) bearbeiteten 30 Kinder mit ADHS und 31 Vergleichskinder eine visuell-räumliche Arbeitsgedächtnisaufgabe. Es zeigte sich, dass die Leistungen der Kinder mit ADHS signifikant schlechter waren als die der Kinder ohne ADHS. Es wurden zudem vier Verstärkerbedingungen umgesetzt, die sich auf die Motivation der Kinder auswirken sollten. In der ersten Bedingung bekamen die Kinder nur die Anweisung, sie sollen ihr Bestes geben und bekamen zum Schluss ein Feedback. In der zweiten Bedingung wurde den Kindern eine Belohnung von einem Euro für eine gute Leistung in Aussicht gestellt, in der dritten Bedingung 10 € und in der vierten Bedingung war die Arbeitsgedächtnisaufgabe in eine Computerspielumgebung übertragen worden. Erwartungsgemäß zeigte sich, dass die Kinder, die nur ein Feedback bekamen, die schlechtesten Leistungen erbrachten. Etwas besser waren die Leistungen der Kinder, die einen Euro Belohnung erhielten. Noch besser und etwa gleich gut schnitten die Kinder ab, die 10 € bekamen, und die Kinder mit der als Computerspiel umgesetzten Aufgabe. Aber auch die Leistungen der beiden besten ADHS-Gruppen blieben unter der Leistung der Kinder ohne ADHS, die für alle vier Belohnungsbedingungen gleich hoch war. Dieser Befund spricht dafür, dass sowohl kognitive als auch motivationale Defizite für die Leistungseinbußen verantwortlich sind. Dies deckt sich mit der Dual Pathway Theorie (Sonuga-Barke, 2002), die sowohl exekutive als auch motivationale Ursachen der ADHS postuliert.

Der Einfluss des Kontextes und insbesondere der externen Motivierung zeigt sich bei Kindern mit ADHS, z. B. darin, dass sie stärkere Leistungseinbußen zeigen, wenn ein längerer Aufschub von Belohnungen erfordert wird (Sonuga-Barke, 2002) oder wenn nach einer kontinuierlichen Verstärkung nur mehr intermittierend verstärkt wird (Sagvolden, Aase, Zeiner & Berger, 1998). Verzögert dargebotene Verstärker wirken rückwirkend weniger. Die Kinder haben einen steileren „delay-of-reinforcement gradient“, d. h. dass die Stärke der Verbindung zwischen ihrer Verhaltensreaktion und darauf folgender Belohnung stark abfällt, je mehr zeitliche Verzögerung dazwischen Reaktion und Verstärkung besteht (Johansen et al., 2009). Weiterhin versuchen Kinder mit ADHS Verzögerungen möglichst zu verringern („aversion of delay“), auch wenn sie dadurch größere Belohnungen, die ein Abwarten erfordern, nicht bekommen (Banaschewski, et al., 2004).

Beeinträchtigte Exekutivfunktionen werden nicht nur bei Kindern, sondern auch bei Jugendlichen und Erwachsenen festgestellt (Seidman, 2006). In einer Studie von Seidman et al. (2005) führten 101 Mädchen und 103 Jungen mit ADHS und 109 weibliche und 70 männliche Kontrollkinder neuropsychologische Aufgaben durch. Die Kinder und Jugendlichen mit ADHS hatten signifikant schlechtere Leistungen in den Aufgaben zu den exekutiven Funktionen, aber unterschieden sich nicht signifikant voneinander.

Dieses Verhältnis zeigte sich sowohl für Kinder im Alter von 9 bis 12 Jahren als auch für Jugendlichen von 13 bis 17 Jahren. Die Autoren schlussfolgern, dass mangelhafte Exekutivfunktionen als Korrelat bei ADHS unabhängig von Geschlecht und Alter zumindest bis in die späte Adoleszenz vorliegen. In einer Metaanalyse fassten Hervey, Epstein und Curry (2004) die Ergebnisse von 33 veröffentlichten Studien zusammen. Dabei wurden nur Studien mit Kontrollgruppen und Teilnehmern im Alter von 18 Jahren oder älter eingeschlossen. Die für das Kindes- und Jugendalter festgestellten schlechteren Leistungen der ADHS-Gruppe zeigten sich auch für die Erwachsenen mit ADHS.

Treffend fassen Banaschewski und Kollegen (2004) den Stand der Dinge zusammen: „Die Mehrzahl der vorliegenden empirischen Befunde scheinen am ehesten mit der Annahme einer *gestörten Fähigkeit zur optimalen Regulation des Aktivierungsniveaus* (kursiv im Original) (bzw. des kontrollierten Einsatzes energetischer Ressourcen) oder einer Abneigung gegen Verstärkungsverzögerung kompatibel zu sein“ (S. 140).

3.6.7 Soziale Interaktionen

Sozialen Interaktionen kommt bei der Entwicklung der ADHS vor allem eine die Problematik aufrechterhaltende bzw. verschärfende Funktion zu (Gawrilow, 2012). Die Auffälligkeiten des Kindes mit ADHS, die sich insbesondere in neuropsychologischen Defiziten äußern, werden von der Umwelt des Kindes in einer bestimmten Weise aufgenommen, dem Kind zurückgemeldet und bestimmen so die Gestaltung zwischenmenschlicher Beziehungen. In wechselseitigen Interaktionen über die Zeit prägen soziale Beziehungen zu Eltern, Lehrern und Gleichaltrigen die Entwicklung der Störung (Johnston & Scoular, 2008).

Interaktionen mit Eltern

Ein problematisches Erziehungsverhalten stellt insbesondere im frühen Kindesalter einen der wichtigsten Risikofaktoren zur Entwicklung externalisierender Verhaltensstörungen dar (Beelmann, Stemmler, Lösel & Jaursch, 2007). Das Auftreten von externalisierenden Verhaltensweisen und ADHS konnte in empirischen Untersuchungen durch inkonsistente und bestrafende elterliche Erziehung vorhergesagt werden (Schreyer & Hampel, 2009). Beelmann et al. (2007) konnten in einer Befragung von Eltern von 448 Kindern im Alter von drei bis sechs Jahren, dass einzelne Merkmale der Erziehung bis zu mittel hoch mit der Entwicklung problematischen Verhaltens korrelierten. Eine Faktorenanalyse ergab zwei Faktoren im elterlichen Erziehungsverhalten, die von den Autoren als „engagierte/ selbstbewusste Elternschaft“ und „problematisches Disziplinierungsverhalten“ bezeichnet wurden. Diese Faktoren korrelierten mit dem Auftreten und der Intensität des Problemverhaltens. Insbesondere trifft dies auf die problematische Disziplinierung zu und zwar sowohl für Mütter als auch für die Väter. Problematisches Erziehungsverhalten, das bei Familien mit Kindern mit externalisierenden Verhaltensauffälligkeiten häufiger gefunden wird, besteht in einer geringeren Supervision und

Aufsicht der Eltern, in einem inkonsistenten und harsch disziplinierendem Verhalten und auch körperlicher Gewalt und Bestrafungen (Beelmann et al., 2007).

Beziehungen und Interaktionen in Familien mit Kindern mit ADHS sind intensiv beforscht worden. Eltern von Kindern mit ADHS erleben mehr Stress, haben ein niedrigeres Selbstwertgefühl und häufiger depressive Symptome (Trautmann-Villalba et al., 2001). Diese Aspekte können in einem interaktiven Modell sowohl als Folgeerscheinungen als auch als ursächliche Faktoren bezogen auf die Probleme der Kinder verstanden werden. Es hat sich zudem gezeigt, dass das Risiko der Chronifizierung der ADHS höher ist, je mehr negative Interaktionen zwischen Eltern und Kindern vorliegen (Schreyer & Hampel, 2009).

Teilweise wurden aber auch widersprüchliche Ergebnisse gefunden (Johnston & Mash, 2001). Oftmals werden querschnittliche Designs verwendet, die Familien mit Kindern mit ADHS mit Familien ohne Kinder mit ADHS vergleichen. Längsschnittliche Verläufe und gemeinsame Entwicklungspfade mit anderen Störungen können somit in den wenigsten Untersuchungen auf kausale Zusammenhänge hin überprüft werden. Unklar ist z. B. wie sich Probleme im Sozialverhalten, die oftmals mit ADHS einhergehen, auf die familiären Interaktionen auswirken. Die wenigen vorliegenden längsschnittlichen Studien deuten darauf hin, dass familiäre Probleme stärker mit der SSV zusammenhängen als mit der ADHS. Es werden aber auch gemeinsame Entwicklungspfade der Störungen beschrieben und die genauen Wirkmechanismen sind noch unklar (Johnson & Mash, 2001).

Übereinstimmend wurde festgestellt, dass sich Eltern von Kindern mit ADHS direkter und stärker kontrollierend verhalten und weniger sozial mit ihren Kindern interagieren und diese weniger belohnen (Johnston & Mash, 2001; Trautmann-Villalba et al., 2001). Die Erziehung von Eltern mit Kindern mit ADHS wird als häufiger inkonsequent beschrieben, Leistungserwartungen und Interesse für das Kind waren geringer, es wurde häufiger heftig bestraft und positives Verhalten wurde weniger häufig positiv verstärkt (Schreyer & Hampel, 2009).

Auf der anderen Seite befolgten Kinder mit ADHS Anforderungen weniger häufig als Vergleichskinder ohne ADHS. Insbesondere scheint dies auf jüngere Kinder zuzutreffen, sowohl auf Jungen als auch auf Mädchen, und sowohl in der Interaktion mit ihren Müttern als auch mit ihren Vätern (Johnson & Mash, 2001). Weiterhin verhielten sich die Kinder mit ADHS vergleichsweise aggressiver, zeigten mehr Trotzverhalten, reagierten nicht auf Instruktionen oder Fragen, brachen Aufgaben häufiger ab und reagierten seltener auf Lob (Danforth, Anderson, Barkley & Stokes, 1991). Dass Interaktionsprobleme stärker bei jüngeren Kindern gefunden werden, kann damit zusammenhängen, dass die Symptomatik, insbesondere Hyperaktivität und Impulsivität, mit steigendem Alter abnimmt. Eine alternative Erklärung ist, dass Eltern aufgrund häufiger Misserfolge in ihrem Erziehungsverhalten Kontrollversuche aufgeben und sich aus der Interakti-

on mit ihrem Kind zurückziehen. Dies kann kurzfristig zu einer Entlastung führen, aber ist langfristig oft wenig hilfreich (Johnson & Mash, 2001).

Winsler (1998) beobachtete 19 Jungen mit ADHS und 20 gematchte Kontrollkinder, die eine Problemlöseaufgabe mit einem Elternteil bearbeiteten. Es stellte sich heraus, dass die Eltern der Jungen mit ADHS mehr verbale Kontrollstrategien einsetzten und ihre Kontrolle im Laufe der Aufgabenbearbeitung weniger zurücknahmen als die Eltern mit Kindern ohne ADHS. Zudem war ihr Scaffolding weniger effektiv und sie konnten ihre Unterstützung und Anforderungen weniger gut an die Fähigkeiten ihrer Kinder anpassen. Die Jungen mit ADHS selbst waren mehr mit Off-task-Verhalten beschäftigt als Jungen ohne ADHS und ihre Selbstanweisungen waren weniger effektiv. Es konnte gezeigt werden, dass der Kontext einen großen Einfluss auf die gegenseitigen Interaktionen hat. So zeigten Kinder in Anforderungssituationen mehr externalisierendes Verhalten als in freien und spielerischen Situationen und das Ausmaß von Konflikten nahm mit steigender Anforderung zu (Trautmann-Villalba et al., 2001). Trautmann-Villalba und Kollegen beobachteten Interaktionen zwischen 298 Kindern mit ADHS und SSV im Alter von acht Jahren und ihren Müttern in einer Leistungssituation. Es zeigte sich, dass die Mütter der Kinder mit ADHS ihren Kindern häufiger restriktive Anweisungen gaben und sich häufiger abwertend äußerten. Sie kontrollierten ihre Kinder mehr, diese Steuerung wurde aber als weniger adäquat beurteilt. Die Kinder mit ADHS wurden als stärker widerstrebend und abwertender der Mutter gegenüber beschrieben und sie konnten sich seltener durchsetzen und verhielten sich weniger hilflos. Darüber hinaus beschrieben die Autoren ein typisches Interaktionsmuster. Sie fanden, dass die Kinder mehr hyperaktive Verhaltensweisen zeigten, wenn sich die Mütter ihnen gegenüber aversiv verhielten. Waren die Mütter nicht aversiv, zeigte sich kein signifikanter Unterschied im Verhalten der Kinder mit ADHS zu Vergleichskindern. Die Autoren erklären dies mit gegenseitigen negativen Verstärkungsprozessen: Die Mütter versuchten hyperkinetisches Verhalten ihrer Kinder durch strenges, einschränkendes Verhalten zu unterbinden, was nicht immer erfolgreich sei. Die Kinder reagierten aus Überforderung mit dieser Einschränkung weiterhin oder zunehmend impulsiv und widerstrebend. So entstehe ein Teufelskreis, der zur Steigerung der Problematik führe.

Wie auch bei oppositionellen Verhaltensweisen von Kindern werden in Familien mit Kindern mit ADHS sog. koersive Interaktionsmuster gefunden (Patterson, 1982). Dabei führen negative Verstärkungseffekte zu erpresserischem Verhaltensweisen, die dissoziale Verhaltensweisen beim Kind fördern und die Interaktionen zwischen Eltern und Kind stark beeinträchtigen (Beelmann et al., 2007). Kinder versuchen dabei durch anhaltendes aversives Verhalten, z. B. Quengeln und Trotzen, ihre Eltern so lange unter Druck zu setzen, bis diese nachgeben. Sind sie damit erfolgreich, wird das unangemessene Verhalten verstärkt. Da sie nicht immer erfolgreich sind, wird das Verhalten zudem intermittierend verstärkt, so dass es eine höhere Lösungsresistenz aufweist (Trautmann-Villalba et al., 2001). Auf der anderen Seite drohen die Eltern so lange mit Bestrafung, bis das Kind nachgibt. Eltern und Kinder versuchen also gegenseitig, das Verhalten des

anderen zu erzwingen. Es entstehen Kreisläufe des sich gegenseitig unter Druck Setzens und Erpressung. Diese werden durch intermittierende negative Verstärkung verfestigt.

Es fanden sich auch Unterschiede in den auf die Erziehung bezogenen Kognitionen von Eltern mit und ohne Kinder mit ADHS. Eltern von Kindern mit ADHS haben eine niedrigere Einschätzung ihrer eigenen Erziehungskompetenz. Darüber hinaus stellten Sonuga-Barke und Goldfoot (1995) fest, dass Mütter von Jungen mit ADHS geringere Erwartungen an die Entwicklung ihres Kindes hatten als Mütter von Kindern ohne ADHS, auch bei nach dem IQ gematchten Gruppen. Dies könnte entweder dafür sprechen, dass die Eltern die Kompetenzen ihres Kindes unterschätzten oder dass die Kinder in der Performanz hinter ihren Möglichkeiten zurück blieben. Bezogen auf die Attribution des Verhaltens der Kinder durch ihre Eltern wird oftmals festgestellt, dass die ADHS-Symptomatik auf unkontrollierbare, stabile Faktoren innerhalb des Kindes zurückgeführt wird und positive Verhaltensweisen als weniger dispositional gesehen werden (Johnson & Mash, 2001).

Bezogen auf die Emotionalität wird dem stärkeren Vorliegen negativer Emotionen sowohl ursächliche als auch aufrecht erhaltende Funktion bei externalisierenden Verhaltensstörungen zugeschrieben. Ärgerliches Reagieren auf trotziges Verhalten des Kindes erschwerten frühzeitiges, proaktives Erziehungsverhalten und die Entwicklung angemessen gleichgewichtiger Interaktionen. Weiterhin kann sich das externalisierende Verhalten verstärken, wenn positive Emotionen fehlten bzw. selten vorkämen (Trautmann-Villalba et al., 2001).

Auch konfliktreiche partnerschaftliche Beziehungen können die Beziehungen mit einem Kind mit ADHS verschärfen bzw. sich gegenseitig. Einerseits kann eine Folge elterlichen Streits ein inkonsistentes Erziehungsverhalten sein, andererseits können Prozesse des Modelllernens dazu führen, dass konfliktreiche Interaktionen und Kommunikationsmuster übernommen werden (Johnson & Mash, 2001). Theule, Wiener, Tannock und Jenkins (2013) fanden in einer Metaanalyse über insgesamt 44 publizierte und nicht publizierte Studien, dass Eltern mit Kindern mit ADHS mehr Stress in ihrer Erziehung erlebten als Eltern von Kindern ohne ADHS und dass der Stress höher war, je deutlicher die Symptomatik ausgeprägt war. Allerdings wurde nicht konsistent gefunden, dass die Trennungs- und Scheidungsrate in Familien mit Kindern mit ADHS höher ist.

Es werden auch Zusammenhänge zwischen ADHS und elterlicher Pathologie diskutiert, u. a. elterliche Alkoholabhängigkeit. Diese scheinen allerdings eher mit kindlichen Problemen im Sozialverhalten in Verbindung zu stehen als mit ADHS. Es ist in diesem Zusammenhang zu bedenken, dass familiäre Probleme einen universellen Risikofaktor darstellen und nicht spezifisch für das Auftreten einer ADHS sind (Johnson & Mash, 2001). Einige Studien finden auch einen Zusammenhang zwischen ADHS und einem geringen sozio-ökonomischen Status der Familie (Brassett-Harknett & Butler, 2007). Auch dabei handelt es sich allerdings eher um einen allgemeinen als einen ADHS spezifischen Risikofaktor.

Trautmann-Villalba et al. (2001) weisen darauf hin, dass es wichtig ist, Kausalzuschreibungen vorsichtig vorzunehmen, z. B. bei der oft untersuchten Rolle der Mutter und ihrer Interaktionen mit dem Kind. Es handelt sich bei den Problemkonstellationen zu meist um komplexe Interaktionen zwischen einem herausfordernden kindlichen Verhalten, das schnell zu einer Überforderung der Mütter bzw. der Eltern führen könne. Somit sei eine generelle Schuldzuschreibung zu Lasten der Mütter bzw. den Eltern nicht angemessen.

Interaktionen mit Lehrern

Kinder und Jugendliche mit ADHS verhalten sich im Unterricht anders als Schüler ohne ADHS. Sie benötigen mehr Aufmerksamkeit durch die Lehrperson, bekommen Anweisungen nicht mit, führen Aufgaben nicht zu Ende und stören häufiger aktiv den Unterricht (Lauth & Naumann, 2009). Ihre Probleme in der Daueraufmerksamkeit und in der Inhibition irrelevanter Reize erschwert es Ihnen, sich unterrichtskonform zu verhalten. Neben aktivem, unterrichtsstörendem Verhalten folgen sie auch weniger dem Unterricht. So konnten Kofler, Rapport und Alderson (2008) in einer Metaanalyse über 23 kontrollierte Vergleichsstudien zeigen, dass sich die Schüler mit ADHS nur in 75 % der Zeit mit unterrichtsrelevanten Dingen beschäftigten – also „on-task“ waren – im Vergleich zu 88 % bei ihren Mitschülern ohne ADHS. Dies entspricht einer Effektstärke von 1.4. Zudem zeigten die Kinder mit ADHS eine größere Variabilität in ihrem Aufmerksamkeitsverhalten. Alle Schüler schwanken im Unterricht zwischen aufmerksamen und unaufmerksamen Zuständen. Allerdings sind die Wechsel zwischen den Zuständen bei Kindern mit ADHS häufiger und ihre Aufmerksamkeitsspanne ist kürzer (Rapport, Kofler, Alderson, Timko & DuPaul, 2009).

In einer Studie von Lauth und Mackowiack (2004) bzw. Lauth, Heubeck und Mackowiack (2006), die auch in die Metaanalyse von Kofler et al. (2008) eingeschlossen wurde, wurden 55 Grundschüler mit ADHS und 55 gematcht zugeordnete Kinder ohne ADHS im Unterricht beobachtet. Auch hier waren die Kinder mit ADHS seltener „on-task“ als ihre Mitschüler. Zudem wurde beobachtet, dass das Aufmerksamkeitsverhalten stark mit dem Unterrichtskontext zusammenhing. Dabei unterschieden sich die Kinder mit ADHS nicht von den Kindern ohne ADHS. Für alle Schüler galt, dass aktiv störendes Verhalten häufiger bei solchen Unterrichtsformen gezeigt wurde, die ohne Interaktionen ablaufen, z. B. bei der stillen Einzelbeschäftigung mit einer Aufgabe. Passives Verhalten wurde dagegen im normalen Unterricht häufiger beobachtet. In interaktiven Situationen verhielten sich beide Schülergruppen häufiger nicht anforderungsgemäß. Für alle Verhaltensweisen des „on-task“ zusammengenommen wurden weder für Kinder mit ADHS noch für Kinder ohne ADHS Unterschiede je nach Kontext festgestellt.

Einige Studien konnten zeigen, dass Lehrer Schüler anders bewerten, wenn Sie wissen, dass eine ADHS vorliegt. In einer kanadischen Studie untersuchten Ohan, Visser, Strain und Allen (2011) die Bewertung von Fallvignetten durch 34 Grundschullehrer und 32 Studierende, um zu überprüfen, ob ein „labeling bias“ vorliegt. Die Vignetten beschrie-

ben Kinder, die eine ADHS-Symptomatik aufwiesen, aber nur in einer Version der Vignetten wurde die Bezeichnung ADHS verwendet. Bei den Vignetten mit ADHS schätzten die Lehrer und Studierenden die Probleme und Einschränkungen der Schüler als gravierender ein. Zudem wurden mehr negative Emotionen in Reaktion auf die Kinder mit der Bezeichnung ADHS berichtet und es wurde weniger Zuversicht auf einen gelingenden weiteren Umgang mit dem Kind geäußert. Demgegenüber war allerdings die Bereitschaft höher, dem Kind eine Förderung zu ermöglichen. Die Autoren weisen darauf hin, dass es sich bei den Einschätzungen nur um einen ersten Eindruck nach Aktenlage handelt, der nicht sehr naturalistisch sei. Trotzdem kommt dem ersten Eindruck von Personen eine große Vorhersagekraft zu, ist dieser negativ, fällt es oft schwer, die Einschätzung zu verändern, wie Studien sozialpsychologische Studien zeigen (Downey & Christensen, 2006).

Schüler mit ADHS haben Schwierigkeiten in der Performanz, d. h. es gelingt Ihnen weniger gut ihre Kompetenzen umzusetzen. Da sie grundsätzlich in der Lage sind, sich anforderungsgemäß zu verhalten, oftmals aber sehr viel mehr Anleitung und Kontrolle benötigen, entsteht bei einigen Lehrern der Eindruck, der ADHS betroffene Schüler wolle lediglich provozieren, da er eigentlich ja wisse was verlangt werde und dies prinzipiell auch umsetzen könne (Lauth & Naumann, 2009). Daher reagieren einige Lehrer stärker mit Bestrafungen, um das Auftreten der vermeintlichen Disziplinlosigkeit zu verringern. Notwendige Anleitung, Überwachung und Lenkung, am besten proaktiv bevor ein Problem auftritt, des Schülers kommt häufig zu kurz. Auch in der Interaktion von Lehrern und Schülern mit ADHS ist anzunehmen, dass sich wie in häuslichen Situationen ungünstige Interaktionskreisläufe abspielen, bei denen unangemessenes Verhalten unbeabsichtigt verstärkt wird. Dies ist z. B. der Fall, wenn eine Anweisung des Lehrers vom Schüler nicht befolgt wird, auch nicht nach mehrmaliger Aufforderung und der Lehrer nach längerer Diskussion und Ermahnung schließlich aufgibt. Durch das Insistieren kann der Schüler schließlich die unangenehme Aufforderung und Ermahnung beenden und somit wird sein widerspenstiges Verhalten verstärkt. Analoges gilt, wenn ein Schüler aufgrund von Störverhalten aus dem Klassenraum geschickt wird. Wenn es sich um ein Fach oder einen Lehrer handelt, den der Schüler nicht mag, und er sich lieber außerhalb der Klasse aufhält, wirkt das Hinausschicken negativ verstärkend für das Störverhalten, weil die unangenehme Unterrichtssituation dadurch beendet wird (Lauth & Naumann, 2009).

Interaktionen mit Gleichaltrigen

In einem Vergleich von 165 Kindern mit und 165 Kindern ohne ADHS fanden Hoza und Kollegen (2005), dass die Kinder mit ADHS von ihren Gleichaltrigen weniger gemocht und präferiert wurden, weniger enge Freunde hatten und häufiger zurückgewiesen wurden. Kinder mit ADHS wurden seltener von solchen Kindern als Freunde ausgesucht, die selbst sozial stark akzeptiert wurden. Bagwell, Molina, Pelham und Hoza (2001) verglichen 111 Jugendliche mit ADHS und 100 Jugendliche ohne ADHS im Alter von 13 bis 18 Jahren. Die Eltern der Jugendlichen berichteten, dass ihre Kinder

weniger enge Freunde hätten und häufiger von Gleichaltrigen zurückgewiesen würden. Die Jugendlichen selbst berichteten, dass ihre Freunde seltener an konventionellen Ereignissen wie Schul- oder Sportveranstaltungen teilnahmen. Damit soll das Ausmaß an prosozialen Aktivitäten von Freunden erfasst werden, welches demnach geringer ist bei Freunden von Jugendlichen mit ADHS. Weiterhin konnte regressionsanalytisch festgestellt werden, dass eine ADHS-Symptomatik im Kindesalter die sozialen Probleme im Jugendalter vorhersagen konnte. Dies war auch dann der Fall, wenn im Jugendalter kein Vollbild der ADHS mehr vorlag. Lag zusätzlich eine Störung des Sozialverhaltens vor, waren die Probleme erwartungsgemäß stärker ausgeprägt. Es zeigte sich weiterhin, dass auch dem Vorliegen aggressiver Symptomatik im Kindesalter eine Bedeutung für soziale Probleme in der Adoleszenz zukam. So war Aggression im Kindesalter prädiktiv für eine geringere von den Jugendlichen selbst wahrgenommene soziale Kompetenz. Die Ergebnisse unterstreichen die Stabilität sozialer Probleme bei vorliegender ADHS vom Kindes- zum Jugendalter, die im Allgemeinen für externalisierende Störungen gefunden werden (z. B. Laird, Jordan, Dodge, Pettit & Bates, 2001; Lansford, Malone, Dodge, Pettit & Bates, 2010).

Murray-Close et al. (2010) analysierten die Stichprobe der MTA-Studie ($n = 536$) und verglichen diese mit einer Kontrollgruppe ($n = 284$) über einen Zeitraum von sechs Jahren. Es zeigte sich, dass die Kinder mit ADHS zu Beginn der Untersuchung in allen erfassten Bereichen Probleme aufwiesen. Diese waren Aggression, soziale Kompetenz, positiver illusorischer Bias und Zurückweisung durch Gleichaltrige. Die Autoren stellen fest, dass es bezogen auf die zeitliche Entwicklung der Probleme Teufelskreise entwickelten. So führte beispielsweise die Zurückweisung durch Gleichaltrige zu geringeren sozialen Kompetenzen, die wiederum spätere Zurückweisung durch Gleichaltrige vorhersagte. Wie sich die sozialen Probleme entwickelten, war nicht unterschiedlich bei der Gruppe ohne ADHS und der Gruppe mit ADHS. Aber das Ausmaß der Probleme war größer bei den Kindern mit ADHS. In einer anderen Auswertung der MTA-Studie konnten und Mrug und Kollegen (2012) zeigen, dass die Zurückweisung durch Gleichaltrige sowohl Zigarettenkonsum, Delinquenz, Ängstlichkeit und allgemeine Einschränkungen sechs Jahre nach der ersten Erhebung vorhersagen konnte. Einen guten Freund zu haben war hingegen kein signifikanter Prädiktor und konnte die negativen Effekte der Zurückweisung durch Gleichaltrige nicht reduzieren.

In einer Untersuchung von 87 Kindern mit ADHS und 46 Vergleichskindern im Alter von 7 bis 13 Jahren wurden Erkenntnisse zu freundschaftlichen Beziehungen gewonnen (Normand et al., 2011). Die Kinder mit ADHS hatten häufiger Freunde, die selbst ebenfalls ADHS-typische und oppositionelle Verhaltensweisen aufwiesen. Die Kinder mit ADHS bekamen weniger positive Rückmeldungen und mehr negative Rückmeldungen und waren weniger zufrieden mit ihren Freundschaften. Verhaltensbeobachtungen zeigten, dass die Kinder mit ADHS in Spielen mit Wettbewerbscharakter sowohl mehr legale als auch illegale Spielzüge umsetzten. In Auseinandersetzungen und Diskussionen mit ihren Freunden waren die Kinder mit ADHS häufiger dominant, unsensibel und machen mehr selbstbezogene Vorschläge.

Im Zusammenhang mit sozialen Schwierigkeiten bei ADHS scheinen Prozesse des Bullyings eine Rolle zu spielen. In einer italienischen Studie mit 195 Kindern im Alter von 8 bis 10 Jahren zeigte sich, dass der in der Studie gefundene Zusammenhang zwischen ADHS-Symptomen und sozialer Beliebtheit komplett über Bullying in der Schule erklärt werden konnte (Bacchini, Affuso & Trotta, 2008). Einige weitere Studien finden einen Zusammenhang zwischen ADHS und Bullying. Dabei sind Kinder und Jugendliche mit ADHS sowohl häufiger Täter als auch Opfer. Dies zeigte sich z. B. in einer schwedischen Untersuchung mit 516 Kindern im Alter von zehn Jahren (Holmberg & Hjern, 2008). Auch Unnever und Cornell (2003) fanden in einer Studie mit 1.315 Mittelschülern, dass Schüler, die berichteten Medikamente für Ihre ADHS zu nehmen, ein höheres Risiko hatten Opfer oder Täter von Bullying zu werden. In Mediationsanalysen zeigte sich, dass der Zusammenhang zwischen ADHS und der Täterrolle beim Bullying vollständig über die gering ausgeprägte Fähigkeit zur Selbstkontrolle erklärt werden konnte. Dies traf nicht für die Opferrolle zu. Auch Wiener und Mak (2009) fanden ein höheres Risiko, Täter und/oder Opfer von Bullying zu sein, bei 52 Kindern mit ADHS im Vergleich zu 52 Kindern ohne ADHS im Alter von 9 bis 14 Jahren. In dieser Studie zeigte sich, dass von den Eltern eingeschätzte oppositionelle Verhaltensweisen als Mediator des Zusammenhangs zwischen ADHS-Symptomen und Bullying fungierten.

Soziale Schwierigkeiten, insbesondere Bullying und oppositionelle Verhaltensweisen, spielen eine wesentliche Rolle in der Verschärfung der ADHS-Problematik. Sie tendieren dazu sich zu verfestigen und haben damit einen störungsaufrechterhaltenen Charakter. Kinder ohne ADHS scheinen Kinder mit ADHS eher zu meiden. Diese haben weniger Freundschaften und es fehlen ihnen daher an Möglichkeiten soziale Kompetenzen zu erlernen und zu verbessern.

3.7 Intervention

In diesem Abschnitt werden Behandlungsmöglichkeiten der ADHS mit einem Schwerpunkt auf Interventionen für Jugendliche vorgestellt. Begonnen wird mit einer kurzen Darstellung medikamentöser Behandlungsmöglichkeiten gefolgt von einem Überblick über psychosoziale Interventionen anhand von Metaanalysen und Reviews. Es folgt eine Vertiefung in Form einer Darstellung des Lerntrainings für Jugendliche mit ADHS, das im empirischen Teil dieser Arbeit evaluiert wird. Das Konzept und die Evidenzbasierung des Trainings werden ausführlich dargestellt.

Zur Behandlung der ADHS im Kindes- und Jugendalter existieren nationale und internationale Behandlungsleitlinien. Übereinstimmend wird eine multimodale Therapie empfohlen, die nahezu ausschließlich ambulant durchgeführt wird. Die Therapie setzt sich aus den Bausteinen Psychoedukation, Pharmakotherapie, Verhaltens- und Kognitive Therapie, Selbsthilfe und Behandlung komorbider Störungen und Problemlagen zusammen (Arbeitsgemeinschaft ADHS der Kinder- und Jugendärzte, 2007; Deutsche Gesellschaft für Kinder- und Jugendpsychiatrie und Psychotherapie, 2007). Dies gilt

ebenso für die Therapie von ADHS im Erwachsenenalter (Ebert, Krause & Roth-Sackenheim, 2003). Ergänzend wird ein Coaching für Erwachsene mit ADHS empfohlen, das u. a. bei der Organisation alltäglicher Aufgaben unterstützen soll (Kooij et al., 2010).

3.7.1 Medikation

In der Öffentlichkeit kontrovers diskutiert wird die medikamentöse Behandlung von Kindern und Jugendlichen mit ADHS. In Deutschland gelten Stimulantien und Atomoxetin, die zumeist in Tablettenform oral eingenommen werden, als Behandlungsmittel der Wahl (Pelz, Banaschewski & Becker, 2008). Am häufigsten wird der Wirkstoff Methylphenidat eingesetzt, der unter verschiedenen Handelsnamen erhältlich ist. Es handelt sich um eine amphetaminähnliche Substanz, die dafür sorgt, dass Dopamin und Noradrenalin langsamer aus dem synaptischen Spalt wiederaufgenommen werden, so dass deren Konzentration im synaptischen Spalt höher ist. Die Wirkung, die fast unmittelbar nach der Einnahme beginnt, nimmt bei vielen Kindern, die das Medikament früh am Morgen einnehmen, schon im späteren Verlauf des Vormittages ab. Um die Wirkdauer zu verlängern gibt es mittlerweile länger wirkende Retardpräparate oder es werden über den Tag verteilt mehrfach Tabletten eingenommen. Die Responderrate, also der Prozentsatz von Kindern und Jugendlichen, die auf die Medikation reagieren, ist mit ca. 75 % sehr hoch (Pliszka, 2003). Seit 2004 ist auch der Wirkstoff Atomoxetin zur Behandlung der ADHS im Kindes- und Jugendalter zugelassen. Er sorgt dafür, dass Noradrenalin langsamer aus dem synaptischen Spalt wiederaufgenommen wird. Es dauert ca. 6-8 Wochen, bis die vollständige Wirkung erreicht wird. Danach scheint die Wirkung über den gesamten Tagesablauf stabil zu sein (Pelz et al., 2008). Auch die Responderrate von Atomoxetin ist mit ca. 60 % als hoch zu bezeichnen (Spencer et al., 2002). Überaus wichtig bei der medikamentösen Behandlung ist die Kontrolle unerwünschter Nebenwirkungen. Diese können u. a. Schlafstörungen, Appetitverlust und Bauch- oder Kopfschmerzen sein (Weber, Bubl & Lüscht, 2003). Es muss individuell passend das richtige Präparat mit der richtigen Dosierung ausgewählt und fortlaufend überprüft werden.

Übereinstimmend wird in den Behandlungsleitlinien darauf hingewiesen, dass die Pharmakotherapie nicht die alleinige Art der Behandlung sein soll. Auch bei sehr hohen Responderraten profitieren nicht alle Kinder und einige leiden unter Nebenwirkungen oder anderen Kontraindikationen. Zudem ist fraglich, ob überdauernde Kompetenzen z. B. im Lernverhalten durch eine reine Medikation aufgebaut werden können. Insbesondere auch bei der Behandlung von Jugendlichen mit ADHS kommt den psychosozialen Interventionen eine große Bedeutung zu. Pharmakotherapeutische Behandlung bei Schulkindern und Jugendlichen stellt nicht die Behandlung der Wahl dar, sondern es wird empfohlen, diese zur Behandlung schwerwiegender Symptomatiken zurückzuhalten bzw. wenn nicht-medikamentöse Behandlungen ineffektiv waren (National Institute for Health and Care Excellence, 2008). Geeignete Interventionen wie Verhaltensthera-

pie werden zur Behandlung von Adoleszenten als Alternative oder als Ergänzung der medikamentösen Behandlung empfohlen (Wolraich et al., 2011).

3.7.2 Psychosoziale Interventionen

Zur Wirksamkeit psychosozialer Interventionen bei ADHS liegen mittlerweile Metaanalysen vor, zusammengefasst z. B. von Bachmann, Bachmann, Rief und Mattejat (2008). Der Grad der Evidenzbasierung im Jugendalter ist im Vergleich zum Erwachsenenalter und insbesondere zum Kindesalter als dürftig zu bezeichnen (Barkley, 2004; Fabiano et al., 2009). So können Young und Amarasinghe (2010) in ihrem Review keine Studien zur Wirksamkeit kognitiv-behavioraler Therapie bei Adoleszenten mit ADHS finden. Auch zur Wirksamkeit von Elterntraining und schulischen Interventionen bei ADHS im Jugendalter finden die Autorinnen nur unkontrollierte bzw. Einzelfall-Studien, die für eine Wirksamkeit sprechen. So verglich Morris (1993) zwölf Jugendliche mit einer Störung des Sozialverhaltens (SSV) mit zwölf Jugendlichen mit ADHS, die in einem klinischen Setting zwölf Wochen lang mit einer rational-emotiven Therapie behandelt wurden. Während sich die Jugendlichen mit SSV signifikant verbessern konnten, zeigte die ADHS-Gruppe keine signifikanten Veränderungen. Es ist allerdings fraglich, ob aufgrund der geringen Stichprobengröße eine ausreichende Teststärke vorlag, um signifikante Effekte festzustellen. Es könnte aber auch vermutet werden, dass der rein kognitive Ansatz für die Jugendlichen mit ADHS nicht geeignet war. In zwei Studien verglich eine Arbeitsgruppe um Barkley die Effekte von einem verhaltensbezogenen Elterntraining (BPT) mit einem kombinierten Problemlöse- und Kommunikationstraining (PSCT) zur Behandlung von Jugendlichen mit ADHS und ihren Familien (Barkley, Edwards, Laneri, Fletcher & Metevia, 2001; Barkley, Guevremont, Anastopoulos & Fletcher, 1992). In der aktuelleren Studie wurden 97 Familien entweder über 18 Termine nur mit dem PSCT behandelt oder über 9 Termine mit dem BPT und im Anschluss für 9 Termine mit dem PSCT. Für beide Interventionen zeigten sich signifikante Verbesserungen in den Konflikten zwischen Adoleszenten und Eltern. Die beiden Interventionen unterschieden sich allerdings nicht signifikant voneinander. Bei der reinen PSCT-Behandlung zeigte sich ein signifikant höherer Drop-out. Insgesamt erreichten 23 % der Familien eine reliable Verbesserung. Die Autoren schließen, dass die festgestellten Verbesserungen auf der Gruppenebene für die Effektivität der Intervention spreche, die Ergebnisse zur reliablen Veränderung allerdings weniger gut ausfallen. Auch Pelham und Fabiano (2008) kommen in einem Review zu dem Schluss, dass PSCT eine vielversprechende Intervention für die Gruppe der Jugendlichen mit ADHS sei, die stärker beforscht werden solle.

Es liegt ein etwas älteres Review vor, das sich mit der Wirksamkeit von Intervention speziell für Jugendliche beschäftigt (Smith, Waschbusch, Willoughby & Evans, 2000). Bezogen auf psychosoziale Interventionen wird das Training des Erstellens von Notizen im Rahmen von Vorlesungen als vermutlich effektiv beurteilt. Als vielversprechend, aber nicht validiert wird die Verhaltensmodifikation im Unterricht und Familientherapie eingestuft. Kognitiv-behaviorale Gruppentherapie wurde als vermutlich nicht effektiv

eingestuft. Das Erstellen von Notizen und die Familientherapie werden in Begriffen der Pharmakoprüfung als sicher und praktikabel bezeichnet. Die Autoren berichten darüber hinaus, dass wenig über die Langzeiteffekte und Compliance der Interventionen bekannt sein, ebenso wenig wie über die Kombinationsbehandlung aus Stimulantienbehandlung und Verhaltenstherapie. Zu diesem Schluss kommt auch ein aktuelleres Review von Chronis, Jones und Raggi (2006).

3.7.3 Das Lerntraining für Jugendliche mit ADHS (LeJA)

Da wirksame Interventionskonzepte speziell zur Behandlung von Jugendlichen mit ADHS bislang fehlen, wurde das Lerntraining für Jugendliche mit ADHS (LeJA) (Linderkamp et al., 2011) entwickelt. Ziele des Trainings sind 1. die Verbesserung des Lern- und Arbeitsverhaltens der Jugendlichen und 2. die Bewältigung aktueller Entwicklungsaufgaben, insbesondere bezogen auf den Übergang ins Berufsleben. Das erste Ziel soll durch kontinuierliche Übungsbehandlungen erreicht werden, das zweite Ziel wird in einem Selbstmanagement- und Coaching-Ansatz verfolgt.

Als ein zentrales Merkmal der ADHS werden Defizite der exekutiven Funktionen angenommen, die zu einer ineffizienten Selbst- bzw. Handlungsteuerung der Jugendlichen beitragen. Diese neuropsychologischen Defizite führen dazu, dass Handlungen, die Aufmerksamkeit und Konzentration erfordern, nicht adäquat durchgeführt werden. Durch Unaufmerksamkeit und Impulsivität werden Handlungen nicht korrekt ausgeführt und Handlungsziele nicht erreicht. Daher werden solche Strategien und Techniken eingeübt, welche die intentionale Handlungsteuerung verbessern. Dieses Vorgehen ist vielversprechender als das reine Trainieren von Konzentrationsaufgaben, das zwar zu einer schnellen und großen Verbesserung führt, sich allerdings nur wenig auf schulische Aufgaben generalisiert (Lauth, 2004).

Konzeptionelle Grundlagen des Trainings sind Techniken der kognitiven Verhaltenstherapie, die sich in empirischen Studien als wirksam zur Förderung von Lernprozessen erwiesen haben (Grünke, 2006). Diese werden auch erfolgreich in der Behandlung von Kindern mit ADHS eingesetzt (Döpfner, Schürmann & Frölich, 2013; Lauth & Schlottke, 2009). Zur Unterstützung bei der Bewältigung anstehender Entwicklungsaufgaben wird ein Coaching zur besseren Selbstorganisation durchgeführt, wie es auch mit guten Effekten in der Therapie von Erwachsenen mit ADHS umgesetzt wird (Ramsay & Rostain, 2007). Persönliche Probleme, die einer erfolgreichen Bewältigung von Entwicklungsaufgaben im Wege stehen, werden im Rahmen eines Selbstmanagement-Ansatzes bearbeitet (Kanfer, Reinecker & Schmelzer, 2006). Grundlegend orientiert sich die Haltung des Trainers an den durch Grawe und anderen identifizierten allgemeinen Wirkfaktoren in Psychotherapien (Grawe, 2005).

Das Training richtet sich an Jugendliche mit ADHS aller Störungssubtypen im Alter von 12 bis 17 Jahren. Liegen schwerwiegende Komorbiditäten vor, ist abzuwägen ob das Training die Behandlung der Wahl ist. Bei akuten Problemlagen sind vermutlich

andere Maßnahmen dem Lerntraining vorzuziehen. Das Lerntraining kann aber auch als ein Baustein in einer multimodalen Behandlung eingesetzt werden. Komorbide Probleme, z. B. im Sozialbereich, können begrenzt auch im Rahmen des Lerntrainings behandelt werden, nehmen aber hinter den aufs Arbeitsverhalten bezogenen Maßnahmen einen geringeren Stellenwert ein.

Die Umsetzung des Trainings findet in 16 bis 20 Einzelsitzungen statt. Wenn möglich soll zumindest ein Termin pro Woche stattfinden. Die Inhalte des Trainings lassen sich grob in vier Blöcke einteilen. Zu Beginn des Trainings wird dem Aufbau einer partnerschaftlichen therapeutischen Beziehung genügend Raum gewidmet. Nach Grawe (2005) kommt der therapeutischen Beziehung als allgemeinem Wirkfaktor eine zentrale Bedeutung zu. In einer Rollenklärung legt der Trainer dar, dass er dem Jugendlichen helfen und ihn unterstützen wird, dieser die angestrebten Ziele letztendlich aber selbst erreichen muss. Zur Verdeutlichung dieses Verhältnisses wird das Bild eines Trainers oder Coaches beim Sport genutzt. Auch dieser trainiert mit den Spielern außerhalb des Spiels, steht beim entscheidenden Punktspiel aber nicht mit auf dem Platz. In einem Trainingsvorgang erklären Jugendliche und Trainer, dass sie sich an verbindliche Absprachen halten werden und sich Mühe geben wollen. So soll die Bereitschaft der Jugendlichen zur aktiven Mitarbeit gesteigert werden und dass therapeutische Bündnis verfestigt werden. Im Ablauf folgen zwei Sitzungen mit psychoedukativen Inhalten. Unter Berücksichtigung der subjektiven Sichtweise des Jugendlichen wird mithilfe von Brainstormings und Modellierungsdialogen ein individuelles Bedingungsmodell der ADHS-Symptomatik und den Lernschwierigkeiten des Jugendlichen erarbeitet. Der Trainer betont, dass vorhandene Probleme veränderbar sind, wenn der Jugendliche an ihn arbeitet. Weiterhin werden auch positive Aspekte genannt und Ressourcen des Jugendlichen festgestellt. Auch die Ressourcenaktivierung wurde nach Grawe (2005) als wichtiger allgemeiner Wirkfaktor therapeutischer Intervention festgestellt. Parallel zu den ersten drei Sitzungen nimmt der Trainer Kontakt zu einem gemeinsam mit dem Jugendlichen ausgewählten Lehrer oder Lehrerin auf. Basierend auf den gesammelten Einschätzungen von Lehrern, Eltern, dem Jugendlichen selbst und der Auswertung vorliegender Arbeitsproben werden in der vierten Sitzung zentrale Problembereiche und dazugehörige Ziele für das Training festgelegt.

Im umfangreichen zweiten Trainingsblock wird die Verbesserung des Lern- und Arbeitsverhaltens des Jugendlichen angestrebt. Im Vordergrund steht die Einübung eines systematischen Vorgehens bei der Bearbeitung von (Schul-)Aufgaben. Unterstützend werden in Anlehnung an Camp und Bash (1981) Signalkarten erarbeitet, auf denen die notwendigen Schritte verdeutlicht sind. Jeder Schritt wird mit einer Selbstanweisung versehen, die im Sinne eines Selbstinstruktionstrainings (Meichenbaum & Goodman, 1971) genutzt werden. Anschließend wird dieser ideale Problemlöseablauf eingeübt. Es wird mit Sudoku- und anderen nicht schulbezogenen Aufgaben begonnen, später und ausgiebiger folgen aktuelle Schulaufgaben, die der Jugendliche mitbringt oder die der Trainer nach telefonischer Absprache mit den Lehrern auswählt. Die Aufgaben sollen eine mittlere Schwierigkeit aufweisen, so dass sie einerseits nicht zu leicht erscheinen

und andererseits nicht überfordern. Der Trainer sorgt über verhaltensnahe Instruktionen und positives Feedback gelungener Verhaltensweisen mittels sozialer Verstärkung dafür, dass der Jugendliche die Aufgaben durch systematisches Vorgehen lösen kann. Individuell passend zu den schulischen Problemen des Jugendlichen werden spezifische und übergreifende Lerntechniken und -strategien trainiert. Dazu stehen im Trainingsmanual Arbeitsblätter zur Verfügung, die genutzt werden können.

In einem kürzeren Trainingsblock sind zwei Termine zur Verbesserung der Lernorganisationen vorgesehen. Es werden gemeinsam mittels Modellierungsdialogen Bedingungen einer günstigen Lernumgebung herausgearbeitet. Anschließend wird die Umsetzung besserer Lernbedingungen für den Jugendlichen in Absprache mit Eltern und Lehrern vorbereitet. Bezogen auf den schulischen Bereich kann z. B. der Sitzplatz des Jugendlichen im Unterricht verändert werden, so dass er weniger abgelenkt wird. Für das häusliche Umfeld wird insbesondere diskutiert, wo der Jugendliche lernt und gegebenenfalls angeregt, dass ein eigener Schreibtisch in einem störungsfreien Umfeld bereitgestellt wird. Neben solchen die Situation strukturierenden Maßnahmen wird auch die Verbesserung der Selbstorganisation des Jugendlichen angestrebt. Ein Ziel ist die Optimierung der Terminplanung, z. B. die effektive und langfristige Vorbereiten auf Klassenarbeiten mit Hilfe von Zeitplänen.

Im letzten Trainingsblock besteht die Möglichkeit persönliche Probleme zu besprechen und Lösungen zu finden. Der Ablauf orientiert sich am Vorgehen der Selbstmanagementtherapie (Kanfer et al., 2006). Diese betont nach einer Aktualisierung der kooperativen Therapiebeziehung den Aufbau von Veränderungsmotivation, um aufbauend auf einer funktionalen Verhaltensanalyse gemeinsam Ziele und problemlöse Strategien zu formulieren. Gemeinsam wird ein Vorgehen überlegt, das nach der Durchführung reflektiert, evaluiert und gegebenenfalls verbessert wird. Trainer und Jugendliche arbeiten an aktuellen Problemen, so dass konkrete Problembewältigungserfahrungen gemacht werden können. In der Forschung von Grawe und Kollegen (2005) konnte gezeigt werden, dass auch die Aktualisierung und Bewältigung vom Problem als allgemeine therapeutische Wirkfaktoren eingeschätzt werden können. Insbesondere für ältere Jugendliche relevant sind zwei folgende Sitzungen zur Vorbereitung einer Berufsperspektive. Ausgehend von den Stärken und Interessen des Jugendlichen wird ein Eignungsprofil erarbeitet. Dieses Eignungsprofil wird mit den Anforderungsprofilen ausgewählter Berufsbilder verglichen. Der Begriff der Passung zwischen individuellen Stärken und Interessen und den Anforderungen verschiedener Berufsbilder wird herausgestellt. Es folgen weiterführende Überlegungen, wie der Prozess der Berufsvorbereitung weitergeführt werden kann, z. B. durch das Suchen eines Praktikumsplatzes. In der letzten Sitzung werden der Trainingsverlauf reflektiert, Erfolge rückgemeldet und eine Perspektive für die Zeit nach dem Training erarbeitet. Optional können Auffrischungstermine vereinbart werden, um eine langfristige Generalisierung der Trainingserfolge sicherzustellen. Dem Jugendlichen und auch den Eltern wird angeboten, sich auch bei zukünftigen Schwierigkeiten erneut zu melden. Tabelle 4 zeigt den Ablauf der Sitzungen mit jeweiligen Inhalten.

Tabelle 4: Ablauf des Lerntrainings für Jugendliche mit ADHS (LeJA)

Block	Sitzung	Inhalte
Beziehungsaufbau, Problem- und Zieldefinition	1	Kennenlernen, Beziehungsaufbau, Rollenklärung, Abschluss eines Trainingsvertrags
	2	Psychoedukation ADHS
	3	Psychoedukation Lernen und Lernstörungen
	4	Problem- und Zieldefinition erarbeiten, Ressourcen aktivieren
Lerntechniken und Lernstrategien	5-12	Lerntraining zur Verbesserung des Lern- und Arbeitsverhaltens, Erarbeitung und Training eines idealen Problemlöseprozesses mit Hilfe von Signalkarten
	13	Gedächtnisstrategien und spezielle Lerntechniken
Lernorganisation	14-15	Lernorganisation, Schaffung einer günstigen Lernumgebung
Coaching/ Selbstmanagement	16-17	Bearbeitung persönlicher Probleme (Eltern, Lehrer, Gleichaltrige, Gefühle, Motivation)
	18-19	Vorbereitung einer Berufsperspektive (Passung von individueller Eignung und Berufsprofilen)
	20	Abschluss, Reflexion, Besprechung der Perspektive nach dem Training, Generalisierung und Transfer von Trainingserfolgen
	Follow-up (optional)	Auffrischung

Anmerkungen: Die Sitzungen 13-19 können bei Bedarf vorgezogen werden. Die Sitzungen 18 und 19 sind vor allem für ältere Jugendliche relevant.

Der Ablauf der einzelnen Sitzungen ist bei jedem Termin gleich, um so eine Routine zu etablieren. Die Sitzungen à 50 Minuten laufen wie folgt ab:

- (1) *Blitzlicht*, in dem der Jugendliche kurz berichtet, was seit der letzten Sitzung geschehen ist, und wie es ihm geht. Liegen akute Probleme vor, werden diese kurz besprochen. (1 min)
- (2) *Kurze Wiederholung* der Inhalte der letzten Sitzung (2 min)
- (3) *Auswahl einer Aufgabe und Erarbeitung des Problemlösewegs* mittels Brainstorming und kognitiven Modellierungsdialogen (6 min)
- (4) *Übungsphase* mit verhaltensleitenden Instruktionen, prozessorientierten Hilfen und geeigneten kognitiven Techniken (30 min)

- (5) *Verhaltensfeedback*, in dem rein positiv gelungene Verhaltensweisen rückgemeldet und so sozial verstärkt werden (1 min)
- (6) *Spielerischer Ausklang* nach Wunsch des Jugendlichen zur Motivierung und Verbesserung der therapeutischen Beziehung (10 min)

Im Mittelpunkt des Trainings steht eine wiederholte, vertiefte Übungsbehandlung. Dabei soll anhand aktueller Schulaufgaben ein systematischer Stil der Aufgabenbearbeitung bzw. Problemlösung eingeübt werden. Dieses wiederholte Üben („explicit practice“) hat sich in Wirksamkeitsstudien als effektiv zur Förderung von Jugendlichen mit Lernstörungen herausgestellt (Swanson & Deshler, 2003). Nach dem Motto „so wenig helfen wie möglich und so viel wie notwendig“ wird das Prinzip des „Scaffolding“ umgesetzt (Wood, Bruner & Ross, 1976). Dabei baut der Trainer dem Jugendlichen ein mentales Gerüst, z. B. durch eine Vorstrukturierung und Komplexitätsreduktion der Aufgaben, durch handlungsleitende Fragen und durch eine Verstärkung richtiger Schritte der Aufgabenbearbeitung. So wird der Jugendliche dazu angeleitet, dass er selbst zur richtigen Lösung findet und eine selbstwertdienliche Attribution des Lernerfolges auf die eigene Fähigkeit und Anstrengung vornehmen kann (Weiner, 1985).

Bei der Gestaltung der Trainingssituation kommen insbesondere verhaltenstherapeutische Techniken zum Einsatz, die sich bei der Behandlung der ADHS im Kindesalter bewährt haben (Fabiano et al., 2009). Als operante Verfahren eignen sich bei Jugendlichen insbesondere die soziale Fremdverstärkung und die Selbstverstärkung (Linderkamp, 2007b; Weisz, Weiss, Han, Granger & Morton, 1995). Der Trainer verstärkt sozial während des Prozesses durch „Prompts“, also kurze verbale Bestätigungen die den Prozess nicht unterbrechen („mhm“, „super“, „weiter so“), bei gelungenen Verhaltensschritten des Jugendlichen und durch ein ausführliches rein positives Feedback nach Abschluss der Aufgabenbearbeitung. Dabei wird eine selbstwertdienliche Attribution unterstützt, indem der Trainer dem Jugendlichen gezeigte Stärken und Kompetenzen zurückgemeldet und ihm nahelegt, stolz auf diese zu sein. Die letzte der Signalkarten unterstützt zudem die Selbstverstärkung. Operante Verstärkung wirkt sich auch förderlich auf die Motivation der Jugendlichen aus (Linderkamp, 2004).

Auch kognitive Techniken der Verhaltenstherapie sind für das Training angemessen, da die Jugendlichen aufgrund ihrer fortgeschrittenen kognitiven Entwicklung vermutlich besser von diesen Techniken profitieren können als Kinder (Durlak, Fuhrman & Lampman, 1991). Bei Jugendlichen ist es u. a. besser möglich, dass das Verhaltensfeedback erst im Anschluss an die Aufgabenbearbeitung durchgeführt wird, es also eine verzögerte Kontingenz zwischen konkreter Verhaltensweise und Verstärkung gibt. Im Training werden mit den Jugendlichen gemeinsam Problemlösewege erarbeitet und diese nicht direktiv vorgegeben. Das Erarbeiten von Strategien und Plänen wird von sokratischen bzw. modellierenden Erkenntnisdialogen begleitet. Dabei arbeitet der Trainer Erkenntnisziele schrittweise mit dem Jugendlichen zusammen heraus. Er moderiert durch Paraphrasieren, gezieltes Fragen in Richtung des angestrebten Erkenntnisziels

und bezieht gegenläufige Äußerungen ein. So erlebt der Jugendliche das Ergebnis des Dialogs als stärker selbstbestimmt und kann das Vorgehen besser verstehen (Linderkamp, 2007a).

Eine weitere kognitive Technik ist das Einüben von Problemlösestrategien. Schon 1949 empfahl Pólya in seiner „Schule des Denkens“ die Einhaltung von vier Schritten zur Lösung von Mathematikaufgaben. Zur Beantwortung der Frage „Wie sucht man die Lösung“ solle zunächst die Aufgabe verstanden werden, dann müsse ein Plan der Aufgabebearbeitung gefasst werden, der darauf folgend kontrolliert ausgeführt werde, und abschließend erfolge eine Kontrolle des Ergebnisses (Pólya, 1949). Damit lieferte er eine Heuristik, die eine Grundlage für das erfolgreiche Lösen von Problemen bildet und nicht nur im mathematischen Bereich eingesetzt werden kann. Eine Anwendung auf alltägliche Probleme und Kombination mit verhaltenstherapeutischen Techniken wurde u. a. von D’Zurilla und Goldfried (1971) vorgeschlagen und auch im Programm „I can problem solve“ von Spivack und Shure (1974) zur Intervention bei sozialen und Verhaltensstörungen. Im Lerntraining LeJA erarbeiten Jugendliche und Trainer gemeinsam den idealen Problemlöseablauf, erstellen für jeden Schritt eine Signalkarte, die bei der Aufgabebearbeitung als Strukturierungshilfe genutzt wird.

Wichtig in der Trainingsgestaltung ist auch das Vorgehen bei der Instruktion bzw. Anleitung des Jugendlichen durch den Trainer. Empirische Studien haben gezeigt, dass zur Förderung bei Lernschwierigkeiten die Strategieinstruktion, die direkte Instruktion, die Selbstinstruktionen und deren Kombinationen besonders effektiv sind (Baer & Nietzel, 1991; Swanson & Hoskyn, 1998). Daher kommen sie im beim Lerntraining LeJA zum Einsatz. Bei der Strategieinstruktion werden bereichsübergreifender und bereichsspezifische Lernstrategien besprochen, modellhaft durch den Trainer demonstriert, dann anhand einfacher und später komplexerer Aufgaben trainiert und schließlich ihre Generalisierung vorbereitet. Bei der direkten Instruktion leitet der Trainer direkt das Verhalten des Jugendlichen an. Diese Instruktion bezieht sich im LeJA-Training weniger auf Inhalte, sondern stärker auf die Prozesse der Aufgabebearbeitung. Der Trainer „bremst“ bei einer vorschnellen Handlungsausführung (z. B. „Stopp – das musst du mir noch einmal erklären“) oder lenkt die Aufmerksamkeit auf kritische Punkte der Aufgabebearbeitung (z. B. „Lies dir die Aufgabenstellung noch einmal durch!“). Bei der Selbstinstruktion nutzt der Jugendliche die auf Grundlage des idealen Problemlöseprozesses erstellten Signalkarten, auf denen handlungsleitende Fragen (z. B. „Was ist meine Aufgabe?“) vermerkt sind. Diese werden zunächst laut geäußert und schrittweise immer mehr internalisiert. Die Kombination von kognitiver Modellierung und Selbstinstruktion ist insbesondere zur Verbesserung der Steuerung und Organisation des eigenen Handelns geeignet (Lauth, 2008).

Wie in den Behandlungsleitlinien empfohlen bekommen auch die Eltern Informationen zum Störungsbild der ADHS in Form einer Psychoedukation. Die Wirksamkeit von Psychoedukation bei ADHS wird in einem Review von Montoya, Colom und Ferrin (2011) dokumentiert. Die Autoren konnten sieben Studien identifizieren, davon waren

vier randomisiert und kontrolliert. Es wurden positive Effekte berichtet, z. B. verbesserte sich das Verhalten der Kinder und Jugendlichen, Eltern und Kindern waren zufriedener, wussten mehr über ADHS und hielten ärztliche Empfehlung zur Medikation eher ein. Speziell für das Jugendalter konnten McCleary und Ridley (1999) in einem Präpost-Design zeigen, dass die Teilnahme von Eltern an einer psychoedukativen Gruppe signifikant Häufigkeit und Intensität von Eltern-Adoleszenten-Konflikten und das Auftreten von Problemverhalten verringern konnte. Die Psychoedukation beim Lerntraining LeJA erfolgt entweder in Einzelkontakten oder in Form von Elternabenden mit anderen Eltern, deren Kinder mit dem Lerntraining behandelt werden. Insbesondere werden folgende Aspekte vermittelt: Die zentrale Problematik bei ADHS sind Defizite der Handlungsorganisation und Handlungssteuerung. Auch wenn es Anstrengung erfordert, sind die Probleme veränderbar und die Jugendlichen können eine Verbesserung durch das Training erzielen, wobei die Eltern sie unterstützen sollten. Dazu ist ein altersangemessenes und unterstützendes Erziehungsverhalten erforderlich. Dieses beinhaltet u. a. Strukturierung im Alltag, gemeinsam getroffene verbindliche Absprachen und Regeln, viel Lob, wenn etwas gut läuft usw. Die Lebensphase der Adoleszenz wird als eine Phase des Umbruchs und der Neueorientierung dargestellt, welche die Jugendlichen vor große Herausforderungen stellt und von einigen Jugendlichen krisenhaft erlebt wird. So soll eine Perspektivübernahme und ein verständnisvoller Umgang der Eltern mit ihren Kindern erleichtert werden. Bei schwerwiegenden Problemen werden auch Krisensitzungen mit beratenden Gesprächen.

Die Trainer nehmen auch Kontakt zu den Lehrern der Jugendlichen auf. Dies geschieht insbesondere telefonisch. Am Anfang fragen die Trainer nach der Einschätzung des Lehrers zu Lern- und Leistungsergebnissen und zum Verhalten des Jugendlichen im Unterricht. Weiterhin wird gefragt, ob aktuelle Schulleistungsaufgaben zur Verfügung gestellt werden können. Es wird angeboten, Informationsmaterial zur Psychoedukation in die Schule zu schicken, wenn dies erwünscht ist. Bei einer hohen Kooperationsbereitschaft der Lehrer können auch standardisierte Instrumente wie Beobachtungsbögen zum Unterrichtsverhalten eingesetzt werden. Es wird angestrebt, einen kontinuierlichen Kontakt mit den Lehrern zu halten und schulische Maßnahmen abzustimmen (z. B. günstiger Sitzplatz im Klassenraum). Dabei ist es wichtig, dem Jugendlichen gegenüber diese Lehrerkontakte transparent zu machen. Welcher Lehrer kontaktiert wird, wird zunächst gemeinsam besprochen. Es wird betont, dass es nicht um eine Kontrolle des Schülers gehe, sondern darum die Lehrer einzubinden, da sie die zentralen für die Schule sind, ohne deren Einbeziehung ein Lerntraining wenig Erfolg versprechend ist. Die Trainer versuchen weiterhin, bei etwaigen Konflikten zwischen Lehrern, Eltern und Schülern zu vermitteln. Die Lehrerkontakte sollen dienen der Generalisierung und dem Transfer der Trainingserfolge in die Schule. Auch Smith und Kollegen (2000) folgern in ihrem Review, dass Schulkontakte ein zentraler Baustein bei der Behandlung von ADHS im Jugendalter seien.

4 Evaluation psychologischer Interventionen

Ein zentrales Erkenntnisinteresse der Erforschung psychologischer Interventionen ist die Klärung ihrer Wirksamkeit. Dahinter verbirgt sich die Frage, ob eine therapeutische Intervention überhaupt wirkt und wenn ja, in welcher Art und Weise. Das Hauptziel der Wirksamkeitsforschung ist damit die genaue Analyse therapeutischer Veränderungsprozesse, um durch die Nutzung der gewonnenen Erkenntnisse einen Beitrag zur Verbesserung und Neuentwicklung wirksamer Verfahren beizusteuern. Die Erforschung bleibt dabei nicht bei einer globalen Bewertung stehen, sondern bezieht Erkenntnisse zur differentiellen Wirksamkeit und Indikation in die Analyse ein, um so ein immer genaueres Bild davon zu gewinnen, welches Verfahren, unter welchen Bedingungen zu welchem Ergebnis für welche Person führt (Lutz & Grawe, 2007).

Dieses Kapitel beginnt mit einem kurzen Abriss zur Tradition der Wirksamkeitsforschung in der Psychotherapie. Es folgen Überlegungen dazu, wie der Erfolg einer Intervention zu definieren und zu messen ist. Anschließend wird der Themenkomplex Veränderungsmessung dargestellt, da diese eine Grundlage des weiteren methodischen Vorgehens bildet. Im Anschluss wird das Konzept der differentiellen Wirksamkeit genauer beleuchtet. Den Abschluss dieses Kapitels bildet die Darstellung empirischer Ergebnisse zur differentiellen Wirksamkeit, von denen im folgenden Kapitel Hypothesen abgeleitet.

4.1 Psychotherapieforschung

Einen großen Beitrag zur Erforschung der Wirksamkeit von Psychotherapien hat Grawe erbracht. Grawe (1992, 1997) unterscheidet vier Phasen psychologischer Interventionsforschung, die sich in ihrer schwerpunktmäßigen Zielsetzung voneinander unterscheiden:

- (1) *Legitimation*: Wirkt Psychotherapie überhaupt?
- (2) *Wettbewerb*: Welche Form der Psychotherapie ist besser oder am besten?
- (3) *Verschreibung*: Welche Form der Psychotherapie ist bei wem indiziert?
- (4) *Prozessforschung*: Auf welche Weise wirkt Psychotherapie?

Die ersten drei Phasen folgen mit Überlappungen chronologisch aufeinander und sind zur Evaluation bestehender Verfahren ausgelegt. Die vierte Phase der Prozessforschung ist nicht als Ablösung der vorangegangenen Phasen zu verstehen, sondern ergänzt diese und ermöglicht als einzige die Entwicklung neuer Verfahren.

Die erste Frage, der sich die Psychotherapie erwehren musste, war diejenige nach ihrer generellen Wirksamkeit. In dieser Phase der Legitimation stellte sich daher die Frage, ob Psychotherapie überhaupt eine Wirkung und einen Nutzen hat. In einer viel disku-

tierten Studie zur Behandlung neurotischer Patienten kam Eysenck (1952) zu dem Schluss, dass die psychoanalytische Behandlung die Effekte der Spontanremission unbehandelter Patienten nicht übertreffen konnte. Eysenck betont den Bedarf nach gut geplanten und durchgeführten experimentellen Studien, die sich von den damals üblichen kasuistischen Fallberichten mit oftmals unsystematischer Erhebung und Darstellung der Daten abgrenzten.

Die daraufhin zunehmende Überprüfung von Psychotherapie in experimentellen Studien führte zu einer Phase des Wettbewerbs zwischen verschiedenen Therapieformen, insbesondere der Psychoanalyse und der Verhaltenstherapie. Die zentrale Frage war, welche Therapieform einer anderen überlegen war und welche schließlich die beste Therapieform sei. Durch besser kontrollierte Studien konnte zwar der allgemeine Nachweis für die Wirksamkeit von Psychotherapie im Allgemeinen plausibel geführt werden, aber es zeichnete sich kein Sieger im Wettstreit der verschiedenen therapeutischen Schulen ab. Es ergaben sich ähnliche Wirksamkeiten und Endergebnisse der verschiedenen Therapieerichtungen und es wurden keine eindeutigen „Verlierer“ gefunden. Dieser Befund wurde als „Äquivalenz-Paradox“ (Meyer, 1990) bezeichnet oder prosaisch ausgedrückt als „Dodo-Verdikt“ nach Lewis Carrolls „Alice in Wonderland“, getreu dem Motto „Everyone has won and all must have prizes“ (Luborsky, Singer & Luborsky, 1975). Dass im Therapievergleich keine Unterschiede gefunden werden konnten, mag einerseits damit zusammenhängen, dass die untersuchten Stichproben zu klein waren, um Effekte statistisch signifikant absichern zu können, oder kann als ein Hinweis auf die Wirkung grundlegende Faktoren interpretiert werden, die in allen therapeutischen Behandlungen realisiert wurden (Grawe, 1992).

Nachdem sich die Psychotherapie allgemein als wirksam behauptet hatte und der Wettstreit der therapeutischen Richtungen nicht entschieden werden konnte, rückte die behandelte Person als Einheit der Differenzierung in den Fokus des Interesses. Ziel dieser Perspektive ist die Vereinfachung und Verbesserung von Entscheidungen darüber, welche Therapie für welchen Klienten am besten geeignet und damit indiziert ist. Analog zur Verschreibung von Medikamenten sollen somit spezifischen Problemlagen und Störungen geeignete Interventionen zugewiesen werden. Vorläufer dieser Phase waren einfache noch an den therapeutischen Schulen verhaftete Verschreibungsschemata mit etwa folgenden Empfehlungen: „Leichte Fälle für die Verhandlungstherapie, schwere für tiefenpsychologische Therapie“ (Pongratz, 1973, S. 378; zitiert nach Grawe, 1992). Nachfolgende Untersuchungen setzten – oftmals ohne theoretisch fundierte Vorannahmen – verschiedene Merkmale der behandelten Patienten in Zusammenhang mit dem Therapieerfolg. Dieses Vorgehen führte zu einer Vielzahl an signifikanten Zusammenhängen, die zumeist nur schwerlich in ein konsistentes Modell integriert werden können, um zu begründen, warum eine Art der Intervention für bestimmte Personen besonders geeignet oder ungeeignet sei (Grawe, 1992). Methodisches Vorgehen ist zumeist die Durchführung von Post-hoc-Analysen von Interventionsstudien, die den Zusammenhang von Klientenvariablen mit dem Trainingserfolg untersuchen. Problematisch erscheint dabei insbesondere, dass Klientenvariablen, die vor der Therapie erhoben

werden, nicht isoliert und konstant wirken. Es ist von weitgehenden Wechselwirkungen auszugehen, die kaum kontrolliert überprüft werden können. Neben Interaktionen der verschiedenen Klientenvariablen untereinander treten Merkmale des Klienten mit denen des Therapeuten in Wechselwirkung und interagieren mit anderen Prozessmerkmalen, so dass sich bis heute wenig konsistente und klare Zusammenhänge nachweisen lassen (Clarkin & Levy, 2004).

Parallel zu den drei ersten Phasen, welche die Wirksamkeit von Therapien als Ganzes überprüfen, widmet sich die Prozessforschung bzw. Prozess-Outcomeforschung der Frage, wie eine Therapie wirkt, also welche Inhalte und Prozesse zu einem Erfolg oder Misserfolg beitragen. Ziel ist die Verbesserung bestehender bzw. Entwicklung neuer psychotherapeutischer Verfahren. Dabei fließen immer mehr naturwissenschaftliche Erkenntnisse in die Entwicklung neuer Interventionen ein und es werden auf Grundlage bestehender Erkenntnisse neue Verfahren entwickelt, die sich nicht mehr unbedingt auf eine Verortung in den traditionellen Therapieschulen berufen (Jacobi, 2006).

Im Fokus der vorliegenden Studie steht die differentielle Wirksamkeit eines manualisierten, standardisiert durchgeführten Trainings. Damit bewegt sich dieser Teil der Untersuchung im Rahmen der Phase der Verschreibung und verfolgt das Ziel, Hinweise dazu zu geben, für welche Subgruppen das Training besonders geeignet ist, und Implikationen zur Optimierung des Trainings für schlechter profitierenden Subgruppen abzuleiten. Um dabei nicht der Gefahr des willkürlichen in Beziehung Setzens von Klientenvariablen und Outcomes zu erliegen, die von Grawe (1992) als „Herumstochern in einem Heuhaufen“ verbildlicht wird, wird wenn möglich Theorie und Hypothesen geleitet vorgegangen.

Zur Evaluation therapeutischer Interventionen, sowohl bei neu entwickelten als auch bei alt bekannten, die aber nicht systematisch überprüft worden sind, wird zunächst die allgemeine oder globale Wirksamkeit überprüft. Vorgehen der Wahl ist dabei der Vergleich relevanter Erfolgskriterien vor und nach der Intervention und zumeist einige Zeit später, als Follow-up zur Überprüfung der Langzeiteffekte. Damit kann gezeigt werden, ob die Intervention im Mittel überhaupt zu einer Verbesserung führt und ob diese nachhaltig ist. Wird eine Verbesserung festgestellt, ist allerdings nicht bewiesen, dass diese nicht auch ohne die Durchführung der Intervention eingetreten wäre, bei Kindern und Jugendlichen z. B. ausgelöst durch altersgemäße Reifung. Um dieses auszuschließen werden Vergleichsgruppen mit ähnlichen Zusammensetzungen als Kontrollbedingungen eingeführt, wobei die Zuordnung zu den Gruppen randomisiert oder parallelisiert erfolgt. Am einfachsten realisierbar ist dabei die Wartegruppe, die keinerlei Intervention bekommt und bei der lediglich zu den Zeiten vor und nach dem Training der Interventionsgruppe die Erfolgskriterien erhoben werden. Zeigt die Interventionsgruppe im Vergleich eine stärkere Verbesserung als die Wartekontrollgruppe, kann geschlussfolgert werden, dass die Intervention dem reinen Abwarten überlegen ist und anderen Effekten, z. B. Reifung, überlegen ist. Nach diesem Überprüfungsschritt ist allerdings immer noch nicht auszuschließen, dass nicht das spezifische Vorgehen, also einer therapeutischen

Richtung und Theorie abgeleiteten Intervention den Effekt herbeigeführt habe, sondern sog. unspezifische Trainingseffekte wie die bloße Zuwendung durch eine Person mit einem offenen Ohr für die individuelle Situation und Problemlage. Um dieses wiederum auszuschließen werden Gruppen eingeführt, die entweder eine sog. unspezifische Behandlung erfahren, also eine Intervention, von der nicht erwartet wird, dass sie sich auf den Interventionserfolg auswirkt, oder eine alternative Behandlung, z. B. die bislang eingesetzte Standardmethode, deren Unterlegenheit im Vergleich zur neuen Behandlung gezeigt werden soll. Erst wenn die zu überprüfende Intervention nachweislich zu einer klinisch relevanten Verbesserung im Vorher-nachher-Vergleich führt und sie den Kontrollbedingungen signifikant überlegen ist, kann ihre Wirksamkeit als relativ sicher bewiesen gelten. Das Vorgehen die Wirksamkeit in einem Vergleichsdesign mit mindestens einer Kontrollgruppe zu überprüfen, wobei die Zuordnung zur Interventions- oder Kontrollgruppe zufällig geschieht, wird als „randomized clinical trials“ (Kraemer, 2003) oder „randomized controlled trials“ (Moher, Schulz & Altman, 2004) bezeichnet und als RCT abgekürzt.

4.2 Operationalisierung von Interventionserfolg

Um zu entscheiden, ob eine Intervention erfolgreich war, muss definiert werden, was als Erfolg verstanden wird, es ist also eine Operationalisierung des Konstrukts Interventionserfolg notwendig. Im Kontext der Psychotherapieforschung besteht weitestgehend Konsens, dass Therapieerfolg multimodal und multimethodal erfasst werden soll (Jacobi, 2006). Laut Schulte (1993) ist wie bei jedem anderen psychologischen Konstrukt zunächst Theorie geleitet festzulegen, mit welchen Indizes oder Variablen das Konstrukt gemessen werden soll und anschließend wie diese Verfahren operationalisiert und gemessen werden können. Schulte (1993) schlägt eine inhaltliche Operationalisierung des Therapieerfolgs auf drei Ebenen vor: 1. Ebene der Krankheitsursachen oder der „Defekte“, 2. auf der Ebene der Symptomatik und 3. auf der Ebene der Krankenrolle bzw. der Krankheitsfolgen im Sinne der „Einschränkung normaler Rollen“.

Bezogen auf Psychotherapie für Kinder und Jugendliche unterscheidet Kazdin (2004, S. 570) drei Bereiche relevanter Outcome-Kriterien: 1. child functioning, 2. parent and family functioning und 3. social impact measures, die je passend zum jeweiligen Problem oder Störung relevant seien. Der Bereich child functioning wird dabei noch unterteilt in symptoms, impairment, prosocial competence und academic functioning. Die funktionellen, auf das Kind bezogenen Maße können weiter differenziert werden in den Ebenen Kognition, Emotion und Verhalten.

Ebenfalls relevant, insbesondere im Kindes- und Jugendalter, ist die Bewältigung anstehender Entwicklungsaufgaben. Psychische Störungen in dieser Altersspanne werden oftmals gerade als eine nicht adäquate Bewältigung altersgemäßer Entwicklungsaufgaben definiert (Heinrichs & Lohaus, 2011). Demnach kann Therapieerfolg auch als Ausmaß der Bewältigung anstehender Entwicklungsaufgaben konzeptualisiert werden.

Zusammenfassend werden die folgenden Dimensionen für die Operationalisierung einer lernfördernden Intervention für Kinder und Jugendliche als relevant bewertet, wobei diese Faktoren nicht unabhängig voneinander sind:

- *Symptomverringering*
- *Besserer Umgang / Coping mit der Problematik*
- *Verbesserung kognitiver Leistungsmaße*
- *Verbesserung akademischer Leistungen*
- *Verringerung der Beeinträchtigung in der Partizipation bzw. in der Übernahme normativer Rollen*
- *Verringerung der allgemeinen Problem- / Stressbelastung*
- *Gesteigertes subjektives Wohlbefinden*
- *Zufriedenheit mit der Intervention*
- *Positive Einschätzung der Ergebnisse der Maßnahme durch relevante Bezugspersonen (v. a. Eltern und Lehrer)*

So differenziert und informativ eine Vielzahl an Outcome-Kriterien ist, so birgt ein Vorgehen mit vielen Erfolgsmaßen auch Probleme. Es ist davon auszugehen, dass unterschiedliche Instrumente unterschiedliche Ergebnisse liefern, die sich teilweise widersprechen. In einem solchen Fall ist schwierig zu entscheiden, welchen Instrumenten mehr zu trauen ist oder wie eine zusammenfassende Abwägung geschehen kann. Die Frage, ob ein Training insgesamt wirksam war oder nicht, lässt sich nur schwerlich beantworten. Einige Autoren empfehlen daher nur eine einzelne Outcome-Variable festzulegen, um die primäre Forschungsfrage zu beantworten (Friedman, Furberg & DeMets, 2010).

Zweben und Cisler (1996) führen die Problematik für den Bereich der Behandlung bei Alkoholabhängigkeit aus und konstatieren „a major dilemma for clinical researchers is interpreting disparities among different outcome measures and the stability or instability of these outcomes over time“ (S. 1784). Die Autoren beschreiben in einem Überblick, dass zumeist Einteilungen des Zustandes nach einer Intervention auf einem Kontinuum mit abgestuften Kategorien vorgenommen werden. Die Pole des Kontinuums stellen demnach Problemfreiheit und damit Interventionserfolg auf der positiven Seite und Rückfall bzw. problemhaftes Verhalten und damit Misserfolg der Intervention auf der negativen Seite dar.

In der medizinischen Forschung werden im Rahmen von klinischen Studien teilweise zusammengesetzte Erfolgsmaße („composite outcomes“) bestimmt. Als deren Vorteile

werden u. a. bessere Chancen statistische Signifikanz zu erreichen, eine Abschätzung des Nettoeffekts und eine einfachere Umsetzbarkeit genannt. Probleme bestehen darin, die Ergebnisse hinsichtlich ihrer klinischer Signifikanz zu interpretieren, es fehlt teilweise an Relevanz für die Patienten und die gleichwertige Behandlung aller eingehenden Komponenten mag fragwürdig erscheinen (Heddle & Cook, 2011).

Im Bereich der medizinischen Forschung sind mit mehreren Outcomes allerdings zu meist unterschiedliche Ereignisse wie Hospitalisierung, Herzinfarkt oder Tod gemeint. Diese werden auch als „endpoints“ bezeichnet und es handelt sich folglich nicht um kontinuierliche sondern dichotome Variablen, die angeben, ob ein bestimmtes kritisches Ergebnis eingetreten ist oder nicht (Friedman et al., 2010). In diesem Zusammenhang erscheint es folglich fragwürdig, ob einzelne Komponenten sinnvoll zusammengefasst werden können. Anders verhält es sich bei der Zusammenfassung mehrerer kontinuierlicher Variablen, die zusammengenommen das Konstrukt Gesamttrainingserfolg konstituieren.

Auch Conners und Kollegen (2001) nennen Gründe für die Vorzüge eines einzelnen Outcome-Maßes. Zunächst sei es statistisch betrachtet sehr viel schwieriger, multiple Effekte nachzuweisen. Dies wurde in der ursprünglichen Auswertung der MTA-Studie mit 579 Kindern mit ADHS in vier unterschiedlichen Behandlungsgruppen versucht (Jensen, Arnold, Richters, Severe, Vereen, Vitiello, Schiller, Hinshaw, Elliott, Conners, Wells, March, Swanson, Wigal, Cantwell, Abikoff, Hechtman, Greenhill, Newcorn, Pelham, Hoza, Kraemer et al., 1999). Für jeden Outcome-Bereich (ADHS-Symptome, Aggressivität, Internalisierende Symptome, Sozialverhalten, Schulleistungen) wurden Regressionsanalysen durchgeführt, in die Durchführungsort, Zeit und Behandlungsgruppe und deren Interaktionen eingingen. Da sich jeder Outcome-Bereich aus mehreren Einzelmessungen zusammensetzte, wurden Bonferroni-Korrekturen durchgeführt. Die Signifikanz im Bereich ADHS-Symptomatik, der mittels fünf Einzelmessungen erfasst wurde, wurde von konventionell $p < 0.05$ auf $p < 0.01$ gesenkt. Im Anschluss wurde bei signifikanten Tests, die sich zunächst auf alle vier Gruppen bezogen, paarweise Vergleiche einzelner Behandlungsgruppen durchgeführt, u. a. reine Medikation vs. Kombinationstherapie. Da sechs solcher Paarvergleiche durchgeführt wurden, wurde das Signifikanzniveau weiter reduziert; im Bereich ADHS-Symptome von $p < 0.01$ auf $p < 0.0017$ ($0.01 \div 6$). Durch dieses Vorgehen konnte einer Alphafehler-Kumulation zwar entgegengewirkt werden, es wurde aber sehr schwer, überhaupt noch signifikante Unterschiede nachzuweisen. Dies führte auch dazu, dass die erwartete Überlegenheit der Kombinationsbehandlung vor der reinen Medikation nicht statistisch nachweisbar war. Irrtümlicherweise wurde dies so interpretiert, dass eine Kombinationsbehandlung keine Vorteile gegenüber der reinen Medikation biete. Durch eine Re-Kalkulation der Daten mit einem aus den Einzelmessungen ermittelten Indexwertes als einziger Outcome-Messung zeigte sich ein statistisch signifikante Überlegenheit der Kombinationstherapie im Vergleich zur reinen Medikation mit einer kleinen Effektgröße (Conners et al., 2001).

Auch inhaltlich bereitet die Interpretation unterschiedlicher Outcome-Maße Schwierigkeiten, z. B. wenn einige Maße eine Verbesserung abbilden und andere nicht. Dies kann damit zusammenhängen, dass die Intervention auf einigen Ebenen wirkt und auf anderen nicht oder dass sich eine Störung im zeitlichen Verlauf unterschiedlich äußert, z. B. nach Tagesform, oder in Abhängigkeit vom jeweiligen Beobachter. Eine solche Variabilität scheint gerade für Kinder mit ADHS typisch zu sein. Ein Gesamtmaß kann somit reliabler sein als die Einzelmessungen, aus denen es zusammengesetzt ist. Viele einzeln zu kleine Effekte, die aber relevant sind, können zusammengenommen eine Auswirkung auf das Gesamtbild abgeben und zufällige Fehler einzelner Messungen werden besser ausgeglichen (Conners et al., 2001).

Vorteile eines einzelnen Outcome-Maßes liegen damit sowohl inhaltlich in der besseren Interpretierbarkeit als auch statistisch begründet in einer größeren Sensitivität für das Aufspüren signifikanter Effekte und einer vermeintlich höheren Reliabilität der Messung, die allerdings zu überprüfen ist. Je nach Art der Fragestellung scheint eine Kombination von einzelnen und einer kombinierten Messung sinnvoll zu sein. Bei einer rein deskriptiven Auswertung bietet die Darstellung auf möglichst vielen unterschiedlichen Ebenen vermutlich den größeren Informationsgehalt, während zu einer generellen Hypothesenprüfung ein Gesamtmaß nützlicher sein kann.

4.3 Veränderungsmessung

Im Kontext der Erfassung des Interventionserfolges spielt die Messung von Veränderung eine zentrale Rolle, z. B. zur Einschätzung des Ausmaßes der Reduktion störungsbezogener Symptome. Nach der Festlegung geeigneter Konstrukte und deren Operationalisierung zur Erfassung stellt sich daher die Frage, wie diese Veränderung gemessen werden soll. Prinzipiell stehen zwei Möglichkeiten der Veränderungsmessung zur Verfügung, die direkte und die indirekte Messung. Bei der direkten Veränderungsmessung werden die Patienten gebeten, die Veränderung im Anschluss an die Behandlung retrospektiv einzuschätzen. Bei der indirekten Veränderungsmessung wird dagegen das relevante Konstrukt (z. B. Symptomstärke) zu mindestens zwei Messzeitpunkten gemessen und aus dem Vergleich der Messwerte auf die Veränderung geschlossen (Michalak, Kosfelder, Meyer & Schulte, 2003).

Während die direkte Veränderungsmessung aufgrund nur eines einzigen Messzeitpunktes einfacher realisierbar ist und das subjektive Erleben der Veränderung gut abbildet, ist sie anfällig gegenüber Verzerrungseffekten und hat einen hohen Aufforderungscharakter, der u. a. soziale Erwünschtheit fördert. Dementgegen ist die indirekte Veränderungsmessung aufwändiger und unterliegt stärker messtechnischen Problemen, z. B. dem Regressionseffekt oder Übungseffekte durch wiederholte Testung (Gollwitzer & Jäger, 2009). Direkte und indirekte Veränderungsmessung können demnach unterschiedliche Ergebnisse liefern.

Als Vorgriff sei an dieser Stelle erwähnt, dass die indirekte Veränderungsmessung im Messkonzept der klassischen Testtheorie, auf der die meisten Testverfahren beruhen, eine inhärente Problematik hinsichtlich der angemessenen Konzeptualisierung von Veränderungsmessung unterliegt. Diese Problematik wird im Folgenden genauer erläutert. Neben diesen Schwierigkeiten spricht die „Nützlichkeit, Brauchbarkeit und die breite Anwendung (...) für den Einsatz einer indirekten Veränderungsmessung (Lutz & Grawe, 2007, S. 739).

