



universität
wien

Diplomarbeit

Titel der Arbeit

Die Entwicklung und Erprobung eines spielbasierten Itempools zur Erfassung der Entwicklung Zweijähriger

unter besonderer Berücksichtigung der kognitiven Entwicklung, der
visuellen Wahrnehmung, des Gedächtnisses und der Arbeitshaltungen

Verfasserin

Elisabeth Sapper

Angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag. rer. nat.)

Wien, im August 2011

Studienkennzahl: 298

Studienrichtung: Psychologie

Betreuerin: Ass.-Prof. Dr. Pia Deimann

Inhaltsverzeichnis

Einleitung	9
I. THEORETISCHER TEIL	11
1 Entwicklungsdiagnostik	13
1.1 Begriffsdefinition	13
1.2 Anforderungen an die Entwicklungsdiagnostik und an die Diagnostik allgemein – Gütekriterien.....	15
1.2.1 Objektivität	16
1.2.2 Reliabilität.....	16
1.2.3 Validität.....	17
1.2.4 Normierung	18
1.2.5 Vergleichbarkeit.....	18
1.2.6 Ökonomie	18
1.3 Entwicklungsdiagnostik im Kleinkindalter	19
1.3.1 Bestehenden Verfahren für die Entwicklungsdiagnostik von Kindern unter 3 Jahren	20
1.3.1.1 Münchener funktionelle Entwicklungsdiagnostik: Zweites und drittes Lebensjahr (MFED 2-3).....	21
1.3.1.2 Bayley Scales of Infant Development, 2nd Ed. (BSID-II)	22
1.3.1.3 Bayley Scales of Infant and Toddler Development, 3rd Ed. (Bayley-III).....	24
1.3.1.4 Entwicklungstest 6 Monate – 6 Jahre (ET 6-6)	25
1.4 Zusammenfassung	28

2 Spielbasierte Diagnostik	28
2.1 Begriffsdefinition.....	29
2.2 Anfänge der spielbasierten Diagnostik.....	30
2.3 Verschiedene Ansätze der spielbasierten Diagnostik.....	31
2.3.1 Ein Ansatz spielbasierter Entwicklungsdiagnostik in der DDR.....	31
2.3.2 Transdisciplinary Play-Based Assessment (TPBA).....	32
2.3.3 Play Assessment Scale (PAS)	33
2.3.4 The Developmental Play Assessment (DPA) Instrument.....	34
2.3.5 The Play Observation Kit (POKIT)	35
2.4 Spielbasierte und klassische Diagnostik bei Kindern von 2;0 bis 2;11 Jahren im Vergleich.....	35
2.4.1 Klassische Diagnostik	36
2.4.2 Vorteile der klassischen Diagnostik	38
2.4.3 Nachteile der klassischen Diagnostik.....	40
2.4.4 Vorteile der spielbasierten Diagnostik.....	43
2.4.5 Nachteile der spielbasierten Diagnostik.....	44
2.4.6 Fazit	45
2.5 Zusammenfassung.....	46

3 In der vorliegenden Studie behandelte Entwicklungsbereiche mit Fokus auf das Alter von 2;0 bis 2;11 Jahren.....	47
3.1 Kognitive Entwicklung.....	48
3.1.1 Theory of Mind	48
3.1.1.1 Erfassung der theory of mind	49
3.1.1.1.1 <i>False-belief – Aufgaben zur falschen Überzeugung – Platzwechsel- Aufgaben</i>	<i>50</i>
3.1.1.1.2 <i>Appearance-reality – Aufgaben zur Unterscheidung von Erscheinungsbild und Wirklichkeit.....</i>	<i>51</i>
3.1.1.2 Entwicklung der theory of mind	52
3.1.1.3 Vorläufer der theory of mind.....	55
3.1.1.3.1 <i>Erkennen des eigenen Spiegelbildes</i>	<i>55</i>
3.1.1.3.2 <i>Als-ob-Spiel</i>	<i>56</i>
3.1.1.3.3 <i>Imitation.....</i>	<i>58</i>
3.1.1.3.4 <i>Intention.....</i>	<i>60</i>
3.1.1.4 Einflüsse auf die theory of mind.....	62
3.1.1.4.1 <i>Exekutivfunktionen.....</i>	<i>62</i>
3.1.1.4.2 <i>Sprachfähigkeiten</i>	<i>64</i>
3.1.2 Numerisches Wissen.....	65
3.1.2.1 Zählen	65
3.1.2.2 Erfassung von Mengen	67
3.1.3 Schlussfolgerndes Denken	69
3.1.3.1 Induktives Denken	69
3.1.3.2 Deduktives Denken.....	70
3.1.3.3 Zusammenhänge / Unterschiede zwischen induktivem und deduktivem Denken	72
3.1.3.4 Analoges Denken.....	73
3.2 Visuelle Wahrnehmung	74
3.2.1 Objekterkennung	74
3.2.1.1 <i>Objektwahrnehmung</i>	<i>74</i>

3.2.1.2 Mustererkennung	75
3.2.2 Farbwahrnehmung / -differenzierung	76
3.2.3 Formwahrnehmung / -differenzierung	77
3.2.4 Größenwahrnehmung / -differenzierung.....	78
3.3 Gedächtnis	78
3.3.1 Ultrakurzzeitgedächtnis	78
3.3.2 Kurzzeit- bzw. Arbeitsgedächtnis.....	79
3.3.3 Langzeitgedächtnis	80
3.4 Arbeitshaltungen.....	83
3.4.1 Aufmerksamkeit	83
3.4.2 Leistungsmotivation	85
3.5 Zusammenfassung.....	86
II. EMPIRISCHER TEIL.....	89
4 Zielsetzung und Fragestellungen	91
4.1 Zielsetzungen	91
4.2 Fragestellungen.....	92
5 Methode	92
5.1 Beschreibung des Itempools	92
5.1.1 Kognitive Entwicklung	94
5.1.1.1 Theory of mind	94
5.1.1.2 Numerisches Wissen	95
5.1.1.3 Schlussfolgerndes Denken	95

5.1.2 Visuelle Wahrnehmung	95
5.1.3 Gedächtnis	97
5.1.4 Arbeitshaltungen	97
5.2 Stichprobe	98
5.3 Durchführung der Untersuchung	98
6 Ergebnisse	100
6.1 Ergebnisse zur kognitiven Entwicklung	101
6.1.1 Ergebnisse zur theory of mind.....	101
6.1.2 Ergebnisse zum numerischen Wissen	104
6.1.3 Ergebnisse zum schlussfolgernden Denken	107
6.2 Ergebnisse zur visuellen Wahrnehmung.....	109
6.3 Ergebnisse zum Gedächtnis	113
6.4 Ergebnisse zu den Arbeitshaltungen.....	115
6.5 Problem der fehlenden Werte	117
7 Diskussion	121
7.1 Aufgabenbereiche.....	122
7.1.1 Kognitive Entwicklung	122
7.1.1.1 Theory of mind.....	122
7.1.1.2 Numerisches Wissen	123
7.1.1.3 Schlussfolgerndes Denken	124
7.1.2 Visuelle Wahrnehmung	125
7.1.3 Gedächtnis	127

7.1.4 Arbeitshaltungen	128
7.2 Testsituation	128
7.3 Zusammenfassung.....	130
III. LITERATURVERZEICHNIS	133
IV. ANHANG	145

Einleitung

Ziel dieser Arbeit ist es, einen Itempool für die Erhebung der allgemeinen Entwicklung von Kindern im Alter von 2;0 bis 2;11 Jahren zu konstruieren. Die Items sollen so gestaltet sein, dass sie dem Kind eine angenehme Testsituation bieten, die einer lockeren und freien Spielatmosphäre entspricht, in der flexibel auf das Kind eingegangen werden kann. Trotzdem soll die Erhebung verschiedener Items möglichst standardisiert erfolgen und so gestaltet sein, dass die Ergebnisse des Kindes schließlich mit den Ergebnissen Gleichaltriger verglichen werden können.

Da dieser gesamte Themenkomplex sehr umfangreich ist, liegt das Augenmerk dieser Arbeit auf den Entwicklungsbereichen kognitive Entwicklung, visuelle Wahrnehmung, Gedächtnis und Arbeitshaltungen. Mit den Bereichen Sprache, Motorik und sozial-emotionale Entwicklung beschäftigt sich Kuchler (in Druck).

Im Theorieteil wird zuerst auf die Entwicklungsdiagnostik im Allgemeinen eingegangen, danach speziell auf die Entwicklungsdiagnostik im Kleinkindalter. Dabei werden gängige Verfahren vorgestellt und kritisch bewertet. Anschließend wird ein Überblick über die spielbasierte Diagnostik gegeben und es werden ebenfalls einige Verfahren dazu vorgestellt. Schließlich werden diejenigen Entwicklungsbereiche behandelt, die für dieses Verfahren von Interesse sind.

Im nachfolgenden empirischen Teil wird der entwickelte Itempool vorgestellt. Danach wird die Untersuchungssituation beschrieben und auf die Ergebnisse, die aus der Erprobung des Verfahrens an 20 Kindern gewonnen wurden, eingegangen. Schließlich werden diese Ergebnisse diskutiert und kritisch betrachtet.

I. THEORETISCHER TEIL

1 Entwicklungsdiagnostik

1.1 Begriffsdefinition

Der Begriff „Entwicklungsdiagnostik“ beschäftigt sich nach Deimann & Kastner-Koller (2007) mit sowohl der quantitativen als auch der qualitativen Erfassung von entwicklungsbedingten Kompetenzen. Ihr Ziel ist die Beschreibung des momentanen Kompetenzniveaus, die Prognose der zukünftigen Entwicklung, sowie die Planung und Evaluation von Interventionen. Unterteilt wird die Gruppe der Entwicklungstests in die Untergruppe der allgemeinen Entwicklungstests und die der speziellen Entwicklungstests. *Allgemeine Entwicklungstests* erfassen die gesamte Bandbreite des Entwicklungsstandes eines Kindes durch die Vorgabe einzelner Subtests, *spezielle Entwicklungstests* hingegen erfassen nur einen bestimmten Ausschnitt der kindlichen Entwicklung. Sie sind lediglich auf einzelne Funktionsbereiche fokussiert. Weiters gibt es auch sogenannte *Entwicklungsscreenings*, welche den Entwicklungsstand eines Kindes sehr ökonomisch, also mit möglichst wenig Aufwand und in möglichst kurzer Zeit (meist 10 bis 20 Minuten) bewerten. Ein Screening eignet sich jedoch lediglich als Vorab-Untersuchung, um eine Klassifikation der Entwicklung in „nicht auffällig“ und „auffällig“ zu treffen. Dies wird anhand von Cut-Off-Werten bestimmt, welche im Vorhinein festgelegt werden. Befinden sich die Werte eines Kindes über dem festgesetzten Cut-Off-Wert, gilt seine Entwicklung als unauffällig, liegt die Leistung darunter, gilt sie als auffällig. In diesem Fall sollte das Kind einer weiterführenden, genaueren Diagnostik unterzogen werden (Petermann & Macha, 2005). Zusätzlich werden Informationen auch sehr oft aus der *Befragung der Bezugspersonen* und aus der *Verhaltensbeobachtung* gewonnen.

Die Annahme, die hinter Entwicklungstestverfahren steht, ist die, dass hinter dem beobachtbaren Verhalten, das ein Kind zeigt, unterschiedliche Dimensionen oder Funktionsbereiche liegen, die durch die Erhebung des Verhaltens indirekt erfasst werden können (Reuner & Pietz, 2006). Diejenigen Bereiche, die so gut wie in jedem Entwicklungstest erhoben

werden, sind die Grobmotorik, die Feinmotorik, die kognitive Entwicklung, die Sprachentwicklung, sowie die sozial-emotionale Entwicklung. Eventuell werden von einzelnen Verfahren ergänzend auch noch andere Bereiche erfasst (Petermann & Macha, 2006).

Das *traditionelle Verständnis von Entwicklung* geht davon aus, dass eine normale Entwicklung in engen Bahnen verläuft, die miteinander zusammenhängen. Dabei bedingen Beeinträchtigungen in einem Bereich auch Beeinträchtigungen in anderen Entwicklungsbereichen („Reifungskonzept“). Durch Beseitigung der Ursache dieser Beeinträchtigung kann die Entwicklung aber wieder korrigiert werden. Hierbei werden häufig Stufenleiterverfahren angewendet (Petermann & Macha, 2006).