4.3.1 Verzerrungseffekte bei Messwiederholungen

Bei der indirekten Veränderungsmessung wird von der Differenz der Messwerte zu zwei Zeitpunkten auf eine zugrunde liegende Veränderung geschlossen. Die Messungen sollen dabei unabhängig voneinander sein und nur das Konstrukt erfassen. Es kann aber auch sein, dass sich die erste Messung auf die zweite Messung auswirkt und so Verzerrungen auftreten. Diese sind nach Gollwitzer und Jäger (2009):

- *Testübung* insbesondere bei Leistungstests
- *Erinnerungseffekte* wenn die Probanden bei Selbstauskünften noch wissen, was sie bei der letzten Messung angegeben haben
- *Zwischenzeitliches Geschehen*, also Ereignisse, die nichts mit der Intervention zu tun haben, sich aber dennoch auf das Antwortverhalten der Probanden auswirken
- *Response Shift*, als veränderte Einstellung zu einem erfassten Merkmal ohne eine Veränderung des Merkmals an sich (z. B. wenn das Merkmal Vergesslichkeit zunächst als problematisch eingestuft wird, später aber nicht mehr)
- *Sensibilisierung* in der Art, dass die Befragten ahnen, welche Antwort von Ihnen erwartet wird

Die Effekte können zumindest teilweise durch die Erstellung des experimentellen Designs oder durch die Gütekriterien der Messinstrumente entschärft werden. So kann über die Maße der Reliabilität eines Leistungstests eruiert werden, ob Retesteffekte zu erwarten sind. Oftmals finden sich auch Empfehlungen der Autoren zu mindestens einzuhaltenen zeitlichen Intervallen zwischen zwei Messungen. Durch ausreichende lange zeitliche Abstände zwischen zwei Messungen können auch Erinnerungseffekte verringert werden. Es steigt allerdings die Wahrscheinlichkeit von response shift, also einer veränderten Bewertung bestimmter Merkmale, und dem Wirken zwischenzeitlicher Geschehen. Diesen Problemen kann durch die Umsetzung experimenteller Designs mit Kontrollgruppen, die dieselben Messungen durchlaufen, entgegengewirkt werden. Wird die Gruppenzuteilung randomisiert vorgenommen und sind die Gruppen ausreichend groß, wird davon ausgegangen, dass sich unspezifische Effekte wie zwischenzeitliche Ereignisse zufällig auf die Gruppen verteilen und somit identifizierbar sind. Schwierig ist die Kontrolle der Effekte von Sensibilisierung für ein Evaluationsinstrument, wel-

ches mit dem Effekte der sozialen Erwünschtheit zusammenhängt. Teilnehmer einer Intervention erwarten in der Regel, dass eine Verbesserung stattgefunden hat, und Teilnehmer einer reinen Wartegruppe erwarten, dass sich nichts verändert, zunächst unabhängig davon, ob sich tatsächlich etwas verändert hat. Durch die Einführung einer alternativen Behandlungsgruppe können die Effekte zumindest teilweise behoben werden, es bleibt aber fraglich, ob mit der alternativen Behandlung tatsächlich die gleichen Erwartungen verknüpft sind wie mit der eigentlichen Intervention.

4.3.2 Reliabilität von Differenzwerten

Neben den genannten Gefährdungen, die durch das Verhalten der Probanden entstehen, bestehen auch Schwierigkeiten messtechnischer Art bei der Auswertung von wiederholten Messungen. Eine grundsätzliche Problematik offenbart sich in der Einschätzung des Gütekriteriums der Reliabilität, welches Aufschluss über den Grad der Genauigkeit bzw. Präzision angibt, mit dem ein Instrument ein Merkmal erfasst (Bortz & Döring, 2006, S. 196). Das einfachste Maß zur Schätzung einer „wahren“ Vorher-nachher-Veränderung ist die Differenz zweier Testwerte desselben Testinstruments zu zwei aufeinander folgenden Messzeitpunkten, z. B. vor und nach einer Intervention. Dieses einfache Differenzmaß ist in der Vergangenheit oftmals kritisiert worden, insbesondere deshalb, weil die Reliabilität der Differenzwerte als ungenügend beurteilt wurde (Cronbach & Furby, 1970; Linn & Slinde, 1977). Die Diskussion zu diesem Thema und eine abschließende Bewertung werden im Anschluss dargestellt.

Gemäß dem ersten Axiom der klassischen Testtheorie setzt sich ein Testergebnis additiv aus einem „wahren Wert“ und einem Messfehler zusammen. Der Messfehler resultiert aus verfälschenden situativen Einflüssen wie z. B. Müdigkeit des Probanden während der Testdurchführung. Je geringer der Messfehler in das Ergebnis einer Messung eingeht, desto höher ist die Reliabilität eines Instrumentes, wobei eine perfekte Reliabilität im Rahmen der psychologischen Testung praktisch nie erreicht werden kann. Die Abweichungen zweier nachfolgender Messungen werden als Folge der Wirkung unsystematischer Messfehler verstanden und bestimmen die sogenannte Fehlervarianz (Bortz & Döring, 2006).

Zur Bestimmung der Reliabilität wird oftmals die Überprüfung der Stabilität herangezogen, um so die Retestrelabilität der Messwerte zu ermitteln. Dazu wird der zu überprüfende Test mit denselben Probanden zweimal aufeinander folgend durchgeführt, mit einem ausreichend großen Abstand zwischen den Testzeitpunkten, so dass keine Erinnerungseffekte zu erwarten sind. Die Korrelation der Messwerte wird als Retestrelabilität definiert. Ein perfekt reliabler Test müsste bei zeitstabilen Merkmalen zweimal exakt dasselbe Ergebnis liefern. Dabei wäre der Anteil der „wahren Varianz“ (S_T^2) an der beobachteten Varianz (S_T^2) mit 1 bzw. 100 % gleichzusetzen und der Anteil der Fehlervarianz (S_E^2) mit 0 bzw. 0 %. Da eine perfekte Korrelation praktisch nicht auftritt, wird der Wert der Korrelation als Anteil der „wahren Varianz“ definiert und die Differenz zu 1 bzw. zu 100 % als Anteil der Fehlervarianz. Die Reliabilität kann also als Anteil der

„wahren Varianz“ an der beobachteten Varianz verstanden werden wie in folgender Formel nach Bortz und Döring (2006, S. 196) veranschaulicht:

$$Rel = \frac{S_T^2}{S_X^2} = \frac{S_T^2}{S_T^2 + S_E^2}$$

Problematisch ist die Bestimmung der Retestreliabilität bei der Messung von Merkmalen, die nicht zeitstabil sind und sich im Laufe der Zeit verändern. Dies können z. B. affektive oder reifungsbedingte Zustände sein oder konkrete Verhaltensweisen, die von dem situativen Kontext abhängen.

Wenn bei einer Wiederholungsmessung Veränderungen auftreten, muss unterschieden werden, ob diese auf die beabsichtigten Veränderungen im Verhalten, also auf beabsichtigte Veränderungen der „wahren Varianz“, zurückzuführen sind oder ob es sich um Zufallsfluktuationen handelt. Dabei ist die Intention der Retestreliabilität eine hohe Korrelation aufeinander folgender Messungen zu erreichen, also zeitstabile Messwerte zu liefern, während die Differenzwertmessung zu Evaluationszwecken gerade Unterschiede und Veränderungen in den Testscores abzubilden sucht. Prinzipiell müsste ein Messinstrument geschaffen werden, dessen Retestreliabilität niedrig und eine andere Reliabilitätsschätzung (z. B. Paralleltestreliabilität) hoch ausfällt. Im Rahmen der klassischen Testtheorie scheint dies aber nicht möglich zu sein, da beide Reliabilitätsschätzungen per Definition das gleiche Konstrukt erfassen sollen (Petermann, 1978, S. 32).

Eine im Zusammenhang mit der Reliabilitätsschätzung von Differenzwerten diskutierte Problematik offenbart sich im sogenannten Reliabilitäts-Validitäts-Dilemma bzw. unreliability-invalidity dilemma (Bereiter, 1963). Demnach besteht nach den Annahmen der klassischen Testtheorie ein gegenläufiger Zusammenhang zwischen der Reliabilität von Differenzwerten und der Korrelation zweier Testwerte, z. B. einer Vorher-nachher-Korrelation in einer Prä-post-Erhebung mit demselben Messinstrument. Dabei gilt: 1. Je höher die Differenzwert-Reliabilität, desto geringer ist die Vorher-nachher-Korrelation und 2. Je höher die Vorher-nachher-Korrelation, desto geringer ist die Differenzwertreliabilität. Im ersten Fall, in dem die Differenzwert-Reliabilität hoch ist, scheint zwar die Veränderung zuverlässig abgebildet zu sein, die geringe Autokorrelation führt jedoch zu Zweifeln an der Validität der Einzelmessungen, da anscheinend nicht das Gleiche gemessen wurde. Im zweiten Fall scheint zwar das Gleiche gemessen worden zu sein bzw. die gleichen Faktoren scheinen sich auf die Messwerte ausgewirkt zu haben, aber die Veränderung ist nicht reliabel abgebildet. Nach dem Reliabilitäts-Validitäts-Dilemma ist es demnach nicht möglich, gleichzeitig eine hohe Validität der Einzelwerte und eine hohe Reliabilität der Differenzwerte herzustellen.

Veranschaulichen lässt sich dieser Zusammenhang anhand folgender Formel zur Bestimmung der Reliabilität von Differenzwerten nach Lienert und Raatz (1998, S. 215):

$$r_{(1-2)(1-2)} = \frac{r_{11} - r_{12}}{1 - r_{12}}$$

Dabei steht $r_{(1-2)(1-2)}$ für die Reliabilität der Differenz zwischen zwei Testwerten, r_{11} für die Reliabilität der Rohwerte, ausgedrückt als Anteil der wahren Varianz an der Gesamtvarianz, und r_{12} für die Korrelation zwischen den beiden Testwerten. Bei der Herleitung der Formel wird angenommen, dass die Reliabilitäten der Einzeltests vorher und nachher gleich sind, da derselbe Test zweimal durchgeführt wird, und dass die Varianzen zu beiden Messzeitpunkten gleich sind.

Veranschaulicht man diesen Zusammenhang als $y = f(x) = (a-x) / (1-x)$ mit a als Parameter mit Werten zwischen 0.1 und 0.9 mit Abstufungen im Abstand 0.1, ergibt sich eine hyperbolische Kurvenschar, die im für die Reliabilitätskoeffizienten relevanten Definitions- und Wertebereich von 0 bis 1 in Abbildung 16 dargestellt ist.

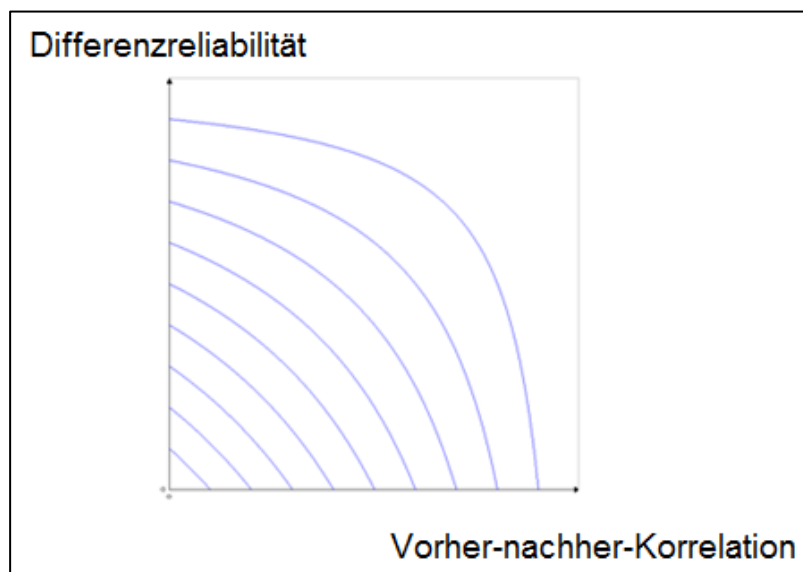


Abbildung 16: Grafische Veranschaulichung des Reliabilitäts-Validitäts-Dilemmas

Auf der Abszissenachse finden sich die Werte für die Vorher-nachher-Korrelation und auf der Ordinatenachse die Werte für die Differenzreliabilität. Der Kurvenausschnitt unten links gehört zum Parameter $a = 0.1$, die weiteren Kurven aufsteigend in x - bzw. y -Richtung jeweils zu einem folgenden Parameter im Abstand von 0.1.

Der im Reliabilitäts-Validitäts-Dilemma beschriebene Zusammenhang wird insofern deutlich, als höhere x -Werte (Vorher-nachher-Korrelation) mit niedrigeren y -Werten (Differenzreliabilität) einhergehen und andersherum. Es wird weiterhin deutlich, dass die Reliabilität der Differenzwerte immer geringer ist als die Reliabilität der Einzelwerte, so liegt beispielsweise die Differenzreliabilität des obersten Kurvenausschnitts im-

mer unter dem zugehörigen Parameter von $a = 0.9$, also dem Wert für die Reliabilität einer Einzelmessung. Schließlich wird deutlich, dass eine hohe Differenzwertreliabilität überhaupt nur bei einer hohen Reliabilität der Einzelwerte (ausgedrückt durch Parameter a) erreicht werden kann.

Das Reliabilitäts-Validitäts-Dilemma könne nach Petermann (1979) im Rahmen der klassischen Testtheorie nicht gelöst werden. Auch Bortz und Döring (2006, S. 553) kommen zu dem Schluss, „dass das Reliabilitätskonzept der klassischen Testtheorie bei der Erfassung der Genauigkeit von Differenzwerten offenbar versagt“.

Ein weiteres Argument für die ungenügende Reliabilität von Differenzwerten ist folgendes: Da schon die Reliabilität einer einzelnen Messung durch einen Messfehler beeinträchtigt sei, müsse die Differenz von zwei Messungen noch stärker fehlerbehaftet sein (Hill & Lambert, 2004). Dabei wird angenommen, dass sich der Messfehler einer Differenz sowohl aus dem Messfehler der ersten Messung als auch aus dem Messfehler der zweiten Messung zusammensetze, so dass die wahre Veränderung quasi doppelt verfälscht gemessen werde. Die Fehlervarianz des Differenzwertes setze sich demnach additiv aus den Fehlervarianzen der Einzelmessungen zusammen.

Diese Auffassung wurde mittlerweile überholt und spezifiziert. Williams und Zimmerman (1996) räumen zwar ein, dass manche zufälligen Fluktuationen als Fehler unabhängig in die erste und in die zweite Messungen eingehen. Es sei aber wahrscheinlich, dass zumindest einige zufällige Einflüsse zeitstabil wirken und somit erste und zweite Messung in gleicher Weise beeinflussten. Wenn die Differenzwertmessung als eine zusammengehörige Einheit verstanden werde, die als ganze von zufälligen Fehlern beeinträchtigt werde, dann sei die Reliabilität weniger fehlerbehaftet als oftmals angenommen. Die Annahme der Unabhängigkeit der Fehlereinflüsse und Additivität der Fehler führe somit zu einer überhöhten Schätzung der Fehlervarianz und einer unberechtigten Unterschätzung der Reliabilität der Differenzmessung. Übereinstimmend zeigt Zimmerman (1997) in einer geometrischen Darstellung, dass die Reliabilität von Differenzwerten nicht jeweils einzeln von den Standardabweichungen vorher und nachher abhängt, sondern insgesamt vom Quotienten bzw. dem Verhältnis der beiden Standardabweichungen und somit nicht doppelt verfälscht ist.

Rogosa (1988) weist darauf hin, dass zur Berechnung der Differenzreliabilität zumeist identische Varianzen (und identische Reliabilitätskoeffizienten) zu beiden Messzeitpunkten angenommen werden. Diese Annahme wurde auch zur Herleitung der Formel zur Darstellung des Reliabilitäts-Validitäts-Dilemma getroffen. Diese Annahme treffe aber nicht immer zu. Die Reliabilität von Differenzwerten ist also nicht generell gering, sondern unter bestimmten Bedingungen und zwar dann, wenn – bei unveränderter Reliabilität des Messinstrumentes – die Varianzen vorher und nachher gleich sind und die Vorher-nachher-Korrelation hoch ist (Bortz & Döring, 2006, S. 553).

Die universelle Annahme der Varianzgleichheit scheint fragwürdig zu sein, da Varianzgleichheit und hohe Vorher-nachher-Korrelation gerade darauf hindeuten, dass es keine

Veränderungen gibt, so dass diese auch nicht von einem Differenzwert entdeckt werden können (Gollwitzer & Jäger, 2009). Auch inhaltlich wurde die Annahme der Varianzgleichheit kritisiert, da sie im Kontext von Wachstumsprozessen nicht immer angemessen zu sein scheint und oftmals nicht zu empirischen Daten passe (Williams & Zimmerman, 1996).

Zusammenfassend lässt sich also schließen, dass Differenzwerte dann reliabel sind, wenn sich bei der Messung mit einem reliablen Messinstrument die Varianzen vorher und nachher unterscheiden und die Vorher-nachher-Korrelation kleiner als 1 ist, also dann, wenn sich die Personen in der Stichprobe in unterschiedlicher Weise verändern (Gollwitzer & Jäger, 2009, S. 73). Rogosa (1988) formuliert dies folgendermaßen: „In sum, when there are individual differences in change, the difference score has decent reliability. ... the difference score is reliable when individual differences in true change exist.” (S. 179). Auch Zimmerman (1997) kommt übereinstimmend in seiner geometrischen Interpretation zu dem Schluss, dass eine hohe Reliabilität von Differenzwerten mit ungleichen Standardabweichungen und ungleichen Reliabilitäten der Einzelmessungen zusammenhänge.

4.3.3 Regression zur Mitte

Ein weiteres Problem der wiederholten Messung besteht in der Wirkung des sogenannten Regressionseffekts oder Effekts der Regression zur Mitte. Der Regressionseffekt besteht darin, dass sich extrem ausgeprägte Messwerte bei nochmaligem Messen der größten Dichte der Verteilung bzw. der Mitte der Merkmalsverteilung annähern (Bortz & Döring, 2006). Anders ausgedrückt: Messwerte einer Person, die bei der ersten Messung stark vom Durchschnitt entfernt sind, weichen bei einer zweiten Messung weniger stark vom Mittelwert ab (Gollwitzer & Jäger, 2009). Diese Veränderungen können fälschlicherweise als Veränderung des untersuchten Merkmals interpretiert werden.

Regression zur Mitte wurde in der Vergangenheit oftmals als unvermeidbar angenommen und als Argument gegen die Verwendung wiederholter Messungen ins Feld geführt (Lord, 1963). Mittlerweile konnte dieses Argument teilweise entschärft werden, indem gezeigt wurde, dass Regression zur Mitte nicht immer auftritt, sondern von spezifischen Bedingungen abhängt. Der Effekt ist nach Gollwitzer und Jäger (2009) nur dann unvermeidbar, wenn die Vorher-nachher-Korrelation der Messwerte kleiner als 1 ist und die Messwertreihen identische Varianzen haben. Die erste Bedingung trete in der Praxis u. a. aufgrund von Messfehlern und Einflüssen der Situation quasi immer auf, die zweite Bedingung der Gleichheit der Varianzen allerdings nicht, es sei denn die Werte werden standardisiert, z. B. durch z-Transformation. Zumeist unterscheidet sich also die Varianz der Werte vorher von denen nachher, da sich die Personen unterschiedlich verändern. Damit ist ein Regressionseffekt nicht zwingend unvermeidbar (Gollwitzer & Jäger, 2009).

Rogosa (1988, 1995) konnte mathematisch zeigen, dass empirisch überprüft werden kann, ob ein Regressionseffekt in den Daten vorliegt. Dazu müsse eine negative Korrelation zwischen Betrag der Veränderung und dem initialem Status vorliegen (Rogosa, 1988, S. 187: „regression to the mean pertains only when the correlation between change and initial status is negative.“). Speer (1992) spricht sich im Anschluss an Rogosas Gedanken dafür aus, jeweils anhand vorliegender Daten empirisch zu überprüfen, ob eine negative Korrelation zwischen Veränderung und initialem Status vorliege, und wenn dem so sei, davon auszugehen, dass ein Regressionseffekt vorliege. Wenn Regression zur Mitte anzunehmen sei, sei zur Schätzung der wahren Veränderung ein Verfahren empfehlenswert, welches eine Korrektur der Differenzen vornehme, und wenn nicht, sei keine Korrektur notwendig und die reinen Differenzwerte seien gut und ausreichend.

Nachtigall und Suhl (2002) weisen auf eine andere Besonderheit des Regressionseffektes hin, nämlich seine Ungerichtetheit. Demnach wirke der Effekt nicht nur von der ersten zur zweiten Messung, sondern auch umgekehrt von der zweiten zur ersten Messung. Damit sei der Differenzwert nicht automatisch verfälscht, sondern in der Regel einen guten Schätzwert für die wahre Veränderung.

Regressionseffekte sind insbesondere dann zu erwarten, wenn Extremgruppen untersucht werden (Bortz & Döring, 2006; Zwingmann & Wirtz, 2005). Wenn aus einer selbst selektierten oder auf andere Weise gezogenen Stichprobe nur Probanden mit einer extremen Merkmalsausprägung ausgesucht werden und beispielsweise nach einer Intervention erneut einer Messung unterzogen werden, sei davon auszugehen, dass die vormals extremen Abweichungen näher am Mittelwert liegen, was fälschlicherweise als Erfolg der Intervention interpretiert werden könne bzw. könne Interventionserfolg und Auswirkung des Regressionseffektes schwerlich unterschieden werden. Nachtigall und Suhl (2002) weisen darauf hin, dass bei Selbstmeldern für therapeutische Verfahren ein erhöhter Leidensdruck und Symptomatik anzunehmen sei als bei der Normalpopulation. Die Gruppe der Selbstmelder stelle demnach eine selbst selektierte Gruppe aus der Gesamtpopulation dar, die im Vergleich zu dieser extreme Merkmalsausprägungen in den jeweiligen Symptomatiken habe. Bandelow und Broocks (2002) weisen in diesem Zusammenhang darauf hin, dass der beschriebene Effekt etwas abgemildert werde, wenn zwischen Anmeldung und Beginn der Maßnahme mit erneuter Messung viel Zeit vergehe.

Eine wichtige Frage bezüglich der Interpretation von Differenzwerten selbstgemeldeter Patienten hängt mit der Definition der zugrunde liegenden Population zusammen. Die Frage ist, ob Selbstmelder in therapeutischen Settings als eine Stichprobe aus einer beeinträchtigten Population verstanden werden oder als Stichprobe aus der unbeeinträchtigten Normalpopulation (Speer, 1992). Angewendet auf das Beispiel Jugendlicher mit ADHS stellt sich die Frage, ob die selbstgemeldeten Jugendlichen mit ADHS eine Stichprobe aus der Population der Jugendlichen mit ADHS darstellen oder eine Stichprobe aus der weitaus größeren Gesamtpopulation aller Jugendlichen ganz allgemein.

Jacobson, Follette und Revenstorf (1986) stellen fest, dass es in diesem Zusammenhang zwei unterschiedliche Interessensgruppen gebe, und zwar diejenigen, die eine Hilfe wegen einer klinische Problemlage bzw. Störung aufsuchen, und diejenigen, die dies nicht tun. Die Hilfe suchende, selbst selektierte Gruppe stamme zwar von einem extremen Ende einer allen Personen zugrunde liegenden Verteilung, die zentrale Frage sei aber diejenige, welches Modell am besten zu den Daten passe. Stelle man die Messungen eines Instrumentes, das die Symptome der jeweiligen Störung spezifisch erfasse, von betroffenen und nicht betroffenen Personen nach ihren Häufigkeiten dar, zeichne sich in Histogrammen eine Charakterisierung nach betroffen und nicht betroffen, also zwei Verteilungen, viel deutlicher ab als eine gemeinsame Verteilung. Die Autoren stellen daher fest, dass ein Model mit zwei Verteilungen zu der Population, für die eine Intervention entwickelt wurde, besser passt als ein Modell mit nur einer Verteilung.

Auch wenn der Einsatz von Kontrollgruppen für viele Fragestellungen besser geeignet ist, so liefern Differenzwerte durchaus aussagekräftige Erkenntnisse. Auch Bortz und Döring (2006) stellen fest: „Für die quasiexperimentelle Überprüfung von Veränderungshypothesen lässt sich zusammenfassend feststellen, dass die einfachen Differenzen zwischen den Messungen verschiedener Messzeitpunkte sinnvolle, unverzerrte Schätzungen für „wahre“ Veränderungen darstellen“ (S. 557). Zum selben Ergebnis kommen Hill und Lambert (2004): „In summary, no standard procedure for assessing change exists, although one is clearly needed ... At this point in time, the raw change score remains the metric of choice for most circumstances in which client change is to be assessed“ (S. 123).

4.3.4 Signifikanzprüfung

Bei der Beurteilung der Wirksamkeit therapeutischer Interventionen werden zumeist im ersten Schritt die Mittelwerte zweier Gruppen verglichen. Dies können entweder die Mittelwerte einer Experimental- und einer Kontrollgruppe sein oder die Mittelwerte derselben Gruppe vor und nach der Intervention. Zur Überprüfung, ob die Differenz der Mittelwerte überzufällig ist, wird ein *t*-Test oder ein anderes geeignetes Verfahren durchgeführt. Damit kann überprüft werden, ob der festgestellte Unterschied überzufällig ist und damit als Beleg der Wirksamkeit der durchgeführten Intervention herangezogen werden kann. Es wird dann von einem statistisch signifikanten Unterschied gesprochen (Leonhart, 2004).

Zur Durchführung der Signifikanzprüfung werden Hypothesen aufgestellt. Der vermutete Effekt dass ein Unterschied vorliegt, befindet sich konventionsgemäß in der Alternativhypothese. Die Nullhypothese, dass kein Unterschied vorliegt, wird überprüft. Finden sich die in den Hypothesen angenommenen Effekte in den Daten bei sehr geringer Irrtumswahrscheinlichkeit, wird der Effekt als signifikant bezeichnet. Die zuvor aufgestellte Nullhypothese, die gegen den vermuteten Effekt spricht, wird verworfen und es wird angenommen, dass die Alternativhypothese auch für die Grundgesamtheit zutref-

fend ist. Als Irrtumswahrscheinlichkeit wird konventionell üblich $\alpha = 5\%$ angenommen (Bortz & Döring, 2006).

Die Durchführung von Signifikanztest ist in der Vergangenheit kritisiert worden (Hunter, 1997). Sie stellt aber immer noch das konventionelle Vorgehen dar. Als Verbesserung wurde empfohlen, nicht nur das Ergebnis des Signifikanztests darzustellen, z. B. in Form von Sternchen, sondern auch die Größe ermittelter Effekte und wenn möglich deren Konfidenzintervalle anzugeben, Teststärken zu berechnen und exakte p -Werte anzugeben (Sedlmeier, 2009; Wilkinson & Task Force on Statistical Inference, 1999).

Die Durchführung eines Signifikanztests dient also vor allem der Überprüfung von Hypothesen. Nach Bortz und Döring (2006) ist es aber auch möglich, explorativ auf Signifikanz zu testen, wenn sich ein interessanter Effekt erst zeigt, wenn die Daten schon erhoben sind und inspiziert werden. Ziel eines solchen „Signifikanztests auf Probe“ (S. 379) sei es, neue Hypothesen abzuleiten, die weiterführend geprüft werden können. Dieses explorative Vorgehen darf allerdings nicht im Nachhinein als Hypothesen prüfendes Vorgehen ausgegeben werden, was als „Pseudohypothesentest“ oder „Hypothesizing After the Results are Known“ („HARKING“, Bortz & Döring, S. 498) bezeichnet wird. Dabei wird in einem Datensatz ohne theoretische Fundierung nach signifikanten Effekten gesucht, zu denen passende Hypothesen im Nachhinein konstruiert werden, um dann die gefundenen Effekte als Bestätigung der Hypothesen zu interpretieren.

Mit dem probeweisen Signifikanztest soll die augenscheinliche Bedeutsamkeit eines Effekts durch eine genaue quantitative Überprüfung ergänzt werden (Bortz & Döring, 2006). Es stellt sich die Frage, welchen Nutzen dieses Signifikanzprüfen ohne Hypothese hat. So mag sich beispielsweise für die vorliegende Untersuchung ergeben, dass sich die Trainingserfolge der Mädchen und der Jungen im Mittel nach Augenschein unterscheiden, ohne dass dies zuvor zu erwarten war. Die Mittelwerte der beiden Gruppen können mit einem t -Test für unabhängige Stichproben verglichen werden. Dabei wird die Wahrscheinlichkeit („ p -Wert“) dafür bestimmt, dass der in den Daten vorliegende Unterschied gefunden wird, wenn in der Grundgesamtheit hingegen kein Unterschied vorliegt. Diese Wahrscheinlichkeit wird als Irrtumswahrscheinlichkeit bezeichnet, da sie ausdrückt, wie wahrscheinlich es ist, sich zu irren, wenn die Annahme, dass kein Unterschied vorliegt, verworfen wird. Ist der p -Wert sehr klein, ist es sehr unwahrscheinlich, dass der in den Daten vorliegende Unterschied gefunden wird, wenn er in der Grundgesamtheit nicht vorliegt. Sinkt der p -Wert unter ein bestimmtes Niveau, wird von einem signifikanten Unterschied gesprochen. Per Konvention ist dieses Niveau, das mit α bezeichnet wird, zumeist auf 5 % oder 1 % festgelegt (Bortz & Döring, 2006).

Ob ein Effekt signifikant wird, hängt von der Größe des Effekts, der Stichprobengröße und dem Signifikanzniveau ab. Signifikanz wird eher erreicht, wenn der Effekt groß ist, die Stichprobe groß ist und eine hohe Irrtumswahrscheinlichkeit akzeptiert wird (Bortz & Schuster, 2010). Ob Unterschiede signifikant werden können, kann mit einer Power-

bzw. Teststärke-Analyse bestimmt werden. Bei der Planung einer neuen Untersuchung wird diese Analyse vor Beginn der Studie durchgeführt, u. a. um eine optimale Stichprobengröße zu bestimmen. Wenn nach der Datenerhebung weitere Analysen durchgeführt werden, kann die Power-Analyse auch post-hoc durchgeführt werden (Sedlmeier & Renkewitz, 2008).

4.3.5 Effektstärken in Gruppenvergleichen

In den vorangegangenen Abschnitten wurden Vor- und Nachteile von Differenzwerten und die Signifikanzprüfung von Unterschieden zur Beurteilung der Wirksamkeit therapeutischer Interventionen vorgestellt. Ein signifikanter Mittelwertsunterschied sagt allerdings nichts darüber aus, wie groß und in welcher Richtung ein Unterschied vorliegt. Daher hat sich weiterführend die Ermittlung von Effektstärken als Mittel der Wahl etabliert (Wilkinson & Task Force on Statistical Inference, 1999). Um die Größe eines Effekts einzuschätzen, werden in der Interventionsforschung verschiedene Möglichkeiten diskutiert. Diese werden in diesem Abschnitt dargestellt.

Der englische Begriff „effect size“ wird zumeist als Effektstärke übersetzt, teilweise auch als Effektgröße, um eine unerwünschte Verbindung zum Begriff der Teststärke zu vermeiden (Hager, 2000). Effektstärken können zum Vergleich zweier Gruppen herangezogen werden, z. B. zum Vergleich der Mittelwerte nach einer Intervention in einem Kontrollgruppendesign. Zur Bestimmung der Größe des Gruppenunterschiedes werden nach Cohen (1988) die Mittelwerte beider Gruppen im interessierenden Outcomekriterium nach der Intervention durch ein Maß der Streuung geteilt:

$$d = \frac{M_1 - M_2}{SD}$$

Der resultierende Wert gibt den Unterschied der Gruppen in Einheiten der Standardabweichung an. Eine Effektstärke von 1 bedeutet, dass sich die Mittelwerte der Gruppen um eine Standardabweichung unterscheiden. Welches Maß der Streuung geeignet ist, hängt vom jeweiligen Kontext ab und wird kontrovers diskutiert. Es kann die Streuung in einer der beiden Gruppen sein, einer anderen Kontrollgruppe oder eine gepoolte Streuung, die aus den Streuungen beider Gruppen ermittelt wird (Leonhart, 2004). Die gepoolte Streuung berücksichtigt unterschiedliche Gruppengrößen und Standardabweichungen in den Gruppen und wird nach Leonhart (2004) folgendermaßen bestimmt:

$$SD_{pool} = \sqrt{\frac{(N_1 - 1)SD_1^2 + (N_2 - 1)SD_2^2}{N_1 + N_2 - 2}}$$

Effektstärken sind deskriptive Maße, die unabhängig vom eingesetzten Messinstrument bestimmt werden können. Sie sind dimensionslos, besitzen also keine Maßeinheit. Daher finden sie breite Anwendung in Metaanalysen, welche die Wirksamkeit von thera-

peutischen Verfahren eruieren und dabei Ergebnisse unterschiedlicher Messinstrumente vereinheitlichen und vergleichen. Bei der Interpretation der ermittelten Effektstärken wird allerdings oftmals nicht berücksichtigt, dass gemessene Effektstärken stark vom jeweiligen Kontext abhängig sind. Es sind beispielsweise bei der Erfassung interventionsnaher Kriterien größere Veränderungen zu erwarten als bei distalen Maßen. Solche Unterschiede werden in metaanalytische Untersuchungen oftmals nicht berücksichtigt. Einige Autoren schließen deshalb, dass Effektstärken aus verschiedenen Untersuchungen nicht miteinander vergleichbar sind (Hager, 2000).

Zur Interpretation von Effektstärken wird von vielen Autoren Cohens (1988) Einteilung von standardisierten Mittelwertunterschieden (d) von 0.2 als kleine, von 0.5 als mittlere und von 0.8 als große Effektstärke verwendet. Dabei wird oftmals außer acht gelassen, dass Cohen diese Einteilung zur Schätzung des Parameters d bei der Durchführung von Poweranalysen vorgeschlagen hat, für den Fall, dass zu den erwartbaren Effekten kaum oder wenig Erkenntnisse vorliegen. Die Einteilung in klein, mittel und groß sei eine konventionelle Definition, welche immer relativ zum jeweiligen Forschungshintergrund und zur eingesetzten Methode zu interpretieren sei und nur angewendet werden solle, wenn es keine bessere Schätzung der Effektstärke gebe (Cohen, 1988). Cohens Einteilung kann als Faustregel und Orientierung sinnvoll sein, sollte aber nur zum Einsatz kommen kann, wenn Informationen zum Kontext der Untersuchung fehlen (Durlak, 2009).

Analog zum Gruppenvergleich können Effektstärken auch in quasi-experimentellen Prä-post-Versuchsplänen mit nur einer Versuchsgruppe bestimmt werden. Dabei wird die Differenz der Werte zum Prä- und Postmesszeitpunkt durch ein Maß der Streuung geteilt. Auch hier eignen sich je nach Kontext die Standardabweichung der Werte zum Prämesszeitpunkt, die gepoolte Standardabweichung von Prä- und Postwerten oder die Standardabweichung der Prä-post-Differenzen (Maier-Riehle & Zwingmann, 2000). Nach Kazis, Anderson und Meenan (1989) empfiehlt es sich, die Streuung der Werte zum Prämesszeitpunkt zu nutzen, da diese bei ausreichender Stichprobengröße am ehesten der Streuung der Gesamtpopulation entspricht, die als Referenz herangezogen werden sollte. Eine Standardisierung anhand der gepoolten Streuung empfiehlt sich nach Leonhard (2004) nur wenn Varianzhomogenität vorliegt. Diese könne z. B. mit dem F-Test nach Fischer überprüft werden. Insgesamt empfiehlt Leonhard (2004) die Verwendung der Streuung des Prämesszeitpunktes im Prä-post-Design mit nur einer Gruppe.

4.3.6 Erfolgsmaße im Einzelfall

Auf Einzelfallebene besteht die Möglichkeit der Standardisierung anhand von Gruppenparametern nicht, da Vergleichswerte von Referenzstichproben zumeist nicht vorliegen. Trotzdem besteht Bedarf nach Maßen bzw. Indices, die Aufschluss darüber geben, ob eine im Einzelfall durchgeführte Intervention erfolgreich war. Dies ist z. B. bei einer Evaluation eines Trainings oder einer Therapie, die nur mit einer einzigen Person durchgeführt wurde, der Fall. Ein individueller Erfolgswert, der jedem Teilnehmer zu-

geordnet werden kann, ist aber auch sinnvoll, wenn in einem Prä-post-Design Subgruppen erfolgreicher und weniger erfolgreicher Interventionsteilnehmer gebildet werden sollen, um diese miteinander zu vergleichen.

Grundsätzlich sind Untersuchungsdesigns mit mehreren Messzeitpunkten aufgrund höherer interner Validität vorzuziehen, aber auch bei nur zwei Messwerten, vor und nach der Intervention, können aussagekräftige Auswertungen durchgeführt werden. Prinzipiell bieten sich folgende Verrechnungsmöglichkeiten zur Quantifizierung von Prä-post-Differenzen im Einzelfall an, die sich jeweils in ihrem Bezugskriterium unterscheiden:

- *Absolute Veränderung der Messwerte.* Beispiel: „Es fand eine Verbesserung der Problematik um 8 Punkte statt.“
- *Standardisierte absolute Veränderung der Messwerte.* Beispiel: „Der Zufriedenheitswert hat sich um eine Standardabweichung verbessert.“
- *Relative Veränderung der Messwerte.* Beispiel: „Der Problemwert hat sich um 25 % reduziert.“
- *Kriteriumsbezogene Veränderung.* Beispiel: „Der kritische Grenzwert von sechs zutreffenden Verhaltensmerkmalen wird nicht mehr erreicht.“

Absolute Veränderungsmaße, die aus der Differenz des Vorher-nachher-Unterschieds gebildet werden, können inhaltlich nur interpretiert werden, wenn die Skalierung des Messinstrumentes bzw. ein Referenzwert bekannt ist. So kann eine Veränderung um 8 Punkte auf einer Skala von 1 bis 10 als sehr viel effektiver beurteilt werden als auf einer Skala von 1 bis 100. Es wurden verschiedene Vorschläge gemacht, um diese Standardisierung herzustellen.

Grawe und Braun (1994) schlagen vor, die absolute Vorher-nachher-Differenz an der Streuung des verwandten Verfahrens zum ersten Messzeitpunkt zu standardisieren, um so die Skalierung der Testwerte zu berücksichtigen. Nachtigall & Suhl (2005, S. 242) drücken dies folgendermaßen aus:

$$\text{Standardisierte Post – Prä – Differenz} = \frac{y - x}{S_x}$$

Mit dieser Berechnung übertragen die Autoren das Prinzip der Effektstärkenberechnung in Gruppenstudien auf Einzelfälle. Die Differenz der Prä-post-Werte wird dabei geteilt durch die Streuung der noch unbeeinflussten Prätestwerte. Problematisch ist dabei, dass diese Effektstärken mit einem höheren Messfehler behaftet seien als solche aus Gruppenstudien. Und es muss darauf geachtet werden, dass solcherart berechnete Effektstärken höher sind als solche aus Gruppenstudien, was an sich unproblematisch ist, so lange nicht unterschiedliche Effektstärkenarten verglichen oder verrechnet werden.

Zur Bestimmung, ob eine Veränderung mehr als zufällig von Null unterschiedlich ist, kann die sogenannte kritische Differenz (Lienert & Raatz, 1998), die auch als Reliable Change Index (RCI; Jacobson & Truax, 1991) bekannt ist, genutzt werden. Nach Nachtigall und Suhl (2005) wird dabei die Vorher-nachher-Differenz ($y - x$) bestimmt und standardisiert, allerdings nicht anhand der Standardabweichung des Präwertes, sondern anhand der Standardabweichung des Differenzwertes (S_{y-x}), wie in folgender Formel abgebildet:

$$RCI = \frac{y - x}{S_{y-x}}$$

RCI und kritische Differenz dienen dem Zweck, die standardisierte Veränderung im nächsten Schritt mit einem kritischen Wert zu vergleichen, um festzustellen, ob die Veränderung signifikant von Null verschieden ist und auszuschließen, dass die Veränderungen durch Zufall oder Messfehler zustande kommen. Dieser zweite Schritt des Vergleiches mit einem kritischen Wert wird nur zu einer kriteriumsbezogenen Einschätzung benötigt. Für die Bestimmung eines individuellen Erfolgswertes für jeden Klienten wird der Vergleich mit dem kritischen Wert nicht benötigt, aber die Kalkulation des Erfolgswertes schon.

Die Verfahren der standardisierten Differenz und der kritischen Differenz bzw. RCI eignen sich zur Bestimmung und standardisierten Veränderung. Es wurde kritisiert, dass diese Berechnungen Messfehler und insbesondere den Effekt der Regression zur Mitte nicht ausreichend berücksichtigen (Steyer, Hannover, Telser & Kriebel, 1997) es wurden Prozeduren vorgeschlagen, um diese Verfälschungen zu bereinigen. Weiterführende Analysen zeigten jedoch, dass die Differenzwerte im Allgemeinen als gute Schätzwerte der wahren Veränderung zu verstehen sind und bessere statistische Eigenschaften aufweisen als die korrigierten Prozeduren (Nachtigall & Suhl, 2005).

Eine weitere Möglichkeit individuelle Erfolgswerte zu berechnen bieten so genannte „residual gain scores“. Diese eignen sich nicht in Einzelfalldesigns, wenn nur eine Person behandelt wird. Aber wenn eine Gruppe im Prä-post-Design behandelt wird, kann jedem Teilnehmer ein Erfolgswert zugeordnet werden. Dazu wird zunächst eine z -Transformation der Werte anhand der Gruppenwerte durchgeführt und es wird die Prä-post-Korrelation der Werte der Gesamtgruppe benötigt. Nach Steketee und Chambless (1992) werden die residualen Zuwachswerte bestimmt, in dem von dem z -standardisierten Wert zum Postmesszeitpunkt (Z_2) das Produkt aus dem z -standardisierten Wert zum Prämesszeitpunkt (Z_1) und der Prä-post-Korrelation (r_{12}) abgezogen wird:

$$\text{Residual Gain Score} = Z_2 - (Z_1 * r_{12})$$

Die z-Standardisierung erfolgt anhand des Mittelwertes der Gesamtgruppe („grand mean“). Dabei wird von jedem Messwert der Mittelwert aus dem Mittelwert von Prä- und Postmesszeitpunkten abgezogen und durch die Standardabweichung zum Prämesszeitpunkt geteilt. Die Werte bekommen dadurch einen Mittelwert von 0 und eine Standardabweichung von 1. Die residual gain scores setzen die individuellen Werte in Relation zu den Veränderungen, die von anderen Teilnehmern des gleichen initialen Niveaus gemacht werden (Lincoln et al., 2005).

Als ein Verfahren, das im Gegensatz zu den bisherigen keine Angaben oder Schätzungen zu Verteilungsparametern benötigt, beschreiben Hiller, Bleichhardt und Schindler (2009) die Methode der prozentualen Wertereduktion, die sich im Kontext der Erforschung von Psychopharmaka bereits durchgesetzt habe. Dabei wird die prozentuale Reduktion des Psychopathologiewertes des Patienten bestimmt. Dieser Wert kann auch zur Definition von Therapierespondern genutzt werden, indem eine Größe festgelegt wird, ab wann ein Patient als Responder auf die Therapie definiert wird. Bei der Behandlung von Depressionen werde zur Einstufung als Responder eine Verringerung um 50 % gefordert, bei schizophrenen Erkrankungen von 25- 30 %. Vorteile dieses Verfahrens seien die Möglichkeit des Vergleichs mit anderen Verfahren und dass Parameter einer zugrunde liegenden Verteilung nicht bekannt sein müssen.

Die prozentuale Wertereduktion kann positiv gewendet auch als prozentuale Verbesserung beschrieben werden und ist als solche zur Bewertung verschiedener therapeutischer Verfahren eingesetzt worden, z. B. bei der Behandlung von Kopfschmerzpatienten (Blanchard & Schwarz, 1988). Die prozentuale Verbesserung lässt sich am einfach ermitteln, indem die Vorher-nachher-Differenz durch den Wert zum Prämesszeitpunkt geteilt wird:

$$\text{Prozentuale Verbesserung} = \frac{\text{Symptomwert}_{\text{prä}} - \text{Symptomwert}_{\text{post}}}{\text{Symptomwert}_{\text{prä}}} \times 100$$

Diese Formel gilt für Messinstrumente, bei denen hohe Werte eine große Problematik abbilden, die durch eine Intervention gesenkt werden sollen. Negative Werte entsprechen einer Verschlechterung, bei der die Problematik größer wird.

4.4 Differentielle Wirksamkeit

Ist die Wirksamkeitsprüfung im Kontrollgruppendesign abgeschlossen, gilt als bewiesen, dass die Intervention im Allgemeinen wirksam ist. Dies gilt allerdings nur für die Überprüfung anhand der gemittelten Werte der Stichprobe, d. h. dass die Teilnehmer der Intervention im Durchschnitt eine Verbesserung erlebt haben. Es ist aber davon auszugehen, dass die Teilnehmer in unterschiedlichem Maße von der Intervention profitiert haben. Manche haben vermutlich eine stärkere Verbesserung erlebt, manche eine weniger starke und bei manchen haben sich vielleicht sogar Verschlechterungen einge-

stellt. Solche individuellen Verläufe werden durch die Verrechnung mit Mittelwerten nicht berücksichtigt. Daher erscheint ein weiterer Schritt in Form der Überprüfung hinsichtlich der differentiellen Wirksamkeit angebracht. Dabei gilt es zu überprüfen, wer im Besonderen von der Intervention profitieren konnte und wer weniger, und ob sich Kriterien ausfindig machen lassen, die sich bei den guten und den schlechten Verläufen unterscheiden und darüber hinaus die Erfolgswahrscheinlichkeit sogar vorhersagbar machen.

Im folgenden Abschnitt wird kurz auf das Verständnis der differentiellen Wirksamkeit bezogen auf die Art der Intervention bzw. deren Inhalte dargestellt. Die Überprüfung der differentiellen Wirksamkeit wird dabei so verstanden, dass wirksame Techniken und Strategien innerhalb einer Intervention identifiziert werden. Da dieses Vorgehen im Rahmen der durchgeführten Untersuchung nicht möglich ist, wird dieser Aspekt nicht ausführlicher behandelt. Dagegen werden in den darauf folgenden Abschnitten das Verständnis differentieller Wirksamkeit als die Frage nach der Wirksamkeit der gleichen Intervention bei unterschiedlichen Teilnehmern ausführlicher behandelt.

4.4.1 Wirksame Teilkomponenten einer Intervention

Während die allgemeine Evaluation nur zu klären vermag, ob eine Intervention insgesamt bzw. im Mittel erfolgreich war oder nicht und in welchem Ausmaß, lohnt es sich im Anschluss oder durch ein anderes Design differentiell zu überprüfen, welche Komponenten der Intervention im Einzelnen zum Erfolg beigetragen haben. Dieser Überprüfung liegt oft die Vorstellung zugrunde, dass der Therapeut aus einer ihm verfügbaren Menge therapeutischer Einzeltechniken auswählt und diese zu einem Gesamtpaket zusammenfügt. Ziel der Analyse ist es, die Wirksamkeit jeder einzelnen Technik zu bestimmen, um in einem nächsten Schritt nur mehr wirksame Techniken zu verwenden. Damit würden nicht allgemein Therapieformen oder Therapieschulen insgesamt als wirksam oder nicht beurteilt werden, sondern es würde dem weiter auf den Grund gegangen mit dem Fernziel, nur wirksame Einzeltechniken zu einem wirksamen Gesamtpaket zu schnüren. Empirisch wurde dieses Vorgehen in einigen Metaanalysen umgesetzt (Swanson & Deshler, 2003; Weisz et al., 1995).

Dieses Verständnis der Zusammensetzung wirksamer Techniken führt allerdings auch zu Schwierigkeiten. Die als wirksam identifizierten Techniken waren zwar im Kontext ihrer Durchführung wirksam, das heißt aber nicht unbedingt, dass sie in einem anderen Kontext ebenso wirken. Eine Komponente mag nur deshalb wirksam gewesen sein, da die Existenz einer anderen Komponente dies ermöglicht hat. Es ist davon auszugehen, dass die therapeutischen Techniken nicht unabhängig voneinander wirken, so dass neben dem Einfluss der Einzeltechniken auf den Interventionserfolg auch die Interaktionen der Faktoren untereinander berücksichtigt werden müssen. Technisch ausgedrückt setzt sich der Gesamterfolg einer Intervention als abhängiger Variable aus der Wirkung der verschiedenen Einzeltechniken und darüber hinaus ihren Interaktionen untereinander als unabhängige Variablen zusammen.

Das metaanalytische Vorgehen eignet sich also nur bedingt zur Identifikation wirksamer Teilkomponenten. Einen höheren Grad an Aussagekraft bieten experimentelle Settings, die sich nur in der Ausprägung einer einzelnen Technik oder in einer überschaubaren Anzahl an Techniken unterscheiden. Im therapeutischen Setting ist dies oft nur schwer umsetzbar. Schwierig ist auch die Operationalisierung der Einzeltechniken. So geben Weisz et al. (1995) beispielsweise die soziale Verstärkung als eine wirksame Einzeltechnik an. Diese Technik wird sicherlich in verhaltenstherapeutischen Interventionen, die den Einsatz von Verhaltensverstärkern ausdrücklich fokussieren explizit genutzt, aber es ist nur schwerlich vorstellbar, dass sich irgendeine therapeutische Intervention gänzlich ohne soziale Verstärkung in Form von Zuspruch, Ermutigung oder Belobigung abspielt. Einzeltechniken sind zumeist nicht in Form von nacheinander einsetzbaren Bausteinen operationalisiert, sondern gehen oft in den Habitus des Therapeuten ein, der sie stets einsetzt und nicht nur punktuell als einen Einzelbaustein.

4.4.2 Moderatoren und Mediatoren

Ein anderes Verständnis von differentieller Wirksamkeit geht der Frage nach, welche Merkmale dazu beitragen, dass manche Klienten besser als andere von einer Intervention profitieren. Das Ziel dieses Ansatzes ist es, solche Variablen ausfindig zu machen, die sich auf den Erfolg der Intervention positiv oder negativ auswirken. Eine solche Variable kann z. B. das Alter sein und die zugehörige Feststellung, dass die älteren Teilnehmer mehr von der Intervention profitieren als die jüngeren. Statistisch-technisch betrachtet wird der Versuch unternommen, Interventionserfolg als abhängige Variable aus einer definierten Menge von Klienten-, Therapeuten- und Settingmerkmalen und deren Interaktion als unabhängige Variablen zu einem großen Anteil vorherzusagen. Diesem Ansatz unterliegt also die Annahme, dass der Erfolg von Interventionen durch bestimmte Konstrukte bzw. sie erfassende Variablen vorhergesagt werden kann. Erfolg wäre demnach nicht entweder zufällig oder nur ipsativ aus dem individuellen Therapieverlauf erklärbar, sondern ließe sich vorhersagen, wenn die passenden vorhersagenden Variablen und ihre Beziehungen untereinander gefunden wären.

Eine klassische Analyse im Sinne dieses variablenzentrierten Ansatzes ist die Suche nach dem Wirken von Moderator- und Mediatorvariablen. Nach Baron und Kenny (1986) beeinflusst ein Moderator die Wirkung einer unabhängigen Variablen auf eine abhängige Variable. Die Wirkung einer Intervention könnte z. B. bei Mädchen erfolgreicher sein als bei Jungen. Geschlecht wäre in diesem Fall eine Moderatorvariable. Bei der Analyse von Moderatoren wird also eine Antwort auf die Frage gegeben, für wen und unter welchen Umständen eine Intervention wirkt. Der Moderationseffekt kann als Interaktionseffekt verstanden werden. Bei der Überprüfung auf einen Moderatoreffekt wird kontrolliert, ob der Einfluss der Interaktion zwischen unabhängiger bzw. Prädiktorvariable und dem Moderator auf die Outcomevariable (Pfad *c*) größer ist als die Einflüsse der unabhängigen Variable bzw. des Prädiktors (Pfad *a*) und des Moderators (Pfad *b*) einzeln (siehe Abbildung 17).

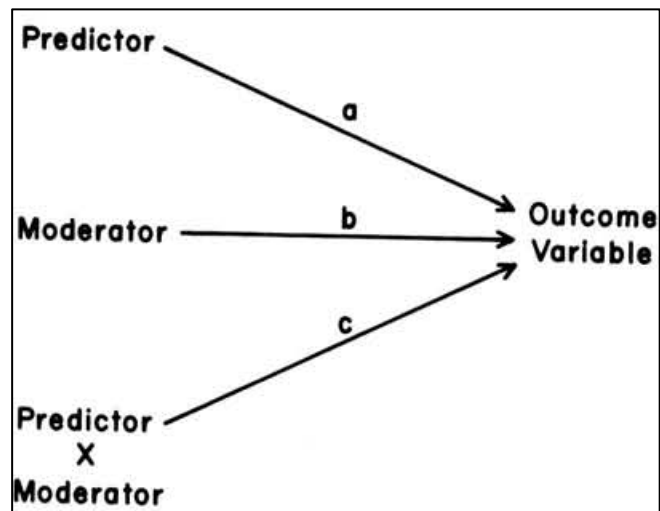


Abbildung 17: Moderatormodell aus Baron und Kenny (1986)

Bei der Anwendung der Moderatoranalyse zur Evaluation einer therapeutischen Maßnahme wäre der Aufbau typischerweise folgender: Die unabhängige Variable wäre die Gruppenzugehörigkeit zu Therapie oder Kontrollgruppe. Bei einem Design mit einer Interventions- und einer Kontrollgruppe wären die Ausprägungen Teilnahme an der Intervention vs. Wartegruppe. Der direkte Zusammenhang zwischen Gruppenzugehörigkeit und dem Therapieerfolg und dem direkten Zusammenhang zwischen einem Moderator, wie z. B. Geschlecht der Teilnehmer, würde mit dem Interaktionseffekt von Gruppe und Geschlecht verglichen werden.

Eine Mediatorvariable vermittelt den Zusammenhang zwischen einer unabhängigen und einer abhängigen Variable (siehe Abbildung 18). So könnte es beispielsweise sein, dass ein gefundener Zusammenhang von Spieltherapie mit akademischen Erfolgen über eine Erhöhung der Selbstwirksamkeitserwartung der Kinder erklärt werden kann. Die Spieltherapie wirkte sich demnach nicht direkt auf die Schulleistungen aus, sondern verbesserte die Erwartungen der Kinder an ihre eigene Kompetenz und dadurch konnten sie wiederum ihre Leistungen verbessern. Mit Mediatoranalysen können also Aussagen darüber getroffen werden, wie Effekte erzielt werden bzw. welche Mechanismen der Wirkung zugrunde liegen (Kendall, Holmbeck, & Verduin, 2004).

Nach Baron und Kenny (1986) ist eine Variable dann ein Mediator, wenn 1. Veränderungen in der unabhängigen Variable signifikant Varianz in der Mediatorvariable erklären können (Pfad *a*), 2. Variationen der Mediatorvariable signifikant Varianz in der abhängigen Variable erklären können (Pfad *b*) und 3. der Zusammenhang zwischen unabhängiger und abhängiger Variable (Pfad *c*), der zuvor einen signifikanten Zusammenhang aufwies, nach Kontrolle der Pfade *a* und *b* nicht mehr signifikant ist. Eine weiterführende Erläuterung findet sich bei Maric, Wiers und Prins (2012).

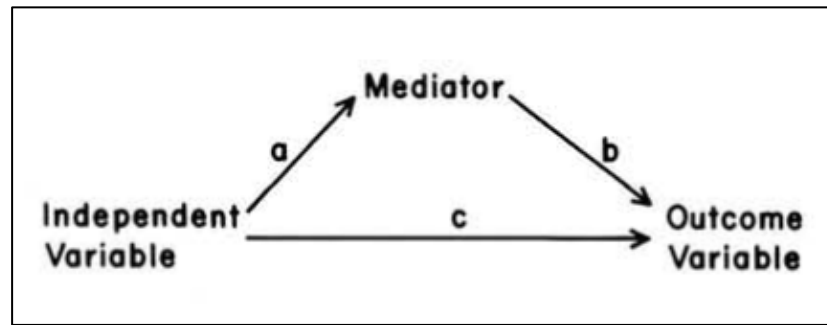


Abbildung 18: Mediatormodell aus Baron und Kenny (1986)

Eine leicht vereinfachte, aber gut verständliche Darstellung von Holmbeck (1997) ist in Abbildung 19 dargestellt.

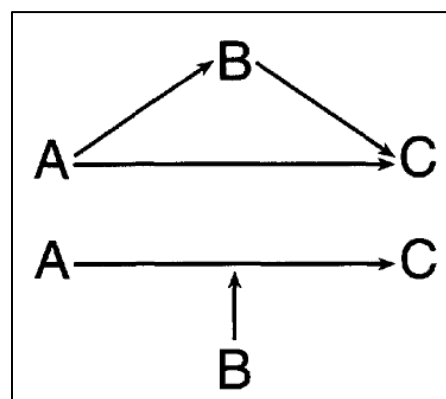


Abbildung 19: Mediator- und Moderatoreffekte aus Holmbeck (1997)

Kraemer, Wilson, Fairburn und Agras (2002) führen den Ansatz von Baron und Kenny (1986) weiter und sorgen für mehr Klarheit in der Unterscheidung zwischen Moderatoren und Mediatoren in randomized clinical trials (RCT). Ein zentrales Unterscheidungskriterium ist gemäß den Autoren, ob die Variable bereits vor der Intervention vorliegt oder erst im Verlauf auftritt. Variablen wie Alter, Geschlecht usw. gehen der Intervention voraus bzw. stehen vor Beginn der Intervention fest und sind demnach Moderatoren. Variablen, die sich erst im Verlauf der Intervention ergeben, z. B. Commitment oder Aspekte der therapeutischen Beziehung, werden als Mediatoren bezeichnet. Ziel der Moderatorenanalyse ist die Identifikation von Subgruppen innerhalb der Population, für die weiterführend unterschiedliche Therapiekonzeptionen sinnvoll sein können. Das Ziel der Mediatoranalyse ist es, die kausalen Zusammenhänge und Wirkweisen von Interventionen zu verstehen. Als weiteren Typus, der im RCT-Design auftreten kann, bezeichnen die Autoren nicht spezifische Prädiktoren. Diese sind Variablen, die schon vor Beginn der Intervention feststehen und einen Effekt auf den Trainingserfolg haben, für die aber kein Interaktionseffekt festgestellt werden kann. Beispielsweise wenn eine Variable den Outcome sowohl der Interventions- als auch der Kontrollgruppe vorhersagen kann. Damit steht sie nicht spezifisch im Zusammenhang mit der Intervention und ist somit kein Moderator. Dies ist beispielsweise der Fall, wenn sich in einer multizentrischen Studie der Ort der Durchführung auf die Outcomes auswirkt, unabhängig davon,

welche Interventionen in den Zentren durchgeführt werden. Unspezifische Prädiktoren können auch Variablen sein, die erst nach dem Training feststehen und keinen Interaktionseffekt aufweisen. Eine solche Variable könnte die Compliance mit der Intervention sein, die sich auf die Outcomes auswirkt unabhängig davon, welche Intervention durchgeführt wurde (Kraemer et al., 2002).

Ein verwandtes Konzept ist das der Aptitude-Treatment-Interaktion. Eine solche Interaktion liegt dann vor, wenn bestimmte Probanden bzw. Gruppen von Probanden in unterschiedlicher Weise auf verschiedene Trainingsvariationen reagieren. Dieses Konzept wurde von Cronbach und Snow (1977) eingeführt und wird insbesondere im Bereich der Bildungsforschung und Didaktik diskutiert. Ein Ziel dieser Forschung ist es, für jeden Lernenden ein passende Lern- bzw. Instruktionmethode zu finden. Die Erforschung von Aptitude-Treatment-Interaktionen eignet sich auch für die Evaluation psychotherapeutischer Interventionen (Snow, 1991) und könnte in diesem Zusammenhang auch als Indikationsforschung bezeichnet werden (Lutz & Grawe, 2007). Das Ziel ist damit die Anpassung der psychotherapeutischen Intervention an den jeweiligen Klienten.

4.4.3 Auswirkung initialer Werte

Über die Auswirkung der Werte vor einer Intervention auf die Werte nach der Intervention ist viel diskutiert worden. Statistisch betrachtet kann diese Auswirkung als eine Prä-post-Korrelation bei einer wiederholten Messung, zwischen der eine Intervention durchgeführt wurde, verstanden werden. Diese Korrelation der initialen Werte mit den Werten nach der Intervention kann auch als „Korrelation Status-Gewinn“ (Klauer, 2001) bezeichnet werden. Sie kann entweder negativ, positiv oder null sein.

Im Zusammenhang mit einer negativen Korrelation zwischen initialer Störungsschwere und Ausmaß der Veränderung wird oftmals das sogenannte Ausgangswertgesetz bzw. Law of Initial Value zitiert (Wilder, 1931). Bei der Beobachtung der Reaktion des vegetativen Nervensystems auf funktionssteigernde Reize wie Adrenalin stellte Wilder fest, dass bei einem höheren Ausgangswert die Reaktion auf den Reiz geringer ausfällt und bei einem niedrigeren Ausgangswert höher. Organe, die bereits stark aktiviert sind, reagieren also stärker auf hemmende als auf aktivierende Reize und schwach aktivierte Organe stärker auf aktivierende als auf hemmende Stimuli reagieren. Es ist umstritten, ob diese Gesetzmäßigkeit, die im Kontext des autonomen Nervensystems formuliert wurde, allgemeingültig ist, allerdings hat sie eine rege Forschungstätigkeit angestoßen (Bortz & Döring, 2006, S. 279). Den gleichen Zusammenhang und negative Korrelation zeigt sich wie beschrieben auch bei der Tendenz zur Regression zur Mitte, so dass es schwierig zu sein scheint, zu unterscheiden, ob es sich um einen statistischen Effekt handelt oder tatsächlich um eine grundlegende Gesetzmäßigkeit.

Bezogen auf den Zugewinn von Kompetenzen im Rahmen kognitiver Trainings unterscheidet Klauer (2001) zwei komplementäre Effekte. Beim kompensatorischen Effekt profitieren die leistungsschwächeren stärker vom Training als die stärkeren. Die schwä-

cheren, für die das Training entwickelt wurde, bekommen die Möglichkeit, ihre Schwäche zu reduzieren, und die Werte der Gesamtgruppe gleichen sich stärker an, so dass die Varianz insgesamt geringer wird. Genau entgegengesetzt ist der „Matthäus-Effekt“ nach dem Bibelvers „Wer hat, dem wird gegeben“ (Klauer, 2001, S. 23). bzw. „them that has, gets“ (Rogosa, 1988, S. 183). Dabei profitieren diejenigen mit einem höheren Funktionslevel mehr von einer Intervention als diejenigen mit einem geringeren Niveau. Dadurch wird insgesamt gesehen die Varianz noch größer, da die Besseren im stärkeren Maße besser werden als die Schwächeren.

Der Begriff „Matthäus-Effekt“ wurde vom Soziologen Merton (1968) eingeführt. Merton beschrieb damit, dass bekannte Wissenschaftler mehr Aufmerksamkeit für ihre Arbeit bekommen als unbekannte Kollegen, auch wenn die Qualität der Arbeiten gleich sei. Der „Matthäus-Effekt“ lässt sich auf viele Kontexte übertragen. Walberg und Tsai (1983) führten ihn als „enhancement, cumulative-advantage or Matthew hypothesis“ (S. 360) in die Bildungsforschung ein. Sie stellen in einer empirischen Untersuchung einen positiven Zusammenhang zwischen frühen Lernerfahrungen, sozioökonomischen Status und späterem Lernerfolg fest. Stanovich (1986) übertrug den „Matthäus-Effekt“ auf den Aufbau von Lesekompetenzen. Um einen unbekanntem Text eigenständig lesen zu können, müssen zumindest einige Wörter bereits bekannt sein. Neue Wörter und die Bedeutung des Textes können besser erfasst werden, wenn bereits zuvor mehrere Wörter bekannt sind. Demnach erscheint es sinnvoll, dass solche Kinder schnellere Fortschritte im Lesen machen, die bereits über einige Fähigkeiten verfügen. Frühe Erfolge beim Lesen begünstigen demnach den weiteren Zuwachs der Lesekompetenzen. Die Leistungen einer Gruppe von Lernern müssten demnach über einen längeren Zeitraum immer weiter voneinander unterscheiden, da die anfänglich Besseren noch stärker besser werden. Daher müsste die Varianz innerhalb der Gruppe größer werden. Dies wird als „fan-spread“ (Shaywitz et al., 1995) oder „Schereneffekt“ bezeichnet und zeigt sich z. B. bei der Entwicklung von IQ-Testergebnissen über längere Zeiträume (Rindermann, 2011).

Die dritte Möglichkeit besteht in einer Nullkorrelation zwischen initialem Status und Veränderung, was nach Anderson (1939) auch als „overlap hypothesis“ bezeichnet wird (Rogosa, 1988, S. 182). Anderson (1939) fand in einer Untersuchung der Entwicklung von IQ-Messungen, dass ein IQ-Wert je eher er gemessen wurde, desto weniger stark er mit einem späteren IQ-Wert korrelierte. Je näher sich zwei Messzeitpunkte waren, desto höher war die Korrelation. Der IQ und sein Wachstum scheinen sich aus verschiedenen Elementen zusammensetzen, deren Wachstumskurven unterschiedlich aussehen und sich überlagern.

4.4.4 Responder und Non-Responder

Zur Beurteilung psychologischer Interventionen ist es wichtig zu bestimmen, welche Maßnahme erfolgreich war und welche nicht. Dazu ist eine Einordnung der Teilnehmer in solche mit günstigen und solche mit ungünstigen Verläufen zu treffen. Es geht also

um die Identifizierung und Analyse von Subgruppen in der Gesamtgruppe der Teilnehmer (Pocock, Assmann, Enos & Kasten, 2002). Nach der Gruppierung nach dem Ausmaß des Trainingserfolges können die gefundenen Gruppen beschrieben werden. Mohr (1995) betont die Wichtigkeit insbesondere auch solche Teilnehmer zu identifizieren, die sich im Laufe der Maßnahme sogar verschlechtert haben. Solche „Non-Responder“ werden im klassischen Kontrollgruppendesign, in dem nur Gruppenmittelwerte verglichen werden, oftmals nicht weiter beachtet. Hiller, Schindler, Andor und Rist (2011) unterscheiden günstige Interventionsverläufe noch weiter in Bezug darauf, ob es nur eine Verbesserung gegeben hat oder ob eine Problematik komplett verschwunden ist. Bei der Verbesserung, also einem positiven Reagieren auf die Intervention, sprechen sie von „Response“ und bei einer vollständigen Genesung von „Remission“.

Ein relativ einfaches Vorgehen wäre es, die Gesamtgruppe in eine Rangfolge nach dem Ausmaß ihres Trainingserfolges zu bringen und dann in der Mitte in zwei Gruppen nach guten und schlechten Trainingsverläufen zu unterteilen. Dabei wäre aber fraglich, warum gerade in der Mitte zu teilen ist. Es ist eine Normalverteilung der Variablen Trainingserfolg anzunehmen, so dass eine Teilung in der Mitte nicht sinnvoll erscheint. Teilnehmer mit sehr ähnlichen Veränderungswerten in der Mitte der Verteilung würden entgegengesetzten Gruppen zugeordnet. Eine bessere Möglichkeit bestünde darin, nur extreme Ausprägungen des Gesamterfolgswertes zu vergleichen, beispielsweise die 15 besten Teilnehmer mit den 15 schlechtesten. Der Extremgruppenvergleich wurde in der Vergangenheit oftmals angewendet (Alf & Abrahams, 1975), stößt heute jedoch weitreichend auf Kritik (Cohen, 1983; MacCallum, Zhang, Preacher & Rucker, 2002; Preacher, Rucker, MacCallum & Nicewander, 2005). Zusammenfassend ist zu sagen, dass eher ungünstig ist, eine Variable zu dichotomisieren, die eigentlich als metrische Variable vorliegt und damit ein höheres Skalenniveau aufweist. Ergebnisse, die erst erscheinen bzw. signifikant werden, wenn das Skalenniveau herabgesetzt wird, sind kritisch zu betrachten, da sich um Artefakte handeln könnte.

Friedman, Furman und DeMets (2010) betonen, dass es besonders dann zu fälschlichen Interpretationen von gesplitteten Erfolgsvariablen kommen kann, wenn die Grenzwerte aus den Daten abgeleitet werden. Eine differenziertere Einstufung kann auf Grundlage des Reliable Change Index (RCI) getroffen werden von Jacobson und Truax (1991). Hierbei werden die Einordnungen nach dem Ausmaß der Veränderung anhand von Daten aus Vergleichsgruppen bzw. Testparametern geschieht. Nach Wise (2004) können vier Gruppen unterschieden werden:

- *recovered*. Es fand eine statistisch signifikante Verbesserung statt und der Nachherwert liegt im Normalbereich.
- *improved*. Es fand eine statistisch signifikante Veränderung statt aber der Nachherwert liegt noch im dysfunktionalen Bereich.
- *unchanged/ indeterminate*. Es fand keine statistisch signifikante Veränderung statt und der Nachherwert liegt noch im dysfunktionalen Bereich.

- *deteriorated*. Es fand eine statistisch signifikante Verschlechterung statt.

Zwei relevante Referenzgrößen kommen in dieser Einordnung zum Einsatz. Erstens wird die Normalpopulation als Vergleichsgruppe herangezogen. Um eine vollständige Genesung („recovery“) feststellen zu können, müssen sich bestimmte störungsbezogene Messwerte nach der Intervention im Normalbereich befinden. Dies kann als ein Aspekt der klinischen Signifikanz verstanden werden. Diese gibt im Gegensatz zur statistischen Signifikanz darüber Aufschluss, ob eine festgestellte Verbesserung im Zuge einer Intervention inhaltlich bedeutsam ist. Zweitens kommt der bereits vorgestellte Reliable Change Index (RCI) zum Einsatz. Übersteigt dieser einen kritischen Wert, so handelt es sich um eine statistisch signifikante Veränderung. Bei einer Irrtumswahrscheinlichkeit von $\alpha = .05$ ist dieser kritische Wert 1.96. Übersteigt der RCI den Wert 1.96 ist es unwahrscheinlich, dass es sich nicht um eine „wahre Veränderung“, sondern nur um zufällige Fluktuationen aufgrund einer ungenauen Messung handelt (Jacobson & Truax, 1991).

In den meisten Studien zur klinischen Signifikanz von Veränderungen kommt eine Einordnung mittels RCI oder normative Vergleiche anhand von Vergleichspopulationen oder eine Einschätzung der prozentualen Verbesserung nach Blanchard und Schwarz (1988) zum Einsatz (Ogles, Lunnen & Bonesteel, 2001). Verschlechterung wird nach RCI zumeist durch eine negative Veränderung, die den kritischen Wert 1.96 übersteigt, bei gleichzeitigem Vorliegen eines auffälligen Symptomscores definiert. Bei der prozentualen Verbesserung, die vor allem im Bereich Gesundheitspsychologie eingesetzt wurde (z. B. Kopfschmerzen, Reizdarmsyndrom), wird Verschlimmerung als Ausbleiben von einer Symptomverbesserung definiert oder als eine Reduktion des Symptomscores unter 50 % (Ogles et al., 2001).

4.5 Empirische Ergebnisse zur differenziellen Wirksamkeit

Im Folgenden werden solche Aspekte vorgestellt, die einen moderierenden Effekt auf den Trainingserfolg haben könnten. Theoretische Grundlagen und empirische Befunde werden zusammenfassend dargestellt. Im folgenden Kapitel werden auf dieser Grundlage Hypothesen aufgestellt.

Geschlecht

Bei beinahe allen psychischen Störungen des Kindes- und Jugendalters gibt es Unterschiede in der Prävalenz bei Jungen und bei Mädchen (Alsaker & Bütikofer, 2005; Kistner, 2009). Psychische Störungen mit einem frühen Beginn wie Autismus, Sprachentwicklungsstörungen, ADHS und LRS treten häufiger bei Jungen auf. Dementgegen sind Störungen mit Beginn im Jugendalter wie emotionale, depressive und Essstörungen bei Mädchen häufiger (Rutter, Caspi & Moffitt, 2003).

Bei einigen Störungen zeigt sich gerade in der Adoleszenz ein Unterschied, z. B. bei depressiven Störungen. Sind Jungen vor der Pubertät noch gleichsam oder sogar stärker betroffen als Mädchen, so ist es in der Pubertät andersherum. Diskutiert wird, dass Mädchen tendenziell mehr Risikofaktoren als Jungen zur Entwicklung einer Depression haben. Im Kontext eines Vulnerabilitäts-Stress-Modells vermuteten Nolen-Hoeksema und Girgus (1994) als eine der ersten, dass Mädchen vor der Adoleszenz mehr Risikofaktoren für die Entwicklung einer Depression aufwiesen als Jungen und dass diese zusammen mit den Veränderungen der einsetzenden Adoleszenz, die für Mädchen noch herausfordernder sein könnten, zur Manifestation einer Depression führten. Mädchen berichten mehr depressive Symptome als Jungen, insbesondere im sozialen Bereich mit Gleichaltrigen. Der Unterschied könnte entweder darauf zurückzuführen sein, dass Mädchen mehr Stressoren ausgesetzt sind als Jungen und deshalb eine stärkere depressive Symptomatik entwickeln oder dass sie stärker mit Depression auf Stressoren reagieren. In einer Studie mit 538 Teilnehmern wurden Hinweise gefunden, dass beides zutrifft (Hankin, Mermelstein & Roesch, 2007).

Geschlechtsbedingte Differenzen hängen mit Unterschieden genetischer Faktoren und daraus folgenden neuronalen und hormonellen Unterschieden zusammen, weiterhin mit unterschiedlicher Reifungsgeschwindigkeit, unterschiedlicher Anfälligkeit für Risiken und mit dem unterschiedlichen Umgang mit Risiken und Typen von Situationen zusammen (Rutter et al., 2003). Gleichsam mit der früheren körperlichen Reifung der Mädchen beginnen auch einige damit verbundene Entwicklungsaufgaben für sie früher und es gibt einige geschlechtsspezifische Entwicklungsaufgaben, die unter anderem mit der Darstellung und Inszenierung von Geschlechtlichkeit im Sinne des „Doing gender“ (West & Zimmerman, 1987) bei Jungen und Mädchen anders auftreten. In der Gestaltung von Maßnahmen zur Prävention und Intervention wird folglich immer stärker auf eine geschlechtergerechte Gestaltung geachtet (z. B. Hinz, 2007).

Empirisch zeigen sich geschlechtsbedingte Unterschiede im Coping-Verhalten von Jugendlichen. Hampel und Petermann (2005) fanden Unterschiede in einer Stichprobe von 1123 Kindern und Jugendlichen zwischen acht und 13 Jahren. Die Mädchen zeigten dabei seltener adaptive Copingstrategien wie positive Selbst-Instruktionen und Ablenkung und häufiger maladaptive Strategien wie Resignation und Rumination. Die Autoren schließen daraus u. a., dass das ungünstige Coping auch zum erhöhten Risiko der Mädchen für die Entwicklung einer depressiven Symptomatik beitragen (vgl. auch Skinner & Zimmer-Gembeck, 2007).

Aufgrund der Unterschiede in der Symptomatik und Prävalenz erscheint es plausibel, dass auch Unterschiede in der Reaktion auf eine psychotherapeutische Behandlung entstehen. Für die Erwachsenenpsychotherapie stellen Clarkin und Levy (2004) allerdings fest, dass in bisherigen Untersuchungen im Wesentlichen keine Unterschiede gefunden wurden, die mit dem Geschlecht der Patienten zusammenhängen. Allerdings seien Studien oftmals nicht dazu geeignet, Unterschiede zu finden, oder es gebe Unterschiede aufgrund unterschiedlicher Prävalenzraten von Störungen.

In einer Metaanalyse über 150 Outcome-Studien wurden höhere Wirksamkeiten für Psychotherapien mit Mädchen als mit Jungen gefunden (Weisz et al., 1995). Dabei zeichnete sich ein Interaktionseffekt zwischen Geschlecht und Alter ab: Die älteren Mädchen erreichten bessere Ergebnisse als die Jungen und auch als die jüngeren Mädchen. In einem metaanalytischen Review zu Präventionsprogrammen von Depression im Jugendalter fanden Stice, Shaw, Bohon, Marti und Rohde (2009) bessere Ergebnisse bei Stichproben mit einem höheren Anteil an Mädchen. Sie bezogen 47 Durchführungen von 32 Programmen in ihre Studie ein. Baskin und Kollegen (2010) fanden hingegen keinen Unterschied für die Wirksamkeit von Beratung und Psychotherapie für Jugendliche in der Schule. Die Autoren untersuchten 107 Studien mit 132 Interventionen. Allerdings waren solche Gruppen effektiver, die nur aus Mädchen bestanden ($d = 0.54$) oder nur aus Jungen ($d = 0.51$), als gemischte Gruppe ($d = 0.33$). Auch Kazdin und Wassell (1999) fanden in einer eigenen Studie mit 200 Kindern mit oppositionellen, aggressiven und antisozialen Verhaltensweisen keine Unterschiede in der Behandlung von Mädchen und Jungen mit kognitiv-verhaltenstherapeutischen Verfahren und Elternt raining.

Bezogen auf ADHS ist das Geschlecht als möglicher Moderator von Interventionserfolgen wenig erforscht. Dies hängt insbesondere damit zusammen, dass sich in den Stichproben aufgrund der geringeren Prävalenzraten oftmals nur sehr wenige Mädchen befinden. Somit haben die Studien zumeist zu wenig Power, um signifikante Unterschiede festzustellen (Pelham & Fabiano, 2008). In der MTA-Studie wurden keine signifikanten geschlechtsbedingten Unterschiede in den Outcomes gefunden (Hinshaw, 2007). Beim Dreijahres-Follow-up stellten Jensen und Kollegen (2007) fest, dass das Geschlecht die Ergebnisse der Behandlung nach 36 Monaten nicht moderierte. Aber zusammen mit der initialen Störungsschwere, Komorbidität, Inanspruchnahme staatlicher Unterstützung und elterlicher Psychopathologie (ADHS) konnten die Outcomes nach 36 Monaten vorhergesagt werden, unabhängig davon, zu welcher Art der Behandlung die Kinder zugeteilt waren. Zur Vorhersage der Outcomes nach drei Jahren trugen die demografischen Variablen Geschlecht und Inanspruchnahme staatlicher Unterstützung mehr bei als die Zuteilung zu den Behandlungsgruppen. Mädchen, die in Familien ohne staatliche Unterstützung aufwuchsen, verbesserten sich in 36 Monaten stärker. Anzumerken ist allerdings auch, dass sich alle Gruppen insgesamt verbesserten, unabhängig von der Zugehörigkeit zu einer bestimmten Intervention.

In einer Metaanalyse konnten 14 Studien zur Wirksamkeit nicht pharmakologischer Interventionen für Kinder mit ADHS gefunden werden (Hodgson, Hutchinson & Denson, 2012). In zehn Studien davon wurden Aussagen zum Geschlecht gemacht. In diesen Studien befanden sich insgesamt $n = 186$ Kinder in der Interventionsbedingung und $n = 152$ in der Kontrollbedingung. Die Autoren kommen zu dem Ergebnis, dass Interventionen bei Mädchen erfolgreicher waren als bei Jungen. Sie begründen dies damit, dass die Anzahl der Mädchen in einer Studie signifikant positiv mit der durchschnittlichen Effektstärke korrelierte ($r = .57$), während die Anzahl der Jungen in einer Studie negativ mit der Effektstärke korrelierte ($r = -.79$). In Wirksamkeitsstudien mit mehr Mädchen in der Evaluationsstichprobe wurden also bessere Ergebnisse erzielt. Es

wurden aber keine Effektstärken getrennt für Mädchen und Jungen über alle Studien bestimmt, was stärkere Aussagen zugelassen hätte. Zudem befanden sich in den Interventionsgruppen jeweils nur $M = 2.5$ Mädchen ($SD = 1.6$) und in den Kontrollgruppen nur $M = 1.9$ Mädchen ($SD = 1.9$) bei einer Spannweite von null bis fünf Mädchen pro Gruppe. Aufgrund der geringen Gruppengröße ist fraglich, ob die Ergebnisse generalisierbar sind.

Bei Interventionen bei Störungen des Sozialverhaltens wird das Geschlecht als ein wichtiger Moderator behandelt (Nangle, Erdley, Carpenter & Newman, 2002). Dies wird insbesondere damit begründet, dass sich Störungen des Sozialverhaltens unterschiedlich bei Jungen und Mädchen äußern. Jungen zeigen häufiger offene und Mädchen eher verdeckte Formen der Aggression. Offene Aggression beinhaltet das Zufügen oder Androhen von physischem Schaden; verdeckte Aggressionen haben oftmals relationalen Charakter und äußern sich beispielsweise durch das Verbreiten von Gerüchten oder den Ausschluss von sozialen Aktivitäten. Diese Unterschiede in der Symptomatik werden in Interventionen bei Störungen des Sozialverhaltens häufig nicht berücksichtigt. Vor diesem Hintergrund erstellten de Boo und Prins (2007) ein Review zur Wirksamkeit sozialer Kompetenztrainings bei Kindern mit ADHS. Sie konnten nur zwei Studien finden, in denen Mädchen und Jungen gleichmäßig über die Interventionsbedingungen verteilt waren, so dass aussagekräftige Ergebnisse gewonnen werden konnten. Entgegen der Annahme fanden sich in keiner der beiden Studien geschlechtsbedingte Unterschiede in den Interventionsergebnissen.

Wie bereits dargestellt ist das Erscheinungsbild der ADHS im Gegensatz zur Störung des Sozialverhaltens bei Jungen und Mädchen eher ähnlich als unterschiedlich. Insbesondere trifft dies auf Lernschwierigkeiten zu, während emotionale Probleme vor allem im Jugend- und Erwachsenenalter stärker bei Mädchen und Frauen auftreten. Die Ähnlichkeit in der Symptomatik nimmt mit steigendem Alter zu. Vor diesem theoretischen Hintergrund ist nicht davon auszugehen, dass das Lerntraining LeJA unterschiedliche Effekte bei Mädchen und Jungen erzielt.

Auch von der Trainingskonzeption her ist nicht davon auszugehen, dass geschlechtsspezifisch unterschiedene Wirksamkeiten auftreten. Die Durchführung findet insbesondere im späteren Coaching- und Selbstmanagementteil individualisiert statt. Die Jugendlichen werden ermutigt, Probleme ins Training einzubringen, die sie aktuell beschäftigen und die sie bearbeiten möchten. So kann es beispielsweise sein, dass die Mädchen stärker als die Jungen soziale Probleme thematisieren und dafür Lösungen gefunden wurden. So sollten durch die Berücksichtigung individueller Bedürfnisse jedes einzelnen Teilnehmers auch eventuell vorhandener geschlechtstypische Bedürfnisse berücksichtigt werden.

Zusammenfassend kann gesagt werden, dass die empirische Studienlage zur Abschätzung des Zusammenhangs vom Geschlecht und dem Interventionserfolg unklar ist. Nussbaum (2012) stellt in ihrem Überblick resümierend fest, dass zwar viele Studien

zur Effektivität von Interventionen bei ADHS vorliegen, aber eine Knappheit an Studien vorherrscht, die Geschlechterunterschiede in den Interventionen berücksichtigen. Allerdings leiden solche Analysen, die eine geschlechtsspezifisch unterschiedliche Wirksamkeit feststellen, oftmals unter einer ungenügenden Studienqualität (Hodgson et al., 2012). Generalisierbare Aussagen können aus den meisten Studien nicht abgeleitet werden, da Mädchen unterrepräsentiert sind. Das Störungsbild ist im Jugendalter bei den Geschlechtern eher ähnlich als unterschiedlich (Novik et al., 2006). Insbesondere gilt dies für das Lernverhalten, welches im Vordergrund der Intervention LeJA steht. Doch selbst wenn die Symptomatik unterschiedlich ist, bedeutet dies nicht, dass die Interventionen geschlechtsspezifisch unterschiedlich wirken, wie am Beispiel der Störung des Sozialverhaltens deutlich wird (de Boo & Prins, 2007). Auch aufgrund der individualisierten Trainingsgestaltung ist also davon auszugehen, dass Geschlecht kein Moderator des Trainingserfolges ist.

An dieser Stelle ist anzumerken, dass es wichtig zu sein scheint, genauer zu untersuchen, welche Mädchen für eine Behandlung angemeldet werden. Es gibt Hinweise darauf, dass es einen Selektionseffekt in klinischen Samples mit Mädchen mit ADHS gibt. Mädchen werden weniger häufig klinisch vorgestellt, obwohl sie oft gleichermaßen beeinträchtigt sind (Novik et al., 2006). Vermutlich werden insbesondere Mädchen mit einer stärkeren Problematik und Komorbidität vorgestellt, so dass ausgeschlossen werden muss, dass vermeintliche Geschlechtseffekte allein über Unterschiede in der Störungsschwere oder Komorbidität erklärbar sind. In diesem Zusammenhang sei darauf hingewiesen, dass die Symptomatik einer ADHS des hyperaktiv-impulsiven Subtypus sehr ähnlich der einer Borderline-Persönlichkeitsstörung ausfallen kann. Es mag sein, dass sich in unserer Stichprobe auch solche Mädchen befinden, die Kriterien einer Borderline-Persönlichkeitsstörung erfüllen bzw. dass die ADHS-Symptomatik als Vorläufer einer Borderline-Persönlichkeitsstörung konzeptualisiert werden kann. Stepp, Burke, Hipwell und Loeber (2012) fanden in diesem Kontext längsschnittliche Zusammenhänge in einer Studie mit 1233 Mädchen. Stärkere Ausprägungen der ADHS und oppositioneller Verhaltensweisen im Alter von acht Jahren sagten eine Borderline-Symptomatik im Alter von 14 Jahren voraus.

Medikation

In der für diese Studie untersuchten Stichprobe finden sich sowohl Jugendliche, die zusätzlich zum Lerntraining medikamentös behandelt werden, als auch solche, die nur mit dem Lerntraining behandelt wurden. Es stellt sich die Frage, ob die zusätzliche Medikation einen Zusammenhang mit dem Trainingserfolg aufweist. Zum Vergleich der Effekte psychosozialer Interventionen und Kombinationsbehandlungen liegen für das Kindesalter Metaanalysen vor, von denen zwei vorgestellt werden. Es folgt ein Ausblick auf das Erwachsenenalter und eine Diskussion der Thematik im Jugendalter, für die bislang keine empirischen Studien vorliegen.

Majewicz-Hefley und Carlson (2007) untersuchten in einer Metaanalyse Studien, die eine rein pharmakologische Behandlung mit einer Kombination von pharmakologischer und psychosozialer Intervention verglichen. Acht Studien wurden in die Metaanalyse aufgenommen, ausschließlich im Kindesalter, mit einer Spannweite von 5 bis 12 Jahren. Die Anzahl der Teilnehmer mit einer Kombinationsbehandlung lag zwischen 18 und 136. Alle Studien berichteten Verbesserungen in beiden Behandlungsgruppen. In drei von acht Studien wurde die Kombinationsbehandlung als erfolgreichste Behandlung eingestuft, in den restlichen fünf Studien wurde kein Unterschied festgestellt. Es wurden zusammenfassende Effektstärken je nach untersuchtem Erfolgskriterium ermittelt. Diese waren höher für die ADHS-Kernsymptomatik (Unaufmerksamkeit: $d = 1.27$, Hyperaktivität: $d = 1.27$, Impulsivität: $d = 0.91$) als für begleitende Merkmale (Sozialverhalten: $d = 0.90$, schulische Leistungen: $d = 0.19$). Da keines der zugehörigen Konfidenzintervalle den Wert 0 beinhaltet, schlussfolgern die Autoren, dass die Kombinationsbehandlung signifikant dazu beitragen kann, sowohl die Kernsymptomatik als auch begleitende Probleme zu verbessern. Weiterhin wird vermutet, dass sich die Medikation vor allem auf die Kernsymptomatik auswirkt und psychosoziale Interventionen stärker zusätzliche Probleme verbessern können, insbesondere im sozialen und im schulischen Bereich. Daher verwundert die geringe Effektstärke für die Schulleistungen. Die Autoren diskutieren, dass die genutzten standardisierten „achievement scores“ eventuell nicht änderungssensibel waren und dass die Zusammenfassung unterschiedlicher schulischer Bereiche zu einer Überlagerung von Effekten geführt haben könnte.

Van der Oord et al. (2008) verglichen die Effektivität von psychosozialer, medikamentöser und Kombinationsbehandlung bei Kindern mit ADHS. Sie schlossen 26 Studien in ihre Metaanalyse ein, die zwischen 1985 und 2006 veröffentlicht wurden. Teilweise waren dies dieselben wie in der Studie von Majewicz-Hefley und Carlson (2007). Eingeschlossen wurden nur Studien mit einem randomisierten kontrollierten Design, die in einem ambulanten Setting mit Kindern von 6 bis 12 Jahren durchgeführt wurden. Sowohl die medikamentöse Behandlung als auch psychosoziale Interventionen können effektiv ADHS-Symptomatik reduzieren. Die psychosozialen Interventionen erreichen zumeist kleinere Effekte und es wurde kein additiver Nutzen bei medikamentöser Behandlung durch zusätzliche psychosoziale Interventionen festgestellt. Zur Verbesserung des Sozialverhaltens scheinen sowohl reine Medikationsbehandlung als auch psychosoziale Interventionen und die Kombinationsbehandlung gleichsam effektiv zu sein. Die Autoren diskutieren, dass möglicherweise kein zusätzlicher Effekt psychosozialer Interventionen deshalb gefunden wurde, da in den meisten Studien zuerst die Medikation dosiert wurde und erst dann psychosoziale Interventionen begonnen wurden. So mag der Spielraum zur Verbesserung der Symptomatik durch die starke Wirkung der Medikation bereits ausgeschöpft worden sein. Es gibt Hinweise darauf, dass größere additive Effekte bei einer geringeren Dosierung der Medikation vorliegen (van der Oord et al., 2008).

Döpfner und Kollegen (2004) untersuchten in einem adaptiven Behandlungsplan die Wirksamkeit behavioraler und medikamentöser Behandlung. 75 Kinder mit ADHS im

Alter von 6 bis 10 Jahren wurden nach einer Psychoedukation zufällig entweder einer Verhaltenstherapie oder medikamentösen Behandlung zugeteilt. Wenn die Behandlungen vollständig erfolgreich waren, wurden sie beendet, wenn nicht wurde ergänzend oder zusätzlich die jeweils andere Behandlungsstrategie eingeführt. Bei 26 % der Kinder mit einer anfänglichen Verhaltenstherapie wurde im späteren Therapieverlauf eine zusätzliche medikamentöse Behandlung eingeführt. In 82 % der Fälle mit anfänglicher Medikationsbehandlung wurde im weiteren Verlauf eine behaviorale Intervention ergänzt. Es wurden große Prä-post-Effektstärken für diese individuell zugeschnittenen Behandlungen ermittelt. Die Studie unterstreicht die Relevanz der Reihenfolge der durchgeführten Interventionen. Bei 74 % der Kinder war eine reine Verhaltenstherapie ohne Medikation ausreichend und für die meisten Kinder war eine reine pharmakologische Behandlung nicht ausreichend.

In einer Studie von Safren und Kollegen (2005) wurden 31 Erwachsene mit ADHS entweder nur medikamentös behandelt oder kombiniert medikamentös und mit kognitiver Verhaltenstherapie (CBT). Die Erwachsenen mit der kombinierten Behandlung zeigten stärkere Verbesserungen in der Symptomatik, sowohl in der Selbst- als auch in der Fremdeinschätzung. Darüber hinaus erzielten sie stärkere Verbesserungen bezogen auf Symptome von Ängstlichkeit und Depression. Unter den Patienten mit kombinierter Behandlung waren 56 % Responder, während in der Gruppe mit der reinen Medikationsbehandlung nur 13 % Responder waren. Die Autoren schließen daher, dass zur Behandlung der ADHS im Erwachsenenalter eine Kombinationsbehandlung aus Medikation und CBT gut umsetzbar ist, gut angenommen wird und gute Wirksamkeit verspricht. Auch Linderkamp und Lauth (2011) fanden in einer Metaanalyse bessere Effektstärken für die psychotherapeutische Behandlung von Erwachsenen mit ADHS ($d = 0.84$) als für die medikamentöse Behandlung ($d = 0.44$). In die Untersuchung wurden 12 Studien zur Wirksamkeit psychotherapeutischer Behandlung und 43 Studien zur Wirksamkeit medikamentöser Behandlung eingeschlossen.

Gerade für Jugendliche scheint es plausibel zu sein, dass eine Kombinationsbehandlung effektiver ist als eine rein medikamentöse Behandlung. Stärker als bei Kindern treten bei Jugendlichen mit ADHS sekundäre Probleme auf, die vermutlich besser durch eine psychosoziale Intervention behandelt werden können. Wie im Modell von Rapport und Kollegen (2001) veranschaulicht, wirken Medikamente insbesondere auf die neurobiologischen Grundlagen der Störung ein und damit nur vermittelt auf Kernsymptomatik und periphere Probleme wie Lernschwierigkeiten. Die sekundäre und Residualsymptomatik kann vermutlich besser durch nichtmedikamentöse Interventionen behandelt werden (Majewicz-Hefley & Carlson, 2007).

ADHS-Subtypus

Die ADHS-Symptomatik äußert sich je nach Störungssubtyp unterschiedlich. Das wichtigste Unterscheidungsmerkmal ist dabei, ob zusätzlich zu den Symptomen der Unaufmerksamkeit hyperaktive und impulsive Verhaltensweisen auftreten. Jugendliche des

vorwiegend unaufmerksamen Subtypus fallen in ihrem Arbeitsverhalten insbesondere durch ihr langsames und schwerfälliges Vorgehen auf. Dies wird auch als „sluggish cognitive tempo“ bezeichnet, das entweder als Aspekt der ADHS oder als eigenständige, komorbide Störung konzeptualisiert werden kann. In einem Lerntraining mit solchen Jugendlichen wirkt der Trainer häufig aktivierend und versucht zusätzliche, zielführende Verhaltensweisen anzubahnen. Jugendliche des hyperaktiv-impulsiven Subtypus fallen dagegen durch ein Übermaß an Aktivität auf und müssen prinzipiell eher gebremst werden. Beide Störungssubtypen weisen Schwierigkeiten der Handlungsorganisation und -steuerung auf, so dass die gleichen therapeutischen Techniken sinnvoll sind. Es ist jedoch anzunehmen, dass die Ausgestaltung eines Trainings mit einem vorrangig unaufmerksamen Jugendlichen im Vergleich zu einem Jugendlichen des kombinierten Subtypus sehr unterschiedlich ausfällt. Aber auch wenn es wahrscheinlich ist, dass sich Trainings mit Jugendlichen unterschiedlicher Subtypen in der Durchführung stark unterscheiden, ist damit noch nicht gesagt, dass sich dies auch in unterschiedlichen Ausmaßen der Wirksamkeit niederschlägt.

Die Frage, welcher ADHS-Subtypus von welcher Intervention am meisten profitiert, ist empirisch nicht abschließend beantwortet. Viele Evaluationsstudien haben die Unterschiede der Subtypen nicht systematisch untersucht oder nur einen Typus in die Studie eingeschlossen, so wie bei der MTA-Studie, die nur den kombinierten Subtypus berücksichtigte (Jensen, Arnold, Richters, Severe, Vereen, Vitiello, Schiller, Hinshaw, Elliott, Conners, Wells, March, Swanson, Wigal, Cantwell, Abikoff, Hechtman, Greenhill, Newcorn, Pelham, Hoza, Kraemer et al., 1999). Es scheint so zu sein, dass Kinder und Jugendliche des kombinierten Subtypus' in stärkerem Maße behandelt werden, sowohl medikamentös als auch nicht-medikamentös (Faraone, Biederman, Weber & Russell, 1998; Weiss, Worling & Wasdell, 2003). Es konnte gezeigt werden, dass alle Subtypen von einer medikamentösen Behandlung profitieren können, wobei Betroffene des kombinierten Subtypus' am besten von mittleren bis starken Dosierungen profitierten, während beim unaufmerksamen Typus geringere Dosen erfolgreich waren (Barkley, Dupaul & McMurray, 1991; Stein et al., 2003). Im Vergleich zu unauffälligen Kontrollgruppen haben ADHS-Betroffene aller Subtypen mehr Unterstützung und Behandlung in Anspruch genommen, im Vergleich der Subgruppen untereinander sind Betroffene des kombinierten Typus' häufiger in Behandlung und profitieren mehr von ihrer aktuellen Behandlung (Gomes, 2011).

Antshel und Remer (2003) führten einer Evaluationsstudie zur Wirksamkeit eines sozialen Kompetenztrainings mit 120 Kindern mit ADHS durch. In der Stichprobe befanden sich 59 Kinder des unaufmerksamen Subtypus und 61 Kinder des kombinierten Subtypus. Die Kinder mit dem unaufmerksamen Störungssubtypus und ihre Eltern berichteten eine signifikant stärkere Verbesserung in der Selbstbehauptung („assertion“) zum Postmesszeitpunkt, allerdings nicht in anderen Bereichen. Dieser Unterschied war auch noch im Follow-up vorhanden, allerdings weniger stark. Die größeren Erfolge der Kinder des unaufmerksamen Subtypus in einem sozialen Kompetenztraining können therotisch folgendermaßen begründet werden: Die Kinder des hyperaktiv-impulsiven Subtypus

fallen durch Probleme der Verhaltensinhibition auf, die sich u. a. in Aggressivität äußern. Dieses ist bei Kindern des vorrangig unaufmerksamen Typus zumeist nicht der Fall, deren Sozialverhalten eher passiv ist und sich durch mangelndes soziales Wissen auszeichnet (de Boo & Prins, 2007). Es kann vermutet werden, dass sie deshalb leichter ein verträgliches Sozialverhalten erlernen können und weniger stark von Gleichaltrigen abgelehnt werden. Auf eine Intervention zur Förderung des Lernverhaltens ist dies allerdings nicht ohne Weiteres übertragbar.

Solanto und Kollegen (2010) evaluierten ein meta-kognitives Gruppenprogramm mit einer Stichprobe von 88 Erwachsenen mit ADHS. Ziele der Intervention waren die Verbesserung von Zeitmanagement, Planung und Selbstorganisation. Die Erwachsenen wurden randomisiert der meta-kognitiven Therapie oder einer unterstützenden Psychotherapie im Gruppensetting zugeordnet. In Regressionsanalysen zeigte sich die Überlegenheit der meta-kognitiven Therapie gegenüber der Kontrollbedingung. Zusätzlich wurden mögliche Moderatoren wie der ADHS-Subtypus in die Analyse eingegeben. Der Haupteffekt der Überlegenheit der Interventionsbedingung blieb auch nach Aufnahme aller potentieller Moderatoren bestehen und es zeichnete sich kein signifikanter Interaktionseffekt ab.

Willcutt und Kollegen (2012) fassen zusammen, dass keine oder nur sehr geringe Unterschiede in der Wirksamkeit von Interventionen für unterschiedliche ADHS-Subtypen gefunden wurden. Ergänzend stellen die Autoren fest, dass bislang keine Evaluationsstudien vorliegen, die den hyperaktiv-impulsiven Subtypus systematisch berücksichtigt haben. Auch aus der Theorie ist nicht eindeutig abzuleiten, für welchen Störungssubtyp sich eine höhere Wirksamkeit abzeichnen sollte. Es geht dabei um die Frage, ob es einfacher ist, das Lernverhalten eines langsamen, schwerfälligen Lerners zu verbessern oder das eines impulsiven, vorschnellen Lerners. Im ersten Fall stehen stärker aktivierende Maßnahmen im Vordergrund, im zweiten Fall eher solche, die das überschießende Verhalten kanalisieren und in die richtigen Bahnen lenken. Welches Vorgehen einfacher ist und zu größeren Erfolgen führt, ist vor dem theoretischen Hintergrund nicht eindeutig ableitbar und wie berichtet empirisch nicht entschieden.

Ein Problem speziell im Jugendalter ist folgendes: Eine klare Zuordnung der Jugendlichen zu einem ADHS-Subtypus ist schwierig, da die Hyperaktivität und Impulsivität in der Adoleszenz zurückgeht und nicht mehr alle zur Diagnose erforderlichen Symptome erfüllt werden. Die Ausprägung der Störung geht immer mehr in Richtung unaufmerksamer Subtypus für alle Jugendlichen. Ob die Subgruppenzugehörigkeit aus der Kindheit über das Jugendalter stabil bleibt, ist fraglich (Larsson, Dilshad, Lichtenstein & Barker, 2011; Normand, Flora, Toplak & Tannock, 2012).

Alter

Da es sich beim LeJA um ein kognitiv-behaviorales Lerntraining handelt, kann die Annahme getroffen werden, dass ältere Teilnehmer aufgrund ihrer weiter vorangeschrittenen kognitiven Reife stärker vom Training profitieren als jüngere. Theoretischer Hin-

tergrund für die Annahme ist die schnelle Entwicklung der Intelligenz, die im Jugendalter ihren Höhepunkt erreicht, sowohl in ihren pragmatischen als auch mechanischen Anteilen (Baltes et al., 1999). Gerade auch noch in der späten Adoleszenz bis in die Zwanzigerjahre zeigen sich deutliche Verbesserung in Entscheidungs- und Planungsprozessen (Albert & Steinberg, 2011a, 2011b; Cauffman & Steinberg, 2000). Dementsprechend kann ein positiver Zusammenhang zwischen Alter und Trainingserfolg vermutet werden.

Viele Methoden der kognitiven Verhaltenstherapie erfordern ein hohes Maß an Selbstreflexion, Metakognition, Vorwegnahme zukünftiger Möglichkeiten und logisches Schlussfolgern. Grundlage dafür sind komplexe symbolische Repräsentationen der Objekte, mit denen operiert wird (Holmbeck, Greenley & Franks, 2003). In der Vergangenheit wurden geringe Wirksamkeiten für stark kognitiv basierte Interventionen wie die Selbstinstruktion festgestellt (Abikoff, 1991). Dies kann damit zusammenhängen, dass die erforderlichen kognitiven Grundlagen bei Kindern noch nicht in ausreichendem Maße vorhanden sind. Somit sollten kognitive Techniken bei Jugendlichen wirksamer sein als bei Kindern. Aber auch im Jugendalter ist die Effektivität - zumindest ausschließlich - kognitiver Techniken eher gering einzuschätzen (Hinshaw, 2006). Dies mag damit zusammenhängen, dass Jugendliche mit ADHS in vielen Entwicklungsbereichen hinter den Gleichaltrigen zurückbleiben und ADHS als eine allgemeine Entwicklungsverzögerung verstanden werden kann (Barkley, 1997). Als ein biologisches Korrelat wird wie bereits beschrieben ein verlangsamter Reifungsprozess der Gehirnentwicklung, insbesondere im präfrontalen Cortex, bei Kindern mit ADHS festgestellt (Shaw et al., 2007). Interventionen, die „top-down“ geleitete Prozesse wie Arbeitsgedächtnis und Aufmerksamkeitsteuerung verbessern sollen, sind daher vermutlich erst im Jugendalter oder im frühen Erwachsenenalter erfolgsversprechend (Rutledge, van den Bos, McClure & Schweitzer, 2012). Zusammenfassend kann gesagt werden, dass Jugendliche mit ADHS im Laufe ihrer Adoleszenz große Zuwächse in ihren kognitiven Fähigkeiten erlangen. Dieses sind sowohl im Vergleich zu Gleichaltrigen verzögert auftretende Zugewinne als auch normativ auftretende Fortschritte in den Entscheidungs- und Planungskompetenzen. Interventionen, die auf kognitive Verbesserungen abzielen, müssten daher mit steigendem Alter als einem Indikator für kognitive Reife zunehmend erfolgreich sein.

In der menschlichen Entwicklung zeigen sich interindividuell große Unterschiede und ein hohes Maß an Variabilität (Siegler, 2007). Entwicklungen laufen diskontinuierlich ab, z. B. bei der früher einsetzenden körperlichen Reifung der Mädchen im Vergleich zu den Jungen. Und auch innerhalb einer abgrenzbaren Gruppe (z. B. bei den Mädchen) ist die Entwicklung und Reifung in unterschiedlichen Entwicklungsbereichen unterschiedlich schnell. Das chronologische Alter determiniert also nicht den Stand der Entwicklung. So wurden auch die Altersangaben in der Zuordnung von Piagets Stufen immer wieder kritisiert und teilweise ganz weggelassen (Siegler et al., 2008). Trotzdem kommt dem chronologischen Alter ein zentraler prädiktiver Wert für die Voraussage des Entwicklungsstandes eines Kindes oder Jugendlichen zu, wie kaum einer anderen Variab-

len. Im Trend oder Gruppenvergleich sind ältere Kinder und Jugendliche weiter entwickelt als jüngere in allen Bereichen der Entwicklung. Die Wahrscheinlichkeit, dass ein älteres Kind weiter entwickelt ist als ein jüngeres ist sehr hoch, aber nicht 100 %. Im Englischen wird dieser Zusammenhang ausgedrückt: „age as a proxy for cognitive developmental level“ (Holmbeck, Devine & Bruno, 2010, S. 30). Wenn es also um einen allgemeinen Indikator für den Stand der Entwicklung in allen ihren Bereichen geht, so ist das chronologische Alter eine gute Wahl.

Systematische und gut kontrollierte Studien, die sich ausschließlich auf Jugendliche mit ADHS beziehen, liegen kaum vor (Pelham & Fabiano, 2008). Daher wird der Stand der empirischen Wirksamkeitsforschung im Folgenden auch anhand verwandter externalisierender Störungen vorgestellt, zu denen Studien vorliegen.

Ein Alterseffekt zeigte sich in einer Metaanalyse über 150 Studien von Weisz, Weiss, Han, Granger und Morton (1995). Die Effektstärken für Psychotherapie mit Jugendlichen mit verschiedenen Störungen waren größer als die für Kinder ($d = 0.65$ vs. $d = 0.45$). Dieser Alterseffekt allerdings wurde allerdings fast ausschließlich bei Mädchen gefunden. Auch im Überblick von Baskin et al. (2010) über 107 Studien zur Wirksamkeit von allgemeiner Beratung und Psychotherapie in der Schule zeigte sich ein Alterseffekt. In den 107 untersuchten Studien mit 132 Einzelmaßnahmen waren die Interventionen für Jugendliche erfolgreicher ($d = 0.59$) als für Kinder ($d = 0.35$). Für einige andere Störungsbilder wie Depression zeigen sich größere Effekte für Programme für Jugendliche (Eric Stice, et al., 2009).

Dush, Hirt und Schroeder (1989) ermittelten in einer Metaanalyse übergreifende Effektstärken aus 48 Outcomestudien zur Behandlung mit Selbstinstruktionsmethoden bei Kindern mit vorrangig aggressiven und hyperaktiv-impulsiven Verhaltensweisen. Aufgeteilt nach Altersgruppen ergaben sich folgende Ergebnisse:

5- 7 Jahre: $d = 0.39$ (Anzahl Einzeleffekte: 35)

8 – 10 Jahre: $d = 0.25$ (Anzahl Einzeleffekte: 112)

11 – 13 Jahre: $d = 0.61$ (Anzahl Einzeleffekte: 23)

14 – 16 Jahre: $d = 0.58$ (Anzahl Einzeleffekte: 32)

Die Autoren heben hervor, dass es keinen linearen Zusammenhang zwischen Alter und Effektstärke zu geben scheint und betonen, dass erwartungswidrig die jüngste Gruppe der 5- bis 7jährigen stärker von den Interventionen profitierte als die Gruppe der 8- bis 10jährigen. Dies mag aber auch mit der vielfach höheren Zahl der untersuchten Studien zusammenhängen. Die Autoren betonen, dass Selbstinstruktionsmethoden für Adoleszenten und Präadoleszenten besonders vielversprechend seien.

Ähnliche Ergebnisse ermittelten Durlak, Fuhrman und Lampman (1991) in einer Metaanalyse zur Effektivität kognitiv-behavioraler Interventionen bei „maladaptive child-

ren“ unter Einbezug von 36 Einzelstudien. Die höchste Effektstärke ergab sich auch hier bei der ältesten untersuchten Altersgruppe. Die Effektstärken für die drei gebildeten Altersgruppen ergaben sich folgendermaßen:

5 – 7 Jahre: $d = 0.57$

7 – 11 Jahre: $d = 0.55$

11 – 13 Jahre: $d = 0.92$

Auch in dieser Studie wurde die mittlere Altersgruppe am meisten beforscht, so dass die mittlere Effektstärke eventuell aufgrund der höheren Zahl einbezogenen Studien am geringsten war.

In einer Metaanalyse zum Vergleich von kognitiver Verhaltenstherapie bei antisozialen Verhalten wurden 30 Studien einbezogen (Bennett & Gibbons, 2000). Das Durchschnittsalter betrug über alle Studien $M = 11.48$ ($SD = 3.23$) bei einer Spannweite von 5.0 bis 18.1 Jahren. Insgesamt wurden kleine bis mittlere Effekte ermittelt (ungewichtet $d = .48$, gewichtet $d = .23$). Das Alter der Probanden korrelierte dabei tendenziell mit $r = .33$ ($p < .10$, zweiseitig) mit den Effektstärken. Die Autoren schlussfolgern, dass aktuelle kognitiv-behaviorale Interventionen in der Behandlung antisozialen Verhaltens für Jugendliche besser geeignet sind als für Kinder.

Auch McCart, Priester, Davies und Azen (2006) konnten einen Alterseffekt in einer Metaanalyse über 71 Studien zum Vergleich behavioraler Elterntrainings (BPT) und CBT-Behandlung bei antisozialen Jugendlichen durch Regressionsanalysen identifizieren. Insgesamt erreichte BPT eine Effektstärke von $d = 0.47$ und CBT den geringeren Effekt von $d = 0.35$ in der Postmessung und von 0.31 im Follow-up. In einer Regressionsanalyse wurde jedoch deutlich, dass das Alter als signifikanter Moderator der Effektstärken wirkte ($\beta = 0.37$) und somit Alter und Effektstärke in einem positiven Zusammenhang stehen, d. h. dass ältere Kinder bzw. Jugendliche stärker von den CBT-Interventionen profitierten als jüngere Kinder.

Toplak, Connors, Shuster, Knezevis und Parks (2008) stellen in einem Überblicksartikel zur Behandlung der ADHS mit CBT, mit rein kognitiven Verfahren und mit Neurofeedback 26 Studien zusammen und kommen zu dem Schluss, dass es deutliche Hinweise auf die Wirksamkeit der CBT-Behandlung gab, allerdings konnten aufgrund der heterogenen Designs der einbezogenen Studien keine übergreifenden Effektstärken bestimmt werden. Die Autoren betonen weiterhin die Relevanz der Entwicklungsperspektive bei der Einschätzung der Wirksamkeit und weisen darauf hin, dass ältere Kinder bzw. Jugendliche vermutlich stärker von einer CBT-Behandlung profitieren als jüngere. Bachmann, Bachmann, Rief und Mattejat (2008) kommen in ihrer Zusammenfassung von 68 Übersichtsarbeiten zur Behandlung von ADHS zu dem Schluss, dass die kognitiv-behaviorale Behandlung kontrovers diskutiert werde, aber sich insgesamt mit Einschränkungen bestätigt habe und altersabhängige Effekte vermutlich relevant sind.

Einige Studien fanden aber auch keinen Zusammenhang zwischen Alter und Trainingserfolg. So z. B. Casey und Berman (1985) in einem Überblick über 75 Kontrollgruppenstudien zur psychotherapeutischen Behandlung von Kindern allgemein oder Kazdin und Wassell (1999) in ihrer Studie zur Behandlung von aggressiven Kindern und ihren Familien. Das Review von Casey und Berman beruht allerdings auf Studien der Jahre 1952 bis 1983 und ist eventuell nicht mehr aktuell. Bei der Untersuchung von Kazdin und Wassell handelte es sich nicht um eine Metaanalyse, sondern eine eigene Untersuchung mit 200 Kindern ohne Kontrollgruppe, so dass die Ergebnisse nicht unbedingt generalisierbar sind.

Zusammenfassen lässt sich, dass die kognitiv-behaviorale Behandlung von ADHS und verwandten Störungen vermutlich bei älteren Kindern bzw. Jugendlichen wirksamer ist als bei jüngeren. Dies erscheint insbesondere auf der Grundlage fortschreitender Kompetenzzuwächse in der kognitiven Entwicklung, z. B. durch reifere Entscheidungs- und Problemlöseprozesse bis ins späte Jugendalter hinein, sinnvoll zu sein. Ältere Jugendliche sind zudem stärker von dem Druck der zu bewältigenden Entwicklungsaufgabe Übergang in das Berufsleben betroffen. Daher ist zu erwarten, dass sie ein Lerntraining, das ihnen genau bei der Bewältigung dieser Herausforderung helfen soll, stärker annehmen, dessen Nutzen höher bewerten und daher motivierter und engagierter sind. Weiterhin unterstützen viele – aber nicht alle – empirische Studien den vermuten Zusammenhang zwischen Alter und Trainingserfolg.

Initiale Störungsschwere

Es stellt sich die Frage, wie die Störungsschwere der Jugendlichen mit dem Erfolg der Maßnahme zusammenhängt. Es wäre denkbar, dass diejenigen mit einer höheren Ausprägung der Problematik stärker vom Training profitieren, weil das Training gerade für ihre Bedürfnisse konzipiert wurde und sie aufgrund der stärkeren Problematik auch mehr Möglichkeiten haben, sich zu verbessern, im Sinne eines kompensatorischen Effektes. Es könnte aber auch genau andersherum vermutet werden, dass gerade diejenigen mit einer geringer ausgeprägten Problematik stärker profitieren, weil die Störung bei ihnen vermeintlich weniger hartnäckig ist und eher auf eine Behandlung anspricht („Matthäus-Effekt“; Klauer, 2001). Ein größere Störungsschwere und höheres Ausmaß funktionaler Beeinträchtigungen wurden als Risikofaktoren für einen ungünstigen Störungsverlauf und höhere Persistenz der ADHS identifiziert (Cherkasova et al., 2013).

Empirische Studien sind uneindeutig. Studien in der Psychotherapieforschung ergaben, dass eine hohe initiale Störungsschwere häufig mit schlechten Interventionserfolgen verbunden ist, unter anderem in der Behandlung von Depression und Suchterkrankungen (Clarkin & Levy, 2004, S. 200). Analog berichtet auch Mohr (1995) von höheren Misserfolgen bei einer hohen Problemstärke, unterscheidet dabei allerdings nach Quelle der Einschätzung. Wird die Problemstärke von einer externen Person, z. B. der Therapeutin als hoch beurteilt, zeigen sich die erwartete negative Auswirkung auf den Therapieerfolg. Wird die Problemstärke allerdings von den Klienten selbst eingeschätzt, so

zeigt sich der Zusammenhang genau anders herum. Die Fremdbewertung scheint stärker auf einer allgemeinen längerfristigen Einschätzung zu basieren, während die Selbsteinschätzung eher spezifische und aktuelle Problemlagen abbildet. Selbsteinschätzungen wurden auch als Maß der Motivation der Klienten herangezogen, wobei eine selbst wahrgenommene höhere eigene Problembelastung eher dazu motiviert, etwas zu verändern, während Klienten mit schwächerer Problematik eher ein höheres Risiko für einen schlechten Therapieverlauf aufweisen (Mohr, 1995). Mohr fasst zusammen, dass Patienten, die ihre eigene Symptomatik als gering einschätzen während die Probleme von anderen Beurteilern als schwerwiegend eingestuft werden, besonders gefährdet sind, Verschlechterungen in der Therapie zu erleben.

Für 1776 Psychotherapien an einer verhaltenstherapeutischen Hochschulambulanz konnten Jacobi et al. (2011) feststellen, dass therapeutischer Misserfolg assoziiert war mit Komorbidität und schlechteren Werten zum Prämesszeitpunkt. Keine Zusammenhänge wurden mit Alter und Geschlecht der Klienten festgestellt. Ein höherer Schulabschluss war mit einer geringeren Mißerfolgsrate verbunden. Keine Zusammenhänge zeigten sich hinsichtlich der Therapieerfahrung der Therapeuten. Längere Therapien waren tendenziell weniger erfolgreich. Für Programme mit depressiven Jugendlichen stellten Stice und Kollegen (2009) allerdings höhere Wirksamkeiten für Jugendliche mit einem höheren Risiko fest. Für die Psychotherapie mit Kindern allgemein vermutet Kazdin (2004, S. 563): „Many characteristics of the child are likely to influence treatment outcome. Severity, chronicity, and scope of dysfunction are obvious contenders. One might expect that treatment would be less effective as a function of how severe, chronic, and pervasive the problem is.”

Ein Matthäus-Effekt zeichnete sich in der MTA-Studie für die 289 Kinder ab, die eine rein medikamentöse oder eine kombinierte Behandlung bekamen (Owens et al., 2003). Je höher die initiale Störungsschwere war, desto weniger gut waren die Interventionserfolge. Ein ähnlicher Effekt zeigte sich auch in einem kognitiven Training mit 120 7-13jährigen Kindern mit aggressiven und antisozialen Verhaltensweisen (Kazdin & Crowley, 1997). Je geringer die Probleme vor Beginn der Intervention (Lern-Leistungsschwierigkeiten und Anzahl von Symptomen), desto größer war der Interventionserfolg.

Es finden sich aber auch Hinweise auf kompensatorische Wirkeffekte z. B. in einer Studie von Hemphill und Littlefield (2006) mit 106 Kindern, die mit einer kognitiv-behavioralen Kurzzeitintervention aufgrund externalisierender Verhaltensprobleme behandelt wurden. Initial höhere emotionale und Verhaltensprobleme waren – bei gleichzeitig niedrigen Aufmerksamkeitsproblemen – Prädiktoren für eine größere Verbesserung des Lern-Leistungsverhaltens neben positiven Interaktionen zwischen Eltern und Kindern. Allerdings fehlte bei dieser Untersuchung eine Kontrollgruppe. Auch in zwei Studien von Kazdin und Whitley (2006) zeigten sich kompensatorische Effekte. Es wurden 183 Kindern mit einer Störung mit oppositionellem Trotzverhalten und 132 Kindern mit einer Störung des Sozialverhaltens mit evidenzbasierten Interventionen

behandelt. Es zeigte sich für beide Störungsbilder, dass die Kinder mit einer stärkeren Störungsschwere die stärkeren Verbesserungen aufwiesen. Allerdings fehlten auch hier Kontrollgruppen. Ebenfalls im Bereich der Förderung exekutiver Funktionen treten kompensatorische Effekte auf. Diamond (2012) stellt unter Bezugnahme auf mehrere empirische Studien fest, dass die Kinder mit den schlechtesten Exekutivfunktionen am meisten von Interventionen und Programmen profitierten, die eben diese Funktionen zu verbessern suchen.

In einem US-amerikanischen RCT im Bereich der allgemeinen medizinischen Erstversorgung („primary care“) wurden 163 Kindern mit Verhaltensproblemen im Alter von 6-11 Jahren untersucht (Kolko, Cheng, Campo & Kelleher, 2011). Die Kinder bekamen zufällig zugeteilt entweder eine Standardbehandlung oder gemäß eines standardisierten Protokolls eine evidenzbasierte Behandlung, die zu großen Teilen aus kognitiv-verhaltenstherapeutischen Techniken bestand. Kinder, die vor Beginn der Intervention ein höheres Niveau an Depression, Ängstlichkeit und Familienkonflikten aufwiesen, erreichten eine stärkere Verbesserung ihrer allgemeinen Gesundheit zwölf Monate später. Kinder mit größeren funktionalen Einschränkungen hatten eine stärkere Reduktion des Schweregrads im Vergleich zu Kindern mit geringerer initialer Problematik. Allerdings hatten die Kinder mit der schwerwiegenderen Problematik vor der Intervention auch nach der Intervention noch schlechtere Werte im Vergleich mit den anderen Kindern. Es fand also eine Kompensation und Varianzverringern in der Gesamtgruppe statt, aber keine Normalisierung der höher belasteten Kinder.

Schließlich zeigen sich auch unklare Ergebnisse und Studien, die keinen Zusammenhang zwischen der initialen Störungsschwere und Interventionserfolgen finden. Dies war z. B. der Fall bei van der Oord, Prins, Osterlaan und Emmelkamp (2008) in einer Studie zur Behandlung von 50 Kindern mit ADHS mit Methylphenidat und einer Kombination mit Verhaltenstherapie. Die von den Lehrern eingeschätzten Ergebnisse konnten – wenn auch nur zu einem geringen Ausmaß – mittels Maßen der Intelligenz und Ängstlichkeit vorhergesagt werden, für die Elterneinschätzung gab es aber keinen signifikanten Prädiktor. Weder initiale Störungsschwere noch elterliche Depressivität hatten einen signifikanten Einfluss auf das Behandlungsergebnis. In dieser Studie fehlte allerdings eine Kontrollgruppe. In einer Metaanalyse fassten Weisz, Jensen-Doss und Hawley (2006) 32 RCTs zu direkten Vergleichen evidenzbasierter Intervention und Standardbehandlungen („usual care“) zusammen. Zielgruppe waren Jugendliche, vorrangig mit Störungen des Sozialverhaltens und Delinquenz. Die evidenzbasierten Verfahren erreichten signifikant besser Erfolge als die Kontrollbedingungen. Die Wirksamkeit wurde dabei nicht geringer, wenn die Probleme vor Beginn der Intervention schwerwiegender waren. So bestimmten die Autoren zusammengefasste Effektstärken für die Studien mit Teilnehmern mit schwerwiegender Problematik ($N_{Studien} = 21$) und mit geringerer initialer Störungsschwere ($N_{Studien} = 11$). Die Kriterien dafür waren stationäre Behandlung, aktuelle oder vorherige Inhaftierung oder Unterbringung in einer Wohneinrichtung. In der hochproblematischen Stichprobe ergab sich eine Effektstärke von $d =$

0.32 und in der gering problematischen Gruppe von $d = 0.27$. Der Unterschied der Effektstärken war dabei nicht signifikant.

Komorbidität

Es liegt die Vermutung nahe, dass Behandlungserfolge bei zusätzlich vorliegenden Komorbiditäten geringer ausfallen. Kadzin und Whitley (2006) sprechen in diesem Zusammenhang von der „complexity hypothesis“, bei der davon ausgegangen wird, dass geringere Interventionserfolge zu erwarten sind je komplexer eine Problematik sich äußert bzw. je mehr Komorbiditäten vorliegen. Neben der theoretischen Annahme dass eine Hauptstörung andere Störungen nach sich zieht, ist nicht auszuschließen, dass sich ein grundlegendes Problem in unterschiedlichen Bereichen und Störungen äußert. Es ist anzunehmen, dass eine breiter ausgeprägte Problematik schwieriger zu behandeln ist, als eine auf einen einzelnen Bereich beschränkte, „reine“ Störung. Ein weiterer „scope of dysfunction“ (Kazdin, 2004, S. 563) ist vermutlich auch mit einer größeren Schwere und stärkerer Chronizität einer Störung verbunden und lässt geringere Behandlungserfolge erwarten. Beispielsweise wurden in der Psychotherapie mit Erwachsenen geringere Verbesserungen im Zuge einer Behandlung festgestellt, wenn gleichzeitig eine komorbide Persönlichkeitsstörung vorlag (Clarkin & Levy, 2004, S. 202). Als Risiko für einen ungünstigen Verlauf der ADHS wurden ab dem Schulalter auch vorliegende Komorbiditäten identifiziert. Insbesondere seien dies Störungen mit oppositionellem Trotzverhalten und Störungen des Sozialverhaltens, wie Cherkasova, Sulla, Dalena, Pondé und Hechtman (2013) auf Grundlage prospektiver Follow-up-Studien zusammenfassen.

Es liegt ein empirischer Überblick zur Komorbidität als Prädiktor und Moderator von Behandlungsergebnissen speziell bei Jugendlichen vor (Ollendick, Jarrett, Grills-Taquechel, Hovey & Wolff, 2008). Die Autoren erfassten 12 RCI-Studien zur Behandlung von ADHS, in denen Angaben zum Zusammenhang von Komorbidität und Trainingserfolg gemacht wurden. Fünf der Untersuchungen stammten von der MTA-Studie. Erwartungswidrig wurde festgestellt, dass in den meisten Untersuchungen nur relativ wenige Hinweise auf schlechtere Behandlungsergebnisse bei höherer Komorbidität berichtet wurden. Die Ergebnisse waren allerdings nicht gänzlich eindeutig. So wurden sogar in verschiedenen Auswertungen der MTA-Daten je nach Auswertungsmodus unterschiedliche Ergebnisse gefunden, die teils für einen moderierenden Effekt der Komorbidität sprachen, teilweise aber auch nicht.

Weiterhin liegen Hinweise darauf vor, dass es relevant ist, welche Art von Komorbidität vorliegt und ob es sich dabei um eine internalisierende oder eine externalisierende Störung handelt. In einer MTA-Auswertung von Hinshaw (2007) zeigte sich, dass die behaviorale Intervention für Kinder mit einer zusätzlichen Angststörung am erfolgreichsten war. Zu einem ähnlichen Ergebnis kommt auch eine aktuellere Studie mit 68 Jugendlichen mit ADHS und Komorbiditäten, die mit einem kognitiv-behavioralen Training behandelt wurden (Antshel, Faraone & Gordon, 2012). Geringere Trainingserfolge

in der Eltern- und Lehrereinschätzung erzielten dabei die Jugendlichen mit einer komorbiden Störung mit oppositionellem Trotzverhalten. Die Jugendlichen mit einer komorbiden Ängstlichkeit oder depressiven Störung hingegen erreichten größere Trainingserfolge. Diese Erfolge der Jugendlichen mit komorbider internalisierender Störung waren für einige Erfolgsvariablen sogar größer als die der Jugendlichen ohne Komorbidität. Die Autoren verweisen auf die Literatur, die gute Wirksamkeiten kognitiv-behavioraler Interventionen bei Ängstlichkeit und Depression im Jugendalter belegen (z. B. Chu & Harrison, 2007).

Trainingsdauer bzw. Anzahl der Sitzungen

Für eine größere Wirksamkeit einer kürzeren Intervention spricht, dass die Inhalte am Stück zusammenhängend vermittelt werden könne. Es lässt sich leichter auf Vorangegangenes aufbauen und Inhalte schneller durcharbeiten. Der Trainer hat einen guten Überblick und stetigen Kontakt zum Jugendlichen, so dass eine gute Überprüfung der Trainingseffekte in Form von Evaluation und Anpassung möglich sein müsste. Stice und Kollegen (2009) fanden größere Effekte für kürzere Präventionsprogramme für depressive Jugendliche. Auch für Psychotherapien im ambulanten Setting waren längere Therapien tendenziell weniger erfolgreich (Frank Jacobi, Uhlmann & Hoyer, 2011).

Eine kurze Trainingsdauer spricht oft auch für eine gute Organisation durch Eltern. Wenn es länger dauert, liegt es oft daran, dass die Eltern es nicht schaffen, die Jugendlichen regelmäßig zum Training zu bringen. Längere Trainingsdauer kann also auch ein Indikator für eine schlechter gelingende Organisation der Eltern sein. Kazdin (1996) spricht in diesem Zusammenhang von einem „burden-of-treatment“ (S. 140). Obwohl Interventionen zu einer Verbesserung der familiären Situation führen sollen, stellt die Teilnahme eine Herausforderung auf verschiedenen Ebenen dar und kann den Stress innerhalb einer Familie – zumindest zunächst – noch weiter erhöhen. Dies kann auch zu Trainingsabbrüchen führen, die nicht unbedingt daraus resultieren, dass die Intervention als nicht angemessen oder hilfreich wahrgenommen wird.

Für eine größere Wirksamkeit längerer Trainings spricht, dass die Jugendlichen über einen längeren Zeitraum begleitet werden und so schwerwiegende Probleme, die nicht schnell lösbar sind, langfristig bearbeitet und verbessert werden können. Eine längere Trainingsdauer entspricht oftmals auch dem Wunsch der Eltern. Größere Effekte für längere Interventionen wurden bei anderen Störungsbildern, z. B. für die Prävention von Nikotinkonsum (Rooney & Murray, 1996) und Essstörungen (Stice & Shaw, 2004), gefunden. Ein positiver Zusammenhang zeigte sich auch in der Studie von Kolko und Kollegen (2011). Von den Kindern mit Verhaltensproblemen, die mit dem kognitiv-behavioralen Programm behandelt wurden, zeigten diejenigen mit mehr absolvierten Therapiestunden größere Verbesserungen ihres allgemeinen Gesundheitsstatus.

Es liegen Metaanalysen für das Kindesalter vor. Diese finden übereinstimmend keine signifikanten Zusammenhänge zwischen Dauer und Erfolg der Interventionen (Fabiano et al., 2009; Hodgson et al., 2012). Van der Oord et al. (2008) sammelten Artikel, die

zwischen 1985 und 2006 veröffentlicht wurden. Es wurden nur Studien eingeschlossen, die Kinder im Alter von 6 bis 12 Jahren mit einer ADHS-Diagnose berücksichtigten, ein randomisiertes, kontrolliertes Design aufwiesen und Eltern- und Lehrereinschätzungen berichteten. Nach diesen Kriterien wurden 24 Studien ausgewählt. Um zu überprüfen, ob die Dauer der Intervention einen Zusammenhang mit den Effekten der Interventionen aufwies, wurden Korrelationen nach Pearson bestimmt. Für die psychosozialen Interventionen wurden folgende Korrelation bestimmt: $r = 0.14$ für die Elterneinschätzung der ADHS-Symptomatik und $r = 0.34$ für die Lehrereinschätzung der Symptomatik. Die Autoren geben an, dass es sich nicht um signifikante Korrelationen handelt. Weiterhin schien die besonders lange MTA-Studie einen besonders starken Einfluss zu haben. Wurde nur diese eine Studie aus der Analyse entfernt, zeigten sich nur noch Korrelationen von $r = -0.02$ die Elterneinschätzung und $r = 0.12$ für die Lehrereinschätzung. Die Autoren kommen daher zu dem Schluss, dass mit oder Berücksichtigung der besonders langen Studie kein Zusammenhang zwischen der Dauer des Trainings und dem Trainingserfolg feststellbar war.

Auch in der Studie von Bennett und Gibbons (2000) zur kognitiv-behavioralen Behandlung bei antisozialem Verhalten zeigte sich kein Zusammenhang der Dauer und des Erfolgs der Interventionen. In dieser Analyse wurde die Trainingsdauer sehr genau differenziert. So finden sich Auswertungen der Anzahl der Sitzungen, Anzahl der Termine pro Woche, die Dauer der Sitzungen in Minuten und die Dauer der Behandlung in Wochen. Aber keine der zugehörigen Variablen korrelierte signifikant oder tendenziell mit den Effektstärken.

5 Untersuchungsplanung

In diesem Kapitel werden die allgemeine Wirksamkeitsprüfung, die differenzielle Wirksamkeitsprüfung und die Datenaufbereitung vorgestellt.

5.1 Allgemeine Wirksamkeitsprüfung

Im Anschluss wird die Evaluation des Lerntrainings LeJA im Kontrollgruppendesign vorgestellt. Die verwendeten Instrumente werden ausführlich dargestellt, da sie auch für die Überprüfung der differenziellen Wirksamkeit verwendet werden. Es wird verdeutlicht, dass es sich um objektive, reliable und valide Instrumente handelt. Die am Ende des Abschnitts vorgestellten Ergebnisse sind als vorläufig zu betrachten und werden deshalb nur kurz dargestellt. Sie finden sich ausführlich in Schramm (im Druck).

5.1.1 Evaluationsdesign

Das Versuchsdesign der allgemeinen Wirksamkeitsprüfung besteht aus drei experimentellen Gruppen. Die erste Gruppe ist eine Interventionsgruppe, die ohne Vorlauf mit dem Lerntaining behandelt wird. Die zweite Gruppe ist eine Wartekontrollgruppe, mit der keine Intervention durchgeführt wird. Die dritte Gruppe ist eine Kontrollgruppe, mit einer alternativen Intervention durchgeführt wird. Dies ist eine progressive Muskelentspannung nach Hofmann (2003), die in kleinen Gruppen von 4 bis 6 Jugendlichen stattfindet. Die Teilnehmer der Wartekontrollgruppe wurden aus ethischen Gründen im Anschluss an ihre Wartezeit mit dem LeJA-Training nachversorgt. Die Datenerhebung erfolgte zu drei Messzeitpunkten. Bei der Interventionsgruppe war der erste Messzeitpunkt (t_0) direkt vor Beginn des Lerntainings, der zweite Messzeitpunkt (t_1) direkt nach Abschluss des Trainings und der dritte Messzeitpunkt (t_2) als Follow-up ca. ein halbes Jahr nach Trainingsende. Bei der Wartekontrollgruppe fand der erste Messzeitpunkt vor Beginn der Warteperiode statt, der zweite Messzeitpunkt nach Abschluss der Wartezeit, also direkt vor Beginn des Lerntainings, und der dritte Messzeitpunkt nach Abschluss des Lerntainings (siehe Tabelle 5).

Tabelle 5: Versuchsdesign der allgemeinen Wirksamkeitsprüfung

Gruppe	Messzeitpunkte				
	T0		T1		T2
Interventionsgruppe (IG)	X	LeJA	X	-	X
Wartekontrollgruppe (WKG)	X	-	X	LeJA	X
Kontrollgruppe mit alternativer Intervention (Progressive Muskelentspannung) (EG)	X	Progressive Muskelentspannung	X	-	-

5.1.2 Eingesetzte Instrumente

Tabelle 6 zeigt die erhobenen Erfolgskriterien und eingesetzten Instrumente. Diese Instrumente wurden sowohl für die allgemeine Wirksamkeitsprüfung genutzt als auch für die später dargestellte Analyse der differentiellen Wirksamkeit. Die Erfolgskriterien können auch englischsprachig als „Outcomes“ bezeichnet werden. Bei der Auswahl der Instrumente wurden die Empfehlungen von Schulte (1993) berücksichtigt. Das Erfolgskriterium „Aufmerksamkeitsleistung“ erfasst neuropsychologische Aspekte der Aufmerksamkeitsfunktionen, die auf der Ebene der Krankheitsursachen zu verorten sind. Das Kriterium „ADHS-Symptomatik“ bezieht sich auf die Ebene der Symptomatik und die Kriterien „Strategiewissen/ Lern- und Arbeitsverhalten“ und „Allgemeine Problembelastung“ zielen auf Krankheitsfolgen.

Tabelle 6: Erfolgskriterien, Messinstrumente und Informationsquellen

Erfolgskriterium	Erhebungsverfahren (Operationalisierung)	Auskunftsquelle	Subskalen
ADHS-Symptomatik	DISYPS-KJ (Diagnostik-System für psychische Störungen im Kindes- und Jugendalter; Döpfner & Lehmkuhl, 2006)	Jugendliche Eltern Lehrer	Gesamtproblemwert Unaufmerksamkeit Impulsivität Hyperaktivität
	WRI (Wender-Reimherr-Interview; Homburger ADHS-Skalen für Erwachsene; Rösler, Retz-Junginger, Retz & Stieglitz, 2008)	Jugendliche	Gesamtwert
Strategiewissen / Lern- und Arbeitsverhalten	WLST 7-12 (Würzburger Lesestrategie-Wissenstest für die Klassen 7-12; Schlagmüller & Schneider, 2007)	Jugendliche	Gesamtwert

	(Unstandardisierter) Einschätzungsbogen	Lehrer Jugendliche	Regelkonformes Unterrichtsverhalten Lern- und Aufmerksamkeitsverhalten Gesamtdurchschnitt
Allgemeine Problembelastung	SDQ (Fragebogen zu Stärken und Schwächen; Goodman, 1997)	Jugendliche Eltern Lehrer	Gesamtwert Emotionale Probleme Verhaltensprobleme Hyperaktivität Verhaltensprobleme mit Gleichaltrigen Prosoziales Verhalten
Aufmerksamkeitsleistung	TAP (Testbatterie zur Aufmerksamkeitsprüfung; Zimmermann & Fimm, 2008)	Jugendliche	Alertness Flexibilität Go/NoGo

Im Folgenden werden die einzelnen Verfahren genauer vorgestellt. Dabei werden Anwendungsbereiche des Instrumentes, zugehörige Materialien, die Durchführungsbedingungen und Angaben zu den Gütekriterien dargestellt. Dazu werden Angaben aus den Manualen der Verfahren und weitere empirische Ergebnisse herangezogen. Weiterhin werden Angaben zur Güte der verwendeten Verfahren anhand der klassischen Testgütekriterien Objektivität, Reliabilität und Validität (Bühner, 2011) getroffen, um so die Auswahl der Verfahren zu begründen. Bezogen auf die Objektivität der Verfahren werden wenn möglich Angaben zu den Aspekten Durchführung, Auswertung und Interpretation gemacht. Zur Bewertung der Reliabilität des Verfahrens werden zumeist Angaben zur internen Konsistenz dargestellt, wobei der Kennwert Cronbachs α genutzt wird. Angaben zur Retest-Reliabilität bzw. Stabilität des Testergebnisses finden sich seltener, werden aber angegeben, falls verfügbar. Angaben zur Validität der Verfahren werden gemacht, soweit sie vorliegen. Diese sind von Verfahren zu Verfahren verschieden und beziehen sich auf unterschiedliche Aspekte der Validität, z. B. den Inhalt, die Faktorenstruktur oder die konvergente bzw. divergente Validität, die sich aus dem Vergleich mit anderen Verfahren ableitet.

DISYPS-II

Das Diagnostik-System für psychische Störungen nach ICD-10 und DSM-IV für Kinder und Jugendliche – II (Döpfner, Görtz-Dorten & Lehmkuhl, 2008) dient der Erfassung psychischer Störungen im Kindes- und Jugendalter nach ICD- und DSM-Kriterien. Es wird zur klinischen, kategorialen Diagnostik eingesetzt. Darüber hinaus ermöglicht es

eine dimensionale Einschätzung des Ausmaßes der Problematik und ist auch zur Verlaufsdokumentation für Forschungsprojekte geeignet (Manual S. 13-14). Die Symptome bzw. Kriterien aus den Klassifikationssystemen werden in Selbst- und Fremdeinschätzung einzeln auf einer vierstufigen Skala danach eingeschätzt wie zutreffend sie jeweils sind („gar nicht“, „ein wenig“, „weitgehend“, „besonders“). Symptome mit einer Ausprägung von „weitgehend“ oder „besonders“ gelten als erfüllt. Wenn nicht bekannt, soll „gar nicht“ angekreuzt werden.

Im vorliegenden Projekt wurden die Bögen zur Selbst- und Fremdeinschätzung (Eltern und Lehrer) zum Bereich ADHS eingesetzt. Da das DISYPS-II zu Beginn der Studie noch nicht vorlag, wurde die Vorgängerversion DISYPS-KJ (Döpfner & Lehmkuhl, 2006) genutzt. Da die zur Auswertung herangezogenen Teile der Fragebögen identisch sind, werden im Folgenden Angaben zur Normierung und psychometrischen Qualität des DISYPS-II dargestellt.

Die 18 Kriterien aus ICD-10 und DSM-IV werden mit 20 Items abgefragt. Analog zur Definition in den Kategoriensystemen werden drei Skalen gebildet: Aufmerksamkeitsstörungen (neun Items), Überaktivität (sieben Items) und Impulsivität (vier Items) und eine Gesamtskala, in die alle 20 Items eingehen (Manual S. 40). Es liegen ferner sechs Kompetenz-Items vor, die für die vorliegende Studie nicht ausgewertet wurden. Es liegt eine repräsentative Normierungsstichprobe mit insgesamt 713 Fremdbeurteilungsbögen (Alter 4-17 Jahre) und 316 Selbstbeurteilungsbögen (Alter 11-17 Jahre) vor.

Zur psychometrischen Qualität der Fragebögen ist zu sagen, dass durch die standardmäßige Darreichung der einheitlichen Bögen und reines Auszählen der Rohwerte eine ausreichende Objektivität sichergestellt ist. Die Reliabilität wurde mittels Konsistenzanalysen überprüft. Es ergaben sich für die Symptomskalen ausreichende bis sehr gute Werte von Cronbachs α zwischen 0.68 und 0.94. Angaben zur Testwiederholung bzw. Retest-Reliabilität liegen nicht vor. Der Fragebogen ist als inhaltlich unmittelbar valide zu bezeichnen, da die Symptomkriterien aus der Störungsdefinition direkt abgefragt werden. Die Zuordnung der Items zu den Subskalen wird in Faktoranalysen mit ausreichender Varianzaufklärung bestätigt (Manual S. 55-62). Signifikante Korrelationen im mittleren Bereich zeigten sich zwischen Selbst- und Fremdeinschätzung der Symptomskalen zwischen $r = 0.56$ und $r = 0.69$ ($n = 277$). Auch die Übereinstimmung der Fragebogenergebnisse mit einem klinischen Urteil wurde überprüft, allerdings nur anhand einer kleinen Stichprobe ($n = 23$). Es ergaben sich geringe bis mittlere Korrelationen zwischen $r = 0.38$ und $r = 0.74$ (Görtz-Dorten & Döpfner, 2009).

Es liegt weiterhin eine Untersuchung der Lehrerversionen des Fragebogens vor (Breuer, Wolff Metternich & Döpfner, 2009). Darin zeigte sich unter anderem anhand einer Inanspruchnahme-Stichprobe einer kinder- und jugendpsychiatrischen Klinik mit $N = 78$ Teilnehmern die vermutete Faktorstruktur mit zwei bzw. drei zu Grunde liegenden Faktoren. Weiterhin füllten 79 Lehrer in einer multizentrischen Studie den DISYPS-II zur Evaluation der Wirksamkeit einer Stimulanzienbehandlung aus. Unter anderem wurde

die Retest-Reliabilität überprüft. Dazu wurde der Fragebogen einmal vormittags und einmal nachmittags ausgefüllt. Es zeigten sich hohe Korrelationen zwischen $r = .90$ und $.94$. Dagegen war die Übereinstimmungen mit einem drei Tage später ausgefüllten Fragebogen nur noch moderat und lagen zwischen $r = .65$ und $.74$. Die Autoren attestieren dem Fragebogen insgesamt eine gute Validität und schlussfolgern, dass der DISYPS-II gut geeignet ist, um die Symptome der ADHS durch Lehrer einschätzen zu lassen.

Auch im Rahmen der BELLA-Studie, die ein Teil der KiGGS-Studie zur Kinder- und Jugendgesundheit in Deutschland war, wurde der DISYPS eingesetzt (Erhart, Döpfner, Ravens-Sieberer & The Bella study group, 2008). Es wurden Fragebögen von 2.863 Eltern ausgewertet. Es zeigte sich eine gute interne Konsistenz der Skalen mit Cronbachs α zwischen 0.73 und 0.90. In explorativen und konfirmatorischen Faktorenanalysen konnte die Faktorstruktur zufriedenstellend abgebildet werden. Weiterhin wurde die Übereinstimmung mit der ebenfalls von den Eltern ausgefüllten Conners' Parent Rating Scale (Conners, Sitarenios, Parker & Epstein, 1998) überprüft, die als ein Standardverfahren der Erfassung von ADHS-Symptomen gilt. Es zeigten sich moderate bis hohe Korrelationen zwischen $r = 0.73$ und 0.90. Dies wird als ein Hinweis auf eine gute konvergente Validität gewertet wird. Insgesamt schlussfolgern die Autoren, dass es sich beim DISYPS um ein reliables und valides Messinstrument handelt.

WRI

Das Wender-Reimherr-Interview (WRI; deutsche Version aus HASE - Homburger ADHS-Skalen für Erwachsene (Rösler, Retz-Junginger, Retz & Stieglitz, 2008) dient der Feststellung einer ADHS im Erwachsenenalter und kann auch zur Verlaufsbeobachtung genutzt werden. Das WRI beruht auf den von Paul Wender formulierten Kriterien zur ADHS im Erwachsenenalter (Wender, 2002). Zusätzlich zu den aus dem Kindesalter bekannten Symptomgruppen Aufmerksamkeitsstörungen, Überaktivität und Impulsivität werden die Bereiche Temperament, Affektive Labilität, Emotionale Überreagibilität und Desorganisation abgefragt. Pro Skala finden sich drei bis fünf Items, die von einem klinischen Interviewer mit der von ADHS betroffenen Person abgefragt werden. Jedes Item soll auf einer dreistufigen Skala (0 = „nein“, 1 = „leicht“, 2 = „mittel bis schwer ausgeprägt“) danach eingeschätzt werden, inwieweit es zutrifft. Zu jedem Bereich kann der klinische Interviewer eine zusätzliche Gesamtbeurteilung abgeben, welche die Symptomschwere und die mit der Symptomatik verbundenen Behinderungen und Einschränkungen berücksichtigen soll. Weiterhin finden sich Zusatzfragen, u. a. zum Verlauf früherer Partnerschaften und zur Verbreitung von ADHS in der Familie. Zur Auswertung werden die Punkte zu jedem Bereich addiert und als Summenscore für jede Skala errechnet. Zusätzlich wird ein Gesamtscore als Summe aller Skalen bzw. aller Items bestimmt. Für die vorliegende Untersuchung wurden sowohl die Skalen- als auch der Gesamtscore ausgewertet. Die Gesamtbeurteilung durch den Interviewer und die Zusatzfragen wurden nicht durchgeführt.

Zur Objektivität des WRI ist zu sagen, dass eine standardmäßige Instruktion durch vorformulierte Fragen im Dokumentationsbogen gegeben ist. In den Hinweisen für den Interviewer wird empfohlen, zunächst die vorformulierten Fragen zu stellen und im Anschluss frei formulierte Ergänzungsfragen, falls sich die Symptomatik nicht eindeutig klären lässt. Die Items bestehen jeweils aus bis zu drei Einzelfragen, z. B. „Haben Sie Probleme sich zu konzentrieren? Sind Sie häufig geistesabwesend? Sind Sie vielfach ein Tagträumer?“ Je nach Schwerpunkt, den die Person mit ADHS bei diesen mehrfachen Fragen setzt, und in etwaiger Beeinflussung durch freie weitere Fragen des Interviewers ist davon auszugehen, dass die Fragen nicht immer gleich gestellt und verstanden werden. Diese nur teilweise zutreffende Standardisierung bedroht die Objektivität der Durchführung. Auch die Auswertungsobjektivität ist nicht selbstverständlich gegeben, da der Interviewer den Schweregrad einschätzt.

Bezogen auf die Objektivität wird im Manual des Verfahrens auf eine gute Interrater-Reliabilität des Gesamtwertes von 0.92 (Intraklassen-Korrelation) hingewiesen, die allerdings nur anhand einer kleinen Stichprobe von 18 Personen überprüft wurde (Rösler et al., 2008). In einer etwas größeren Stichprobe von 31 Patienten aus einer ADHS-Spezialsprechstunde in der Schweiz wird ebenso eine hohe Interrater-Reliabilität berichtet, die mittels Cohens κ überprüft wurde. Lediglich 2 von 28 Items lagen im Bereich einer mittleren Übereinstimmung ($\kappa = 0.41-0.60$), alle anderen Items darüber, 12 Items sogar mit $\kappa > 0.81$ (Corbisiero, Buchli-Kammermann & Stieglitz, 2010). In einer Stichprobe mit 168 Patienten einer ADHS-Ambulanz wurde eine Übereinstimmung von $\kappa = 1.0$ für die Feststellung der ADHS-Diagnose bestimmt. Für die Summenwerte des WRI wurde eine sehr hohe Intraklassen-Korrelation von 0.98 ermittelt (Rösler et al., 2008).

An dieser Stelle sei zur Begriffsklärung darauf hingewiesen, dass die Interrater-Reliabilität dem Gütekriterium der Objektivität zuzuordnen ist. Es wird überprüft, ob zwei Beurteiler zu einer übereinstimmenden Auswertung und Interpretation einer Beobachtung gelangen, also ob Auswertung und Interpretation unabhängig vom Beurteiler und damit objektiv sind (Fisseni, 2004).

Die Überprüfung der internen Konsistenz als Maß der Reliabilität wurde ebenso anhand der Stichprobe von 168 Patienten einer ADHS-Ambulanz bestimmt (Rösler et al., 2008). Für die sieben Subskalen ergaben sich Werte von Cronbachs α zwischen 0.60 und 0.72 und für die Gesamtskala $\alpha = 0.82$. Corbisiero und Kollegen (2010) untersuchten 540 Erwachsenen mit ADHS, die sich bei einer ADHS-Spezialsprechstunde in der Schweiz meldeten. Für die Gesamtskala wurde ein Cronbachs α von 0.88, für die Einzelskalen zwischen 0.64 und 0.80.

Das Verfahren wird von den Testautoren auch zur Verlaufsbeobachtung empfohlen, Angaben zur Retestreliabilität bzw. Stabilität finden sich aber nicht.

Zur Überprüfung der konvergenten Validität des WRI wurde die Übereinstimmung mit einem ADHS-Selbsteinschätzungsbogen der Testautorengruppe ermittelt (ADHS-SB; (Rösler et al., 2008). Es ergab sich zwischen WRI-Gesamtwert und Summenwert des

ADHS-SB eine Produkt-Moment-Korrelation von $r = 0.70$. Auch Corbisiero et al. (2010) berichten mittlere bis hohe Übereinstimmungen mit anderen Verfahren als Aspekt der konvergenten Validität (Spearman Korrelationen zwischen $\rho = 0.47$ und 0.79). Corbisiero et al. (2010) konnten die Faktorstruktur mit den sieben Faktoren gemäß der Kriterien nach Wender (2002) replizieren.

WLST

Der Würzburger Lesestrategie-Wissenstest für die Klassen 7-12 (WLST; Schlagmüller & Schneider, 2007) ist ein standardisiertes Testverfahren zur Erfassung des Lesestrategiewissens von Schülern der siebten bis zwölften Klassenstufe. Er kann als Einzel- oder Gruppentest durchgeführt werden und besteht aus sechs Fragen zu verschiedenen Lernsituationen. Zu jeder Frage werden Auswahlmöglichkeiten vorgegeben, die von den Schülern mittels Schulnoten danach bewertet werden sollen, wie gut sie geeignet scheinen, um die zuvor geschilderte Aufgabe zu meistern (z. B. „Die Lehrkraft fragt die Schülerinnen und Schüler, was sie tun, um einen Text möglichst vollständig zu verstehen. Hier sind einige Antworten, die du bewerten sollst: 1. Ich lese zunächst den letzten Abschnitt des Textes und gehe den Text dann von vorne durch.“). Die Antwortmöglichkeiten wurden durch Experten in eine Rangfolge gebracht. Im Vergleich mit dieser Rangreihe erhalten die Schüler Punkte für ihre Antworten.

Es liegen Normen für die Klassen 7 bis 12 aus einer nicht-klinischen Gesamtstichprobe von 4490 Schülern aller Schulformen bis auf Förderschulen vor. Verteilungsparameter werden getrennt nach Klasse, Geschlecht und Schulform angegeben. Der Rohwert beträgt im Mittel $M = 60.7$ bei einer Standardabweichung von $SD = 11.6$. Mädchen erzielten signifikant bessere Werte ($p < 0.001$, $d = 0.39$). Daher wurden auch nach Geschlecht getrennte Parameter bestimmt. Die Jungen erzielten einen Mittelwert von 58.4 ($SD = 12.2$, $n^1 = 2131$), die Mädchen 62.8 ($SD = 10.5$, $n = 2359$).

Die Objektivität wird durch einheitliche Fragebögen, standardisierte Instruktionen und Auswertung und Interpretation anhand der Normen sichergestellt. Die Reliabilität wurde mittels Split-Half nach Spearman-Brown ($r_{tt} = 0.88$, $N = 4490$), interner Konsistenz (Cronbachs $\alpha = 0,88$, $N = 4490$) und Retestung nach sechs Wochen mit einer Teilstichprobe ($r_{re} = 0,81$, $n = 137$) als zufriedenstellend überprüft. Lerneffekte bei mehrmaliger Testdurchführung sind nicht zu erwarten, da die Probanden keine Rückmeldung über richtige und falsche Antworten erhalten.

Die inhaltliche Validität wurde mittels Expertenurteil sichergestellt. Zur Überprüfung der kriterienbezogenen Validität wurde die Übereinstimmung des WLST mit verschiedenen Außenkriterien ermittelt. Es ergaben sich geringe bis mittlere Korrelationen mit dem im PISA-2003 erhobenen Leseverständnis ($r = 0.48$, $N = 971$), dem Lesegeschwin-

¹Anmerkung: Im Folgenden wird der Kleinbuchstabe n zur Angabe der Teilnehmeranzahl genutzt, wenn nur ein Teil der Gesamtstichprobe gemeint ist. Der Großbuchstabe N bezieht sich auf die Gesamtstichprobe (Deutsche Gesellschaft für Psychologie, 2007).

digkeits- und verständnistest LGVT ($r = 0.46$, $N = 809$) und dem Intelligenztest KFT ($r = 0.41$, $N = 976$). Die Autoren schlussfolgern, dass diese mittleren Zusammenhänge erwartungsgemäß dem entsprechen, was zum Zusammenhang zwischen mit der mit dem WLST erhobenen Metakognition und den anderen Konstrukten zu erwarten sei, und attestieren eine hinreichend hohe Validität des WLST (Schlagmüller & Schneider, 2007).

Weitere Studien zur Überprüfung der Qualität des Gesamttests konnten aktuell nicht gefunden werden. In einer Untersuchung von (Schütte, Wirth & Leutner, 2012) wurde eine auf drei Fragen gekürzte Version des WLST eingesetzt ($N = 559$ Gymnasiasten der neunten Klassenstufe), für die ein zufriedenstellendes Cronbachs α von 0.82 ermittelt wurde.

SDQ

Der Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) bzw. Fragebogen zu Stärken und Schwächen (SDQ-Deu; Goodman, 2005) ist ein allgemeines Screening-Instrument für Kinder und Jugendliche mit einer Selbsteinschätzungsversion und zwei Fremdeinschätzungsversionen für Eltern und Lehrer. Die Fragebögen bestehen aus 25 Items, die vier Problemskalen (Emotionale Probleme, Verhaltensprobleme, Hyperaktivität, Verhaltensprobleme mit Gleichaltrigen) und einer Kompetenzskala (Prosoziales Verhalten) zugeordnet sind. Alle Skalen bestehen aus fünf Items. Die vier Problemskalen können zu einem Gesamtproblemwert zusammengefasst werden. Jedes Item soll auf einer dreistufigen Skala danach eingeschätzt werden, wie zutreffend es ist (gar nicht, teilweise, eindeutig). Es finden sich weiterführende Fragen zu den Schwierigkeiten, z. B. nach zeitlicher Dauer, beeinträchtigten Bereichen und Leidensdruck, die für die vorliegende Studie nicht ausgewertet wurden. Es liegt auch eine Follow-up-Version des Fragebogens vor, in der Fragen nach der Entwicklung seit dem letzten Termin gestellt werden und eine Einschätzung abfragt wird, als wie hilfreich der letzte Termin empfunden wurde. Die 25 Items sind mit denen der normalen Version identisch. Es sind Cut-off-Werte angegeben, nach denen eine Einordnung in „unauffällig“, „grenzwertig“ und „auffällig“ vorgenommen werden kann (siehe auch Becker, 2007)

Die psychometrische Güte der deutschen Version des SDQ wurde in mehreren Studien untersucht. Die Objektivität wird durch die einheitliche Darreichung in Fragebogenform, einfache Auszählung der Rohwerte und inhaltliche Interpretation der Skalenwerte durch Angabe von Cut-off-Werten sichergestellt. Zur Überprüfung der Reliabilität des Verfahrens wurden insbesondere Konsistenzanalysen durchgeführt, die zumeist eine hohe Reliabilität für den Gesamtproblemwert feststellen und etwas weniger hohe, aber zumeist noch zufriedenstellende Werte für die Einzelskalen. Tabelle 7 gibt einen Überblick über Angaben zur internen Konsistenz der drei deutschsprachigen SDQ-Versionen aus unterschiedlichen Studien.

Tabelle 7: Studien zur internen Konsistenz der deutschen Version des SDQ

Fragebogen	Studie	Cronbachs α Gesamtproblemwert	Cronbachs α Subskalen	Stichprobe
Eltern	Bettge, Ravens-Sieberer, Wietzker & Hölling, 2002	< 0.70 (nicht genauer bezeichnet)	< 0.70 für alle Skalen bis auf Aufmerksamkeitsprobleme (0.77)	N = 404, nicht-klinisch
	Woerner et al., 2002; Woerner, Becker & Rothenberger, 2004	0.82	0.58 – 0.76	N = 930, nicht-klinisch
	Becker, Hagenberg, Roessner, Woerner & Rothenberger, 2004	0.82	0.73 – 0.82	N = 124, klinisch
	Becker, Woerner, Hasselhorn, Banaschewski & Rothenberger, 2004	0.83	0.72 – 0.81	N = 543, klinisch
	Rothenberger et al., 2008	0.82	0.58 – 0.79	N = 2406, nicht-klinisch
Lehrer	Becker, Hagenberg, Roessner, Woerner & Rothenberger, 2004	0.82	0.75 – 0.87	N = 124, klinisch
	Becker, Woerner, et al., 2004	0.83	0.75 – 0.83	N = 543, klinisch
	Saile, 2007	0.86	0.74 – 0.86	N = 642, nicht-klinisch
Selbst	Bettge, et al., 2002	0.71	0.46 – 0.68	N = 404, nicht-klinisch
	Becker, Hagenberg et al., 2004	0.78	0.58 – 0.78	N = 124, klinisch
	Altendorfer-Kling, Ardelt-Gattinger & Thun-Hohenstein, 2007	0.63	0.65 und 0.67 für Prosoziales Verhalten und Emotionale Probleme, übrige Skalen < 0.4	N = 2529, nicht-klinisch, Österreich

Zur Retest-Reliabilität bzw. Stabilität der deutschsprachigen Version des SDQ finden sich keine Angaben. Internationale Studien zeigen zumeist gute Retestkorrelationen, die kleiner werden, je weiter die Testzeitpunkte zeitlich auseinander liegen. Für einen sehr kurzen Abstand von zwei Wochen wurde für den Gesamtproblemwert in der Elterneinschätzung eine Korrelation von $r = 0.96$ gefunden (Smedje, Broman, Hetta & von Knorring, 1999). Bei einer Testwiederholung nach drei bis vier Wochen wurden Restkorrelationen von $r = 0.85$ für den Gesamtproblemwert und zwischen $r = 0.74$ und 0.83 für die Subskalen festgestellt (Goodman & Scott, 1999). Muris, Meesters und van den

Berg (2003) fanden bei einem Abstand von zwei Monaten Korrelationen für die Elterneinschätzung zwischen $r = 0.76$ und 0.91 und für die Selbsteinschätzung zwischen $r = 0.59$ und 0.88 , wobei nur die Werte der Skala Prosoziales Verhalten unter $r = 0.70$ lagen. Goodman (2001) fand für einen Abstand von vier bis sechs Monaten eine Korrelation von $r = 0.62$ für die Elterneinschätzung, $r = 0.73$ für die Lehrereinschätzung und $r = 0.51$ für die Selbsteinschätzung (siehe auch Becker, 2007).

Die meisten faktoranalytischen Untersuchungen bilden die ursprüngliche Skalenstruktur ab, so z. B. Woerner (2002, 2004) für die Elterneinschätzung; Becker, Woerner et al. (2004) für die Eltern- und Lehrereinschätzung; Saile (2007) für den Lehrerfragebogen und Becker, Hagenberg et al. (2004) für die Selbsteinschätzung. Nicht ideal konnte die Faktorstruktur bei einer großen österreichischen Stichprobe mit $N = 2529$ gefunden werden (Altendorfer-Kling et al., 2007).

Zur Überprüfung der konkurrenten Validität wurde die Übereinstimmung des SDQ mit der Child Behavior Checklist ermittelt (CBCL, deutsche Version; Döpfner, Schmeck & Berner, 1994) bzw. des zugehörigen Lehrerfragebogens (TRF, deutsche Version; Döpfner, Berner & Lehmkuhl, 1994a) und Selbsteinschätzungsbogens (YSR, deutsche Version; Döpfner, Berner & Lehmkuhl, 1994b). Die im Vergleich zum SDQ umfangreichere CBCL wurde in zahlreichen Studien evaluiert und kann als „Goldstandard“ herangezogen werden (Klasen et al., 2000). Studien fanden zumeist hohe Übereinstimmungen des SDQ mit der CBCL, z. B. Klasen et al. (2000); Woerner et al. (2004); Becker, Hagenberg et al. (2004) und Becker, Woerner et al. (2004).

Saile (2007) fand sehr gute Übereinstimmungen des SDQ-Lehrerfragebogens mit soziometrischen Kennwerten (Wahl von Sitznachbarn in der Schulklasse). Bettge und Kollegen (2002) überprüften die diagnostische Qualität der SDQ-Eltern- und Selbsteinschätzung durch die Übereinstimmung der Einordnung als psychisch auffällig versus unauffällig mittels SDQ mit einer Klassifikation aufgrund eines klinischen Interviews ($N = 155$). Der SDQ konnte richtig positiv 46 % (Elternurteil) bzw. 43 % (Selbsturteil) und richtig negativ 87 % (Elternurteil) bzw. 93 % (Selbsturteil) erkennen. Dabei war der SDQ der CBCL in etwa ebenbürtig, sowohl in der Einschätzung, ob überhaupt eine psychische Störung vorliegt oder nicht, als auch bezüglich einzelner Störungsbilder. Auch Becker, Hagenberg et al. (2004) stellen fest, dass die diskriminative Validität des SDQ-Selbst vergleichbar mit der des längeren Fragebogens YSR war und die beide Verfahren gleich gut Patienten mit und ohne psychiatrischen Diagnosen unterscheiden konnten.

Zusammenfassend kann gesagt werden, dass insbesondere der Gesamtwert des SDQ eine hohe Konsistenz aufweist. Die Subskalen erreichen meist niedrigere Werte für Cronbachs α , was auch damit zusammenhängen mag, dass nur jeweils fünf Items zu einer Skala gehören. Faktoranalysen bilden die Skalen meist gut ab. Die Übereinstimmung mit anderen Verfahren im Sinne einer konkurrenten Validität sind im Wesentlichen als gut zu bezeichnen. Der SDQ zeigt damit insgesamt gute psychometrische Eigenschaften, die sowohl für die deutschsprachige Version überprüft wurden und sich

mit internationalen Forschungsergebnissen decken. Der SDQ kann somit als ein Standardverfahren zur Erfassung psychischer Auffälligkeiten des Kindes- und Jugendalters bezeichnet werden.

TAP

Die Testbatterie zur Aufmerksamkeitsprüfung (TAP; Zimmermann & Fimm, 2009) ist ein computergestütztes Verfahren zur Erfassung von Aufmerksamkeitsfunktionen. Ursprünglich wurde die Testbatterie zur neuropsychologischen Diagnostik mit Erwachsenen mit Hirnverletzungen entwickelt. Mittlerweile liegt die TAP in Version 2.2 vor. Für Kinder wurde die Kinderversion der Testbatterie zur Aufmerksamkeitsprüfung (KiTAP; Zimmermann, Gondan & Fimm, 2002) entwickelt, bei der die Aufgaben in Geschichten eingebettet sind, um die Motivation bei der Testbearbeitung sicherzustellen. Die KiTAP eignet sich nicht für den Einsatz bei Jugendlichen, da die Geschichten nicht mehr altersangemessen sind. Für die vorliegende Untersuchung wurde die TAP-Version 2.1 genutzt (Zimmermann & Fimm, 2008). Die TAP besteht aus 13 Untertests, die verschiedene Komponenten der Aufmerksamkeit erfassen. Aufmerksamkeit wird vor dem Hintergrund eines Informationsverarbeitungsparadigmas als Zusammenspiel von Teilleistungen verstanden, die entsprechend separat erfasst werden. Um sicherzustellen, dass die Instruktion verstanden wurde, beginnt jeder Subtest mit einem kurzen Vortest, der nicht ausgewertet wird.

Für die vorliegende Untersuchung wurden die Untertests Alertness, Flexibilität und Go/NoGo durchgeführt. Untersuchungen haben gezeigt, dass in diesen Bereichen besondere Schwierigkeiten von Probanden mit ADHS zu erwarten sind. Die Befundlage zur klinischen Validität der TAP ist allerdings nicht ganz einheitlich und es ist nicht endgültig geklärt, welche Aufmerksamkeitskomponenten resp. Subtests der TAP zwischen Personen mit und ohne ADHS unterscheiden. In empirischen Studien wurden zumeist nur kleinere Stichproben erhoben, teilweise mit jüngeren Probanden mit der KiTAP (Drechsler, Rizzo & Steinhausen, 2009; Dreisörner & Georgiadis, 2011; Földényi, Imhof & Steinhausen, 2000; Kaufmann et al., 2010; Tucha et al., 2006).

Die Normierung der TAP wurde von den Autoren selbst und in unterschiedlichen anderen Projekten durchgeführt und ist somit uneinheitlich was die Größe und Zusammensetzung der Stichprobe angeht. Für die meisten Untertests liefert das Programm Prozentränge, die nach Alter, Geschlecht und Schulbildung (bei Erwachsenen) bestimmt werden.

Beim Subtest Alertness wird über die Messung der Reaktionszeiten die allgemeine Wachheit erfasst. Die Probanden sollen möglichst schnell auf eine Taste drücken, sobald ein Kreuz auf dem Bildschirm erscheint („intrinsische Alertness“). In einer zusätzlichen Bedingung wird als Hinweisreiz ein Warnton vorgegeben bevor das Kreuz erscheint („phasisches Arousal“). Die Messung beginnt mit einem Durchlauf ohne Warnton, es folgen zwei Durchläufe mit Warnton und den Abschluss bildet erneut ein Durchgang ohne Ton. Die Dauer der Messung beträgt ca. 4,5 min. Es werden Mittelwert, Me-

dian, Standardabweichung der Reaktionszeit, Anzahl der richtigen Lösungen und Auslassungen jeweils für die Durchgänge mit und ohne Warnton ausgegeben. Der Median der Reaktionszeiten ohne Warnsignal gibt Auskunft über die allgemeine Reaktionszeit und mögliche generelle Verlangsamungen. Um der Schiefe der Verteilung gerecht zu werden, ist der Median aussagekräftiger als Mittelwert. Die Standardabweichung der mittleren Reaktionszeiten ohne Warnton geben Auskunft über Schwankungen der Aufmerksamkeit und somit die Stabilität der Aufmerksamkeitsleistung. Die Differenz der Reaktionszeiten für Durchgänge mit und ohne Warnton gibt Auskunft für die phasische Alertness. Weiteren Aufschluss geben Ausreißer und Antizipationen mit Warnton.

Im Untertest Flexibilität wird die Fähigkeit sich auf neue Situation einzustellen erfasst, was auch als „set shifting“ bezeichnet wird. In der für diese Studie eingesetzten Testvariante („Wechsel nonverbal“) werden immer gleichzeitig zwei Formen auf dem Bildschirm dargeboten, eine auf der linken und eine auf der rechten Seite. Von den beiden Formen ist jeweils eine eckig und die andere rund; welche Form auf welcher Seite erscheint, ändert sich zufällig. Die Probanden bekommen die Anweisung immer abwechselnd die eckige und die runde Form zu identifizieren. Sie bekommen dazu jeweils eine Taste für die linke Hand und eine für die rechte Hand. Bei den ersten beiden Formen, die auf dem Bildschirm erscheinen, soll die Taste auf der Seite mit der eckigen Form betätigt werden, bei den darauf folgenden beiden Formen die Taste auf der Seite mit der runden Form, dann wieder auf der Seite mit der eckigen Form usf. Wird die falsche Taste betätigt, zeigt bei den folgenden zwei Formen ein Rahmen die richtige Seite an. Die Testdauer beträgt mindestens 3,5 min, oftmals jedoch länger, da bei falschem Antwortverhalten zusätzliche Items vorgegeben werden. Es werden der Mittelwert der Reaktionszeit, Median und Standardabweichung ausgegeben sowie die Anzahl korrekter Reaktionen, Fehler und Ausreißer. Zusätzlich wird ein Index zum „speed-accuracy-tradeoff“ bestimmt, in den Genauigkeit und Reaktionszeit eingehen, die sich je nach Strategie des Probanden gegenseitig beeinflussen. Wird möglichst schnell reagiert, werden mehr Fehler gemacht und wenn mehr Wert auf Genauigkeit und richtiges Reagieren gelegt wird, ergeben sich längere Reaktionszeiten.

Beim Subtest Go/NoGo wird die Fähigkeit schnelle Reaktionen zu zeigen bei gleichzeitiger Unterdrückung inadäquater Reaktionen erfasst. Dazu werden zufällig abwechselnd Zielreize, die eine Reaktion erfordern, und Distraktoren, auf die keine Reaktion gezeigt werden darf, vorgegeben. In der für die Studie durchgeführten Testvariante („1 aus 2“) bekommen die Probanden die Anweisung, eine Taste zu betätigen, wenn ein Kreuz (x) auf dem Bildschirm erscheint, und nicht zu drücken, wenn ein Pluszeichen (+) erscheint. Die Stimuli sind sich also sehr ähnlich und nur in ihrer Drehung unterschieden. Die Durchführung nimmt 2 min in Anspruch. Ausgegeben wird für die Reaktionszeit der Mittelwert, Median, Standardabweichung und die Anzahl korrekter Antworten, Fehler, Auslassungen und Ausreißer. Insbesondere die Anzahl der Fehler, die entstehen, wenn die Taste bei Darbietung der Distraktoren betätigt wird, gibt Aufschluss über Probleme in der Impulskontrolle.

Die Durchführungsobjektivität der TAP scheint auf einen ersten Blick sehr gut zu sein, da die Instruktionen und Items standardmäßig am Bildschirm vorgegeben werden. Weiterhin werden vor jedem Haupttest kurze Vortests mit den gleichen Aufgaben durchgeführt, so dass sichergestellt werden kann, dass die Instruktionen verstanden wurden und angewendet werden können. Erfahrungen in der Testanwendung zeigen allerdings, dass die Instruktionen am Bildschirm oftmals nicht ausreichen, um die Anweisung richtig zu verstehen, so dass zusätzliche Erklärungen des Testleiters notwendig sind. Dreisörner und Georgiadis (2011) bemängeln in diesem Zusammenhang neben der Notwendigkeit zusätzlicher freier Instruktionen, dass die Wichtigkeit des Verhältnisses von Reaktionszeit und Genauigkeit nicht genügend erklärt werde. So bleibe unklar, ob möglichst schnell oder möglichst genau vorgegangen werden solle, mit der Folge dass Fehler nicht eindeutig zugeordnet werden können. Weiterhin bleibe unklar, wie der Testleiter reagieren solle, wenn Kinder unmotiviert seien oder sich entgegen den Anweisungen verhielten.

Die Auswertungsobjektivität der TAP ist als hoch zu bezeichnen, da die Auswertung vom Programm übernommen wird, welches die relevanten Kennwerte liefert. Die Interpretationsobjektivität ist etwas eingeschränkt, da die Normstichprobe nicht einheitlich ist. Für die weitere Auswertung in der vorliegenden Untersuchung werden daher die Rohwerte statt der Normwerte verwendet.

Als ein Aspekt der Reliabilität wurde von den Testautoren die Konsistenz der Skalen anhand der Odd-even-Reliabilität überprüft. Bei der Odd-even-Methode als einer Variante der Testhalbierung (split half) werden pro Skala bzw. Subtest Korrelationen der geraden mit den ungeraden Testitems angegeben. Im Testmanual werden für die Reaktionszeiten für alle Subtests zumeist sehr hohe Split-half-Reliabilitäten mit zumeist $r > 0.90$ angegeben. Für die Fehler und Auslassungen sind die Reliabilitäten geringer, teilweise nicht ausreichend (Zimmermann & Fimm, 2009, S. 93). Für relevante Subskalen für die vorliegende Untersuchung siehe ausführlich Tabelle 8.

Tabelle 8: Odd-even-Reliabilität der Testbatterie zur Aufmerksamkeitsprüfung (TAP) in für die Studie relevanten Subtests (Zimmermann & Fimm, 2009)

Stichprobe	Subtest	Kennwert	Odd-even-Reliabilität
6-19 Jahre; $N = 381$	Alertness	alle	> 0.99
6-19 Jahre; $N = 184$	Go/NoGo („1aus2“)	Median der Reaktionszeit	0.995
		Fehlerreaktionen	0.588
		Auslassungen	0.69
6-12 Jahre; $N = 86$	Flexibilität nonverbal (eckige und runde Form)	Median der Reaktionszeit - Gesamttest	0.997
		Fehlreaktionen – Gesamttest	0.634

Als weiterer wichtiger Aspekt der Reliabilität ist die Retest-Reliabilität bzw. Stabilität der Testergebnisse relevant. Diese wurde von den Testautoren nur an einer sehr kleinen Stichprobe von 36 Probanden erfasst worden. Die Ergebnisse dieser und weiterer Untersuchungen sind in Tabelle 9 abgebildet. Ebenfalls eine aktuelle Studie zur KiTAP.

Tabelle 9: Studien zur Retest-Reliabilität relevanter Subtests der Testbatterie zur Aufmerksamkeitsprüfung (TAP)

Studie	Retest-Intervall und Stichprobe	Subtest	Kennwert	Retest-Korrelation (Pearson)
Zimmermann & Fimm, 2010	durchschnittlich 25 Tagen (Spannweite 18-43 Tage) 36 jüngere, gesunde Probanden	Alertness	Median der Reaktionszeit Gesamttest	0.81**
			Kennwert der phasischen Alertness	0.07
		Go/NoGo („1aus2“)	Median der Reaktionszeit	0.56**
			Fehlerreaktionen	0.73**
			Auslassungen	-0.09
		Zoccolotti et al., 2000	ein Monat 62 hirngeschädigte Patienten	Alertness
Median der Reaktionszeit mit Warnton	0.86**			
Kennwert der phasischen Alertness	0.15			
Go/NoGo („2aus5“)	Median der Reaktionszeiten			0.53**
	Fehlerreaktionen			0.59**
	Auslassungen			0.16
Földényi, Giovanoli, Tagwerker-Neuenschwander, Schallberger & Steinhausen, 2000	durchschnittlich 40 Tage Spannweite 16-81) 95 7-10jährige Kinder	Alertness	Median der Reaktionszeit Gesamttest	0.61**
			Kennwert der phasischen Alertness	0.16
		Go/NoGo („1aus2“)	Median der Reaktionszeiten	0.56*
			Fehlerreaktionen	0.53**
			Auslassungen	0.55**
		Günther, Herpertz-Dahlmann & Konrad, 2005	6 Wochen 33 gesunde Kinder	Go/NoGo („1aus2“)
Standardabweichung [ms]	0.32 (n. s.)			
Fehler	0.25 (n. s.)			

			Auslassungen	0.22 (n. s.)
Renner, Lessing, Krampen & Irblich, 2012 KiTAP wurde durchgeführt	14 – 28 Tage 64 sechs- und siebenjährige Kinder	Altertness	Median der Reaktionszeit Gesamttest	0.70 (keine Signifikanzprüfung)
			Standardabweichung der Reaktionszeit Gesamttest	0.45
		Go/NoGo („1aus2“)	Fehler	0.54
			Auslassungen	0.32
			Reaktionszeit	0.53
		Flexibilität	Fehler	0.67
			Median der Reaktionszeit Gesamttest	0.69

Es ist anzumerken, dass bei Konzentrations- und Aufmerksamkeitstests wie der TAP Lerneffekte bei wiederholter Testdurchführung zu erwarten sind. Westhoff (1989) ging für Konzentrationstests im Allgemeinen von einer Leistungssteigerung von ca. 25 % zwischen erster und zweiter Testdurchführung aus. Bühner, Ziegler, Bohnes und Lauterbach (2006) führen den TAP-Tests Go/NoGo viermal mit fünfminütigen Abständen mit 125 gesunden jungen Erwachsenen durch. Trotz unrealistisch schneller und häufiger Testwiederholung ergaben sich nur geringe bis moderate Übungseffekte für die Reaktionszeitmediane. Die Autoren führen weiterführend aus, dass noch geringere Übungseffekte bei längeren Intervallen zwischen den Testungen und nur einmaliger Wiederholung zu erwarten seien. Weiterhin wurde festgestellt, dass die Übungseffekte in der TAP deutlich geringer waren als im parallel durchgeführten Aufmerksamkeits-Belastungs-Test d2 (Brickenkamp, 2002). Für 33 Patienten mit Schlaganfall berichten Sturm et al. (2003) leichte signifikante Verbesserungen für den Test Go/NoGO für eine vierwöchige Baselinephase.