Stufenleitern, wie etwa die „Münchener funktionelle Entwicklungsdiagnostik: Zweites und drittes Lebensjahr“ (MFED 2-3; Hellbrügge, 1994), sind Verfahren, in welchen die Aufgaben, die zu einem Entwicklungsbereich gehören, in der Reihenfolge ihrer Schwierigkeit angeordnet werden. Aus dem Alter des Kindes wird abgeleitet, mit welcher Einstiegsaufgabe begonnen wird. Diese Verfahren gehen von einer starren Abfolge der Entwicklung aus, was empirisch häufig nicht zu belegen ist. Ebenso ergeben sich aus dieser Annahme Schwierigkeiten in der Testkonstruktion, da sich Fertigkeiten mit unterschiedlichen Geschwindigkeiten entwickeln, weswegen eine Aufgabe nicht für jedes Kind und nicht in jedem Lebensalter leichter oder schwieriger als eine andere ist. Daher kann von der Tatsache, dass eine Fertigkeit mit einem gewissen Alter noch nicht erlernt wurde, keinesfalls schon auf eine Retardierung geschlossen werden, denn diese Variabilität könnte durchaus noch in einem unauffälligen Bereich liegen. Aufgrund dessen werden die Kriterien für einen Testabbruch eher weit gefasst, was aber wiederum der Genauigkeit Abbruch tut. Das Ergebnis von Stufenleitern bildet meist ein Entwicklungsalter (die Leistung im Vergleich zum Lebensalter) oder ein Entwicklungsquotient (die Leistung im Vergleich zur durchschnittlichen Leistung Gleichaltriger). Hier stellt sich allerdings die Frage, wie aussagekräftig ein solcher Gesamtwert ist, da Beeinträchtigungen oder Verzögerungen oft nur auf einen oder wenige Bereiche beschränkt sind (Petermann & Macha, 2003).

Nach dem *modernen Entwicklungsverständnis* passiert die normale Entwicklung hingegen innerhalb eines genetisch vorgegebenen Rahmens, wobei die Bedingungen der individuellen Lebensumwelt aber eine große Rolle für die dann tatsächlich stattfindenden Entwicklungsverläufe spielen. Die verschiedenen Bereiche der Entwicklung berühren sich zunächst nicht, können aber mit der Zeit komplexe Muster von Entwicklungsabweichungen ausbilden („adaptive Theorien“). Gegen abweichende Entwicklungsverläufe wird durch eine ganzheitliche Perspektive vorgegangen. In diesem Entwicklungsverständnis werden meist Inventare oder Testbatterien durchgeführt (Petermann & Macha, 2006).

Testbatterien bestehen mehr oder weniger aus einer Aneinanderreihung von in sich homogenen Einzeltests, die je einen Entwicklungsbereich abbilden. Im Vergleich zu Stufenleitern oder Inventaren wird jeder Bereich mit einer größeren Anzahl an Aufgaben erhoben und daher genauer abgebildet. Ein Beispiel hierfür wäre der „Wiener Entwicklungstest“ (WET; Kastner-Koller & Deimann, 2002).

Inventare, wie zum Beispiel der „Entwicklungstest 6 Monate bis 6 Jahre“ (ET 6-6; Petermann, Stein & Macha, 2004), wollen möglichst vielfältige Aspekte der Entwicklung erfassen. Die Überprüfung der Gütekriterien ist dabei nicht immer einfach, aber Expertenurteile schreiben ihnen oft eine hohe Gültigkeit zu (Petermann & Macha, 2003).

1.2 Anforderungen an die Entwicklungsdiagnostik und an die Diagnostik allgemein – Gütekriterien

Natürlich werden an die Diagnostik der Entwicklung aber auch gewisse Anforderungen gestellt. Lienert und Ratz (1998) definierten sogenannte Haupt- und Nebengütekriterien, die die Qualität diagnostischer Untersuchungen gewährleisten sollen. Im Folgenden sollen diese Gütekriterien kurz vorgestellt werden. Sowohl die drei Hauptgütekriterien „Objektivität“, „Reliabilität“ und „Validität“, als auch die Nebengütekriterien „Normierung“, „Vergleichbarkeit“, „Ökonomie“ und „Nützlichkeit“ werden dabei behandelt (vgl. Kubinger, 2009; Lienert & Ratz, 1998).

1.2.1 Objektivität

Die Objektivität eines Verfahrens bezeichnet die Unabhängigkeit desselbigen vom Testleiter, der die Untersuchung durchführt, sodass unterschiedliche Untersucher zu denselben Testergebnissen gelangen können. Differenzieren lässt sich hierbei zwischen der Durchführungs-, der Auswertungs-, sowie der Interpretationsobjektivität.

Die *Durchführungsobjektivität* beschreibt die Unabhängigkeit der Ergebnisse von Variationen im Verhalten des Testleiters und wird durch eine Standardisierung der Testvorgabe möglichst hoch gehalten.

Auswertungsobjektivität, also inwieweit verschiedene Auswerter zu denselben Ergebnissen kommen, wird durch Standardisierung in der Auswertung, welche meist im Testmanual genau beschrieben ist, ermöglicht.

Schließlich gibt die *Interpretationsobjektivität* an, inwieweit verschiedene Personen aufgrund derselben Ergebnisse zu denselben Schlussfolgerungen gelangen. Dies wird etwa durch die Umrechnung der Ergebnisse in standardisierte und infolge dessen eindeutig interpretierbare Werte erreicht (vgl. Kubinger, 2009; Lienert & Raatz, 1998).

1.2.2 Reliabilität

Die Reliabilität eines Verfahrens gibt an, wie genau dieses ein bestimmtes Merkmal misst. Dabei ist es irrelevant, ob es das Ziel des Verfahrens ist, exakt dieses Merkmal zu messen. Bei der Reliabilität wird unterschieden zwischen Paralleltest-Reliabilität, Retest-Reliabilität und innerer Konsistenz.

Paralleltest-Reliabilität wird erreicht, wenn zwei Verfahren, die dieselbe Eigenschaft erfassen, tatsächlich dasselbe messen.

Die *Retest-Reliabilität* gibt an, inwieweit ein einmal angewendetes Verfahren bei einer erneuten Anwendung nach einer längeren Zeitspanne wieder zu demselben Ergebnis gelangt.

Zu guter Letzt bestimmt die *innere Konsistenz*, inwieweit einzelne Teile eines Verfahrens dasselbe messen. Eine Möglichkeit, dies zu berechnen, bietet die „Split-half“-Methode, bei welcher das Testverfahren halbiert und die beiden Teile miteinander korreliert werden (vgl. Kubinger, 2009; Lienert & Raatz, 1998).

1.2.3 Validität

Die Validität eines Verfahrens gibt an, inwieweit das Verfahren wirklich dasjenige Merkmal misst, das es zu messen vorgibt. Unterschieden wird hierbei zwischen inhaltlicher Validität, Konstruktvalidität und Kriteriumsvalidität.

Die *inhaltliche Validität* sagt aus, inwieweit die Testaufgaben auch tatsächlich die Gesamtheit des Merkmals abbilden, das sie messen sollen. Sie ist am höchsten, wenn das Testverfahren selbst die bestmögliche Operationalisierung des Merkmals darstellt, das es zu messen beansprucht. Meist wird sie durch ein Experten-Rating bestimmt.

Bei der *Konstruktvalidität* handelt es sich um den Anspruch an ein Testverfahren, das zugrundeliegende theoretische Konstrukt, das es messen soll, auch wirklich zu erfassen. Entweder wird hierbei der Anspruch gestellt, dass die Messdaten zweier Verfahren, die dasselbe Konstrukt messen, hoch miteinander korrelieren (sog. *Konvergenzvalidität*), oder zwei Verfahren, die unterschiedliche Konstrukte abbilden, sollen möglichst gering miteinander korrelieren (sog. *Diskriminanzvalidität*).

Zur Angabe der *Kriteriumsvalidität* wird ein Außenkriterium, etwa ein Testverfahren, das dasselbe Konstrukt erfassen soll, mit den Testaufgaben korreliert. Entweder kann die Erfassung dieses Außenkriteriums gleichzeitig mit der Testvorgabe erfolgen, dann spricht man von *Übereinstimmungsvalidität*, oder die Erfassung des Außenkriteriums findet später statt, die Messdaten sagen also das Außenkriterium zu einem gewissen Grad vorher, dann handelt es sich um *Vorhersagevalidität* (vgl. Kubinger, 2009; Lienert & Raatz, 1998).

1.2.4 Normierung

Das Gütekriterium der Normierung ist durch folgende Bedingungen gekennzeichnet: Die Normen, also die Daten der Stichprobe, die als Vergleich für die Testergebnisse einer Person dient, dürfen nicht veraltet sein, am besten nicht älter als 10 Jahre. Denn zumindest für Intelligenzleistungen konnte herausgefunden werden, dass die durchschnittliche Leistung innerhalb von 10 Jahren etwa um 5 IQ-Punkte zunimmt (Flynn, 1987). Daher würden veraltete Normen die Leistungsfähigkeit einer Testperson überschätzen. Aber auch Effekte, die in die andere Richtung gehen, können auftreten, wie es etwa in den Bereichen Motorik oder Sprache zu vermuten ist (Petermann & Macha, 2005). Weiters muss eine Definition der Population, für welche die Normen gelten, vorhanden sein, und die Stichprobe muss für eben diese Gesamtpopulation, in der das Verfahren angewandt wird, repräsentativ sein (vgl. Kubinger, 2009; Lienert & Raatz, 1998).

1.2.5 Vergleichbarkeit

Existieren zu einem Verfahren Parallelförmige oder validitätsähnliche andere Verfahren und ist somit eine Reliabilitäts- oder Validitätskontrolle möglich, ist das Gütekriterium der Vergleichbarkeit erfüllt (vgl. Kubinger, 2009; Lienert & Raatz, 1998).

1.2.6 Ökonomie

Ökonomie ist dann gegeben, wenn das Verfahren im Verhältnis zu seinem diagnostischen Nutzen möglichst wenig Zeit, Geld, oder Materialien verbraucht. Weiters gilt ein Verfahren als ökonomisch, wenn es als Gruppentest durchführbar und einfach und schnell auswertbar ist (vgl. Kubinger, 2009; Lienert & Raatz, 1998).

1.2.7 Nützlichkeit

Wenn es schließlich auch tatsächlich praktisch relevant ist, das Merkmal, das ein bestimmtes Verfahren misst, zu erfassen und bei diesem Verfahren der Nutzen gegenüber dem Schaden überwiegt, dann ist das Gütekriterium der Nützlichkeit erfüllt (vgl. Kubinger, 2009; Lienert & Raatz, 1998).

1.3 Entwicklungsdiagnostik im Kleinkindalter

Die Untersuchung von Kleinkindern nimmt in der Diagnostik eine spezielle Rolle ein, da wohl Personen keines anderen Altersbereiches während der Untersuchung so viel Zuwendung benötigen und so viel Konzentration vom Diagnostiker abverlangen. Erweist sich die bloße Beobachtung eines Kleinkindes intuitiv noch als halbwegs einfach, stellt die Vorgabe eines Testverfahrens an diesem wohl schon ein schwierigeres Unterfangen dar. Vor allem die Anwendung traditioneller diagnostischer Verfahren an Kleinkindern stellt so manchen Diagnostiker / manche Diagnostikerin wahrscheinlich vor eine Herausforderung. Denn gerade mit dieser Art von Verfahren scheint es schwierig zu sein, auf die so unterschiedlichen Bedürfnisse von Kleinkindern eingehen zu können. Kinder, die schon zumindest den Kindergarten besuchen, haben bereits Erfahrung mit fremdgestellten Aufgaben gemacht. Gerade jüngere Kinder sind es jedoch meist gewohnt, ihre Spielinhalte selbst zu wählen und den Ablauf eines Spieles selbst gestalten zu können. Folglich wird es einiger Raffinesse und Erfahrung im Umgang mit Kleinkindern bedürfen, um zum Beispiel ein zweijähriges Kind dazu zu bringen, „das Spiel“ des Testleiters / der Testleiterin zu Ende zu spielen, wenn es das absolut nicht mehr möchte. Dies wird sich noch schwieriger gestalten, wenn es sich bei dem „Spiel“ um ein wenig kindgerechtes, für das Kind langweiliges Verfahren handelt.