Die faktorielle Validität wurde anhand unterschiedlicher Stichproben überprüft. Es konnte keine vollständige Überprüfung der Gültigkeit des Verfahrens vorgenommen werden, da in den Überprüfungsstichproben unterschiedliche Teilmengen der Tests und unterschiedliche Stichproben untersucht wurden (Zimmermann & Fimm, 2009, S. 105). Übereinstimmend zeigt sich in Hauptkomponentenanalysen die Mehrdimensionalität der TAP. Allerdings sind die gefundenen Faktoren je nach Stichprobe nicht völlig übereinstimmend. Ergebnisse aus verschiedenen Subtests laden auf unterschiedliche Faktoren; die einzelnen Tests der TAP sind also nicht als unterschiedliche Skalen bzw. Dimensionen zu verstehen.

Die klinische Validität wurde in einigen Studien hinsichtlich der Differenzierungsfähigkeit der TAP zwischen Kindern mit und ohne ADHS überprüft. (Földényi, Imhof et al., 2000) verglichen 20 Kinder mit ADHD mit 20 Kontrollkindern mittels sechs TAP-

Tests. Mit dem Test Alertness konnten 71 % der Kinder richtig klassifiziert werden. Für die Prognose waren die Reaktionszeitschwankungen bedeutsam. Weiterhin waren die Schwankungen der Reaktionszeiten im Go/NoGo-Test relevant und die Fehlerzahlen im Test Flexibilität (bis Testversion 1.7 unter dem Namen „Reaktionswechsel“). Durch die letzten beiden Parameter und das Testalter konnten 90 % der Kontrollkinder richtig klassifiziert werden. Die positive prädiktive Validität wird mit 86 %, die negative mit 93 % beziffert. Auch in einer Untersuchung von Drechsler, Brandeis, Földényi, Imhof und Steinhausen (2005) unterschied die Variabilität der Reaktionszeit im Test Alertness – nicht aber die Reaktionszeit an sich – zwischen 28 Kindern mit ADHS und 25 Kontrollkindern. Allerdings ist nicht ersichtlich, ob es sich bei den beiden Stichproben teilweise um dieselben Kinder handelt, da es sich um dieselbe Arbeitsgruppe handelt und die Fallzahlen ähnlich sind.

In einem Vergleich von 19 Jugendlichen mit ADHS aus der Stichprobe der vorliegenden Untersuchung und 19 hinsichtlich Alter, Geschlecht, Schulform, Klasse und Medikation parallelisiert zugeordneten Jugendlichen wurden die Tests Alertness, Go/NoGo und Flexibilität durchgeführt. Die Leistungen der Jugendlichen mit ADHS waren durchweg schlechter, signifikant unterschiedlich waren die Unterschiede allerdings nur für die Schwankungen der Reaktionszeiten im Test Go/NoGo und fast in der Fehlerzahl in diesem Test. In einer ROC-Analyse konnten beide Parameter des Tests Go/NoGo signifikant zwischen den Gruppen trennen (Börnert, 2011).

Auch die KiTAP wurde hinsichtlich ihrer Differenzierungsfähigkeit untersucht. Drechsler et al. (2009) fanden signifikante Gruppenunterschiede zwischen 50 Kindern mit ADHS und 50 Kindern ohne ADHS in den Tests Go/NoGo und Alertness. Dreisörner und Georgiadis (2011) verglichen 24 Kinder mit ADHS und 24 Kontrollkinder im Alter von acht bis zehn Jahren. Alle Kennwerte des Tests Alertness und Median und Standardabweichung der Reaktionszeiten im Test Go/NoGo unterschieden die Gruppen signifikant. Eine Vorhersage des Vorliegens einer ADHS konnte am besten durch die Parameter Standardabweichung der Reaktionszeit im Untertest Alertness (ohne Warnton) durch den Median der Reaktionszeiten im Untertest Reaktionswechsel (Flexibilität) gemacht werden.

Angaben zur Übereinstimmung mit anderen Verfahren im Sinne einer konkurrenten werden im Manual nicht gemacht.

Die Studienlage zur Güte der TAP ist noch unklar. Oftmals sind die Aussagen durchgeführter Studien nicht generalisierbar, da die Stichprobengrößen zu gering sind oder da unterschiedliche Altersgruppen untersucht wurden oder da eine unterschiedliche Auswahl durchgeführter Tests gemacht wurde. Zur klinischen Validität und damit Differenzierungsfähigkeit zwischen Kindern und Jugendlichen mit und ohne ADHS scheinen insbesondere die Tests Alertness (insbesondere Schwankungen der Reaktionszeiten), Go/NoGo und Flexibilität aussagekräftig zu sein.

Als Erfolgsvariablen aus der TAP werden die Mediane und Fehlerzahlen ausgesucht. Damit sind eine Geschwindigkeitskomponente und eine Komponente der Genauigkeit berücksichtigt. Die Standardabweichungen differenzieren zwar recht gut zwischen ADHS betroffenen und unauffälligen Kindern, dies ist für den Trainingserfolg aber weniger wichtig. Wichtiger ist, dass die Jugendlichen zügig arbeiten und dabei wenig Fehler machen.

AVL

Der Fragebogen zum Lern- und Arbeitsverhalten (AVL) wird von den Lehrern ausgefüllt. Er besteht aus elf Items mit erwünschten Verhaltensweisen, die in Prozent danach eingeschätzt werden sollen, wie sehr sie in der letzten Woche zutrafen. Der Fragebogen wurde erstmalig von Lauth und Fellner (2004) zur Einzelfallevaulation einer Förderung von Kindern mit ADHS eingesetzt und von Linderkamp, Hennig und Schramm (2011) modifiziert.

Für eine gute Durchführungsobjektivität des AVL spricht die standardmäßig dargebotene Instruktion auf dem Fragebogen und dass sich die Items auf proximale schulalltägliche Verhaltenskriterien beziehen, die von den Lehrern direkt eingeschätzt werden können. Ein Manko an der Formulierung der Instruktionen ist, dass nicht vorgegeben oder abgefragt wird, welche Bezugsnorm der einschätzende Lehrer anlegen soll bzw. angelegt hat. So ist denkbar, dass der Lehrer die aktuellen Verhaltensweisen mit vorangehendem Verhalten des gleichen Schülers vergleicht (ipsativ), mit dem Verhalten der Mitschüler (sozial) oder mit dem, was er eigentlich als zu erwarten voraussetzt (kriterial). Wird der AVL allerdings zur Verlaufsevaluation bei demselben Schüler vom selben Lehrer genutzt, erscheint dies wenig problematisch, da zu erwarten ist, dass stets dieselbe Bezugsnorm angelegt wird.

Eine Auswertung im herkömmlichen Sinne ist nicht notwendig, da es sich bei allen Items um positiv formulierte erwünschte Verhaltensweisen handelt. Je höher die Merkmalsausprägung, desto besser. Eine Interpretation ist damit ebenso unmittelbar gegeben. Es liegen keine Normwerte vor, die eine Einordnung der individuellen Ergebnisse in eine Vergleichsstichprobe gleichen Alters, Geschlecht und Schulform ermöglichen. Der Nutzen einer solchen Normierungsstichprobe ist allerdings fraglich, da zum momentanen Zeitpunkt keine qualitativ vergleichende Einschätzung vorgenommen werden soll, sondern ein individueller und über Individuen zusammengefasster Vorher-nachher-Vergleich.

Die endgültige Validierung des Fragebogens steht noch aus. In einer vorläufigen Analyse zur Überprüfung der psychometrischen Güte wurde eine nicht-klinische Stichprobe mit 89 Schülern im Alter von 12 bis 17 Jahren untersucht (Wietfeldt, 2010). In einer Hauptachsenanalyse wurden zwei Faktoren extrahiert, die gemäß Kaiser-Kriterium einen Eigenwert von > 1 aufweisen und auch nach Screeplot sinnvoll erscheinen. Der erste Faktor hat einen Eigenwert von 7.1 und klärt mit 64.6 % deutlich mehr Varianz auf als der zweite Faktor mit 12.5 % und einem Eigenwert von 1.378. Zusammen klären

die Faktoren 77 % der Gesamtvarianz auf. Die Items des Faktors mit dem geringeren Eigenwert und geringerer Varianzaufklärung beziehen sich inhaltlich vor allem auf ein angemessenes Sozialverhalten und Einhaltung von Regeln im Unterricht (Items 1-4), während sich die Items des Faktors mit höherem Eigenwert und Varianzaufklärung auf konkrete Lernhandlungen und Aufmerksamkeit beziehen (Items 5-11). Die durchschnittliche Item-Korrelation innerhalb der Skalen beträgt 0.66 bzw. 0.70 und ist damit deutlich höher als die durchschnittliche Item-Korrelation beider Skalen mit 0.51. Die Annahme von zwei Faktoren erscheint somit auch inhaltlich als sinnvoll interpretierbar und valide.

Nach den Ergebnissen der Faktorenanalyse erscheint es sinnvoll, für beide Faktoren getrennt je einen Indexwert zu bestimmen. Damit ergibt sich ein durchschnittlicher Wert zum prozentualen Gelingen regelkonformen Verhaltens im Unterricht und einer zu gelungenem Lern- und Aufmerksamkeitsverhalten. Diese beiden Werte lassen sich wiederum zu einem Gesamtwert zusammenfassen.

Zur Reliabilität des Fragebogens wurde eine Konsistenzanalyse für beide Faktoren resp. Skalen getrennt durchgeführt. Die interne Konsistenz der Skala „Lern- und Aufmerksamkeitsverhalten“ ist mit Cronbachs $\alpha = 0.93$ etwas höher als die der Skala „Regelkonformes Unterrichtsverhalten“ mit Cronbachs $\alpha = 0.89$, was auch mit der größeren Itemzahl des ersten Faktors zusammenhängen kann. Gerade im Anbetracht der geringen Itemanzahl der Skalen sind die ermittelten Konsistenzen als hoch zu bezeichnen (Schermelleh-Engel & Werner, 2012).

Angaben zur Retest-Reliabilität liegen bislang nicht vor. Es ist zu vermuten, dass Korrelationen zwischen mehreren Testzeitpunkten im mittleren Bereich zu erwarten sind. Es soll aktuelles Verhalten eingeschätzt werden, das je nach Tagesform, Zeitpunkt im Schulkalender (z. B. kurz vor oder nach Ferienzeiten) und nach den jeweiligen schulischen Inhalten variieren kann, so dass keine sehr große Stabilität anzunehmen ist. Dementgegen ist allerdings auch nicht davon auszugehen, dass sehr geringe Korrelationen auftreten, da immerhin das Verhalten einer gesamten Woche eingeschätzt werden soll, wenn möglich von einer Lehrperson, die viel Unterricht mit dem einzuschätzenden Schüler hat, z. B. die Klassenlehrerin.

Zur Überprüfung der klinischen Validität des Fragebogens wurden 35 Jugendliche mit ADHS aus der Stichprobe der vorliegenden Untersuchung mit 35 unauffälligen Jugendlichen verglichen, um die Differenzierungsfähigkeit des Fragebogens zu überprüfen. Da aufgrund des kleinen Stichprobenumfangs eine Radomisierung zu keiner sicheren Kontrolle von Störvariablen geführt hätte, wurden die unauffälligen Jugendlichen nach Alter, Geschlecht und Schulform parallelisiert den Jugendlichen mit ADHS zugeordnet. In *t*-Tests für abhängige Stichproben zeigten sich keine signifikanten Unterschiede für die Skala „Regelkonformes Unterrichtsverhalten“, allerdings schon für die Skala „Lern- und Aufmerksamkeitsverhalten“ bei einem kleinen Effekt von Cohens $d = 0.41$.

5.1.3 Gesamtproblemmaß

Aufgrund der Vielzahl an Outcomevariablen empfiehlt sich eine Datenreduktion durch die Bestimmung eines Gesamtproblemmaßes. Dieses Vorgehen wurde beispielsweise auch in der Auswertung der MTA-Studie von Conners et al. (2001) durchgeführt. Ein Gesamtproblemmaß kann zu einer zusammenfassenden Abschätzung des Trainingserfolges herangezogen werden. Die Datenreduktion verfolgt folgende Ziele (siehe auch Kapitel 4.2):

- *Verringerung von Redundanz* durch Zusammenfassung solcher Messinstrumente, welche dasselbe Konstrukt erfassen;
- *Gesamteinschätzung des Trainingserfolges* durch Abstraktion von möglicherweise widersprüchlichen Einzelmessungen bzw. Ausgleich von einzelnen extremen und verzerrenden Einschätzungen;
- *Erhöhung der Teststärke („Power“)*, da keine Absenkung des Signifikanzniveaus (z. B. Bonferroni-Korrektur) wegen wiederholten Hypothesentestens durchgeführt werden muss und
- *Ausgleich fehlender Werte*, da auch bei fehlenden Einzelmessungen aus den übrigen Messwerten ein Gesamtmaß ermittelt werden kann.

Zur Zusammenfassung der einzelnen Messinstrumente stehen prinzipiell zwei Strategien zur Verfügung. In einem deduktiven Vorgehen werden Theorie geleitet solche Messinstrumente zusammengefasst werden, die vermeintlich dieselben Konstrukte erfassen. Ein solches Konstrukt wäre z. B. die ADHS-Symptomstärke, die zusammenfassend aus den Einschätzungen von Eltern, Lehrern und Jugendlichen in deren DISYPS-II-Urteilen und dem WRI ermittelt würden. Dabei würde allerdings nicht geprüft, ob sich in unterschiedlichen zusammengefassten Variablen redundante Information befinden. So könnten zwei zusammengefasste Variablen, die vermeintlich unabhängig voneinander sind, lediglich beide ein weiteres drittes Konstrukt abbilden. Weiterhin wird nicht kontrolliert, ob die zusammengefassten Einzelmessungen gleichförmige Ausprägungen annehmen oder ob sie sich entgegen laufen. So könnte es sein, dass sich die Selbst- und Fremdeinschätzungen der ADHS-Symptomstärke widersprechen und dies durch bloße Zusammenfassung kaschiert würde oder zu Varianz- und Informationsverlust führte. In einem induktiven Vorgehen werden die erhobenen Daten bezüglich der Erfolgskriterien auf zu Grunde liegende Strukturen überprüft. Dazu ist ein faktorenanalytisches Vorgehen angemessen, da es sich zur Ermittlung von Faktoren bzw. latenten Variablen eignet, welche die in den Daten vorliegenden Korrelation erklären (Bühner, 2011). Nach der Identifikation solcher Items, die scheinbar dieselben Konstrukte erfassen, was durch eine hohe Korrelation deutlich wird, werden diese zusammengefasst.

Sowohl das deduktive als auch das induktive Vorgehen zur Datenreduktion sind Erfolg versprechend und gut umsetzbar. Beide ermöglichen eine zusammenfassende Aussage

über den Gesamterfolg der Maßnahme und erhöhen die Teststärke. Für diese Ausarbeitung wird das induktive Vorgehen vorgezogen. So erscheint es besser geeignet zu sein, redundante Informationen zu erkennen, und Einzelmessungen so zusammenzufassen, dass Variablen mit gleichförmigen Ausprägungen zusammengefasst werden, um ein höheres Ausmaß an Information und Varianz bei der abstrahierenden Zusammenfassung zu erhalten. Analog zur Auswertung der MTA-Studie nach Connors et al. (2001) soll daher ein zusammengesetztes Gesamtmaß bestimmt werden. Zunächst wird eine explorative Faktorenanalyse auf Grundlage der Prämessung der Erfolgskriterien durchgeführt.

Einige Voraussetzungen sollten zur Durchführung einer Faktorenanalyse erfüllt sein. Backhaus, Erichson, Plinke und Weiber (2011) empfehlen mindestens Intervallskalierung und eine Fallzahl in mindestens dreifacher Höhe der Anzahl an Variablen bzw. größer als 50. Weiterhin wichtig ist, dass die Variablen gute Reliabilitäten aufweisen. Für Skalenwerte und Testkennwerte aus Leistungstests reliabler Messinstrumente wie in der vorliegenden Auswertung kann dies als gegeben angenommen werden (Bühner, 2011). Die Überprüfung der Dateneignung wird mittels Kaiser-Meyer-Olkin-Koeffizienten (KMO), Measure of Sample Adequacy-Koeffizienten (MSA) und dem Bartlett-Test überprüft. Der Bartlett-Test prüft, ob die Korrelationen der Korrelationsmatrix signifikant von null verschieden sind, da sonst keine Faktoren gefunden werden könnten.

Es kommen mehrere Methoden zur Extraktion der Faktoren infrage. In der MTA-Auswertung wurde eine Hauptkomponentenanalyse durchgeführt, die von Autoren wie Bühner (2011) nicht als eigentliches faktorenanalytisches Vorgehen eingeordnet wird, sondern als Verfahren der Datenreduktion zu wenigen Komponenten bzw. Zusammenfassungen von Einzelitems. Die Hauptkomponentenanalyse wird häufig durchgeführt und erscheint für das Ziel dieser Analyse gut geeignet zu sein. Alternativ kommen die Hauptachsenanalyse und die Maximum-Likelihood-Analyse infrage. Es wird empfohlen die Ergebnisse unterschiedlicher Extraktionsmethoden miteinander zu vergleichen (Bühner, 2011), daher werden in der weiteren Auswertung die Ergebnisse der Hauptkomponentenanalyse ausführlich dargestellt und mit den Ergebnissen der Hauptachsen- und Maximum-Likelihood-Methode verglichen.

Zur Überprüfung der angemessenen Anzahl an Faktoren wird sehr häufig das Kaiser-Kriterium herangezogen, welches grafisch durch Inspektion des Screeplots abgesichert wird. Beim Kaiser-Kriterium werden nur Faktoren mit Eigenwerten > 1 ausgewählt. Beim Scree-Test werden die Eigenwerte in absteigender Reihenfolge abgebildet und dann überprüft, an welcher Stelle ein „Knick“ vorliegt, d.h. ab wann sich die Eigenwerte asymptotisch der Abszissen-Achse annähern (Backhaus, Erichson, Plinke et al., 2011).

Neben der Frage nach einer angemessenen Extraktionsmethode stellt sich die Frage nach einer passenden Rotationsmethode. Grundsätzlich ist eine orthogonale Rotation

dann sinnvoll, wenn die zu Grunde liegenden Faktoren als nicht miteinander korreliert angenommen werden oder wenn keine theoretischen Annahmen über die Dimensionalität der eingehenden Variablen vorliegen. Ist hingegen eine Korrelation der Faktoren untereinander anzunehmen, ist eine oblique Rotation vorzuziehen (Moosbrugger & Schermelleh-Engel, 2012). Bezogen auf das Ziel der vorliegenden Analyse ist modellhaft nicht klar, ob die zu findenden Erfolgsfaktoren besser als korreliert oder nicht korreliert anzunehmen sind. Prinzipiell wird davon ausgegangen, dass es ein Gesamterfolgsmaß gibt, so dass einzelne Erfolgsfaktoren Korrelationen aufweisen müssten. Denkbar ist aber auch, dass mehrheitlich nur in bestimmten Bereichen Verbesserungen erzielt wurden, in anderen aber nicht. Bühner (2011) schlägt bei einer solchen Unklarheit vor, zuerst eine oblique Rotation durchzuführen, am besten mit der Promax-Methode, und bei geringer Korrelation der Faktoren untereinander orthogonal zu rotieren. Zur orthogonalen Rotation wird die Varimax-Methode empfohlen (Bühner, 2011). Analog zu Connors (2001) werden kleine Ladungen (< 0.40) unterdrückt und ein paarweiser Fallausschluss vorgenommen.

Zentral bei der Interpretation der Ergebnisse ist die inhaltliche Validität der ermittelten Faktorenlösungen. Es wird empfohlen auch Items mit geringen oder uneindeutigen Ladungen zu behalten, wenn sie inhaltlich wichtig erscheinen (Bühner, 2011). Für die vorliegende Auswertung werden daher keine Items eliminiert, da es sich nicht um Einzeli-tems zur Zusammenstellung eines Testverfahrens handelt, sondern um ausgewählte (Skalen-)Werte, die zur Evaluation der Maßnahme zusammengestellt wurden. Die Items werden dem Faktor zugeordnet, auf den sie am höchsten laden.

Um mit den Faktorwerten weiter rechnen zu können und sie z. B. auf Zusammenhänge mit möglichen Moderatorvariablen zu untersuchen, sollen Bartlett-Schätzer bestimmt werden. Diese Schätzer ordnen jedem Teilnehmer einen Wert zu, der aus den auf den jeweiligen Faktor ladenden Variablen gebildet wird. Es werden also neue Variablen generiert durch Zusammenfassung und Gewichtung der in die Faktorenanalyse eingehenden Variablen (Bühl, 2010). Weiterhin sollen manuell Faktorwerte wie bei Connors et al. (2001) bestimmt werden. Die Zusammenhänge der Bartlett-Schätzer und den manuell bestimmten Werten werden im Anschluss überprüft. Zeigen sich sehr hohe Übereinstimmungen, wird davon ausgegangen, dass die Maße dasselbe erfassen. Dann wird nur mit der manuellen Variante weitergerechnet, da sie am einfachsten und am besten nachvollziehbar ist.

Die ermittelten Faktoren können als abhängige Variablen zur Überprüfung allgemeiner und differenzieller Trainingseffekte genutzt werden. Sie stellen globale Erfolgsmaße dar, die eine Abschätzung der Nettoeffekte zulassen, so wie es auch in medizinischen klinischen Studien umgesetzt wird (Cleophas, Zwinderman, Cleophas & Cleophas, 2009). Zu diesem Zweck wird für jeden Teilnehmer der Mittelwert der Erfolgsfaktoren bestimmt. Bei fehlenden Werten für einen Faktor wird das Gesamterfolgsmaß aus den übrigen vorhandenen Faktoren bestimmt. Die Bestimmung des Gesamterfolgsmaßes mit Ausgleich für fehlende Werte hat den Vorteil, dass für jeden Fall ein Erfolgsmaß be-

stimmt werden kann, so dass alle Fälle miteinander verglichen werden können. Das Gesamtmaß wird im Folgenden als „Gesamtproblematik“ bezeichnet.

5.1.4 Veränderungsmessung

Damit aus den Erfolgsmaßen auf eine Veränderung geschlossen werden kann, werden die Werte des Prä- und des Postmesszeitpunktes auf unterschiedliche Weise miteinander verrechnet. Beispielsweise wird ein Differenzwert der Gesamtproblematik ermittelt, indem vom Wert des Postmesszeitpunktes der Wert des Prämesszeitpunktes subtrahiert wird. Aus der Differenz wird auf eine zugrunde liegende Veränderung geschlossen. Es handelt sich also um eine indirekte Erfolgsmessung (Gollwitzer & Jäger, 2009).

Eine indirekte Erfolgsmessung ist für die vorliegende Untersuchung einer direkten Erfolgsmessung, z. B. über Befragung der Jugendlichen direkt nach Trainingsabschluss, aus einigen Gründen vorzuziehen: Das Problem der Verfälschung der Ergebnisse bei einer direkten retrospektiven Veränderungseinschätzung stellt sich bei Jugendlichen mit ADHS im Besonderen, da sie dazu neigen, ihre Probleme zu unterschätzen und ihre eigenen Fähigkeiten zu überschätzen (Kaiser & Hoza, 2008). So könnte eine reine Befragung der Jugendlichen nach dem Training ergeben, dass gar keine Probleme vorliegen, unabhängig davon, welche Veränderungen durch das Training stattgefunden haben. Weiterhin zeigen sich bei vielen Jugendlichen mit ADHS Beeinträchtigungen von Exekutiv- und Gedächtnisfunktionen (Willcutt et al., 2005). Somit ist unklar, ob sie sich noch valide an den Zustand vor dem Training erinnern und die Bewertung vornehmen können. Schließlich fußt das Training auf einer intensiven, positiv getönten Beziehung der Trainer zu den Jugendlichen, die wahrscheinlich macht, dass eine sehr positive Einschätzung des Trainings in Folge sozialer Erwünschtheit vorgenommen wird. Die soziale Erwünschtheit mag sich – wenn auch weniger stark – ebenso in der Einschätzung der Eltern und Lehrer niederschlagen und kann nur schwerlich kontrolliert werden. Im Folgenden wird daher die indirekte Veränderungsmessung vorgezogen, da diese den oben beschriebenen Problemen nicht bzw. weniger stark unterliegt.

Bei der indirekten Erfolgsmessung bestehen einige messtheoretische Probleme, z. B. der Effekt der Regression zur Mitte. Trotzdem wird die indirekte Erfolgsmessung in der Wirksamkeitsforschung vorrangig verwendet und als nützlich und sinnvoll beurteilt (Lutz & Grawe, 2007). Wie bereits dargestellt werden einige Kritikpunkte, die in der Vergangenheit stark diskutiert wurden, im Licht neuerer Forschung als weniger problematisch eingeschätzt.

Die theoretische Diskussion wird an dieser Stelle nicht erneut vertieft, sondern zusätzlich ein empirisches Argument ins Feld geführt: Aus der allgemeinen Wirksamkeitsstudie liegen Ergebnisse für eine Wartekontrollgruppe vor (Schramm, im Druck). Diese Gruppe besteht aus $n = 34$ Teilnehmern. Wie aus dem Evaluationsdesign der allgemeinen Wirksamkeitsprüfung ersichtlich wird, wurden mit dieser Gruppe zwei Messzeitpunkte umgesetzt, zwischen denen keine Intervention stattfand. Somit können mit den

Ergebnissen dieser Gruppe Aussagen zu Effekten der reinen Messwiederholung getroffen werden. Insgesamt zeigten sich keine bzw. sehr geringe Veränderungen (Schramm, im Druck). Da nach Rogosa (1988) ein Regressionseffekt zur Mitte nur dann auftritt, wenn die Korrelation zwischen initialem Status und Veränderungen negativ ist, kann ein solcher Effekt hier ausgeschlossen werden, da gar keine Veränderung aufgetreten ist. Darüber hinaus müsste sich bei vorliegendem Regressionseffekt eine Varianzreduktion abzeichnen, da sich extreme Messwerte bei wiederholter Messung stärker dem Mittelwert angleichen. Dies ist in den Daten nicht der Fall und es liegen zu beiden Messzeitpunkten in etwa gleich große Varianzen vor. Somit ist aus den Daten der Wartekontrollgruppe abzuleiten, dass es keine Hinweise darauf gibt, dass ein Regressionseffekt zur Mitte aufgetreten ist.

Aber selbst dann, wenn auch in der Wartegruppe Verbesserungen auftreten würden, können weiterführende Analysen potentielle Effekte der Intervention aufzeigen. Es müsste dazu die inkrementelle Wirksamkeit überprüft werden, also der zusätzliche Effekt der Intervention über den reinen Effekt der Messwiederholung hinaus. Dazu werden die Effektstärken von Wartekontrollgruppe und Interventionsgruppe verglichen. Ist der Effekt in der Interventionsgruppe größer als in der Wartekontrollgruppe, so ist die Differenz der Effektstärken der eigentliche Gewinn durch die Intervention.

Wie bereits beschrieben, bieten sich die folgenden drei einfachen Verrechnungsmöglichkeiten an: Differenzwerte, „residual gain scores“ und prozentuale Veränderungen. Nach diesen drei Auswertungsmöglichkeiten werden jeweils Erfolgsmaße bestimmt. Wenn sich herausstellt, dass ein Maß nicht geeignet ist, wird es von der weiteren Auswertung ausgeschlossen. Wenn sich herausstellt, dass mit allen Maßen dieselbe Information erfasst wird, wird nur ein geeignetes Maß weiterverwendet.

5.1.5 Vorläufige Ergebnisse

Die Ergebnisse zur allgemeinen Wirksamkeitsprüfung sind derzeit im Druck (Schramm, im Druck). Daher können an dieser Stelle keine detaillierten Aussagen dazu gemacht werden. Zusammenfassend kann bereits gesagt werden, dass mit dem Lerntraining für Jugendliche mit ADHS im Vergleich zu einer Wartekontrollbedingung signifikante Verbesserungen erzielt wurden. Im Vergleich zur alternativen Behandlung mit der progressiven Muskelentspannung wurde allerdings keine signifikante Überlegenheit des LeJA-Trainings festgestellt. Es zeigen sich leicht bessere Werte für das LeJA-Training im Bereich kleiner Effekte verglichen mit dem Entspannungstraining, es wird aber keine Signifikanz erreicht.

5.2 Differentielle Wirksamkeit

Im Anschluss an die allgemeine Wirksamkeitsprüfung im Kontrollgruppendesign stellt sich die Frage nach der differentiellen Wirksamkeit des Lerntrainings LeJA. Damit ist an dieser Stelle die Frage danach gemeint, ob die einzelnen Teilnehmer in unterschied-

licher Weise vom Training profitiert haben, um somit eine Aussage über zusammenfassende Gruppenmittelwertvergleiche hinaus treffen zu können. Die Frage nach der differentiellen Wirksamkeit wird in dieser Untersuchung anhand einer vergrößerten Stichprobe als in der Kontrollgruppenstudie beantwortet werden. Die Daten werden post hoc weiterführend ausgewertet. Dieses Vorgehen verfolgt das Ziel, zusätzliche Erkenntnisse aus den Daten zu gewinnen und für die weitere Optimierung zu nutzen. Kazdin (2010, S. 477) fasst den Nutzen solcher Analysen treffend zusammen: "Post hoc analyses provide useful leads about predicting (i. e., for the next study) who is likely to respond to treatment, to identify more intense or stronger treatments for such individuals, and to test whether one can improve responsiveness of these individuals."

Zur Überprüfung der differentiellen Wirksamkeit sind auf konzeptueller Ebene zwei Vorgehensweisen denkbar und werden auch umgesetzt. Durch eine Herangehensweise, die stärker deduktiven Charakter hat, werden Hypothesen geleitet Auswirkungen möglicher Moderatorvariablen auf den Trainingserfolg überprüft. Dies sind sowohl dichotome Merkmale wie Geschlecht oder Medikation als auch metrische Variablen wie das Alter der Teilnehmer. Durch eine stärker induktive Herangehensweise werden zunächst solche Teilnehmer der Intervention identifiziert, die einen besonders guten Trainingserfolg vorweisen können, um diese Gruppe genauer zu beschreiben. Ergänzend wird eine Gruppe mit besonders schlechtem Trainingserfolg identifiziert („Non-Responder“), um auch diese Gruppe genauer zu beschreiben und mögliche Risikofaktoren für schlechte Trainingsverläufe zu finden. Beide Vorgehensweisen nähern sich damit demselben Ziel, jedoch von der jeweils anderen Richtung her. Der eher deduktive Ansatz schaut von Moderatorvariablen zum Erfolg, den diese erklären können, der eher induktive Ansatz sucht ausgehend von den Erfolgsvariablen erklärende Moderatoren (siehe Abbildung 20 zur Veranschaulichung). Die Begriffe „deduktiv“ und „induktiv“ sind in der Abbildung jeweils in Anführungszeichen gesetzt, da beide Herangehensweisen sowohl induktive als auch deduktive Elemente aufweisen (Sedlmeier & Renkewitz, 2008). Es handelt sich also um eine Akzentuierung der jeweiligen Vorgehensweise.

Die Ergebnisse der Analysen sollen dazu genutzt werden, optimale Trainingsbedingungen für Subgruppen innerhalb der Zielgruppe zu identifizieren. So soll eine Grundlage für die Anpassung und Optimierung der Trainingskonzeption für solche Teilnehmer oder Konstellationen gefunden werden, in denen das Training bislang wenig erfolgreich gewirkt hat. Die Stoßrichtung der Untersuchung orientiert sich damit an der „Verschreibungsphase“ in der Psychotherapieforschung (Grawe, 1997), die fragt, für welchen Patienten welche Therapieform in welchem Setting am erfolgversprechendsten ist. Diese Art der Forschung wird auch als Indikationsforschung bezeichnet (Lutz & Grawe, 2007) und ist der Treatment-Aptitude-Forschung (Snow, 1991) sehr ähnlich, die bereits beschrieben wurde.

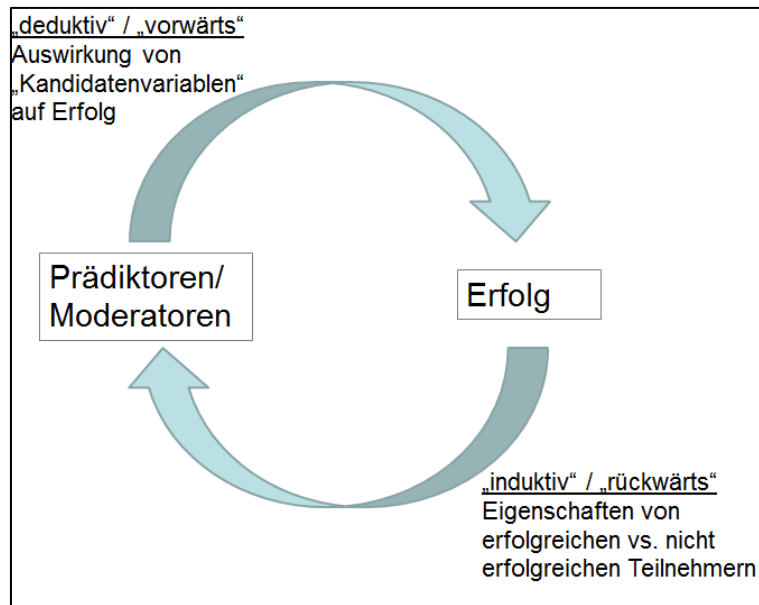


Abbildung 20: Veranschaulichung der stärker deduktiven und der stärker induktiven Vorgehensweise zur Beantwortung der Fragestellung nach der differentiellen Wirksamkeit des Lerntrainings

5.2.1 Fragestellung

Die allgemeine Fragestellung bezüglich der differentiellen Wirksamkeit des Trainings kann unter Berücksichtigung der oben genannten Aspekte folgendermaßen zusammengefasst werden:

Wie hängen Merkmale der Teilnehmer und der Trainingsdurchführung mit dem Trainingserfolg zusammen?

Hierbei werden in der empirischen Prüfung folgende Merkmale untersucht:

- (1) Geschlecht
- (2) Medikation
- (3) ADHS-Subtypus
- (4) Alter
- (5) Initiale Störungsschwere
- (6) Komorbidität
- (7) Trainingsdauer

Es werden genau diese Merkmale untersucht, da es vor dem Hintergrund theoretischer Überlegungen und empirischer Erkenntnisse plausibel ist, dass sie einen moderierenden Effekt auf den Trainingserfolg ausüben.

Die allgemeine Fragestellung wird weiter spezifiziert, damit sie mit geeigneten Methoden beantwortet werden kann. Folgende Ausdifferenzierung wird vorgenommen, die im Anschluss erläutert wird:

Fragestellung 1 a: *Wie hängen die Merkmale der Teilnehmer und der Trainingsdurchführung jeweils einzeln mit den verschiedenen Aspekten des Trainingserfolges zusammen?*

Fragestellung 1 b: *Wie hängen die Merkmale der Teilnehmer und der Trainingsdurchführung – wenn sie gleichzeitig berücksichtigt werden – mit den verschiedenen Aspekten des Trainingserfolges zusammen?*

Fragestellung 1 c: *Wie hängen die Merkmale der Teilnehmer und der Trainingsdurchführung – wenn sie gleichzeitig berücksichtigt werden – mit den zusammengefassten Aspekten des Trainingserfolges zusammen?*

Fragestellung 2 a: *Lassen sich Gruppen von Teilnehmern mit besonders starken Verbesserungen und weniger starken bzw. ohne Verbesserungen durch das Training identifizieren und wenn ja, welche Unterschiede weisen diese Gruppen jeweils auf?*

Fragestellung 2 b: *Lassen sich im Vergleich der Teilnehmer untereinander – bei gleichzeitiger Berücksichtigung mehrerer Einzelaspekte des Gesamttrainingserfolgs – Gruppen mit ähnlichen Veränderungsmustern feststellen und wenn ja, welche Unterschiede weisen diese Gruppen jeweils auf?*

Die Spezifizierung der Fragestellung begründet sich wie folgt: Die Fragestellungen 1 a – c beziehen sich auf die deduktive bzw. vorwärts gerichtete Analyse. Diese grenzen sich von den Fragestellungen 2 a und b ab, die der induktiven bzw. rückwärtsgerichteten Vorgehensweise zuzuordnen sind. In der Beantwortung von Fragestellung 1 a sollen die einfachen Zusammenhänge der Merkmale und der Aspekte des Erfolgs jeweils einzeln nacheinander überprüft werden. In Fragestellung 1 b sollen die Merkmale gleichzeitig auf ihre Zusammenhänge mit den unterschiedlichen Aspekten des Trainingserfolges analysiert werden sollen. Damit können die Zusammenhänge der Merkmale untereinander berücksichtigt werden. Es wäre beispielsweise denkbar, dass in der Einzelauswertung mehrere Zusammenhänge gefunden werden, die aber redundante Informationen enthalten (Sedlmeier & Renkewitz, 2008). Dieses würde in der gleichzeitigen Analyse sichtbar. In Fragestellung 1 c werden nicht nur die Einzelmerkmale der Teilnehmer und der Durchführung gleichzeitig berücksichtigt, sondern auch alle Aspekte des Trainingserfolges gleichzeitig. Zur Beantwortung dieser Fragestellung ist ein multivariates Vorgehen angeraten. Damit können komplexe Zusammenhänge aufgedeckt werden, die in Einzelauswertungen nicht sichtbar sind (Tabachnick & Fidell, 2007). Die Fragestellungen 2 a und b implizieren eine Einordnung der Teilnehmer der Intervention nach ihrem individuellen Trainingserfolg. Nach dieser Einordnung werden die gefundenen Gruppen auf ihre Unterschiede hin untersucht. Dazu werden Merkmale der Teilnehmer und der Durchführung und andere Maße genutzt, die vor Beginn des Trainings vorliegen. Bei

Fragestellung 2 a soll die Gruppenbildung nach einem statistischen Kriterium vorgenommen werden. Dazu wird der „Reliable Change Index“ (RCI; Jacobson & Truax, 1991) als Maß für eine reliable Veränderung zur Gruppenzuordnung genutzt. Zur Beantwortung von Fragestellung 2 b soll die Gruppenzuordnung nach der Ähnlichkeit der Teilnehmer untereinander bezogen auf mehrere Einzelaspekte des Trainingserfolges stattfinden (vgl. Lauth & Minsel, 2011).

5.2.2 Hypothesen

Basierend auf den zuvor dargestellten theoretischen Grundlagen und empirischen Ergebnissen zur differenziellen Wirksamkeit werden im Folgenden Hypothesen aufgestellt. Die Hypothesen beziehen sich ausschließlich auf Fragestellung 1 a, da die anderen Fragestellungen entweder zu komplex sind und sich daher keine begründeten Hypothesen ableiten lassen oder stärker explorativen Charakter haben, insbesondere bei den Fragen 2 a und b. Da die Hintergründe der Fragestellungen zu den unterschiedlichen Bereichen zuvor ausführlich erläutert wurden, werden die bereits dargestellten Quellen an dieser Stelle nicht wiederholt, um Redundanz zu vermeiden. Falls theoretische Überlegungen und empirische Ergebnisse keine eindeutige Ableitung von Hypothesen zulassen, werden im Weiteren nur deskriptive Auswertungen vorgenommen. Diese beschreibenden Erkenntnisse können als Grundlage für weiterführende Forschung genutzt werden im Sinne eines hypothesengenerierenden Vorgehens (Kraemer et al., 2002).

Konventionsgemäß befindet sich beim Aufstellen der Hypothesen der vermutete Effekt in der Alternativhypothese und in der Nullhypothese das Gegenteil (Bortz & Döring, 2006). Die Überprüfung bezieht sich auf die Nullhypothese. Wird diese abgelehnt, wird angenommen, dass die Alternativhypothese zutreffend ist. Das Signifikanzniveau wird wie konventionell üblich auf $\alpha = 5\%$ festgelegt.

Wenn statistische Hypothesen formuliert werden, so beziehen sich diese auf die Veränderung der Gesamtproblematik als Maß des Trainingserfolges. Das Gesamtproblemmaß wird aus den einzelnen erhobenen Erfolgsmaßen bestimmt. Um die Veränderung im Laufe des Trainings abzubilden, werden die Werte des Gesamtproblemmaßes vom Prä- und Postmesszeitpunkt miteinander verrechnet, so dass ein Veränderungswert entsteht, der eine neue Variable bildet. Als Verrechnungsmöglichkeiten kommen, wie beschrieben, einfache Differenzwerte, prozentuale Veränderungen oder Residual Gain Scores in Betracht. Der Veränderungswert wird so gebildet, dass negative Werte eine Reduzierung der Gesamtproblematik anzeigen. Das heißt, negative Werte deuten auf ein erfolgreiches Training hin und dass Probleme verringert werden konnten. So wird beispielsweise der Differenzwert ermittelt, indem der Gesamtproblemwert zum Prämesszeitpunkt vom Wert des Postmesszeitpunkts subtrahiert wird. Bei einem erfolgreichen Training ist der Gesamtproblemwert zum Postmesszeitpunkt geringer als zum Prämesszeitpunkt, so dass sich ein negativer Differenzwert ergibt. Bei einem erfolgreichen Training ergibt sich also ein negativer Differenzwert.

Geschlecht

Die empirische Studienlage zur Abschätzung des Zusammenhangs vom Geschlecht und dem Interventionserfolg bei ADHS ist unklar. Nussbaum (2012) stellt in ihrem Überblick resümierend fest, dass zwar viele Studien zur Effektivität von Interventionen bei ADHS vorliegen, aber nur wenige die Geschlechterunterschiede in den Interventionen berücksichtigen. Das Störungsbild ist im Jugendalter bei den Geschlechtern eher ähnlich als unterschiedlich, vor allem beim Lernverhalten. Auch aufgrund der individualisierten Trainingsgestaltung ist davon auszugehen, dass Geschlecht kein Moderator des Trainingserfolges ist. Trotzdem stellt das Geschlecht einen zentralen unterscheidenden Aspekt dar. Daher werden die Daten der Mädchen und Jungen vergleichend gegenübergestellt.

Medikation

Wie bereits dargestellt ist eine Kombinationsbehandlung für Jugendliche mit ADHS vielversprechend. Die Ausgangslage der vorliegenden Studie ist allerdings eine besondere. Es werden zwar Jugendliche mit rein psychosozialer Intervention mit Jugendlichen mit einer Kombinationsbehandlung verglichen, es handelt sich bei letzteren aber um Jugendliche, die schon seit längerem stabil medikamentös eingestellt sind. Es wird also keine Kombinationsbehandlung neu begonnen, sondern der medikamentösen Behandlung wird eine psychosoziale Intervention additiv hinzugefügt. Somit werden die Effekte des kognitiv-behavioralen Lerntrainings mit dem additiven Effekt dieses Trainings zu einer medikamentösen Behandlung verglichen.

Unterschiedliche Annahmen erscheinen vor diesem Hintergrund plausibel. Es scheint so zu sein, dass die Jugendlichen mit stabiler Medikation unter residualen Symptomen und Problemen leiden, die durch die Medikamente in der Vergangenheit nicht komplett behoben werden konnten. Dies mag daran liegen, dass es sich um eine besonders gravierende und schwer zu verändernde Problematik handelt. Wenn diese Annahme der besonders stabilen Symptomatik bei Teilnehmern mit Medikation zutrifft, so müssten die Trainingseffekte für diese Teilnehmer geringer ausfallen. In dieser Argumentationslinie könnte weiterhin vermutet werden, dass der Spielraum für Verbesserungen durch die Effekte der Medikation bereits ausgeschöpft ist. Es könnten dann nur noch kleinere zusätzliche Gewinne erreicht werden, aber der Großteil des Gesamtverbesserungspotentials wäre bereits erschöpft. Das Ausmaß der Problematik sollte anhand der Werte des Messzeitpunkts vor Beginn des Lerntrainings überprüft werden. Teilnehmer mit Medikation hätten eine weniger ausgeprägte Problematik, wenn diese Annahme zuträfe.

Andererseits scheinen sich die Jugendlichen mit stabiler Medikation bzw. ihre Eltern bewusst für die Teilnahme an dem Lerntraining entschieden zu haben. Da sie bereits behandelt wurden, haben sie vermutlich realistische Vorstellungen, welches Ausmaß an Veränderung realistisch ist, und die Entscheidung getroffen, dass gerade diese Förderung für sie erfolgversprechend ist. Jugendliche ohne Medikation und ihre Eltern, die zum ersten Mal eine Behandlung in Anspruch nehmen, haben vermutlich weniger rea-

listische Vorstellungen, welche Verbesserungen durch das Lerntraining erreicht werden können, und überschätzten möglicherweise den zu erwartenden Trainingserfolg. Verglichen mit ihren überhöhten Erwartungen der Effekte fällt die Einschätzung des Realerfolgs dann geringer aus. Die Erfolge wären demnach höher für Teilnehmer mit Medikation.

Auf die gleiche Annahme läuft folgende Überlegung hinaus. In der Behandlung von Kindern mit ADHS wird eine primäre medikamentöse Behandlung oft als notwendige Grundlage verstanden, damit überhaupt eine psychosoziale Behandlung stattfinden kann (Deutsche Gesellschaft für Kinder- und Jugendpsychiatrie und Psychotherapie u. a., 2007). Für Schulkinder und Jugendliche wird eine primäre medikamentöse Behandlung empfohlen, wenn eine so stark ausgeprägte Symptomatik - insbesondere der Hyperaktivität – vorliegt, dass situationsübergreifend erhebliche Beeinträchtigungen bestehen oder eine krisenhafte Zuspitzung zu befürchten ist (Arbeitsgemeinschaft ADHS der Kinder- und Jugendärzte, 2007). Einige Kinder, die von besonders starker Symptomatik betroffen sind, können die Anforderungen des Lerntrainings, das ein überdachtes, langsames und konzentriertes Arbeitsverhalten anstrebt, ohne eine Dämpfung ihrer Kernsymptomatik durch ein Medikament nicht umsetzen. Im Jugendalter ist aufgrund der sich abschwächenden Hyperaktivität und Impulsivität anzunehmen, dass dies weniger problematisch ist als im Kindesalter. Trotzdem kann nicht ausgeschlossen werden, dass sich in der Gruppe der Teilnehmer ohne Medikation solche Jugendlichen befinden, die das Training aufgrund stark ausgeprägter Kernproblematik nicht optimal nutzen können und bei denen eine gleichzeitige pharmakologische Behandlung angeraten wäre. Auch im adaptiven Behandlungsplan der Studie von Döpfner et al. (2004) reichte die reine Verhaltenstherapie bei 26 % der Kinder nicht aus und wurde schließlich durch zusätzliche Medikation ergänzt.

Zusammenfassend können sowohl Argumente für eine höhere Wirksamkeit des Lerntrainings bei gleichzeitiger Medikation gefunden werden als auch für eine geringere Wirksamkeit. Darüber hinaus könnten auch gegenläufige Effekte auftreten, die sich neutralisieren, so dass in der Summe gar kein Unterschied gefunden wird. Ob in der vorliegenden Untersuchung die Verbesserung der Residualsymptomatik der Jugendlichen mit Medikation den Trainingserfolg der Teilnehmer ohne Medikation übersteigt oder andersherum, ist daher kaum einzuschätzen. Insgesamt ist eher davon auszugehen, dass für beide Gruppen mit ähnlich großen Erfolgen zu rechnen ist. Es wird daher nicht Hypothesen prüfend vorgegangen, sondern ein deskriptiver Vergleich durchgeführt. Plausibel ist anzunehmen, dass die ADHS-Symptomatik der Jugendlichen mit Medikation zum Prämesszeitpunkt geringer ausgeprägt ist, da sie schon eine Reduktion der Symptome durch die Medikation erlebt haben.

ADHS-Subtypus

Aus theoretischen Überlegungen ist nicht eindeutig abzuleiten, ob Jugendliche des vorwiegend unaufmerksamen Typus größere oder weniger große Trainingserfolge erzielen

als Jugendliche des vorrangig hyperaktiv-impulsiven Subtypus. Empirische Untersuchungen zur differenziellen Trainingswirksamkeit in Bezug zum ADHS-Subtypus deuten darauf hin, dass nur sehr geringe bzw. keine Unterschiede zu erwarten sind. Deshalb wird an dieser Stelle keine Hypothese aufgestellt, sondern deskriptiv verglichen.

Alter

Zusammenfassen lässt sich, dass die kognitiv-behaviorale Behandlung von ADHS und verwandten Störungen vermutlich bei älteren Kindern bzw. Jugendlichen wirksamer ist als bei jüngeren. Die kognitiven Aspekte und Techniken der Therapie erfordern ein gewisses Maß der kognitiven Reife, das komplexe Prozesse wie Selbstreflexion und Metakognition erst möglich macht. Noch bis ins späte Jugendalter hinein zeigen sich starke Verbesserungen kognitiver Fähigkeiten wie Entscheidungs- und Problemlösekompetenzen. Biologische Grundlage ist u. a. die voranschreitende Gehirnreifung. Bei Jugendlichen mit ADHS treten einige Entwicklungsprozesse verzögert ein, so dass sich einige Kompetenzzuwächse, die ihre Gleichaltrigen schon im Kindesalter durchlaufen haben mögen, auf die Adoleszenz verschieben. Ältere Jugendliche sind zudem stärker von dem Druck der zu bewältigenden Entwicklungsaufgabe des Übergangs in das Berufsleben betroffen. Daher ist zu erwarten, dass sie ein Lerntraining, das ihnen genau bei der Bewältigung dieser Herausforderung helfen soll, stärker annehmen, dessen Nutzen höher bewerten und daher motivierter und engagierter sind. Weiterhin unterstützen viele – aber nicht alle – empirische Studien den vermuten Zusammenhang zwischen Alter und Trainingserfolg. An dieser Stelle ist zu wiederholen, dass chronologische Alter an sich nicht als ein kausaler Faktor zu betrachten ist, sondern als ein Indikator für die kognitive Reife bzw. das kognitive Entwicklungsniveau herangezogen wird.

Wie genau dieser Zusammenhang aussieht, ist bislang unklar. Die in den Studien von Dush, Hirt und Schroeder (1989) und Durlak, Fuhrman und Lampman (1991) gebildeten Gruppen sind eventuell zu groß, um genauere Zusammenhänge abzubilden. Zudem wurden relativ wenige Jugendliche in der mittleren Adoleszenz, etwa ab 14 Jahren, in die Untersuchungen eingeschlossen. Auch wenn es fraglich ist, ob der Zusammenhang zwischen Alter und Trainingserfolg linear ist, wird die Hypothese eines linearen korrelativen Zusammenhangs überprüft. Dies erscheint sinnvoll, da sich auch kein quadratischer oder andersartiger Zusammenhang stimmig aus Theorie und Empirie ableiten lassen. So wird zunächst nach einer linearen Korrelation überprüft, so wie es der Konvention entspricht. Bei Nichtzutreffen wird weiterführend deskriptiv ausgewertet.

Folgende inhaltliche Hypothese wird überprüft: Ältere Teilnehmer profitieren mehr vom Training und haben eine größere Reduktion ihrer Problematik. Das bedeutet, dass der Veränderungswert des Gesamtproblemwertes, der positive und negative Ausprägungen annehmen kann, bei ihnen kleiner bzw. stärker negativ ist. Negative Werte zeigen eine Problemreduktion und damit einen Trainingserfolg an. Die Null- und Alternativhypothese lauten dementsprechend:

$$H_1: \rho_{\text{Alter, VeränderungswertGesamtproblematik}} < 0$$

$$H_0: \rho_{\text{Alter, VeränderungswertGesamtproblematik}} \geq 0.$$

Initiale Störungsschwere

In den meisten Studien wird ein Zusammenhang zwischen initialer Störungsschwere und Interventionserfolg festgestellt. Wie dieser Zusammenhang gerichtet ist, ist nicht klar. Daher wird die ungerichtete Hypothese aufgestellt, dass ein Zusammenhang vorhanden ist, um dann weitere explorative Auswertungsschritte anzuschließen. Die ungerichtete Hypothese lautet dementsprechend:

$$H_0: \rho_{\text{InitialeStörungsschwere, VeränderungswertGesamtproblematik}} = 0$$

$$H_1: \rho_{\text{InitialeStörungsschwere, VeränderungswertGesamtproblematik}} \neq 0$$

Komorbidität

Trotz uneinheitlicher empirischer Ergebnisse wird im Sinne der „complexity hypothesis“ an der Annahme festgehalten, dass größere Komorbiditätsraten schlechtere Interventionserfolge erwarten lassen. Dies erscheint plausibel, da für unterschiedliche Problemlagen vermutlich mehrere Interventionen zum Einsatz kommen müssen, damit sich in der Gesamtproblematik eine deutliche Verbesserung einstellt. Das Lerntraining zielt insbesondere auf eine Verbesserung des Lern- und Leistungsverhaltens. Wenn daneben viele weitere Probleme vorliegen, können diese im Rahmen des LeJA-Trainings nur in geringem Umfang berücksichtigt werden.

Um die Ergebnisse der von der Altersgruppe und Intervention sehr ähnlichen Studie von Antshel et al. (2012) zu berücksichtigen, wird eine Differenzierung nach externalisierenden und internalisierenden Komorbiditäten berücksichtigt. Die „complexity hypothesis“ gilt demnach nur für externalisierende Komorbiditäten. Es wird daher angenommen, dass ein höheres Ausmaß an externalisierender Komorbidität mit geringerem Trainingserfolg bzw. mit stärker positiven Veränderungswerten in Verbindung steht und ein höheres Ausmaß an internalisierender Komorbidität mit besseren Trainingserfolgen bzw. mit stärker negativen Veränderungswerten. Die Hypothesen lauten:

$$H_1: \rho_{\text{KomorbiditätExt, VeränderungswertGesamtproblematik}} > 0$$

$$H_0: \rho_{\text{KomorbiditätExt, VeränderungswertGesamtproblematik}} \leq 0$$

$$H_1: \rho_{\text{KomorbiditätInt, VeränderungswertGesamtproblematik}} \leq 0$$

$$H_0: \rho_{\text{KomorbiditätInt, VeränderungswertGesamtproblematik}} > 0$$

Trainingsdauer

Es stellt sich die Frage, ob längere oder kürzere Trainings effektiver sind oder kein Zusammenhang zwischen Dauer und Trainingserfolg anzunehmen ist. In dieser Frage ist zwischen der Dauer bzw. zeitlichen Erstreckung des Trainings in Monaten und der Anzahl der durchgeführten Sitzungen zu unterscheiden. Diese Aspekte hängen zwar insofern zusammen, als anzunehmen ist, dass Trainings mit mehr Sitzungen vermutlich auch länger sind, aber es dennoch Unterschiede gibt, insbesondere aufgrund der zeitlichen Organisation des Ablaufes. So wurden in einigen Fällen zwei Sitzungen pro Woche durchgeführt, so dass viele Sitzungen in einer kurzen Zeitspanne durchgeführt wurden, aber es gab auch Verläufe mit größeren Pausen, z. B. wegen Krankheit, so dass eine lange Trainingsdauer vorlag, ohne dass mehr Sitzungen durchgeführt wurden.

Unter Abwägung der konzeptuellen Überlegungen und empirischen Ergebnisse wird insgesamt davon ausgegangen, dass eine kürzere Trainingsdauer und weniger Sitzungen mit einem größeren Trainingserfolg zusammenhängen, da in einer kürzeren Dauer die Inhalte aneinanderhängend und gut aufeinander aufbauend vermittelt werden können. Eine kurze Trainingsdauer wird als Indikator für eine gut gelingende Trainingsorganisation und plangemäße Durchführung interpretiert, die dem Trainingserfolg zuträglich ist. Dauer und Sitzungsanzahl werden demnach nicht als ursächliche Faktoren verstanden, sondern haben eine Hinweisfunktion auf unterschiedliche dahinterliegende Aspekte.

Bezugnehmend auf Metaanalysen, die keinen Zusammenhang zwischen Dauer und Erfolg der Intervention bei ADHS feststellen (z. B. van der Oord et al., 2008), ist Folgendes zu sagen: In metaanalytischen Untersuchungen wird die Dauer des Trainings zumeist als ein Moderator konzeptualisiert. Das bedeutet, dass die Dauer des Trainings bzw. Anzahl der Sitzungen bei Beginn der Maßnahme bereits feststeht. Im metaanalytischen Vergleich werden die Angaben zur Dauer aus der Trainingskonzeption oder die durchschnittliche Länge innerhalb einer Gruppe mit den Angaben aus den Wirksamkeitsstudien anderer Interventionen verglichen. In der hier vorliegenden Auswertung wird aber die Dauer des Trainings innerhalb einer Gruppe verglichen. Es steht zwar im Voraus der Rahmen von 16 bis 20 Terminen fest, die möglichst schnell, einmal wöchentlich durchgeführt werden sollen, dies kann aber nicht immer planmäßig umgesetzt werden. Aufgrund von Krankheit, Ferien, Schwierigkeiten in der Organisation u. ä. ist damit zu rechnen, dass sich die Trainingsdauer stark verlängern kann. Manchmal müssen mehr Sitzungen als vorgesehen durchgeführt werden, wenn nach einem längeren Ausfall Inhalte wiederholt werden müssen.

In der hier vorliegenden Untersuchung wird die Dauer des Trainings daher nicht als eine Moderatorvariable, sondern als eine Mediatorvariable konzeptualisiert. Die Dauer des Trainings ergibt sich im Trainingsverlauf aufgrund unterschiedlicher Erwägungen. Wenn möglich entscheiden der Jugendliche, der Trainer und die Eltern im Konsens, welche Anzahl an Sitzungen und welcher Durchführungsmodus sinnvoll erscheinen.

Bezogen auf die zeitliche Erstreckung des Trainings spricht eine kürzere Dauer für eine besser gelingende Trainingsorganisation, die vermutlich zu einem größeren Trainingserfolg beiträgt. Trainingsinhalte können kontinuierlich aufeinander aufbauend behandelt werden. Bezogen auf die Anzahl der Trainingssitzungen sprechen wenige Sitzungen dafür, dass eine schnelle Verbesserung eingetreten ist und entschieden wurde, dass keine weiteren Termine notwendig sind. Da es sich um eine Completeranalyse handelt, werden etwaige Trainingsabbrüche von der Analyse ausgeschlossen. So wird sichergestellt, dass keine Verzerrung durch diese Fälle mit wenigen Sitzungen und geringem bzw. ohne Trainingserfolg vorliegt.

Bei der Überprüfung der Trainingsdauer sind also sowohl die Anzahl der Sitzungen als auch die zeitliche Erstreckung des Trainings von Interesse. Beide Variablen werden überprüft. Es wird angenommen, dass Trainings mit einer kürzeren Dauer und weniger Sitzungen erfolgreicher sind. Damit müssten bei kürzeren Trainings stärker negative Veränderungswerte auftreten. Die gerichteten Hypothesen lauten deshalb wie folgt:

$$H_1: \rho_{\text{Trainingsdauer, VeränderungswertGesamtproblematik}} > 0$$

$$H_0: \rho_{\text{Trainingsdauer, VeränderungswertGesamtproblematik}} \leq 0$$

$$H_1: \rho_{\text{AnzahlSitzungen, VeränderungswertGesamtproblematik}} > 0$$

$$H_0: \rho_{\text{AnzahlSitzungen, VeränderungswertGesamtproblematik}} \leq 0$$

5.2.3 Operationalisierung

Die zentrale Frage nach dem Zusammenhang zwischen Merkmalen der Jugendlichen und der Trainingsdurchführung mit dem Trainingserfolg kann auf unterschiedliche Weise untersucht werden. Grundsätzlich geht es auf einer operationalisierten Ebene um die Analyse der Zusammenhänge mehrerer Variablen. Welche Variable dabei unabhängig und welche abhängig ist, ist je nach Forschungskontext festzulegen (Tabachnick & Fidell, 2007). Eine Variable kann in einem Forschungszusammenhang die unabhängige Variable sein und in einem anderen die abhängige.

Für die vorliegende Untersuchung ist das zentrale Konstrukt, das auf seine Zusammenhänge mit anderen Konstrukten überprüft werden soll, der Trainingserfolg. Daher werden alle Variablen, die den Trainingserfolg erfassen, als abhängige Variablen definiert. Diese sind in einigen Auswertungen die Veränderungswerte (Differenzen, prozentuale Veränderungen oder residual gain scores) oder die Werte der Erfolgskriterien zum Postmesszeitpunkt. Die vorgestellten Merkmale der Teilnehmer und der Durchführung, die potentiell mit dem Trainingserfolg zusammenhängen, werden als unabhängige oder als Moderatorvariablen definiert.

Das Begriffspaar unabhängige und abhängige Variable wird im Paradigma der Regressionsanalyse alternativ auch folgendermaßen bezeichnet: Prädiktorvariable – Prognose-

variable, erklärende – erklärte Variable oder Regressor – Regressand (Backhaus, Erichson, Plinke et al., 2011). Tabachnick und Fidell (2007) schlagen ergänzend schlicht „input-output“ (S. 2) vor.

Neben einer Operationalisierung der abhängigen Variablen des Trainingserfolges ist auch bei einigen unabhängigen Variablen, also bei Merkmalen der Teilnehmer und der Durchführung, eine Operationalisierung vorzunehmen. Dies trifft auf die Merkmale ADHS-Subtypus und Komorbidität zu.

Die Zuordnung der Teilnehmer zu einem ADHS-Subtypus ist nicht unmittelbar gegeben und kann nur schwerlich getroffen werden. Gemäß DSM-IV-TR müssten die erfüllten Symptome der Hyperaktivität und Impulsivität zusammengezählt werden und bei weniger als sechs erfüllten Symptomen wäre ein unaufmerksamer Subtypus zutreffend und bei mehr erfüllten Symptomen der hyperaktiv-impulsive Subtypus. Aufgrund der rückläufigen Hyperaktivität und Impulsivität in der Adoleszenz scheint diese Grenze aber nicht angemessen zu sein. Verlaufsstudien zeigen, dass die Subgruppenzugehörigkeit aus der Kindheit über das Jugendalter oftmals nicht stabil erhalten bleibt (Larsson et al., 2011; Normand et al., 2012). Da Hyperaktivität und Impulsivität im Jugendalter abnehmen, ist eine Gruppenzuordnung auf Grundlage des momentanen Status schwierig. Die für das Kindesalter formulierten Symptomkriterien sind für das Jugendalter teilweise unpassend. Deshalb werden die für die Subgruppenzuordnung zu erfüllenden Symptomkriterien nicht immer erreicht. Im DSM-V werden daher die Anzahl der zu erfüllenden Kriterien für eine Diagnose und Zuordnung zu einem Störungssubtypus der ADHS bei Erwachsenen reduziert (Kooij et al., 2010). Zudem fällt es schwer, bei einer über längere Zeit stabilen Medikation, welche insbesondere die hyperaktiv-impulsiven Symptomatik reduziert, eine Einschätzung dieser Problematik vorzunehmen. Werden alternativ retrospektive Einschätzungen herangezogen, können Verzerrungen auftreten.

Die Teilnehmer könnten nach den Einschätzungen aus dem DISYPS-II einem Subtypus zugeordnet werden. Doch es ist schwer zu entscheiden, wie Cut-off-Werte festgelegt werden sollten, die angeben, ab wann ein Teilnehmer welchem Subtypus zugeordnet werden soll. Es wird daher aufgegeben, eine Gruppenzuordnung vorzunehmen. Stattdessen werden Hyperaktivität und Impulsivität als kontinuierliche Merkmale in die Auswertung einbezogen. Dieses Vorgehen ist auch besser mit der Annahme kontinuierlich verteilter psychischer Konditionen, wie z. B. im Modell von Rommelse et al. (2008) bezogen auf die Endophänotypen in Einklang zu bringen. Die Wirkung der Medikation betreffend kann argumentiert werden, dass die aktuelle Ausprägung der Symptomatik im Vordergrund steht, unabhängig davon, ob der jetzige Zustand mit oder ohne Medikation erreicht wurde.

Die Komorbidität wird über die Einschätzungen zum Prämesszeitpunkt in den jeweiligen Subskalen des SDQ vorgenommen. Es handelt sich also nicht um diagnostizierte weitere Störungen neben der ADHS, sondern um kontinuierlich erfasste Symptome, die jeweils einer Komorbidität zuzuordnen sind. Wie bereits dargestellt weist der SDQ auch

auf Subskalenebene eine gute psychometrische Qualität auf. So können die Subskalen einzeln genutzt werden, auch wenn sie nur aus jeweils fünf Items bestehen. Bezogen auf die Erfassung internalisierendes Verhaltens wird die Subskala „Emotionale Probleme“ herangezogen. Diese besteht aus fünf Items, welche die Häufigkeit von Kopfschmerzen, sich Sorgen machen, unglücklich sein, nervöses und anklammerndes Verhalten und Ängsten abfragen. Das externalisierende Verhalten wird über die Subskala „Verhaltensprobleme“ erfasst. Darin werden mit fünf Items die Häufigkeit von Wutanfällen, Folgsamkeit, häufiges Streiten, Lügen und Mogeln und Diebstahl zuhause abgefragt. Zur Erfassung der externalisierenden Komorbidität würde sich prinzipiell auch die Subskala „Verhaltensprobleme mit Gleichaltrigen“ anbieten. Inhaltlich und empirisch zeigt sich aber, dass diese Skala nicht passend ist. Es werden zwar die sozialen Beziehungen erfasst, allerdings bezogen auf die Anbindung und Stellung des Jugendlichen unter seinen Gleichaltrigen. In fünf Fragen wird erfasst, ob der Jugendliche eher ein Einzelgänger ist, zumindest einen guten Freund hat, beliebt ist, gehänselt wird und ob er besser mit Erwachsenen auskommt. Es geht aber nicht um oppositionelle oder aggressive Verhaltensweisen. In einer großen britischen Studie mit 18.222 Kindern und Jugendlichen zwischen 5 und 16 Jahren zeigte sich, dass die Skala „Verhaltensprobleme mit Gleichaltrigen“ hohe Korrelationen mit der Skala „Emotionale Probleme“ aufwies (Goodman, Lamping & Ploubidis, 2010). Basierend auf Faktorenanalysen sprechen sich die Autoren daher dafür aus, dass diese beiden Subskalen zu einem Faktor als internalisierende Verhaltensweisen zusammengefasst werden können.

5.2.4 Methodisches Vorgehen

In der vorliegenden Auswertung werden fünf Zugänge der Datenauswertung umgesetzt, die sich auf die fünf spezifizierten Fragestellungen beziehen. Es handelt sich dabei sowohl um univariate als auch multivariate Auswertungen. Univariaten Ansätzen ist gemein, dass sie nur Zusammenhänge von Variablen überprüfen, wenn nur eine einzige abhängige Variable vorliegt. Bei multivariaten Analysen können Zusammenhänge mit mehreren abhängigen Variablen überprüft werden (Bortz & Schuster, 2010).

Als zentrale abhängige Variable wird der Trainingserfolg festgelegt. Dieser kann entweder als ein Gesamtkonstrukt konzeptualisiert und operationalisiert werden oder als mehrere Teilaspekte, die jeweils unabhängig sind und einzeln erfasst werden. Die formulierten Hypothesen zu Fragestellung 1 a) beziehen sich nur auf den Trainingserfolg als Gesamtkonstrukt, welches über die Veränderung des Gesamtwertes operationalisiert wird. Dies hat den Vorteil, dass eine globale Einschätzung getroffen werden kann, die aus den Einzelaspekten abstrahiert ist. In weiterführenden Schritten erscheint es sinnvoll, den Trainingserfolg in seinen Einzelaspekten zu betrachten und mit den unabhängigen Variablen in Beziehung zu setzen, um so differenziertere Aussagen treffen zu können. Dieses Vorgehen wird im Weiteren vertieft erklärt. In Bezug auf die vorliegende Auswertung handelt es sich demnach um eine univariate Auswertung, wenn Zusammenhänge mit dem Erfolg als Gesamtkonstrukt überprüft werden oder wenn Zusammenhänge mit den Teilaspekten des Erfolges jeweils einzeln analysiert werden.

Wenn der Erfolg mittels Teilaspekten erfasst werden soll und diese Aspekte gleichzeitig analysiert werden, dann handelt es sich um eine multivariate Auswertung.

Als weiteres Unterscheidungskriterium der Auswertungsmöglichkeiten bietet sich die Anzahl der unabhängigen Variablen an. Analog zur Terminologie der Regressionsanalyse (Backhaus, Erichson, Plinke et al., 2011) werden Auswertungen mit nur einer unabhängigen Variablen als einfach bezeichnet und solche mit mehr als einer unabhängigen Variablen als multipel.

Schließlich bietet sich die Unterscheidung nach variablenzentrierten und personenzentrierten Ansätzen der Datenanalyse an (Muthén & Muthén, 2000). Beim variablenzentrierten Ansatz geht es um Zusammenhänge zwischen Variablen, z. B. darum eine abhängige Variable auf Grundlage mehrerer unabhängiger Variablen vorherzusagen. Beim personenzentrierten Vorgehen werden hingegen Zusammenhänge zwischen den Individuen bzw. den Fällen in der Stichprobe in den Blick genommen. Diese werden beispielsweise in einer Clusteranalyse nach ihrer Ähnlichkeit oder Unähnlichkeit in Gruppen eingeordnet.

In der weiteren Auswertung werden die folgenden fünf Auswertungsschritte umgesetzt, die im Anschluss genauer beschrieben werden. Sie beziehen sich der Reihenfolge nach auf die zuvor formulierten fünf Fragestellungen.

- (1) *Variablenzentrierte einfache univariate Auswertung*: Es wird gleichzeitig immer nur eine Variable auf ihren Zusammenhang mit dem Gesamterfolgsmaß überprüft. Nacheinander werden hierfür die Einzelerfolgsmaße herangezogen.
- (2) *Variablenzentrierte multiple univariate Auswertung*: Es wird der Zusammenhang mehrerer Variablen gleichzeitig mit dem Gesamterfolgsmaß und danach einzeln mit den Einzelerfolgsmaßen überprüft. Dies wird so umgesetzt, dass die Werte der Erfolgsvariablen zum Postmesszeitpunkt aus den Werten des Prämesszeitpunktes und anderen Variablen mittels einer multiplen Regression vorhergesagt werden.
- (3) *Variablenzentrierte multivariate Auswertung*: Es wird der Zusammenhang mehrerer Variablen mit allen Einzelerfolgsmaßen mittels einer multivariaten multiplen Regression gleichzeitig überprüft.
- (4) *Personenzentrierte univariate Auswertung*: Teilnehmer werden entsprechend dem Kriterium, ob sie eine reliable Verbesserung oder keine Veränderung bzw. Verschlechterung ihrer Werte aufweisen, in Gruppen eingeordnet. Dazu wird der „Reliable Change Index“ (RCI; Jacobson & Truax, 1991) genutzt. Im Anschluss werden die Gruppen verglichen.
- (5) *Personenzentrierte multivariate Auswertung*: Die Teilnehmer werden aufgrund ihrer Ähnlichkeit in den zusammengefassten Einzelerfolgsmaßen mittels einer Clusteranalyse in Gruppen eingeordnet. Danach werden die Gruppen verglichen.

Variablenzentrierte einfache univariate Auswertung

Fragestellung 1 a: *Wie hängen die Merkmale der Teilnehmer und der Trainingsdurchführung jeweils einzeln mit den verschiedenen Aspekten des Trainingserfolges zusammen?*

In diesem Auswertungsschritt werden die Merkmale der Jugendlichen bzw. der Trainingsdurchführung jeweils einzeln mit dem Trainingserfolg in Beziehung gesetzt. Als Grundlage dieser Auswertung muss eine neue Variable bestimmt werden, die den Trainingserfolg abbildet. Aus dieser werden Differenzwerte, Residual Gain Scores und prozentuale Veränderungen bestimmt. Anhand der Daten wird im Anschluss überprüft, wie geeignet die Maße für die weitere Auswertung erscheinen. Wenn sich herausstellt, dass mit allen Maßen dieselbe Information erfasst wird, wird nur ein geeignetes Maß weiterverwendet.

Bei den dichotomen Variablen werden die Teilnehmer nach der jeweiligen Variablen in zwei Gruppen aufgeteilt, deren Werte verglichen werden, z. B. der durchschnittliche Trainingserfolg der teilnehmenden Mädchen versus den der Jungen. Bei metrischen Variablen werden Korrelationen berechnet, z. B. zwischen Alter und Trainingserfolg. Bevor die Zusammenhänge mit den Veränderungsmaßen bestimmt werden, werden die Daten des Prämesszeitpunktes analysiert, um zu überprüfen, ob schon vor dem Lerntraining Unterschiede in der jeweiligen Variablen vorlagen.

Zur Überprüfung und Einschätzung der Generalisierbarkeit der Effekte werden Signifikanztests durchgeführt. Die Durchführung eines Signifikanztests dient vor allem der Überprüfung von Hypothesen. Nach Bortz und Döring (2006) ist es wie bereits beschrieben aber auch möglich, explorativ auf Signifikanz zu testen, sollte sich ein interessanter Effekt erst zeigen, wenn die Daten schon erhoben sind und inspiziert werden.

Zum Einsatz von Signifikanztests gilt für die vorliegende Untersuchung: Die zuvor aufgestellten Hypothesen werden als Erstes mittels Signifikanztests überprüft. Der vermutete Effekt befindet sich in der Alternativhypothese und in der Nullhypothese das Gegenteil. Die Nullhypothese wird überprüft. Finden sich die in den Hypothesen angenommenen Effekte in den Daten bei sehr geringer Irrtumswahrscheinlichkeit, wird der Effekt als signifikant bezeichnet. Die zuvor aufgestellte Nullhypothese, die gegen den vermuteten Effekt spricht, wird verworfen und es wird angenommen, dass die Alternativhypothese auch für die Grundgesamtheit zutreffend ist. Als Irrtumswahrscheinlichkeit wird $\alpha = 5\%$ angenommen. Weiterhin gilt: Falls sich interessante Effekte erst während der Dateninspektion zeigen, so werden auch diese mittels Signifikanztest überprüft. Stellt sich heraus, dass es sich um signifikante Effekte handelt, kann davon ausgegangen werden, dass der gefundene Effekt sehr wahrscheinlich auch in der Grundgesamtheit vorliegt. Da zuvor aber keine Null- und Alternativhypothese aufgestellt wurden, kann das Ergebnis nicht als Bestätigung oder Widerspruch zu einer Annahme gelten. Dieses könnte erst in einer nachfolgenden Untersuchung überprüft werden (Bortz & Döring, 2006).

Bei den dichotomen Variablen werden t -Tests für unabhängige Stichproben durchgeführt, da zwei Mittelwerte unabhängiger Gruppen verglichen werden sollen. Für die metrischen Variablen werden Korrelationen berechnet, deren Signifikanz ebenfalls mittels t -Tests überprüft wird (Bühner & Ziegler, 2009). Welche Voraussetzungen für die Durchführung von t -Tests erfüllt sein müssen bzw. welche Verletzungen der Annahmen toleriert werden können, wird in der Literatur kontrovers diskutiert. Einigkeit besteht darüber, dass die Stichproben unabhängig voneinander und intervallskaliert sein müssen (Sedlmeier & Renkewitz, 2008). Der t -Test ist allerdings relativ robust gegen eine Verletzung der Normalverteilungsvoraussetzung (Bortz & Schuster, 2010), insbesondere bei Stichproben mit $n \geq 30$ (Kubinger, Rasch & Moder, 2009).

Wenn bei einer Variable jedoch sehr kleine Subgruppen mit $n < 30$ entstehen, so wird überprüft, ob die entsprechende Variable normalverteilt ist. Zur Überprüfung auf Normalverteilung wird der Shapiro-Wilk-Test empfohlen (Brosius, 2011; Yazici & Yolacan, 2007). Falls keine Normalverteilung vorliegt, wird zum Vergleich von Mittelwerten bei dichotomen Merkmalen der Mann-Whitney-Test dem t -Test vorgezogen. Bei metrischen Variablen wird bei nicht vorliegender Normalverteilung zur Überprüfung von Zusammenhängen das robustere Maß Kendalls τ dem Pearson-Korrelationskoeffizienten r vorgezogen (Bortz, Lienert & Boehnke, 2008).

Weiterhin werden für das Alter der Teilnahme wie bei Durlak und Kollegen (1991) und Dush et al. (1989) Effektstärken für drei Altersgruppen bestimmt. Dies begründet sich damit, dass insbesondere die Studie von Durlak et al. häufig als Beleg dafür herangezogen wird, dass kognitiv-behaviorale Interventionen stärker bei älteren Kindern bzw. Jugendlichen wirken. Daher soll überprüft werden, ob die Ergebnisse dieser wegweisenden Studie mit derselben Methodik repliziert werden können. Zur Effektstärkenbestimmung wird dazu analog zu Cohens d (1988) die Post-prä-Differenz durch die Standardabweichung der Werte zum Prämesszeitpunkt geteilt.

Das Vorgehen der einfachen univariaten Auswertung lässt sich zusammenfassend folgendermaßen skizzieren:

- (1) Analyse der Werte zum Prämesszeitpunkt,
- (2) Analyse der Veränderungswerte, Überprüfung der aufgestellten Hypothesen mittels geeignetem Signifikanztest bzw. deskriptive Auswertung bei Variablen ohne Hypothese und gegebenenfalls Signifikanztest auf Probe.

Variablenzentrierte multiple univariate Auswertung

Fragestellung 1 b: *Wie hängen die Merkmale der Teilnehmer und der Trainingsdurchführung – wenn sie gleichzeitig berücksichtigt werden – mit den verschiedenen Aspekten des Trainingserfolges zusammen?*

In diesem Auswertungsschritt gehen die Merkmale der Jugendlichen und der Durchführung gleichzeitig in die Auswertung ein. Das Ziel dieser Auswertung ist die Vorhersage des Trainingserfolges aus den Merkmalen der Teilnehmer und der Trainingsdurchführung. Methode der Wahl ist daher die Regressionsanalyse, die in ähnlichen Untersuchungen häufig zum Einsatz kommt (Borge, Hoffart & Sexton, 2010; Lincoln et al., 2005; Scholing & Emmelkamp, 1999; Schuurmans et al., 2009).

Abhängige Variable (AV) bei dieser Auswertung ist der Trainingserfolg, unabhängige Variablen (UV) sind die Merkmale der Teilnehmer und der Durchführung. Der Trainingserfolg wird in diesem Schritt univariat erfasst, das heißt entweder als ein zusammengefasstes Gesamterfolgsmaß oder die unterschiedlichen Aspekte des Trainingserfolges werden jeweils einzeln ausgewertet. Als AV werden entweder der Wert des Gesamtproblemmaßes zum Postmesszeitpunkt oder die Veränderungswerte, die schon in der einfachen univariaten Auswertung genutzt wurden, herangezogen.

Wenn wie in dieser Auswertung nicht klar ist, wie viele und welche Variablen in die Regression eingeschlossen werden sollen, stellt sich die Frage, ob ein Modell mit vielen Variablen bessere Vorhersagen treffen kann als ein Modell mit wenigen Variablen (Yan & Su, 2009). Grundlegendes Prinzip ist dabei, ein komplettes Modell („full/ complete model“) mit einem oder mehreren reduzierten Modellen zu vergleichen und zu entscheiden, welches die bessere Vorhersage der AV ermöglicht, ohne jedoch zu viele oder redundante Variablen aufzunehmen (Kleinbaum, Kupper, Nizam & Muller, 2008). Prinzipiell stehen zwei Vorgehensweisen zur Verfügung. Die eine ist mit einem einfachen Modell anzufangen und mögliche Kandidatenvariablen nach und nach hinzuzufügen. Die andere Möglichkeit ist es, zunächst alle verfügbaren Variablen als Prädiktoren einzubeziehen, also mit dem kompletten Modell zu beginnen, und dann schrittweise die Kandidaten zu eliminieren, die keinen Zusammenhang mit der AV aufweisen oder nicht zur Varianzaufklärung beitragen. Wenn durch Hinzunahme oder Weglassen von Variablen keine zusätzliche oder nur sehr geringe zusätzliche Varianz aufgeklärt wird, so wird das reduzierte Modell bevorzugt. Die Varianzaufklärung (R^2) gibt Auskunft über die Güte der Vorhersage mittels der ermittelten Regressionsgleichung (Bühl, 2010).

Für die vorliegende Auswertung werden in Frage kommende Variablen als UV schrittweise mit der Statistiksoftware SPSS 20 überprüft. Das Programm nimmt bei der Einstellung „schrittweise“ aus den UV diejenige mit der höchsten partiellen Korrelation mit der AV in die Regressionsgleichung auf und geht so schrittweise weiter vor. Nach jeder Neuaufnahme werden die bereits aufgenommenen Variablen nach der Rückwärts-Methode untersucht. Dabei wird die Variable mit dem kleinsten partiellen Korrelationskoeffizienten wieder ausgeschlossen, allerdings nur, wenn keine Signifikanz vorliegt (Bühl, 2010).

Bevor die unabhängigen Variablen für die Regressionsanalyse genutzt werden, wird eine Zentrierung durchgeführt. Dabei wird jede Variable anhand eines Referenzwertes umkodiert. Bei metrischen Variablen wird der Mittelwert als Referenzwert empfohlen.

Dieser wird von den Variablen jeweils abgezogen, so dass die umkodierten Werte die Abweichung vom Mittelwert angeben. Bei dichotomen Variablen wird eine Kodierung mit $-1/2$ und $+1/2$ empfohlen (Kraemer & Blasey, 2004). So kann dem Problem der Multikollinearität vorgebeugt werden, das bei starken Korrelationen der Variablen untereinander auftreten kann, die in empirischen Daten oftmals vorliegen. Die Parameterschätzungen und die Modellgüte verschlechtern sich bei großer Multikollinearität. Die Variablen des Trainingserfolges bzw. Gesamtproblemmaßes werden keiner Zentrierung unterzogen, da es keine Vorteile bietet, die abhängigen Variablen zu zentrieren (Tabachnick & Fidell, 2007).

Damit gute Schätzungen mittels Regressionsanalysen gemacht werden können, müssen bestimmte Modellprämissen erfüllt sein: Mittels visueller Inspektion der Residuen wird überprüft, ob Heteroskedastizität vorliegt. Wenn dies ausgeschlossen werden kann, ist davon auszugehen, dass die Residuen gleiche Varianzen haben, was eine Voraussetzung der Regressionsanalyse ist. Weiterhin wird überprüft, ob eine Autokorrelation bzw. Korrelation der Residuen vorliegt. Dazu wird die Durbin-Watson-Statistik herangezogen. Werte dieser Statistik, die nahe beim optimalen Wert von 2 liegen, zeigen an, dass Autokorrelation ausgeschlossen werden kann. Weiterhin soll Multikollinearität ausgeschlossen werden, also eine hohe zwischen unabhängigen Variablen. Dazu werden der Toleranzwert und der Varianzinflationsfaktor bzw. VIF-Wert bestimmt, die beide nahe bei 1 liegen sollten (Urban & Mayerl, 2011).

Das multiple univariate Vorgehen lässt sich also folgendermaßen beschreiben:

- (1) Zentrierung der unabhängigen Variablen anhand des Mittelwertes bei metrischen Variablen bzw. Umkodierung in die Ausprägungen $-1/2$ und $+1/2$ bei dichotomen Variablen,
- (2) Durchführung einer schrittweisen hierarchischen Regressionsanalyse mit den Werten des Postmesszeitpunktes als abhängige Variablen,
- (3) Durchführung einer schrittweisen hierarchischen Regressionsanalyse mit den Veränderungswerten als abhängige Variablen,
- (4) Überprüfung der Modellprämissen der Regressionsanalysen.

Variablenzentrierte multivariate Auswertung

Fragestellung 1 c: *Wie hängen die Merkmale der Teilnehmer und der Trainingsdurchführung – wenn sie gleichzeitig berücksichtigt werden – mit den zusammengefassten Aspekten des Trainingserfolges zusammen?*

Bei der multiplen univariaten Regression können zwar mehrere unabhängige Variablen berücksichtigt werden, aber nur eine abhängige Variable. Um auch mehrere abhängige Variablen gleichzeitig betrachten zu können, wird eine multiple multivariate Regression durchgeführt. Eine solche multivariate Auswertung kann zu signifikanten Ergebnissen

führen, auch wenn die zugehörigen univariaten Analysen nicht zu signifikanten Ergebnissen führen (Tabachnick & Fidell, 2007).

Eine multivariate Regression durchzuführen ist nur dann sinnvoll, wenn die abhängigen Variablen korreliert sind (Bühl, 2010). Diese Überprüfung muss also nach der Bestimmung der zusammengefassten Erfolgsmaße geschehen. Wenn dies nicht der Fall ist, ist es ausreichend, univariate Analysen für jede abhängige Variable durchzuführen (Backhaus, Erichson, Plinke & Weiber, 2011).

Die multivariate Regression kann mit SPSS 20 durchgeführt werden. Sie wird dort als multivariate Varianzanalyse eingeordnet, da beide Verfahren dem allgemeinen linearen Modell zuzuordnen sind (Field, 2009). Die Varianzanalyse kann somit als eine Form der multiplen Regression angesehen werden, bei der die unabhängigen Variablen Ausprägungen von diskreten Variablen sind und keine kontinuierliche Variablen wie sonst bei der Regression üblich (Tabachnick & Fidell, 2007).

Das Vorgehen der variablenzentrierten multivariaten Auswertung gestaltet sich demnach folgendermaßen:

- (1) Überprüfung der abhängigen Variablen des Trainingserfolges auf signifikante Korrelationen,
- (2) Durchführung einer multivariaten Regression mit den signifikant korrelierenden abhängigen Variablen.

Personenzentrierte univariate Auswertung

Fragestellung 2 a: Lassen sich Gruppen von Teilnehmern mit besonders starken Verbesserungen und weniger starken bzw. ohne Verbesserungen durch das Training identifizieren und wenn ja, welche Unterschiede weisen diese Gruppen jeweils auf?

In dieser personenzentrierten Auswertung werden die Teilnehmer anhand ihres Trainingserfolges gruppiert. Grundlegend könnten zwei Gruppen gebildet werden, Teilnehmer mit erfolgreichen und Teilnehmer mit wenig bzw. nicht erfolgreichen Trainingsverläufen, die im Anschluss verglichen werden (Cleophas et al., 2009). Ziel dieses Auswertungsschrittes ist also die Identifizierung und Analyse von Subgruppen in der Gesamtgruppe der Teilnehmer (Pocock et al., 2002). Eine Möglichkeit ist an dieser Stelle, die Stichprobe in eine Rangfolge nach dem Ausmaß ihres Trainingserfolges zu bringen und bei einer kritischen Grenze in zwei Gruppen nach guten und schlechten Trainingsverläufen zu unterteilen. Möglich wäre, die Mitte der Verteilung als kritische Grenze anzunehmen. Da eine Normalverteilung der Variablen Trainingserfolg anzunehmen ist, erscheint eine Teilung in der Mitte nicht besonders sinnvoll. Teilnehmer mit sehr ähnlichen Veränderungswerten in der Mitte der Verteilung würden entgegengesetzten Gruppen zugeordnet. Eine bessere Möglichkeit bestünde darin, nur extreme Ausprägungen des Gesamterfolgswertes zu vergleichen, beispielsweise die 15 besten Teilnehmer mit

den 15 schlechtesten. Dieses Verfahren wurde in der Vergangenheit oftmals angewendet (Alf & Abrahams, 1975), stößt heute jedoch weitreichend auf Kritik (Cohen, 1983; MacCallum et al., 2002; Preacher et al., 2005). Wenn ein Effekt nur dann festgestellt wird, wenn die Mitte der Verteilung nicht beachtet wird und nur die extremen Fälle verglichen werden, ist kritisch zu überprüfen, ob es sich um ein Artefakt handelt.

Ein besser geeignetes Kriterium zur Gruppenbildung ist der „Reliable Change Index“ (RCI; Jacobson & Truax, 1991; Wise, 2004). Dieser stellt einen kritischen Wert dar, der angibt, ab wann eine individuelle Veränderung überzufällig von Null verschieden ist. Anhand dieser kritischen Grenze sollen die Teilnehmer als „Responder“ oder „Non-Responder“ (Hiller et al., 2009; Mohr, 1995) klassifiziert werden. Es folgt eine Beschreibung der beiden Gruppen anhand ihrer Verteilungsparameter bezogen auf Merkmale der Teilnehmer und der Durchführung und der Erfolgsmaße zum Prämesszeitpunkt. Durch die Gruppenbeschreibung und den Vergleich soll überprüft werden, ob es Maße gibt, die in beiden Gruppen unterschiedlich verteilt sind. Bei diesen könnte es sich um Moderatoren des Trainingserfolges handeln. Dabei sind keine grundlegend abweichenden Ergebnisse im Vergleich zu den vorangehenden Auswertungsschritten zu erwarten. Es handelt sich modellhaft um eine Annäherung an dieselbe Fragestellung, jedoch ausgehend von den abhängigen Variablen in Richtung der unabhängigen Variablen.

Das Vorgehen der personenzentrierten univariaten Auswertung sieht also folgendermaßen aus:

- (1) Bestimmung des RCI,
- (2) Aufteilung der Teilnehmer in zwei Gruppen: Responder – Non-Responder,
- (3) Vergleichende Beschreibung der Gruppen anhand von Mittelwerten der Merkmale der Teilnehmer und der Durchführung und Ausprägung von Erfolgskriterien zum Prämesszeitpunkt,
- (4) Signifikanztest auf Probe, wenn sich Unterschiede zeigen.

Personenzentrierte multivariate Auswertung

Fragestellung 2 b: Lassen sich im Vergleich der Teilnehmer untereinander – bei gleichzeitiger Berücksichtigung mehrerer Einzelaspekte des Gesamttrainingserfolgs – Gruppen mit ähnlichen Veränderungsmustern feststellen und wenn ja, welche Unterschiede weisen diese Gruppen jeweils auf?

Im vorangegangenen Auswertungsschritt wurden die Teilnehmer nach dem Ausmaß ihres Gesamttrainingserfolgs in zwei Gruppen nach erfolgreichen und nicht erfolgreichen Trainingsabsolventen eingeordnet. In diesem Auswertungsschritt wird weiterführend und ergänzend eine multiple Gruppierung auf Grundlage mehrerer Erfolgsmaße durchgeführt. Nach der Bestimmung zusammengefasster Veränderungsmaße stellt sich

die Frage, ob charakteristische Muster der Veränderung feststellbar sind. So wäre denkbar, dass sich hinter großen Verbesserungen im Gesamtmaß unterschiedliche Muster auf der darunter liegenden Ebene verbergen. Ein moderates Gesamtveränderungsmaß kann z. B. auf gleichförmig moderate Verbesserungen in allen Einzelmaßen zurückzuführen sein, aber auch auf extreme Veränderungen bei einigen Maßen und genau gegenläufigen Tendenzen in anderen. Daher werden die Teilnehmer nach ihren Veränderungswerten in mehreren zusammengefassten Veränderungsmaßen gruppiert. So können auf einer deskriptiven Ebene Veränderungsmuster festgestellt werden. Diese Gruppen können mit Hilfe einer Clusteranalyse über die Erfolgsmaße bestimmt werden, so wie es beispielsweise auch bei Lauth und Minsel (2011) zur Evaluation einer Therapie für Erwachsene mit ADHS umgesetzt wurde.

Im Anschluss werden die Clusterzuordnungen der Teilnehmer genutzt, um die gefundenen Cluster genauer zu beschreiben. Dazu werden die Mittelwerte der Cluster anhand der Personenvariablen der Teilnehmer und Outcomes zum Prämesszeitpunkt verglichen. So sollen Besonderheiten innerhalb der Cluster und im Vergleich mit den anderen Clustern deskriptiv festgestellt werden. Diese Auswertung bleibt aufgrund der geringen Zellengröße und nicht vorformulierter Hypothesen auf einer deskriptiven, explorativen Ebene.

Zur Gruppierung der Teilnehmer wird eine hierarchische Clusteranalyse durchgeführt. Dabei werden die Teilnehmer aufgrund ihrer Ähnlichkeit anhand bestimmter Variablen zu Gruppen zusammengefasst. Die Gruppen, die als Cluster bezeichnet werden, sollen in sich möglichst homogen und ähnlich sein sowie gleichzeitig möglichst wenig Ähnlichkeit mit den anderen Gruppen aufweisen (Backhaus, Erichson, Plinke et al., 2011). Es werden nur Teilnehmer mit vollständigen Gesamtwerten für alle Problemmaße einbezogen und es wird davon abgesehen für fehlende Werte die jeweiligen Mittelwerte einzusetzen, da es zu Verzerrungen und einer Varianzverringerng führen kann. Für die meisten Anwendungen der Clusteranalyse ist daher keine Ersetzung fehlender Werte durch Mittelwerte angeraten (Backhaus, Erichson, Plinke et al., 2011).

Das Vorgehen der personenzentrierten multivariaten Auswertung lässt sich also folgendermaßen skizzieren:

- (1) Gruppierung der Teilnehmer nach ihren Veränderungswerten in den unterschiedlichen Problemmaßen mittels hierarchischer Clusteranalyse,
- (2) Vergleichende Beschreibung der Cluster anhand der Mittelwerte der Merkmale von Teilnehmern und der Durchführung und Ausprägung der Erfolgskriterien des Prämesszeitpunktes.

5.2.5 Stichprobe

Bei Projektstart wurden zur Rekrutierung von Teilnehmern alle ortsnahen Schulen, Kinderärzte, Kinder- und Jugendlichenpsychiater und -psychotherapeuten, Kinder- und

Jugendlichenpsychiatrien, Beratungsstellen und andere Fördereinrichtungen angeschrieben mit der Bitte geeignete Jugendliche zu überweisen. Zusätzlich wurde eine Pressemitteilung herausgegeben, die in der lokalen Presse sowohl in kostenpflichtigen als auch in kostenlosen Zeitungen abgedruckt wurde und auch online erschien. Interessenten für eine Teilnahme wurden auf einer Warteliste gesammelt und stratifiziert randomisiert den experimentellen Gruppen der allgemeinen Wirksamkeitsprüfung zugewiesen. Es handelt sich damit um eine Inanspruchnahmestichprobe (Bortz, Lienert, Barskova, Leitner & Oesterreich, 2008).

Einschlusskriterien für die Teilnahme an der Studie waren:

- *Alter:* 12-17 Jahre.
- *ADHS liegt vor:* Eine ADHS-Diagnose wurde entweder in der Vorbehandlung gestellt und konnte im Erstgespräch in der Universität anamnestisch bestätigt werden oder wurde erstmals in der Universität gestellt. Dazu kamen standardisierte Ratinginstrumente wie das Diagnostik-System für psychische Störungen im Kindes- und Jugendalter (Döpfner & Lehmkuhl, 2006) und der Fragebogen zu Stärken und Schwächen (Goodman, 1997) zum Einsatz, die von den Jugendlichen selbst, von den Eltern und von Lehrern ausgefüllt wurden. Darüber hinaus wurde das Wender-Reimherr-Interview aus den Homburger ADHS-Skalen für Erwachsene (Michael Rösler, et al., 2008) durchgeführt und die Testbatterie zur Aufmerksamkeitsprüfung (Zimmermann & Fimm, 2008). Die Eltern wurden zudem gebeten, Gutachten über vorherige psychiatrische Untersuchungen und therapeutische Behandlungen einzureichen, falls diese vorlagen.

Ausschlusskriterien von der Teilnahme an der Studie waren:

- *Akute Krisensituation oder primär behandlungsbedürftige komorbide Störungen:* Eine Aufnahme in das Lerntraining war nicht indiziert, wenn eine andere Problematik so präsent war, dass diese primär behandlungswürdig erschien. Darunter fällt eine akute Selbst- oder Fremdgefährdung des Jugendlichen, z. B. bei selbstverletzenden und suizidalen Verhaltensweisen oder sehr stark physisch aggressives Verhalten. Weiterhin war eine Teilnahme in akuten Krisensituationen abgeraten, z. B. bei gravierenden Erziehungsschwierigkeiten, die zu einer Entnahme des Jugendlichen aus der Familie führen könnten.
- *Ablehnung der Teilnahme durch den Jugendlichen:* Wenn nur die Eltern eine Teilnahme am Training wünschten, sich die Jugendlichen aber selbst entschieden dagegen aussprachen, wurde gegen die Aufnahme in das Training entschieden.

In die vorliegende Auswertung wurden alle Teilnehmer aufgenommen, die das Lerntraining LeJA vollständig durchlaufen haben. Es handelt sich damit um eine Completeranalyse (Hedeker & Gibbons, 2006). Es wurden die Daten der Interventionsgruppe der allgemeinen Wirksamkeitsstudie genutzt ($n = 43$) und die Daten von Jugendlichen, die

in der allgemeinen Wirksamkeitsprüfung der Wartegruppe angehörten und erst nach der Wartezeit ein LeJA-Training durchliefen ($n = 24$). Bei der Interventionsgruppe wurden die Daten der Messzeitpunkte t_0 und t_1 verwendet, bei der Wartekontrollgruppe die Daten zu den Messzeitpunkten t_1 und t_2 .

Drei begonnene Trainings konnten nicht ordnungsgemäß abgeschlossen werden. Bei einem Jugendlichen fehlte nach einigen Trainingssitzungen die Bereitschaft das Training fortzusetzen, so dass es abgebrochen werden musste. Bei zwei Jugendlichen musste das Training abgebrochen werden, weil sich unabhängig vom Training eine starke familiäre Problematik ergeben hatte. In beiden Fällen verlief das Training anfänglich gut, dann ergaben sich zunächst organisatorische Probleme und es konnten keine regelmäßigen Termine mehr gefunden werden, da die beiden allein erziehenden Mütter kaum erreichbar waren. Schließlich stellte sich heraus, dass es so starke Erziehungsschwierigkeiten gab, dass die beiden Jugendlichen nicht mehr im Elternhaus verblieben und nicht mehr erreicht werden konnten.

Die resultierende Stichprobe besteht insgesamt aus 67 Jugendlichen, 13 Mädchen (19 %) und 54 Jungen (81 %). Zum Messzeitpunkt vor dem Training bekamen 37 Teilnehmer keine ADHS bezogene Medikation (55 %), 30 Teilnehmer nahmen ein Medikament ein (45 %), zwei davon Atomoxetin, der Rest Methylphenidat. Alle für die Zielgruppe infrage kommenden Schulformen waren in der Stichprobe vertreten. Die meisten Jugendlichen in der Stichprobe besuchten eine Realschule ($n = 26$), gefolgt vom Besuch des Gymnasiums ($n = 13$), Hauptschule ($n = 11$), Gesamtschule ($n = 10$), Förderschule ($n = 5$; davon $n = 2$ Förderbedarf Lernen, $n = 2$ Förderbedarf Hören, $n = 1$ Förderbedarf körperliche und motorische Entwicklung) und jeweils einmal Waldorfschule und Berufsschule (siehe Tabelle 10).

Tabelle 10: Stichprobenverteilung nach Geschlecht, Medikamenteneinnahme zum Prämesszeitpunkt und Schulform ($N = 67$)

		Häufigkeit	Prozent
Geschlecht	männlich	54	81
	weiblich	13	19
Medikamenteneinnahme zu t_0	nein	37	55
	ja	30	45
Schulform	Hauptschule	11	16
	Realschule	26	39
	Gymnasium	13	19
	Gesamtschule	10	15
	Förderschule	5	8
	Waldorfschule	1	2
	Berufsschule	1	2

Die Teilnehmer waren zwischen 12;0 und 17;9 Jahren alt, im Durchschnitt 14;2 Jahre bei einer Standardabweichung von 18 Monaten. Die meisten Teilnehmer waren 13 Jahre alt, die zweitmeisten waren 12 Jahre alt, dann folgten in absteigender Reihenfolge die Altersstufen 14 bis 17 (siehe Abbildung 21).

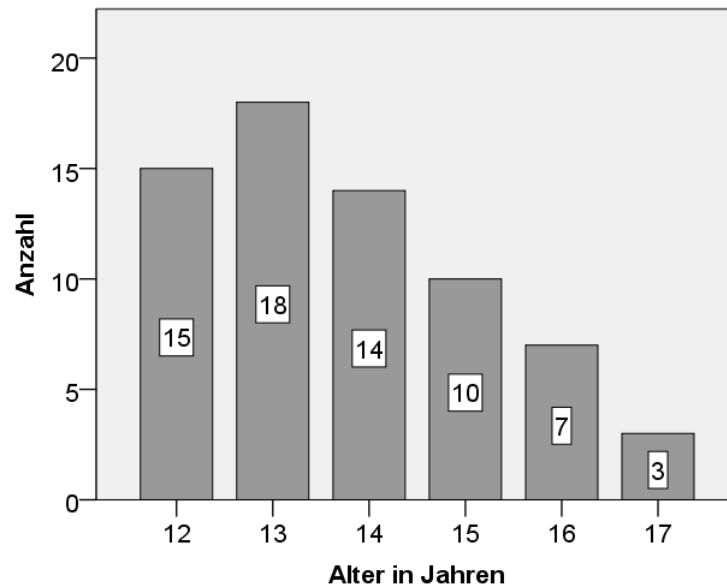


Abbildung 21: Altersverteilung der Stichprobe in Jahren

5.2.6 Durchführung

Interessierte Teilnehmer wurden in eine Warteliste aufgenommen. Im Rahmen der allgemeinen Wirksamkeitsprüfung fand zunächst eine stratifizierte Randomisierung nach Geschlecht und Alter statt. Dabei wurde ein „trickle process“ (Braucht & Reichardt, 1993) umgesetzt, der angepasst an die kontinuierlich eintreffenden Weiteranmeldungen fortgeführt wurde. Dieses Vorgehen ist in der empirischen klinischen Forschung häufig, da bei Studienbeginn oftmals nicht die gesamte Stichprobe vorliegt und randomisiert zugeordnet werden kann. Deshalb wurde in unterschiedlich großen Blöcken nach und nach randomisierte Gruppenzuordnungen mittels eines Zufallsgenerators vorgenommen.

Nach der Gruppenzuordnung wurde mit Eltern und Jugendlichen ein erstes anamnestisches Gespräch vom Autoren oder einem anderen Diplom-Psychologen durchgeführt. Darin wurde überprüft, ob die Einschlusskriterien zur Teilnahme an der Studie zutrafen, und Informationen zum Ablauf des Trainings vermittelt. Bei Eignung und Bereitschaft zur Teilnahme wurde den Jugendlichen der Interventionsgruppe eine Trainerin oder ein Trainer zugewiesen und das Training startete unmittelbar im Anschluss an das Erstgespräch. Bei den Jugendlichen der Wartekontrollgruppe begann eine ca. sechsmonatige Wartezeit. Danach wurden sie und ihre Eltern zu einem weiteren Gespräch eingeladen, in welchem die Outcomemaße erneut erhoben und besondere Ereignisse während der Wartezeit abgefragt wurden. Im Anschluss an dieses Gespräch begann das Training.

Die Durchführung der Trainings erfolgte im Zeitraum von Januar 2009 bis Juli 2012. Die Trainings wurden vom Autor, einem Kollegen Diplom-Psychologen und zum größten Anteil von angeleiteten Studierenden der Sonder- und Rehabilitationspädagogik der Universität Oldenburg durchgeführt. Das Training wurde gemäß Manual (Linderkamp, et al., 2011) mit wöchentlicher Supervision durchgeführt. Begleitend wurden drei Elternabende angeboten, bei denen Informationen zu ADHS und zum Lerntaining vermittelt wurden.

Direkt vor Beginn der Trainings und direkt nach dem Abschluss wurden alle relevanten Erfolgskriterien erfasst. Im Durchschnitt vergingen zwischen Prä- und Postmessung sechs Monate ($SD = 1.7$). Das kürzeste Training war nach drei Monaten abgeschlossen, das längste Training nach zehn Monaten (siehe Abbildung 22).

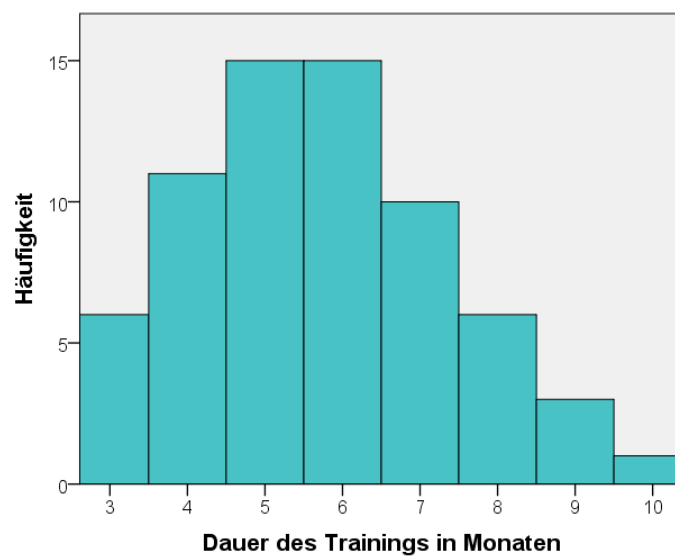


Abbildung 22: Verteilung der Häufigkeiten der Dauer des Trainings in Monaten

Laut Trainingskonzeption sollen 16 bis 20 Trainingssitzungen durchgeführt werden, je nach Bedarf des Jugendlichen. Zumeist wurden 20 Sitzungen durchgeführt, der Mittelwert beträgt 18 ($SD = 2.1$). Das kürzeste Training wurde nach 13 Sitzungen beendet, das längste dauerte 23 Sitzungen (siehe Abbildung 23). Diese Fälle wurden nicht von der Auswertung ausgeschlossen, sondern die Trainingsdauer als eine mögliche Moderatorvariable des Trainingserfolges untersucht. Somit kann eine Aussage zur Auswirkung der Länge des Trainings gemacht werden. In zwei Fällen wurde die genaue Sitzungsanzahl durch die Trainer entgegen der Instruktion nicht protokolliert und konnte nach Trainingsabschluss nicht mehr rekonstruiert werden.

Nach Beendigung des Trainings fand ein gemeinsames Abschlussgespräch mit dem Jugendlichen, den Eltern, den Trainern und dem Autor oder seinem Kollegen statt. Darin wurden die im Training behandelten Inhalte besprochen und die weitere Perspektive reflektiert. Die meisten Jugendlichen nahmen das weiterführende Angebot eines Nachsorgetermins einige Wochen nach Trainingsende an. Die Jugendlichen der Interventi-

onsgruppe wurden mindestens ein halbes Jahr nach Trainingsabschluss zu einem Follow-up eingeladen.

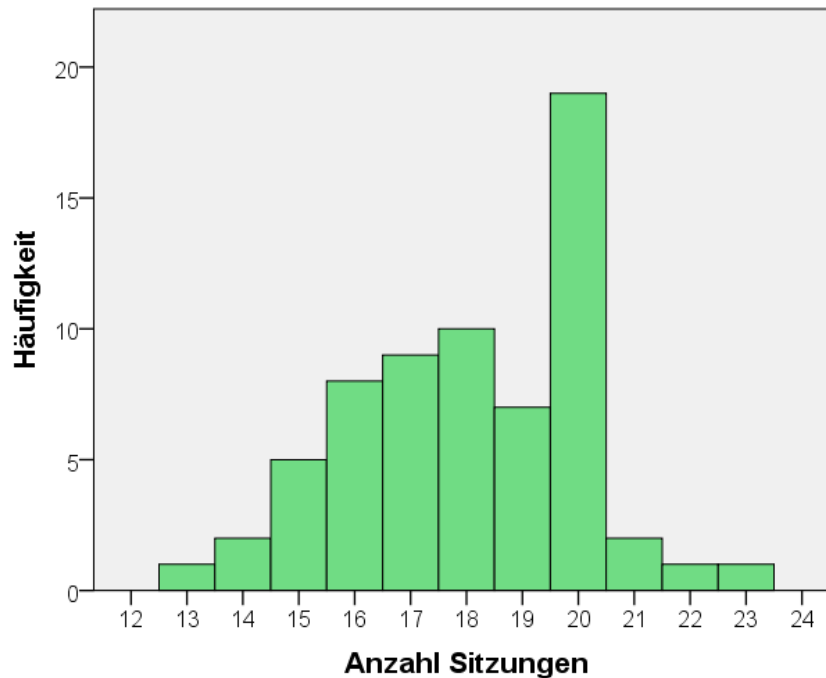


Abbildung 23: Verteilung der Häufigkeiten der Anzahl der durchgeführten Sitzungen pro Teilnehmer (bei zwei Trainings konnte die genaue Sitzungsanzahl nicht rekonstruiert werden)

5.3 Aufbereitung der Daten

Bevor die Hypothesenprüfung und weitere Auswertungsschritte durchgeführt werden, wird eine Datenaufbereitung durchgeführt. Diese beginnt mit einer Inspektion der Veränderungen, die vom Prämesszeitpunkt zum Postmesszeitpunkt stattgefunden haben. Danach werden die für eine empirische Untersuchung typischen fehlenden Werte analysiert und gegebenenfalls ersetzt. Im Anschluss wird ein Gesamtproblemmaß aus mehreren Einzelvariablen zusammengestellt, welches die Grundlage für die weitere Auswertung bildet. Schließlich werden die Werte des Prä- und Postmesszeitpunktes zu Veränderungswerten verrechnet.

5.3.1 Prä-post-Veränderungen auf Variablenebene

Bevor die Hypothesenprüfung und weitere Auswertungen beginnen, werden alle Erfolgskriterien deskriptiv dargestellt. Von besonderem Interesse ist dabei, ob sich Decken- oder Bodeneffekte offenbaren, die eine Veränderungsmessung beeinträchtigen könnten. Es wäre z. B. problematisch, wenn die Symptomschwere in der Prämessung bereits so niedrig wäre, dass sie kaum noch verbessert werden könnte. Weiterhin können durch die Bestimmung von Effektstärken anhand der jeweiligen Standardabweichung unterschiedliche Instrumente miteinander verglichen werden.

Es folgt eine deskriptive Auswertung der Prä- und Postmesswerte. Zusammengehörige Messungen sind direkt untereinander dargestellt. Angegeben sind in Tabelle 11 Gruppengröße (N), kleinster und größter Wert (Min , Max), Mittelwert (M), Standardabweichung (SD) und Effektstärken, standardisiert an der Streuung der Präwerte ($ES_{prä}$) und anhand einer gepoolten Größe aus der Vorher- und der Nachher-Streuung (ES_{pool}) (Maier-Riehle & Zwingmann, 2000). Die gepoolte Effektstärke, die unterschiedliche Gruppengrößen und Standardabweichungen berücksichtigt, wurde nach Hartmann, Herzog und Drinkmann (1992) folgendermaßen bestimmt:

$$ES_{pool} = \frac{M_{t1} - M_{t2}}{\sqrt{\frac{(N_{t1} - 1)SD_{t1}^2 + (N_{t2} - 1)SD_{t2}^2}{N_{t1} + N_{t2} - 2}}}$$

Tabelle 11: Deskriptive Auswertung der Outcomemaße vor und nach dem Training

Instrument/ Gesamtskalen (Rohwerte)	Werte	N	Min	Max	M	SD	$ES_{prä}$	ES_{pool}
DISYPS-II Eltern prä	0-3	67	0.6	2.7	1.55	0.50		
DISYPS-II Eltern post	0-3	65	0.2	2.2	1.18	0.52	-0,74	-0,73
DISYPS-II Lehrer prä	0-3	57	0.0	3.0	1.16	0.62		
DISYPS-II Lehrer post	0-3	55	0.0	2.3	0.87	0.52	-0,47	-0,51
DISYPS-II Selbst prä	0-3	67	0.3	2.4	1.13	0.59		
DISYPS-II Selbst post	0-3	66	0.3	2.1	0.94	0.46	-0,32	-0,36
WRI prä	0-56	67	2	47	24.57	10.35		
WRI post	0-56	67	2	46	20.39	9.50	-0,40	-0,42
WLST prä	0-80	67	32	75	55.28	11.13		
WLST post	0-80	67	32	76	58.63	9.75	0,30	0,32
SDQ Eltern prä	0-40	67	7	31	19.60	5.60		
SDQ Eltern post	0-40	66	4	29	15.41	6.30	-0,75	-0,70
SDQ Lehrer prä	0-40	57	1	34	13.37	6.13		
SDQ Lehrer post	0-40	55	0	26	10.93	5.83	-0,40	-0,41
SDQ Selbst prä	0-40	67	4	29	16.00	6.31		
SDQ Selbst post	0-40	65	2	28	11.54	5.35	-0,71	-0,76
TAP Alertness RZ prä	0-	67	206	425	254.13	40.58		
TAP Alertness RZ post	0-	66	203	431	248.50	30.14	-0,14	-0,16
TAP Flexibilität RZ prä	0-	67	353	1945	848.42	295.74		
TAP Flexibilität RZ post	0-	66	463	1348	722.85	191.60	-0,42	-0,50
TAP Flexibilität FA prä	0-	67	1	35	8.25	7.18		
TAP Flexibilität FA post	0-	66	0	29	6.29	5.72	-0,27	-0,30
TAP Go/NoGo RZ prä	0-	67	225	582	402.06	97.58		
TAP Go/NoGo RZ post	0-	66	219	590	397.64	80.82	-0,05	-0,05

TAP Go/NoGo FA prä	0-	67	0	16	4.85	3.82		
TAP Go/NoGo FA post	0-	66	0	16	4.83	3.37	-0,01	-0,01
AVL prä	0-100	57	10	100	63.96	18.56		
AVL post	0-100	55	23	100	71.87	17.828	0,43	0,43

Anmerkungen: prä = Prämesszeitpunkt, post = Postmesszeitpunkt, RZ = Reaktionszeit, FA = Fehleranzahl. Weitere Erläuterungen im Fließtext

Anhand der deskriptiven Darstellung der Daten wird deutlich, dass sich in allen Erfolgsvariablen Verbesserungen in Form von Mittelwertänderungen in die erwünschte Richtung abbilden. Weiterhin zeigt sich eine Varianzreduktion für alle Variablen außer den beiden Elterneinschätzungen (DISYPS-II und SDQ). Die größten Verbesserungen nach den gepoolten Effektstärken wurden in den Einschätzungen der Jugendlichen und ihrer Eltern bezogen auf die Einschätzung der Gesamtproblematik und der ADHS-Symptomschwere erzielt. Die Einschätzungen der Lehrer fallen etwas geringer aus, auch die Ergebnisse im TAP-Test Flexibilität. Die kleinsten Verbesserungen zeigten sich für die TAP-Tests Alertness und Go/NoGo.

In drei Einzelmessungen war ein Decken- bzw. Bodeneffekt festzustellen. In einer Lehrereinschätzung wurde eine ADHS-Symptomschwere (DISYPS-II) von 0 (= „nicht vorhanden“) vor dem Training angegeben. In einer Lehrereinschätzung des Lern- und Arbeitsverhaltens (AVL) wurde der beste Wert von 100 % vor dem Training angegeben. Ein Jugendlicher erreichte in der TAP bei der Aufgabe Go/NoGo vor dem Training eine Fehlerzahl von 0. In diesen drei Einzelmessungen konnten keine Verbesserungen erzielt werden, da die optimalen Werte schon vor dem Training erreicht wurden. Da es sich nur um 3 von insgesamt 908 Einzelmessungen handelt, wird die Veränderungsmessung wie geplant durchgeführt. Bei der folgenden Bestimmung von Gesamterfolgswerten werden mehrere Einzelmessungen zusammengefasst, so dass schließlich keine Decken- und Bodeneffekte mehr vorliegen werden.

Die Effektstärken nach der gepoolten Methode unterscheiden sich von denen nach der prä-Methode nur in ihren absoluten Werten. Die Korrelation beträgt $r = 0.996$ ($p < .01$). Die tendenziell etwas höheren Werte nach der gepoolten Methode resultieren unter anderem daraus, dass Varianzreduktionen von prä nach post aufgetreten sind, die durch das gepoolte Verfahren „belohnt“ werden (Maier-Riehle & Zwingmann, 2000).

An dieser Stelle geht es nicht darum, nachzuweisen, dass die gefundenen Effekte auf das Lerntraining zurückzuführen sind. Diese Fragestellung wurde in der vorangegangenen allgemeinen Wirksamkeitsstudie überprüft. Daher werden hier keine Kontrollgruppen oder kritischen Werte zum Vergleich herangezogen. Das Ziel dieser Analyse ist der Vergleich der unterschiedlichen Messinstrumente und Konstrukte untereinander. Da die angegebenen Effektstärken aus einem Eingruppen-prä-post-Design stammen, ist es nicht unproblematisch, sie mit solchen zu vergleichen, die aus einem Kontrollgruppenvergleich stammen. Sie sind jedoch dazu geeignet, die Veränderung der verschiedenen Messinstrumente und der dahinter liegenden Erfolgskriterien zu vergleichen.

Es ist je nach Ebene des Erfolges mit unterschiedlich schnellen Veränderungen zu rechnen. Erkenntnisse aus der Psychotherapieforschung legen nahe, dass mit einer schnellen Verbesserung des psychischen Wohlbefindens zu rechnen ist, gefolgt von einer Symptomreduktion und erst langsamen Veränderung der Anpassung und Funktionsfähigkeit (Hartmann & Herzog, 1995). Erwartungsgemäß zeigt sich für das Lerntraining die Tendenz, dass sich allgemeine Problemmaße, über die sich vermutlich auch Zufriedenheit und allgemeines Wohlbefinden äußern, schneller verbessern als Symptommaße und diese wiederum schneller als funktionale Maße wie Konzentrationsaufgaben. Positiv anzumerken sind die relativ starken Verbesserungen in der Lehrereinschätzung, die sich auch explizit auf das Lern- und Unterrichtsverhalten beziehen, also funktionale Messungen von zentralen Trainingszielen darstellen.

5.3.2 Fehlende Werte

Um fehlende Werte zu vermeiden, wurde in einem ersten Schritt eine einfache Imputation durchgeführt, bei der alle drei Messzeitpunkte der allgemeinen Wirksamkeitsprüfung genutzt wurden. Diese wurden als Längsschnittdaten einer Zeitreihe interpretiert, so dass ein fehlender Wert durch den vorausgehenden oder nachfolgenden Wert derselben Person ersetzt werden kann. Für Teilnehmer der Interventionsgruppe mit fehlenden Werten direkt nach dem Training, wurde überprüft, ob in dieser Variablen Follow-up-Daten verfügbar waren. Falls vorhanden wurden die Follow-up-Daten von ca. einem halben Jahr nach Trainingsende im Sinne eines „next observation carried backward“ (NOCB; Engels & Diehr, 2003) als Wert direkt nach dem Training gesetzt. Der Follow-up-Wert wird als bester vorhandener Schätzer für den nicht vorhandenen Wert des Postmesszeitpunkts angenommen. Das Verfahren wurde für eine Versuchsperson angewendet, bei der die fehlende Lehrereinschätzung zu t_1 (SDQ, DISYPS-II und AVL) durch die vorliegende Einschätzung zu t_2 ersetzt wurde.

Bei Teilnehmern der Wartegruppe der allgemeinen Wirksamkeitsprüfung, die im Anschluss an die Wartezeit das Training durchlaufen haben, wurde bei fehlenden Werten direkt vor dem Training überprüft, ob Werte der fehlenden Variablen vor der Wartezeit vorlagen. Falls vorhanden wurden diese Werte im Sinne eines „last observation carried forward“ (LOCF; Engels & Diehr, 2003) übertragen. Damit wurden die fehlenden Werte vor der Warteperiode als Schätzer für den Messzeitpunkt direkt vor dem Training genommen. Dieses Vorgehen wurde für zwei Versuchspersonen durchgeführt, einmal wurde eine fehlende Lehrereinschätzung (SDQ, DISYPS-II und AVL) ersetzt und einmal fehlende Werte des WRI.

Die Übernahme von Daten von anderen Messzeitpunkten in diesem Sinne lässt sich inhaltlich mit den Ergebnissen der allgemeinen Wirksamkeitsprüfung rechtfertigen. Einerseits konnte gezeigt werden, dass sich in der Wartezeit keine bedeutsamen Veränderungen ergaben, so dass die Werte vor der Warteperiode als gute Schätzungen der Werte nach der Wartezeit herangezogen werden können. Andererseits zeigte sich eine Stabilität der Ergebnisse in der Interventionsgruppe von Trainingsabschluss zum

Follow-up, so dass die Werte nach dem Training gute Schätzungen der Werte ein halbes Jahr später sind (Schramm, im Druck).

Schwieriger zu entscheiden ist, wie mit fehlenden Daten umgegangen werden soll, die nicht wie zuvor beschrieben ausgeglichen werden können. Insbesondere dann, wenn nur Daten direkt vor der Intervention vorliegen, aber keine Daten danach. Engels und Diehr (2003) beschreiben unterschiedliche Imputationsverfahren. Für den Zweck der vorliegenden Auswertung ist es nicht sinnvoll, eine Imputation durchzuführen, die auf Populationskennwerten beruht, z. B. indem der Mittelwert aller Teilnehmer der jeweiligen Variable für fehlende Werte gesetzt wird. Ein solches Vorgehen würde die Verrechnung von Prä- und Postwerten auf Individuumsebene stark verzerren. Weiterhin ist es nicht sinnvoll auf Imputationen wie das „hot deck“ oder „Nächster-Nachbar-Prozeduren“ zurückzugreifen, bei denen ähnliche Versuchspersonen mit kompletten Datensätzen gesucht werden, deren Daten übernommen werden (Longford, 2005). Andere Ansätze benötigen keine Vergleichspopulation, z. B. Regressionsmethoden. Fehlende Werte werden dabei aus verschiedenen Variablen, die im Voraus vorliegen, geschätzt (Engels & Diehr, 2003). Zu diesem Vorgehen müssen allerdings detaillierte Annahmen dazu vorliegen, wie der geschätzte Wert vorhergesagt werden kann. Für die vorliegende Untersuchung ist aber unklar, wie solche Zusammenhänge aussehen und das Ziel ist genau die Analyse und Identifikation solcher Zusammenhänge. Somit sind auch Regressionsmethoden nicht angemessen.

Es wird daher im Folgenden auf eine Ersetzung fehlender Werte verzichtet. Die Auswertungen erfolgten jeweils nur für vollständige Datenpaare mit vorliegendem Wert zum Prämesszeitpunkt und zum Postmesszeitpunkt für die jeweilige Analyse. Es wird also ein paarweiser Fallausschluss vorgenommen. Dieses Vorgehen hat den Vorteil, dass möglichst viele Fälle für jede einzelne Analyse beibehalten werden. Somit können möglichst viele Informationen aus den vorliegenden Daten gewonnen werden und der als nicht sehr groß zu bezeichnende Datensatz wird nicht verkleinert. Allerdings resultiert eine unterschiedliche Stichprobengröße bei den unterschiedlichen Auswertungen. Andere Verfahren hätten aber stärkere Nachteile. So z. B. der listenweise Fallausschluss, bei dem nur absolut vollständige Fälle in die Analysen eingehen, also ohne einen einzigen fehlenden Wert in einer einzigen Variablen. Dadurch verkleinert sich der Datensatz sehr stark, wenn es sich um empirische Untersuchungen in praxisnahen Feldern handelt, in denen mit fehlenden Werten zu rechnen ist. Ersetzungsverfahren, z. B. durch den Mittelwert oder Median einer Verteilung, haben den Nachteil, dass die Variabilität in den Daten reduziert wird und Kovarianz- und Korrelationsschätzungen schlechter werden (Hawthorne & Elliott, 2005; Schafer & Graham, 2002).

Die unterschiedlich großen Fallzahlen der Auswertungsschritte resultieren dementsprechend aus fehlenden Werten in einzelnen Variablen. Vorausgreifend soll festgestellt werden, dass zur weiteren Auswertung einzelne Messinstrumente zu Faktoren zusammengefasst werden und daraus auch ein Gesamterfolgsmaß bestimmt wird. Dies führt

dazu, dass fehlende Werte durch andere Variablen innerhalb der Faktoren ausgeglichen werden können, so dass einzelne fehlende Werte nicht gravierend ins Gewicht fallen.

Anzahl und Prozent der fehlenden Werte zum Messzeitpunkt vor und nach dem Training sind in Tabelle 12 abgebildet. Bei den fehlenden Lehrereinschätzungen fehlen jeweils alle drei Lehrer-Outcomemaße für einen Fall.

Tabelle 12: Fehlende Werte der Outcomemaße vor und nach dem Training nach der ersten einfachen Imputation

	Prämesszeitpunkt		Postmesszeitpunkt	
	Fehlend		Fehlend	
N = 67	Anzahl	Prozent	Anzahl	Prozent
DISYPS-II Eltern	0	0	2	3.0
DISYPS-II Lehrer	10	14.9	12	17.9
DISYPS-II Selbst	0	0	1	1.5
SDQ Eltern	0	0	1	1.5
SDQ Lehrer	10	14.9	12	17.9
SDQ Selbst	0	0	2	3.0
WLST	0	0	0	0
WRI	0	0	0	0
AVL	10	14.9	12	17.9
TAP Alertness RZ	0	0	1	1.5
TAP Flexibilität RZ	0	0	1	1.5
TAP Flexibilität FA	0	0	1	1.5
TAP Go/NoGo RZ	0	0	1	1.5
TAP Go/NoGo FA	0	0	1	1.5
Gesamt	30	3.2	47	5.0

Beim Prämesszeitpunkt fehlen knapp 15 % der Lehrereinschätzungen. Alle anderen Werte liegen vor. Der Prozentsatz fehlender Werte zum Postmesszeitpunkt ist mit knapp 18 % für die Lehrereinschätzung am höchsten. Insgesamt wurden 14 Outcomemaße von 67 Teilnehmern erhoben, also sollten insgesamt 938 Einzelmessungen vorliegen. Insgesamt fehlen zum Prämesszeitpunkt 3.2 % der Werte und zum Postmesszeitpunkt 5.0 % aller Werte. Diese Prozentsätze sind als sehr gering zu bezeichnen. In vergleichbaren Studien liegen die Raten fehlender Werte durchschnittlich über 30 % (McKnight, McKnight, Sidani & Figueredo, 2007).

Es zeigen sich in den übrigen Problemmaßen keine signifikanten Unterschiede zwischen den Teilnehmern mit vorliegender und mit fehlender Lehrereinschätzung (siehe Anhang). Es liegt also kein Grund zur Annahme vor, dass die Werte selektiv fehlen, z. B. weil Lehrer aufgrund einer besonders gravierenden Problematik nicht bereit waren, an der Evaluation der Maßnahme teilzunehmen. Daher wird davon ausgegangen, dass die fehlenden Einschätzungen vor allem auf mangelnde zeitliche Ressourcen der Lehrer

zurückzuführen sind. Dies entspricht auch den Rückmeldungen bei Nachfrage nach fehlenden Fragebögen.

5.3.3 Bestimmung eines Gesamtproblemmaßes

Analog zur Auswertung der MTA-Studie nach Conners et al. (2001) wird ein zusammengesetztes Gesamtmaß bestimmt. Zunächst wird eine explorative Faktorenanalyse auf Grundlage der Prämessung der 14 Erfolgsmaße (siehe Tabelle 11) durchgeführt.

Die Kriterien nach Backhaus, Erichson, Plinke und Weiber (2011) zur Durchführung einer Faktorenanalyse werden erfüllt: Intervallskalierung und gute Reliabilitäten können für die genutzten standardisierten Skalen und Tests als gegeben angenommen werden (Bühner, 2011). Die Fallzahl beträgt $N = 67$ und ist somit als ausreichend zu bezeichnen.

Die Daten werden mittels Kaiser-Meyer-Olkin-Koeffizienten (KMO), Measure of Sample Adequacy-Koeffizienten (MSA) und dem Bartlett-Test auf ihre Eignung überprüft. Damit Faktoren bestimmt werden können, müssen substantielle Korrelationen vorhanden sein. Dazu wird der Bartlett-Test herangezogen, der überprüft, ob die Korrelationen der Korrelationsmatrix signifikant von null verschieden sind. Die Prüfgröße Chi-Quadrat beträgt für die vorliegenden Daten $\chi^2 = 333.22$ bei 91 Freiheitsgraden und ist damit signifikant ($p < 0.001$). Es liegen also substantielle Korrelationen vor, die zur Ermittlung von Faktoren erforderlich sind. Der KMO-Koeffizient ist mit 0.68 nach Bühner (2011) als „mäßig“ zu bezeichnen (ab 0.70 „mittel“). Die MSA-Koeffizienten liegen im annehmbaren mittleren Bereich. Nur für ein Item (Fehleranzahl im TAP-Test Go/NoGo) liegt der Wert mit 0.47 knapp unter der von Bühner (2011) vorgeschlagenen unteren Grenze von 0.50. Da das Item als inhaltlich sehr relevant eingeschätzt wird, verbleibt es in der weiteren Auswertung. Auch Bühner (2011) spricht sich dagegen aus, Items lediglich aufgrund statistischer Kennwerte zu entfernen. Die Daten sind damit zur weiteren faktorenanalytischen Auswertung geeignet.

Es wird zunächst eine Hauptkomponentenanalyse durchgeführt und mit den Ergebnissen einer Hauptachsen- und Maximum-Likelihood-Methode verglichen. Als Kriterien zur angemessenen Anzahl der Faktoren werden das Kaiser-Kriterium und der Scree-Test herangezogen. Gemäß Kaiser-Kriterium werden nur Faktoren mit Eigenwerten > 1 ausgewählt. Nach dem Scree-Test werden die Eigenwerte in absteigender Reihenfolge abgebildet und dann überprüft, an welcher Stelle ein „Knick“ vorliegt, d.h. ab wann sich die Eigenwerte asymptotisch der Abszissen-Achse annähern (Backhaus, Erichson, Plinke, et al., 2011). Abbildung 24 zeigt den Screeplot zur vorliegenden Auswertung. Übereinstimmend legen Kaiser-Kriterium und Scree-Test eine Lösung mit vier Faktoren nahe, da vier Faktoren Eigenwerte > 1 aufweisen und vier Faktoren vor dem „Knick“ liegen.

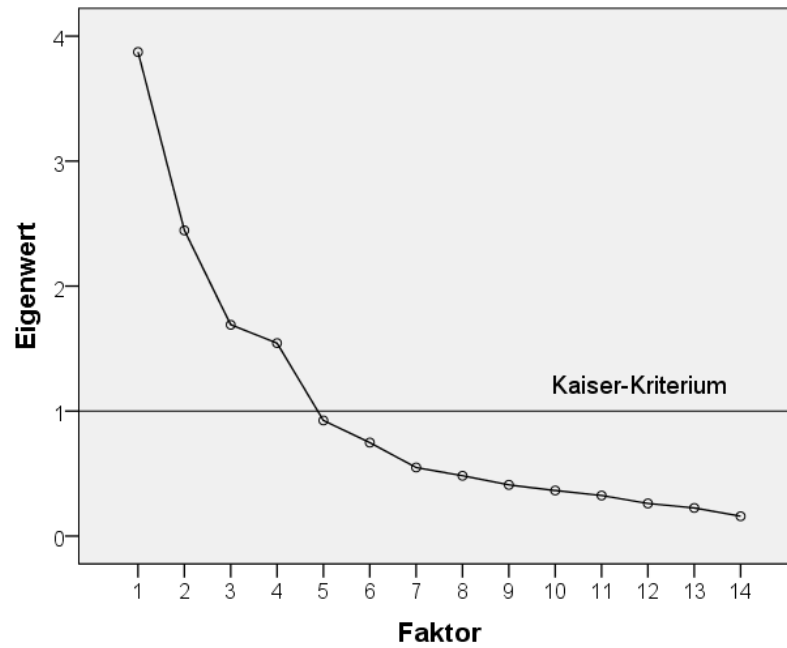


Abbildung 24: Screeplot der Outcomemaße vor Trainingsbeginn

Es wird zunächst eine Hauptkomponentenanalyse mit Promax-Rotation, dann zum Vergleich mit einer Varimax-Rotation durchgeführt. Die Ergebnisse beider Analysen sind in Tabelle 13 dargestellt. Wie bei zu Connors et al. (2001) werden kleine Ladungen (< 0.40) unterdrückt und ein paarweiser Fallausschluss vorgenommen.

Tabelle 13: Hauptkomponentenanalyse (Promax- und Varimax-rotiert) der Outcome-Maße zum Zeitpunkt vor Trainingsbeginn ($N = 67$ bzw. 57 für die Lehrerauskünfte)

Outcome-Variable	Strukturmatrix Hauptkomponentenanalyse mit Promax-Rotation				Komponentenmatrix Hauptkomponentenanalyse mit Varimax-Rotation			
	Faktor				Faktor			
	1	2	3	4	1	2	3	4
WRI	,788				,817			
DISYPS-II Selbst	,774				,764			
SDQ Selbst	,771				,758			
SDQ Eltern	,749	,444			,705			
DISYPS-II Eltern	,737	,532			,677			
AVL		-,869				-,884		
DISYPS-II Lehrer		,867				,832		
SDQ Lehrer	,400	,829				,786		
TAP Go/NoGo RZ			,815				,811	
TAP Alertness RZ			,806				,806	
TAP Flexibilität RZ			,791				,789	
TAP Flexibilität FA				,704				,707
TAP Go/NoGo FA			-,568	,659			-,573	,644
WLST				-,652				-,654

Für die Ergebnisse der Promax-Rotation werden folgende Korrelation der Komponenten bestimmt: Am höchsten korrelieren Faktor 1 und Faktor 2 mit $r = 0.39$. Faktor 2 und Faktor 3 korrelieren mit $r = 0.12$. Die übrigen Komponenten korrelieren nur sehr gering ($r < 0.03$). Für die Ergebnisse der Varimax-Rotation wird der Anteil der erklärten Gesamtvarianz bestimmt: Faktor 1 erklärt 21.25 %, Faktor 2 erklärt 17.60 %, Faktor 3 erklärt 16.92 % und Faktor 4 erklärt 12.46 % der Gesamtvarianz. Insgesamt werden durch die vier Faktoren 68.23 % der Gesamtvarianz erklärt. Es werden zum Vergleich auch Hauptachsen- und Maximum-Likelihood-Analysen mit Varimax- und Oblimin-Rotation durchgeführt. Es ergibt sich dieselbe Faktorstruktur mit denselben Itemzuordnungen.

Es werden keine Items mit geringen oder uneindeutigen Ladungen eliminiert, da alle erhobenen Maße inhaltlich wichtig erscheinen. Die Items werden dem Faktor zugeordnet, auf den sie am höchsten laden. Somit ergeben sich aus den Daten vier Erfolgsfaktoren, die auch inhaltlich stimmig erscheinen. Der erste Faktor, der aus fünf Items besteht, beschreibt die ADHS-Symptomatik und die allgemeine Problembelastung sowohl aus Sicht der Eltern als auch aus der Sicht der Jugendlichen selbst. Der zweite Faktor besteht aus drei Items und beinhaltet alle Lehrereinschätzungen, sowohl zur ADHS-Symptomatik als auch zur allgemeinen Problembelastung und zum Lern- und Arbeitsverhalten. In den dritten Faktor mit drei Items gehen die durchschnittlichen Reaktionszeiten der TAP ein. In den vierten Faktor mit ebenfalls drei Items gehen die Fehleran-

zahlen der TAP-Tests Flexibilität und Go/NoGo und der Gesamtwert des WLST ein. Die negativen Vorzeichen der Ladungen im AVL und WLST ergeben sich daraus, dass diese beiden Variablen Kompetenzen abbilden, in denen positive Werte erwünscht sind, während die übrigen Messungen das Ausmaß von Schwierigkeiten und Problemen erfassen und somit möglichst niedrig bzw. negativ ausgeprägt sein sollen.

Auffällig ist, dass die Lehrereinschätzung einen eigenen Faktor bildet und sich nicht die erwarteten Outcomevariablen ADHS-Symptomatik und Allgemeine Problembelastung abzeichnen, auf die jeweils Eltern-, Jugendlichen- und Lehrereinschätzung laden. Schwache Zusammenhänge in der Einschätzung von Eltern und Lehrern bezogen auf Verhaltens- und emotionale Probleme wurden auch in Metaanalysen gefunden, z. B. bei Achenbach, McConaughy und Howell (1987), die in 119 Studien einen schwachen signifikanten Zusammenhang der beiden Auskunftquellen von $r = 0.28$ feststellten. Bezogen auf ADHS stellen Willcutt et al. (2012) eine Übereinstimmung von $r = .43$ für Symptome der Unaufmerksamkeit und von $r = .42$ für hyperaktiv-impulsive Symptome fest. Diese Feststellung entspricht dem Bild der in den vorliegenden Daten gefundenen zwei eigenständigen Faktoren 1 und 2, die nach der obliquen Promax-Rotation zu $r = 0.39$ korrelieren. Die Unterschiede der Eltern- und Lehrereinschätzung seien nach Willcutt und Kollegen zumindest teilweise dadurch zu erklären, dass die Kinder sich in der Schule und zuhause tatsächlich unterschiedlich verhalten. Beide Informationsquellen scheinen jeweils einzeln unterschiedliche Informationen mit klinischer Relevanz zu liefern (Hart, Lahey, Loeber & Hanson, 1994).

Um mit Faktorwerten weiter rechnen zu können, wurden mit SPSS 20 Bartlett-Schätzer für die Varimax- und die Promax-Rotation bestimmt. Die Werte nach beiden Verfahren korrelieren erwartungsgemäß sehr hoch miteinander ($r \geq 0.98$). Weiterhin wurden manuell Faktorwerte analog zu Connors (2001) ermittelt. Dazu wurden die Item-Zuordnungen der durchgeführten Faktoranalysen genutzt. Alle Items wurden zunächst anhand ihrer Streuungen und Mittelwerte z -standardisiert. Die Werte des AVL und WLST wurden aufgrund der negativen Ladungen und inhaltlich gegensätzlichen Bedeutung umgepolt bzw. rekodiert, damit eine Reduktion von Werten stets eine Verbesserung ausdrückt.

In Faktor 1 gehen die ersten fünf Items der Faktorenanalyse ein. Er wird mit „Symptomatik“ bezeichnet. In Faktor 2, der als „Lehrereinschätzung“ benannt wird, gehen die Items 6 bis 8 ein. Die Items 9 bis 11 werden als Faktor 3 unter der Bezeichnung „Reaktionszeit“ zusammengefasst. Die letzten vier Items gehen in Faktor 4 ein, der als „Fehler“ bezeichnet wird. Anschließend werden die Mittelwerte über die standardisierten Items für jeden der vier Faktoren errechnet. Wenn in einem Fall ein Wert fehlt, so wird der Mittelwert über die vorhandenen anderen Variablen des jeweiligen Faktors errechnet. Fehlende Werte können somit ausgeglichen werden.

Die so errechneten Faktorwerte korrelieren ebenfalls sehr hoch mit den zuvor bestimmten Bartlett-Schätzern ($r = 0.95-0.99$). Aufgrund der hohen Korrelationen wird nur die

manuelle Variante fortgeführt. Dieses Vorgehen ist einfach nachzuvollziehen und ermöglicht eine analoge Bestimmung der Werte zum Postmesszeitpunkt. Die Werte des Postmesszeitpunkts werden anhand der Mittelwerte und Standardabweichungen der vom Training unbetroffenen Werte des Prämesszeitpunktes z -standardisiert. Es fehlen in einem Fall zur Berechnung des Nachherwerts des Faktors 1 ein Eltern- und ein Selbsteinschätzungsfragebogen. Der Faktorwert wird aus den vorliegenden drei anderen Fragebögen, die in diesen Faktor eingehen, bestimmt. In vier Fällen fehlt nach dem Training einer der drei Lehrerfragebögen zur Berechnung von Faktor 2. In einem Fall liegt nur einer von drei Lehrerfragebögen vor. In diesen Fällen wird der Faktorwert aus den vorliegenden zwei bzw. dem vorliegenden einen Fragebogen ermittelt.

Die vier ermittelten Faktoren können als abhängige Variablen zur Überprüfung allgemeiner und differenzieller Trainingseffekte genutzt werden. Auch wenn aufgrund der geringen Korrelationen der Faktoren untereinander nicht davon auszugehen ist, dass sie gleichmäßig variieren, wird zusätzlich ein übergeordnetes, globales Gesamterfolgsmaß bestimmt. Dazu wird für jeden Teilnehmer der Mittelwert der vier Erfolgsfaktoren bestimmt. Die Faktoren sind damit im Gesamtmaß gleichwertig zu je einem Viertel vertreten, unabhängig davon, wie viele Items in den jeweiligen Faktor eingehen. Dieses Vorgehen trägt der Überlegung Rechnung, dass einige Variablen dasselbe Konstrukt zu messen scheinen. Bei der Auswahl der Messinstrumente wurden viele Verfahren ausgewählt, welche die Symptomatik (Faktor 1) abbilden. Würde der Mittelwert direkt über alle Items errechnet, so erhielte die Symptomatik ein zu großes Gewicht in der Gesamteinschätzung.

Bei fehlenden Werten für einen Faktor wird das Gesamterfolgsmaß aus den übrigen vorhandenen Faktoren bestimmt. Zum Prämesszeitpunkt fehlt zehnmal das Lehrerurteil, zum Postmesszeitpunkt fehlt ebenfalls zehnmal das Lehrerurteil, einmal die Einschätzung der Symptome und in einem Fall sowohl Faktor 3 als auch Faktor 4. Die Bestimmung des Gesamterfolgsmaßes mit Ausgleich für fehlende Werte hat den Vorteil, dass für jeden der 67 Fälle ein Erfolgsmaß bestimmt werden kann, so dass alle Fälle miteinander verglichen werden können. Das Gesamtmaß wird im Folgenden als „Gesamtproblematik“ bezeichnet.

Die Verteilungsparameter der Faktoren und der Gesamtproblematik sind in Tabelle 14 dargestellt. Aufgrund der z -Standardisierung sind alle Mittelwerte zum Prämesszeitpunkt $M = 0$.

Tabelle 14: Verteilungsparameter der vier Faktoren bzw. Problemaße und des Gesamtproblemaßes

Faktoren/ Problemaße	N	Min	Max	M	SD
Symptomatik prä	67	-1.30	1.48	0	0.77
Symptomatik post	66	-1.99	1.23	-0.58	0.75
Lehrereinschätzung prä	57	-1.95	2.44	0	0.87
Lehrereinschätzung post	57	-2.00	2.04	-0.43	0.83
Reaktionszeit prä	67	-0.98	3.25	0	0.83
Reaktionszeit post	66	-1.11	2.25	-0.20	0.64
Fehler prä	67	-1.10	1.96	0	0.70
Fehler post	66	-1.05	1.94	-0.20	0.58
Gesamtproblematik prä	67	-0.99	1.26	0	0.46
Gesamtproblematik post	67	-1.28	0.91	-0.34	0.42

5.3.4 Bestimmung von Veränderungsmaßen

Zur weiteren Auswertung werden nicht nur Prä- und Postwerte benötigt, sondern diese sollen auch zu Veränderungsmaßen verrechnet werden, um eine Variable bestimmen zu können, welche das Ausmaß des Trainingserfolges ausdrückt. Prinzipiell kommen drei Verrechnungsmöglichkeiten bzw. Veränderungsmaße in Frage:

- *Differenzwerte* durch Abzug des Wertes der Prämessung vom Wert der Postmessung als Maß der absoluten individuellen Veränderung,
- *Prozentuale Veränderungen* als Maß der relativen Veränderung im Verhältnis zum individuellen Wert der Prämessung,
- *Residuale Zuwächse* als Maß der individuellen Veränderung in Relation zur Veränderung anderer Teilnehmer mit ähnlichen Werten zum Prämesszeitpunkt.

Die drei Veränderungsmaße haben unterschiedliche Vor- und Nachteile. So ist es für Teilnehmer mit initial eher geringen Problemwerten schwierig, große Veränderungen in absoluten Differenzen zu erreichen. Für diese Teilnehmer sind prozentuale Veränderungswerte besser geeignet, um Veränderungen auszudrücken. Teilnehmern mit anfänglich hohen Problemwerten können sich prozentual aber nur wenig verbessern. Für sie sind absolute Differenzwerte besser geeignet. Die residualen Zuwachswerte berücksichtigen darüber hinaus Veränderungen anderer Teilnehmer und drücken die individuellen Veränderungen in Bezug dazu aus (Steketee & Chambless, 1992).

Veränderungsmaße können nur bestimmt werden, wenn für einen Fall beide Werte vom Prä- und Postmesszeitpunkt vorliegen. Deshalb kommt es zu einigen fehlenden Werten, insbesondere in der Lehrereinschätzung. Es können für 52 Fälle (78 %) Veränderungsmaße der Lehrereinschätzung bestimmt werden. Fünf Lehrereinschätzungen fehlen

komplett und in zehn Fällen liegen entweder nur die Werte des Prämesszeitpunktes oder des Postmesszeitpunktes vor.

Die absoluten Differenzen werden durch Subtraktion des Vorherwertes vom Nachherwert ermittelt. Da es sich um Problemmaße handelt, ist eine Verringerung des Wertes wünschenswert. Die Variablen drücken die Veränderung als Differenz der Vorher- und Nachherwerte aus. Negative Werte bedeuten eine Verringerung der Problematik, positive Werte eine Verschlechterung.

Bei der Bestimmung prozentualer Veränderungen sind die Verteilungen der Variablen zu berücksichtigen. In der vorliegenden Auswertung verteilen sich die Problemfaktoren um den Mittelwert $M = 0$. Auf Grundlage eines Präwertes von 0 lässt sich keine prozentuale Veränderung bestimmen, da sie sich aus dem Quotienten der Differenz und des Wertes der Prämessung zusammensetzt. Bei sehr kleinen Werten um 0 ergeben sich darüber hinaus sehr große prozentuale Veränderungen, so dass der Trainingserfolg überschätzt wird. Aufgrund der vorliegenden Verteilung der Problemmaße ist die prozentuale Werteveränderung nicht zur Auswertung geeignet. Es daher werden keine prozentualen Veränderungswerte bestimmt.

Residuale Zuwächse (RZ) berücksichtigen im Gegensatz zu reinen Differenzwerten die Präwerte und die Vorher-nachher-Korrelation der Werte in der Stichprobe. Damit setzen sie die individuellen Werte in Relation zu den Veränderungen, die von anderen Teilnehmern des gleichen initialen Niveaus gemacht werden (Lincoln, et al., 2005). Prä- und Postwerte werden anhand der Gesamtgruppe („grand mean“) z -standardisiert, so dass sie einen Mittelwert von 0 und eine Standardabweichung von 1 haben. Anschließend wird der mit dem Wert der Vorher-nachher-Korrelation multiplizierte Prä-Wert vom Postwert abgezogen. Residuale Veränderungswerte sollen dadurch sowohl für initiale Veränderungen zwischen den Teilnehmern kontrollieren als auch für Messfehler, die bei der Durchführung wiederholter Messungen entstehen (Steketee & Chambless, 1992). Sie werden nach der folgenden Formel bestimmt: $RZ = Z_{post} - (Z_{prä} * r_{12})$.

Die Differenzwerte und residualen Zuwächse werden im Anschluss auf ihren Zusammenhang überprüft. Beide Maße korrelieren hoch mit $r = 0.89-0.97$ ($p = .00$). Da es sich bei beiden Maßen im Wesentlichen um redundante Informationen zu handeln scheint, werden für die weitere Auswertung lediglich die absoluten Differenzwerte genutzt. Diese sind die einfacher zu bestimmenden Maße und nachvollziehbar, weshalb sie bei gleichem Informationsgehalt vorgezogen werden.

In Tabelle 15 sind die Verteilungsparameter für die Differenzwerte abgebildet. Es lässt sich anhand der negativen Mittelwerte ablesen, dass sich für alle Problemmaße Verbesserungen ergeben haben. Die größten Verbesserung wurden im Maß Symptomatik erreicht, gefolgt von der Lehrereinschätzung.

Tabelle 15: Verteilungsparameter der Differenzwerte der vier Faktoren bzw. Problemmaße und des Gesamtproblemmaßes

Faktoren/ Problemmaße	N	Minimum	Maximum	M	SD
Symptomatik	66	-1.86	114	-0.58	0.57
Lehrereinschätzung	52	-3.24	1.29	-0.41	0.84
Reaktionszeit	66	-2.02	0.73	-0.21	0.42
Fehler	66	-1.62	1.30	-0.17	0.46
Gesamtproblematik	67	-1.44	0.44	-0.34	0.32

Die Differenzwerte der Problemmaße korrelieren untereinander kaum, sehr ähnlich den Vorherwerten. Die einzige signifikante Korrelation besteht zwischen der Symptomatik und der Lehrereinschätzung mit $r = 0.34$ ($p = .02$).

6 Ergebnisse

Die Darstellung der Ergebnisse folgt der Aufteilung der Fragestellung in fünf Unterfragen. Zum Abschluss des Kapitels wird eine weiterführende Auswertung dargestellt, die sich im Laufe der Analyse als sinnvoll herausgestellt hat.

6.1 Variablenzentrierte einfache univariate Auswertung

Fragestellung 1 a: *Wie hängen die Merkmale der Teilnehmer und der Trainingsdurchführung jeweils einzeln mit den verschiedenen Aspekten des Trainingserfolges zusammen?*

Bei der Vorstellung der Ergebnisse wird zunächst die Verteilung des interessierenden Merkmals vorgestellt. Dann erfolgt die Analyse der Problemmaße zum Prämesszeitpunkt. Bei dichotomen Merkmalen werden die Mittelwerte der Gruppen verglichen und Effektstärken für die Größe des Unterschieds angegeben. Bei metrischen Variablen werden Korrelationen zwischen den Problemmaßen und dem jeweiligen Merkmal bestimmt. Anschließend erfolgt die Darstellung der Differenzmaße als Maße des Trainingserfolges. Wurde eine Hypothese zum jeweiligen Merkmal aufgestellt, wird diese zunächst anhand des Gesamtproblemmaßes überprüft. Im Anschluss werden die übrigen Problemmaße analysiert. Dabei wird nach Bortz und Döring (2006) probeweise auf Signifikanz getestet, d. h. es werden jeweils p -Werte für die Signifikanz des zugehörigen Tests angegeben. Das Signifikanzniveau wird bei allen Tests wie konventionell üblich auf $\alpha = 5\%$ festgelegt.

Unterschiedliche Gruppengrößen resultieren aus fehlenden Werten. Höhere Werte in den Problemmaßen bedeuten eine stärkere bzw. gravierendere Problematik. Negative Werte in den Differenzwerten zeigen eine Verbesserung an.

Geschlecht

Es haben 13 Mädchen (19 %) und 54 Jungen (81 %) das Training durchlaufen.

Vergleich der Werte zum Prämesszeitpunkt

Tabelle 16 zeigt die Mittelwerte der Problemmaße für Mädchen und Jungen. Die Effektstärke des Unterschiedes zwischen Jungen und Mädchen wurde analog zu Cohens d (Cohen, 1988) anhand der gepoolten Streuung nach Hartmann, Herzog und Drinkmann (1992) bestimmt.

Tabelle 16: Vergleich der Problemmaße von Mädchen und Jungen zum Prämesszeitpunkt

Problemmaß	Geschlecht	<i>N</i>	<i>M</i>	<i>SD</i>	Mittelwert- differenz	<i>ES_{pool}</i>
Symptomatik	weiblich	13	0.31	0.81	0.39	0.50
	männlich	54	-0.07	0.75		
Lehrereinschätzung	weiblich	11	0.07	0.82	0.09	0.10
	männlich	46	-0.02	0.89		
Reaktionszeit	weiblich	13	0.12	1.10	0.15	0.18
	männlich	54	-0.03	0.76		
Fehler	weiblich	13	0.23	0.91	0.28	0.41
	männlich	54	-0.05	0.63		
Gesamtproblematik	weiblich	13	0.17	0.55	0.22	0.46
	männlich	54	-0.04	0.43		

Im Durchschnitt weisen die Mädchen vor dem Training schlechtere Werte in allen Merkmalen auf. Sie erreichen höhere Werte, d. h. ihre Problematik wird als schwerwiegender eingeschätzt und ihre Leistungen in den Tests fallen schlechter aus. In der Symptomatik liegen die Werte der Mädchen und Jungen um eine halbe Standardabweichung auseinander, in der Gesamtproblematik fast um eine halbe Standardabweichung.

Bei der weiteren Exploration der initialen Unterschiede wird in den Häufigkeitsverteilungen deutlich, dass von den vier Teilnehmern mit der höchsten Gesamtproblematik vor dem Training drei Mädchen sind. Keines der Mädchen hat sehr geringe Werte unter -0.5 (siehe Abbildung 25).

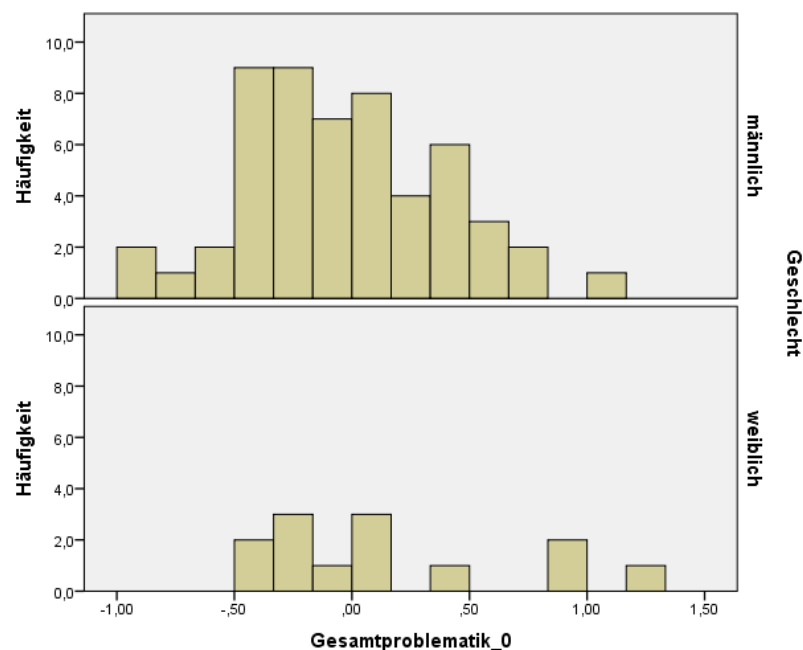


Abbildung 25: Vergleich der Gesamtproblematik bei Mädchen und Jungen vor Trainingsbeginn

Im Boxplot zeigt sich, dass die Mediane im Problemmaß Fehler gleich hoch liegen und dass es bei den Mädchen drei extreme Ausreißer mit besonders hohen Werten gibt (siehe Abbildung 26). Diese Ausreißer tragen wesentlich zu den Mittelwertunterschieden bei. Keines der Mädchen hat besonders schlechte Werte. Im Problemmaß Symptomatik zeigt sich bei nahezu identischen Spannweiten, dass der Median und das untere und das obere Quartil bei den Mädchen höher liegen (siehe Abbildung 26).

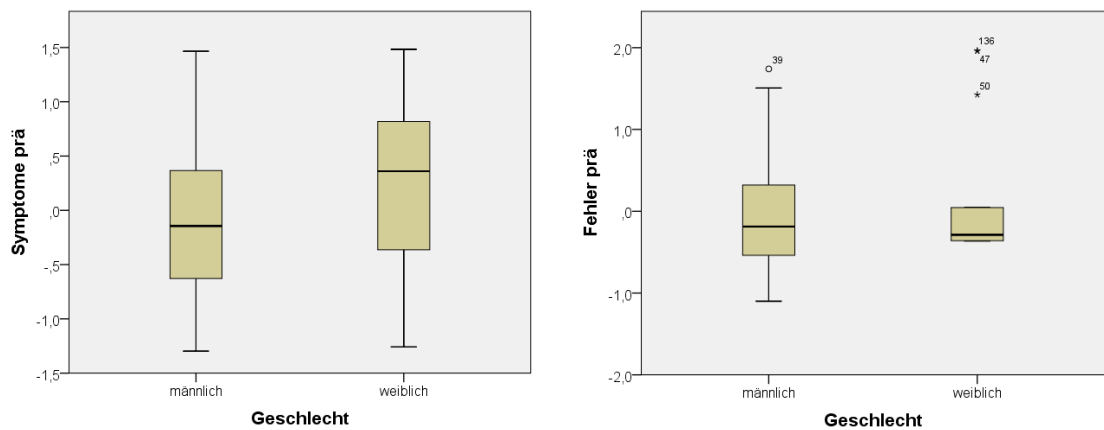


Abbildung 26: Vergleich der Problemmaße Symptomatik und Fehler vor Trainingsbeginn bei Mädchen und Jungen in Boxplots

Bei genauerer Analyse des Problemmaßes Symptomatik zeigt sich, dass die höheren Werte der Mädchen insbesondere auf die Selbstbeurteilungen zurückzuführen sind. Die Mädchen geben eine höhere Symptomatik und Problemschwere an als die Jungen. Die Unterschiede im Problemmaß Fehler lassen sich auf Unterschiede im TAP-Subtest Flexibilität zurückführen. Dort machen die drei Mädchen mit den extrem hohen Werten sehr viele Fehler. Die Unterschiede in der Gesamtproblematik sind also vornehmlich auf Unterschiede in den Faktoren Symptomatik und Fehler zurückzuführen. Die Unterschiede in den Maßen Lehrereinschätzung und Reaktionszeit fallen weniger stark ins Gewicht.

Um im Sinne des Signifikanztestens auf Probe nach Bortz und Döring (2006) zu überprüfen, ob die gefundenen Unterschiede in den absoluten Werten statistisch signifikant sind, wird zunächst überprüft, ob die Daten normalverteilt sind. Bei kleinen Stichproben, etwa $N < 50$, wird der Shapiro-Wilk-Test zur Überprüfung auf Normalverteilung empfohlen (Brosius, 2011; Yazici & Yolacan, 2007). Da es sich insbesondere bei den Mädchen um eine sehr kleine Gruppe handelt, wird der Shapiro-Wilk-Test durchgeführt.

In den Problemmaßen Reaktionszeit und Fehler liegen Verletzungen der Normalverteilungsannahme vor (siehe Anhang). Bei diesen Variablen sind aufgrund der kleinen Gruppengröße nicht-parametrische Tests, wie der Wilcoxon-Mann-Whitney-Test, solchen Tests vorzuziehen, die eine Normalverteilung voraussetzen, wie der t -Test (Bortz

& Schuster, 2010). Zur Überprüfung auf Signifikanz der Mittelwertunterschiede wurden für die normalverteilten Problemaße t -Tests und für die nicht normalverteilten Maße Mann-Whitney-Tests durchgeführt.

Es liegen keine signifikanten Mittelwertunterschiede vor:

Symptomatik: $t(65) = -1.65, p = .10$

Lehrereinschätzung: $t(55) = -0.29, p = .77$

Reaktionszeit: $U = 350, Z = -0.02, p = .99$

Fehler: $U = 295, Z = -0.89, p = .38$

Gesamtproblematik: $t(65) = -1.53, p = .13$

Vergleich der Differenzwerte

Für die Problemaße Symptomatik und Gesamtproblematik kann Normalverteilung angenommen werden, für die übrigen Problemaße nicht (siehe Anhang). Zum Vergleich der Differenzwerte von Mädchen und Jungen werden für die normalverteilten Variablen t -Tests für unabhängige Stichproben und für die nicht normalverteilten Variablen Mann-Whitney-U-Tests durchgeführt. Es liegen keine signifikanten Unterschiede in den Veränderungswerten vor. Die Effektgrößen zum Vergleich der Mittelwerte sind gering (siehe Tabelle 17).

Tabelle 17: Vergleich der Differenzwerte von Mädchen und Jungen

Problem- maß	Ge- schlecht	<i>N</i>	<i>M</i>	<i>SD</i>	Mittel- wert- differenz ^a	<i>ES_{pool}</i> ^a	<i>t</i> -Test	Mann- Whitney- U-Test
Symptomatik	weiblich	13	-0.52	0.64	0.07	0.12	$t(64) = -0.40$ $p = .69$	
	männlich	53	-0.59	0.56				
Lehrer- einschätzung	weiblich	9	-0.50	0.67	-0.12	-0.13		$U = 179$ $Z = -0.35$ $p = .73$
	männlich	43	-0.39	0.87				
Reaktionszeit	weiblich	12	-0.12	0.40	0.11	0.24		$U = 296$ $Z = -0.47$ $p = .64$
	männlich	54	-0.22	0.43				
Fehler	weiblich	12	-0.28	0.47	-0.13	-0.28		$U = 273,$ $Z = -0.85$ $p = .40$
	männlich	54	-0.15	0.46				
Gesamt- problematik	weiblich	13	-0.32	0.33	0.03	0.10	$t(65) = -0.28$ $p = .78$	
	männlich	54	-0.35	0.31				

Anmerkung: ^aDie Werte der Jungen wurden von denen der Mädchen abgezogen. Negative Werte bedeuten, dass die Mädchen stärkere Verbesserungen aufwiesen.

Es ist zusammenzufassen, dass sich die Ergebnisse von Mädchen und Jungen nicht signifikant unterscheiden. Die Mädchen haben vor dem Training leicht höhere Problemwerte und schlechtere Testergebnisse. Die Unterschiede sind insbesondere auf drei Ausreißerwerte im TAP-Test Flexibilität zurückzuführen und darauf, dass die Mädchen in der Selbstauskunft eine höhere Symptomatik angeben. Die Differenzwerte als Maß der Veränderung während des Trainings unterscheiden sich nur minimal und sind nicht statistisch signifikant.

Medikation

In der Stichprobe nehmen 37 Teilnehmer (55 %) Medikamente zur Behandlung Ihrer ADHS ein. Die übrigen 30 Teilnehmer (45 %) nehmen keine Medikamente ein. Da die Gruppengrößen mit mindestens $n = 30$ hinreichend groß sind (Kubinger, Rasch & Morder, 2009), wird keine Überprüfung auf Normalverteilung vorgenommen und der t -Test zum Vergleich der Mittelwerte durchgeführt.

Vergleich der Werte zum Prämesszeitpunkt

Tabelle 18 zeigt den Vergleich der Mittelwerte von Teilnehmern mit und ohne Medikation. Die Mittelwerte unterscheiden sich nicht signifikant voneinander. Die größten Unterschiede mit einer Effektgröße von fast einer halben Standardabweichung liegen in der Lehrereinschätzung vor. Jugendliche mit Medikation haben dabei die besseren Werte.

Tabelle 18: Vergleich der Werte von Teilnehmern mit und ohne Medikation zum Prämesszeitpunkt

Problemmaß	Medikation	<i>N</i>	<i>M</i>	<i>SD</i>	Mittelwertdifferenz ^a	<i>ES_{pool}</i> ^a	<i>t</i> -Test
Symptomatik	nein	37	0.01	0.70	0.03	0.03	$t(65) = 0.13, p = .90$
	ja	30	-0.01	0.85			
Lehrereinschätzung	nein	30	0.18	0.89	0.38	0.44	$t(55) = 1.69, p = .10$
	ja	27	-0.20	0.82			
Reaktionszeit	nein	37	0.06	0.78	0.14	0.17	$t(65) = 0.67, p = .51$
	ja	30	-0.08	0.89			
Fehler	nein	37	0.00	0.71	0.07	0.00	$t(65) = 0.04, p = .97$
	ja	30	0.00	0.70			
Gesamtproblematik	nein	37	0.05	0.48	0.12	0.26	$t(65) = 1.08, p = .29$
	ja	30	-0.07	0.43			

Anmerkung: ^aDie Werte der Teilnehmer mit Medikation wurden von denen ohne Medikation abgezogen.

Vergleich der Differenzwerte

Tabelle 19 zeigt den Vergleich der Differenzwerte der Teilnehmer mit und ohne Medikamenteneinnahme. Es liegen keine signifikanten Unterschiede vor.

Tabelle 19: Vergleich der Differenzwerte von Teilnehmern mit und ohne Medikation

Problemmaß	Medikation	N	M	SD	Mittelwertdifferenz ^a	ES _{pool} ^a	t-Test
Symptomatik	nein	37	-0.59	0.57	-0.03	-0.05	t(64) = -0.20, p = .85
	ja	29	-0.56	0.59			
Lehrereinschätzung	nein	28	-0.34	0.97	0.14	0.17	t(50) = 0.62, p = .54
	ja	24	-0.48	0.66			
Reaktionszeit	nein	36	-0.14	0.39	0.14	0.33	t(64) = 1.32, p = .19
	ja	30	-0.28	0.45			
Fehler	nein	36	-0.22	0.47	-0.10	-0.22	t(64) = -0.88, p = .39
	ja	30	-0.12	0.44			
Gesamtproblematik	nein	37	-0.32	0.36	0.05	0.16	t(65) = 0.61, p = .55
	ja	30	-0.37	0.25			

Anmerkung: ^aDie Werte der Teilnehmer mit Medikation wurden von denen ohne Medikation abgezogen. Negative Werte bedeuten, dass die Teilnehmer ohne Medikation stärkere Verbesserungen aufwiesen.

ADHS-Subtypus

Es wird keine Gruppenzuordnung nach ADHS-Subtypus vorgenommen, sondern die kontinuierliche Einschätzung der Symptome der Hyperaktivität und Impulsivität im DISYPS-II genutzt. Die Eltern-, Lehrer- und Selbsteinschätzungen werden jeweils separat ausgewertet.

Zusammenhang mit den Werten zum Prämesszeitpunkt

Tabelle 20 zeigt die Korrelationen der Hyperaktivität und Impulsivität vor dem Training mit den Problemmaßen zum Prämesszeitpunkt. Es wird Kendalls τ als Maß des Zusammenhangs bestimmt.

Tabelle 20: Zusammenhang der Hyperaktivität und Impulsivität in der Elterneinschätzung mit den Problemmaßen vor dem Training

		Symptomatik	Lehrereinschätzung	Reaktionszeit	Fehler	Gesamtproblematik
Hyperakt.-Impuls. Eltern	τ	.51	.30	.02	.06	.37
	p	.00	.00	.82	.47	.00
	N	67	57	67	67	67
Hyperakt.-Impuls.. Lehrer	τ	.23	.52	.09	.06	.37
	p	.01	.00	.33	.49	.00
	N	57	57	57	57	57
Hyperakt.-Impuls. Selbst	τ	.59	.25	-.04	.12	.42
	p	.00	.01	.67	.16	.00
	N	67	57	67	67	67

Alle Einschätzungen der Hyperaktivität und Impulsivität korrelieren signifikant positiv mit der Einschätzung der Symptomatik, der Lehrereinschätzung und der Gesamtproblematik. Weitere signifikante Korrelationen liegen nicht vor.

Zusammenhang mit den Differenzwerten

Tabelle 21 zeigt die Zusammenhänge der Hyperaktivität und Impulsivität mit den Differenzwerten.

Tabelle 21: Zusammenhang der Hyperaktivität und Impulsivität in der Elterneinschätzung vor dem Training mit den Differenzwerten

		Symptomatik	Lehrer- einschätzung	Reaktions- zeit	Fehler	Gesamt- problematik
Hyperakt.-Impuls. Eltern	τ	-.12	-.07	.01	-.07	-.08
	ρ	.17	.50	.92	.40	.33
	N	66	52	66	66	67
Hyperakt.-Impuls.. Lehrer	τ	.08	-.04	.02	.01	-.04
	ρ	.40	.69	.80	.96	.65
	N	56	52	56	56	57
Hyperakt.-Impuls. Selbst	τ	-.20	-.15	-.02	-.05	-.11
	ρ	.02	.13	.81	.56	.21
	N	66	52	66	66	67

Es zeigt sich eine signifikante negative Korrelation zwischen der Hyperaktivität und Impulsivität im Selbsturteil und der Differenz in der Symptomatik vor. Sie ist mit $\tau = -.20$ als gering zu bezeichnen. Schätzen die Jugendlichen ihre Hyperaktivität und Impulsivität vor dem Training also als höher ein, hatten sie tendenziell eine stärkere Reduktion der Symptomatik. Für die übrigen Variablen zeigen sich keine signifikanten Korrelationen.

Alter

Die Teilnehmer sind zwischen 144 und 213 Monaten alt. Das entspricht einem Alter zwischen 12;0 und 17;9 Jahren. Im Durchschnitt sind die Jugendlichen $M = 171$ Monate alt ($SD = 18$). Nach Shapiro-Wilk-Test kann keine Normalverteilung angenommen werden: $W(67) = .96$, $p = .03$. Die korrelativen Zusammenhänge werden deshalb mit Kendalls τ überprüft.

Die inhaltliche Hypothese, dass ältere Teilnehmer mehr vom Training profitieren und eine größere Reduktion des Gesamtproblemwertes aufweisen, wird anhand nachfolgender Hypothesen überprüft: Alternativhypothese Es gibt einen negativen Zusammenhang zwischen dem Alter und den Differenzwerten im Gesamtproblemmaß. Nullhypothese: Es gibt einen positiven oder keinen Zusammenhang zwischen dem Alter und den Differenzwerten im Gesamtproblemmaß.

$$H_1: \rho_{\text{Alter, VeränderungswertGesamtproblematik}} < 0$$

$$H_0: \rho_{\text{Alter, VeränderungswertGesamtproblematik}} \geq 0.$$

Zusammenhang mit den Werten zum Prämesszeitpunkt

Tabelle 22 zeigt die Zusammenhänge des Alters der Teilnehmer mit den Präwerten der Problemmaße. Da keine Normalverteilung gegeben ist, wird Kendalls τ als Maß des Zusammenhanges herangezogen.

Tabelle 22: Zusammenhang des Alters der Teilnehmer mit den Werten der Problemmaße zum Prämesszeitpunkt

		Symptomatik	Lehrer- einschätzung	Reaktions- zeit	Fehler	Gesamt- problematik
Alter	τ	.00	-.14	-.19	-.04	-.17
	ρ	.98	.14	.03	.65	.05
	N	67	57	67	67	67

Es zeigt sich eine signifikante negative Korrelation zwischen dem Alter und der Gesamtproblematik von $\tau = -.17$. Die älteren Teilnehmer haben demnach eine tendenziell weniger starke Problematik vor dem Training. Ein signifikanter Zusammenhang zeigt sich auch mit dem Problemmaß Reaktionszeit ($\tau = -.19$).

Zusammenhang mit den Differenzwerten

Zur Hypothesenprüfung wird Kendalls τ bestimmt und auf Signifikanz geprüft. Es ergibt sich für eine zweiseitige Signifikanz $\tau = .00$, $p = .99$ ($N = 67$). Die Nullhypothese wird damit nicht verworfen. Abbildung 27 zeigt Alter und die Differenzen der Gesamtproblematik im Streudiagramm.

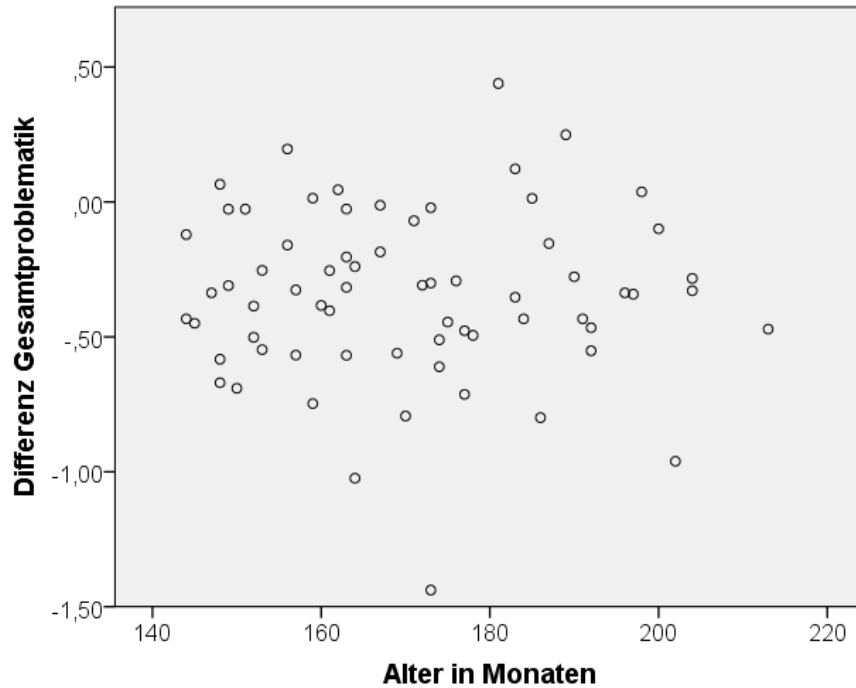


Abbildung 27: Alter der Teilnehmer und Differenzwerte im Gesamtproblemmaß

Tabelle 23 zeigt die Korrelationen von Alter mit den Differenzwerten in den vier Problemmaßen.

Tabelle 23: Zusammenhang des Alters der Teilnehmer mit den Differenzwerten

		Symptomatik	Lehrereinschätzung	Reaktionszeit	Fehler	Gesamtproblematik
Alter	r	-.11	.05	.06	-.07	.00
	p	.20	.60	.49	.39	.99
	N	66	52	66	66	67

Für die vier Problembereiche liegen beim „Testen auf Probe“ keine signifikanten Zusammenhänge vor.

Tabelle 24 zeigt die Effektstärken nach Altersgruppen analog zu Durlak et al. (1991) und Dush et al. (1989). Dazu wird die Post-prä-Differenz durch die Standardabweichung der Werte zum Prämesszeitpunkt geteilt.

Tabelle 24: Prä-post-Effektstärken der Gesamtproblematik nach Altersgruppen

Altersgruppe	Effektstärke (diff/SDprä)
11-13 (n=33)	-0.72
14-15 (n=24)	-0.79
16-17 (n=10)	-0.68

Initiale Störungsschwere

Es wird überprüft, ob ein Zusammenhang zwischen der Gesamtproblematik vor dem Training und dem Gewinn durch das Training, der durch das Differenzmaß operationalisiert wird, besteht. Da keine einheitliche Annahme zur Gerichtetheit des Zusammenhangs abgeleitet werden kann, wird ungerichtet getestet. Als Maß des Zusammenhangs wird Kendalls τ bestimmt.

$$H_0: \rho_{\text{InitialeStörungsschwere, VeränderungswertGesamtproblematik}} = 0$$

$$H_1: \rho_{\text{InitialeStörungsschwere, VeränderungswertGesamtproblematik}} \neq 0$$

Die Hypothesenprüfung ergibt $\tau = -.23$, $p = .01$, $N = 67$. Damit kann die Nullhypothese mit $\alpha = 5\%$ verworfen werden. Die initiale Störungsschwere und die Differenzen der Gesamtproblematik stehen in einem signifikanten Zusammenhang. Dieser ist als gering zu bezeichnen.

Zur deskriptiven Veranschaulichung sind in Abbildung 28 die Werte des Gesamtproblemmaßes zum Prä- und Postmesszeitpunkt dargestellt. Gestrichelt eingezeichnet ist die erste Winkelhalbierende ($y = x$), alle Punkte darüber zeigen eine Verschlechterung an, alle Werte darauf keine Veränderung und alle Werte darunter eine Verbesserung. Weiterhin ist die Regressionsgerade mit einer Varianzaufklärung von $R^2 = 56\%$ eingezeichnet. Aus der Grafik lässt sich ablesen, dass Teilnehmer mit höheren Anfangswerten eine stärkere Verbesserung erfahren haben als solche mit kleineren Anfangswerten, da die Regressionsgerade flacher und unterhalb der Winkelhalbierenden verläuft.