Nach Macha, Proske und Petermann (2005) sollten im Kleinkindalter die Aufgabenstellungen möglichst kurz, sowie die Anzahl der Testaufgaben möglichst niedrig sein, da Kinder in diesem Alter einen Drang haben, ihre Umwelt zu erkunden, und daher impulsiv und ablenkbar sind. Es ist für sie

schwierig, die Aufmerksamkeit über längere Zeit auf eine Aufgabe zu richten. Daher ist es auch wichtig, eine möglichst reizarme Testsituation zu gestalten. Ein weiteres Problem der Diagnostik im Kleinkindalter stellen die großen Unterschiede bezüglich der Sprachentwicklung dar. Dies alles beeinflusst die Standardisierung der Testsituation, denn dadurch sind oft mehrere Anläufe oder unterschiedliche Arten der Vorgabe der Instruktion notwendig, was sich natürlich auf die Objektivität der Untersuchung auswirkt.

Auch die große Variabilität der Entwicklungsverläufe erschwert die Beurteilung des Entwicklungsstandes eines Kindes. Das Wissen, dass Kinder durchschnittlich in einem bestimmten Alter eine gewisse Fertigkeit beherrschen, darf nicht automatisch dazu führen, aus den gebildeten Mittelwerten eine Entwicklungsreihe zu erstellen. Dies ist empirisch nicht begründbar, da viele Kinder diese Fertigkeiten auch in umgekehrter Reihenfolge erwerben (Petermann & Macha (2003)).

1.3.1 Bestehenden Verfahren für die Entwicklungsdiagnostik von Kindern unter 3 Jahren

Nach Grimm und Aktas (2002), die eine Studie zur Anwendung standardisierter psychologisch-diagnostischer Verfahren für die Entwicklungsdiagnostik in Deutschland durchführten, werden bei Kindern bis 3 Jahren nur selten entwicklungsdiagnostische Verfahren angewendet. Ein möglicher Grund dafür wird darin gesehen, dass es „kaum geeignete standardisierte Verfahren für eine adäquate und nützliche präventive Diagnostik“ (S. 167) gibt.

Im Folgenden werden nun diejenigen Entwicklungstestverfahren dargestellt, die nach Grimm und Aktas in Deutschland in der Praxis am häufigsten angewendet werden und die den hier relevanten Altersbereich von 2;0 bis 2;11 Jahren vollständig abdecken. Dabei handelt es sich um die „Münchener funktionelle Entwicklungsdiagnostik“ (MFED; Hellbrügge, 1994) (in der Untersuchung von Grimm und Aktas wurde zwar die Version aus dem Jahr 1978 genannt, aufgrund des Vorhandenseins einer neueren Version soll hier jedoch diese angeführt werden), sowie die „Bayley Scales of Infant

Development, Second Edition“ (BSID II; Bayley, 1993/2007) (auf die Beschreibung der ebenfalls häufig verwendeten „Bayley Scales of Infant Development“ (BSID; Bayley, 1969), wird verzichtet, da die „Bayley Scales of Infant Development, Second Edition“ eine aktuellere Alternative dazu darstellen). Außerdem werden noch die erst nach dem Verfassen des zitierten Artikels erschienenen „Bayley Scales of Infant and Toddler Development, Third Edition“ (Bayley III; Bayley, 2006), sowie der „Entwicklungstest 6 Jahre – 6 Monate“ (ET 6-6; Petermann, Stein & Macha, 2004), der zu diesem Zeitpunkt erst kürzlich erschienen war, vorgestellt.

1.3.1.1 Münchener funktionelle Entwicklungsdiagnostik: Zweites und drittes Lebensjahr (MFED 2-3)

Das Ziel des Stufenleiterverfahrens „Münchener funktionelle Entwicklungsdiagnostik: Zweites und drittes Lebensjahr“ (MFED 2-3; Hellbrügge, 1994) ist es, den momentanen Entwicklungsstand eines Kindes zu erheben, indem ein für einen Entwicklungsbereich repräsentativer Verhaltensausschnitt erfasst wird. Hat ein Kind mindestens 3 aufeinanderfolgende Aufgaben richtig gelöst und außerdem 3 aufeinanderfolgende schwierigere Aufgaben nicht mehr richtig gelöst, gilt ein Entwicklungsbereich als abgeklärt. Die Durchführungsdauer des Verfahrens beträgt etwa 50 Minuten. Als Ergebnis wird für jeden Entwicklungsbereich ein Entwicklungsalter berechnet. Auf die Angabe eines gesamten Entwicklungsquotienten wird verzichtet. Folgende Bereiche werden mit der MFED 2-3 erhoben:

- Laufen (Körperbewegung)
- Handgeschicklichkeit
- Perzeption (Erfassen von Zusammenhängen)
- Sprechen (aktive Sprache)
- Sprachverständnis
- Soziales Verhalten
- Selbstständigkeit

Die Aufgaben, die dem Kind hierzu gestellt werden, sind nach Reuner und Pietz (2006) aus den „Entwicklungsphysiologischen Tabellen für das Säuglingsalter“ (Hellbrügge & Pechstein, 1968), sowie aus Beobachtungen der Autoren und anderen Entwicklungstests abgeleitet.

Als Stärken des Verfahrens werden seine Übersichtlichkeit (Macha, Proske & Petermann, 2005) und seine Vollstandardisierung bezüglich des Testmaterials und der Durchführungsanweisungen (Petermann & Macha, 2003) genannt, wodurch die Durchführungsobjektivität als gegeben angesehen werden kann.

Allerdings ist die Kritik an der MFED 2-3 zahlreich. So werden die Normen, welche als längst veraltet und nicht hinreichend angesehen werden, oftmals kritisiert (siehe etwa Hagmann von Arx, Meyer & Grob, 2008; Macha, Proske & Petermann, 2005; Reuner & Pietz, 2006). Außerdem ist zu bemängeln, dass das dahinter stehende Stufenkonzept bereits einer veralteten Sichtweise der Entwicklung entspricht, und dass im Handbuch keine Angaben zur Validierung gemacht werden (Macha, Proske & Petermann, 2005). Generell wird die Anwendung der MFED 2-3 in der Praxis von den meisten Kritikern nicht empfohlen.

1.3.1.2 Bayley Scales of Infant Development, 2nd Ed. (BSID-II)

Bei den „Bayley Scales of Infant Development, Second Edition“ (BSID-II; Bayley, 1993/2007) handelt es sich um die zweite Auflage der „Bayley Scales of Infant Development“, die seit 2007 für die Anwendung im deutschsprachigen Raum vorliegt. Erfasst wird damit der Entwicklungsstand von Kindern zwischen 1 und 42 Monaten. Die Durchführungsdauer beträgt je nach Alter des Kindes 25 bis 60 Minuten. Da Bayley bei der Konstruktion der Skalen davon ausging, dass die Fähigkeiten, die in den beiden ersten Lebensjahren erworben werden, nicht verschiedenen Faktoren zugeordnet werden können, werden bei den BSID-II, wie auch schon in der ersten Auflage, lediglich eine kognitive und eine motorische Skala, sowie eine Skala zur standardisierten Beurteilung des Verhaltens gebildet. In der

deutschsprachigen Fassung sind allerdings nur die kognitive und die motorische Skala enthalten. Die *kognitive Skala* umfasst dabei:

- frühe Gedächtnisleistungen
- Habituation
- Problemlösefähigkeiten
- frühe Zahlkonzepte
- Klassifikation und Kategorisierungsfähigkeit
- Vokalisation
- sprachliche Kompetenzen
- frühe sozial-kommunikative Fähigkeiten

Die *motorische Skala* umfasst:

- zunehmende Handlungskontrolle
- grob- und feinmotorische Koordination (Bewegungskontrolle beim Krabbeln, Kriechen, Sitzen, Stehen und Rennen)
- feinmotorische Manipulation beim Greifen
- altersgemäßer Gebrauch von Stiften
- Imitieren von Handbewegungen

Die Items jeder Skala sind nach aufsteigender Schwierigkeit gereiht. Es sind Einstiegs- und Ausstiegsregeln vorhanden, anhand von denen festgemacht wird, welche Aufgabengruppe ein Kind vorgelegt bekommt. Diejenige Aufgabengruppe, bei der ein Kind mindestens 4 oder 5 Items lösen kann, wird vorgegeben. Werden in einer Aufgabengruppe alle bis auf 1 oder 2 Items gelöst, wird die nächst schwierigere Item-Gruppe vorgegeben, bis das Kind 2 oder 3 Items nicht mehr lösen kann. Die Reihenfolge der Aufgaben wird dem Untersuchungsleiter freigestellt, im Manual finden sich allerdings Vorschläge dazu. Für die Auswertung wird aus der Anzahl der richtig gelösten Items pro Skala ein Rohwert gebildet, der zu einem Indexwert umgewandelt wird und mit den Normen verglichen werden kann. Darüber hinaus kann für jede Skala ein Entwicklungsalter berechnet werden.

Die psychometrischen Charakteristiken der BSID-II gelten als adäquat (Johnson & Marlow, 2006), vor allem ihre Validierung ist angemessen

(Macha, Proske & Petermann, 2005). Johnson und Marlow (2006) bescheinigen dem Verfahren moderate bis hohe Korrelationen für verschiedene Messungen der Reliabilität und der Kriteriumsvalidität, vermissen allerdings zusätzliche Informationen zur prädiktiven Validität.

Die BSID-II weisen den Nachteil auf, dass ihre Normen mittlerweile veraltet sind, da sie bei der deutschen Bearbeitung nicht neu normiert wurden. Die Normierungsstichprobe stammt schon aus der Zeit vor 1993 und ist nicht an den deutschen Sprachraum angepasst. Weiters mögen die Freiheiten in der Durchführung der Aufgaben zwar die Vorgabe für den Untersucher einfacher machen, doch verringern sie die Objektivität des Verfahrens. Außerdem sind die Anweisungen komplex, was eine Vertrautheit des Untersuchers mit den BSID-II unabdingbar macht. Schließlich gilt das dahinterstehende Entwicklungsmodell (bei den BSID-II handelt es sich um ein Stufenleiterverfahren) mittlerweile als veraltet.

1.3.1.3 Bayley Scales of Infant and Toddler Development, 3rd Ed. (Bayley-III)

Die „Bayley Scales of Infant and Toddler Development, Third Edition“ (Bayley-III; Bayley, 2006), die mittlerweile dritte Auflage der „Bayley Scales of Infant Development“, sind für Kinder zwischen 1 und 41 Monaten konzipiert und erfassen über 5 Skalen folgende Entwicklungsbereiche:

- Kognitive Entwicklung
- Sprachliche Entwicklung (rezeptiv und expressiv)
- Motorische Entwicklung (feinmotorisch und grobmotorisch)
- Sozial-emotionale Entwicklung
- Adaptive Entwicklung

Die beiden letzteren Bereiche werden dabei über eine Befragung der Bezugspersonen erhoben. Hauptanliegen der Bayley-III ist es, Entwicklungsverzögerungen herauszufinden und Informationen für die Interventionsplanung zur Verfügung zu stellen. Die Durchführungsdauer beträgt je nach Alter des Testkindes 50 bis 90 Minuten.

Für die Bayley-III existiert noch keine Übersetzung in die deutsche Sprache, was sie für den deutschsprachigen Raum noch nicht verwendbar macht. Daher kann hier nur auf die Vorgängerversion, die „Bayley Scales of Infant Development, Second Edition“ (Bayley, 1993/2007), von welcher eine Übersetzung ins Deutsche erhältlich ist, zurückgegriffen werden (siehe oben).

Die Kritiken zu den Bayley-III sind im Allgemeinen gut. Gelobt werden größtenteils die aktuellen und (zumindest für die USA) repräsentativen Normen (zum Beispiel Reuner & Pietz, 2006), sowie die guten psychometrischen Charakteristiken, die alle zumindest minimalen Anforderungen entsprechen (Albers & Grieve, 2007; Reuner & Pietz, 2006). So sind bezüglich der Reliabilität zufriedenstellende Daten zur Retest-Reliabilität und zur inneren Konsistenz vorhanden. Was die Validität betrifft, so sind Korrelationen zu den BSID-II, der „Wechsler Preschool and Primary Scales of Intelligence – Third edition“ (Wechsler, 2002), der „Preschool Language Scale – Fourth edition“ (Zimmermann, Steiner & Pond, 2002), den „Peabody Developmental Motor Skills – Second edition“ (Folio & Fewell, 2000), sowie den „Vineland Adaptive Behavior Scales – Interview edition“ (Sparrow, Balla & Cicchetti, 1984) vorhanden.

1.3.1.4 Entwicklungstest 6 Monate – 6 Jahre (ET 6-6)

Der „Entwicklungstest 6 Monate – 6 Jahre“ von Petermann, Stein und Macha (2004) umfasst 6 Beschreibungsdimensionen, welche insgesamt 13 Entwicklungsdimensionen abbilden: Körpermotorik, Handmotorik, kognitive Entwicklung (Gedächtnis, Handlungsstrategien, Kategorisieren, Körperbewusstsein), Sprachentwicklung (rezeptive Sprachentwicklung, expressive Sprachentwicklung), Sozialentwicklung (Interaktion mit Erwachsenen, Interaktion mit Gleichaltrigen, Verhalten in Gruppen, soziale Eigenständigkeit) und emotionale Entwicklung.