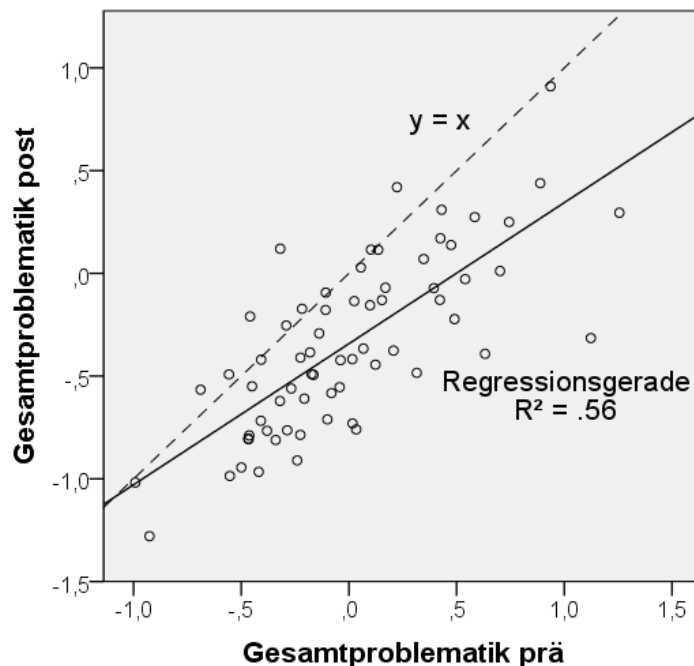


Abbildung 28: Zusammenhang der Werte der Gesamtproblematik vor und nach dem Training

Ergänzend zeigt Tabelle 25 die Zusammenhänge der initialen Störungsschwere des jeweiligen Bereichs mit dem zugehörigen Differenzwert in den vier Problemmaßen.

Tabelle 25: Zusammenhang der initialen Störungsschwere mit den Differenzwerten

		Symptomatik	Lehrer- einschätzung	Reaktions- zeit	Fehler	Gesamt- problematik
Initiale Störungsschwere	τ	-.23	-.24	-.41	-.31	-.23
	ρ	.01	.01	.00	.00	.01
	N	66	52	66	66	67

Für alle Bereiche ergeben sich signifikante negative Korrelationen zwischen der initialen Störungsschwere bzw. dem Status und den Differenzen bzw. Gewinn. Je höher die Werte vor dem Training ausgeprägt sind, desto höher ist die Reduktion der Werte. Die Zusammenhänge sind gering für die Symptomatik ($\tau = -.23$), Lehrereinschätzung ($\tau = -.24$) und Gesamtproblematik ($\tau = -.23$) und etwas höher für die Fehler ($\tau = -.31$) und am höchsten für die Reaktionszeit ($\tau = -.41$).

Komorbidität

Es wird angenommen, dass eine stärkere externalisierende komorbide Problematik mit geringeren Trainingserfolgen verbunden ist und eine stärkere internalisierende Problematik mit stärkeren Trainingserfolgen. Die externalisierenden Probleme werden mit der SDQ-Subskala „Verhaltensprobleme“ erfasst und die internalisierenden Probleme mit der SDQ-Subskala „Emotionale Probleme“. Es ergeben sich folgende Hypothesenpaare:

$$H_1: \rho_{\text{KomorbiditätExt, VeränderungswertGesamtproblematik}} > 0$$

$$H_0: \rho_{\text{KomorbiditätExt, VeränderungswertGesamtproblematik}} \leq 0$$

$$H_1: \rho_{\text{KomorbiditätInt, VeränderungswertGesamtproblematik}} \leq 0$$

$$H_0: \rho_{\text{KomorbiditätInt, VeränderungswertGesamtproblematik}} > 0$$

Da nicht für alle Variablen eine Normalverteilung angenommen werden kann, wird das robustere Kendalls τ als Maß des Zusammenhangs genutzt.

Vergleich der Werte des Prämesszeitpunktes

Tabelle 26 zeigt den Zusammenhang zwischen den Einschätzungen der externalisierenden Komorbidität und den Werten der Problemaße zum Prämesszeitpunkt.

Tabelle 26: Zusammenhang der externalisierenden Komorbidität mit den Problemmaßen vor dem Training

		Symptoma- tik	Lehrer- einschät- zung	Reaktions- zeit	Fehler	Gesamt- problematik
Verhaltensprobleme Elterneinschätzung	r	.27	.27	.07	.00	.28
	p	.00	.01	.45	.98	.00
	N	67	57	67	67	67
Verhaltensprobleme Lehrereinschätzung	r	.13	.52	.19	.00	.37
	p	.20	.00	.06	.99	.00
	N	57	57	57	57	57
Verhaltensprobleme Selbsteinschätzung	r	.41	.22	-.05	.10	.29
	p	.00	.02	.60	.26	.00
	N	67	57	67	67	67

Die Einschätzungen der Verhaltensprobleme durch die Eltern korrelieren signifikant positiv mit der Symptomatik, der Lehrereinschätzung und der Gesamtproblematik. Die Lehrereinschätzung der Verhaltensprobleme korreliert signifikant positiv mit der (Gesamt-)Lehrereinschätzung und der Gesamtproblematik. Die Selbsteinschätzung der Verhaltensprobleme korreliert signifikant positiv mit der Symptomatik, der Lehrereinschätzung und der Gesamtproblematik. Tabelle 27 zeigt den Zusammenhang zwischen den Einschätzungen der internalisierenden Komorbidität und den Werten der Problemaße zum Prämesszeitpunkt.

Tabelle 27: Zusammenhang der internalisierenden Komorbidität mit den Problemmaßen vor dem Training

		Symptoma- tik	Lehrer- einschät- zung	Reaktions- zeit	Fehler	Gesamt- problematik
Emot. Probleme Elterneinschätzung	r	.31	-.03	-.15	-.07	.07
	p	.00	.79	.08	.46	.43
	N	67	57	67	67	67
Emot. Probleme Lehrereinschätzung	r	.21	.14	.01	.00	.18
	p	.03	.16	.91	.97	.06
	N	57	57	57	57	57
Emot. Probleme Selbsteinschätzung	r	.36	-.03	-.22	.01	.08
	p	.00	.79	.01	.88	.37
	N	67	57	67	67	67

Die Einschätzungen der emotionalen Probleme durch die Eltern, Lehrer und Jugendliche korrelieren signifikant positiv mit der Symptomatik. Weiterhin korreliert die Selbsteinschätzung negativ mit dem Problemaß Reaktionszeit.

Vergleich der Differenzwerte

Bezogen auf die Verhaltensprobleme ergibt sich für die Elterneinschätzung der Gesamtproblematik $\tau = -.09$, $p = .31$, $N = 67$. Für die Lehrereinschätzung ergibt sich $\tau = -.10$, $p = .30$, $N = 57$. Für die Selbsteinschätzung ergibt sich $\tau = -.14$, $p = .11$, $N = 67$. Die Nullhypothese wird damit nicht verworfen. Es liegt kein signifikanter Zusammenhang zwischen dem Ausmaß der externalisierenden Komorbidität und den Differenzwerten vor. Tabelle 28 zeigt weiterführend die Zusammenhänge zwischen den Einschätzungen der externalisierenden Komorbidität und den Differenzwerten auf der Ebene der Problemmaße.

Tabelle 28: Zusammenhang der externalisierenden Komorbidität mit den Differenzwerten

		Symptomatik	Lehrereinschätzung	Reaktionszeit	Fehler	Gesamtproblematik
Verhaltensprobleme Elterneinschätzung	τ	-.07	-.11	-.03	.11	-.09
	p	.45	.27	.73	.21	.31
	N	66	52	66	66	67
Verhaltensprobleme Lehrereinschätzung	τ	.10	-.14	-.10	.08	-.10
	p	.30	.16	.33	.41	.30
	N	56	52	56	56	57
Verhaltensprobleme Selbsteinschätzung	τ	-.03	-.20	.01	.07	-.14
	p	.71	.04	.88	.43	.11
	N	66	52	66	66	67

Es besteht eine signifikante Korrelation zwischen der Selbsteinschätzung und den Differenzen des Lehrerurteils. Es handelt sich um einen schwachen Zusammenhang ($\tau = -.20$).

Bezogen auf die emotionalen Probleme ergibt sich für die Elterneinschätzung der Gesamtproblematik eine Korrelation von $\tau = -.14$, $p = .11$, $N = 67$. Für die Lehrereinschätzung ergibt sich $\tau = -.19$, $p = .05$, $N = 57$. Für die Selbsteinschätzung ergibt sich $\tau = -.10$, $p = .26$, $N = 67$. Die Nullhypothese kann damit für die Eltern- und Selbsteinschätzung nicht verworfen werden, für die Lehrereinschätzung aber schon. Es liegt kein signifikanter Zusammenhang zwischen dem Ausmaß der internalisierenden Komorbidität im Eltern- und Selbsturteil vor, aber ein signifikanter Zusammenhang zwischen Lehrerurteil der emotionalen Probleme und den Differenzwerten. Die Stärke des Zusammenhangs ist mit $\tau = -.19$ als gering einzuschätzen. Tabelle 29 zeigt weiterführend die Zusammenhänge zwischen den Einschätzungen der internalisierenden Komorbidität und den Differenzwerten auf der Ebene der Problemmaße.

Tabelle 29: Zusammenhang der internalisierenden Komorbidität mit den Differenzwerten

		Symptoma- tik	Lehrer- einschät- zung	Reaktions- zeit	Fehler	Gesamt- problematik
Emot. Probleme Elterneinschätzung	τ	-.18	-.09	.07	-.05	-.14
	p	.04	.40	.41	.56	.11
	N	66	52	66	66	67
Emot. Probleme Lehrereinschätzung	τ	-.20	-.23	.03	.08	-.19
	p	.04	.02	.78	.37	.05
	N	56	52	56	56	57
Emot. Probleme Selbsteinschätzung	τ	-.18	-.14	.16	-.09	-.10
	p	.05	.17	.07	.30	.26
	N	66	52	66	66	67

Bezogen auf die Gesamtproblematik zeigt sich eine signifikante negative Korrelation mit der Lehrereinschätzung der emotionalen Probleme. Der Zusammenhang ist mit $\tau = -.19$ als gering zu bezeichnen. Die Lehrereinschätzung korreliert auch signifikant negativ mit den Differenzwerten der Symptomatik ($\tau = -.20$) und den Differenzwerten der Lehrereinschätzung ($\tau = -.23$). Weiterhin zeigt sich eine signifikante Korrelation zwischen der Elterneinschätzung und den Differenzen in der Symptomatik von ($\tau = -.18$). Auch die Selbsteinschätzung korreliert signifikant negativ mit den Differenzen in der Symptomatik ($\tau = -.18$). Die Stärke des Zusammenhangs ist als schwach zu bezeichnen.

Trainingsdauer

Es wird angenommen, dass eine kürzere Trainingsdauer mit größerem Erfolg verbunden ist bzw. mit stärker negativen Differenzwerten. Die zu prüfenden Hypothesen lauten:

$$H_1: \rho_{\text{Trainingsdauer, VeränderungswertGesamtproblematik}} > 0$$

$$H_0: \rho_{\text{Trainingsdauer, VeränderungswertGesamtproblematik}} \leq 0$$

Als Maß für die Trainingsdauer wird das Zeitintervall zwischen der Vorher- und Nachhermessung herangezogen. Dieses Intervall ist etwas länger als die Trainingsdauer, es wurde aber darauf geachtet, dass zwischen Messung und Training nicht viel Zeit vergeht. Die mittlere Trainingsdauer beträgt $M = 5.7$ Monate ($SD = 1.67$).

Vergleich der Werte zum Prämesszeitpunkt

Tabelle 30 zeigt die Zusammenhänge der Trainingsdauer mit den Problemmaßen zum Prämesszeitpunkt. Da bei den Problemmaßen nicht von einer Normalverteilung ausgegangen werden kann, wird Kendalls τ als Maß des Zusammenhangs bestimmt.

Tabelle 30: Zusammenhang der Trainingsdauer mit den Problemmaßen vor dem Training

		Symptomatik	Lehrer- einschätzung	Reaktions- zeit	Fehler	Gesamt- problematik
Trainingsdauer	τ	.18	.22	.07	.15	.28
	p	.05	.03	.45	.10	.00
	N	67	57	67	67	67

Es zeigen sich signifikante positive Korrelationen der Symptomatik, der Lehrereinschätzung und der Gesamtproblematik mit der Trainingsdauer. Diese sind im kleinen Bereich.

Vergleich der Differenzwerte

Anhand der Gesamtproblematik wird die aufgestellte Hypothese überprüft. Es ergibt zwar ein signifikantes Ergebnis mit $\tau = -.22$ ($p = .01$, $N = 67$), dieses läuft aber in die entgegengesetzte Richtung als angenommen wurde. Damit wird die Nullhypothese nicht abgelehnt. Tabelle 31 zeigt die Zusammenhänge der Trainingsdauer mit allen Differenzwerten.

Tabelle 31: Zusammenhang der Trainingsdauer mit den Differenzwerten

		Symptomatik	Lehrer- einschätzung	Reaktions- zeit	Fehler	Gesamt- problematik
Trainingsdauer	τ	-.09	-.12	-.19	-.09	-.22
	p	.32	.24	.04	.35	.01
	N	66	52	66	66	67

Es zeigen sich signifikante negative Korrelationen der Differenzwerte der Reaktionszeit und der Gesamtproblematik mit der Trainingsdauer. Diese liegen im kleinen Bereich. Tendenziell gilt also, dass eine längere Trainingsdauer mit höheren Differenzwerten zusammenhängt.

Anzahl der Sitzungen

Es wird angenommen, dass Trainings mit weniger Sitzungen erfolgreicher waren und kleinere Differenzwerte aufweisen. Die Hypothesen lauten:

$$H_1: \rho_{\text{AnzahlSitzungen, VeränderungswertGesamtproblematik}} > 0$$

$$H_0: \rho_{\text{AnzahlSitzungen, VeränderungswertGesamtproblematik}} \leq 0$$

Die meisten Trainings nehmen 20 Sitzungen in Anspruch. Der Mittelwert liegt bei $M = 18.11$ ($SD = 2.11$). Die Anzahl der Sitzungen korreliert nicht signifikant mit der Trainingsdauer ($\tau = .16$, $p = .11$, $N = 65$).

Vergleich der Werte zum Prämesszeitpunkt

Tabelle 32 zeigt die Zusammenhänge zwischen der Anzahl der Sitzungen und den Problemmaßen zum Prämesszeitpunkt. Es zeigen sich keine signifikanten Zusammenhänge.

Tabelle 32: Zusammenhang der Anzahl der Sitzungen mit den Werten zum Prämesszeitpunkt

		Symptomatik	Lehrer- einschätzung	Reaktions- zeit	Fehler	Gesamt- problematik
Anzahl Sitzungen	r	-.08	-.05	-.09	.01	-.08
	p	.36	.64	.30	.96	.39
	N	65	55	65	65	65

Vergleich der Differenzwerte

Es zeigt sich kein signifikanter Zusammenhang zwischen Anzahl der Sitzungen und dem Differenzwert der Gesamtproblematik ($p = .16$, $N = 65$). Damit kann die Nullhypothese nicht abgelehnt werden. Tabelle 33 zeigt die Zusammenhänge zwischen der Anzahl der Sitzungen und allen Differenzwerten. Es zeigen sich keine signifikanten Zusammenhänge.

Tabelle 33: Zusammenhang der Anzahl der Sitzungen mit den Differenzwerten

		Symptomatik	Lehrer- einschätzung	Reaktions- zeit	Fehler	Gesamt- problematik
Anzahl Sitzungen	r	.11	.11	.04	.10	.13
	p	.22	.31	.70	.28	.16
	N	64	50	64	64	65

6.2 Variablenzentrierte multiple univariate Auswertung

Fragestellung 1 b: *Wie hängen die Merkmale der Teilnehmer und der Trainingsdurchführung – wenn sie gleichzeitig berücksichtigt werden – mit den verschiedenen Aspekten des Trainingserfolges zusammen?*

Im Folgenden werden die Ergebnisse der schrittweisen Regressionsanalyse mit SPSS 20 zur Vorhersage der Werte nach dem Training dargestellt. Abhängige Variablen sind die Daten zum Postmesszeitpunkt und die Differenzwerte. Zunächst wird das Gesamtproblemmaß analysiert, danach einzeln die vier Problemfaktoren. Es handelt sich also um univariate Auswertungen. Die unabhängigen Variablen werden wie zuvor beschrieben zentriert, um dem Problem der Multikollinearität vorzubeugen (Kraemer & Blasey, 2004). Fehlende Werte werden paarweise ausgeschlossen, d. h. nur wenn der Wert für die jeweilige Analyse fehlte. Unterschiedliche Stichprobengrößen bei den Auswertungen resultieren aus fehlenden Werten. Um Verzerrungen zu vermeiden wird darauf verzichtet, den Mittelwert für fehlende Werte einzusetzen. Um den Datensatz nicht zu sehr

zu verkleinern wird kein listenweiser Fallausschluss vorgenommen, bei dem Fälle mit zumindest einem fehlenden Wert aus allen Analysen ausgeschlossen werden. Es ist darauf zu achten, dass nicht zu viele unabhängige Variablen in die Regressionsanalyse eingehen (Tabachnick & Fidell, 2007).

Folgende Merkmale und Ausprägungen werden genutzt:

- (1) *Geschlecht*: weiblich, männlich
- (2) *Medikamente*: ja, nein
- (3) *ADHS-Subtypus*: Elterneinschätzung der Symptome von Hyperaktivität und Impulsivität im DISYPS-II. Lehrer- und Selbsteinschätzungen werden an dieser Stelle nicht berücksichtigt, da sich sonst zu viele unabhängige Variablen ergeben.
- (4) *Alter* in Monaten
- (5) *Initiale Störungsschwere*: Werte des Prämesszeitpunktes von Gesamtproblematik bzw. der vier Problemmaße, passend zur jeweiligen abhängigen Variable
- (6) *Externalisierende Komorbidität*: Selbsteinschätzung der Jugendlichen der SDQ-Subskala „Verhaltensprobleme“. Lehrer- und Elterneinschätzungen werden an dieser Stelle nicht berücksichtigt, da sich sonst zu viele unabhängige Variablen ergeben und sich die Selbsteinschätzung in der univariaten Analyse als möglicherweise relevant herausgestellt hat.
- (7) *Internalisierende Komorbidität*: Lehrereinschätzung der SDQ-Subskala „Emotionale Probleme“. Eltern- und Selbsteinschätzungen werden an dieser Stelle nicht berücksichtigt, da sich sonst zu viele unabhängige Variablen ergeben und sich die Lehrereinschätzung in der univariaten Analyse als möglicherweise relevant herausgestellt hat.
- (8) *Dauer des Trainings* in Monaten
- (9) *Anzahl der durchgeführten Sitzungen*

AV Gesamtproblematik post

In Tabelle 34 sind die Ergebnisse der schrittweisen Regressionsanalyse mit SPSS 20 dargestellt.

Tabelle 34: Ergebnisse der schrittweisen linearen Regressionsanalyse. Abhängige Variable: Gesamtproblematik post ($N = 67$)

Unabhängige Variable(n)	β	t	p	R^2
Gesamtproblematik prä	.75	8.22	.00	.56
Ausgeschlossene Variablen				
Geschlecht		1.00	.32	
Medikamente		-1.12	.27	
Hyperaktivität-Impulsivität		0.40	.69	
Alter		-.71	.48	
Externalisierende Komorbidität		-.04	.97	
Internalisierende Komorbidität		-.91	.37	
Dauer des Trainings		-1.32	.19	
Anzahl der Sitzungen		1.63	.11	

In der schrittweisen Regression kann die Gesamtproblematik zum Prämesszeitpunkt als einzige Variable signifikant Varianz in der Gesamtproblematik zum Postmesszeitpunkt aufklären. Die übrigen Variablen werden ausgeschlossen, da sie nicht signifikant mit der abhängigen Variablen zusammenhängen oder keine zusätzliche Varianz aufklären können. Das Modell erklärt 56 % der Varianz. Der Regressionskoeffizient $\beta = .75$ zeigt an, dass eine höhere Gesamtproblematik vor dem Training mit einer höheren Gesamtproblematik nach dem Training verbunden ist. Die relevanten Modellprämissen werden erfüllt. So zeigt die visuelle Inspektion der Residuen, dass keine Heteroskedastizität vorliegt. Die Durbin-Watson-Statistik liegt bei 2.04 und damit sehr nahe bei dem optimalen Wert von 2.00. Es werden ein Toleranzwert von 1.00 und ein VIF-Wert von 1.00 ermittelt.

AV Differenz Gesamtproblematik

Tabelle 35 zeigt die Ergebnisse der schrittweisen Regressionsanalyse mit der abhängigen Variable Differenzwerte der Gesamtproblematik.

Tabelle 35: Ergebnisse der schrittweisen linearen Regressionsanalyse. Abhängige Variable: Differenz Gesamtproblematik ($N = 67$)

Unabhängige Variable(n)	β	t	p	R^2
Gesamtproblematik prä	-.46	-3.76	.00	.21
Ausgeschlossene Variablen				
Geschlecht		1.00	.32	
Medikamente		-1.12	.27	
Hyperaktivität-Impulsivität		0.40	.69	
Alter		-.71	.48	
Externalisierende Komorbidität		-.04	.97	
Internalisierende Komorbidität		-.89	.38	
Dauer des Trainings		-1.32	.19	
Anzahl der Sitzungen		1.71	.09	

Es zeigen sich beinahe identische Ergebnisse wie bei der Vorhersage der Werte des Postmesszeitpunktes. Die Gesamtproblematik zum Prämesszeitpunkt kann 21 % der Varianz in den Differenzwerten erklären. Das negative Gewicht von $\beta = -.46$ zeigt an, dass bei höherem Wert zum Prämesszeitpunkt stärker negative Differenzwerte erreicht werden und damit größere Verbesserungen. Die Modellprämissen werden ebenso mit einer Durbin-Watson-Statistik von 2.04 und Toleranzwert und VIF-Wert von 1.00 erfüllt.

Es folgen die Regressionsanalysen für die vier Problemmaße. Zuerst werden als abhängige Variable die Werte des jeweiligen Problemmaßes zum Postmesszeitpunkt herangezogen, danach die Differenzwerte. Zugunsten der Übersichtlichkeit wird im Weiteren darauf verzichtet, die ausgeschlossenen Variablen jeweils einzeln darzustellen, da sie keine zusätzlichen Informationen liefern. Es werden nur diejenigen Variablen abgebildet, die signifikant Varianz aufklären konnten. Im Anhang finden sich genaue Angaben zu den ausgeschlossenen Variablen.

AV Symptomatik postTabelle 36: Schrittweise lineare Regressionsanalyse. Abhängige Variable: Symptomatik post ($N = 66$)

Modell	β	t	p	R^2
Unabhängige Variablen				
1 Symptomatik prä	.72	7.49	.00	.51

Die Symptomatik zum Prämesszeitpunkt kann 51 % der Varianz in der Symptomatik zum Postmesszeitpunkt aufklären (siehe Tabelle 36). Die übrigen Variablen werden ausgeschlossen, da sie keine zusätzliche Varianz aufklären. Je höher die Symptomatik zum Prämesszeitpunkt ausgeprägt war, desto höher ist sie auch zum Postmesszeitpunkt. Die Modellprämissen werden bei einer Durbin-Watson-Statistik von 1.84 und Toleranz- und VIF-Wert von 1.00 erfüllt.

AV Differenz SymptomatikTabelle 37: Schrittweise lineare Regressionsanalyse. Abhängige Variable: Differenz Symptomatik ($N = 66$)

Modell	β	t	p	R^2
Unabhängige Variablen				
1 Symptomatik prä	-.41	-3.27	.00	.17

Die Symptomatik zum Prämesszeitpunkt kann 17 % der Varianz in den Differenzwerten der Symptomatik aufklären (siehe Tabelle 37). Die übrigen Variablen werden ausgeschlossen. Je höher die Symptomatik zum Prämesszeitpunkt ausgeprägt ist, desto stärker negativ sind die Differenzwerte. Die Modellprämissen werden bei einer Durbin-Watson-Statistik von 1.85 und Toleranz- und VIF-Wert von 1.00 erfüllt.

AV Lehrereinschätzung postTabelle 38: Schrittweise lineare Regressionsanalyse. Abhängige Variable: Lehrereinschätzung post ($N = 57$)

Modell	β	t	p	R^2
Unabhängige Variablen				
1 Lehrereinschätzung prä	.53	-4.16	.00	.28
2 Lehrereinschätzung prä	.47	3.93	.00	
Medikamente	-.24	-2.03	.05	.33

Bei der Vorhersage der Lehrereinschätzung zum Postmesszeitpunkt werden zwei Modelle ausgegeben (siehe Tabelle 38). Im ersten Modell, bei dem 28 % der Varianz aufgeklärt wird, geht nur die Lehrereinschätzung zum Prämesszeitpunkt ein. Je höher die Lehrereinschätzung zum Prämesszeitpunkt ausgeprägt ist, desto höher ist sie auch zum Postmesszeitpunkt. Im zweiten Modell, das insgesamt 33 % Varianz aufklären kann, geht zusätzlich die unabhängige Variable Medikamente ein. Dadurch werden 5 % zusätzliche Varianz aufgeklärt. Die negative Gewichtung ($\beta = -.24$) zeigt an, dass die Lehrereinschätzung zum Postmesszeitpunkt geringer ist für Jugendliche, die eine Medikation bekommen, da keine Medikation mit $-1/2$ kodiert wurde und vorliegende Medikation mit $+1/2$. Die Modellprämissen werden bei einer Durbin-Watson-Statistik von 2.06 und Toleranz- und VIF-Werten von 1.00 für Modell 1 und mit 0.95 bzw. 1.05 für Modell 2 erfüllt.

AV Differenz Lehrereinschätzung

Tabelle 39: Schrittweise lineare Regressionsanalyse. Abhängige Variable: Differenz Lehrereinschätzung ($N = 52$)

Modell	β	t	p	R^2
Unabhängige Variablen				
1				
Lehrereinschätzung prä	-.54	-4.39	.00	.29

Die Lehrereinschätzung zum Prämesszeitpunkt kann 29 % der Varianz in den Differenzwerten der Lehrereinschätzung aufklären (siehe Tabelle 39). Die übrigen Variablen werden ausgeschlossen. Je höher die Lehrereinschätzung zum Prämesszeitpunkt ausgeprägt ist, desto stärker negativ sind die Differenzwerte. Die Modellprämissen werden bei einer Durbin-Watson-Statistik von 2.12 und Toleranz- und VIF-Wert von 1.00 erfüllt.

AV Reaktionszeit post

Tabelle 40: Schrittweise lineare Regressionsanalyse. Abhängige Variable: Reaktionszeit post ($N = 66$)

Modell	β	t	p	R^2
Unabhängige Variablen				
1				
Reaktionszeit prä	.87	12.78	.00	.76
2				
Reaktionszeit prä	.86	13.02	.00	
Medikamente	-.15	-2.23	.03	.78

Bei der Vorhersage der Reaktionszeit zum Postmesszeitpunkt werden zwei Modelle ausgegeben (siehe Tabelle 40). Im ersten Modell, bei dem 76 % der Varianz aufgeklärt wird, geht nur die Reaktionszeit zum Prämesszeitpunkt ein. Je höher die Reaktionszeit

zum Prämesszeitpunkt, desto höher ist sie auch zum Postmesszeitpunkt. Im zweiten Modell, das insgesamt 78 % Varianz aufklären kann, geht zusätzlich die unabhängige Variable Medikamente ein. Durch diese Hinzunahme werden 2 % zusätzliche Varianz aufgeklärt. Die negative Gewichtung ($\beta = -.15$) zeigt an, dass die Reaktionszeit zum Postmesszeitpunkt geringer für Jugendliche ist, die eine Medikation bekommen. Die Modellprämissen werden bei einer Durbin-Watson-Statistik von 2.20 und Toleranz- und VIF-Werten von 1.00 für Modell 1 und mit 0.99 bzw. 1.01 für Modell 2 erfüllt.

AV Differenz Reaktionszeit

Tabelle 41: Schrittweise lineare Regressionsanalyse. Abhängige Variable: Differenz Reaktionszeit ($N = 66$)

Modell	Unabhängige Variablen	β	t	p	R^2
1	Reaktionszeit prä	-.65	-6.29	.00	.43
2	Reaktionszeit prä	-.67	-6.66	.00	.48
	Medikation	-.22	-2.17	.04	
3	Reaktionszeit prä	-.65	-6.53	.00	.52
	Medikamente	-.21	-2.13	.04	
	Dauer des Trainings	-.20	-2.04	.05	

Bei der Vorhersage der Differenzwerte der Reaktionszeit werden drei Modelle ausgegeben (siehe Tabelle 41). Im ersten Modell, bei dem 43 % der Varianz aufgeklärt wird, geht nur die Reaktionszeit zum Prämesszeitpunkt ein. Je höher die Reaktionszeit zum Prämesszeitpunkt ist, desto stärker negativ sind die Differenzwerte. Im zweiten Modell, das insgesamt 48 % Varianz aufklären kann, geht zusätzlich die unabhängige Variable Medikamente ein. Durch diese Hinzunahme werden 5 % zusätzliche Varianz aufgeklärt. Die negative Gewichtung ($\beta = -.22$) zeigt an, dass die Differenz der Reaktionszeit stärker negativ für Jugendliche ist, die eine Medikation bekommen. Im dritten Modell werden 52 % der Varianz aufgeklärt. Zusätzlich zur Reaktionszeit des Prämesszeitpunkt und der Medikation geht die Dauer des Trainings ein, die 4 % zusätzliche Varianz aufklären kann. Die negative Gewichtung ($\beta = -.20$) zeigt an, dass bei einer längeren Trainingsdauer stärker negative Differenzwerte erreicht wurden. Die Modellprämissen werden erfüllt mit Durbin-Watson-Statistik mit 2.24 und Toleranz- bzw. VIF-Werten von je 1.00 für Modell 1, 0.99 bzw. 1.01 für Modell 2 und 0.98 – 0.99 bzw. 1.01 – 1.03 für Modell 3.

AV Fehler postTabelle 42: Schrittweise lineare Regressionsanalyse. Abhängige Variable: Fehler post ($N = 66$)

Modell	β	t	p	R^2
Unabhängige Variablen				
1 Fehler prä	.73	7.86	.00	.54

Die Fehlervariable zum Prämesszeitpunkt kann 54 % der Varianz in den Werten der Fehlervariable zum Postmesszeitpunkt aufklären (siehe Tabelle 42). Die übrigen Variablen werden ausgeschlossen. Je höher die Fehlervariable zum Prämesszeitpunkt ausgeprägt ist, desto höher sind Werte zum Postmesszeitpunkt. Die Modellprämissen werden bei einer Durbin-Watson-Statistik von 1.92 und Toleranz- und VIF-Wert von 1.00 erfüllt.

AV Differenz FehlerTabelle 43: Schrittweise lineare Regressionsanalyse. Abhängige Variable: Differenz Fehler ($N = 66$)

Modell	β	t	p	R^2
Unabhängige Variablen				
1 Fehler prä	-.51	-4.28	.00	.26

Die Fehlervariable zum Prämesszeitpunkt kann 26 % der Varianz in den Differenzwerten der Fehlervariable aufklären (siehe Tabelle 43). Die übrigen Variablen werden ausgeschlossen. Je höher die Fehlervariable zum Prämesszeitpunkt ausgeprägt ist, desto stärker negativ sind die Differenzwerte. Die Modellprämissen werden bei einer Durbin-Watson-Statistik von 1.97 und Toleranz- und VIF-Wert von 1.00 erfüllt.

6.3 Variablenzentrierte multivariate Auswertung

Fragestellung 1 c: *Wie hängen die Merkmale der Teilnehmer und der Trainingsdurchführung – wenn sie gleichzeitig berücksichtigt werden – mit den zusammengefassten Aspekten des Trainingserfolges zusammen?*

Bei der multivariaten Auswertung mit SPSS werden nur vollständige Datensätze einbezogen. Es wird also ein listenweiser Fallausschluss vorgenommen (Backhaus, Erichson, Plinke & Weiber, 2011). Es liegen $N = 49$ vollständige Datensätze für die relevanten Variablen vor.

Abhängige Variablen sind die Problemmaße Symptomatik und Lehrereinschätzung zum Postmesszeitpunkt, da diese als einzige der abhängigen Variablen des Trainingserfolgs signifikant korrelieren ($r = .36$, $p = .01$, $N = 56$). Als feste Faktoren gehen die zentrier-

ten dichotomen Variablen Geschlecht und Medikamenteneinnahme in die Auswertung mit SPSS 20 ein. Als Kovariaten werden die Werte der beiden Problemmaße zum Prämesszeitpunkt ein und die zentrierten kontinuierlichen Variablen Alter, Anzahl der Sitzungen und Trainingsdauer berücksichtigt. Tabelle 44 zeigt die Ergebnisse des globalen multivariaten Tests.

Tabelle 44: Ergebnisse der multivariaten Varianzanalyse (Pillai-Spur). Abhängige Variablen: Symptomatik und Lehrereinschätzung zum Postmesszeitpunkt ($N = 49$)

	Wert	F	p	Partielles η^2
Symptomatik prä	.21	4.84	.01	.21
Lehrereinschätzung prä	.28	6.86	.00	.28
Alter in Monaten	.02	0.34	.71	.02
Anzahl der Sitzungen	.05	0.91	.41	.05
Trainingsdauer in Monaten	.00	0.06	.95	.00
Hyperaktivität-Impulsivität	.07	1.36	.27	.07
Internalisierende Komorbidität	.03	.51	.61	.03
Externalisierende Komorbidität	.00	.08	.93	.00
Geschlecht	.10	2.00	.15	.10
Medikamenteneinnahme	.06	1.12	.34	.06
Geschlecht*Medikamentein.	.00	.02	.98	.00

Die Symptomatik und die Lehrereinschätzung zum Prämesszeitpunkt können signifikant Varianz in den Werten zum Postmesszeitpunkt aufklären. Weitere univariate Auswertungen zeigen, dass sich erwartungskonform die Präwerte auf die Postwerte sowohl für die Symptomatik als auch für die Lehrereinschätzung auswirken. Auf die Angabe der Parameterschätzer bzw. β -Gewichte für die Regressionsgleichung wird an dieser Stelle verzichtet, da sie denen der univariaten Auswertung entsprechen. Es wird auch darauf verzichtet, die Ergebnisse der Analyse mit den Differenzwerten als abhängige Variablen darzustellen, da sich bis auf die genauen Werte identische Ergebnisse zeigen.

6.4 Personenzentrierte univariate Auswertung

Fragestellung 2 a: *Lassen sich Gruppen von Teilnehmern mit besonders starken Verbesserungen und weniger starken bzw. ohne Verbesserungen durch das Training identifizieren und wenn ja, welche Unterschiede weisen diese Gruppen jeweils auf?*

Als Abschluss werden die Teilnehmer nach dem Gesamterfolgsmaß nach dem Reliable Change Index (RCI; Jacobson & Truax, 1991) eingruppiert. Dies kann als personenzentriertes univariates Vorgehen bezeichnet werden. Der RCI wird nach Nachtigall und Suhl (2005) bestimmt.

Die Standardabweichung der Differenzen des Gesamtproblemmaßes beträgt $S_{y-x} = 0.32$. Damit eine reliable Veränderung festgestellt werden kann, muss der RCI 1.96 überstei-

gen. Dies ist der Fall, wenn sich im Gesamtproblemmaß Veränderungen von mindestens 0.62 ergeben.

Abbildung 29 zeigt die Werte des Gesamtproblemmaßes zum Prä- und Postmesszeitpunkt. Die durchgezogene Linie zeigt an, dass keine Veränderung vorliegt. Werte darunter zeigen eine Verbesserung an, Werte darüber eine Verschlechterung. Als gestrichelte Linien sind die Grenzen durch den RCI eingezeichnet. Als zweites Kriterium bei der Bewertung des RCI soll ein Normvergleich dienen. Es soll überprüft werden, ob die Werte der Teilnehmer noch im dysfunktionalen oder schon im Normalbereich liegen. Dieses Kriterium kann hier nicht herangezogen werden, da es sich um ein für diese Untersuchung ermitteltes Gesamtproblemmaß handelt, zu dem keine Normwerte vorliegen.

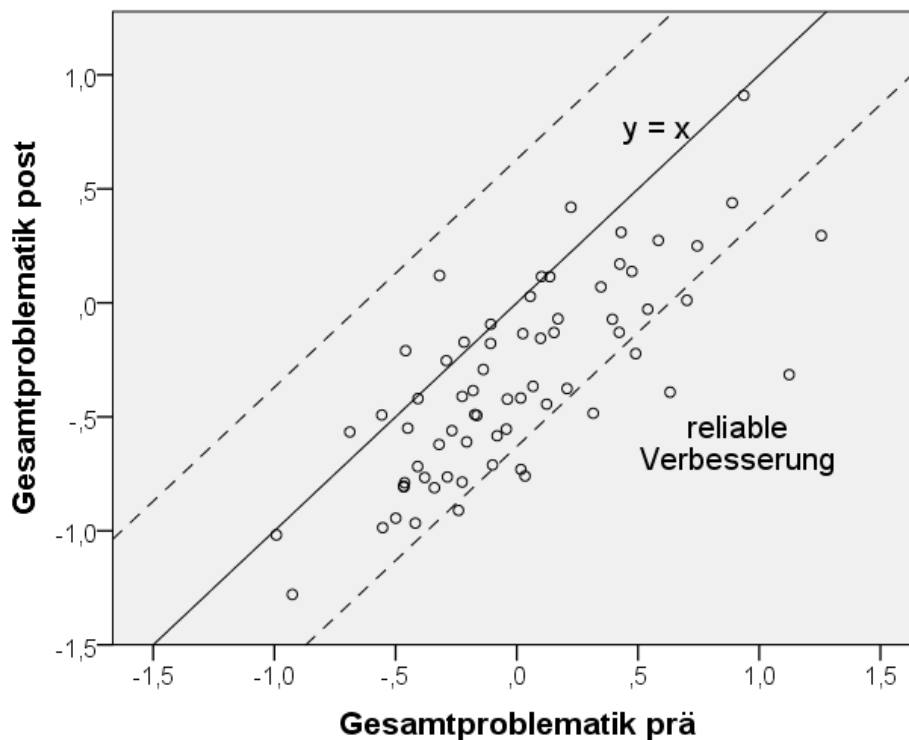


Abbildung 29: Gesamtproblematik vor und nach dem Training mit Reliable Change Index (RCI)

Von den 67 Teilnehmern erzielen 58 (87 %) nach dem Training bessere Werte als vorher. Neun Teilnehmer (13 %) erreichen nach dem Training schlechtere Werte als zuvor. Kein Teilnehmer weist eine reliable Verschlechterung nach RCI auf. Neun Teilnehmer (13 %) haben sich reliabel verbessert.

Tabelle 45 und Tabelle 46 zeigen eine Gegenüberstellung von Respondern und Non-Respondern. Es werden Mittelwerte und Standardabweichungen der beiden Gruppen zu den Merkmalen der Personen, Umstände der Durchführung und relevanten Instrumenten zum Prämesszeitpunkt angegeben. Die Darstellung erfolgt möglichst breit mit vielen abgebildeten Variablen, um explorativ solche Variablen zu finden, die in beiden Gruppen unterschiedlich verteilt sind.

Tabelle 45: Deskriptive Ergebnisse für Responder und Non-Responder nach RCI (dichotome Variablen)

	Gruppe		
	Non-Responder (<i>n</i> = 58)	Responder (<i>n</i> = 9)	Total (<i>N</i> = 67)
<i>N</i> Mädchen (%)	10 (17 %)	3 (33 %)	13 (19 %)
<i>N</i> keine Medikation (%)	31 (53 %)	6 (67 %)	37 (55 %)

Tabelle 46: Deskriptive Ergebnisse für Responder und Non-Responder nach RCI (metrische Variablen)

Variablen (Prämesszeitpunkt)	Gruppe		
	Non-Responder (<i>n</i> = 58)	Responder (<i>n</i> = 9)	Total (<i>N</i> = 67)
	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)
Alter in Monaten	170 (18)	170 (17)	171 (18)
Anzahl der Sitzungen	18 (2)	17 (2)	18 (2)
Trainingsdauer in Monaten	6 (1)	7 (2)	6 (2)
DISYPS-II gesamt Eltern	1.49 (0.48)	1.91 (0.56)	1.55 (0.50)
DISYPS-II gesamt Lehrer	1.08 (0.56)	1.63 (0.76)	1.16 (0.62)
DISYPS-II gesamt selbst	1.08 (0.59)	1.43 (0.53)	1.13 (0.59)
Hyperakt.-Impuls. Eltern	1.08 (0.61)	1.61 (0.74)	1.15 (0.65)
Hyperakt.-Impuls. Lehrer	0.72 (0.68)	1.14 (0.98)	0.78 (0.73)
Hyperakt.-Impuls. selbst	0.86 (0.64)	1.24 (0.55)	0.91 (0.64)
SDQ gesamt Eltern	19.05 (5.48)	23.11 (5.28)	19.60 (5.59)
SDQ gesamt Lehrer	12.33 (5.36)	19.75 (7.07)	13.37 (6.13)
SDQ gesamt selbst	15.48 (6.00)	19.33 (7.71)	16.00 (6.31)
SDQ Verhaltensprob. Eltern	3.71 (1.91)	4.11 (1.90)	3.76 (1.90)
SDQ Verhaltensprob. Lehrer	1.69 (1.67)	3.00 (2.88)	1.88 (1.91)
SDQ Verhaltensprob. selbst	2.98 (1.97)	3.44 (1.42)	3.04 (1.90)
SDQ Emot. Prob. Eltern	4.74 (2.63)	5.56 (3.28)	4.85 (2.71)
SDQ Emot. Prob. Lehrer	2.16 (1.86)	4.25 (3.41)	2.46 (2.23)
SDQ Emot. Prob. selbst	3.57 (2.64)	4.22 (2.95)	3.66 (2.67)
WLST Gesamtwert	55.97 (10.97)	50.89 (11.76)	55.28 (11.13)
WRI Gesamtwert	24.28 (10.43)	26.44 (10.20)	24.57 (10.35)
Regelkonformes Unterrichtsverhalten	75.64 (17.32)	59.81 (27.99)	73.42 (19.64)
Lern- & Aufmerksamkeitsverhalten	62.00 (19.14)	37.77 (22.18)	58.59 (21.16)
Arbeits- & Lernverhalten gesamt Lehrer	66.97 (16.43)	45.48 (21.24)	63.96 (18.56)
TAP Alertness Reaktionszeit	261 (44)	273 (48)	263 (45)
TAP Flexibilität Reaktionszeit	864 (276)	1021 (434)	885 (303)
TAP Flexibilität Fehleranzahl	8 (6)	11 (11)	8 (7)
TAP Go/NoGo Reaktionszeit	408 (84)	459 (96)	415 (87)
TAP Go/NoGo Fehleranzahl	5 (4)	4 (4)	5 (4)

Die Responder haben in allen Variablen zum Prämesszeitpunkt schlechtere Werte. Dies gilt übereinstimmend für die Einschätzungen durch die Eltern, Lehrer und Jugendlichen selbst. Nur in den Fehlern bei der Go/NoGo-Aufgabe machen sie im Durchschnitt einen Fehler weniger.

Die größten Unterschiede zeigen sich in den Gesamtwerten des DISYPS und des SDQ in der Eltern- und Lehrereinschätzung und im AVL. Beim probeweisen Signifikanztesten zeigt sich, dass für folgende Variablen signifikante Mittelwertunterschiede zwischen Respondern und Non-Respondern vorliegen: Elterneinschätzung Hyperaktivität-Impulsivität ($p = .02$, $ES_{pool} = 0.85$), DISYPS-II gesamt ($p = .02$, $ES_{pool} = 0.87$) und SDQ Gesamtproblemwert ($p = .04$, $ES_{pool} = 0.74$). Lehrereinschätzung DISYPS-II gesamt ($p = .02$, $ES_{pool} = 0.92$), SDQ emotionale Probleme ($p = .01$, $ES_{pool} = 1.01$) SDQ Gesamtproblemwert ($p = .00$, $ES_{pool} = 1.33$), Lern- und Aufmerksamkeitsverhalten ($p = .00$, $ES_{pool} = 1.24$) und AVL gesamt ($p = .00$, $ES_{pool} = 1.26$). Die Effektstärken werden aus dem Quotienten der Differenz und der gepoolten Standardabweichung nach Hartmann, Herzog und Drinkmann (1992) bestimmt.

6.5 Personenzentrierte multivariate Auswertung

Fragestellung 2 a: Lassen sich Gruppen von Teilnehmern mit besonders starken Verbesserungen und weniger starken bzw. ohne Verbesserungen durch das Training identifizieren und wenn ja, welche Unterschiede weisen diese Gruppen jeweils auf?

Für die vorliegende Auswertung werden nach den Differenzwerten der vier Problemmaße Symptomatik, Lehrereinschätzung, Reaktionszeit und Fehler gruppiert. So können gleiche „Erfolgsmuster“ identifiziert werden, also Teilnehmer mit einer gleichen Entwicklung in den unterschiedlichen Problemmaßen. In die Auswertung werden nur Teilnehmer mit vollständigen Gesamtwerten für alle Problemmaße einbezogen (75 %). Es wird davon abgesehen für fehlende Werte die jeweiligen Mittelwerte einzusetzen, da dies zu Verzerrungen und einer Varianzverringerng führen kann. Für die meisten Anwendungen der Clusteranalyse ist daher keine Ersetzung fehlender Werte durch Mittelwerte angeraten (Backhaus, Erichson, Plinke et al., 2011).

In Tabelle 47 ist das Ergebnis der Clusteranalyse abgebildet, bei der als Ähnlichkeitsmaß die einfache euklidische Distanz herangezogen wurde und als Fusionierungsmethode das Verfahren „Linkage innerhalb der Gruppen“. Um Unterschiede in den Skalierungen der Differenzmaße zu berücksichtigen, gehen die Werte z -standardisiert in die Analyse ein (Bühl, 2010). Tabelle 47 zeigt die Größe der Cluster, die Mittelwerte und die Standardabweichungen (in Klammern), die sich für die Cluster jeweils ergeben.

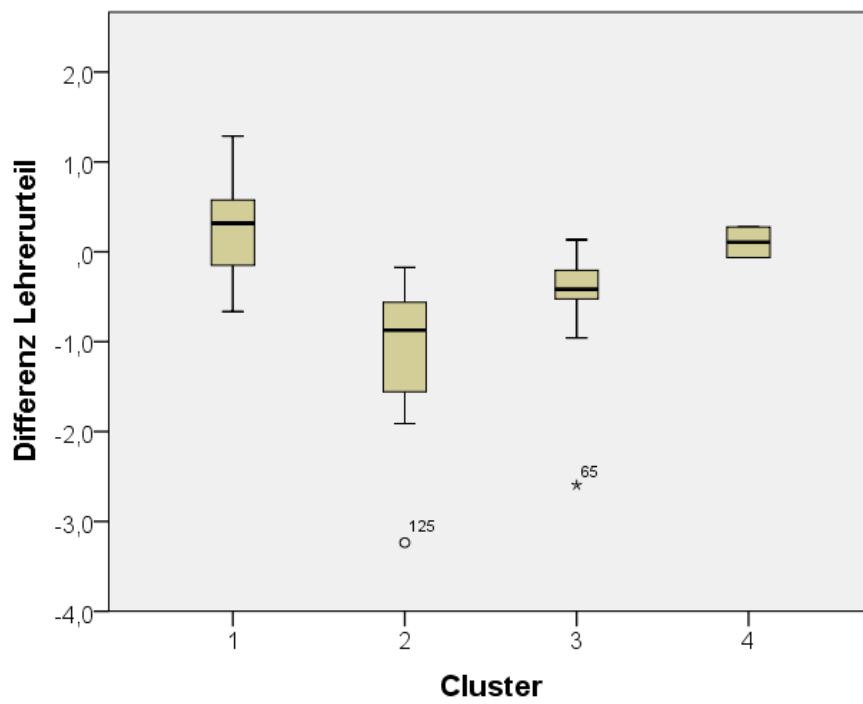
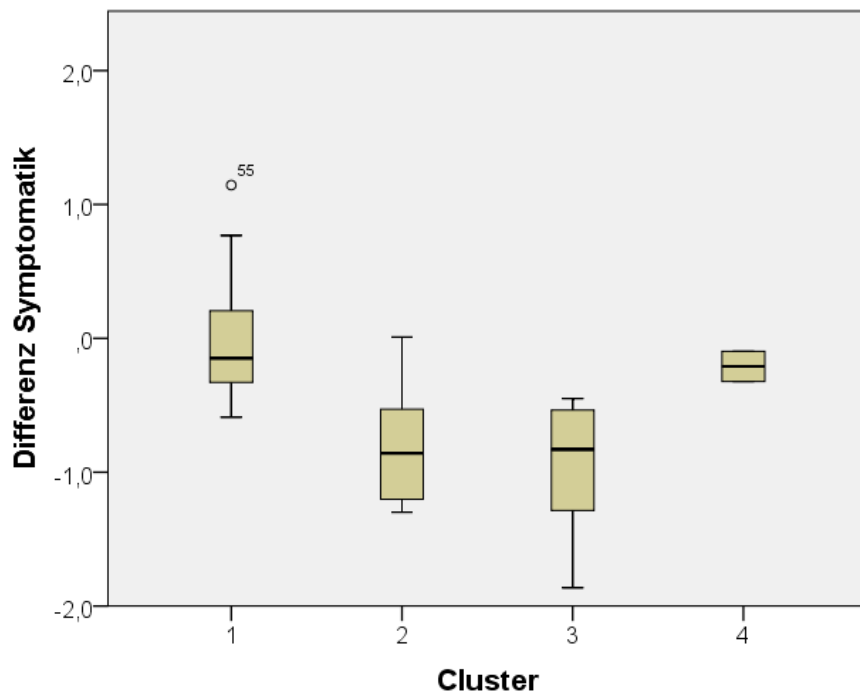
Tabelle 47: Hierarchische Clusteranalyse nach Differenzwerten der Problemmaße (Linkage innerhalb der Gruppen)

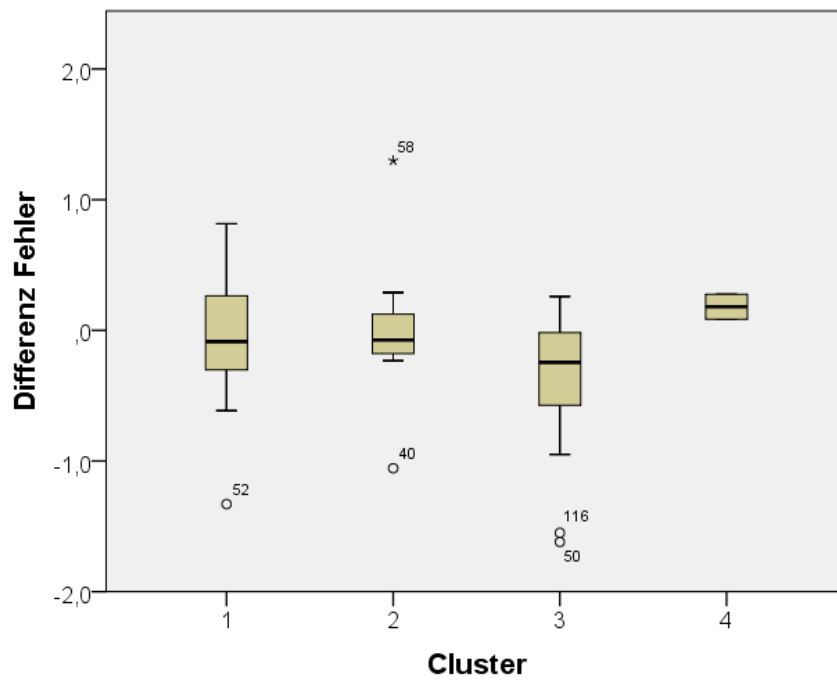
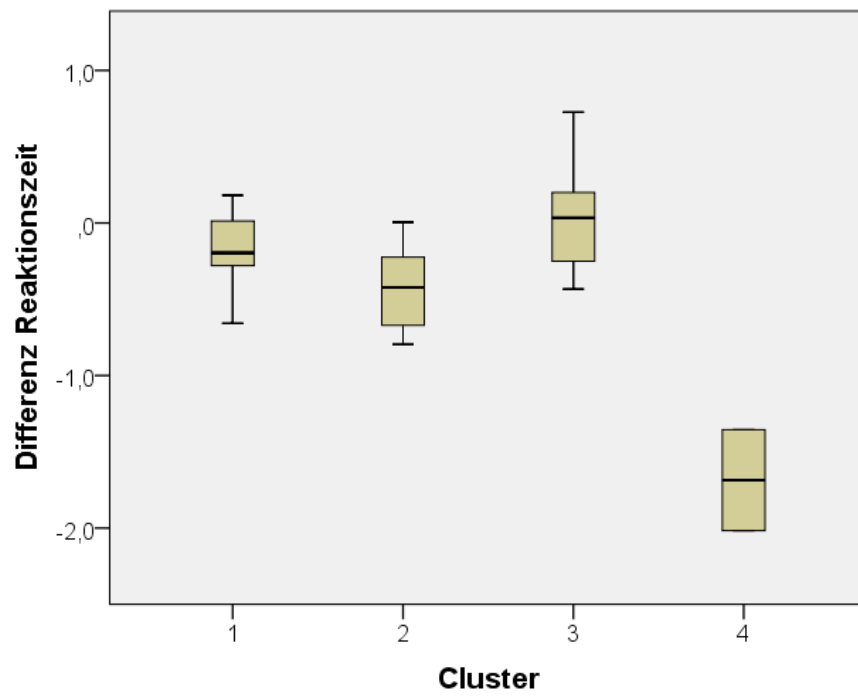
Differenz- werte ^a		Symptomatik	Lehrerurteil	Reaktionszeit	Fehler	Gesamt- problematik
1 (n = 17)	M (SD)	.00 (.49)	.25 (.57)	-.17 (.21)	-.08 (.47)	.00 (.19)
2 (n = 14)	M (SD)	-.81 (.43)	-1.13 (.82)	-.40 (.27)	-.01 (.49)	-.59 (.21)
3 (n = 17)	M (SD)	-.95 (.47)	-.50 (.62)	.04 (.35)	-.39 (.55)	-.45 (.33)
4 (n = 2)	M (SD)	-.21 (.16)	.11 (.24)	-1.69 (.47)	.18 (.13)	-.40 (.13)
Total (N = 50 ^b)	M (SD)	-.56 (.62)	-.40 (.85)	-.22 (.45)	-.16 (.52)	-.34 (.35)

Anmerkungen: ^aNegative Werte bedeuten eine Verbesserung. ^bEs werden nur Teilnehmer mit Differenzwerten in allen vier Problemmaße berücksichtigt

In der abgebildeten Lösung werden die Teilnehmer vier Clustern zugeordnet. Die ersten drei Cluster haben mit $n = 14$ bzw. $n = 17$ ähnliche Gruppengrößen, während sich in Cluster 4 nur zwei Teilnehmer befinden. Im Vergleich der Gesamtproblematik wird deutlich, dass in Cluster 1 die Teilnehmer mit den geringsten Veränderungen zusammengefasst sind. In Cluster 2 befinden sich hingegen die Teilnehmer mit den stärksten Verbesserungen. Teilnehmer in Cluster 3 und Cluster 4 zeigten im Mittel moderate Verbesserungen.

Eine genauere Inspektion der Cluster ermöglicht die Darstellung in Form von Boxplots, die in Abbildung 30 dargestellt sind.





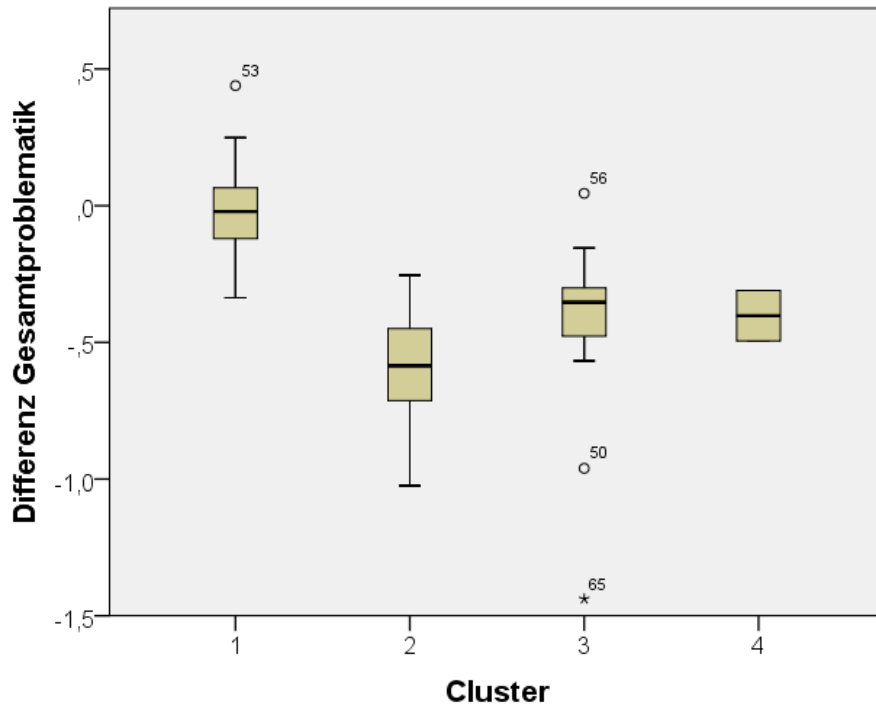


Abbildung 30: Darstellung der zu Clustern zusammengefassten Veränderungsmaße der Teilnehmer in Form von Boxplots

Schließlich werden die Mittelwerte der Cluster in den relevanten Personen- und Durchführungsvariablen sowie der Einschätz- und Testmaße zum Prämesszeitpunkt gegenübergestellt. So kann deskriptiv der Brückenschlag vom „Output“ zum „Input“ bzw. relevanten Variablen, welche die Gruppenzuordnung erklären könnten, ermöglicht werden. Die Auswahl der abgebildeten Variablen ist an dieser Stelle möglichst breit und geht über die Auswahl in den vorangegangenen Auswertungsschritten hinaus. An dieser Stelle muss keine geringe Variablenzahl wie beispielsweise bei der Regressionsanalyse eingehalten werden. Im Sinne einer Hypothesengenerierung wird eine möglichst breite Inspektion verschiedener Aspekte angestrebt. Tabelle 48 zeigt die deskriptiven Ergebnisse der Cluster für dichotome Variablen, Tabelle 49 für die metrischen Variablen mit Mittelwerten und Standardabweichungen.

Tabelle 48: Deskriptive Ergebnisse für die Cluster (dichotome Variablen)

	Cluster				
	1 (n = 17)	2 (n = 14)	3 (n = 17)	4 (n = 2)	Total (N = 50)
N Mädchen (%)	2 (12 %)	3 (21 %)	3 (18 %)	0 (0 %)	8 (16 %)
N keine Medikation (%)	11 (65 %)	5 (36 %)	10 (59 %)	1 (50 %)	27 (54 %)

Tabelle 49: Deskriptive Ergebnisse für die Cluster (metrische Variablen)

Variablen (Prämesszeitpunkt)	Cluster				
	1 (n = 17)	2 (n = 14)	3 (n = 17)	4 (n = 2)	Total (N = 50)
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Alter in Monaten	167 (16)	169 (15)	180 (18)	164 (21)	172 (17)
Anzahl der Sitzungen	18 (2)	17 (2)	18 (2)	20 (0)	18 (2)
Trainingsdauer in Monaten	5 (2)	7 (2)	6 (1)	7 (1)	6 (2)
DISYPS-II gesamt Eltern	1.46 (0.58)	1.73 (0.49)	1.57 (0.54)	1.65 (0.50)	1.58 (0.53)
DISYPS-II gesamt Lehrer	1.16 (0.61)	1.44 (0.47)	0.92 (0.67)	1.50 (0.28)	1.17 (0.61)
DISYPS-II gesamt selbst	0.79 (0.39)	1.49 (0.59)	1.17 (0.60)	1.15 (0.92)	1.13 (0.60)
Hyperakt.-Impuls. Eltern	1.15 (0.67)	1.32 (0.67)	1.11 (0.78)	1.50 (0.64)	1.20 (0.70)
Hyperakt.-Impuls. Lehrer	0.88 (0.74)	0.83 (0.69)	0.56 (0.76)	1.30 (0.49)	0.77 (0.73)
Hyperakt.-Impuls. selbst	0.66 (0.42)	1.23 (0.60)	0.89 (0.73)	1.23 (1.10)	0.92 (0.64)
SDQ gesamt Eltern	18.18 (5.00)	21.21 (6.10)	18.18 (5.74)	21.00 (12.73)	19.14 (5.84)
SDQ gesamt Lehrer	12.41 (5.28)	17.86 (7.00)	11.00 (4.73)	16.00 (2.83)	13.60 (6.15)
SDQ gesamt selbst	13.35 (6.60)	18.50 (7.22)	16.35 (5.44)	19.50 (10.61)	16.06 (6.69)
SDQ Verhaltensprob. Eltern	3.47 (1.88)	4.36 (2.06)	3.37 (1.77)	3.00 (1.41)	3.70 (1.88)
SDQ Verhaltensprob. Lehrer	2.00 (1.62)	2.50 (2.53)	1.47 (1.77)	3.50 (0.71)	2.02 (1.96)
SDQ Verhaltensprob. selbst	2.65 (1.90)	3.64 (2.13)	2.94 (1.78)	3.50 (2.12)	3.06 (1.92)
SDQ Emot. Prob. Eltern	4.12 (2.34)	5.14 (2.85)	4.82 (2.56)	5.00 (5.66)	4.68 (2.63)
SDQ Emot. Prob. Lehrer	1.06 (1.09)	4.14 (2.71)	2.41 (1.70)	3.00 (4.24)	2.46 (2.27)
SDQ Emot. Prob. selbst	3.06 (2.75)	3.50 (2.62)	4.47 (2.72)	3.50 (3.54)	3.68 (2.71)
WLST Gesamtwert	59 (12)	56 (10)	57 (13)	53 (9)	57 (12)
WRI Gesamtwert	21 (9)	25 (9)	27 (12)	34 (2)	25 (10)
Regelkonformes Unterrichts- verhalten	76 (21)	68 (21)	76 (21)	66 (9)	73 (20)
Lern- und Aufmerksamkeits- verhalten	60 (21)	41 (19)	68 (19)	67 (10)	58 (22)
Arbeits- und Lernverhalten ge- samt Lehrer	66 (18)	51 (17)	71 (18)	67 (3)	63 (19)
TAP Alertness Reaktionszeit	250 (43)	254 (33)	249 (35)	352 (103)	255 (44)
TAP Flexibilität Reaktionszeit	865 (319)	867 (220)	822 (324)	1450 (700)	874 (324)
TAP Flexibilität Fehleranzahl	8 (5)	8 (8)	10 (9)	9 (2)	9 (7)
TAP Go/NoGo Reaktionszeit	392 (93)	420 (111)	370 (85)	560 (31)	399 (99)
TAP Go/NoGo Fehleranzahl	5 (3)	4 (5)	5 (3)	4 (2)	5 (4)

Die Interpretation der Mittelwertvergleiche erfolgt im Kapitel Diskussion. Aufgrund der geringen Gruppengröße wird keine probeweise Signifikanzprüfung der Mittelwertunterschiede vorgenommen.

6.6 Weiterführende Auswertung: Zusammenhänge der Ausgangswerte

An dieser Stelle folgt ein Exkurs zum Zusammenhang der initialen Störungsschwere bzw. den Werten zum Prämesszeitpunkt mit den Differenzwerten, da die bisherige Auswertung gezeigt hat, dass den initialen Werten eine starke Bedeutung zukommt. Es besteht eine negative Korrelation von initialen und Differenzwerten. Um zu unterscheiden, ob es sich um einen Zusammenhang der wahren Werte handelt und um von statistischen Schwierigkeiten wie der Regression zur Mitte abzugrenzen, werden die gefundenen Korrelationen mit denen der Wartekontrollgruppe aus der allgemeinen Wirksamkeitsüberprüfung verglichen (vgl. Schramm, im Druck). Die Wartegruppe ist zwar kleiner ($n = 34$) als die untersuchte Stichprobe, bietet aber trotzdem eine gute Schätzung der Effekte der wiederholten Messung bei der Zielgruppe. Einschränkend ist zu erwähnen, dass die Gruppen nicht völlig unabhängig sind. Einige der Teilnehmer (36 %) waren zunächst in der Wartegruppe und wurden im Anschluss mit dem Lerntraining behandelt. Ihre Daten vor und nach der Wartezeit gehen in die Berechnung der Kontrollgruppe ein. Für die vorliegende Untersuchung entsprechen die Daten nach der Wartezeit allerdings denen vor der Behandlung mit dem Lerntraining und gehen somit doppelt in die Auswertung ein. Trotzdem ermöglicht die Heranziehung der Wartegruppe eine gute Möglichkeit zur Schätzung der Effekte der Wiederholungsmessung.

Zum Vergleich der Korrelationen wird die Prüfgröße z aus einer Fisher- z -Transformation der Korrelationskoeffizienten unter Berücksichtigung der unterschiedlichen Gruppengrößen beider Stichproben nach folgenden Formeln bestimmt (Bühner & Ziegler, 2009, S. 612):

$$z = \frac{r'_1 - r'_2}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}}$$

$$r'_i = (0.5) \cdot \ln \left| \frac{1 + r_i}{1 - r_i} \right|$$

Die Effektstärke von Unterschieden von Korrelationen lässt sich nach Cohen (1988) mit folgender Formel bestimmen:

$$q = r'_1 - r'_2 = (0.5) \cdot \ln \left| \frac{1 + r_1}{1 - r_1} \right| - (0.5) \cdot \ln \left| \frac{1 + r_2}{1 - r_2} \right|$$

Cohen (1988) schlägt folgende Einordnung vor: $q = 0.10$ kennzeichnet einen geringen Unterschied, $q = 0.30$ einen moderaten Unterschied und $q = 0.50$ kennzeichnet einen großen Unterschied.

Tabelle 50 zeigt die Korrelationen für die beiden Gruppe, die Prüfgröße z und die Effektstärke q für den Unterschied der Korrelationen.

Tabelle 50: Vergleich der Zusammenhänge der Ausgangswerte (Prämesszeitpunkt) und der Differenzwerte

	LeJA^a	WKG^b	z	q
Gesamtproblematik	$r = -.46$ $R^2 = .21 (N = 67)$	$r = -.23$ $R^2 = .05 (N = 34)$	1.20	.26
Symptomatik	$r = -.41$ $R^2 = .17 (N = 66^c)$	$r = -.42$ $R^2 = .18 (N = 34)$	0.07	0.01
Lehrereinschätzung	$r = -.54$ $R^2 = .29 (N = 52)$	$r = -.09$ $R^2 = .01 (N = 29)$	2.11	0.51
Reaktionszeit	$r = -.65$ $R^2 = .43 (N = 66)$	$r = -.26$ $R^2 = .07 (N = 34)$	2.36	0.51
Fehler	$r = -.51$ $R^2 = .26 (N = 66)$	$r = -.57$ $R^2 = .33 (N = 34)$	0.43	0.08

Anmerkungen: ^aLerntraining. ^bWartekontrollgruppe. ^cUnterschiedliche Stichprobengrößen innerhalb einer Gruppe resultieren aus fehlenden Werten.

Für die Gesamtproblematik zeigt sich mit $q = 0.26$ ein geringer bis moderater Unterschied in den Korrelationen. Bei genauerer Analyse der Faktoren, aus denen der Gesamtproblemwert gebildet wird, wird deutlich, wie dieser Effekt zu Stande kommt. Für die Problemmaße Symptomatik und Fehler zeigen sich praktisch gar keine Unterschiede. Dafür ergeben sich in den Maßen Lehrereinschätzung und Reaktionszeit große Unterschiede mit jeweils $q = 0.51$.

7 Diskussion

In diesem Kapitel werden die Befunde der unterschiedlichen Auswertungsschritte interpretiert. Die Diskussion ist nach den Merkmalen der Personen und der Durchführung geordnet und fasst Befunde der unterschiedlichen Auswertungsschritte zusammen. Dadurch entsteht ein möglichst vollständiges Bild und es wird eine abschließende Bewertung der Relevanz des jeweiligen Aspektes möglich. Zudem werden die Methoden einer kritischen Würdigung unterzogen.

7.1 Interpretation der Befunde

Wie hängen Merkmale der Teilnehmer und der Trainingsdurchführung mit dem Trainingserfolg zusammen?

Die Darstellung der Befunde ist nach den Merkmalen der Teilnehmer und der Durchführung geordnet. Zu jedem Merkmal werden die Ergebnisse mehrerer Auswertungsschritte herangezogen, um so eine möglichst umfassende Einschätzung der Relevanz des jeweiligen Aspektes vornehmen zu können. Die Ergebnisse der personenzentrierten Auswertungen werden dabei mit einbezogen und am Ende des Abschnitts noch einmal einzeln ausführlich dargestellt.

Ein Problem für die Vergleichbarkeit der unterschiedlichen Analysen und der Generalisierbarkeit ist die unterschiedliche Stichprobengröße in den unterschiedlichen Auswertungen. Diese ist auf fehlende Werte zurückzuführen, die für eine empirische Untersuchung typisch sind. Fehlende Werte liegen insbesondere in der Lehrereinschätzung vor. Im Vergleich von Teilnehmern mit vorliegenden und fehlenden Lehrereinschätzungen zeigen sich allerdings keine signifikanten Unterschiede, wie zuvor berichtet wurde. In den meisten Auswertungsschritten konnten fehlende Werte durch die zusammengefassten Problemmaße ausgeglichen werden.

Bei der variablenzentrierten multivariaten Auswertung können allerdings nur vollständige Datensätze auf Ebene der Problemmaße verwendet werden. Die Stichprobengröße in dieser Auswertung beträgt mit $N = 49$ nur 73 % der Gesamtstichprobe. In diesem Auswertungsschritt wurden aber keine zusätzlichen Effekte festgestellt, sondern die Ergebnisse decken sich mit denen der beiden vorangegangenen Auswertungsschritten. Keine der Interpretationen beruft sich allein auf die variablenzentrierte multivariate Auswertung. Daher ist die geringere Fallzahl an dieser Stelle für die weitere Argumentation irrelevant. Dass diese Auswertung keinen weiteren Informationsgewinn im Vergleich zu den vorangegangenen Analysen liefert, liegt unter anderem daran, dass die abhängigen Variablen, also die vier Problemmaße, kaum miteinander korrelieren, was eine Voraussetzung für die multivariate Auswertung ist. Die geringen Korrelationen rühren daher, dass die Problemmaße nach einer Faktorenanalyse zusammengestellt worden sind. Ziel der Faktorenanalyse ist es, solche Variablen zusammenzufassen, die untereinander korrelieren, und solche Variablen zu trennen, die nicht miteinander korre-

lieren. Das Ergebnis sind Faktoren, die wenig miteinander korrelieren. Es wäre nicht empfehlenswert gewesen, alle 14 Outcome-Maße als abhängige Variablen in die Analyse einzubeziehen, da nicht zu viele abhängigen Variablen ausgewählt werden sollen (Tabachnick & Fidell, 2007). Dies hat einerseits methodische Gründe und andererseits wäre die Interpretation schwierig, wenn sich uneinheitliche Befunde ergäben.

Schließlich ist anzumerken, dass das Vorgehen der schrittweisen Regression kritisiert wird, da die Auswahl der unabhängigen Variablen aufgrund eines statistischen Kriteriums vorgenommen wird (Bühner & Ziegler, 2009). Ob ein Prädiktor aufgenommen wird, hängt davon ab, ob er signifikant Varianz aufklären kann. Die Signifikanz hängt wiederum mit der Stichprobengröße zusammen. Wünschenswert ist, dass inhaltliche Argumente über die Auswahl der Prädiktoren entscheiden. Dies ist in der vorliegenden Untersuchung auch der Fall, da nur solche Prädiktoren gewählt wurden, die nach theoretischen Überlegungen und empirischen Erkenntnissen möglicherweise Varianz aufklären können. Nur über die Reihenfolge bzw. die Zusammenhänge der Prädiktoren untereinander bestehen aufgrund der Komplexität möglicher Verbindungen keine Hypothesen. Das schrittweise Vorgehen bietet den Vorteil, dass Variablen mit demselben Informationsgehalt identifiziert werden, indem nur Variablen in die Regressionsgleichung aufgenommen werden, die zusätzliche Varianz aufklären. Weiterführende Untersuchungen können elaboriertere Modelle überprüfen. In dieser Auswertung ist zunächst nur eine explorative Analyse möglich.

Geschlecht

Es finden sich keine Hinweise auf einen moderierenden Effekt des Geschlechtes. Die Unterschiede in den Veränderungen im Training erfasst über die Differenzwerte sind gering. Vor Beginn des Trainings sind alle Problemmaße in der Gruppe der Mädchen stärker ausgeprägt als in der Gruppe der Jungen. Dies ist insbesondere darauf zurückzuführen, dass bei den Mädchen im TAP-Test Flexibilität Ausreißerwerte vorliegen und dass die Mädchen in der Selbsteinschätzung eine stärkere Symptomatik angeben. Der Unterschied zum Prämesszeitpunkt ist aber nicht so groß, dass er statistische Signifikanz erreicht. Dies kann aber auch mit der geringen Stichprobengröße der Gruppe der Mädchen zusammenhängen. In den Regressionsanalysen hat das Geschlecht keinen signifikanten Einfluss und wird durch das schrittweise Verfahren in jeder der Auswertungen aus dem Modell ausgeschlossen. Auch im personenzentrierten Vorgehen zeigt sich eine ähnliche Verteilung der Mädchen auf die Cluster bzw. Einteilung in Responder und Non-Responder wie für die Jungen.

Bei der Einschätzung des Aspekts des Geschlechts ist zu berücksichtigen, dass die Gruppe der Mädchen mit $n = 13$ klein ist. Dies kann dazu beigetragen haben, dass Mittelwertunterschiede zwischen Jungen und Mädchen nicht signifikant werden. Je größer eine Stichprobe ist, desto eher werden Effekte signifikant (Bortz & Schuster, 2010). Da es sich um eine Post-hoc-Auswertung handelt, ist der Stichprobenumfang nicht vor der Durchführung der Analysen festgelegt worden. Daher kann es sein, dass die Auswer-

tung für diesen Aspekt „underpowered“ ist, d. h. die Teststärke ist zu gering, um einen Effekt zu finden (Cohen, 1988). Die Werte der Mädchen und Jungen unterscheiden sich insbesondere zum Prämesszeitpunkt, weniger in den Differenzwerten. Die Unterschiede zum Prämesszeitpunkt könnten mit einer größeren Stichprobe oder Anhebung der Irrtumswahrscheinlichkeit signifikant werden, bei den Differenzwerten wäre dies fraglich. Trotzdem können die Ergebnisse nicht generalisiert werden. Dafür wäre eine größere Anzahl von Mädchen in der Stichprobe notwendig.

Medikation

Ob die teilnehmenden Jugendlichen parallel zum Training eine Medikation bekommen oder nicht, hat bezogen auf das Gesamtproblemmaß keinen moderieren Einfluss auf den Trainingserfolg. Die Differenzwerte unterscheiden sich nicht signifikant voneinander. Auch in der multiplen Regression mit der Gesamtproblematik als abhängiger Variable und der multivariaten Regression zeigt sich kein signifikanter Effekt. In den personen-zentrierten Vorgehensweisen zeigen sich keine Besonderheiten hinsichtlich der Medikation.

Die Medikation scheint jedoch bezogen auf die Lehrereinschätzung und die Reaktionszeit relevant zu sein. Es zeigt sich im Vergleich der Werte zum Prämesszeitpunkt, dass die Lehrer die Jugendlichen mit gleichzeitiger Medikation als weniger problembelastet einschätzen. Auch für die Lehrereinschätzung zum Postmesszeitpunkt zeigt sich in der multiplen Regression, dass die Einschätzung geringer ist für Jugendliche, die eine Medikation bekommen. Dies ist nicht damit zu erklären, dass die Lehrereinschätzung für diese Jugendlichen schon zum Prämesszeitpunkt geringer ausgeprägt ist, da die Lehrereinschätzung zum Prämesszeitpunkt als erste unabhängige Variable in das Regressionsmodell eingeht. Dass die Medikation als zweite Variable in das Modell aufgenommen wird, weist auf einen zusätzlichen Informationsgehalt hin. Kein signifikanter Zusammenhang zeigt sich hingegen mit der Differenz der Lehrereinschätzungen in der Regressionsanalyse. Die Signifikanz wird mit $p = .08$ knapp verfehlt (siehe Anhang). Die Jugendlichen mit einer Medikation haben sich tendenziell in der Lehrereinschätzung etwas stärker verbessert.

Das Problemmaß Symptomatik besteht aus den Einschätzungen der Jugendlichen und ihrer Eltern. Dass diese keinen Einfluss der Medikation feststellen, die Lehrer aber schon, mag mit dem Zeitpunkt der Einnahme und der Dauer der Wirkung zusammenhängen. Die Jugendlichen nehmen morgens ihre Medikamente ein, die so eingestellt sind, dass sie möglichst so lange wirken, bis die Schule vorbei ist. Wenn die Schüler mittags oder am Nachmittag wieder nachhause kommen, lässt die Wirkung der Medikation nach oder ist schon beendet. Daher ist die Lehrereinschätzung als überaus relevant zu bezeichnen, da die Lehrer die Schüler dann erleben, wenn die Medikamente wirken.

Die Ergebnisse weisen darauf hin, dass das Lerntraining bei gleichzeitiger Medikation in der Lehrereinschätzung etwas erfolgreicher war. Dies kann unter Bezugnahme auf das Modell von Rapport und Kollegen (2001) interpretiert werden. Die Medikation hat

eine Wirkung auf das der Störung zu Grunde liegende biologische Substrat. Dieses bildet eine Grundlage für die Ausgestaltung der Symptomatik. Durch die kognitiv-behaviorale Behandlung werden die Symptome Unaufmerksamkeit, Hyperaktivität und Impulsivität behandelt und zusätzlich komorbide Probleme, z. B. im Sozialverhalten. Die Kombination aus Medikation und Lerntraining setzt somit an mehreren Stellen an.

Alternativ kann vermutet werden, dass Erwartungen der Lehrer dazu beitragen, dass sie die Verbesserungen der Schüler ohne Medikation als geringer einschätzen. Die Erwartung mag so aussehen, dass die Kombinationsbehandlung als am wirkungsvollsten eingeschätzt wird. Wenn diese vorliegt, so erwarten die Lehrer, dass eine stärkere Verbesserung eintritt. Bei Schülern, die nur mit dem Lerntraining behandelt werden, könnten die Lehrer die Einschätzung haben, dass das reine Lerntraining nicht ausreichend sei, um die Probleme zu beheben, und eine zusätzliche Medikation erforderlich sei. Einige internationale Studien belegen, dass Lehrer Kombinationsbehandlungen bei ADHS als effektiver als rein medikamentöse oder rein verhaltensbezogene Interventionen einschätzen (Glass & Weigar, 2000; Havey, 2007).

Ein ähnlicher Effekt zeigt sich für die Reaktionszeit. Die Jugendlichen mit Medikation haben zum Postmesszeitpunkt bessere Reaktionszeiten und auch Ihre Verbesserung in den Reaktionszeiten ist stärker als für die Jugendlichen ohne Medikation. Es handelt sich dabei um kleine Effekte. Eine Interpretation dieses Effekts fällt schwer. Es geht nicht darum zu begründen, dass die Jugendlichen mit einer Medikation kürzere Reaktionszeiten haben als Jugendliche ohne Medikation, sondern warum die Verbesserung der Reaktionszeiten bei Ersteren stärker ausgeprägt ist. Es ist festzustellen, dass die Messungen der Reaktionszeit am Nachmittag erfolgten. Es ist fraglich, ob zu diesem Zeitpunkt überhaupt noch eine Wirkung des Medikaments vorhanden war. Da es sich um Ergebnisse von Aufmerksamkeitstests und keine subjektiven Einschätzungen handelt, sind Erwartungseffekte auszuschließen. Eine Deutung des Effektes kann an dieser Stelle nicht vorgenommen werden.

Es ist festzuhalten, dass der Status der Medikation in der Gesamteinschätzung des Trainings unerheblich ist. Jugendliche mit und ohne eine stabile Einstellung auf Medikamente profitieren gleichsam vom Training. Für die Lehrereinschätzungen scheint die Medikation allerdings eine Rolle zu spielen. Tendenziell schätzen Lehrer die Wirksamkeit der Kombinationsbehandlung etwas höher ein als ein Lerntraining ohne gleichzeitige Medikation. Dies deckt sich mit Ergebnissen empirischer Studien zur Erwartung der Wirksamkeit von Interventionen bei ADHS durch Lehrer.

ADHS-Subtypus

Erwartungsgemäß korrelieren alle Einschätzungen der Hyperaktivität und Impulsivität mit den Problemmaßen Symptomatik, Lehrereinschätzung und der Gesamtproblematik, da alle diese Maße Aspekte der Symptomatik erfassen. Zusammenhänge mit den Veränderungswerten sind jedoch nur gering bis gar nicht vorhanden. In der einfachen univariaten Auswertung zeigt sich eine signifikante Korrelation zwischen der Hyperaktivi-

tät-Impulsivität in der Selbsteinschätzung und den Differenzen in der Symptomatik. Jugendliche, die sich selbst als stärker überaktiv und impulsiv einschätzen, haben eine etwas stärkere Reduktion der Symptomatik. Es handelt sich dabei um einen schwachen Zusammenhang. In den weiteren Regressionsanalysen zeigt sich kein signifikanter Effekt. Vermutlich hängt dies mit der hohen Korrelationen mit den Problemmaßen zum Prämesszeitpunkt zusammen. Die initialen Werte der Problemmaße gehen jeweils als erste unabhängige Variablen in die Regressionsanalysen ein und die Hyperaktivität und Impulsivität kann keine zusätzliche Varianz mehr aufklären.

Im Vergleich von Respondern und Non-Respondern zeigt sich, dass in der Gruppe der Responder deutlich höhere Werte der Hyperaktivität-Impulsivität angegeben werden als in der Gruppe der Non-Responder. Dieses gilt ebenso für die Gesamteinschätzung im DISYPS-II und den SDQ. Es scheint sich um einen Zusammenhang mit den initialen Werten zu handeln und nicht um einen auf die Hyperaktivität-Impulsivität beschränkten Effekt.

Zusammenfassend ist zu sagen, dass sich kein Hinweis darauf finden lässt, dass das Training für einen der ADHS-Subtypen besonders geeignet oder ungeeignet ist.

Alter

Die Hypothese, dass ältere Teilnehmer stärker vom Training profitieren, wird nicht bestätigt. Es scheint keinen Zusammenhang zwischen Alter und Trainingserfolg zu geben. Deskriptiv im Streudiagramm und auch nach Gruppenbildung finden sich keinerlei Hinweise auf einen Zusammenhang, auch nicht auf nicht-lineare Zusammenhänge. Auch in den weiteren Regressionsanalysen trägt das Alter der Teilnehmer nicht signifikant zur Varianzaufklärung bei. In den personenzentrierten Auswertungen ist das Alter ähnlich über die Gruppen verteilt.

Im Rückblick auf die zur Ableitung der Hypothese herangezogenen Studien fällt auf, dass diese sich vor allem auf das frühe Jugendalter beziehen. So ist bei Durlak et. al. (1991) die älteste Altersgruppe die der 11- bis 13-jährigen und das Durchschnittsalter in der Studie von McCart et al. (2006) lag bei elf Jahren. Lediglich bei Dush et al. (1989) wurden auch 14- bis 16-jährige untersucht. Die Alterstrends werden folglich in der Literatur vor allem für jüngere Jugendliche bzw. Kinder gefunden. Es mag sein, dass die Unterschiede der kognitiven Fähigkeiten zwischen Jugendlichen unterschiedlichen Alters im Vergleich zu den Unterschieden zwischen Jugendlichen und Kindern eher zu vernachlässigen sind.

Weiterhin fällt bei den zitierten Studien auf, dass die größten Effekte für die Gruppen mit den wenigsten Studien bzw. Einzeleffekten gefunden werden. Vermutlich handelt es sich beim postulierten Zusammenhang von Alter und Interventionserfolg zumindest teilweise um ein Artefakt. In Untersuchungen mit kleinen Stichprobengrößen werden häufiger große Effektstärken berichtet, u. a. aufgrund eines „publication bias“ (Slavin & Smith, 2008). Das bedeutet, dass Studien, die signifikante Effekte feststellen, häufiger

publiziert werden als Studien, die keine Effekte dokumentieren. Dies wird bildhaft auch als „file drawer problem“ (Rosenthal, 1979) bezeichnet. Untersuchungen ohne signifikante Ergebnisse bleiben dabei „in der Schublade“ und können nicht in Fachzeitschriften veröffentlicht werden. Somit gehen sie auch nicht in zusammenfassende Metaanalysen und Reviews ein. Bei Untersuchungen mit kleinen Stichprobengrößen ist die Generalisierbarkeit der Ergebnisse eingeschränkt, da nicht sichergestellt werden kann, dass es sich um eine repräsentative Stichprobe handelt. Es besteht das Risiko, dass nur bestimmte Teilnehmer aus der Grundgesamtheit einbezogen wurden. Es gibt insgesamt weniger Evaluationsforschung zu Interventionen mit Jugendlichen als mit Kindern (Beelmann & Schneider, 2003). Daher mag sich der „publication bias“ bei der Evaluationsforschung im Jugendalter stärker auswirken. Kleinere Studien mit Jugendlichen, die signifikante Effekte feststellen, werden vermutlich noch veröffentlicht, während kleinere Studien ohne signifikante Effekte vermutlich eher fallen gelassen werden. Ob ein Effekt Signifikanz erreicht, hängt neben der Größe des Effekts und der gewählten Irrtumswahrscheinlichkeit von der Stichprobengröße ab. Effekte werden eher signifikant bei größeren Stichproben. Somit werden bei den größer angelegten Evaluationsstudien mit Kindern eher Effekte erreicht, die allerdings tendenziell kleiner sind. Aufgrund der kleinen Stichproben in Untersuchungen mit Jugendlichen ist es schwerer, signifikante Effekte festzustellen. Daher greift das „Schubladenproblem“ hier vermutlich stärker und es werden nur solche Studien veröffentlicht, die größere Effekte feststellen. Zusammengefasst heißt das, dass im Kindesbereich mehr signifikante aber kleinere Effektstärken festgestellt werden. Demgegenüber werden im Jugendbereich selektiv die Studien publiziert, die größere Effekte feststellen, was mit kleineren und damit nicht immer repräsentativen Stichproben zusammenhängen mag.

In der Stichprobe der vorliegenden Untersuchung befinden sich mehr jüngere als ältere Jugendliche. So sind nur zehn Jugendliche 16 oder 17 Jahre alt, was lediglich 15 % der gesamten Stichprobe entspricht. Bei einer Wiederholung oder ähnlichen Untersuchung sollte darauf geachtet werden, dass eine gleichmäßige Verteilung über die Altersgruppen umgesetzt wird.

Es ist plausibel, dass die einzelnen Trainingsverläufe jeweils individuell und altersentsprechend ausgestaltet werden. In einer weiterführenden inhaltsanalytischen Auswertung könnte eruiert werden, ob in den Trainings mit jüngeren Kindern andere Trainingsinhalte und therapeutische Techniken im Vordergrund stehen als in Trainings mit älteren Jugendlichen. Somit wird über die individuelle Anpassung der Trainingskonzeption an den jeweiligen Jugendlichen schon eine Anpassung an das Alter vorgenommen.