Zusätzlich gibt es einen Subtest „Nachzeichnen“, der ab dem 4. Lebensjahr eingesetzt wird. Ziel des ET 6-6 ist es, ein möglichst aussagekräftiges Entwicklungsprofil eines Kindes zu erstellen, um Auffälligkeiten in der

Entwicklung differenziert zu erfassen, Stärken beziehungsweise Entwicklungsdefizite herauszufinden und dadurch Entwicklungsprognosen zu unterstützen. Das Verfahren ist für Kinder von 6 Monaten bis 6 Jahre gedacht. Da dies eine relativ große Altersspanne darstellt, ist nicht jede der Dimensionen für jede Altersgruppe gleich wichtig. Ist im Säuglingsalter etwa noch die Körper- oder die Handmotorik von größter Bedeutung, müssen im Kleinkindalter auch alle anderen Bereiche genau erhoben werden. Der gesamte Altersbereich wird in 12 Altersgruppen eingeteilt. Im Alter von 6 bis 24 Monaten erfolgt eine Einteilung in Drei-Monats-Schritten, der Altersbereich von 24 Monaten und 1 Tag bis 48 Monaten wird in Sechs-Monats-Schritte eingeteilt und die Altersgruppe von 48 Monaten und 1 Tag bis 60 Monaten wird in Zwölf-Monats-Schritte unterteilt. Für jede Altersgruppe ist eine eigene, altersentsprechende Aufgabenzusammenstellung vorhanden, wobei sich die Items sehr wohl überschneiden können. Das Testverfahren umfasst insgesamt eine systematische Verhaltensbeobachtung in der Testsituation sowie eine Befragung der Eltern mittels Fragebogen, da nicht alle Verhaltensweisen, die von Interesse sind, in der Testsituation erfasst werden können. Die Durchführungsdauer hängt vom Alter des Kindes ab (sie steigt ab ca. 24 Monaten kontinuierlich an) und beträgt in etwa 30 bis 90 Minuten. Für die Auswertung werden die Items jeder Dimension summiert und in einen Testwert transformiert. Ergebnis ist ein Entwicklungsprofil, das mit den Normen verglichen werden kann. Als über- oder unterdurchschnittliche Leistungen werden Ergebnisse gesehen, die mehr als 2 Standardabweichungen vom Referenzgruppenmittelwert abweichen.

Positiv wird am ET 6-6 gesehen, dass die Durchführungs-, Auswertungs- und Interpretationsobjektivität aufgrund der genauen Durchführungsanweisungen und der Standardisierung gegeben ist (Lücking & Scheithauer, 2006). Zur Validität wurde eine Untersuchung durchgeführt, welche das Ergebnis brachte, dass das Verfahren in der Lage ist, frühgeborene von reifen Kindern zu unterscheiden (Macha, Proske & Petermann, 2005). Auch die Kriteriumsvalidität ist zum Beispiel durch eine vorhandene Korrelation mit der „Kaufman Assessment Battery for Children“ (K-ABC; Melchers & Preuß,

1994) gegeben. Weiters zeigt die reine Fragebogenversion des ET 6-6 insgesamt sehr gute Übereinstimmungen mit der Originalversion (Lissmann, Domsch & Lohaus, 2006). Bezüglich der Reliabilität sind geringe bis mittlere Skalengkorrlationen (Hagmann von Arx, Meyer & Grob, 2008) und Stabilitäten im kognitiven und emotionalen Bereich von 6 Monaten auf 12 Monate, nicht aber von 6 auf 24 oder von 12 auf 24 Monate (Lissmann, Domsch & Lohaus, 2006; Lissmann, Korntheuer & Lohaus, 2007), zu beobachten.

Als Schwäche des ET 6-6 gilt etwa die Tatsache, dass die Durchführung aufwendig und unübersichtlich ist, da abhängig vom Alter des Kindes eine unterschiedliche Anzahl von Items vorgegeben wird (Sarimski, 2002). Reuner und Pietz (2006) merken an, dass ein Vergleich der Ergebnisse mit der Normstichprobe im Säuglingsalter aufgrund der Drei-Monats-Schritte wenig differenzierend ist. Außerdem ist die Normstichprobe nicht umfangreich genug (Naggl, 2007). Notwendig wären nach Lücking und Scheithauer (2006) neben einer Erweiterung der Normstichprobe auch zusätzliche Befunde zur Reliabilität und Validität. Dazu ist allerdings zu sagen, dass die Berechnung der Reliabilitäten aufgrund der Konstruktionsmerkmale des ET 6-6 von den Autoren nicht vorgesehen ist. Da zu erwarten ist, dass sich die Testleistungen zu einem späteren Messzeitpunkt verbessert haben, ist etwa die Bestimmung einer Retest-Reliabilität nicht sinnvoll. Auch die Berechnung der Splithalf-Reliabilität ist nicht möglich, da bewusst heterogene Skalen konstruiert wurden (Lücking & Scheithauer, 2006). Darüber hinaus wird kritisiert, dass die Genauigkeit der Differenzierung zwischen Normal- und Risikobereich fehlt. Diese Zuordnung wird beim ET 6-6 der Interpretation des Auswerters überlassen. Außerdem enthalten einzelne Entwicklungsdimensionen nur einige wenige Items (Naggl, 2007). Ein weiterer Kritikpunkt am ET 6-6 ist, dass Skalierung, Homogenität und Messgenauigkeit der Skalen nicht überprüft wurden, aber dennoch die gelösten Aufgaben zu Summenscores zusammengefasst und interpretiert werden. Da dies eigentlich nicht zulässig ist, kann der ET 6-6 streng genommen nicht als Entwicklungstest betrachtet werden (Kastner-Koller & Deimann, in Druck).

Keines der eben vorgestellten Verfahren überzeugt durchgehend für den Einsatz in der Praxis. Entweder gilt das Entwicklungskonzept als veraltet (MFED 2-3, BSID-II), die Normen sind nicht repräsentativ, beziehungsweise ebenfalls veraltet (MFED 2-3, BSID-II), das Verfahren ist für die Anwendung im deutschen Sprachraum nicht geeignet (Bayley-III), oder es kann streng genommen gar nicht als Entwicklungstest betrachtet werden (ET 6-6). Außerdem besteht das Problem, dass einige Verfahren die praktische Anwendung sehr kompliziert gestalten. Die Empfehlung eines der angeführten Testverfahren für die Anwendung bei der Diagnostik von Kleinkindern gestaltet sich daher als schwierig.

1.4 Zusammenfassung

Im vergangenen Kapitel wurde erläutert, was unter „Entwicklungsdiagnostik“ genau zu verstehen ist, welche Ziele diese verfolgt und mithilfe von welchen Mitteln sie durchgeführt wird. Danach wurden die Anforderungen, die an eine qualitativ hochwertige Entwicklungsdiagnostik gestellt werden, angeführt, und schließlich auf die Entwicklungsdiagnostik speziell im Alter von 2;0 bis 2;11 Jahren eingegangen. Im Zuge dessen wurden einige Verfahren, die für diesen Altersbereich gedacht und verfügbar sind, vorgestellt und kritisch betrachtet. Dabei handelte es sich um die „Münchener funktionelle Entwicklungsdiagnostik“ (MFED; Hellbrügge, 1994), die „Bayley Scales of Infant Development, Second Edition“ (BSID II; Bayley, 1993/2007), die „Bayley Scales of Infant and Toddler Development, Third Edition“ (Bayley III; Bayley, 2006), sowie um den „Entwicklungstest 6 Jahre – 6 Monate“ (ET 6-6; Petermann, Stein & Macha, 2004).

2 Spielbasierte Diagnostik

Unter anderem aus den Problemen bei der Entwicklungsdiagnostik im Kleinkindalter heraus wurde der Ansatz der spielbasierten Diagnostik geboren. Damit sollte ein Instrument geschaffen werden, das Kinder nicht in

ihren Handlungen einschränkt, sondern zulässt, dass sie ihre Fähigkeiten entfalten und ihnen die Möglichkeit gibt, diese auch in der Untersuchungssituation zu zeigen.

2.1 Begriffsdefinition

Unter dem Begriff „Spielbasierte Diagnostik“ wird die Beobachtung des Spielverhaltens eines Kindes als diagnostisches Instrument verstanden. Zum einen fällt darunter die Beobachtung des interaktiven Spiels mit Erwachsenen oder Peers, zum anderen das Spiel mit Spielzeug oder anderen Objekten, das sogenannte object play (Vig, 2007). Vor allem durch die Beobachtung dieser Art von Spiel lassen sich Rückschlüsse auf den kognitiven Entwicklungsstand eines Kindes ziehen, denn das object play fordert vom Kind, eine Reihe von kognitiven Leistungen aufzubringen. Es muss zum einen seine Aufmerksamkeit auf das Spielzeug fokussieren, zum anderen Motivation zeigen, schließlich aber auch metakognitive Prozesse anwenden, um etwa Spielthemen zu entwickeln (Vig, 2007; Vig & Sanders, 2007). Im Normalfall zeigen Kinder im Spiel ihren höchsten Fähigkeitslevel innerhalb von fünf Minuten (Vig, 2007). Aber auch kognitive Einschränkungen, die das Kind möglicherweise aufweist, können im Spiel beobachtet werden. Sie zeigen sich vor allem durch eine unspezifische Handhabung des Spielzeugs, weniger Kombinationen verschiedener Spielsachen, Wiederholung der immer selben Handlungen, oder auch durch Passivität. Schließlich kann die Beobachtung des Spielverhaltens aber auch als Möglichkeit zur Beobachtung des spontanen Sprachverhaltens eines Kindes dienen (Vig & Sanders, 2007).

Die spielbasierte Diagnostik kann aber sehr wohl auch mit anderen Verfahren kombiniert werden. In Kombination mit standardisierten, klassischen Verfahren kann die Beobachtung des Spielverhaltens als Screening (Vig, 2007), oder als Validitätskontrolle der Ergebnisse, die in standardisierten Verfahren erzielt wurden (Vig & Sanders, 2007), angewendet werden. Sie kann aber auch Informationen über diejenigen Kinder liefern, die von standardisierten Verfahren als „untestbar“ deklariert

wurden (Vig, 2007; Vig & Sanders, 2007). Zur „Untestbarkeit“ mancher Kinder soll im Abschnitt „Nachteile der klassischen Diagnostik“ noch näher eingegangen werden.

2.2 Anfänge der spielbasierten Diagnostik

Nach Witzlack (2001) entwickelte sich die spielbasierte Entwicklungsdiagnostik aus der spielanalytischen Entwicklungsdiagnostik heraus und hat als theoretischen Ausgangspunkt die Spieltheorie von Sigmund Freud.

Als Begründerin der Spieldiagnostik sieht Witzlack (2001) Melanie Klein an, die das Spielverhalten von Kindern psychoanalytisch deutete, indem sie Spielgegenstände als Symbole zur Triebbefriedigung betrachtete. Melanie Klein richtete ein eigenes Spielzimmer zur Beobachtung der Kinder ein, weder Durchführung noch Instruktion waren bei diesen Beobachtungen allerdings standardisiert.

Stark beeinflusst wurde die spielbasierte Diagnostik von Margaret Lowenfeld, der Leiterin einer Ausbildungseinrichtung für Kinderpsychotherapie. Sie ließ in ihrer „Welttechnik“ die zu beobachtenden Kinder mit verschiedenem Spielmaterial eine Welt aus Spielsachen aufbauen, die sie dann anhand der eigenen Worte des Kindes deutete. Jedoch wurde diese Methode nicht als Testverfahren verwendet, sondern um im Verlauf einer Therapie psychische Veränderung innerhalb eines Kindes festzustellen. Charlotte Bühler entwickelte dieses Verfahren weiter, indem sie es einer Reihe von nicht entwicklungsauffälligen Kindern verschiedener Altersstufen vorgab und sowohl die Instruktion als auch die Auswertung des Verfahrens anhand dieser standardisierte. Gösta Harding vereinigte diese beiden Ansätze schließlich im sogenannten „Dreiweltentest“ (vgl. Witzlack, 2001).

2.3 Verschiedene Ansätze der spielbasierten Diagnostik

Für ein besseres Verständnis dafür, was die spielbasierte Diagnostik nun umfasst, welche Vorgehensweisen sie beinhaltet, wodurch sie charakterisiert ist, vor allem aber dafür, wie verschiedenartig sie aufgebaut sein kann, sollen nachfolgend einige Ansätze vorgestellt werden, welche den zentralen Gedanken der spielbasierten Diagnostik beinhalten und umsetzen.

2.3.1 Ein Ansatz spielbasierter Entwicklungsdiagnostik in der DDR

Laut Witzlack (2001) wurde im Zuge einer Reform der Einschulungsdiagnostik in der DDR in den 1980er-Jahren vermehrt Augenmerk auf das kindliche Spielverhalten gelegt, da das Spiel die Haupttätigkeit eines Vorschulkindes darstellt. Aufgrund dessen ging man dazu über, das Spielverhalten von Vorschulkindern zu analysieren. Witzlack (1988, zitiert nach Witzlack, 2001) entwickelte dazu verschiedene Rollenspiele, wie etwa den Einschulungstest „Kaufmannsladen“. Im Zentrum stand bei dessen Entwicklung die Frage, ob es möglich ist, einen objektiven Entwicklungstest zu entwickeln, indem Rollenspiele standardisiert werden. Witzlack ist der Meinung, dass Standardisierung erreicht wird, indem fünf Handlungskomponenten systematisch variiert werden und somit der Inhalt, der Verlauf und die Richtung des Spiels gesteuert werden können. Diese fünf Komponenten sind:

- Rollenspielinhalt (das Thema des Spiels, z.B. Kaufmannsladen, Kindergeburtstag, Besuch im Zoo, u.ä.)
- Spielrolle des Kindes (davon sind die Handlungsmöglichkeiten des Kindes abhängig)
- Spielrolle des Spielpartners (als Spielpartner fungiert der Untersucher)
- Art und Anzahl der Spielgegenstände

- Sukzession des Angebots der Spielgegenstände (es werden immer wieder neue Handlungssequenzen eingeleitet)

2.3.2 Transdisciplinary Play-Based Assessment (TPBA)

Das TPBA von Linder (1990) dient laut dessen Autorin der Einschätzung der Entwicklung eines Kindes durch Beobachtung. Transdisziplinär bedeutet, dass die Beobachtung von mehreren Fachleuten aus verschiedenen Bereichen vorgenommen wird und diese Fachleute im Zuge dessen eng zusammenarbeiten. Ein Verständnis jedes Mitarbeiters für die jeweiligen Fachbereiche der anderen Mitarbeiter ist dabei essentiell. Dieser Ansatz unterscheidet sich vom interdisziplinären Ansatz insofern, dass bei diesem das Kind von verschiedenen Fachleuten separat begutachtet wird, diese allerdings gemeinsam über die Ergebnisse des Kindes diskutieren. Beim multidisziplinären Ansatz hingegen wird das Kind von verschiedenen, voneinander völlig unabhängigen Fachleuten begutachtet und es erfolgt danach keine gemeinsame Besprechung der Ergebnisse.

Das zu beobachtende Kind muss für die Anwendung des TPBA einen Entwicklungsstand von 6 Monaten bis 6 Jahre aufweisen. Die Beobachtung, die etwa 60 bis 90 Minuten lang dauert, kann so gut wie überall stattfinden und um diese zu vereinheitlichen, sind Richtlinien vorhanden, durch welche die Beobachtung geleitet wird. Generell werden die sensumotorische, die sprachliche, die kognitive, die emotionale, sowie die soziale Entwicklung beobachtet. Zusätzlich zu den Fähigkeiten des Kindes in diesen Entwicklungsbereichen werden aber auch die zugrundeliegenden Entwicklungsprozesse, der Lernstil und die Interaktion des Kindes erfasst. Verwendet wird das TPBA sowohl für die Diagnostik, als auch für die Interventionsplanung, es kann jedoch auch für die formative oder summative Evaluation eines Interventionsprogrammes eingesetzt werden.

Eine der mitarbeitenden Personen stellt dabei den Spiel-Moderator dar, also diejenige Person, die mit dem Kind interagiert, es anspricht, ihm Struktur vorgibt und sein Verhalten spiegelt. Eine Person wird als Familien-Moderator eingesetzt. Diese Person arbeitet mit den Eltern des Kindes und

kommuniziert mit ihnen. Eine andere Person filmt die Spielszene, alle übrigen führen Live-Beobachtungen durch. Die überaus wichtige Rolle der Eltern gestaltet sich so, dass sie einen Fragebogen über das Verhalten, welches das Kind zu Hause zeigt, ausfüllen, damit das Verhalten und das Können des Kindes während der Beobachtung damit verglichen werden können. Weiters spielen sie auch mit dem Kind, damit die Eltern-Kind-Interaktion beobachtet werden kann. Die Eltern sind zentraler Bestandteil der gesamten Untersuchung und nehmen an allen Treffen und Besprechungen teil.

Der Ablauf der Beobachtung sieht folgendermaßen aus: Die erste Spielsequenz beinhaltet unstrukturiertes Spiel, das durch das Kind geleitet wird. Danach folgt eine Phase strukturierten Spiels. Dabei wird das Verhalten, welches das Kind spontan noch nicht gezeigt hat, durch den Spiel-Moderator eingeleitet. Die nächste Sequenz beinhaltet das Spiel mit einem Peer, um die Kind-Kind-Interaktion zu beobachten. Anschließend folgt eine Phase unstrukturierten sowie strukturierten Spiels mit einem Elternteil, um einen Eindruck von der Eltern-Kind-Interaktion zu bekommen. Danach wird unstrukturiertes sowie strukturiertes motorisches Spiel eingeleitet. Zum Schluss bekommt das Kind einen Snack, damit auch die oral-motorischen Fähigkeiten beobachtet werden können. Nach der Beobachtungseinheit findet ein Post-Session-Meeting statt. Die Videoaufnahme wird analysiert und es wird kontrolliert, ob die vorgegebenen Richtlinien bei der Beobachtung eingehalten wurden. Danach werden die Stärken und Schwächen des Kindes zusammengefasst, wobei für jeden Entwicklungsbereich ein Entwicklungsalter angegeben wird, in dem sich das Kind präsentiert hat. In einem Programmplanungs-Meeting wird schließlich ein Interventionsprogramm überlegt und die Ergebnisse, sowie alle weiterführenden Empfehlungen, werden in einem Bericht zusammengefasst.

2.3.3 Play Assessment Scale (PAS)

Die PAS von Fewell (1992, zitiert nach Athanasiou, 2007) ist ein Beobachtungsverfahren, das für Kinder zwischen 2 und 36 Monaten gedacht

ist. Die Beobachtung kann in jeder beliebigen Umgebung stattfinden, solange am Boden genügend Platz für das Spiel des Kindes besteht. Verschiedene Gruppen von Spielzeug, die auf das jeweilige Alter des Kindes abgestimmt sind, werden dabei in das Spiel mit einbezogen. In einem ersten Durchgang wird das Spielverhalten beobachtet, das das Kind spontan zeigt, danach werden ihm Items vorgegeben, um das noch nicht beobachtete Verhalten zu erfassen und eine Obergrenze der Entwicklung zu bestimmen. Aus der Beobachtung des spontanen Spielverhaltens wird schließlich ein allgemeines Spielalter errechnet und die Beobachtung des strukturierten Spieldurchganges dient der qualitativen Beschreibung des Entwicklungsstandes des Kindes (Athanasiou, 2007).

2.3.4 The Developmental Play Assessment (DPA) Instrument

Das DPA Instrument (Lifter, Edwards, Avery, Anderson & Sulzer-Azaroff, 1988, zitiert nach Lifter, 2000; Lifter, 1996, zitiert nach Lifter, 2000) umfasst eine Beobachtung des Spielverhaltens von Kindern mit kognitiven Verzögerungen oder Behinderungen. Das Ziel besteht darin, deren Können und Nicht-Können unabhängig von der Sprachentwicklung zu bestimmen. Das DPA Instrument dient sowohl der Diagnostik als auch der Intervention. Dabei nimmt das Spiel in diesem Ansatz eine bedeutende Rolle ein, da für die Autoren die Entwicklung des Spielverhaltens eines Kindes genauso wichtig ist wie etwa die Entwicklung der Sprache oder der sozialen Interaktionen.

Etwa 30 Minuten lang wird das unstrukturierte Spiel des Kindes in Gegenwart eines bekannten Erwachsenen und in einer bekannten Umgebung beobachtet und auf Video aufgezeichnet. Vier Gruppen von Spielzeug werden verwendet, um dem Kind möglichst vielfältige Handlungsmöglichkeiten zu bieten, wobei sich das Kind ungefähr sieben bis acht Minuten mit jeder Spielzeug-Gruppe beschäftigen soll. Die mit dem Kind interagierende Person hat die Aufgabe, das Verhalten des Kindes zu spiegeln und verbal zu reflektieren, sowie die Aufmerksamkeit des Kindes von einer Spielzeug-Gruppe zur nächsten zu lenken. Im Nachhinein wird die

Häufigkeit des Vorkommens aller spontanen Spielverhaltensweisen des Kindes notiert. Danach werden diese Verhaltensweisen in Kategorien zusammengefasst und die Aktivitäten des Kindes werden dahingehend bestimmt, ob das Kind diese schon richtig durchführen kann, sie gerade lernt, oder noch nicht kann. Daraus wird schließlich der Entwicklungsstand des Kindes abgeleitet (Lifter, 2000).

2.3.5 The Play Observation Kit (POKIT)

Das POKIT (Mogford-Bevan, 2000) dient der Beobachtung des spontanen, interaktiven Spielverhaltens von Kindern zwischen 12 und 48 Monaten, bei welchen schon ein Defizit in zumindest einem Entwicklungsbereich festgestellt wurde und für die andere diagnostische Verfahren ungeeignet wären. Verwendet wird ein Set von standardisiertem Spielzeug, mit welchem das Kind etwa 15 bis 20 Minuten lang spielen soll. Fünf Checklisten mit Items, von welchen jedes einer Verhaltensweise im Spiel mit dem Spielzeug entspricht, unterstützen die Beobachtung. Zu jedem Item ist ein Altersbereich angeführt, in welchem das Verhalten bei einer unauffällig entwickelten Gruppe von Kindern beobachtet wurde. Auch ein Elternteil wird in die Analyse der Beobachtungen mit einbezogen, um das Verhalten, welches das Kind in der Beobachtungssituation zeigt, in Bezug zu dem Verhalten zu setzen, das es tagtäglich zeigt. Der Elternteil soll auch mit dem Kind interagieren, um ein typisches Spielverhalten des Kindes hervorzurufen. Schließlich wird aus der Beobachtung ein Urteil über den allgemeinen kognitiven Entwicklungsstand des Kindes gebildet (Mogford-Bevan, 2000).

2.4 Spielbasierte und klassische Diagnostik bei Kindern von 2;0 bis 2;11 Jahren im Vergleich

Im Folgenden werden die Inhalte der klassischen Diagnostik und die verschiedenen Ansätze der spielbasierten Diagnostik, sowie deren Vor- und Nachteile bei der Anwendung an Kleinkindern thematisiert. Wie schon

erwähnt, interessiert dabei vorwiegend die Anwendung dieser Ansätze an Kindern, die noch keinen Kindergarten besuchen, das sind meist Kinder in einem Alter von 0 bis 3 Jahren. Ein Grund dafür ist, dass für Kinder ab 3 Jahren schon Entwicklungstestverfahren auf dem Markt sind, wie etwa der Wiener Entwicklungstest (WET; Kastner-Koller & Deimann, 2002), die den Bedürfnissen von Kindern im Alter von 3 bis 6 Jahren gerecht werden. Da dies für Kinder unter 3 Jahren nicht der Fall ist, soll hier auf gerade diese Kinder eingegangen werden. Weil die Entwicklungsfortschritte in diesem Alter allerdings so zahlreich und vielfältig sind, dass ein Altersbereich, der drei Jahren umfasst, zu unüberschaubar wäre, wurde für diese Studie der Altersbereich von 24 bis 36 Monaten ausgewählt, also das Jahr bevor im Normalfall der Kindergarteneintritt erfolgt.

Unter dem Begriff „Klassische Diagnostik“ wird nachfolgend die Anwendung von Tests im engeren Sinne als Untergruppe der psychologisch-diagnostischen Verfahren verstanden, wie es bei Kubinger (2009) zu finden ist. „Spielbasierte Diagnostik“ beschreibt im Gegensatz dazu die Beobachtung des Spielverhaltens eines Kindes, verwendet als diagnostisches Instrument. Denn nach Athanasiou (2007) steht das Spielverhalten in einer wechselseitigen Beziehung mit der kognitiven, sozial-emotionalen, sprachlichen, sowie der motorischen Entwicklung.