Schließlich stellt sich wie bereits beschrieben die Frage, wie gut das chronologische Alter als ein Indikator der kognitiven Reife geeignet ist. Es kann sein, dass die interindividuelle Variabilität in der kognitiven Entwicklung im Jugendalter so groß ist, dass die Zusammenhänge mit dem Alter sehr gering ausfallen. Da in der vorliegenden Untersuchung kein anderes Maß der kognitiven Reife erfasst wurde, kann die Übereinstimmung unterschiedlicher Verfahren nicht überprüft werden. In nachfolgende Untersu-

chungen sollten alternative Messungen kognitiver Fertigkeiten einbezogen werden. Infrage kommen dazu IQ-Messungen und Verfahren, die Aspekte wie das Entscheidungsverhalten, Antizipation von Handlungsfolgen, Perspektivübernahme oder soziale Informationsverarbeitung erfassen (Holmbeck, O'Mahar, Abad, Colder & Updegrave, 2006).

Zusammengefasst kann gesagt werden, dass sich das Lerntraining für Jugendliche mit ADHS zur Anwendung über die gesamte Altersspanne von 12 bis 17 Jahren eignet.

Initiale Störungsschwere

Die Hypothese, dass es einen Zusammenhang zwischen der initialen Störungsschwere und dem Trainingserfolg gibt, bestätigt sich. Teilnehmer mit einer stärkeren Problematik vor dem Training erreichen eine größere Reduktion ihrer Problematik. Die Stärke des Zusammenhangs ist mit $\tau = -.23$ bzw. $r = -.43$ nach Cohen (1988) als mittel hoch zu bezeichnen.

Insgesamt gilt, dass die Werte zum Postmesszeitpunkt am besten durch die entsprechenden Werte des Prämesszeitpunktes vorhergesagt werden können. Teilnehmer mit höheren Werten vor dem Training haben auch nach dem Training höhere Werte. Mit Ausnahme der Lehrereinschätzung werden durch die Werte des Prämesszeitpunktes in allen Problembereichen mehr als 50 % der Varianz in den Werten zum Postmesszeitpunkt aufgeklärt. Der Zusammenhang ist am stärksten für das Problemmaß Reaktionszeit, bei dem 76 % der Varianz aufgeklärt wird. Die Reaktionszeiten scheinen somit besonders stabil zu sein. Weitere in die Regressionsgleichung aufgenommene Variablen können nur wenig zusätzliche Varianz aufklären. Am geringsten ist die Varianzaufklärung durch die Werte der Prämessung für das Lehrerurteil mit 28 %.

Bei der Vorhersage der Differenzwerte zeigt sich für alle Bereiche ein negativer Zusammenhang zwischen den Ausgangswerten und der Veränderung. Teilnehmer mit größeren Problemwerten vor dem Training haben tendenziell einen größeren Trainingserfolg. Es liegt also ein kompensatorischer Effekt des Trainings vor, bei dem sich die Teilnehmer eher angleichen anstatt sich weiter voneinander zu entfernen wie bei einem Matthäus-Effekt. Auch die Betrachtung der Varianzen zum Prä- und Postmesszeitpunkt zeigt, dass die Varianzen insgesamt etwas geringer werden und somit kein Schereneffekt vorliegt, wie es bei einem Matthäus-Effekt zu erwarten wäre.

Die weiterführende Auswertung durch einen Vergleich mit der Wartekontrollgruppe der allgemeinen Wirksamkeitsprüfung zeigt, dass der Kompensationseffekt bezogen auf die Gesamtproblematik auch in der Wartekontrollbedingung vorliegt, dort allerdings weniger stark. Anhand der Wartekontrollgruppe wird also deutlich, dass es zwar einen Effekt der Messwiederholung gibt, dieser allerdings kleiner ist als der Effekt in der Interventionsgruppe. Die negative Korrelation zwischen initialen Werten und den Veränderungswerten in der LeJA-Gruppe ist damit vermutlich zusammengesetzt aus Effekten der Messwiederholung und eines kompensatorischen Effekts des Trainings.

Bezogen auf die vier Problembereiche wird deutlich, dass der kompensatorische Effekt in der Wartegruppe nur für die Symptomatik und die Fehler ähnlich hoch ist wie in der LeJA-Gruppe. In der Lehrereinschätzung, in der Reaktionszeit und insgesamt auch in der Gesamtproblematik liegt der Kompensationseffekt deutlich stärker in der Trainingsgruppe vor.

Es kann vermutet werden, dass der Kompensationseffekt für die Symptomatik damit zusammenhängt, dass Eltern und Jugendliche die Symptomatik und damit verbundenen Probleme zum ersten Messzeitpunkt als sehr stark einschätzen. Zu diesem Zeitpunkt wussten sie noch nicht, ob sie an dem Training teilnehmen können, da dieser Messzeitpunkt auch der diagnostischen Abklärung diene. Vermutlich sind die Einschätzungen deshalb eher drastischer als moderater ausgefallen, um den Bedarf an Unterstützung deutlich zu zeigen. Mit allen Teilnehmern der Wartegruppe, die dies wünschten, wurde nach der Wartezeit ein Training durchgeführt. Deshalb sind die Einschätzungen nach der Warteperiode vermutlich moderater, da klar ist, dass bald eine Förderung stattfinden wird. So mögen zum ersten Messzeitpunkt Erwägungen eine Rolle spielen, eher größere Problemwerte anzugeben, um aufgrund der stärkeren Bedürftigkeit in das Training aufgenommen zu werden. Die Einschätzung nach der Wartezeit bzw. direkt vor dem Training mag moderater ausfallen, da eine baldige Förderung und zu erwartende Verbesserung ansteht oder da bereits das bloße Gespräch vor der Wartezeit schon eine Verbesserung bewirkt hat.

Für die Einschätzungen der Lehrer spielten Erwägungen über die Aufnahme in das Training keine Rolle. Die Lehrer wurden erst nach dem ersten diagnostischen Gespräch mit den Eltern kontaktiert, wenn schon feststand, dass ein Training durchgeführt wird. Den Lehrern wurde mitgeteilt, dass ein Lerntraining mit dem jeweiligen Schüler stattfinden werde, aber nicht, wann genau es starten solle. Daher ist anzunehmen, dass die Lehrereinschätzungen bezogen auf Erwägungen, die bei den Eltern vorgelegen haben könnten, unverfälscht sind. Die Lehrer berichten eine stärkere Verbesserung für diejenigen Schüler mit einer größeren Problematik und das nur in der Interventionsbedingung, nicht aber in der Wartebedingung, zu der mit $q = .51$ diesbezüglich ein großer Unterschied vorliegt.

Der kompensatorische Effekt ist ebenso stark für das Problemmaß Reaktionszeit. Es scheint so zu sein, dass langsame Reaktionszeiten von den Teilnehmern der Trainingsgruppe am einfachsten zu verbessern waren. Warum in der Wartekontrollgruppe für dieses Problemmaß nur ein sehr schwacher Kompensationseffekt vorliegt, ist unklar. Die Fähigkeit schnell auf Reize zu reagieren und kurze Reaktionszeiten zu erreichen, wurde im Training nicht systematisch trainiert. Der kompensatorische Effekt im Problemmaß Fehler tritt dagegen sowohl in der Trainings- als auch in der Wartekontrollgruppe auf.

Zusammenfassend kann gesagt werden, dass sich in der vorliegenden Untersuchung keine Hinweise auf einen Matthäus-Effekt finden, sondern dass sich im Gegenteil ein

kompensatorischer Effekt abzeichnet. Jugendliche mit stärkeren Problematiken hatten größere Trainingserfolge. Dieser Befund deckt sich mit den Ergebnissen von Kazdin und Whitley (2006) sowie Hemphill und Littlefield (2006) zur Behandlung von Verhaltensstörungen und mit Diamond (2012) bezogen auf die Förderung von Exekutivfunktionen. Die weiterführende Auswertung durch den Vergleich mit der Wartekontrollgruppe der allgemeinen Wirksamkeitsprüfung unterstützt diese Interpretation. Der kompensatorische Effekt wird besonders deutlich im Lehrerurteil, was besonders wertzuschätzen ist, da die Einschätzungen der Lehrer im Gegensatz zu den Elterneinschätzungen frei von Erwägungen sind, die mit der Aufnahme in das Training zusammenhängen. Übereinstimmend kann mit Mohr (1995) eine als höhere eingeschätzte Störungsschwere durch die Jugendlichen selbst als ein Indiz für eine größere Veränderungsmotivation verstanden werden.

Komorbidität

Die Annahme, dass eine stärkere externalisierende Komorbidität mit geringeren Trainingserfolgen verbunden ist, hat sich nicht bestätigt. Es liegen im Gegenteil Hinweise vor, dass stärkere Verhaltensprobleme mit stärkeren Verbesserungen verbunden sind. Die Einschätzungen der Verhaltensprobleme korrelieren erwartungsgemäß positiv mit der Symptomatik, der Lehrereinschätzung und der Gesamtproblematik. Damit liegt die Vermutung nahe, dass es sich bei dem Zusammenhang zwischen stärkerer externalisierender Komorbidität und größeren Trainingserfolgen um denselben Zusammenhang wie bei der initialen Störungsschwere handelt. Dafür spricht auch, dass in den Regressionsanalysen die externalisierende Komorbidität keinen signifikanten Beitrag zur Varianzaufklärung liefern konnte, wenn die initialen Werte der jeweiligen abhängigen Variable bereits in die Regression eingegangen sind.

Es liegt eine einzige signifikante Korrelation zwischen den Selbsteinschätzungen und den Verbesserungen in der Lehrereinschätzung vor. Wenn die Jugendlichen selbst ihre Verhaltensprobleme als gravierender einschätzen, erreichen sie stärkere Verbesserungen in der Einschätzung ihrer Lehrer. Hierbei handelt es sich um einen interessanten Zusammenhang, der allerdings nicht überinterpretiert werden darf, da es sich um ein signifikantes Ergebnis aus einem probeweisen Signifikanztest post-hoc handelt. Es kann vermutet werden, dass die Jugendlichen, die sich eigene gravierende Verhaltensprobleme eingestehen, in stärkerem Maße dazu bereit sind, etwas zu verändern und dieses auch gelingt. Im Sinne von Mohr (1995) kann die Selbsteinschätzung hier als ein Indikator für die Motivation zur Mitarbeit und zur Veränderungen der Jugendlichen interpretiert werden. Interessant ist der Effekt auch deshalb, da es sich um einen Zusammenhang zwischen den voneinander unabhängigen Lehrer- und Selbsteinschätzungen handelt. Somit scheint dieser Effekt weniger stark mit der Fragestellung nach der Auswirkung der initialen Störungsschwere konfundiert zu sein. In einer etwaigen anschließenden Untersuchung sollte überprüft werden, ob dieser Effekt einer Hypothesenprüfung standhalten kann.

Wichtig ist es an dieser Stelle zu betonen, dass die Ergebnisse nicht dafür sprechen, dass die externalisierende Komorbidität nicht relevant sei. So beinhaltet auch die initiale Störungsschwere Aspekte der Komorbidität. Die SDQ-Gesamtwerte gehen auch in die Berechnung der Problemmaße ein. Die Ergebnisse sprechen dafür, dass je gravierender eine Problematik war, sowohl bezogen auf die Kernsymptomatik als auch auf komorbide Symptome, tendenziell stärkere Trainingserfolge erzielt werden können. Die Komplexitätshypothese wird eindeutig nicht bestätigt. Dieses Ergebnis zeigte sich ebenso in den zwei Studien von Kazdin und Whitley (2006) zur Behandlung bei einer Störung mit oppositionellem Trotzverhalten bzw. bei einer Störung des Sozialverhaltens. Auch hierbei handelt es sich um Completer-Analysen ohne Kontrollgruppe. Es zeigte sich für beide Störungsbilder, dass die Kinder mit mehr Komorbiditäten und einer stärkeren Störungsschwere die stärkeren Verbesserungen aufwiesen.

Die Hypothese dass eine stärkere internalisierende Problematik mit größeren Trainingserfolgen verbunden ist, hat sich für die Lehrereinschätzungen bestätigt, für die Eltern- und Selbsteinschätzung jedoch nicht. Wenn die Lehrer die emotionalen Probleme als gravierender einschätzen, sind die Verbesserungen etwas stärker. Auch in diesem Fall gilt, dass keine zusätzliche Varianz aufgeklärt werden konnte, wenn in die Regressionsanalysen zuvor die initialen Werte eingegangen waren, so dass auch dieses post-hoc signifikante Ergebnis nur vorsichtig interpretiert werden sollte. Interessant ist hier ebenfalls, dass die Lehrereinschätzung der Verhaltensprobleme nicht nur mit der Lehrereinschätzung der Verbesserung und der Gesamtproblematik zusammenhängt, sondern auch mit der unabhängigen Einschätzung der Symptomatik durch Eltern und Jugendliche. Die Schule ist für Jugendliche ein überaus wichtiger Lebensbereich. Probleme in der Schule können deshalb gravierende Auswirkungen auf die emotionale Befindlichkeit der Schüler haben. Dies trifft sowohl auf den Lern-Leistungsbereich als auch auf den Sozialbereich zu. So konnte beispielsweise Seiffge-Krenke (2008) zeigen, dass schulische Probleme bei Jugendlichen nach Zukunftsängsten den meisten Stress verursachen, sogar noch vor Problemen mit Eltern und romantischen Beziehungen. Lehrer können die emotionalen Probleme vermutlich gut einschätzen. Im Gegensatz zu den Eltern sind sie in den schulischen Situationen anwesend und im Gegensatz zu den Jugendlichen haben sie vermutlich eine objektivere Einschätzung. Da es sich um einen schambesetzten Bereich handelt und Jugendliche mit ADHS Verzerrungstendenzen bezüglich ihrer Selbsteinschätzung aufweisen, geben sie emotionale Probleme vermutlich weniger gern zu bzw. spielen diese herunter. Die Lehrer haben zudem gute Vergleichsmöglichkeiten mit Gleichaltrigen. Wenn man davon ausgeht, dass die Einschätzungen der emotionalen Befindlichkeit durch die Lehrer valide sind, dann spricht der gefundene Zusammenhang dafür, dass eine höhere internalisierende Problematik mit größeren Trainingserfolgen in Verbindung steht.

Dieses Ergebnis deckt sich mit den Ergebnissen der MTA-Studie nach Hinshaw (2007). Die Gruppe von Kindern mit komorbider Angststörung zeigte bessere Ergebnisse bei der behavioralen Behandlung als die übrigen Kinder. Hinshaw (2007) empfiehlt daher abweichend von den sonstigen Empfehlungen, die eine Medikation oder Kombinations-

behandlung als Intervention der Wahl favorisieren, bei dieser Subgruppe mit einer Behandlung ohne Medikation und reiner Verhaltenstherapie zu beginnen. Auch die Studie von Antshel, Faraone und Gordon (2012) kommt zu dem Ergebnis, dass eine internalisierende Komorbidität mit größeren Interventionserfolgen verbunden ist. Die Autoren verweisen auf Studien, die eine gute Wirksamkeit kognitiv-behavioraler Interventionen bei der Behandlung von Ängstlichkeit und Depression im Jugendalter dokumentieren. Die kognitiv-behaviorale Trainingskonzeption eignet sich vermutlich gut für die Behandlung von Jugendlichen, die komorbide internalisierende Symptome aufweisen.

Abschließend ist daran zu erinnern, dass die Komorbidität über die jeweiligen SDQ-Subskalen erfasst wurde. Dies entspricht der Erfassung von Symptomen der Komorbiditäten, aber nicht kompletten komorbid vorliegenden Störungsbildern. Die gefundenen Zusammenhänge sind folglich Zusammenhänge von Symptomausprägungen. Dies entspricht stärker der Vorstellung kontinuierlich verteilter psychischer Konditionen, die sich phänotypisch unterschiedlich äußern (Rommelse et al., 2008), und damit den Annahmen der Entwicklungspsychopathologie (Sroufe, 1997).

Trainingsdauer

Die Hypothese, dass kürzere Trainings mit einem größeren Trainingserfolg verbunden sind, konnte nicht bestätigt werden. Die weiterführenden Auswertungen weisen auf einen entgegengesetzten Zusammenhang hin. Längere Trainingsverläufe sind tendenziell erfolgreicher.

Es gilt jedoch auch, dass eine schwerer wiegende Problematik mit einer längeren Trainingsdauer zusammenhängt. Es ist zu vermuten, dass nicht die längere Dauer der Trainings, sondern die damit verbunden höhere initiale Störungsschwere mit den stärkeren Verbesserungen zusammen hängt. Dass diese Annahme plausibel ist, wird in den multiplen Regressionsanalysen bestätigt. Nach Aufnahme der initialen Störungsschwere in die Regressionsgleichung klärt die Dauer des Trainings keine zusätzliche Varianz in der Veränderung der Gesamtproblematik mehr auf. Lediglich bei der Vorhersage der Verbesserung der Reaktionszeiten liefert die Trainingsdauer einen kleinen Beitrag von 4 % zusätzlicher Varianzaufklärung, wenn die Reaktionszeit zum Prämesszeitpunkt und die Medikation als unabhängige Variablen bereits berücksichtigt sind. Dabei gilt, dass Teilnehmer mit einem längeren Training ihre Reaktionszeiten stärker verbessern. Insgesamt deuten die Ergebnisse darauf hin, dass der Zusammenhang der Trainingsdauer mit dem Trainingserfolg über die gleichzeitige Korrelation mit der initialen Störungsschwere erklärbar ist. Dafür spricht weiterhin, dass die Anzahl der Sitzungen, die nicht mit den Werten des Prämesszeitpunktes korreliert, auch nicht mit den Differenzwerten korreliert. Dieser Befund deckt sich mit den Ergebnissen aktueller Metaanalysen (Fabiano et al., 2009; Hodgson et al., 2012; van der Oord et al., 2008).

Anzahl der Sitzungen

Die Hypothese, dass Teilnehmer mit weniger Sitzungen größere Trainingserfolge haben, konnte nicht bestätigt werden. Es scheint keinen Zusammenhang zwischen der Anzahl der Sitzungen und dem Trainingserfolg zu geben, wie auch Metaanalysen feststellen (Bennett & Gibbons, 2000; Fabiano et al., 2009; Hodgson et al., 2012; van der Oord et al., 2008). Die Annahme, dass weniger durchgeführte Sitzungen ein Indikator für eine gut gelingende Trainingsorganisationen sind, scheint nicht zuzutreffen oder die Organisation des Trainings hat keinen Zusammenhang mit dem Trainingserfolg.

Ergebnisse Responder und Non-Responder

Die neun Responder haben schlechtere initiale Werte. In der Lehrereinschätzung ist der Unterschied besonders deutlich, aber auch in der Elterneinschätzung ist die Diskrepanz erkennbar. Etwas schwächer sind die Unterschiede im Selbsturteil. Dies deckt sich mit den zuvor dargestellten Ergebnissen zur initialen Störungsschwere. Keiner der Teilnehmer hat eine reliable Verschlechterung im Laufe des Trainings erlebt. Für die Güte der Intervention ist dies positiv einzuschätzen, es bietet sich aber kein Anknüpfungspunkt für eine weiterführende differenzierte Auswertung.

Die Analyse der Non-Responder ist nicht zielführend, da es sich nach der vorgenommenen Auswertung um 87 % der Stichprobe handelt. Dieser hohe Prozentsatz widerspricht den Erwartungen und der klinischen Einschätzung der Verläufe. In der grafischen Darstellung zeigt sich, dass einige der Teilnehmer nah an der Grenze zu einer reliablen Verbesserung waren, diese aber knapp nicht erreichen konnten. Eine kleine Absenkung der Grenze hätte die Gruppenzuordnung dieser Teilnehmer verändert. Vermutlich ist die Irrtumswahrscheinlichkeit mit $\alpha = 5\%$ in Anbetracht der kleinen Stichprobengröße zu gering gewählt. Die Festlegung der Irrtumswahrscheinlichkeit ist im Rahmen des jeweiligen Untersuchungskontextes zu treffen und ist nicht festgelegt auf einen Standardwert (Kazdin, 2010). So wäre für die vorliegende Fragestellung auch $\alpha = 10\%$ akzeptabel. Damit wären deutlich mehr Teilnehmer in die Gruppe der Responder gelangt. Durch die Erhöhung des α -Niveaus würden zwar mehr Teilnehmer fälschlicherweise als Responder eingestuft, die eigentlich keine Responder sind, dafür wären aber kleine Effekte eher feststellbar, die sonst aufgrund der kleinen Stichprobe übersehen würden. In einem Hypothesen generierenden Vorgehen erscheint eine falsch positive Einschätzung eher akzeptabel als in einem Hypothesen prüfenden Vorgehen, da es der Aufstellung neuer Annahmen dient, die in einem folgenden Vorgehen überprüft werden. Für die anschließende Prüfung ist eine strengere α -Setzung angemessen.

Es liegt ein ähnlicher Befund aus der Studie von Barkley und Kollegen (2001) mit 97 Familien mit Jugendlichen mit ADHS und oppositionellem Trotzverhalten vor. Von den Familien, die an einem Kommunikations- und Problemlösetraining teilnahmen, erreichten je nach Auskunftsource maximal 23 % eine reliable Verbesserung, je nachdem welches Outcome-Maß herangezogen wird. In der Einschätzung der Väter liegen die Prozentsätze zwischen 6 und 22 % und in der Einschätzung der Mütter zwischen 12 und

23 %. Für die Familien, die eine Kombinationsbehandlung der ersten Maßnahme mit einer Verhaltenstherapie bekamen, waren die Prozentsätze noch etwas geringer. In der Auskunft der Väter liegen reliable Verbesserungen bei 0 bis 18 % und in der Auskunft der Mütter bei 4 bis 23 % der Gruppe vor. In dieser Studie zeigen sich auch reliable Verschlechterungen. Diese liegen bei der ersten Behandlung zwischen 0 und 10 % in der Einschätzung der Väter und zwischen 0 und 3 % in der Einschätzung der Mütter. Bei der Kombinationsbehandlung sind nur in der Einschätzung der Mütter für ein einziges Outcome-Maß Verschlechterungen in Höhe von 7 % vorhanden.

Es ist daher als fraglich zu beurteilen, ob der RCI an dieser Stelle ein gutes Maß ist, um klinisch signifikante Trainingserfolge festzustellen. Gut handhabbar zur Unterscheidung von Subgruppen innerhalb der Evaluationsstichprobe ist er anscheinend nicht. Eine Anhebung des Signifikanzniveaus für nachfolgende ähnliche Untersuchungen könnte sinnvoll sein. Auch Mohr (1995) berichtet von einer Anpassung des RCI zur Identifikation von Non-Respondern. So reduziert er die in der ursprünglichen Version des RCI (Jacobson & Revenstorf, 1988) geforderten zwei Standardmessfehler als kritische Grenze auf einen Standardmessfehler.

Ergebnisse der Clusteranalyse

In Cluster 1 befinden sich die Teilnehmer mit den geringsten Veränderungen. Teilnehmer dieses Clusters können auch als Non-Responder bezeichnet werden, auch wenn viele von ihnen leichte Verbesserungen erreichen, die allerdings sehr gering sind. Durch die Clusteranalyse können zudem zwei Teilnehmer identifiziert werden, die zwar in der Gesamtsymptomatik moderate Verbesserungen erreichen, was ausschließlich auf Verbesserungen im Problemmaß Reaktionszeit zurückzuführen ist (Cluster 4). Auch diese beiden Teilnehmer sind den Non-Respondern zuzuordnen, da sie lediglich schneller werden, ohne jedoch weniger Fehler zu machen oder sich in der Fremdeinschätzung zu verbessern.

In den Clustern 2 und 3 befinden sich Teilnehmer mit starken Verbesserungen. Nach der Gesamtsymptomatik sind die 14 Teilnehmer in Cluster 2 diejenigen mit den stärksten Verbesserungen, insbesondere im Lehrerurteil. In Cluster 3 sind die Verbesserungen aus Lehrersicht zwar geringer, doch zeigen sich für diese Teilnehmer im Vergleich mit den anderen Clustern die stärksten Verbesserungen im Faktor Fehler. Dieser Umstand ist besonders zu würdigen, da für diesen Faktor in der gesamten Stichprobe die geringsten Verbesserungen erreicht wurden. Aus inhaltlichen Erwägungen ist dieser Faktor besonders wichtig, da er abbildet, wie korrekt und sorgfältig Aufgaben gelöst werden. Die Teilnehmer aus Cluster 3 erzielten gleichzeitig als einzige im Mittel keine Verbesserungen im Faktor Reaktionszeit. Es scheint so zu sein, dass diese Teilnehmer genau das Ziel der Aneignung eines sorgfältigen, langsamen und bedachten Arbeitsstils erreichen: Sie bearbeiten die Aufgaben langsamer und machen daher weniger Fehler.

Erfolgreiche und weniger erfolgreiche Teilnehmer lassen sich gut anhand des Problemmaßes Symptomatik unterscheiden. Nach dieser Einschätzung sind sich die Cluster 2

und 3 ähnlicher als Cluster 1 und 4. Die Einteilung nach der Symptomatik gelingt besser als nach dem Gesamtproblemmaß, da hiernach auch die beiden Teilnehmer des vierten Clusters gute Ergebnisse erzielen. Diese liegen allerdings ausschließlich an der schnelleren Aufgabenbearbeitung, die an sich nicht als Verbesserung zu bewerten ist.

Zusammenfassend liefert die Clusteranalyse folgende Ergebnisse zu Mustern in der Veränderung: Die wenig bzw. nicht erfolgreichen Teilnehmer erreichen im Wesentlichen durchgängig schlechte Ergebnisse in allen vier Problembereichen. Als ein spezifisches Misserfolgsmuster wird eine reine Verbesserung der Reaktionszeit bei gleichzeitiger Nichtverbesserung der anderen Problemmaße gefunden. Bei der Beurteilung des Trainingserfolges ist also darauf zu achten, bei Teilnehmern mit großen Verbesserungen in der Reaktionszeit insbesondere die anderen Outcomevariablen zu überprüfen. Bei den erfolgreichen Teilnehmern zeigen sich zwei Erfolgsmuster: Eine Gruppe verbessert sich am stärksten in der Lehrereinschätzung, die andere zeigt im Vergleich der Cluster die stärksten Verbesserungen in der Güte ihrer Aufgabenbearbeitung.

Die Teilnehmer, die ihre Aufgaben besser bearbeiten, erfüllen insbesondere das Ziel des Lerntrainings. Vermutlich kann der verbesserte Arbeitsstil noch nicht von allen Teilnehmern dieser Gruppe auf konkrete Unterrichtssituationen übertragen werden, daher die etwas geringeren Verbesserungen im Lehrerurteil. Es ist zu überprüfen, ob bei dieser Gruppe verzögert Verbesserungen eintreten. Das Muster der Verbesserungen im Lehrerurteil bei durchschnittlich geringen Verbesserungen der Aufgabenbearbeitung mag damit zusammenhängen, dass diese Teilnehmer andere von den Lehrern wahrgenommene Probleme verbessern. Diese Probleme bzw. Ziele können sich auf das Verhalten im Unterricht, Beteiligung, Sozialverhalten usw. beziehen und somit nicht unmittelbar mit dem Arbeitsstil zusammenhängen. Beide Erfolgsmuster sind für sich genommen als sehr positiv zu bewerten. Es scheint sowohl schwierig zu sein, einen festgefahrenen Arbeitsstil zu verbessern, als auch feststehende Einschätzungen und Ansichten von Lehrern zu verändern.

Es folgt der Mittelwertvergleich für die Cluster 1 bis 3. Da die Zellengrößen klein sind, wird keine Signifikanzprüfung durchgeführt.

Bei den dichotomen Variablen fällt auf, dass in Cluster 2 prozentual die meisten Teilnehmer mit einer medikamentösen Behandlung sind, während das Verhältnis von Mädchen und Jungen in etwa gleich ist. Bei den metrischen Variablen zeigt sich, dass die Teilnehmer von Cluster 3 etwas älter sind als die Teilnehmer von Cluster 1 und 2. Die Anzahl der Sitzungen ist für alle Cluster annähernd gleich. Die Trainingsdauer ist in Cluster 2 etwas länger als in Cluster 3 und noch etwas länger als in Cluster 1.

Bei der Einschätzung der ADHS-Symptomatik mittels DISYPS-II fällt auf, dass die am wenigsten erfolgreichen Teilnehmer in Cluster 1 in der Eltern- und Selbsteinschätzung die geringsten Symptomwerte zum Prämesszeitpunkt aufweisen. Besonders auffällig ist die Diskrepanz in der Selbsteinschätzung. Die Jugendlichen des Clusters 1 schätzen ihre Symptomatik deutlich geringer ein als ihre Eltern und deutlich geringer als die Teil-

nehmer der übrigen Cluster, insbesondere im Vergleich zum zweiten Cluster. Der Unterschied in der Elterneinschätzung zwischen Cluster 1 und Cluster 2 beträgt in Werten des DISYPS $d = 0.27$ (entspricht 0.51 Einheiten der Standardabweichung der Gesamtgruppe). Zwischen Cluster 1 und Cluster 3 beträgt die Differenz $d = 0.11$ (entspricht 0.21 Einheiten der Standardabweichung der Gesamtgruppe). In der Selbsteinschätzung findet sich hingegen im Vergleich zwischen Cluster 1 und Cluster 2 eine Differenz von $d = 0.7$ (entspricht 1.17 Einheiten der Standardabweichung) und zwischen Cluster 1 und Cluster 3 ein Unterschied von $d = 0.38$ (entspricht 0.63 Einheiten der Standardabweichung). Diese Zusammenhänge zeigen sich gleichsam im SDQ und im WRI, allerdings weniger stark ausgeprägt. Der auffälligste Unterschied zwischen den erfolgreichen Clustern 2 und 3 und dem wenig erfolgreichen Cluster 1 liegt also darin, dass die Jugendlichen selbst ihre Problematik als weniger gravierend einschätzen.

Die Teilnehmer des dritten Clusters, die in der Lehrereinschätzung die geringsten Veränderungen erreichen, haben auch die geringsten Problemwerte im Lehrerurteil zum Prämesszeitpunkt. Die Teilnehmer in Cluster 2 haben die geringsten Werte im Lern- und Arbeitsverhalten. Dies wird insbesondere deutlich, wenn beide Skalen des Fragebogens vergleichend gegenübergestellt werden. Für das Lern- und Aufmerksamkeitsverhalten sind die Einschätzungen für Cluster 2 besonders niedrig. Die Verbesserungen für diesen Cluster sind für die Lehrereinschätzung am größten.

Im Vergleich der Leistungen in der TAP zeigen sich fast keine Unterschiede. Tendenziell sind die Teilnehmer von Cluster 3 etwas schneller und machen etwas mehr Fehler zum Prämesszeitpunkt.

Zusammenfassend lassen sich aus den Mittelwertvergleichen der Cluster zwei Tendenzen ablesen. Erstens zeigen sich stärkere Verbesserungen, wenn die Problematik zum Prämesszeitpunkt schwerwiegender ausgeprägt ist. So haben die Teilnehmer des insgesamt sehr erfolgreichen Clusters 2 in den meisten Outcomevariablen die höchsten Problemwerte bzw. die niedrigsten Kompetenzwerte zum Prämesszeitpunkt. Für Cluster 3 zeigt sich der Zusammenhang in einer gering verbesserten Lehrereinschätzung nach einer durch den Lehrer als gering eingeschätzten Problematik vor dem Training.

Zweitens scheint den Selbsteinschätzungen großes Gewicht beizukommen. Die Jugendlichen des wenig erfolgreichen Clusters 1 geben selbst eine deutlich geringere Problematik zum Prämesszeitpunkt an als die Jugendlichen der anderen Cluster, während sich die Elterneinschätzungen der drei Cluster zwar auch voneinander unterscheiden, aber weniger stark. Dieser Befund stimmt mit Ergebnissen von Mohr (1995) über ein, wonach eine geringe Selbsteinschätzung der Problematik bei gleichzeitig hoher Fremdeinschätzung als ein besonderes Risiko für einen wenig erfolgreichen Therapieverlauf zu sehen ist. Die Selbsteinschätzung der Problematik zum Prämesszeitpunkt scheint demnach insbesondere auch die Motivation der Jugendlichen zur Teilnahme am Training zu erfassen. Die Teilnehmer mit der höchsten selbst eingeschätzten Symptomatik in Cluster 2 erreichen die größten Verbesserungen. Da sie selbst angeben, gravierende Probleme

me zu haben, ist davon auszugehen, dass bei ihnen ein höheres Ausmaß an Leidensdruck und darüber hinaus Veränderungsmotivation vorliegt. Dieser Befund passt zum Modell von Sonuga-Barke (2002), das die Rolle der Motivation bei ADHS herausstellt. Geben die Jugendlichen selbst an, dass es Probleme gibt, sind sie motivierter etwas zu verändern. Damit wird eine bereits vollzogene oder geplante Veränderung ihres motivationalen Stils deutlich. Die im Modell postulierte erworbene generalisierte Verzögerungsaversion wird aufgebrochen, so dass die ADHS-Symptomatik verändert werden kann.

Die Ergebnisse der vorgestellten Clusteranalyse ermöglichen eine nachvollziehbare Zusammenfassung der Veränderungsmaße, auch wenn sich die Cluster teilweise überschneiden. Deskriptiv können plausible Veränderungsmuster erkannt werden. Allerdings ist eine Überlappung der Cluster festzustellen, was anhand der Standardabweichungen deutlich wird. Es handelt sich nicht um stark voneinander abgegrenzte Gruppen. Weiterhin ist anzumerken, dass die vorgestellte Lösung mit anderen Ähnlichkeitsmaßen und Fusionierungsalgorithmen nicht repliziert werden kann. Zur Entscheidung für die optimale Clusteranzahl ist die graphische Inspektion der Entwicklung des Heterogenitätsmaßes („Koeffizienten“ bei SPSS) mittels Scree-Plot analog zur Faktorenanalyse wenig aussagekräftig. So kann kein „Ellenbogen“ festgestellt werden. Daher ist die optimale Clusteranzahl primär nach sachlogischen Erwägungen getroffen, die nach (Backhaus, Erichson, Plinke et al., 2011) zur Bestimmung der Anzahl der Cluster herangezogen werden können. Die Lösung mit den vier Clustern ist inhaltlich plausibel und ermöglicht eine gute Darstellung und Zusammenfassung der Teilnehmer anhand ihrer Veränderungsdaten. Somit wird das Ziel der Analyse zufriedenstellend erreicht.

7.2 Fazit und klinische Implikationen

Zusammenfassend kann festgestellt werden, dass die beste Vorhersage der Werte zum Postmesszeitpunkt durch die Werte zum Prämesszeitpunkt möglich ist. Je höher die Werte vor dem Training ausgeprägt sind, desto höher sind sie nach dem Training. Weiterhin haben Teilnehmer mit einer größeren Problemschwere die besseren Erfolgsaussichten. Dieses ist das zentrale Ergebnis der vorliegenden Untersuchung und zieht sich durch alle Auswertungsschritte hindurch. Es liegen keine Hinweise auf einen Matthäus-Effekt vor. Ein kleiner Anteil des kompensatorischen Effekts mag durch den Regressionseffekt zur Mitte zu erklären sein, aber wie der weiterführende Vergleich mit den Ergebnissen der Kontrollgruppe zeigt, macht dies nur einen geringen Teil aus. Der kompensatorische Effekt zeichnet sich insgesamt deutlich ab. Andere Variablen können nur gering zur Güte der Vorhersage beitragen. Das Training scheint so wie es ist für die gesamte Zielgruppe geeignet zu sein. Insgesamt werden moderate Verbesserungen erreicht. Kontraindikationen, bei denen reliable Verschlechterungen auftreten, liegen nicht vor.

Die Ergebnisse der Untersuchung machen Hoffnung. Sie zeigen, dass auch bei schwerwiegenden Problemlagen gute Erfolge durch das Lerntraining erreicht werden können. Dies sollte den Eltern und Jugendlichen in der Psychoedukation zu Beginn des Trainings vermittelt werden, um Hoffnung zu induzieren, die im Zuge einer langen Geschichte von Misserfolgen oftmals verloren gegangen ist.

Dass Jugendliche mit Medikation stärkere Verbesserungen in der Lehrereinschätzung aufweisen, kann zumindest zwei Gründe haben. Einerseits vermag die medikamentöse Wirkung auf das „neurobiological substrate“ (Rapport et al., 2001) der Störung eine grundlegende Aufmerksamkeitsfähigkeit hergestellt haben, die ein zielführendes Mitarbeiten im Training überhaupt erst ermöglicht hat. So interpretiert spricht der Effekt für eine bessere Wirksamkeit der Kombinationsbehandlung. Alternativ könnte es sich um einen Effekt der Erwartung der Lehrer handeln, dass eine Kombinationsbehandlung die beste Intervention bei ADHS sei, wie einige empirische Studien feststellen. Dies unterstreicht die Wichtigkeit der Psychoedukation und des Schulkontaktes mit Lehrern, wie auch von Smith et al. (2000) festgestellt wird. Den Lehrern könnte stärker vermittelt werden, dass verhaltensbezogene und kognitive Interventionen ohne auch gleichzeitige Medikation effektiv sind. Bei vielen Schülern ist nach anfänglicher psychosozialer Intervention keine zusätzliche Medikation erforderlich (Döpfner et al., 2004). Insgesamt ist aber zu betonen, dass die Medikation mit dem Gesamtproblemmaß des Trainings, also bei Einbezug der Eltern- und Selbsteinschätzung, keinen Zusammenhang aufweist. Jugendliche mit und ohne eine gleichzeitige medikamentöse Behandlung profitieren in der Gesamteinschätzung gleich gut vom Training.

Das Training eignet sich insbesondere für Jugendliche mit einer größeren Problematik. Es hat einen kompensatorischen Effekt und ist geeignet, die bestehenden Probleme auszugleichen. Insbesondere die Ergebnisse der Clusteranalyse unterstreichen die Wichtigkeit der Selbsteinschätzungen der Jugendlichen. Diese sind auch als Maße der (Veränderungsmotivation zu verstehen. Jugendliche, die ihre eigene Problematik vor dem Training als gravierender einschätzen, hatten prinzipiell die größeren Erfolge im Training. Die Elterneinschätzung der Problematik war hingegen fast in allen Fällen hoch.

Manche Eltern stellen sich die Frage, ob sie ihr jugendliches Kind für die Teilnahme am Lerntraining anmelden sollen, wenn die Problematik gerade wenig akut ist. Die Eltern berichten, dass sie eine optimale Unterstützung ihres Kindes ermöglichen wollen, aber das Kind auch nicht überfordern möchten. Hierzu ist zu sagen, dass geringe Erfolge des Trainings zu erwarten sind, wenn die Problematik sowohl in der Selbst- als auch in der Fremdeinschätzung gering ausgeprägt ist. In einer solchen Situation ist auch die Meinung des Jugendlichen heranzuziehen. Wenn dieser eine Teilnahme für nicht notwendig hält und die Eltern und Lehrer übereinstimmend keine starke aktuelle Problematik berichten, so ist von einer Teilnahme eher abzuraten.

Eine andere Situation liegt vor, wenn die Jugendlichen selbst ihre Probleme als gering einschätzen, relevante Bezugspersonen wie Eltern und Lehrer aber eine gravierende

Problematik sehen. Hier könnte die Tendenz der Jugendlichen mit ADHS zutage treten, eigene Probleme eher herunterzuspielen und die eigenen Kompetenzen zu überschätzen (Kaiser & Hoza, 2008). Oder es wirkt sich die Tendenz zur Verzögerungsaversion nach dem Dual-Pathway-Modell aus (Sonuga-Barke, 2002), da die Jugendlichen wissen, dass sie, wenn sie Probleme zugeben, in einen anstrengenden, länger währenden Arbeitsprozess verwickelt werden. Damit wäre es so, dass Probleme und Leidensdruck zwar vorliegen, von den Jugendlichen aber nicht wahrgenommen oder geleugnet werden.

Es könnte sinnvoll sein, bei solchen Diskrepanzen zunächst an der Motivation der Jugendlichen zur Teilnahme und Veränderung der Probleme zu arbeiten. Dazu könnten Techniken der Motivierenden Gesprächsführung nach Miller und Rollnick (2009) zum Einsatz kommen. Bussing et al. (2012) stellen fest, dass die Motivation an einer ADHS bezogenen Intervention teilzunehmen und die „willingness to engage in treatment“ (S. 92) bei den Jugendlichen selbst deutlich geringer ausgeprägt sind als in der Einschätzung ihrer Eltern. Schlussfolgernd stellen sie den Bedarf nach solchen Interventionen heraus, die das Engagement der Jugendlichen mit ADHS verbessern.

Zur Erfassung von „wahren“ Veränderungen des Verhaltens und was als Erfolg einer Intervention mit Jugendlichen mit ADHS zu verstehen ist, sei noch Folgendes gesagt: Selbst wenn es sich nicht um „wahre“ Verhaltensänderungen der Jugendlichen handelt, kann die Verbesserung in der Einschätzung des Verhaltens trotzdem als Erfolg verstanden werden. Eine veränderte Einschätzung der Problematik kann auch darin begründet sein, dass bessere Umgangsweisen mit einem Problem im Sinne von Coping gefunden worden sind oder dass durch die erste Einschätzung ein Denk- und Umdenkprozess ausgelöst worden ist. Dies sollte als großer Erfolg gewertet werden.

Wenn die Einschätzungen relevanter Personen einer gravierenden manifesten Problematik durch bloße Wiederholung der Vorlage eines Fragebogens verändert werden könnten, dann wäre dies erst einmal verwunderlich, aber wenn es so wäre, wäre auch schon viel erreicht. In der vergleichenden Interventionsforschung steht die Frage danach im Vordergrund, welche die aktiven Elemente einer Behandlung sind und welche Intervention besser wirkt als eine andere. In der Fragestellung dieser Untersuchung geht es aber um die differentielle Wirksamkeit gerade dieser Maßnahme. Dazu ist zu sagen, dass sie besser wirkt bei größerer initialer Störungsschwere. Um dies mit einem Beispiel zu illustrieren: Fragte eine Mutter, ob das Training für ihren Sohn, der ein besonders „schwerer Fall“ sei, ob das Training auch ihm etwas bringen könne oder ob eine andere, intensivere Behandlung notwendig sei, dann wäre es ihr vermutlich egal, woran es liege, dass ihre Einschätzung, die ihres Sohnes und des Lehrers nach dem Training besser sein wird als zuvor. Schließlich ist festzustellen, dass der in dieser Untersuchung festgestellte kompensatorische Effekt nicht selbstverständlich ist. So werden in analogen Prä-post-Designs eben auch Matthäus-Effekte gefunden, z. B. bei der Behandlung von Depression (Jacobi et al., 2011).

Zur Beantwortung der Frage nach der differenziellen Wirksamkeit kann festgestellt werden, dass die Jugendlichen in der Stichprobe sich eher ähnlich sind als unterschiedlich. Es zeigt sich kein Veränderungsbedarf der Trainingskonzeption. Weiterführend könnten solche Aspekte untersucht werden, die in der vorliegenden quantitativen Auswertung nicht untersucht werden konnten. Es könnte in Mikroanalysen der Trainingssitzungen analysiert werden, in welchem Ausmaß die Wirkfaktoren jeweils realisiert wurden. Dazu könnten wie in der Forschung von Grawe (2005) die Videos der Sitzungen nach dem Ausmaß der Umsetzung der Wirkfaktoren bewertet werden. Außerdem könnten die Trainingsdokumentationen inhaltsanalytisch danach ausgewertet werden, welche Hinweise sich auf die Umsetzung der Wirkfaktoren finden. Auch die Untersuchung der Probleme der Organisation des Trainings bzw. Barrieren zur Teilnahme könnte ausgeweitet werden. In der dargestellten Auswertung zeigt sich kein Zusammenhang der Dauer des Trainings, die als Indikator einer gut gelingenden Trainingsorganisation durch die Eltern gesehen wird, mit dem Trainingserfolg. Es könnte sein, dass nicht die bloße Länge des zeitlichen Intervalls relevant ist, sondern die Kontinuität der Durchführung innerhalb dieses Intervalls. Die Trainingsdokumentationen könnten weiterführend danach ausgewertet werden, ob bei einer langen Trainingsdauer kontinuierlich mit größeren Abständen trainiert wurde oder ob größere Unterbrechungen vorlagen. Es ist zu vermuten, dass eine kontinuierliche aber seltenere Trainingsdurchführung einer Trainingsdurchführung mit längeren Unterbrechungen überlegen ist.

Für das Geschlecht der Jugendlichen ergeben sich in keiner der Auswertungen signifikante Effekte. Die Gruppe der Mädchen ist klein, daher kann nicht generalisiert werden. Wenn auch nicht signifikant, so ist bei den Mädchen die Symptomatik in der Selbsteinschätzung gravierender ausgeprägt und es liegen einige Ausreißerwerte in anderen Maßen vor. Eine Vermutung ist nach wie vor, dass die meisten Mädchen den Jungen sehr ähnlich sind, aber es darüber hinaus einige Mädchen mit besonders gravierenden Problemen gibt. Für diese sind Komorbiditäten, z. B. mit der Borderline-Persönlichkeitsstörung, denkbar (Stepp et al., 2012). Daher erscheint es vielversprechend, die Gruppe der Mädchen mit ADHS in nachfolgenden Untersuchungen genauer zu betrachten.

Als weiterführende Auswertungsmöglichkeit bei einer größeren Stichprobe würde sich der Einsatz sog. Künstlicher Neuronaler Netze eignen. Diese berücksichtigen eine Vielzahl von Variablen, deren Beziehungen untereinander vielfältig und nicht-linear sein können. Künstliche Neuronale Netze eignen sich insbesondere, wenn aufgrund komplexer Zusammenhänge keine begründeten Hypothesen abgeleitet werden können, so wie auch im Falle der vorliegenden Analyse. Neuronale Netze bieten eine Alternative für viele klassische Auswertungsmethoden, z. B. für die Regressionsanalyse. Ziel des Vorgehens ist dabei nicht vorrangig das Aufdecken von Wirkungszusammenhängen, sondern das „Trainieren“ des Netzes um zukünftige Ergebnisse bei bekannter Ausgangslage vorhersagen zu können. Um eine Aussage über die Güte der Vorhersagen treffen zu können, werden die Daten in einen größeren Trainingsdatensatz und in einen Testdatensatz eingeteilt. Das Training des Netzes wird nur mit dem Trainingsdatensatz durchgeführt und die Vorhersagegüte wird anhand des Testdatensatzes überprüft. Ein gut trai-

niertes Netz kann zur Vorhersage zukünftiger Outcomes genutzt werden. Die genauen Wirkzusammenhänge zwischen Input und Output werden nicht genau geklärt, aber es können Aussagen über die Bedeutsamkeit einzelner Variablen getroffen werden. So kann die relative Wichtigkeit einer Input-Variablen für den Output geschätzt werden. Der vorliegende Datensatz ist allerdings für ein solches Vorgehen zu klein, es werden einige hundert Fälle benötigt (Backhaus, Erichson & Weiber, 2011).

Abschließend ist darauf hinzuweisen, dass die vorliegende Untersuchung einen explorativen Charakter hat, bei dem Daten post-hoc ausgewertet wurden. Die meisten Auswertungen beruhen auf korrelativen Zusammenhängen, so dass Kausalschlüsse nicht möglich sind. In nachfolgenden Untersuchungen sollten solche Designs umgesetzt werden, in denen relevante Aspekte experimentell variiert werden, um so ihren Einfluss besser einschätzen zu können. Beispielsweise könnte bei der Zuordnung zu den experimentellen Gruppen einer Untersuchung die initiale Störungsschwere als eine stratifizierende Variable herangezogen werden. Weitere Beachtung verdient auch die Selbsteinschätzung der externalisierenden Komorbidität und die Lehrereinschätzung der internalisierenden Probleme. Welchen prädiktiven Wert diese Aspekte haben, sollte ebenfalls in weiterführenden Studien eruiert werden.

Literatur

- Abikoff, H. (1991). Cognitive training in ADHD children: Less to it than meets the eye. *Journal of Learning Disabilities, 24*(4), 205-209.
- Abikoff, H. B., Jensen, P. S., Arnold, L. L. E., Hoza, B., Hechtman, L., Pollack, S. et al. (2002). Observed classroom behavior of children with ADHD: Relationship to gender and comorbidity. *Journal of Abnormal Child Psychology, 30*(4), 349-359.
- Achenbach, T. M., McConaughy, S. H. & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*(2), 213-232.
- Albert, D. & Steinberg, L. (2011a). Age Differences in Strategic Planning as Indexed by the Tower of London. *Child Development, 82*(5), 1501-1517.
- Albert, D. & Steinberg, L. (2011b). Judgment and Decision Making in Adolescence. *Journal of Research on Adolescence, 21*(1), 211-224.
- Alf, E. F. & Abrahams, N. M. (1975). The use of extreme groups in assessing relationships. *Psychometrika, 40*(4), 563-572.
- Alsaker, F. D. & Bütikofer, A. (2005). Geschlechtsunterschiede im Auftreten von psychischen und Verhaltensstörungen im Jugendalter. *Kindheit und Entwicklung, 14*(3), 169-180.
- Altendorfer-Kling, U., Ardel-Gattinger, E. & Thun-Hohenstein, L. (2007). Der Selbstbeurteilungsbogen des SDQ anhand einer österreichischen Feldstichprobe. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie, 35*(4), 265-271.
- Anderson, J. E. (1939). The limitations of infant and preschool tests in the measurement of intelligence. *Journal of Psychology, 8*, 351-379.
- Antshel, K. M., Faraone, S. V. & Gordon, M. (2012). Cognitive behavioral treatment outcomes in adolescent ADHD. *Journal of Attention Disorders, XX*(X), 1-XX.
- Antshel, K. M. & Remer, R. (2003). Social skills training in children with attention deficit hyperactivity disorder: A randomized-controlled clinical trial. *Journal of Clinical Child and Adolescent Psychology, 32*(1), 153-165.
- Arbeitsgemeinschaft ADHS der Kinder- und Jugendärzte. (2007). *ADHS bei Kindern und Jugendlichen*: Verfügbar unter: www.agadhs.de/uploads/Leitlinie2009.pdf [06.05.2013].
- Arnett, J. J. (1999). Adolescent storm and stress, reconsidered. *American Psychologist, 54*(5), 317-326.
- Arnett, J. J. (2000). Emerging adulthood - A theory of development from the late teens through the twenties. *American Psychologist, 55*(5), 469-480.
- Babinski, D. E., Pelham, W. E., Molina, B. S. G., Waschbusch, D. A., Gnagy, E. M., Yu, J. H. et al. (2011). Women with childhood ADHD: Comparisons by diagnostic group and gender. *Journal of Psychopathology and Behavioral Assessment, 33*(4), 420-429.

- Bacchini, D., Affuso, G. & Trotta, T. (2008). Temperament, ADHD and peer relations among schoolchildren: The mediating role of school bullying. *Aggressive Behavior, 34*(5), 447-459.
- Bachmann, M., Bachmann, C., Rief, W. & Mattejat, F. (2008). Wirksamkeit psychiatrischer und psychotherapeutischer Behandlungen bei psychischen Störungen von Kindern und Jugendlichen - Eine systematische Auswertung der Ergebnisse von Metaanalysen und Reviews. Teil II: ADHS und Störungen des Sozialverhaltens. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie, 36*(5), 321-333.
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2011). *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung* (13., überarb. Aufl.). Heidelberg: Springer.
- Backhaus, K., Erichson, B. & Weiber, R. (2011). *Fortgeschrittene Multivariate Analysemethoden: Ein anwendungsorientierte Einführung*. Heidelberg: Springer.
- Baddeley, A. (2012). Working Memory: Theories, Models, and Controversies. In S. T. Fiske, D. L. Schacter & S. E. Taylor (Eds.), *Annual Review of Psychology, Vol 63* (Vol. 63, pp. 1-29). Palo Alto: Annual Reviews.
- Baer, R. A. & Nietzel, M. T. (1991). Cognitive and behavioral treatment of impulsivity in children: A meta-analytic review of the outcome literature. *Journal of Clinical Child Psychology, 20*(4), 400-412.
- Bagwell, C. L., Molina, B. S. G., Pelham, W. E. & Hoza, B. (2001). Attention-deficit hyperactivity disorder and problems in peer relations: Predictions from childhood to adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*(11), 1285-1292.
- Ball, S. W., Gilman, S. E., Mick, E., Fitzmaurice, G., Ganz, M. L., Seidman, L. J. et al. (2010). Revisiting the association between maternal smoking during pregnancy and ADHD. *Journal of Psychiatric Research, 44*(15), 1058-1062.
- Baltes, P. B., Staudinger, U. M. & Lindenberger, U. (1999). Lifespan psychology: Theory and application to intellectual functioning. *Annual Review of Psychology, 50*, 471-507.
- Banaschewski, T., Roessner, V., Uebel, H. & Rothenberger, A. (2004). Neurobiologie der Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS). *Kindheit und Entwicklung, 13*(3), 137-147.
- Bandelow, B. & Broocks, A. (2002). Wirksamkeitsuntersuchungen in der Psychotherapieforschung. *Verhaltenstherapie, 12*(3), 205-215.
- Barbarese, W. J., Katusic, S. K., Colligan, R. C., Weaver, A. L. & Jacobsen, S. J. (2007). Long-term school outcomes for children with attention-deficit/hyperactivity disorder: A population-based perspective. *Journal of Developmental and Behavioral Pediatrics, 28*(4), 265-273.
- Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121*(1), 65-94.

- Barkley, R. A. (2004). Adolescents with attention-deficit/hyperactivity disorder: An overview of empirically based treatments. *Journal of Psychiatric Practice*, 10(1), 39-56.
- Barkley, R. A. (2006). *Attention deficit hyperactivity disorder: A handbook for diagnosis and treatment* (3 ed.). New York: Guilford Press.
- Barkley, R. A. (2012). Distinguishing sluggish cognitive tempo from Attention-deficit/hyperactivity disorder in adults. *Journal of Abnormal Psychology*, 121(4), 978-990.
- Barkley, R. A., Dupaul, G. J. & McMurray, M. B. (1991). Attention-deficit disorder with and without hyperactivity: Clinical response to three dose levels of methylphenidate. *Pediatrics*, 87(4).
- Barkley, R. A., Edwards, G., Laneri, M., Fletcher, K. & Metevia, L. (2001). The efficacy of problem-solving communication training alone, behavior management training alone, and their combination for parent-adolescent conflict in teenagers with ADHD and ODD. *Journal of Consulting and Clinical Psychology*, 69(6), 926-941.
- Barkley, R. A., Fischer, M., Smallish, L. & Fletcher, K. (2006). Young adult outcome of hyperactive children: Adaptive functioning in major life activities. *Journal of the American Academy of Child and Adolescent Psychiatry*, 45(2), 192-202.
- Barkley, R. A., Guevremont, D. C., Anastopoulos, A. D. & Fletcher, K. E. (1992). A comparison of 3 family-therapy programs for treating family conflicts in adolescents with attention-deficit hyperactivity disorder. *Journal of Consulting and Clinical Psychology*, 60(3), 450-462.
- Baron, R. M. & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173-1182.
- Baskin, T. W., Slaten, C. D., Crosby, N. R., Pufahl, T., Schneller, C. L. & Ladell, M. (2010). Efficacy of Counseling and Psychotherapy in Schools: A Meta-Analytic Review of Treatment Outcome Studies. *Counseling Psychologist*, 38(7).
- Bauermeister, J. J., Barkley, R. A., Bauermeister, J. A., Martinez, J. V. & McBurnett, K. (2012). Validity of the sluggish cognitive tempo, inattention, and hyperactivity symptom dimensions: Neuropsychological and psychosocial correlates. *Journal of Abnormal Child Psychology*, 40(5), 683-697.
- Bauermeister, J. J., Shrout, P. E., Ramirez, R., Bravo, M., Alegria, M., Martinez-Taboas, A. et al. (2007). ADHD correlates, comorbidity, and impairment in community and treated samples of children and adolescents. *Journal of Abnormal Child Psychology*, 35(6), 883-898.
- Becker, A. (2007). *Strengths and Difficulties Questionnaire (SDQ). Evaluations and applications*. Dissertationsschrift Georg-August-Universität zu Göttingen, Göttingen.
- Becker, A., Hagenberg, N., Roessner, V., Woerner, W. & Rothenberger, A. (2004). Evaluation of the self-reported SDQ in a clinical setting: Do self-reports tell us more than ratings by adult informants? *European Child & Adolescent Psychiatry*, 13, 17-24.

- Becker, A., Woerner, W., Hasselhorn, M., Banaschewski, T. & Rothenberger, A. (2004). Validation of the parent and teacher SDQ in a clinical sample. *European Child & Adolescent Psychiatry, 13*, 11-16.
- Becker, K., El-Faddagh, M., Schmidt, M. H. & Laucht, M. (2007). Dopaminerge Polymorphismen und frühkindliche Regulationsprobleme. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie, 35*(2), 145-151.
- Beelmann, A. & Schneider, N. (2003). Wirksamkeit von Psychotherapie bei Kindern und Jugendlichen. Eine Übersicht und Meta-Analyse zum Bestand und zu Ergebnissen der deutschsprachigen Effektivitätsforschung. *Zeitschrift für Klinische Psychologie und Psychotherapie, 32*(2), 129-143.
- Beelmann, A., Stemmler, M., Lösel, F. & Jaurisch, S. (2007). Zur Entwicklung externalisierender Verhaltensprobleme im Übergang vom Vor- zum Grundschulalter. *Kindheit und Entwicklung, 16*(4), 229-239.
- Bennett, D. S. & Gibbons, T. A. (2000). Efficacy of child cognitive-behavioral interventions for antisocial behavior: A meta-analysis. *Child & Family Behavior Therapy, 22*(1), 1-15.
- Bereiter, C. (1963). Some persisting dilemmas in the measurement of change. In C. W. Harris (Ed.), *Problems in measuring change* (pp. 3-20). Madison: University of Wisconsin Press.
- Berk, L. E. (2011). *Entwicklungspsychologie*. München: Pearson.
- Best, J. R. & Miller, P. H. (2010). A Developmental Perspective on Executive Function. *Child Development, 81*(6).
- Bettge, S., Ravens-Sieberer, U., Wietzker, A. & Hölling, H. (2002). Ein Methodenvergleich der Child Behavior Checklist und des Strengths and Difficulties Questionnaire. *Das Gesundheitswesen, 64*.
- Biederman, J., Ball, S. W., Monuteaux, M. C., Surman, C. B., Johnson, J. L. & Zeitlin, S. (2007). Are girls with ADHD at risk for eating disorders? Results from a controlled, five-year prospective study. *Journal of Developmental and Behavioral Pediatrics, 28*(4), 302-307.
- Biederman, J., Faraone, S. V. & Monuteaux, M. C. (2002). Differential effect of environmental adversity by gender: Rutter's index of adversity in a group of boys and girls with and without ADHD. *American Journal of Psychiatry, 159*(9), 1556-1562.
- Biederman, J., Kwon, A., Aleardi, M., Chouinard, V. A., Marino, T., Cole, H. et al. (2005). Absence of gender effects on attention deficit hyperactivity disorder: findings in nonreferred subjects *Am J Psychiatry* (Vol. 162, pp. 1083-1089). United States.
- Biederman, J., Mick, E. & Faraone, S. V. (2000). Age-dependent decline of symptoms of attention deficit hyperactivity disorder: Impact of remission definition and symptom type. *American Journal of Psychiatry, 157*(5), 816-818.
- Biederman, J., Monuteaux, M. C., Mick, E., Spencer, T., Wilens, T. E., Silva, J. M. et al. (2006). Young adult outcome of attention deficit hyperactivity disorder: a controlled 10-year follow-up study. *Psychological Medicine, 36*(2), 167-179.

- Biederman, J., Petty, C. R., Evans, M., Small, J. & Faraone, S. V. (2010). How persistent is ADHD? A controlled 10-year follow-up study of boys with ADHD. *Psychiatry Research*, 177(3), 299-304.
- Biederman, J., Petty, C. R., O'Connor, K. B., Hyder, L. L. & Faraone, S. V. (2012). Predictors of persistence in girls with attention deficit hyperactivity disorder: results from an 11-year controlled follow-up study. *Acta Psychiatrica Scandinavica*, 125(2), 147-156.
- Biederman, J., Petty, C. R., Woodworth, K. Y., Lomedico, A., Hyder, L. L. & Faraone, S. V. (2012). Adult outcome of attention-deficit/hyperactivity disorder: a controlled 16-year follow-up study. *The Journal of clinical psychiatry*, 73(7), 941-950.
- Birchwood, J. & Daley, D. (2012). Brief report: The impact of Attention Deficit Hyperactivity Disorder (ADHD) symptoms on academic performance in an adolescent community sample. *Journal of Adolescence*, 35(1), 225-231.
- Blanchard, E. B. & Schwarz, S. P. (1988). Clinically significant changes in behavioral medicine. *Behavioral Assessment*, 10(2), 171-188.
- Borge, F.-M., Hoffart, A. & Sexton, H. (2010). Predictors of outcome in residential cognitive and interpersonal treatment for social phobia: Do cognitive and social dysfunction moderate treatment outcome? *Journal of Behavior Therapy and Experimental Psychiatry*, 41(3), 212-219.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation: für Human- und Sozialwissenschaftler* (4., überarb. Aufl.). Berlin: Springer.
- Bortz, J., Lienert, G. A., Barskova, T., Leitner, K. & Oesterreich, R. (2008). *Kurzgefasste Statistik für die klinische Forschung: Leitfaden für die verteilungsfreie Analyse kleiner Stichproben* (3., aktual. u. bearb. Aufl.). Berlin: Springer.
- Bortz, J., Lienert, G. A. & Boehnke, K. (2008). *Verteilungsfreie Verfahren in der Biostatistik* (3. korr. Aufl.). Heidelberg: Springer.
- Bortz, J. & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler* (7., vollst. überarb. u. erw. Aufl.). Heidelberg: Springer.
- Brandstädter, J. & Lindenberger, U. (Hrsg.). (2007). *Entwicklungspsychologie der Lebensspanne*. Stuttgart: Kohlhammer.
- Brassett-Harknett, A. & Butler, N. (2007). Attention-deficit/hyperactivity disorder: An overview of the etiology and a review of the literature relating to the correlates and lifecourse outcomes for men and women. *Clinical Psychology Review*, 27(2), 188-210.
- Braucht, G. N. & Reichardt, C. S. (1993). A Computerized Approach To Trickle-Process, Random Assignment. *Evaluation Review*, 17(1), 79-90.
- Breuer, D., Wolff Metternich, T. & Döpfner, M. (2009). Die Erfassung von Merkmalen von Aufmerksamkeitsdefizit-/Hyperaktivitätsstörungen (ADHS) anhand von Lehrerurteilen. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, 37(5), 431-440.
- Brickenkamp, R. (2002). Test d2. Aufmerksamkeits-Belastungs-Test (9. Auflage ed.). Göttingen: Hogrefe.

- Brook, J. S., Brook, D. W., Zhang, C., Seltzer, N. & Finch, S. J. (2013). Adolescent ADHD and adult physical and mental health, work performance, and financial stress. *Pediatrics*, *131*(1), 5-13.
- Brosius, F. (2011). *SPSS 19*. Heidelberg: mitp.
- Bruchmüller, K., Margraf, J. & Schneider, S. (2012). Is ADHD diagnosed in accord with diagnostic criteria? Overdiagnosis and influence of client gender on diagnosis. *Journal of Consulting and Clinical Psychology*, *80*(1), 128-138.
- Burgess, G. C., Depue, B. E., Ruzic, L., Willcutt, E. G., Du, Y. P. P. & Banich, M. T. (2010). Attentional control activation relates to working memory in Attention-Deficit/Hyperactivity Disorder. *Biological Psychiatry*, *67*(7), 632-640.
- Bush, G. (2011). Cingulate, frontal, and parietal cortical dysfunction in Attention-deficit/hyperactivity disorder. *Biological Psychiatry*, *69*(12), 1160-1167.
- Bush, G., Frazier, J. A., Rauch, S. L., Seidman, L. J., Whalen, P. J., Jenike, M. A. et al. (1999). Anterior cingulate cortex dysfunction in attention-deficit/hyperactivity disorder revealed by fMRI and the counting stroop. *Biological Psychiatry*, *45*(12), 1542-1552.
- Bussing, R., Koro-Ljungberg, M., Noguchi, K., Mason, D., Mayerson, G. & Garvan, C. W. (2012). Willingness to use ADHD treatments: A mixed methods study of perceptions by adolescents, parents, health professionals and teachers. *Social Science & Medicine*, *74*(1), 92-100.
- Börnert, M. (2011). *Klinische Validität der Testbatterie zur Aufmerksamkeitsprüfung*. Universität Oldenburg: unveröffentlichte Bachelorarbeit.
- Bühl, A. (2010). *SPSS 18: Einführung in die moderne Datenanalyse* (12., aktual. Aufl.). München: Pearson.
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion* (3. akt. Aufl.). München: Pearson.
- Bühner, M. & Ziegler, M. (2009). *Statistik für Psychologen und Sozialwissenschaftler*. München: Pearson.
- Bühner, M., Ziegler, M., Bohnes, B. & Lauterbach, K. (2006). Übungseffekte in den TAP Untertests Test Go/Nogo und Geteilte Aufmerksamkeit sowie dem Aufmerksamkeits-Belastungstest (d2). *Zeitschrift für Neuropsychologie*, *17*(3), 191-199.
- Camp, B. W. & Bash, M. A. S. (1981). *Think aloud: Increasing social and cognitive skills - A problem-solving program for children*. Champaign: Research Press.
- Carlson, C. L. & Mann, M. (2002). Sluggish cognitive tempo predicts a different pattern of impairment in the attention deficit hyperactivity disorder, predominantly inattentive type. *Journal of Clinical Child and Adolescent Psychology*, *31*(1), 123-129.
- Case, R. (1992). *The mind's staircase: Exploring the conceptual underpinnings of children's thought and knowledge*. Hillsdale, N.J.: Erlbaum.
- Casey, B. J., Tottenham, N., Liston, C. & Durston, S. (2005). Imaging the developing brain: what have we learned about cognitive development? *Trends in Cognitive Sciences*, *9*(3).

- Casey, R. J. & Berman, J. S. (1985). The outcome of psychotherapy with children. *Psychological Bulletin*, 98(2).
- Castellanos, F. X. & Tannock, R. (2002). Neuroscience of attention-deficit/hyperactivity disorder: The search for endophenotypes. *Nature Reviews Neuroscience*, 3(8), 617-628.
- Cattell, R. B. (1963). Theory of fluid and crystallized intelligence - A critical experiment. *Journal of Educational Psychology*, 54(1), 1-22.
- Cauffman, E. & Steinberg, L. (2000). (Im)maturity of judgment in adolescence: Why adolescents may be less culpable than adults. *Behavioral Sciences & the Law*, 18(6), 741-760.
- Cherkasova, M., Sulla, E. M., Dalena, K. L., Ponde, M. P. & Hechtman, L. (2013). Developmental course of attention deficit hyperactivity disorder and its predictors. *Journal of the Canadian Academy of Child and Adolescent Psychiatry = Journal de l'Academie canadienne de psychiatrie de l'enfant et de l'adolescent*, 22(1), 47-54.
- Chronis, A. M., Jones, H. A. & Raggi, V. L. (2006). Evidence-based psychosocial treatments for children and adolescents with attention-deficit/hyperactivity disorder. *Clinical Psychology Review*, 26(4), 486-502.
- Chronis-Tuscano, A., Molina, B. S. G., Pelham, W. E., Applegate, B., Dahlke, A., Overmyer, M. et al. (2010). Very early predictors of adolescent depression and suicide attempts in children with Attention-deficit/hyperactivity disorder. *Archives of General Psychiatry*, 67(10), 1044-1051.
- Chu, B. C. & Harrison, T. L. (2007). Disorder-specific effects of CBT for anxious and depressed youth: A meta-analysis of candidate mediators of change. *Clinical Child and Family Psychology Review*, 10(4), 352-372.
- Cicchetti, D. & Rogosch, F. A. (2002). A developmental psychopathology perspective on adolescence. *Journal of Consulting and Clinical Psychology*, 70(1).
- Clarkin, J. F. & Levy, K. N. (2004). The influence of client variables on psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (Vol. 5, pp. 194-226). New York: Wiley.
- Cleophas, T. J., Zwinderman, A. H., Cleophas, T. F. & Cleophas, E. P. (2009). *Statistics applied to clinical trials*: Dordrecht: Springer.
- Coghill, D., Nigg, J., Rothenberger, A., Sonuga-Barke, E. & Tannock, R. (2005). Whither causal models in the neuroscience of ADHD? *Developmental Science*, 8(2), 105-114.
- Coghill, D. & Seth, S. (2011). Do the diagnostic criteria for ADHD need to change? Comments on the preliminary proposals of the DSM-5 ADHD and Disruptive Behavior Disorders Committee. *European Child & Adolescent Psychiatry*, 20(2), 75-81.
- Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement*, 7(3), 249-253.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

- Coles, E. K., Slavec, J., Bernstein, M. & Baroni, E. (2012). Exploring the gender gap in referrals for children with ADHD and other disruptive behavior disorders. *Journal of Attention Disorders, 16*(2), 101-108.
- Conners, C. K., Epstein, J. N., March, J. S., Angold, A., Wells, K. C., Klaric, J. et al. (2001). Multimodal treatment of ADHD in the MTA: An alternative outcome analysis. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*(2), 159-167.
- Conners, C. K., Sitarenios, G., Parker, J. D. A. & Epstein, J. N. (1998). The revised Conners' Parent Rating Scale (CPRS-R): Factor structure, reliability, and criterion validity. *Journal of Abnormal Child Psychology, 26*(4), 257-268.
- Connor, D. F. & Ford, J. D. (2012). Comorbid symptom severity in attention-deficit/hyperactivity disorder: a clinical study. *The Journal of clinical psychiatry, 73*(5), 711-717.
- Corbisiero, S., Buchli-Kammermann, J. & Stieglitz, R.-D. (2010). Reliabilität und Validität des Wender-Reimherr-Interviews (WRI). *Zeitschrift für Psychiatrie, Psychologie und Psychotherapie, 58*(4), 323-331.
- Costello, E. J., Copeland, W. & Angold, A. (2011). Trends in psychopathology across the adolescent years: What changes when children become adolescents, and when adolescents become adults? *Journal of Child Psychology and Psychiatry, 52*(10), 1015-1025.
- Costello, E. J., Mustillo, S., Erkanli, A., Keeler, G. & Angold, A. (2003). Prevalence and development of psychiatric disorders in childhood and adolescence. *Archives of General Psychiatry, 60*(8), 837-844.
- Coyle, T. R., Pillow, D. R., Snyder, A. C. & Kochunov, P. (2011). Processing Speed Mediates the Development of General Intelligence (g) in Adolescence. *Psychological Science, 22*(10), 1265-1269.
- Cronbach, L. J. & Furby, L. (1970). How we should measure "change" - or should we? *Psychological Bulletin, 74*(1), 68-80.
- Cronbach, L. J. & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. New York: Irvington.
- Cumyn, L., French, L. & Hechtman, L. (2009). Comorbidity in adults with attention-deficit hyperactivity disorder. *Canadian Journal of Psychiatry-Revue Canadienne De Psychiatrie, 54*(10), 673-683.
- D'Zurilla, T. J., & Goldfried, M. R. (1971). Problem solving and behavior modification. *Journal of Abnormal Psychology, 78*(1), 107-126.
- Danforth, J. S., Anderson, L. P., Barkley, R. A. & Stokes, T. F. (1991). Observations of parent-child interactions with hyperactive children: Research and clinical implications. *Clinical Psychology Review, 11*(6), 703-727.
- de Boo, G. M. & Prins, P. J. M. (2007). Social incompetence in children with ADHD: Possible moderators and mediators in social-skills training. *Clinical Psychology Review, 27*(1), 78-97.
- de Zwaan, M., Gruss, B., Mueller, A., Graap, H., Martin, A., Glaesmer, H. et al. (2012). The estimated prevalence and correlates of adult ADHD in a German

- community sample. *European Archives of Psychiatry and Clinical Neuroscience*, 262(1), 79-86.
- Derks, E. M., Hudziak, J. J. & Boomsma, D. I. (2007). Why more boys than girls with ADHD receive treatment: A study of Dutch twins. *Twin Research and Human Genetics*, 10(5), 765-770.
- Deutsche Gesellschaft für Kinder- und Jugendpsychiatrie und Psychotherapie u. a. (Hrsg.). (2007). *Leitlinien zur Diagnostik und Therapie von psychischen Störungen im Säuglings-, Kindes- und Jugendalter* (3. überarb. Aufl.) Deutscher Ärzte Verlag.
- Deutsche Gesellschaft für Psychologie. (2007). *Richtlinien zur Manuskriptgestaltung* (3., überarb. u. erweiter. Aufl.). Göttingen: Hogrefe.
- Diamond, A. (2005). Attention-deficit disorder (attention-deficit/hyperactivity disorder without hyperactivity): A neurobiologic ally and behaviorally distinct disorder from attention-deficit/hyperactivity disorder (with hyperactivity). *Development and Psychopathology*, 17(3), 807-825.
- Diamond, A. (2012). Activities and programs that improve children's executive functions. *Current Directions in Psychological Science*, 21(5), 335-341.
- Dilling, H., Mombour, W. & Schmidt, M. H. (2004). *Internationale Klassifikation psychischer Störungen. ICD-10 Kapitel V (F). Diagnostische Kriterien für Forschung und Praxis*. Bern: Huber.
- Douglas, V. I. (1972). Stop, look, and listen: The problem of sustained attention and impulse control in hyperactive and normal children. *Canadian Journal of Behavioural Science*(4), 259-282.
- Dovis, S., van der Oord, S., Wiers, R. W. & Prins, P. J. M. (2012). Can motivation normalize working memory and task persistence in children with Attention-deficit/hyperactivity disorder? The effects of money and computer-gaming. *Journal of Abnormal Child Psychology*, 40(5), 669-681.
- Downey, J. L. & Christensen, L. (2006). Belief persistence in impression formation. *North American Journal of Psychology*, 8(3), 479-488.
- Doyle, A. E., Willcutt, E. G., Seidman, L. J., Biederman, J., Chouinard, V. A., Silva, J. et al. (2005). Attention-deficit/hyperactivity disorder endophenotypes. *Biological Psychiatry*, 57(11), 1324-1335.
- Drechsler, R., Brandeis, D., Földényi, M., Imhof, K., & Steinhausen, H.-C. (2005). The course of neuropsychological functions in children with attention deficit hyperactivity disorder from late childhood to early adolescence. *Journal of Child Psychology and Psychiatry*, 46(8), 824-836.
- Drechsler, R., Rizzo, P., & Steinhausen, H.-C. (2009). Zur klinischen Validität einer computergestützten Aufmerksamkeitstestbatterie für Kinder (KITAP) bei 7- bis 10-jährigen Kindern mit ADHS. *Kindheit und Entwicklung*, 18(3), 153 – 161.
- Dreisörner, T. & Georgiadis, J. (2011). Sensitivität und Spezifität computergestützter Verfahren zur Diagnostik von Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS) im Kindes- und Jugendalter – Die Testbatterie zur Aufmerksamkeitsprüfung (TAP) und Testbatterie zur Aufmerksamkeitsprüfung für Kinder (KITAP). *Empirische Sonderpädagogik*, 1, 3-19.

- Durlak, J. A. (2009). How to select, calculate, and interpret effect sizes. *Journal of Pediatric Psychology, 34*(9), 917-928.
- Durlak, J. A., Fuhrman, T. & Lampman, C. (1991). Effectiveness of cognitive behavior-therapy for maladapting children. A metaanalysis. *Psychological Bulletin, 110*(2), 204-214.
- Durlak, J. A., Wells, A. M., Cotten, J. K. & Johnson, S. (1995). Analysis of selected methodological issues in child psychotherapy research. *Journal of Clinical Child Psychology, 24*(2), 141-148.
- Durston, S. (2003). A review of the biological bases of ADHD: What have we learned from imaging studies? *Mental Retardation and Developmental Disabilities Research Reviews, 9*(3), 184-195.
- Dush, D. M., Hirt, M. L. & Schroeder, H. E. (1989). Self-statement modification in the treatment of child behavior disorders: A meta-analysis. *Psychological Bulletin, 106*(1), 97-106.
- Döpfner, M., Banaschewski, T. & Sonuga-Barke, E. (2008). Aufmerksamkeitsdefizit-/Hyperaktivitätsstörungen (ADHS). In F. Petermann (Hrsg.), *Lehrbuch der Klinischen Kinderpsychologie* (6., vollst. überarb. Aufl., S. 257-276). Göttingen: Hogrefe.
- Döpfner, M., Berner, W. & Lehmkuhl, G. (1994a). *Handbuch: Lehrerfragebogen über das Verhalten von Kindern und Jugendlichen. Forschungsergebnisse zur deutschen Fassung der Teacher's Report Form (TRF) der Child Behavior Checklist*. Köln: Arbeitsgruppe Kinder-, Jugend- und Familiendiagnostik (KJFD).
- Döpfner, M., Berner, W. & Lehmkuhl, G. (1994b). *Handbuch: Fragebogen für Jugendliche. Forschungsergebnisse zur deutschen Fassung der Youth Self-Report Form (YSR) der Child Behavior Checklist*. Köln: Arbeitsgruppe Kinder-, Jugend- und Familiendiagnostik (KJFD).
- Döpfner, M., Breuer, D., Schürmann, S., Wolff Metternich, T., Rademacher, C. & Lehmkuhl, G. (2004). Effectiveness of an adaptive multimodal treatment in children with attention-deficit hyperactivity disorder - global outcome. *European Child & Adolescent Psychiatry, 13*(supplement 1), I/117-I/129.
- Döpfner, M., Görtz-Dorten, A. & Lehmkuhl, G. (2008). *DISYPS-II: Diagnostik-System für Psychische Störungen nach ICD-10 und DSM-IV für Kinder und Jugendliche - II*. Bern: Huber.
- Döpfner, M. & Lehmkuhl, G. (2006). Diagnostik-System für psychische Störungen im Kindes- und Jugendalter nach ICD-10 und DSM-IV (DISYPS-KJ) (Nachdruck der 2., korr. u. erg. Aufl.). Bern: Huber.
- Döpfner, M., Schmeck, K. & Berner, W. (1994). *Handbuch: Elternfragebogen über das Verhalten von Kindern und Jugendlichen. Forschungsergebnisse zur deutschen Fassung der Child Behavior Checklist (CBCL/4-18)*. Köln: Arbeitsgruppe Kinder-, Jugend- und Familiendiagnostik (KJFD).
- Döpfner, M., Schürmann, S. & Frölich, J. (2013). *Therapieprogramm für Kinder mit hyperkinetischem und oppositionellem Problemverhalten: THOP* (5., überarb. u. erw. Aufl.). Weinheim: Beltz PVU.

- Ebert, D., Krause, J. & Roth-Sackenheim, C. (2003). ADHS im Erwachsenenalter - Leitlinien auf der Basis eines Expertenkonsensus mit Unterstützung der DGPPN. *Nervenarzt*, 74(10), 939-946.
- Elkind, D. & Sameroff, A. (1970). Developmental Psychology. *Annual Review of Psychology*, 21(1), 191-238.
- Ellison-Wright, I., Ellison-Wright, Z. & Bullmore, E. (2008). Structural brain change in Attention Deficit Hyperactivity Disorder identified by meta-analysis. *Bmc Psychiatry*, 8, 8.
- Engels, J. M. & Diehr, P. (2003). Imputation of missing longitudinal data: a comparison of methods. *Journal of clinical epidemiology*, 56(10), 968-976.
- Erhart, M., Döpfner, M., Ravens-Sieberer, U. & The Bella study group. (2008). Psychometric properties of two ADHD questionnaires: comparing the Conners' scale and the FBB-HKS in the general population of German children and adolescents – results of the BELLA study. *European Child & Adolescent Psychiatry*, 17(0), 106-115.
- Erikson, E. H. (1968). *Identity, Youth, and Crisis*. New York: Norton.
- Erikson, E. H. (1970). *Jugend und Krise. Die Psychodynamik im sozialen Wandel*. Stuttgart: Klett.
- Eysenck, H. J. (1952). The Effects of Psychotherapy: An Evaluation. *Journal of Consulting Psychology*, 16(5), 319-324.
- Fabiano, G. A., Pelham, W. E., Coles, E. K., Gnagy, E. M., Chronis-Tuscano, A. & O'Connor, B. C. (2009). A meta-analysis of behavioral treatments for attention-deficit/hyperactivity disorder. *Clinical Psychology Review*, 29(2), 129-140.
- Faraone, S. V., Biederman, J. & Mick, E. (2006). The age-dependent decline of attention deficit hyperactivity disorder: a meta-analysis of follow-up studies. *Psychological Medicine*, 36(2), 159-165.
- Faraone, S. V., Biederman, J., Weber, W. & Russell, R. L. (1998). Psychiatric, neuropsychological, and psychosocial features of DSM-IV subtypes of attention-deficit/hyperactivity disorder: Results from a clinically referred sample. *Journal of the American Academy of Child and Adolescent Psychiatry*, 37(2), 185-193.
- Faraone, S. V., Perlis, R. H., Doyle, A. E., Smoller, J. W., Goralnick, J. J., Holmgren, M. A. et al. (2005). Molecular genetics of attention-deficit/hyperactivity disorder. *Biological Psychiatry*, 57(11), 1313-1323.
- Fayyad, J., De Graaf, R., Kessler, R., Alonso, J., Angermeyer, M., Demyttenaere, K. et al. (2007). Cross-national prevalence and correlates of adult attention-deficit hyperactivity disorder. *British Journal of Psychiatry*, 190, 402-409.
- Fedele, D. A., Lefler, E. K., Hartung, C. M. & Canu, W. H. (2012). Sex differences in the manifestation of ADHD in emerging adults. *Journal of Attention Disorders*, 16(2), 109-117.
- Feinstein, A. R. (1970). Pre-therapeutic classification of co-morbidity in chronic disease. *Journal of Chronic Diseases*, 23(7), 455-468.
- Fend, H. (2005). *Entwicklungspsychologie des Jugendalters* (3., durchgesehene Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.

- Ferrer, M., Andion, O., Matali, J., Valero, S., Navarro, J. A., Ramos-Quiroga, J. A. et al. (2010). Comorbid Attention-deficit/hyperactivity disorder in borderline patients defines an impulsive subtype of borderline personality disorder. *Journal of Personality Disorders*, 24(6), 812-822.
- Field, A. (2009). *Discovering Statistics using SPSS*. Los Angeles: Sage.
- Fisseni, H.-J. (2004). *Lehrbuch der psychologischen Diagnostik* (3. überarb. u. erweit. Aufl.). Göttingen: Hogrefe.
- Flammer, A. & Alsaker, F. D. (2002). *Entwicklungspsychologie der Adoleszenz. Die Erschließung innerer und äußerer Welten im Jugendalter*. Bern: Huber.
- Friedman, L. M., Furberg, C. D. & DeMets, D. L. (2010). *Fundamentals of Clinical Trials* (4 ed.). New York: Springer.
- Földényi, M., Giovanoli, A., Tagwerker-Neuenschwander, F., Schallberger, U. & Steinhausen, H.-C. (2000). Reliabilität und Retest-Stabilität der Testleistungen von 7-10jährigen Kindern in der computerunterstützten TAP. *Zeitschrift für Neuropsychologie*, 11(1), 1-11.
- Földényi, M., Imhof, K. & Steinhausen, H.-C. (2000). Klinische Validität der computerunterstützten TAP bei Kindern mit Aufmerksamkeits-/Hyperaktivitätsstörungen. *Zeitschrift für Neuropsychologie*, 11(3), 154-167.
- Garcia, C. R., Bau, C. H. D., Silva, K. L., Callegari-Jacques, S. M., Salgado, C. A. I., Fischer, A. G. et al. (2012). The burdened life of adults with ADHD: Impairment beyond comorbidity. *European Psychiatry*, 27(5), 309-313.
- Garner, A. A., Marceaux, J. C., Mrug, S., Patterson, C. & Hodgens, B. (2010). Dimensions and correlates of Attention deficit/hyperactivity disorder and sluggish cognitive tempo. *Journal of Abnormal Child Psychology*, 38(8), 1097-1107.
- Gaub, M. & Carlson, C. L. (1997). Gender differences in ADHD: A meta-analysis and critical review. *Journal of the American Academy of Child and Adolescent Psychiatry*, 36(8), 1036-1045.
- Gawrilow, C. (2012). *Lehrbuch ADHS*. München: Reinhardt UTB.
- Gerjets, P., Graw, T., Heise, E., Westermann, R. & Rothenberger, A. (2002). Handlungskontrolldefizite und störungsspezifische Zielintentionen bei der Hyperkinetischen Störung. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 31(2), 89-98.
- Gershon, J. (2002). A meta-analytic review of gender differences in ADHD. *Journal of Attention Disorders*, 5(3), 143-154.
- Glass, C. S. & Weigar, K. (2000). Teacher perceptions of the incidence and management of the attention deficit hyperactivity disorder. *Education*, 121(2), 412-420.
- Glass, K. & Flory, K. (2012). Are symptoms of ADHD related to substance use among college students? *Psychology of Addictive Behaviors*, 26(1), 124-132.
- Gleissner, U. (2007). Lern- und Merkfähigkeit. In L. Kaufmann, H.-C. Nuerk, K. Konrad & K. Willmes (Hrsg.), *Kognitive Entwicklungsneuropsychologie* (S. 177-199). Göttingen: Hogrefe.