2.4.1 Klassische Diagnostik

Meist bestehen psychometrische Tests aus einer Reihe von Aufgaben, wobei jede Aufgabe ein Testitem darstellt. Sie sind größtenteils norm-orientiert, die Ergebnisse, die eine Person in einem Testverfahren erreicht, werden also zu einer bestimmten Stichprobe in Beziehung gesetzt. Die klassische Diagnostik, die in dieser Arbeit definiert ist als die Anwendung von Testverfahren im engeren Sinne als Untergruppe der psychologisch-diagnostischen Verfahren (vgl. Kubinger, 2006), ist das, was umgangssprachlich unter dem Begriff „testen“ verstanden wird.

Nach Cicchetti und Wagner (1990) dient eine Diagnostik der Erfassung der Position eines Individuums in Relation zu einer größeren, definierten Population hinsichtlich bestimmter psychologischer Konstrukte. Aufgrund der Position, die für die jeweilige Person bestimmt wird, werden im Anschluss an die Diagnostik Interventionsmaßnahmen vorbereitet. Um schließlich im Zuge der Intervention die weitere Entwicklung zu erfassen, also den Erfolg der Interventionsmaßnahmen kontrollieren zu können, kann jederzeit wieder eine Diagnostik durchgeführt werden.

Ein Großteil der klassischen psychometrischen Verfahren erfasst die Fähigkeiten der untersuchten Person über die Sprache. Da dies natürlich erst möglich ist, sobald sich Sprachproduktion und Sprachverständnis bis zu einem gewissen Grad in einem Kind entwickelt haben, sind die meisten bestehenden Verfahren frühestens ab einem Alter von zwei Jahren einsetzbar (Preston, 2005). In diesem Alter, das ja genau das hier relevante Alter darstellt, werden bei einer Diagnostik vorwiegend Entwicklungstests eingesetzt, da das psychologische Interesse an einem Kleinkind meist darin liegt, zu erfassen, wie weit fortgeschritten die Entwicklung des Kindes zu einem gewissen Zeitpunkt ist.

Appl (2000) nennt vier Gründe für die Informationserfassung mittels Diagnostik. Diese beziehen sich zwar konkret auf den Altersbereich der Drei- bis Fünfjährigen, haben aber mit Sicherheit auch Gültigkeit in der hier relevanten Altersgruppe der zwei- bis dreijährigen Kinder:

- Screening (Identifikation von Kindern, die eine weiterführende Diagnostik benötigen; dabei sollten die Information aus verschiedenen Quellen stammen; meist erfolgt ein Vergleich mit einer Norm-Stichprobe)
- Genauere diagnostische Untersuchung (Vergleich der Ergebnisse mit einer Norm-Stichprobe)
- Planung individueller Programme (die Ergebnisse der Diagnostik werden dazu verwendet, Ziele für nachfolgende Interventionsmaßnahmen festzulegen und die Intervention anhand dieser Ziele zu planen)
- Fortschrittskontrolle (wird verwendet, um die Schlussfolgerungen der Diagnostik zu validieren und Prognosen für die Zukunft zu treffen).

Schließlich meint Bagnato (2005), was die Zukunft der Diagnostik betrifft, dass sich derzeit ein Wandel bei der Untersuchung von Klein- und Vorschulkindern anbahnt, und die Tendenz immer mehr in Richtung einer Beobachtung der kindlichen Entwicklung geht. Dieser Wandel ist seiner Meinung nach auf drei Charakteristiken aufgebaut. Zum einen geht es immer mehr um Authentizität, also darum, dass Fähigkeiten erhoben werden, die Kinder in einem natürlichen Umfeld zeigen. Weiters geht es vermehrt um den Nutzen der Untersuchung, also darum, dass durch die Erhebung der Fähigkeiten eines Kindes Interventionen geplant und Ziele für Therapien gesetzt werden können. Und nicht zuletzt spielt die Allgemeingültigkeit eine Rolle, also dass es möglich sein sollte, Verfahren bei allen Kindern, ungeachtet etwaiger Einschränkungen, einzusetzen.

2.4.2 Vorteile der klassischen Diagnostik

Explizite Vorteile der klassischen Diagnostik sind in der Literatur nur wenige zu finden. Vorteile, die sich speziell auf die Anwendung klassischer Verfahren an der Altersgruppe der Zwei- bis Dreijährigen beziehen, finden sich überhaupt nicht. Was traditionelle, meist norm-orientierte Verfahren anderen Ansätzen jedoch voraus haben, ist, dass sie im Optimalfall eine hohe psychometrische Integrität besitzen (Bagnato, 2005). Sie erfüllen meist verschiedene Gütekriterien, wie zum Beispiel die Gütekriterien der Reliabilität, Validität, oder der Objektivität. Etwa sind sehr oft die Durchführung, Auswertung, beziehungsweise die Interpretation der Ergebnisse eines Verfahrens standardisiert, was dem Gütekriterium der Objektivität entspricht (vgl. Kubinger, 2009). Aber auch noch andere Charakteristiken, die sie gegenüber anderen Verfahren auszeichnen, weisen die klassischen Verfahren auf. Durch einen Vergleich mit einer Norm-Stichprobe ist ein direkter Vergleich der Ergebnisse eines Kindes mit den Ergebnissen Gleichaltriger möglich. Auch können in unterschiedlichen Verfahren erzielte Werte einander gegenübergestellt werden, indem die jeweiligen Rohwerte in Standardwerte transformiert und so vergleichbar gemacht werden können (Preston, 2005).

Ansonsten sind in der Literatur kaum Vorteile klassischer Verfahren zu finden. Zumindest aber gibt es in einigen Testhandbüchern spezielle Hinweise darauf, welche kinderfreundlichen Charakteristiken das jeweilige Testverfahren aufweist, oder wie die Testung gestaltet werden kann, damit sie auf die Bedürfnisse von Kleinkindern zugeschnitten ist.

Im Handbuch des Verfahrens „Entwicklungstest 6 Monate bis 6 Jahre“ (ET 6-6; Petermann, Stein & Macha, 2004) werden explizite Hinweise auf die Gestaltung des Untersuchungsraums, auf die Kontaktaufnahme mit dem Kind, auf das Verhalten des Testleiters gegenüber dem Kind und auf den Umgang mit den Eltern gegeben. Der Untersuchungsraum sollte beispielsweise Kindermöbel und außer dem Testmaterial keine anderen für das Kind attraktiven Gegenstände beinhalten. Der Testleiter sollte ruhig und freundlich mit dem Kind sprechen und ihm die Möglichkeit geben, den Testraum zu erkunden. Er sollte Blickkontakt suchen und dem Kind während der gesamten Testung nicht seine Leistung, sondern seine Mitarbeit rückmelden.

Bayley (1993/2007) gibt im Handbuch zu den „Bayley Scales of Infant Development, Second Edition“ in Jahresschritten unterteilt genaue Informationen zur Kontaktaufnahme mit dem Kind. Bei den 1;6- bis 2;6-Jährigen sollte etwa genügend Zeit zum Aufbau von Vertrauen vorhanden sein, die Bezugsperson sollte mit eingebunden werden und das Kind sollte ermutigt werden, sich sprachlich zu äußern. Bei 2;6 bis 3;6 Jahre alten Kindern wird vorgeschlagen, die Kontaktaufnahme mit dem Kind mit einem Gespräch über ein Spielzeug zu beginnen.

Obwohl klassische diagnostische Verfahren also, wie bereits erwähnt, keine expliziten Vorteile bei der Anwendung an Zwei- bis Dreijährigen aufweisen, ist standardisiertes Testen vor allem aufgrund der schon genannten Erfüllung von Gütekriterien sehr wohl nützlich und wichtig. Nach Vig und Sanders (2007) ist es dabei allerdings notwendig, auf die Auswahl passender Verfahren zu achten. Sollte kein perfekt passendes Verfahren zur Verfügung stehen, sollte der klinische Nutzen des jeweiligen Verfahrens über dessen psychometrischer Qualität stehen. Generell sollten für den Untersucher nach Möglichkeit immer verschiedene unterschiedliche Instrumente zur Verfügung

stehen, um eine Auswahl aus diesen zu ermöglichen. Darüber hinaus sollte die Anwendung der klassischen Diagnostik möglichst durch andere Ansätze unterstützt werden. Nicht zuletzt sollte der Untersucher bei der Anwendung der Verfahren an Kleinkindern mit dieser Altersgruppe vertraut sein und schon Erfahrung gesammelt haben. Worauf bei der Anwendung klassischer Testverfahren noch geachtet werden sollte, ist laut Harrison (2009), dass die psychometrischen Charakteristiken vieler Verfahren durch bestimmte Dinge beeinflusst werden können. Zum einen sind bei Kindern oft große interindividuelle Unterschiede bezüglich ihrer Entwicklung vorhanden. Zum anderen können unvorhersehbare Schwankungen im Verhalten des Kindes, sowie Ermüdung, fehlende Motivation oder Angst vor dem Untersucher die Leistung des Kindes beeinflussen. Auch Situationsvariablen, wie etwa Lärm, können einen Einfluss auf die Leistung haben. Außerdem bestehen zwischen Kindern oft große Unterschiede bezüglich der Erfahrungen, die sie mit bestimmten Materialien, die in der Testung verwendet werden, schon gemacht haben.

2.4.3 Nachteile der klassischen Diagnostik

Hier soll zunächst eine Studie von Bagnato und Neisworth (1994) über die Verwendung von Intelligenztests bei Kleinkindern vorgestellt werden. Die Autoren fanden heraus, dass in diesem Bereich folgende Verfahren am häufigsten angewendet werden (gereiht nach der Häufigkeit der Anwendung):

- Stanford-Binet IV
- Bayley Scales of Infant Development
- Wechsler Preschool and Primary Scale of Intelligence – Revised
- Kaufmann Assessment Battery for Children
- McCarthy Scales of Children's Abilities

Sie fanden weiters heraus, dass 42,96% der Kinder, die mit einem dieser Verfahren getestet wurden, von den Testleitern als „untestbar“ deklariert wurden. Die Autoren vertreten zu diesem Thema die Meinung, dass kein Kind untestbar ist, sondern diese Bezeichnung nur die Beschränkung

traditioneller, standardisierter Intelligenztestverfahren betitelt. Der Begriff „untestbar“ wird von Bagnato und Neisworth (1994) definiert als:

failure to establish any of the following: basal or floor level; complete scores on all necessary subtests or subdomains; clearly scorable responses on a sufficient number of individual assessment tasks; interpretable standard scores without the necessity of “creatively” (clinical approach) altering scoring procedures, modifying the administration procedures of tasks, stimulus properties of items, nor accommodating the child’s unique response modes. (S. 87)

Gründe für eine Einschätzung als „untestbar“ können laut den Autoren einerseits vorhandene Sprachdefizite bei den jeweiligen Kindern sein, oder andererseits zu hohe Sprachanforderungen der verwendeten Verfahren an die Kinder, denn immerhin bauen die meisten dieser Verfahren auf der Sprache auf. Auch kann das Verhalten des Kindes während der Testung dessen Leistung beeinflussen. Weiters sind klassische Verfahren meist rigide, sodass sie keine auf das Kind zugeschnittenen Änderungen erlauben. Außerdem können hohe item floors, also zu hoch angesetzte untere Leistungsgrenzen, aufgrund derer keine adäquate Differenzierung im unteren Leistungsbereich möglich ist, der Grund für eine „Untestbarkeit“ eines Kindes sein. Dabei ist eine zu geringe Anzahl von Aufgaben vorhanden, um auch einen niedrigen Entwicklungsstand zu erfassen.

Standardisierte Verfahren haben als Ergebnis meist einen Testscore, der die erfassten Fähigkeiten im Vergleich zu einer Normgruppe quantitativ wiedergibt. Dies erweist sich nach Preston (2005) allerdings nicht als hinreichend, da es auch Fähigkeiten gibt, die sich nicht objektiv quantifizieren lassen und über einen Score-Wert hinausgehen. Psychometrische Testverfahren liefern demnach nur ein unvollständiges Bild der untersuchten Person.

Zusätzlich fanden sich in der Literatur als explizite Nachteile der klassischen Diagnostik Verweise auf die unnatürliche Umgebung in der Testsituation, da das Setting meist hoch strukturiert ist und die Testung in Form eines Frage-Antwort-Formates stattfindet (Kelly-Vance, Ryalls & Glover, 2002). Für die Standardisierung ist dies essentiell, da so die Bedingungen bei jeder Testung und für jedes Kind die gleichen sind. Das Kind wird außerdem nicht

abgelenkt und kann sich voll und ganz auf die Testmaterialien konzentrieren (Linder, 1990). Allerdings hat diese so geschaffene Unnatürlichkeit der Umgebung zur Folge, dass das Kind seltener diejenigen Fähigkeiten zeigt, die es auch im Alltag einsetzt (Bagnato, 2005). Auch Neisworth und Bagnato (2004) sind der Meinung, dass bei der Anwendung klassischer Testverfahren nicht die Alltagsfähigkeiten eines Kindes hervorgerufen werden und daher auch nicht deren wahre Fähigkeiten erfasst werden können.

Ungeeignete Normen führen nach Linder (1990) darüber hinaus zu floor- oder ceiling-Effekten, also zu einer ungenügenden Differenzierung im oberen oder unteren Leistungsbereich, wobei gerade Kinder von zwei bis dreieinhalb Jahren zwischen die Testverfahren für Kleinkinder und die für Vorschulkinder fallen.

Kelly-Vance et al. (2002) nennen außerdem als Kritikpunkt, dass klassische diagnostische Verfahren oft rigide sind und nicht auf die Bedürfnisse des untersuchten Kindes eingehen. Dem Kind wird somit keine Gelegenheit gegeben, sein Leistungsoptimum zu zeigen. Als Ergebnis erhält man außerdem nur Scorewerte. Eine Beschreibung der Leistung des Kindes wird dabei außer Acht gelassen.

Tzuriel (2000) fasst einige zusätzliche Kritikpunkte an standardisierten psychometrischen Testverfahren zusammen. Zum einen geben sie keine Informationen über Lernprozesse, spezifische kognitive Funktionen, oder Strategien, die das Lernen ermöglichen würden. Außerdem werden keine Informationen über vorhandene metakognitive Prozesse erhoben, oder Informationen darüber, was das Kind mit der Unterstützung oder Anleitung eines Erwachsenen erreichen könnte. Zum anderen sagt ein niedriger Leistungslevel nichts über das Lernpotenzial aus, das in einem Kind vorhanden ist. Die kognitiven Prozesse, die mit dem Lernen in Verbindung stehen, werden nicht klar beschrieben und mit nicht intelligenzmäßigen Faktoren, etwa mit der intrinsischen Motivation, wird kein Zusammenhang gesehen, obwohl diese sehr wohl die kognitive Leistung beeinflussen können.

2.4.4 Vorteile der spielbasierten Diagnostik

Kelly-Vance und Ryalls (2005) meinen, spielbasierte Ansätze seien bei der Anwendung an Klein- und Vorschulkindern schon allein der Intuition nach vorzuziehen, da deren Hauptbeschäftigung das Spielen ist und sie durch das Spielen vieles lernen.

Was in der Literatur als großer Vorteil, den spielbasierte Ansätze gegenüber den meisten klassischen diagnostischen Verfahren besitzen, zu finden ist, ist, dass auch diejenigen Kinder, die von den eben genannten Verfahren als „untestbar“ deklariert wurden, begutachtet werden können, denn wie Bagnato und Neisworth (1994), sowie Linder (1990) formulierten, jedes Kind ist testbar.

Ebenfalls einen Pluspunkt stellt die natürliche Umgebung dar, in der die Untersuchung stattfindet, da eine Beobachtung des Spielverhaltens entweder in einer Umgebung stattfinden kann, die dem Kind bekannt ist, etwa zu Hause, oder zumindest Elemente beinhalten kann, die das Kind schon kennt. So können beispielsweise Spielsachen, die dem Kind gehören, mit einbezogen werden (Athanasίου, 2007; Kelly-Vance et al., 2002; Linder, 1990).

Weiters sind sie flexibel, etwa was die Material-Zusammenstellung oder Variationen im Ablauf betrifft, und erlauben es, individuell auf das jeweilige Kind einzugehen (Kelly-Vance et al., 2002; Linder, 1990).

Außerdem werden bei der Anwendung spielbasierter Ansätze meist die Eltern mit einbezogen, was zum einen dem Kind ein Gefühl von Sicherheit gibt, zum anderen eine umfassende Begutachtung des Kindes ermöglicht (Athanasίου, 2007; Linder, 1990). Abgesehen davon, dass die Eltern zusätzliche hilfreiche Informationen über das Kind in die Untersuchung einbringen können, kann die Eltern-Kind-Interaktion betrachtet werden und das Kind kann durch die Gegenwart der Eltern in der Untersuchungssituation dazu angespornt werden, typische Verhaltensweisen zu zeigen.

Nach Kelly-Vance et al. (2002) ist das Ergebnis einer Beobachtung des Spielverhaltens eines Kindes außerdem eine qualitative Beschreibung seiner

Stärken und Schwächen, also mehr als ein bloßer Score, der den Fähigkeitslevel des Kindes ausdrückt.

Witzlack (2001) ist schließlich der Meinung, dass bei dieser Art von Begutachtung keine Misserfolgserlebnisse beim Kind auftreten, da für das Kind das Ganze ja nur ein Spiel ist.

2.4.5 Nachteile der spielbasierten Diagnostik

Was in der Literatur als wohl größter Nachteil der spielbasierten Diagnostik zu finden ist, ist der nur ungenügende Nachweis von Reliabilität und Validität bei diesen Ansätzen (Athanasίου, 2007; Linder, 1990).

Weiters geben Athanasίου (2007) und Linder (1990) an, dass die Ergebnisse nicht standardisiert und daher auch nicht mit denen anderer Kinder verglichen werden können.

Außerdem sind keine Richtlinien vorhanden, die Rücksicht auf den Einfluss nehmen würden, den Kultur oder Behinderungen auf das Spielverhalten eines Kindes ausüben (Athanasίου, 2007).

Als ebenfalls großen Kritikpunkt nennen Eisert und Lamorey (1996) die Tatsache, dass die meisten spielbasierten Ansätze noch sehr experimentell gehalten sind, noch gar nicht veröffentlicht wurden, oder nur in der Theorie existieren, jedoch in der Praxis als Verfahren nicht erhältlich sind. Die einzige Ausnahme stellt laut den Autoren hier das TPBA von Linder (1990) dar, welches weiter oben schon vorgestellt wurde.

Darüber hinaus muss der Zusammenhang des Spielverhaltens eines Kindes mit Kognitionen, Sprache oder Motorik noch genauer untersucht werden und die Ergebnisse spielbasierter Diagnostik müssen noch hinreichend mit klassischen, standardisierten Verfahren verglichen werden (Eisert & Lamorey, 1996).

2.4.6 Fazit

Beide vorgestellten Formen der Diagnostik weisen also mehr oder weniger zahlreiche Vorteile und Nachteile auf. Das Problem liegt darin, dass ein einziger Ansatz, beziehungsweise ein Verfahren allein, nicht alle Wünsche des Diagnostikers erfüllen kann und daher entschieden werden muss, worauf das Augenmerk bei der diagnostischen Untersuchung gelegt werden soll. Man wird um die klassische Diagnostik nicht umhinkommen, möchte man standardisierte und vergleichbare Ergebnisse erhalten. Steht jedoch das Anliegen im Zentrum, das zu untersuchende Kind möglichst wenig zu belasten und die Untersuchungssituation so angenehm wie möglich zu gestalten, bietet sich eher ein spielbasierter Ansatz an.

Die Frage, ob eine Form der diagnostischen Untersuchung der anderen bei der Anwendung an Kleinkindern wirklich überlegen und daher vorzuziehen ist, kann aufgrund dessen keinesfalls hundertprozentig beantwortet werden. Eine endgültige Entscheidung drüber muss ohnehin situations- und kindspezifisch gefällt werden und sollte, abgesehen von dem Wissen über die Stärken und Schwächen des jeweiligen Ansatzes, vor allem auf der Erfahrung und der Intuition des Untersuchers / der Untersucherin beruhen. Idealerweise bezieht eine umfassende Begutachtung eines Kleinkindes ihre Informationen ohnehin aus mehreren unterschiedlichen Quellen, nämlich zum einen aus den Fähigkeiten, die vom Kind spontan gezeigt werden, zum anderen aus den Fähigkeiten, die unter strukturierten Bedingungen gezeigt werden, aber auch aus dem, was die Bezugsperson berichtet (Vig & Sanders, 2007).

Was sich hier allerdings nahezu aufdrängt, ist die Forderung nach einem Verfahren, das sowohl die Vorteile der klassischen Diagnostik (Gütekriterien, Vergleichbarkeit), als auch die der spielbasierten Diagnostik (Flexibilität, Kinderfreundlichkeit) vereint. Da es so etwas aktuell noch nicht gibt, ist es das Ziel dieser Studie, einen Itempool für ein solches Verfahren zu entwickeln, um damit den allgemeinen Entwicklungsstand im von Kindern von 24 bis 36 Monaten erfassen zu können.

2.5 Zusammenfassung

In diesem Kapitel wurde zuerst darauf eingegangen, worum es sich bei der spielbasierten Diagnostik handelt, nämlich um die Beobachtung des Spielverhaltens eines Kindes als diagnostisches Instrument, und welche Bedeutung diese für die Diagnostik im Kleinkindalter aufweist. Danach wurden die Anfänge der spielbasierten Diagnostik beginnend mit Melanie Klein veranschaulicht und verschiedene Ansätze vorgestellt, wie der „Kaufmannsladen“ (Witzlack, 1988, zitiert nach Witzlack, 2001), das „Transdisciplinary Play-Based Assessment“ (TPBA; Linder, 1990), die „Play Assessment Scale“ (PAS; Fewell, 1992, zitiert nach Athanasiou, 2007), das „Developmental Play Assessment (DPA) Instrument“ (Lifter et al., 1988, zitiert nach Lifter, 2000; Lifter, 1996, zitiert nach Lifter, 2000), oder der „Play Observation KIT“ (POKIT; Mogford-Bevan, 2000). Schließlich wurde ein Vergleich zwischen der klassischen und der spielbasierten Diagnostik gezogen.

Als Vorteile der klassischen Diagnostik kristallisierten sich die oftmalige Erfüllung von Gütekriterien heraus, oder der mögliche Vergleich der Ergebnisse mit den Ergebnissen Gleichaltriger. Als Nachteile fanden sich etwa ihre Rigidität oder Unnatürlichkeit. Die spielbasierte Diagnostik weist hingegen den Vorteil auf, dass sie flexibel ist und ein Eingehen auf das jeweilige Kind erlaubt. Dafür wird als wohl größter Nachteil die ungenügende Erfüllung von Gütekriterien genannt. Außerdem können Kinder desselben Alters nicht miteinander verglichen werden. Daraus wurde die Schlussfolgerung gezogen, dass eine Kombination von klassischen Ansätzen mit der spielbasierten Diagnostik ideal wäre, wenn damit einerseits eine kinderfreundliche Untersuchungssituation geschaffen wird, andererseits aber die Erfüllung von Gütekriterien und die Vergleichbarkeit von Kindern untereinander möglich ist. Ziel dieser Arbeit ist es daher, den Itempool für ein solches Verfahren theoretisch zu entwickeln und empirisch zu erproben.

3 In der vorliegenden Studie behandelte Entwicklungsbereiche mit Fokus auf das Alter von 2;0 bis 2;11 Jahren

Das Augenmerk in dieser Arbeit liegt vorwiegend auf der Erfassung der kognitiven Entwicklung. Daraus ergibt sich gleich zu Beginn die Notwendigkeit zu definieren, welche Bereiche der kindlichen Entwicklung die kognitive Entwicklung genau umfasst. Hierbei tritt allerdings schon die erste Schwierigkeit auf, da es in der Literatur keine einheitliche Meinung dazu gibt, welche Fähigkeiten und Fertigkeiten unter den Bereich der kognitiven Entwicklung fallen und welche einen eigenständigen Entwicklungsbereich bilden. Da es in dieser Arbeit um die Entwicklung und Erprobung des Itempools für ein spielbasiertes Entwicklungstestverfahren geht, soll die Zuordnung einzelner Bereiche zur kognitiven Entwicklung an die Zuordnung in anderen entwicklungsdiagnostischen Verfahren, zum Beispiel im „Wiener Entwicklungstest“ (WET, Kastner-Koller & Deimann, 2002), angelehnt werden.

Folgende Bereiche werden aufgrund dessen zur kognitiven Entwicklung gezählt: die theory of mind, das numerisches Wissen, sowie das logische Schlussfolgern. Das Gedächtnis, die visuelle Wahrnehmung und die Arbeitshaltungen beschreiben einen jeweils eigenständigen Entwicklungsbereich.