- Goethe, J. W. (1986, Original 1774). *Die Leiden des jungen Werther*. Stuttgart: Reclam.
- Gollwitzer, M. & Jäger, R. S. (2009). *Evaluation kompakt*. Weinheim: Beltz PVU.
- Gomes, L. (2011). *Attention and memory in boys with predominantly inattentive and combined subtypes of ADHD*. (unpublished doctoral dissertation). Australian National University.
- Goodman, A., Lamping, D. L. & Ploubidis, G. B. (2010). When to use broader internalising and externalising subscales instead of the hypothesised five subscales on the Strengths and Difficulties Questionnaire (SDQ): Data from British parents, teachers and children. *Journal of Abnormal Child Psychology*, 38(8), 1179-1191.
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry*, 38, 581-586.
- Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry*, 40(11), 1337-1345.
- Goodman, R. (2005). Fragebogen zu Stärken und Schwächen (SDQ-Deu). Eltern 4-16. Retrieved 01.10., 2007, from <http://www.sdqinfo.com/questionnaires/german/s3.pdf>
- Goodman, R. & Scott, S. (1999). Comparing the strengths and difficulties questionnaire and the child behavior checklist: Is small beautiful? *Journal of Abnormal Child Psychology*, 27(1), 17-24.
- Gottesman, I. I. & Gould, T. D. (2003). The endophenotype concept in psychiatry: Etymology and strategic intentions. *American Journal of Psychiatry*, 160(4), 636-645.
- Graetz, B. W., Sawyer, M. G. & Baghurst, P. (2005). Gender differences among children with DSM-IV ADHD in Australia. *Journal of the American Academy of Child and Adolescent Psychiatry*, 44(2), 159-168.
- Grawe, K. (1992). Psychotherapieforschung zu Beginn der neunziger Jahre. *Psychologische Rundschau*, 43(3), 132-162.
- Grawe, K. (1997). Research-informed psychotherapy. *Psychotherapy Research*, 7(1), 1-19.
- Grawe, K. (2005). (Wie) kann Psychotherapie durch empirische Validierung wirksamer werden? *Psychotherapeutenjournal*, 1, 4-11.
- Greene, R. W., Biederman, J., Faraone, S. V., Monuteaux, M. C., Mick, E., DuPre, E. P. et al. (2001). Social impairment in girls with ADHD: Patterns, gender comparisons, and correlates. *Journal of the American Academy of Child and Adolescent Psychiatry*, 40(6), 704-710.
- Greven, C. U., Rijdsdijk, F. V., Asherson, P. & Plomin, R. (2012). A longitudinal twin study on the association between ADHD symptoms and reading. *Journal of Child Psychology and Psychiatry*, 53(3), 234-242.
- Grevet, E. H., Bau, C. H. D., Salgado, C. A. I., Fischer, A. G., Kalil, K., Victor, M. M. et al. (2006). Lack of gender effects on subtype outcomes in adults with

- attention-deficit/hyperactivity disorder. *European Archives of Psychiatry and Clinical Neuroscience*, 256(5), 311-319.
- Grob, A. & Jaschinski, U. (2003). *Erwachsen werden. Entwicklungspsychologie des Jugendalters*. Weinheim: Beltz PVU.
- Grünke, M. (2006). Zur Effektivität von Fördermethoden bei Kindern und Jugendlichen mit Lernstörungen. Eine Synopse vorliegender Metaanalysen. *Kindheit und Entwicklung*, 15(4), 239-254.
- Gudjonsson, G. H., Sigurdsson, J. F., Sigfusdottir, I. D. & Young, S. (2012). An epidemiological study of ADHD symptoms among young persons and the relationship with cigarette smoking, alcohol consumption and illicit drug use. *Journal of Child Psychology and Psychiatry*, 53(3), 304-312.
- Görtz-Dorten, A. & Döpfner, M. (2009). Aufmerksamkeitsdefizit-/Hyperaktivitätsstörungen von Kindern und Jugendlichen im Elternurteil. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, 37(3), 183-194.
- Günther, T., Herpertz-Dahlmann, B. & Konrad, K. (2005). Reliabilität von Aufmerksamkeits- und verbalen Gedächtnistests bei gesunden Kindern und Jugendlichen - Implikationen für die klinische Praxis. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, 33(3), 169-179.
- Hager, W. (2000). Zur Wirksamkeit von Interventionsprogrammen: Allgemeine Kriterien der Wirksamkeit von Programmen in einzelnen Untersuchungen. In W. Hager, J.-L. Patry & H. Brezing (Eds.), *Evaluation psychologischer Interventionsmaßnahmen* (pp. 153-168). Bern: Hans Huber.
- Hall, G. S. (1904). *Adolescence: Its psychology and its relation to physiology, anthropology, sociology, sex, crime, religion, and education*. New York: Appleton.
- Halmoy, A., Klungsoyr, K., Skjaerven, R. & Haavik, J. (2012). Pre- and Perinatal Risk Factors in Adults with Attention-Deficit/Hyperactivity Disorder. *Biological Psychiatry*, 71(5), 474-481.
- Hampel, P. & Petermann, F. (2005). Age and Gender Effects on Coping in Children and Adolescents. *Journal of Youth and Adolescence*, 34(2), 73-83.
- Hankin, B. L., Mermelstein, R. & Roesch, L. (2007). Sex Differences in Adolescent Depression: Stress Exposure and Reactivity Models. *Child Development*, 78(1), 279-295.
- Hart, E. L., Lahey, B. B., Loeber, R., Applegate, B. & Frick, P. J. (1995). Developmental change in attention-deficit hyperactivity disorder in boys: A four-year longitudinal study. *Journal of Abnormal Child Psychology*, 23(6), 729-749.
- Hart, E. L., Lahey, B. B., Loeber, R. & Hanson, K. S. (1994). Criterion validity of informants in the diagnosis of disruptive behavior disorders in children: A preliminary study. *Journal of Consulting and Clinical Psychology*, 62(2), 410-414.
- Hartman, C. A., Willcutt, E. G., Rhee, S. H. & Pennington, B. F. (2004). The relation between sluggish cognitive tempo and DSM-IV ADHD. *Journal of Abnormal Child Psychology*, 32(5), 491-503.

- Hartmann, A. & Herzog, T. (1995). Varianten der Effektstärkenberechnung in Meta-Analysen: Kommt es zu variablen Ergebnissen? *Zeitschrift für Klinische Psychologie*, 24(4), 337-343.
- Hartmann, A., Herzog, T. & Drinkmann, A. (1992). Psychotherapy of bulimia nervosa: What is effective? A meta-analysis. *Journal of Psychosomatic Research*, 36(2), 159-167.
- Havey, J. M. (2007). A comparison of Dutch and US teachers' perceptions of the incidence and management of ADHD. *School Psychology International*, 28(1), 46-52.
- Havighurst, R. J. (1948). *Developmental tasks and education*. Chicago: University of Chicago Press.
- Hawthorne, G. & Elliott, P. (2005). Imputing cross-sectional missing data: Comparison of common techniques. *Australian and New Zealand Journal of Psychiatry*, 39(7), 583-590.
- Heddle, N. M. & Cook, R. d. J. (2011). Composite outcomes in clinical trials: what are they and when should they be used? *Transfusion*, 51(1), 11-13.
- Hedeker, G. & Gibbons, R. D. (2006). *Longitudinal Data Analysis*. Hoboken: Wiley.
- Heinrichs, N. & Lohaus, A. (2011). *Klinische Entwicklungspsychologie kompakt*. Weinheim: Beltz PVU.
- Hemphill, S. A. & Littlefield, L. (2006). Child and family predictors of therapy outcome for children with Behavioral and emotional problems. *Child Psychiatry & Human Development*, 36(3), 329-349.
- Hennig, T. (2008). *Lerntraining für Jugendliche mit Aufmerksamkeitsstörungen: Konzipierung und Evaluation im Einzelfalldesign*. Universität Oldenburg: unveröffentlichte Diplomarbeit.
- Herbert, M. (1989). *"Ich bin kein Kind mehr!"*. Bern: Huber.
- Hervey, A. S., Epstein, J. N. & Curry, J. F. (2004). Neuropsychology of adults with attention-deficit/hyperactivity disorder: A meta-analytic review. *Neuropsychology*, 18(3), 485-503.
- Hill, C. E. & Lambert, M. J. (2004). Methodological Issues in Studying Psychotherapy Process and Outcomes. In M. J. Lambert (Ed.), *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (Vol. 5, pp. 84-135). New York: Wiley.
- Hiller, W., Bleichhardt, G. & Schindler, A. (2009). Evaluation von Psychotherapien aus der Perspektive von Qualitätssicherung und Qualitätsmanagement. *Zeitschrift fuer Psychiatrie, Psychologie und Psychotherapie*, 57(1), 7-22.
- Hiller, W., Schindler, A., Andor, T. & Rist, F. (2011). Vorschläge zur Evaluation regulärer Psychotherapien an Hochschulambulanzen im Sinne der Phase-IV-Therapieforschung. *Zeitschrift fuer Klinische Psychologie und Psychotherapie*, 40(1), 22-32.
- Hinshaw, S. P. (2006). Treatment for children and adolescents with Attention-deficit/hyperactivity disorder. In P. C. Kendall (Ed.), *Child and adolescent therapy: Cognitive-behavioral procedures* (3 ed., pp. 82-113). New York: Guilford Press.

- Hinshaw, S. P. (2007). Moderators and mediators of treatment outcome for youth with ADHD: Understanding for whom and how interventions work. *Journal of Pediatric Psychology, 32*(6), 664-675.
- Hinshaw, S. P., Owens, E. B., Sami, N. & Fargeon, S. (2006). Prospective follow-up of girls with attention-deficit/hyperactivity disorder into adolescence: Evidence for continuing cross-domain impairment. *Journal of Consulting and Clinical Psychology, 74*(3), 489-499.
- Hinz, A. (2007). Prävention durch Unterstützung bei der Bewältigung geschlechtsspezifischer Entwicklungsaufgaben. *Zeitschrift für Pädagogische Psychologie, 21*(2), 145-155.
- Hodgson, K., Hutchinson, A. D. & Denson, L. (2012). Nonpharmacological treatments for ADHD: A meta-analytic review. *Journal of Attention Disorders, 16*(1), 1-11.
- Hofmann, E. (2003). *Progressive Muskelentspannung: Ein Trainingsprogramm* (2. korr. u. erg. Aufl.). Göttingen: Hogrefe.
- Holmbeck, G. N. (1997). Toward terminological, conceptual, and statistical clarity in the study of mediators and moderators: Examples from the child-clinical and pediatric psychology literatures. *Journal of Consulting and Clinical Psychology, 65*(4), 599-610.
- Holmbeck, G. N., Devine, K. A. & Bruno, E. F. (2010). Developmental issues and considerations in research and practice. In J. R. Weisz & A. E. Kazdin (Eds.), *Evidence-based psychotherapies for children and adolescents* (2 ed., pp. 28-39). New York: Guilford Press.
- Holmbeck, G. N., Greenley, R. N. & Franks, E. A. (2003). Developmental issues and considerations in research and practice. In A. E. Kazdin & J. R. Weisz (Eds.), *Evidence-based psychotherapies for children and adolescents* (pp. 21-41). New York: Guilford Press.
- Holmbeck, G. N., O'Mahar, K., Abad, M., Colder, C. & Updegrave, A. (2006). Cognitive-behavioral therapy with adolescents. In P. C. Kendall (Ed.), *Child and adolescent therapy: Cognitive-behavioral procedures* (3 ed., pp. 419-464). New York: Guilford Press.
- Holmberg, K. & Hjern, A. (2008). Bullying and attention-deficit-hyperactivity disorder in 10-year-olds in a Swedish community. *Developmental Medicine and Child Neurology, 50*(2), 134-138.
- Hoza, B., Mrug, S., Gerdes, A. C., Hinshaw, S. P., Bukowski, W. M. & Gold, J. A. (2005). What aspects of peer relationships are impaired in children with attention-deficit/hyperactivity disorder? *Journal of Consulting and Clinical Psychology, 73*(3), 411-423.
- Hunter, J. E. (1997). Needed: A ban on the significance test. *Psychological Science, 8*(1), 3-7.
- Hurrelmann, K. (2007). *Lebensphase Jugend. Eine Einführung in die sozialwissenschaftliche Jugendforschung* (9., aktual. Aufl.). Weinheim: Juventa.
- Hurrelmann, K. & Quenzel, G. (2012). *Lebensphase Jugend: Eine Einführung in die sozialwissenschaftliche Jugendforschung* (11., vollst. überarb. Aufl.). Weinheim: Beltz Juventa.

- Hurtig, T., Taanila, A., Moilanen, I., Nordstrom, T. & Ebeling, H. (2012). Suicidal and self-harm behaviour associated with adolescent attention deficit hyperactivity disorder-A study in the Northern Finland Birth Cohort 1986. *Nordic Journal of Psychiatry*, 66(5), 320-328.
- Huss, M., Hölling, H., Kurth, B. M. & Schlack, R. (2008). How often are German children and adolescents diagnosed with ADHD? Prevalence based on the judgment of health care professionals: results of the German health and examination survey (KiGGS). *European Child & Adolescent Psychiatry*, 17, 52-58.
- Ihle, W. & Esser, G. (2002). Epidemiologie psychischer Störungen im Kindes- und Jugendalter: Prävalenz, Verlauf, Komorbidität und Geschlechtsunterschiede. *Psychologische Rundschau*, 53(4), 159-169.
- Impey, M. & Heun, R. (2012). Completed suicide, ideation and attempt in attention deficit hyperactivity disorder. *Acta Psychiatrica Scandinavica*, 125(2), 93-102.
- Jacobi, F. (2006). Entwicklung und Beurteilung therapeutischer Interventionen. In H.-U. Wittchen & J. Hoyer (Hrsg.), *Klinische Psychologie & Psychotherapie* (S. 553-580). Heidelberg: Springer.
- Jacobi, F., Uhmann, S. & Hoyer, J. (2011). Wie häufig ist therapeutischer Misserfolg in der ambulanten Psychotherapie? *Zeitschrift für Klinische Psychologie und Psychotherapie*, 40(4), 246-256.
- Jacobson, N. S., Follette, W. C. & Revenstorf, D. (1986). Toward a Standard Definition of Clinically Significant Change. *Behavior Therapy*, 17(3), 308-311.
- Jacobson, N. S. & Revenstorf, D. (1988). Statistics for assessing the clinical significance of psychotherapy techniques: Issues, problems, and new developments. *Behavioral Assessment*, 10(2), 133-145.
- Jacobson, N. S. & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59(1), 12-19.
- Jensen, P. S., Arnold, L. E., Richters, J. E., Severe, J. B., Vereen, D., Vitiello, B. et al. (1999). A 14-month randomized clinical trial of treatment strategies for attention-deficit/hyperactivity disorder. *Archives of General Psychiatry*, 56(12), 1073-1086.
- Jensen, P. S., Arnold, L. E., Richters, J. E., Severe, J. B., Vereen, D., Vitiello, B. et al. (1999). A 14-month randomized clinical trial of treatment strategies for attention-deficit/hyperactivity disorder. *Archives of General Psychiatry*, 56(12), 1073-1086.
- Jensen, P. S., Arnold, L. E., Swanson, J. M., Vitiello, B., Abikoff, H. B., Greenhill, L. L. et al. (2007). 3-year follow-up of the NIMH MTA study. *J Am Acad Child Adolesc Psychiatry*, 46(8), 989-1002.
- Johansen, E., Killeen, P., Russell, V., Tripp, G., Wickens, J., Tannock, R. et al. (2009). Origins of altered reinforcement effects in ADHD. *Behavioral and Brain Functions*, 5(7).
- Johnston, C. & Mash, E. J. (2001). Families of children with attention-deficit/hyperactivity disorder: Review and recommendations for future research. *Clinical Child and Family Psychology Review*, 4(3), 183-207.

- Johnston, C. & Scoular, D. (2008). The family context of ADHD. In K. McBurnett & L. Pfiffner (Eds.), *Attention deficit hyperactivity disorder: Concepts, controversies, new directions* (pp. 41-49). New York: informa.
- Kaiser, N. M. & Hoza, B. (2008). Self-esteem and self-perceptions in ADHD. In K. McBurnett & L. Pfiffner (Eds.), *Attention deficit hyperactivity disorder: Concepts, controversies, new directions* (pp. 29-39). New York: informa.
- Kanfer, F. H., Reinecker, H. & Schmelzer, D. (2006). *Selbstmanagement-Therapie* (4. Aufl.). Heidelberg: Springer.
- Kaufmann, L., Zieren, N., Zotter, S., Karall, D., Scholl-Bürgi, S., Haberlandt, E. et al. (2010). Predictive validity of attentional functions in differentiating children with and without ADHD: a componential analysis. *Developmental Medicine & Child Neurology*, 52(4), 371-378.
- Kawabata, Y., Tseng, W.-L. & Gau, S. S.-F. (2012). Symptoms of Attention-deficit/hyperactivity disorder and social and school adjustment: The moderating roles of age and parenting. *Journal of Abnormal Child Psychology*, 40(2), 177-188.
- Kazdin, A. (1996). Dropping Out of Child Psychotherapy: Issues for Research and Implications for Practice. *Clinical Child Psychology and Psychiatry*, 1(1), 133-156.
- Kazdin, A. & Wassell, G. (1999). Barriers to Treatment Participation and Therapeutic Change Among Children Referred for Conduct Disorder. *Journal of Clinical Child and Adolescent Psychology*, 28(2).
- Kazdin, A. E. (2004). Psychotherapy for children and adolescents. In M. J. Lambert (Ed.), *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (Vol. 5, pp. 543-589). New York: Wiley.
- Kazdin, A. E. (2010). *Research design in clinical psychology* (4. internat. ed.). Boston: Allyn & Bacon.
- Kazdin, A. E. & Crowley, M. J. (1997). Moderators of treatment outcome in cognitively based treatment of antisocial children. *Cognitive Therapy and Research*, 21(2), 185-207.
- Kazdin, A. E. & Whitley, M. K. (2006). Comorbidity, case complexity, and effects of evidence-based treatment for children referred for disruptive behavior. *Journal of consulting and clinical psychology*, 74(3), 455-467.
- Kazis, L. E., Anderson, J. J. & Meenan, R. F. (1989). Effect sizes for interpreting changes in health status. *Medical Care*, 27(3), S178-S189.
- Kebir, O. & Joobar, R. (2011). Neuropsychological endophenotypes in attention-deficit/hyperactivity disorder: a review of genetic association studies. *European Archives of Psychiatry and Clinical Neuroscience*, 261(8), 583-594.
- Kendall, P. C., Holmbeck, G. & Verduin, T. (2004). Methodology, Design, and Evaluation in Psychotherapy Research. In M. J. Lambert (Ed.), *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (5 ed., pp. 16-43). New York: Wiley.
- Kessler, R. C., Adler, L., Barkley, R., Biederman, J., Conners, C. K., Demler, O. et al. (2006). The prevalence and correlates of adult ADHD in the United States:

- Results from the National Comorbidity Survey Replication. *American Journal of Psychiatry*, 163(4), 716-723.
- Kistner, J. (2009). Sex Differences in Child and Adolescent Psychopathology: An Introduction to the Special Section. *Journal of Clinical Child and Adolescent Psychology*, 38(4), 453-460.
- Klasen, H., Woerner, W., Wolke, D., Meyer, R., Overmeyer, S., Kaschnitz, W. et al. (2000). Comparing the German versions of the Strengths and Difficulties Questionnaire (SDQ-Deu) and the Child Behavior Checklist. *European Child & Adolescent Psychiatry*, 9(4), 271-276.
- Klauer, K. J. (2001). *Handbuch Kognitives Training* (2., überarb. u. erweit. Aufl.). Göttingen: Hogrefe.
- Kleinbaum, D. G., Kupper, L. L., Nizam, A. & Muller, K. E. (2008). *Applied Regression Analysis and Other Multivariable Methods* (4 ed.). Belmont: Thomson.
- Knopik, V. S. (2009). Maternal smoking during pregnancy and child outcomes: Real or spurious effect? *Developmental Neuropsychology*, 34(1), 1-36.
- Kofler, M. J., Rapport, M. D. & Alderson, R. M. (2008). Quantifying ADHD classroom inattentiveness, its moderators, and variability: A meta-analytic review. *Journal of Child Psychology and Psychiatry*, 49(1), 59-69.
- Kolko, D. J., Cheng, Y., Campo, J. V. & Kelleher, K. (2011). Moderators and predictors of clinical outcome in a randomized trial for behavior problems in pediatric primary care. *Journal of Pediatric Psychology*, 36(7), 753-765.
- Konrad, K. (2007). Entwicklung von Exekutivfunktionen und Arbeitsgedächtnisleistungen. In L. Kaufmann, H.-C. Nuerk, K. Konrad & K. Willmes (Eds.), *Kognitive Entwicklungsneuropsychologie* (S. 300-320). Göttingen: Hogrefe.
- Kooij, S. J. J., Bejerot, S., Blackwell, A., Caci, H., Casas-Brugue, M., Carpentier, P. J. et al. (2010). European consensus statement on diagnosis and treatment of adult ADHD: The European Network Adult ADHD. *Bmc Psychiatry*, 10.
- Kraemer, H. C. (2003). "Rules" of evidence in assessing the efficacy and effectiveness of treatments. *Developmental Neuropsychology*, 24(2-3), 705-718.
- Kraemer, H. C. & Blasey, C. M. (2004). Centring in regression analyses: a strategy to prevent errors in statistical inference. *International Journal of Methods in Psychiatric Research*, 13(3), 141-151.
- Kraemer, H. C., Wilson, G. T., Fairburn, C. G. & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry*, 59(10), 877-883.
- Krain, A. L. & Castellanos, F. X. (2006). Brain development and ADHD. *Clinical Psychology Review*, 26(4), 433-444.
- Kubinger, K. D., Rasch, D. & Moder, K. (2009). Zur Legende der Voraussetzungen des t-Tests für unabhängige Stichproben. *Psychologische Rundschau*, 60(1), 26-27.

- Lahey, B. B., Pelham, W. E., Loney, J., Lee, S. S. & Willcutt, E. (2005). Instability of the DSM-IV subtypes of ADHD from preschool through elementary school. *Archives of General Psychiatry*, 62(8), 896-902.
- Laird, R. D., Jordan, K. Y., Dodge, K. A., Pettit, G. S. & Bates, J. E. (2001). Peer rejection in childhood, involvement with antisocial peers in early adolescence, and the development of externalizing behavior problems. *Development and Psychopathology*, 13(2), 337-354.
- Langley, K., Heron, J., Smith, G. D. & Thapar, A. (2012). Maternal and paternal smoking during pregnancy and risk of ADHD symptoms in offspring: Testing for intrauterine effects. *American Journal of Epidemiology*, 176(3), 261-268.
- Lansford, J. E., Malone, P. S., Dodge, K. A., Pettit, G. S. & Bates, J. E. (2010). Developmental cascades of peer rejection, social information processing biases, and aggression during middle childhood. *Development and Psychopathology*, 22(3), 593-602.
- Larson, K., Russ, S. A., Kahn, R. S. & Halfon, N. (2011). Patterns of comorbidity, functioning, and service use for US children with ADHD, 2007. *Pediatrics*, 127(3), 462-470.
- Larsson, H., Dilshad, R., Lichtenstein, P. & Barker, E. D. (2011). Developmental trajectories of DSM-IV symptoms of attention-deficit/hyperactivity disorder: genetic effects, family risk and associated psychopathology. *Journal of Child Psychology and Psychiatry*, 52(9), 954-963.
- Larsson, H., Lichtenstein, P. & Larsson, J.-O. (2006). Genetic contributions to the development of ADHD subtypes from childhood to adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry*, 45(8), 973-981.
- Larsson, J. O., Larsson, H. & Lichtenstein, P. (2004). Genetic and environmental contributions to stability and change of ADHD symptoms between 8 and 13 years of age: A longitudinal twin study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 43(10), 1267-1275.
- Lauth, G. W. (2004). Förderung von Aufmerksamkeit und Konzentration. In G. W. Lauth, M. Grünke & J. C. Brunstein (Hrsg.), *Interventionen bei Lernstörungen* (S. 239–248). Göttingen: Hogrefe.
- Lauth, G. W. (2008). Selbstkontrollverfahren, kognitives Modellieren und Selbstinstruktionstraining. In G. W. Lauth, F. Linderkamp, S. Schneider & U. Brack (Hrsg.), *Verhaltenstherapie mit Kindern und Jugendlichen* (2., vollst. überarb. Aufl., S. 715-724). Weinheim: Beltz PVU.
- Lauth, G. W. & Fellner, C. (2004). Therapieverlauf und Langzeiteffekt eines multimodalen Trainingsprogramms bei Aufmerksamkeitsdefizit-/Hyperaktivitätsstörungen. *Kindheit und Entwicklung*, 13(3), 167-179.
- Lauth, G. W., Heubeck, B. G. & Mackowiak, K. (2006). Observation of children with attention-deficit hyperactivity (ADHD) problems in three natural classroom contexts. *British Journal of Educational Psychology*, 76, 385-404.
- Lauth, G. W. & Mackowiack, K. (2004). Unterrichtsverhalten von Kindern mit Aufmerksamkeitsdefizit-/Hyperaktivitätsstörungen. *Kindheit und Entwicklung*, 13(3), 158-166.

- Lauth, G. W. & Minsel, W.-R. (2009). *ADHS bei Erwachsenen: Diagnostik und Behandlung von Aufmerksamkeits-/Hyperaktivitätsstörungen*. Göttingen: Hogrefe.
- Lauth, G. W. & Minsel, W.-R. (2011). ADHS bei Erwachsenen - empirische Evaluation einer behavioralen Therapie nach Lauth und Minsel. In F. Linderkamp (Hrsg.), *ADHS im Jugend- und Erwachsenenalter: Bedingungsgefüge und Therapiekonzepte* (S. 55-71). Tübingen: dgvt-Verlag.
- Lauth, G. W. & Naumann, K. (2009). *ADHS in der Schule: Übungsprogramm für Lehrer*. Weinheim: Beltz.
- Lauth, G. W. & Schlotzke, P. F. (2009). *Training mit aufmerksamkeitsgestörten Kindern* (6. vollst. überarb. Aufl.). Weinheim: Beltz PVU.
- Leonhart, R. (2004). Effektgrößenberechnung bei Interventionsstudien. *Rehabilitation*, 43, 241-246.
- Li, D. W., Sham, P. C., Owen, M. J. & He, L. (2006). Meta-analysis shows significant association between dopamine system genes and attention deficit hyperactivity disorder (ADHD). *Human Molecular Genetics*, 15(14), 2276-2284.
- Lienert, G. A. & Raatz, U. (1998). *Testaufbau und Testanalyse* (6. Aufl.). Weinheim: Beltz PVU.
- Lincoln, T. M., Rief, W., Hahlweg, K., Frank, M., Von Witzleben, I., Schroeder, B. et al. (2005). Who Comes, Who Stays, Who Profits? Predicting refusal, dropout, success, and relapse in a short intervention for social phobia. *Psychotherapy Research*, 15(3), 210-225.
- Linderkamp, F. (2004). Motivierung (durch operante Verstärkung). In G. Lauth, M. Grünke & J. C. Brunstein (Hrsg.), *Interventionen bei Lernstörungen* (S. 124-133). Göttingen: Hogrefe.
- Linderkamp, F. (2007a). Kognitiv-behaviorale Therapieverfahren. In F. Linderkamp & M. Grünke (Hrsg.), *Lern- und Verhaltensstörungen. Genese - Diagnostik - Intervention* (S. 166-174). Weinheim: Beltz PVU.
- Linderkamp, F. (2007b). Operante Methoden. In F. Linderkamp & M. Grünke (Hrsg.), *Lern- und Verhaltensstörungen. Genese - Diagnostik - Intervention* (S. 156-165). Weinheim: Beltz PVU.
- Linderkamp, F., Hennig, T. & Schramm, S. A. (2011). *ADHS bei Jugendlichen. Das Lerntraining LeJA*. Weinheim: Beltz PVU.
- Linderkamp, F. & Lauth, G. (2011). Efficacy of pharmacological versus psychotherapeutic therapies in adults with Attention Deficit/Hyperactivity Disorder (ADHD): An empirical meta-analysis. *Verhaltenstherapie*, 21(4).
- Linn, R. L. & Slinde, J. A. (1977). The Determination of the Significance of Change Between Pre- and Posttesting Periods. *Review of Educational Research*, 47(1), 121-150.
- Linnet, K. M., Dalsgaard, S., Obel, C., Wisborg, K., Henriksen, T. B., Rodriguez, A. et al. (2003). Maternal lifestyle factors in pregnancy risk of attention deficit hyperactivity disorder and associated behaviors: Review of the current evidence. *American Journal of Psychiatry*, 160(6), 1028-1040.

- Longford, N. (2005). *Missing data and small-area estimation: Modern analytical equipment for the survey statistician*. New York: Springer.
- Lord, F. M. (1963). Elementary models for measuring change. In C. W. Harris (Ed.), *Problems in measuring change* (pp. 21-38). Madison: University of Wisconsin Press.
- Luborsky, L., Singer, B. & Luborsky, L. (1975). Comparative studies of psychotherapies: Is it true that "everyone has won and all must have prizes?". *Archives of General Psychiatry*, 32(8), 995-1008.
- Luciana, M., Conklin, H. M., Hooper, C. J. & Yarger, R. S. (2005). The Development of Nonverbal Working Memory and Executive Control Processes in Adolescents. *Child Development*, 76(3), 697-712.
- Lutz, W. & Grawe, K. (2007). Psychotherapieforschung: Grundlagen, Konzepte und neue Trends. In B. Strauß, F. Hohagen & F. Caspar (Hrsg.), *Lehrbuch Psychotherapie. Teilband 1* (S. 727-768). Göttingen: Hogrefe.
- MacCallum, R. C., Zhang, S. B., Preacher, K. J. & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, 7(1), 19-40.
- Maier-Riehle, B. & Zwingmann, C. (2000). Effektstärkevarianten beim Eingruppen-Prä-Post-Design: Eine kritische Betrachtung. *Die Rehabilitation*, 39(4), 189-199.
- Majewicz-Hefley, A. & Carlson, J. S. (2007). A meta-analysis of combined treatments for children diagnosed with ADHD. *Journal of attention disorders*, 10(3), 239-250.
- Maric, M., Wiers, R. W. & Prins, P. J. M. (2012). Ten ways to improve the use of statistical mediation analysis in the practice of child and adolescent treatment research. *Clinical Child and Family Psychology Review*, 15(3), 177-191.
- Max, J. E., Fox, P. T., Lancaster, J. L., Kochunov, P., Mathews, K., Manes, F. F. et al. (2002). Putamen lesions and the development of attention-deficit/hyperactivity symptomatology. *Journal of the American Academy of Child and Adolescent Psychiatry*, 41(5), 563-571.
- McBurnett, K. (2008). Sluggish cognitive tempo: The promise and problems of measuring syndromes in the attention spectrum. In K. McBurnett & L. Pfiffner (Eds.), *Attention deficit hyperactivity disorder: Concepts, controversies, new directions* (pp. 351-357). New York: informa.
- McBurnett, K., Pfiffner, L. J. & Frick, P. J. (2001). Symptom properties as a function of ADHD type: An argument for continued study of sluggish cognitive tempo. *Journal of Abnormal Child Psychology*, 29(3), 207-213.
- McCart, M. R., Priester, P. E., Davies, W. H. & Azen, R. (2006). Differential effectiveness of behavioral parent-training and cognitive-behavioral therapy for antisocial youth: A meta-analysis. *Journal of Abnormal Child Psychology*, 34(4), 527-543.
- McCleary, L. & Ridley, T. (1999). Parenting adolescents with ADHD: Evaluation of a psychoeducation group. *Patient Education and Counseling*, 38(1), 3-10.

- McGough, J. J. & Barkley, R. A. (2004). Diagnostic controversies in adult attention deficit hyperactivity disorder. *American Journal of Psychiatry*, 161(11), 1948-1956.
- McGough, J. J., Smalley, S. L., McCracken, J. T., Yang, M., Del'Homme, M., Lynn, D. E. et al. (2005). Psychiatric comorbidity in adult attention deficit hyperactivity disorder: Findings from multiplex families. *American Journal of Psychiatry*, 162(9), 1621-1627.
- McKnight, P. E., McKnight, K. M., Sidani, S. & Figueredo, A. J. (2007). *Missing Data - A Gentle Introduction*. New York: The Guilford Press.
- Meichenbaum, D. & Goodman, J. (1971). Training impulsive children to talk to themselves: A means of developing self-control. *Journal of Abnormal Psychology*, 77(2), 115-126.
- Merton, R. K. (1968). The Matthew effect in science. *Science*, 159, 56-63.
- Meyer, A.-E. (1990). Eine Taxonomie der bisherigen Psychotherapieforschung. Editorial. *Zeitschrift für Klinische Psychologie*, 19(4), 287-291.
- Michalak, J., Kosfelder, J., Meyer, F. & Schulte, D. (2003). Messung des Therapieerfolgs: Veränderungsmaße oder retrospektive Erfolgsbeurteilung. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 32(2), 94-103.
- Michielsen, M., Semeijn, E., Comijs, H. C., van de Ven, P., Beekman, A. T. F., Deeg, D. J. H. et al. (2012). Prevalence of attention-deficit hyperactivity disorder in older adults in The Netherlands. *British Journal of Psychiatry*, 201(4), 298-305.
- Mick, E., Biederman, J., Faraone, S. V., Sayer, J. & Kleinman, S. (2002). Case-control study of attention-deficit hyperactivity disorder and maternal smoking, alcohol use, and drug use during pregnancy. *Journal of the American Academy of Child and Adolescent Psychiatry*, 41(4), 378-385.
- Mick, E., Biederman, J., Prince, J., Fischer, M. J. & Faraone, S. V. (2002). Impact of low birth weight on attention-deficit hyperactivity disorder. *Journal of Developmental and Behavioral Pediatrics*, 23(1), 16-22.
- Mikami, A. Y. & Hinshaw, S. P. (2008). ADHD in girls. In K. McBurnett & L. Pfiffner (Eds.), *Attention deficit hyperactivity disorder: Concepts, controversies, new directions* (pp. 259-272). New York: informa.
- Milich, R., Balentine, A. C. & Lynam, D. R. (2001). ADHD combined type and ADHD predominantly inattentive type are distinct and unrelated disorders. *Clinical Psychology-Science and Practice*, 8(4), 463-488.
- Miller, W. R. & Rollnick, S. (2009). *Motivierende Gesprächsführung* (3., unveränd. Aufl.). Freiburg: Lambertus.
- Millichap, J. G. (2008). Etiologic classification of attention-deficit/hyperactivity disorder. *Pediatrics*, 121(2), E358-E365.
- Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, 100(4), 674-701.
- Moher, D., Schulz, K. F. & Altman, D. G. (2004). Das CONSORT Statement: Überarbeitete Empfehlungen zur Qualitätsverbesserung von Reports

- randomisierter Studien im Parallel-Design. *Deutsche medizinische Wochenschrift*, 129(S 3), T16-T20.
- Mohr, D. C. (1995). Negative outcome in psychotherapy: A critical review. *Clinical Psychology-Science and Practice*, 2(1), 1-27.
- Montoya, A., Colom, F. & Ferrin, M. (2011). Is psychoeducation for parents and teachers of children and adolescents with ADHD efficacious? A systematic literature review. *European Psychiatry*, 26(3), 166-175.
- Moosbrugger, H. & Schermelleh-Engel, K. (2012). Exploratorische (EFA) und Konfirmatorische Faktorenanalyse (CFA). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 325-343). Berlin: Springer.
- Morris, G. B. (1993). A rational-emotive treatment program with conduct disorder and attention-deficit hyperactivity disorder adolescents. *Journal of Rational-Emotive and Cognitive-Behavior Therapy*, 11(3), 123-134.
- Morton, J. & Frith, U. (1995). Causal modeling: A structural approach to developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology* (pp. 357-390). New York: Wiley.
- Mrug, S., Molina, B. S. G., Hoza, B., Gerdes, A. C., Hinshaw, S. P., Hechtman, L. et al. (2012). Peer rejection and friendships in children with Attention-deficit/hyperactivity disorder: Contributions to long-term outcomes. *Journal of Abnormal Child Psychology*, 40(6), 1013-1026.
- Muris, P., Meesters, C. & van den Berg, F. (2003). The Strengths and Difficulties Questionnaire (SDQ) - Further evidence for its reliability and validity in a community sample of Dutch children and adolescents. *European Child & Adolescent Psychiatry*, 12(1), 1-8.
- Murray-Close, D., Hoza, B., Hinshaw, S. P., Arnold, L. E., Swanson, J., Jensen, P. S. et al. (2010). Developmental processes in peer problems of children with attention-deficit/hyperactivity disorder in The Multimodal Treatment Study of Children With ADHD: Developmental cascades and vicious cycles. *Development and Psychopathology*, 22(4), 785-802.
- Muthén, B. & Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism-Clinical and Experimental Research*, 24(6).
- Nachtigall, C. & Suhl, U. (2002). Der Regressionseffekt: Mythos und Wirklichkeit. *methevalreport*, 4(2).
- Nachtigall, C. & Suhl, U. (2005). Evaluation intraindividuelle Veränderung. Ein Vergleich verschiedener Veränderungskennwerte. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 34(4), 241-247.
- Nangle, D. W., Erdley, C., Carpenter, E. M. & Newman, J. E. (2002). Social skills training as a treatment for aggressive children and adolescents: A developmental-clinical integration. *Aggression and Violent Behavior*, 7(2), 169-199.
- National Institute for Health and Care Excellence. (2008). *Attention deficit hyperactivity disorder: Diagnosis and management of ADHD in children, young people and adults [CG72]*. London: National Institute for Health and Care Excellence.

- Nave-Herz, R. (2007). *Familie heute. Wandel der Familienstrukturen und Folgen für die Erziehung*. Darmstadt: Primus.
- Newcorn, J. H. (2008). Co-morbidity in adults with ADHD. *CNS Spectrums*, 13(8), 12-15.
- Newcorn, J. H., Halperin, J. M., Jensen, P. S., Abikoff, H. B., Arnold, L. E., Cantwell, D. P. et al. (2001). Symptom profiles in children with ADHD: Effects of comorbidity and gender. *Journal of the American Academy of Child and Adolescent Psychiatry*, 40(2), 137-146.
- Nigg, J. T., Willcutt, E. G., Doyle, A. E. & Sonuga-Barke, E. J. S. (2005). Causal heterogeneity in attention-deficit/hyperactivity disorder: Do we need neuropsychologically impaired subtypes? *Biological Psychiatry*, 57(11), 1224-1230.
- Nolen-Hoeksema, S. & Girgus, J. S. (1994). The emergence of gender differences in depression during adolescence. *Psychological Bulletin*, 115(3), 424-443.
- Normand, S., Flora, D., Toplak, M. & Tannock, R. (2012). Evidence for a General ADHD Factor from a Longitudinal General School Population Study. *Journal of Abnormal Child Psychology*, 40(4), 555-567.
- Normand, S., Schneider, B. H., Lee, M. D., Maisonneuve, M.-F., Kuehn, S. M. & Robaey, P. (2011). How do children with ADHD (mis)manage their real-life dyadic friendships? A multi-method investigation. *Journal of Abnormal Child Psychology*, 39(2), 293-305.
- Novik, T. S., Hervas, A., Ralston, S. J., Dalsgaard, S., Pereira, R. R., Lorenzo, M. J. et al. (2006). Influence of gender on Attention-Deficit/Hyperactivity Disorder in Europe - ADORE. *European Child & Adolescent Psychiatry*, 15, 15-24.
- Nussbaum, N. L. (2012). ADHD and female specific concerns: A review of the literature and clinical implications. *Journal of Attention Disorders*, 16(2), 87-100.
- O'Halloran, C. J., Kinsella, G. J. & Storey, E. (2011). The cerebellum and neuropsychological functioning: A critical review. *Journal of Clinical and Experimental Neuropsychology*, 34(1), 35-56.
- Oerter, R. & Dreher, E. (2008). Jugendalter. In R. Oerter & L. Montada (Hrsg.), *Entwicklungspsychologie* (6., vollst. überarb. Aufl.). Weinheim: Beltz.
- Ogles, B. M., Lunnen, K. M. & Bonesteel, K. (2001). Clinical significance: History, application, and current practice. *Clinical Psychology Review*, 21(3), 421-446.
- Ohan, J. L., Visser, T. A. W., Strain, M. C. & Allen, L. (2011). Teachers' and education students' perceptions of and reactions to children with and without the diagnostic label "ADHD". *Journal of School Psychology*, 49(1), 81-105.
- Ollendick, T. H., Jarrett, M. A., Grills-Taquechel, A. E., Hovey, L. D. & Wolff, J. C. (2008). Comorbidity as a predictor and moderator of treatment outcome in youth with anxiety, affective, attention deficit/hyperactivity disorder, and oppositional/conduct disorders. *Clinical Psychology Review*, 28(8), 1447-1471.
- Owens, E. B., Hinshaw, S. P., Kraemer, H. C., Arnold, L. E., Abikoff, H. B., Cantwell, D. P. et al. (2003). Which treatment for whom for ADHD? Moderators of

- treatment response in the MTA. *Journal of Consulting and Clinical Psychology*, 71(3), 540-552.
- Patterson, G. R. (1982). *Coercive family process*. Eugene: Castalia.
- Pelham, W. E. & Fabiano, G. A. (2008). Evidence-based psychosocial treatments for attention-deficit/hyperactivity disorder. *Journal of Clinical Child and Adolescent Psychology*, 37(1), 184-214.
- Pelz, R., Banaschewski, T. & Becker, K. (2008). Pharmakotherapie bei Kindern und Jugendlichen mit ADHS. *Monatsschrift Kinderheilkunde*, 156(8), 768-775.
- Petermann, F. (1978). *Veränderungsmessung*. Stuttgart: Kohlhammer.
- Petermann, F., Niebank, K. & Scheithauer, H. (2004). *Entwicklungswissenschaft*. Berlin: Springer.
- Petermann, F. & Toussaint, A. (2009). Neuropsychologische Diagnostik bei Kindern mit ADHS. *Kindheit und Entwicklung*, 18(2), 83-94.
- Piaget, J. (1947). *Psychologie der Intelligenz*. Zürich: Rascher.
- Piaget, J. (1969). The intellectual development of the adolescent. In G. Caplan & S. Lebovici (Eds.), *Adolescence: Psychosocial perspectives*. New York: Basic Books.
- Pliszka, S. R. (2003). Non-stimulant treatment of attention-deficit/hyperactivity disorder. *CNS Spectrums*, 8(4), 253-258.
- Pocock, S. J., Assmann, S. E., Enos, L. E. & Kasten, L. E. (2002). Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. *Statistics in Medicine*, 21(19), 2917-2930.
- Polanczyk, G., de Lima, M. S., Horta, B. L., Biederman, J. & Rohde, L. A. (2007). The worldwide prevalence of ADHD: A systematic review and meta-regression analysis. *American Journal of Psychiatry*, 164(6), 942-948.
- Polanczyk, G. V. Personal communication, May 05, 2009.
- Preacher, K. J., Rucker, D. D., MacCallum, R. C. & Nicewander, W. A. (2005). Use of the extreme groups approach: A critical reexamination and new recommendations. *Psychological Methods*, 10(2), 178-192.
- Pólya, G. (1949). *Schule des Denkens: Vom Lösen mathematischer Probleme*. Bern: Francke.
- Ramsay, J. R. & Rostain, A. L. (2007). Psychosocial treatments for attention-deficit/hyperactivity disorder in adults: Current evidence and future directions. *Professional Psychology-Research and Practice*, 38(4), 338-346.
- Rapport, M. D., Chung, K. M., Shore, G. & Isaacs, P. (2001). A conceptual model of child psychopathology: Implications for understanding attention deficit hyperactivity disorder and treatment efficacy. *Journal of Clinical Child Psychology*, 30(1), 48-58.
- Rapport, M. D., Kofler, M. J., Alderson, R. M., Timko, T. M. & DuPaul, G. J. (2009). Variability of attention processes in ADHD observations from the classroom. *Journal of Attention Disorders*, 12(6), 563-573.

- Renner, G., Lessing, T., Krampen, G. & Irblich, D. (2012). Reliabilität und Retest-Stabilität der „Testbatterie zur Aufmerksamkeitsprüfung für Kinder“ (KITAP) bei 6- bis 7-jährigen Kindern. *Zeitschrift für Neuropsychologie*, 23(1), 27-36.
- Resch, F. (1999). *Entwicklungspsychopathologie des Kindes- und Jugendalters*. Weinheim: Beltz Psychologie Verlags Union.
- Resch, F., Parzer, P. & Brunner, R. (2008). Self-mutilation and suicidal behaviour in children and adolescents: prevalence and psychosocial correlates: results of the BELLA study. *European Child & Adolescent Psychiatry*, 17, 92-98.
- Retz, W., Stieglitz, R.-D., Corbisiero, S., Retz-Junginger, P. & Roesler, M. (2012). Emotional dysregulation in adult ADHD: what is the empirical evidence? *Expert Review of Neurotherapeutics*, 12(10), 1241-1251.
- Retz-Junginger, P., Retz, W., Schneider, M., Schwitzgebel, P., Steinbach, E., Henges, G. et al. (2007). Der Einfluss des Geschlechts auf die Selbstbeschreibung kindlicher ADHS-Symptome. *Der Nervenarzt*, 78(9), 1046-1051.
- Retz-Junginger, P., Sobanski, E., Alm, B., Retz, W. & Rösler, M. (2008). Alters- und geschlechtsspezifische Besonderheiten der Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung. *Der Nervenarzt*, 79(7), 809-819.
- Rice, F., Harold, G. T., Boivin, J., van den Bree, M., Hay, D. F. & Thapar, A. (2010). The links between prenatal stress and offspring development and psychopathology: disentangling environmental and inherited influences. *Psychological Medicine*, 40(2), 335-345.
- Rindermann, H. (2011). Intelligenzwachstum in Kindheit und Jugend. *Psychologie in Erziehung und Unterricht*, 58, 210-224.
- Robin, A. L. (1998). *ADHD in adolescents: Diagnosis and treatment*. New York: Guilford.
- Robison, R. J., Reimherr, F. W., Marchant, B. K., Faraone, S. V., Adler, L. A. & West, S. A. (2008). Gender differences in 2 clinical trials of adults with attention-deficit/hyperactivity disorder: A retrospective data analysis. *Journal of Clinical Psychiatry*, 69(2), 213-221.
- Rogosa, D. (1988). Myths about Longitudinal Research. In K. W. Schaie, R. T. Campbell, W. Meredith & S. C. Rawlings (Eds.), *Methodological Issues in Aging Research* (pp. 171-209). New York: Springer.
- Rommelse, N. N. J., Altink, M. E., Martin, N. C., Buschgens, C. J. M., Faraone, S. V., Buitelaar, J. K. et al. (2008). Relationship between endophenotype and phenotype in ADHD. *Behavioral and Brain Functions*, 4.
- Rooney, B. L. & Murray, D. M. (1996). A meta-analysis of smoking prevention programs after adjustment for errors in the unit of analysis. *Health Education Quarterly*, 23(1), 48-64.
- Rooney, M., Chronis-Tuscano, A. & Yoon, Y. (2012). Substance use in college students with ADHD. *Journal of Attention Disorders*, 16(3), 221-234.
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86(3), 638-641.

- Rothenberger, A., Becker, A., Erhart, M., Wille, N., Ravens-Sieberer, U. & The Bella study group. (2008). Psychometric properties of the parent strengths and difficulties questionnaire in the general population of German children and adolescents: Results of the BELLA study. *European Child & Adolescent Psychiatry*, 17(0), 99-105.
- Rucklidge, J. J. & Tannock, R. (2002). Neuropsychological profiles of adolescents with ADHD: Effects of reading difficulties and gender. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 43(8), 988-1003.
- Rutledge, K. J., van den Bos, W., McClure, S. M. & Schweitzer, J. B. (2012). Training cognition in ADHD: Current findings, borrowed concepts, and future directions. *Neurotherapeutics*, 9(3), 542-558.
- Rutter, M. (1989). Pathways from childhood to adult life. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 30(1), 23-51.
- Rutter, M. (2005). Multiple meanings of a developmental perspective on psychopathology. *European Journal of Developmental Psychology*, 2(3), 221-252.
- Rutter, M., Caspi, A. & Moffitt, T. E. (2003). Using sex differences in psychopathology to study causal mechanisms: unifying issues and research strategies. *Journal of Child Psychology and Psychiatry*, 44(8), 1092-1115.
- Rutter, M., Kim-Cohen, J. & Maughan, B. (2006). Continuities and discontinuities in psychopathology between childhood and adult life. [Review]. *Journal of Child Psychology and Psychiatry*, 47(3-4), 276-295.
- Rösler, M., Retz, W., Retz-Junginger, P., Stieglitz, R. D., Kessler, H., Reimherr, F. et al. (2008). ADHS-Diagnose bei Erwachsenen. *Der Nervenarzt*, 79(3), 320-327.
- Rösler, M., Retz-Junginger, P., Retz, W. & Stieglitz, R.-D. (2008). *HASE. Homburger ADHS-Skalen für Erwachsene*. Göttingen: Hogrefe.
- Safren, S. A., Otto, M. W., Sprich, S., Winett, C. L., Wilens, T. E. & Biederman, J. (2005). Cognitive-behavioral therapy for ADHD in medication-treated adults with continued symptoms. *Behaviour Research and Therapy*, 43(7), 831-842.
- Sagvolden, T., Aase, H., Zeiner, P. & Berger, D. (1998). Altered reinforcement mechanisms in attention-deficit/hyperactivity disorder. *Behavioural Brain Research*, 94(1), 61-71.
- Saile, H. (2007). Psychometrische Befunde zur Lehrerversion des "Strengths and Difficulties Questionnaire" (SDQ-L). Validierung anhand soziometrischer Indizes. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 39(1), 25-32.
- Sameroff, A. J. (2000). Developmental systems and psychopathology. *Development and Psychopathology*, 12(3).
- Sanders, A. F. (1983). Towards a model of stress and human performance. *Acta Psychologica*, 53(1), 61-97.
- Sarris, J., Kean, J., Schweitzer, I. & Lake, J. (2011). Complementary medicines (herbal and nutritional products) in the treatment of Attention Deficit Hyperactivity Disorder (ADHD): A systematic review of the evidence. *Complementary Therapies in Medicine*, 19(4), 216-227.

- Saß, H., Wittchen, H.-U., Zaudig, M. & Houben, I. (2003). *Diagnostisches und Statistisches Manual Psychischer Störungen (DSM-IV-TR)*. Göttingen: Hogrefe.
- Schachar, R. & Crosbie, J. (2008). Endophenotypes in ADHD: Rational and progress. In K. McBurnett & L. Pfiffner (Eds.), *Attention deficit hyperactivity disorder: Concepts, controversies, new directions* (pp. 373-395). New York: informa.
- Schafer, J. L. & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7(2), 147-177.
- Schermelleh-Engel, K. & Werner, C. S. (2012). Methoden der Reliabilitätsbestimmung. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 119-141). Berlin: Springer.
- Schlack, R., Hölling, H., Kurth, B. M. & Huss, M. (2007). Die Prävalenz der Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS) bei Kindern und Jugendlichen in Deutschland. Erste Ergebnisse aus dem Kinder- und Jugendgesundheitsurvey (KiGGS). *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*, 50(5-6), 827-835.
- Schlagmüller, M. & Schneider, W. (2007). *Würzburger Lesestrategie-Wissenstest für die Klassen 7-12 (WLST 7-12)*. Göttingen: Hogrefe.
- Schmidt, S., Brähler, E., Petermann, F. & Koglin, U. (2012). Komorbide Belastungen bei Jugendlichen und jungen Erwachsenen mit ADHS. *Zeitschrift für Psychiatrie, Psychologie und Psychotherapie*, 60(1), 015-026.
- Schmidt, S. & Petermann, F. (2008). Entwicklungspsychopathologie der ADHS. *Zeitschrift für Psychiatrie, Psychologie und Psychotherapie*, 56(4), 265-274.
- Scholing, A. & Emmelkamp, P. M. G. (1999). Prediction of treatment outcome in social phobia: a cross-validation. *Behaviour Research and Therapy*, 37(7), 659-670.
- Schramm, S. A. (im Druck). *Intervention bei Aufmerksamkeitsstörungen im Jugendalter - Konzeption und empirische Evaluation des Lerntrainings für Jugendliche mit Aufmerksamkeitsstörungen (LeJA)*. Universität Oldenburg: unveröffentlichte Dissertation.
- Schreyer, I. & Hampel, P. (2009). ADHS bei Jungen im Kindesalter - Lebensqualität und Erziehungsverhalten. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, 37(1), 69-75.
- Schuermans, J., Comijs, H., Emmelkamp, P. M. G., Weijnen, I. J. C., van den Hout, M. & van Dyck, R. (2009). Long-term effectiveness and prediction of treatment outcome in cognitive behavioral therapy and sertraline for late-life anxiety disorders. *International Psychogeriatrics*, 21(6), 1148-1159.
- Schütte, M., Wirth, J. & Leutner, D. (2012). Lernstrategische Teilkompetenzen für das selbstregulierte Lernen aus Sachtexten. *Psychologische Rundschau*, 63(1), 26-33.
- Sciutto, M. J., Nolfi, C. J. & Bluhm, C. (2004). Effects of child gender and symptom type on referrals for ADHD by elementary school teachers. *Journal of Emotional and Behavioral Disorders*, 12(4), 247-253.
- Sedlmeier, P. (2009). Beyond the significance test ritual: What is there? *Zeitschrift für Psychologie/Journal of Psychology*, 217(1).

- Sedlmeier, P. & Renkewitz, F. (2008). *Forschungsmethoden und Statistik in der Psychologie*. München: Pearson.
- Seidman, L. J. (2006). Neuropsychological functioning in people with ADHD across the lifespan. *Clinical Psychology Review, 26*(4), 466-485.
- Seidman, L. J., Biederman, J., Monuteaux, M. C., Valera, E., Doyle, A. E. & Faraone, S. V. (2005). Impact of gender and age on executive functioning: Do girls and boys with and without attention deficit hyperactivity disorder differ neuropsychologically in preteen and teenage years? *Developmental Neuropsychology, 27*(1), 79-105.
- Seidman, L. J., Valera, E. M. & Makris, N. (2005). Structural brain imaging of attention-deficit/hyperactivity disorder. *Biological Psychiatry, 57*(11), 1263-1272.
- Seiffge-Krenke, I. (1995). *Stress, coping, and relationships in adolescence*. Mahwah: Erlbaum.
- Seiffge-Krenke, I. (2008). Schulstress in Deutschland: Ursachen, Häufigkeiten und internationale Verortung. *Praxis der Kinderpsychologie und Kinderpsychiatrie, 57*, 3-19.
- Sergeant, J. (2000). The cognitive-energetic model: An empirical approach to Attention-deficit hyperactivity disorder. *Neuroscience and Biobehavioral Reviews, 24*(1), 7-12.
- Sergeant, J. A. (2005). Modeling attention-deficit/hyperactivity disorder: A critical appraisal of the cognitive-energetic model. *Biological Psychiatry, 57*(11), 1248-1255.
- Seymour, K. E., Chronis-Tuscano, A., Halldorsdottir, T., Stupica, B., Owens, K. & Sacks, T. (2012). Emotion regulation mediates the relationship between ADHD and depressive symptoms in youth. *Journal of Abnormal Child Psychology, 40*(4), 595-606.
- Shaw, P., Eckstrand, K., Sharp, W., Blumenthal, J., Lerch, J. P., Greenstein, D. et al. (2007). Attention-deficit/hyperactivity disorder is characterized by a delay in cortical maturation. *Proceedings of the National Academy of Sciences of the United States of America, 104*(49), 19649-19654.
- Shaywitz, B. A., Holford, T. R., Holahan, J. M., Fletcher, J. M., Stuebing, K. K., Francis, D. J. et al. (1995). A Matthew effect for IQ but not for reading: Results from a longitudinal study. *Reading Research Quarterly, 30*(4), 894-906.
- Siegler, R. (1996). *Emerging minds. The process of change in children's thinking*. New York: Oxford University Press.
- Siegler, R., DeLoache, J. & Eisenberg, N. (2008). *Entwicklungspsychologie im Kindes- und Jugendalter*. Heidelberg: Spektrum.
- Siegler, R. S. (2007). Cognitive variability. *Developmental Science, 10*(1), 104-109.
- Simon, V., Czobor, P., Balint, S., Meszaros, A. & Bitter, I. (2009). Prevalence and correlates of adult attention-deficit hyperactivity disorder: meta-analysis. *British Journal of Psychiatry, 194*(3), 204-211.

- Skinner, E. A. & Zimmer-Gembeckz, M. J. (2007). The development of coping *Annual Review of Psychology* (Vol. 58, pp. 119-144). Palo Alto: Annual Reviews.
- Slavin, R. E. & Smith, D. (2008). *Effects of sample size on effect size in systematic reviews in education*. Crystal City, Virginia: Annual meetings of the Society for Research on Effective Education.
- Smedje, H., Broman, J. E., Hetta, J. & von Knorring, A. L. (1999). Psychometric properties of a Swedish version of the "Strengths and Difficulties Questionnaire". *European Child & Adolescent Psychiatry*, 8(2), 63-70.
- Smetana, J. G., Campione-Barr, N. & Metzger, A. (2006). Adolescent development in interpersonal and societal contexts. *Annual Review of Psychology*, 57, 255-284.
- Smith, B. H., Waschbusch, D. A., Willoughby, M. T. & Evans, S. (2000). The efficacy, safety, and practicality of treatments for adolescents with attention-deficit/hyperactivity disorder (ADHD). *Clinical child and family psychology review*, 3(4), 243-267.
- Snow, R. E. (1991). Aptitude-treatment interaction as a framework for research on individual differences in psychotherapy. *Journal of Consulting and Clinical Psychology*, 59(2), 205-216.
- Sobanski, E. (2006). Psychiatric comorbidity in adults with attention-deficit/hyperactivity disorder (ADHD). *European Archives of Psychiatry and Clinical Neuroscience*, 256, 26-31.
- Sobanski, E. & Alm, B. (2004). Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung (ADHS) bei Erwachsenen. Ein Überblick. *Nervenarzt*, 75(7), 697-715.
- Sodian, B. (1998). Theorien der kognitiven Entwicklung. In H. Keller (Ed.), *Lehrbuch Entwicklungspsychologie* (pp. 147-169).
- Solanto, M. V., Marks, D. J., Wasserstein, J., Mitchell, K., Abikoff, H., Alvir, J. M. J. et al. (2010). Efficacy of meta-cognitive therapy for adult ADHD. *American Journal of Psychiatry*, 167(8), 958-968.
- Sonuga-Barke, E. J. S. (2002). Psychological heterogeneity in AD/HD - a dual pathway model of behaviour and cognition. *Behavioural Brain Research*, 130(1-2), 29-36.
- Sonuga-Barke, E. J. S. (2003). The dual pathway model of AD/HD: An elaboration of neuro-developmental characteristics. *Neuroscience & Biobehavioral Reviews*, 27(7), 593-604.
- Sonuga-Barke, E. J. S. & Goldfoot, M. T. (1995). The effect of child hyperactivity on mothers' expectations for development. *Child Care Health and Development*, 21(1), 17-29.
- Speer, D. C. (1992). Clinically Significant Change: Jacobson and Truax (1991) Revisited. *Journal of Consulting and Clinical Psychology*, 60(3), 402-408.
- Spencer, T., Heiligenstein, J. H., Biederman, J., Faries, D. E., Kratochvil, C. J., Conners, C. K. et al. (2002). Results from 2 proof-of-concept, placebo-controlled studies of atomoxetine in children with attention-deficit/hyperactivity disorder. *Journal of Clinical Psychiatry*, 63(12), 1140-1147.

- Spivack, G. & Shure, M. B. (1974). *Social adjustment of young children: A cognitive approach to solving real-life problems*. San Francisco: Jossey-Bass.
- Sroufe, L. A. (1997). Psychopathology as an outcome of development. *Development and Psychopathology*, 9(2).
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21(4), 360–407.
- Stein, M. A., Sarampote, C. S., Waldman, I. D., Robb, A. S., Conlon, C., Pearl, P. L. et al. (2003). A Dose-Response Study of OROS Methylphenidate in Children With Attention-Deficit/Hyperactivity Disorder. *Pediatrics*, 112(5), e404.
- Steinberg, L. (2005a). *Adolescence*. New York: McGrawHill.
- Steinberg, L. (2005b). Cognitive and affective development in adolescence. *Trends in Cognitive Sciences*, 9(2), 69-74.
- Steinberg, L. & Morris, A. S. (2001). Adolescent development. *Annual Review of Psychology*, 52, 83-110.
- Steketee, G. & Chambless, D. L. (1992). Methodological issues in prediction of treatment outcome. *Clinical Psychology Review*, 12(4), 387-400.
- Stepp, S. D., Burke, J. D., Hipwell, A. E. & Loeber, R. (2012). Trajectories of Attention Deficit Hyperactivity Disorder and Oppositional Defiant Disorder Symptoms as Precursors of Borderline Personality Disorder Symptoms in Adolescent Girls. *Journal of Abnormal Child Psychology*, 40(1), 7-20.
- Steyer, R., Hannover, W., Telsler, C. & Kriebel, R. (1997). Zur Evaluation intraindividuelle Veränderung. *Zeitschrift für Klinische Psychologie*, 26(4), 291-299.
- Stice, E. & Shaw, H. (2004). Eating disorder prevention programs: A meta-analytic review. *Psychological Bulletin*, 130(2).
- Stice, E., Shaw, H., Bohon, C., Marti, C. N. & Rohde, P. (2009). A Meta-Analytic Review of Depression Prevention Programs for Children and Adolescents: Factors That Predict Magnitude of Intervention Effects. *Journal of Consulting and Clinical Psychology*, 77(3).
- Strang-Karlsson, S., Raikkonen, K., Pesonen, A.-K., Kajantie, E., Paavonen, E. J., Lahti, J. et al. (2008). Very low birth weight and behavioral symptoms of attention deficit hyperactivity disorder in young adulthood: The Helsinki study of very-low-birth-weight adults. *American Journal of Psychiatry*, 165(10), 1345-1353.
- Sturm, W., Fimm, B., Cantagallo, A., Cremel, N., North, P., Passadori, A. et al. (2003). Specific Computerized Attention Training in Stroke and Traumatic Brain-Injured Patients. *Zeitschrift für Neuropsychologie*, 14(4), 283-292.
- Swanson, H. L. & Deshler, D. (2003). Instructing adolescents with learning disabilities: Converting a meta-analysis to practice. *Journal of Learning Disabilities*, 36(2), 124-135.

- Swanson, H. L. & Hoskyn, M. (1998). Experimental intervention research on students with learning disabilities: A meta-analysis of treatment outcomes. *Review of Educational Research*, 68(3), 277-321.
- Tabachnick, B. G. & Fidell, L. S. (2007). *Using Multivariate Statistics* (5 ed.). Boston: Pearson.
- Taurines, R., Schmitt, J., Renner, T., Conner, A., Warnke, A. & Romanos, M. (2010). Developmental comorbidity in attention-deficit/hyperactivity disorder. *ADHD Attention Deficit and Hyperactivity Disorders*, 2(4), 267-289.
- Teicher, M. H., Polcari, A., Foulrigas, N., Vitaliano, G. & Navalta, C. P. (2012). Hyperactivity persists in male and female adults with ADHD and remains a highly discriminative feature of the disorder: A case-control study. *BMC Psychiatry*, 12, 190.
- Thakur, G. A., Sengupta, S. M., Grizenko, N., Schmitz, N., Page, V. & Joober, R. (2013). Maternal smoking during pregnancy and ADHD: A comprehensive clinical and neurocognitive characterization. *Nicotine & Tobacco Research*, 15(1), 149-157.
- Thapar, A., Cooper, M., Eyre, O. & Langley, K. (2013). Practitioner Review: What have we learnt about the causes of ADHD? *Journal of Child Psychology and Psychiatry*, 54(1), 3-16.
- Thapar, A., Cooper, M., Jefferies, R. & Stergiakouli, E. (2012). What causes attention deficit hyperactivity disorder? *Archives of Disease in Childhood*, 97(3), 260-265.
- Thapar, A., O'Donovan, M. & Owen, M. J. (2005). The genetics of attention deficit hyperactivity disorder. *Human Molecular Genetics*, 14, R275-R282.
- Theule, J., Wiener, J., Tannock, R. & Jenkins, J. M. (2013). Parenting stress in families of children with ADHD: A meta-analysis. *Journal of Emotional and Behavioral Disorders*, 21(1), 3-17.
- Todd, R. D., Huang, H., Todorov, A. A., Neuman, R. J., Reiersen, A. M., Henderson, C. A. et al. (2008). Predictors of stability of attention-deficit/hyperactivity disorder subtypes from childhood to young adulthood. *Journal of the American Academy of Child and Adolescent Psychiatry*, 47(1), 76-85.
- Toplak, M. E., Connors, L., Shuster, J., Knezevic, B. & Parks, S. (2008). Review of cognitive, cognitive-behavioral, and neural-based interventions for Attention-Deficit/Hyperactivity Disorder (ADHD). *Clinical Psychology Review*, 28(5), 801-823.
- Toplak, M. E., Sorge, G. B., Flora, D. B., Chen, W., Banaschewski, T., Buitelaar, J. et al. (2012). The hierarchical factor model of ADHD: invariant across age and national groupings? *Journal of Child Psychology and Psychiatry*, 53(3), 292-303.
- Trautmann-Villalba, P., Gerhold, M., Polowczyk, M., Dinter-Jörg, M., Laucht, M., Esser, G. et al. (2001). Mutter-Kind-Interaktion und externalisierende Störungen bei Kindern im Grundschulalter. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, 29(4), 263-273.
- Tucha, O., Walitza, S., Mecklinger, L., Sontag, T. A., Kubber, S., Linder, M. et al. (2006). Attentional functioning in children with ADHD - predominantly

- hyperactive-impulsive type and children with ADHD - combined type. *J Neural Transm*, 113(12), 1943-1953.
- Turgay, A., Goodman, D. W., Asherson, P., Lasser, R. A., Babcock, T. F., Pucci, M. L. et al. (2012). Lifespan persistence of ADHD: The life transition model and its application. *Journal of Clinical Psychiatry*, 73(2), 192-201.
- Unnever, J. D. & Cornell, D. G. (2003). Bullying, self-control, and ADHD. *Journal of Interpersonal Violence*, 18(2), 129-147.
- Urban, D. & Mayerl, J. (2011). *Regressionsanalyse: Theorie, Technik und Anwendung* (4., überarb. u. erw. Aufl.). Wiesbaden: VS Verlag.
- Valera, E. M., Faraone, S. V., Murray, K. E. & Seidman, L. J. (2007). Meta-analysis of structural imaging findings in attention-deficit/hyperactivity disorder. *Biological Psychiatry*, 61(12), 1361-1369.
- Van der Oord, S., Prins, P. J. M., Oosterlaan, J. & Emmelkamp, P. M. G. (2008). Efficacy of methylphenidate, psychosocial treatments and their combination in school-aged children with ADHD: A meta-analysis. *Clinical Psychology Review*, 28(5).
- van der Oord, S., Prins, P. J. M., Oosterlaan, J. & Emmelkamp, P. M. G. (2008). Treatment of attention deficit hyperactivity disorder in children. *European Child & Adolescent Psychiatry*, 17(2), 73-81.
- Voeller, K. K. S. (2004). Attention-deficit hyperactivity disorder (ADHD). *Journal of Child Neurology*, 19(10), 798-814.
- Walberg, H. J. & Tsai, S. L. (1983). Matthew effects in education. *American Educational Research Journal*, 20(3), 359-373.
- Waldman, I. D. & Gizer, I. R. (2006). The genetics of attention deficit hyperactivity disorder. *Clinical Psychology Review*, 26(4), 396-432.
- Weber, P., Bubl, R. & Lütschg, J. (2003). Nebenwirkungen einer Methylphenidat-Therapie bei Schulkindern. *Monatsschrift Kinderheilkunde*, 151(4), 399-404.
- Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, 92(4), 548-573.
- Weinstock, M. (2008). The long-term behavioural consequences of prenatal stress. *Neuroscience and Biobehavioral Reviews*, 32(6), 1073-1086.
- Weiss, M., Worling, D. & Wasdell, M. (2003). A chart review study of the inattentive and combined types of ADHD. *Journal of attention disorders*, 7(1).
- Weisz, J. R., Jensen-Doss, A. & Hawley, K. M. (2006). Evidence-based youth psychotherapies versus usual clinical care - A meta-analysis of direct comparisons. *American Psychologist*, 61(7), 671-689.
- Weisz, J. R., Weiss, B., Han, S. S., Granger, D. A. & Morton, T. (1995). Effects of psychotherapy with children and adolescents revisited: A meta-analysis of treatment outcome studies. *Psychological Bulletin*, 117(3), 450-468.
- Wender, P. H. (2002). *Aufmerksamkeits- und Aktivitätsstörungen bei Kindern, Jugendlichen und Erwachsenen: Ein Ratgeber für Betroffene und Helfer*. Stuttgart: Kohlhammer.

- Werner, E. E. (1993). Risk, resilience, and recovery: Perspectives from the Kauai longitudinal study. *Development and Psychopathology*, 5(4).
- West, C. & Zimmerman, D. H. (1987). Doing gender. *Gender & Society*, 1(2), 125-151.
- Westhoff, K. (1989). Übungsabhängigkeit von Leistungen in Konzentrationstests. *Diagnostica*, 35, 122-130.
- Wiener, J. & Mak, M. (2009). Peer victimization in children with attention-deficit/hyperactivity disorder. *Psychology in the Schools*, 46(2), 116-131.
- Wietfeldt, A. (2010). *Untersuchung zur psychometrischen Güte und Validierung des Lehrerfragebogens zum Lern- und Arbeitsverhalten von Schülern mit und ohne ADHS*. Universität Bremen: unveröffentlichte Diplomarbeit.
- Wilder, J. (1931). Das „Ausgangswert-Gesetz“, ein unbeachtetes biologisches Gesetz und seine Bedeutung für Forschung und Praxis. *Zeitschrift für die gesamte Neurologie und Psychiatrie*, 137, 317-338.
- Wilens, T. E., Prince, J. & Biederman, J. (2008). ADHD in adults. In K. McBurnett & L. Pfiffner (Eds.), *Attention deficit hyperactivity disorder: Concepts, controversies, new directions* (pp. 273-290). New York: informa.
- Wilkinson, L. & Task Force on Statistical Inference. (1999). Statistical methods in psychology journals - Guidelines and explanations. *American Psychologist*, 54(8).
- Willcutt, E. (2012). The prevalence of DSM-IV Attention-deficit/hyperactivity disorder: A meta-analytic review. *Neurotherapeutics*, 9(3), 490-499.
- Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V. & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: A meta-analytic review. *Biological Psychiatry*, 57(11), 1336-1346.
- Willcutt, E. G., Nigg, J. T., Pennington, B. F., Solanto, M. V., Rohde, L. A., Tannock, R. et al. (2012). Validity of DSM-IV Attention deficit/hyperactivity disorder symptom dimensions and subtypes. *Journal of Abnormal Psychology*, 121(4), 991-1010.
- Williams, R. H. & Zimmerman, D. W. (1996). Are simple gain scores obsolete? *Applied Psychological Measurement*, 20(1), 59-69.
- Winsler, A. (1998). Parent-child interaction and private speech in boys with ADHD. *Applied Developmental Science*, 2(1), 17-39.
- Wise, E. A. (2004). Methods for analyzing psychotherapy outcomes: A review of clinical significance, reliable change, and recommendations for future directions. *Journal of Personality Assessment*, 82(1), 50-59.
- Woerner, W., Becker, A., Friedrich, C., Klasen, H., Goodman, R. & Rothenberger, A. (2002). Normierung und Evaluation der deutschen Elternversion des Strengths and Difficulties Questionnaire (SDQ): Ergebnisse einer repräsentativen Felderhebung. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie*, 30(2), 105-112.
- Woerner, W., Becker, A. & Rothenberger, A. (2004). Normative data and scale properties of the German parent SDQ. *European Child & Adolescent Psychiatry*, 13, 3-10.

- Wolke, D., Rizzo, P. & Woods, S. (2002). Persistent infant crying and hyperactivity problems in middle childhood. *Pediatrics*, 109(6), 1054-1060.
- Wolraich, M., Brown, L., Brown, R. T., DuPaul, G., Earls, M., Feldman, H. M. et al. (2011). ADHD: Clinical practice guideline for the diagnosis, evaluation, and treatment of attention-deficit/hyperactivity disorder in children and adolescents. *Pediatrics*, 128(5), 1007-1022.
- Wood, D., Bruner, J. S. & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 17(2), 89-100.
- Wu, S. Y. & Gau, S. S. F. (2013). Correlates for academic performance and school functioning among youths with and without persistent attention-deficit/hyperactivity disorder. *Research in Developmental Disabilities*, 34(1), 505-515.
- Yan, X. & Su, X. G. (2009). *Linear Regression Analysis: Theory and Computing*. Singapore: World Scientific Publishing.
- Yazici, B. & Yolacan, S. (2007). A comparison of various tests of normality. *Journal of Statistical Computation and Simulation*, 77(2).
- Yoshimasu, K., Barbaresi, W. J., Colligan, R. C., Voigt, R. G., Killian, J. M., Weaver, A. L. et al. (2012). Childhood ADHD is strongly associated with a broad range of psychiatric disorders during adolescence: a population-based birth cohort study. *Journal of Child Psychology and Psychiatry*, 53(10), 1036-1043.
- Young, S. & Amarasinghe, J. M. (2010). Practitioner review: Non-pharmacological treatments for ADHD: A lifespan approach. *Journal of Child Psychology and Psychiatry*, 51(2), 116-133.
- Zentall, S. S. (1975). Optimal stimulation as theoretical basis of hyperactivity. *American Journal of Orthopsychiatry*, 45(4), 549-563.
- Zentall, S. S. & Zentall, T. R. (1983). Optimal stimulation: A model of disordered activity and performance in normal and deviant children. *Psychological Bulletin*, 94(3), 446-471.
- Ziegler, W. & Hegerl, U. (2002). Der Werther-Effekt. Bedeutung, Mechanismen, Konsequenzen. *Der Nervenarzt*, 73(1), 41-49.
- Zimmerman, D. W. (1997). A geometric interpretation of the validity and reliability of difference scores. *British Journal of Mathematical & Statistical Psychology*, 50, 73-80.
- Zimmermann, P. & Fimm, B. (2008). *Testbatterie zur Aufmerksamkeitsprüfung (TAP) (Version 2.1)*. Herzogenrath: Psytest.
- Zimmermann, P. & Fimm, B. (2009). *Testbatterie zur Aufmerksamkeitsprüfung (TAP) (Version 2.2)*. Herzogenrath: Psytest.
- Zimmermann, P. & Fimm, B. (2010). *Testbatterie zur Aufmerksamkeitsprüfung (TAP) (Version 2.2) (2. Aufl.)*. Herzogenrath: Psytest.
- Zimmermann, P., Gondan, M. & Fimm, B. (2002). *Testbatterie zur Aufmerksamkeitsprüfung für Kinder (KiTAP)*. Herzogenrath: Psytest.

- Zobel, A. & Maier, W. (2004). Endophänotypen - ein neues Konzept zur biologischen Charakterisierung psychischer Störungen. *Der Nervenarzt*, 75(3), 205-214.
- Zoccolotti, P., Matano, A., Deloche, G., Cantagallo, A., Passadori, A., Leclercq, M. et al. (2000). Patterns of attentional impairment following closed head injury: A collaborative European study. *Cortex*, 36(1), 93-107.
- Zweben, A. & Cisler, R. (1996). Composite outcome measures in alcoholism treatment research: Problems and potentialities. *Substance Use & Misuse*, 31(13), 1783-1805.
- Zwingmann, C. & Wirtz, M. (2005). Regression zur Mitte. *Die Rehabilitation*, 44(4), 244-251.

Anhang

- A.1 Erklärungen
- A.2 Vergleich fehlende Lehrereinschätzung
- A.3 Überprüfungen auf Normalverteilung
- A.4 Weitere Ergebnisse der univariaten Regressionsanalysen
- A.5 Weitere Ergebnisse der multivariaten Regressionsanalyse

A.1 Erklärungen

Hiermit versichere ich, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Hilfsmittel und Quellen benutzt habe. Der Inhalt wurde nicht schon für eine Diplom- oder ähnliche Prüfungsarbeit verwendet.

Hiermit versichere ich, dass die Leitlinien guter wissenschaftlicher Praxis an der Carl von Ossietzky Universität Oldenburg befolgt worden sind

Oldenburg, 08.07.2013

A.2 Vergleich fehlende Lehrereinschätzung

Vergleich der Problemmaße und des Gesamterfolgsmaß der Teilnehmer mit vorliegender Lehrereinschätzung und mit fehlender Lehrereinschätzung zum Prämesszeitpunkt

	Lehrereinschätzung vorhanden (ja = 1, nein = 0)	N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes		
Symptomatik	0	10	-,2425	,65064	,20575		
	1	57	,0425	,78372	,10381		
Reaktionszeit	0	10	-,0953	,51757	,16367		
	1	57	,0167	,87407	,11577		
Fehler	0	10	,0862	,68403	,21631		
	1	57	-,0151	,70391	,09324		
Gesamtproblematik	0	10	-,071130	,3768227	,1191618		
	1	57	,011035	,4761874	,0630726		
			Levene-Test der Varianzgleichheit				
			F	Signifikanz	T	df	Sig. (2-seitig)
Symptomatik	Varianzen gleich		2,639	,109	-1,085	65	,282
	Varianzen nicht gleich				-1,237	14,019	,236
Reaktionszeit	Varianzen gleich		1,425	,237	-,392	65	,697
	Varianzen nicht gleich				-,559	19,476	,583
Fehler	Varianzen gleich		,345	,559	,421	65	,675
	Varianzen nicht gleich				,430	12,585	,674
Gesamtproblematik	Varianzen gleich		,423	,518	-,517	65	,607
	Varianzen nicht gleich				-,609	14,566	,552

Vergleich der Problemaße und des Gesamterfolgsmaß der Teilnehmer mit vorliegender Lehrereinschätzung und mit fehlender Lehrereinschätzung zum Postmesszeitpunkt

	Lehrereinschätzung vorhanden (ja = 1, nein = 0)	N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes		
Symptomatik	0	12	-,5112	,82708	,23876		
	1	54	-,5895	,74051	,10077		
Reaktionszeit	0	11	-,2773	,49874	,15037		
	1	55	-,1880	,67177	,09058		
Fehler	0	11	-,2829	,59173	,17841		
	1	55	-,1844	,58271	,07857		
Gesamtproblematik	0	12	-,302156	,4801201	,1385987		
	1	55	-,351908	,4136316	,0557741		
			Levene-Test der Varianzgleichheit				
			F	Signifikanz	T	df	Sig. (2-seitig)
Symptomatik	Varianzen gleich		,050	,823	,325	64	,746
	Varianzen nicht gleich				,302	15,168	,766
Reaktionszeit	Varianzen gleich		1,460	,231	-,418	64	,678
	Varianzen nicht gleich				-,509	18,132	,617
Fehler	Varianzen gleich		,250	,619	-,511	64	,611
	Varianzen nicht gleich				-,505	14,157	,621
Gesamtproblematik	Varianzen gleich		,048	,828	,367	65	,715
	Varianzen nicht gleich				,333	14,772	,744

A.3 Überprüfungen auf Normalverteilung

Überprüfung auf Normalverteilung der Problemfaktoren und des Gesamtproblemmaßes für die gesamte Stichprobe

Prämesszeitpunkt	Shapiro-Wilk		
	Statistik	df	Signifikanz
Symptomatik	,967	67	,070
Lehrereinschätzung	,970	57	,174
Reaktionszeit	,862	67	,000
Fehler	,917	67	,000
Gesamtproblematik	,973	67	,156
Postmesszeitpunkt	Shapiro-Wilk		
	Statistik	df	Signifikanz
Symptomatik	,982	66	,472
Lehrereinschätzung	,969	57	,148
Reaktionszeit	,909	66	,000
Fehler	,913	66	,000
Gesamtproblematik	,986	67	,629
Differenzwerte	Shapiro-Wilk		
	Statistik	df	Signifikanz
Symptomatik	,984	66	,535
Lehrereinschätzung	,936	52	,008
Reaktionszeit	,912	66	,000
Fehler	,905	66	,000
Gesamtproblematik	,976	67	,230

Überprüfung auf Normalverteilung der Problemfaktoren und des Gesamtproblemmaßes von Mädchen und Jungen vor Trainingsbeginn

Faktor	Geschlecht	Shapiro-Wilk		
		Statistik	df	Signifikanz
Symptomatik	männlich	,963	54	,099
	weiblich	,965	13	,828
Lehrereinschätzung	männlich	,968	46	,238
	weiblich	,910	11	,246
Reaktionszeit	männlich	,874	54	,000
	weiblich	,805	13	,008
Fehler	männlich	,958	54	,054
	weiblich	,669	13	,000
Gesamtproblematik	männlich	,986	54	,795
	weiblich	,880	13	,071

Überprüfung auf Normalverteilung der Differenzwerte von Mädchen und Jungen

Faktor	Shapiro-Wilk		
	Statistik	df	Signifikanz
Symptomatik	,984	66	,535
Lehrereinschätzung	,936	52	,008
Reaktionszeit	,912	66	,000
Fehler	,905	66	,000
Gesamtproblematik	,976	67	,230

A.4 Weitere Ergebnisse der univariaten Regressionsanalysen

Ausgeschlossene Variablen in der schrittweisen Regressionsanalyse im Problemmaß Symptomatik

AV: Symptomatik post Modell		β	t	p	Partielle Korrelation	Kollinearitätsstatistik Toleranz
1	Geschlecht	,105	1,072	,289	,147	,960
	Medikamente	,016	,168	,867	,023	1,000
	Alter	-,075	-,783	,437	-,108	1,000
	Anzahl der Sitzungen	,105	1,095	,279	,150	,989
	Dauer des Trainings	-,030	-,298	,767	-,041	,917
	Hyperaktivität-Impulsivität	,165	1,205	,234	,165	,483
	Internal. Komorbidität.	-,102	-1,016	,314	-,140	,911
	External. Komorbidität	,146	1,289	,203	,176	,705

AV: Symptomatik Differenz Modell		β	t	p	Partielle Korrelation	Kollinearitätsstatistik Toleranz
1	Geschlecht	,138	1,080	,285	,148	,960
	Medikamente	,018	,141	,888	,020	1,000
	Alter	-,095	-,757	,452	-,104	1,000
	Anzahl der Sitzungen	,140	1,111	,272	,152	,989
	Dauer des Trainings	-,039	-,297	,768	-,041	,917
	Hyperaktivität-Impulsivität	,214	1,193	,238	,163	,483
	Internal. Komorbidität.	-,129	-,986	,329	-,135	,911
	External. Komorbidität	,178	1,196	,237	,164	,705

Ausgeschlossene Variablen in der schrittweisen Regressionsanalyse im Problemmaß Lehrereinschätzung

AV: Lehrereinschätzung post Modell		β	t	p	Partielle Korrelation	Kollinearitätsstatistik
						Toleranz
1	Geschlecht	,027	,218	,828	,031	,998
	Medikamente	-,243	-2,029	,048	-,278	,950
	Alter	-,067	-,538	,593	-,077	,958
	Anzahl der Sitzungen	,161	1,344	,185	,189	,994
	Dauer des Trainings	-,026	-,199	,843	-,028	,853
	Hyperaktivität-Impulsivität	,082	,612	,544	,087	,826
	Internal. Komorbidität.	-,178	-1,400	,168	-,196	,879
	External. Komorbidität	-,057	-,443	,659	-,063	,881
2	Geschlecht	-,005	-,043	,966	-,006	,981
	Alter	-,054	-,446	,657	-,064	,955
	Anzahl der Sitzungen	,111	,923	,361	,132	,939
	Dauer des Trainings	,008	,066	,948	,009	,837
	Hyperaktivität-Impulsivität	,089	,690	,494	,099	,826
	Internal. Komorbidität.	-,163	-1,316	,194	-,187	,876
	External. Komorbidität	,016	,125	,901	,018	,808

AV: Lehrereinschätzung Differenz Modell		β	t	p	Partielle Korrelation	Kollinearitätsstatistik
						Toleranz
1	Geschlecht	-,033	-,265	,792	-,039	,998
	Medikamente	-,217	-1,770	,083	-,250	,950
	Alter	-,086	-,688	,495	-,100	,958
	Anzahl der Sitzungen	,175	1,447	,155	,206	,994
	Dauer des Trainings	-,061	-,458	,649	-,067	,853
	Hyperaktivität-Impulsivität	,011	,078	,938	,011	,826
	Internal. Komorbidität.	-,193	-1,505	,139	-,214	,879
	External. Komorbidität	-,124	-,952	,346	-,137	,881

Ausgeschlossene Variablen in der schrittweisen Regressionsanalyse im Problemmaß Reaktionszeit

AV: Reaktionszeit post Modell		β	t	p	Partielle Korrelation	Kollinearitätsstatistik
						Toleranz
1	Geschlecht	,106	1,579	,120	,214	,995
	Medikamente	-,147	-2,231	,030	-,296	,993
	Alter	,017	,243	,809	,034	,987
	Anzahl der Sitzungen	-,046	-,670	,506	-,092	,999
	Dauer des Trainings	-,134	-2,011	,050	-,269	,983
	Hyperaktivität-Impulsivität	,050	,732	,467	,101	,993
	Internal. Komorbidität.	,029	,424	,673	,059	,992
	External. Komorbidität	-,038	-,559	,578	-,077	,994
2	Geschlecht	,088	1,340	,186	,184	,977
	Alter	,030	,451	,654	,063	,979
	Anzahl der Sitzungen	-,081	-1,209	,232	-,167	,953
	Dauer des Trainings	-,127	-1,978	,053	-,267	,981
	Hyperaktivität-Impulsivität	,041	,625	,535	,087	,989
	Internal. Komorbidität.	,027	,409	,684	,057	,992
	External. Komorbidität	-,012	-,183	,856	-,026	,962

Ausgeschlossene Variablen in der schrittweisen Regressionsanalyse im Problemmaß Reaktionszeit (Fortsetzung)

AV: Reaktionszeit Differenz		β	t	p	Partielle Korrelation	Kollinearitätsstatistik
Modell						Toleranz
1	Geschlecht	,146	1,415	,163	,193	,995
	Medikamente	-,218	-2,166	,035	-,288	,993
	Alter	,034	,323	,748	,045	,987
	Anzahl der Sitzungen	-,079	-,761	,450	-,105	,999
	Dauer des Trainings	-,211	-2,071	,043	-,276	,983
	Hyperaktivität-Impulsivität	,070	,672	,505	,093	,993
	Internal. Komorbidität.	,029	,272	,787	,038	,992
	External. Komorbidität	-,063	-,601	,550	-,083	,994
2	Geschlecht	,119	1,176	,245	,163	,977
	Alter	,054	,528	,600	,074	,979
	Anzahl der Sitzungen	-,132	-1,292	,202	-,178	,953
	Dauer des Trainings	-,201	-2,038	,047	-,274	,981
	Hyperaktivität-Impulsivität	,057	,565	,575	,079	,989
	Internal. Komorbidität.	,026	,251	,802	,035	,992
	External. Komorbidität	-,025	-,237	,813	-,033	,962
3	Geschlecht	,138	1,402	,167	,195	,970
	Alter	,077	,771	,444	,108	,967
	Anzahl der Sitzungen	-,093	-,912	,366	-,128	,911
	Hyperaktivität-Impulsivität	,101	1,006	,319	,141	,951
	Internal. Komorbidität.	,104	1,000	,322	,140	,881
	External. Komorbidität	,003	,026	,979	,004	,945

Ausgeschlossene Variablen in der schrittweisen Regressionsanalyse im Problemmaß Fehler

AV: Fehler post Modell		β	t	p	Partielle Korrelation	Kollinearitätsstatistik
						Toleranz
1	Geschlecht	-,117	-1,238	,221	-,169	,974
	Medikamente	,130	1,408	,165	,192	1,000
	Alter	-,076	-,814	,420	-,112	,992
	Anzahl der Sitzungen	,045	,472	,639	,065	,985
	Dauer des Trainings	-,022	-,225	,823	-,031	,943
	Hyperaktivität-Impulsivität	-,083	-,877	,384	-,121	,981
	Internal. Komorbidität.	,066	,698	,488	,096	,997
	External. Komorbidität	,026	,266	,791	,037	,961

AV: Fehler Differenz Modell		β	t	p	Partielle Korrelation	Kollinearitätsstatistik
						Toleranz
1	Geschlecht	-,030	-,246	,807	-,034	,974
	Medikamente	,106	,895	,375	,123	1,000
	Alter	-,162	-1,372	,176	-,187	,992
	Anzahl der Sitzungen	,169	1,428	,159	,194	,985
	Dauer des Trainings	,001	,010	,992	,001	,943
	Hyperaktivität-Impulsivität	-,057	-,476	,636	-,066	,981
	Internal. Komorbidität.	,190	1,628	,110	,220	,997
	External. Komorbidität	,081	,671	,505	,093	,961

A.5 Weitere Ergebnisse der multivariaten Regressionsanalyse

Zusammenhang der Symptomatik und Lehrereinschätzung zum Prämesszeitpunkt mit den Werten zum Postmesszeitpunkt in der multivariaten Auswertung.

Quelle	Abhängige Variable	Quadrat- summe vom Typ III	df	Mittel der Quadrate	F	Sig.	Partielles Eta-Quadrat
Symptomatik	Symptomatik	7,239	1	7,239	24,361	,000	,384
	Lehrereinschätzung	,152	1	,152	,316	,577	,008
	Reaktionszeit	,238	1	,238	,481	,492	,012
	Fehler	,340	1	,340	1,024	,318	,026
Lehrerein- schätzung	Symptomatik	,256	1	,256	,861	,359	,022
	Lehrereinschätzung	7,412	1	7,412	15,391	,000	,283
	Reaktionszeit	,001	1	,001	,001	,973	,000
	Fehler	,797	1	,797	2,401	,129	,058