Alle genannten Bereiche werden im Folgenden dargestellt und näher erläutert. Die – ebenfalls eigenständigen – Entwicklungsbereiche Sprache, Motorik und sozial-emotionale Entwicklung bilden den Schwerpunkt in der Arbeit von Kuchler (in Druck).

Einschränkend sei erwähnt, dass die nun gebildeten Bereiche die für Kinder im 3. Lebensjahr relevanten Entwicklungsbereiche sicherlich nicht vollständig abbilden, jedoch war trotz theoretischer Überlegungen nicht immer eine praktische Umsetzung möglich.

3.1 Kognitive Entwicklung

3.1.1 Theory of Mind

Die theory of mind ist die Fähigkeit eines Menschen, einem anderen Menschen verschiedene Gefühle, Wünsche, Intentionen, oder Überzeugungen zuzuschreiben, auch wenn diese sich von den eigenen Gefühlen, Wünschen, Intentionen, oder Überzeugungen unterscheiden, und aus diesen heraus erfolgreich Handlungen vorherzusagen. Allgemein wird davon ausgegangen, dass sich die theory of mind etwa ab dem Alter von 3 Jahren zu entwickeln beginnt (siehe zum Beispiel Carlson, Moses & Breton, 2002; Carlson, Moses & Claxton, 2004; Wellman, Cross & Watson, 2001).

Für psychologisch-diagnostische Untersuchungen bedeutet dies, dass ein Kind, das die Fähigkeit zur theory of mind noch nicht erlangt hat, nicht wissen kann, was von ihm verlangt wird, weil es über die Intentionen und Wünsche des Testleiters / der Testleiterin schlichtweg nicht Bescheid weiß. Daher wird es meist auch nicht die notwendige Anstrengungsbereitschaft und Motivation an den Tag legen, da es nicht über das Wissen verfügt, dass es sein Bestes geben soll. Umso wichtiger sind die Vorläuferkomponenten der theory of mind, da sie einen Hinweis darauf liefern, wie weit die Entwicklung der theory of mind in einem Kind schon vorangeschritten ist und was daher wohl von ihm erwartet werden kann.

Da die theory of mind also eine Fähigkeit ist, die für psychologisch-diagnostische Untersuchungen eine zentrale Bedeutung besitzt, und gerade ihre Vorläufer in der vorliegenden Untersuchung aufgrund des Altersbereichs der Kinder von 2;0 bis 2;11 Jahren eine noch wichtigere Rolle spielen, soll der gesamte Themenbereich „theory of mind“ im Folgenden eingehend behandelt werden.

Für Perner (1991) ist die mentale Repräsentation von Dingen der Wirklichkeit grundlegend für die theory of mind. Er unterscheidet drei Stufen von Repräsentationen.

Zuerst ist ein Kind fähig zu *primären Repräsentationen*, es kann also etwas aus der Realität in den eigenen Gedanken abbilden („etwas ist so“; siehe auch Leslie, 1987).

In einem zweiten Schritt sind etwa im Laufe des 2. Lebensjahres *sekundäre Repräsentationen* möglich. Hierbei wird etwas nicht so repräsentiert, wie es wirklich ist, sondern die Repräsentation ist losgelöst von der Realität. Diese sekundären Repräsentationen sind es, die es erlauben, über etwas nachzudenken, das von der momentanen Realität abweicht, wie etwa die Vergangenheit, die Zukunft oder Hypothetisches (zum Beispiel: „Ich glaube, etwas ist so“). Allerdings sind sie eng an das Vorhandensein von primären Repräsentationen gebunden, da diese ihnen den notwendigen Bezug zur Realität verschaffen.

Schließlich verfügt ein Kind in der letzten Stufe über *Metarepräsentationen*, also die Fähigkeit, zu repräsentieren, dass etwas (von jemand anderem) repräsentiert wird (zum Beispiel „Ich glaube, dass jemand anderer glaubt, es ist so“). Dazu wird das Verständnis benötigt, zwischen dem, was repräsentiert wird, und dem, als was es repräsentiert wird, zu unterscheiden, was wiederum nur mithilfe der sekundären Repräsentationen möglich ist. Speziell diese Stufe ist nach Perner (1991) ausschlaggebend für das Vorhandensein einer theory of mind.

3.1.1.1 Erfassung der theory of mind

Die theory of mind wird am häufigsten mit false-belief-Aufgaben oder appearance-reality-Aufgaben erfasst. Bei false-belief-Aufgaben wird gemessen, inwieweit eine Person in der Lage ist, jemandes falsche Überzeugung zu begreifen, appearance-reality-Aufgaben untersuchen, inwieweit jemand den Unterschied zwischen der Wirklichkeit und der Erscheinung eines Objekts unterscheiden kann. Die häufigste Form von false-belief-Aufgaben sind sogenannte Platzwechsel-Aufgaben, welche stellvertretend für alle verschiedenen Variationen von false-belief-Aufgaben nachfolgend dargestellt werden sollen.

3.1.1.1.1 *False-belief – Aufgaben zur falschen Überzeugung – Platzwechsel-Aufgaben*

Bei diesen Aufgaben wird dem Kind eine Geschichte erzählt und anschließend eine Frage dazu gestellt. Eine solche Geschichte lautet sinngemäß folgendermaßen: „Eine Person versteckt ein Objekt am Platz A. Dann verlässt sie den Raum und eine zweite Person transferiert das Objekt von A nach B. Danach kommt die erste Person wieder zurück.“ Anschließend wird die Frage gestellt: „Wo wird sie das Objekt suchen, beziehungsweise wo wird sie glauben, dass das Objekt versteckt ist?“ Versteht das Kind schon, dass andere Personen Überzeugungen haben, die unabhängig vom eigenen Wissen (dass das Objekt im Versteck B ist) sind, wird es die Frage korrekt beantworten, nämlich dass die Person das Objekt dort suchen wird, wo sie es versteckt hat, nämlich im Versteck A. Versteht das Kind dies noch nicht, wird es als Antwort das Versteck B nennen, also den Ort, an dem sich das Objekt tatsächlich befindet.

Die bekannteste Platzwechsel-Aufgabe ist wohl die von Wimmer und Perner (1983) entwickelte Geschichte von Maxi und der Schokolade. Bei ihrer ersten Anwendung dieser Aufgabe konnten 78% der 4- bis 5-Jährigen die richtige Antwort finden, jedoch nur 15% der 3- bis 4-Jährigen, was für einen plötzlichen Zuwachs in der Fähigkeit zum Wissen um falsche Überzeugungen im Alter von etwa 4 Jahren spricht, oder anders ausgedrückt für eine stetige Zunahme dieser Fähigkeit vom Alter von 3 bis ins Alter von 5 Jahren. Dieses Bild zieht sich durch den Großteil der Studien zu false-belief-Aufgaben hindurch (siehe zum Beispiel Carlson, Moses & Breton, 2002; Carlson, Moses & Claxton, 2004; oder Wellman, Cross & Watson, 2001).

Eine Metaanalyse von Wellman et al. (2001) brachte das Ergebnis, dass Aufgabenvariationen dabei keinen Einfluss auf die Leistung der Kinder bei false-belief-Aufgaben haben. Lediglich das Alter wirkt sich auf ihre Leistung aus.

Ziv und Frye (2003) vermuteten, dass die Probleme Dreijähriger bei dieser Art von Aufgaben dadurch zustande kommen, dass sie bei der Beantwortung der Frage auf das Begehren der Person in der Geschichte Bezug nehmen und in Übereinstimmung mit diesem antworten. Also stellten sie Kindern die

Aufgabe mit nicht begehrten Objekten. Es zeigte sich, dass das Begehren der Person in der Geschichte keinen Einfluss auf die Zuschreibung der Überzeugung durch die Kinder hatte.

3.1.1.1.2 Appearance-reality – Aufgaben zur Unterscheidung von Erscheinungsbild und Wirklichkeit

Als Pioniere auf dem Gebiet der Unterscheidung zwischen Erscheinungsbild und Wirklichkeit gelten Flavell und Kollegen (z.B. Flavell, 1986; Taylor & Flavell, 1984). Ihre Vorgehensweise zur Erhebung dieser Fähigkeit sieht so aus, dass den Versuchspersonen ein Gegenstand gezeigt wird, der nicht das ist, was er zu sein scheint (z.B. ein Stein, der aussieht wie ein Ei). Die Kinder werden gefragt, wonach dieser Gegenstand aussieht, danach wird ihnen gezeigt, was der Gegenstand in Wirklichkeit ist. Anschließend werden sie gefragt, wonach er aussieht (wie ein Stein oder wie ein Ei?) und was er wirklich ist (ein Stein oder ein Ei?). Es zeigte sich, dass Dreijährige nicht zwischen Anschein und Wirklichkeit unterscheiden können. Entweder antworten sie, der Gegenstand sehe aus wie und sei ein Ei, ein sogenannter Erscheinungsfehler, oder er sehe aus wie und sei ein Stein, ein sogenannter intellektueller Realitätsfehler. Vier- und Fünfjährige hingegen können diese Aufgabe im Normalfall lösen.

Um herauszufinden, ob dieses Problem vielleicht daran liegt, dass Dreijährige schlichtweg Probleme bei den verbalen Antworten haben, führten Sapp, Lee und Muir (2000) eine Studie durch. Sie fanden Belege dafür, dass die Dreijährigen, die Probleme mit der verbalen Lösung der Aufgaben hatten, die Aufgaben in einem nonverbalen Durchgang sehr wohl lösen konnten, eine richtige Repräsentation der Erscheinung des Gegenstandes hatten und auch seine wirklichen Eigenschaften richtig repräsentierten. Was sie jedoch auch herausfanden, war ein Reihenfolge-Effekt. Wurde die Aufgabe zuerst verbal gestellt und anschließend nonverbal, war die Leistung bei der nonverbalen Aufgabe schlechter als wenn die nonverbale vor der verbalen Aufgabe gestellt wurde. Auf die Leistung bei der verbalen Aufgabe hatte die Reihenfolge keinen Einfluss. Dies deutet darauf hin, dass Dreijährige zwar schon sehr wohl zwischen Erscheinung und Wirklichkeit unterscheiden

können, die Repräsentation sowohl der Erscheinung als auch die der wirklichen Eigenschaften eines Objektes allerdings nur sehr unsicher gespeichert und noch nicht artikulierbar sind.

Hansen und Markman (2005) wiederum führten die Probleme in diesem Bereich auf die Doppelbedeutung von „sieht aus wie“ zurück. Zum einen wird dieser Ausdruck zum Ausdruck einer äußeren Erscheinung verwendet, zum anderen verweist es auf etwas, das „wahrscheinlich so ist“. Deshalb müsste auch die Frage „Wonach sieht das aus?“ zu Problemen führen. In ihrer Studie änderten sie unter anderem die Frage nach der Erscheinung „Wonach sieht das aus?“ auf „Wonach sieht dieser Schwamm aus?“ und erzielten damit bei 90% der getesteten Dreijährigen richtige Antworten, im Vergleich zu 47% bei traditioneller Aufgabenstellung.

3.1.1.2 Entwicklung der theory of mind

Nach Lewis (2003) vollzieht sich die Entwicklung der theory of mind in den folgenden 4 Schritten:

1. *Ich weiß*: Diese Stufe reicht von der Geburt bis in die Mitte des 2. Lebensjahres. Sie basiert auf dem impliziten Bewusstsein, nicht aber auf dem expliziten, und beansprucht keine oder nur wenige sprachliche Fähigkeiten.
2. *Ich weiß, dass ich weiß*: Dies stellt eine Art Meta-Repräsentation (nicht zu verwechseln mit der Metarepräsentation nach Perner, 1991) dar und erlaubt es demnach, sich darüber Gedanken zu machen, was man selbst weiß. Nach Lewis (1990, zitiert nach Lewis, 2003) tritt diese Stufe etwa in der Mitte des zweiten Lebensjahres auf. Das explizite Bewusstsein ist hierbei schon notwendig, genauso wie Verhalten, das auf einen selbst bezogen ist.
3. *Ich weiß, dass du weißt*: Hier erreicht das Kind die Einsicht, dass es nicht nur selbst etwas weiß, sondern dass dies auch andere Personen wissen. Mit 2;6 Jahren ist diese Einsicht schon vorhanden. Allerdings finden in dieser Stufe noch sogenannte egozentrische Fehler statt, das Kind ist also der Meinung, dass andere dasselbe Wissen haben wie es selbst.

4. *Ich weiß, dass du weißt, dass ich weiß*: Diese Stufe ist diejenige, auf der sich im Normalfall Erwachsene befinden. Sie sind sich darüber im Klaren, dass jede Person eine eigene Perspektive hat und dass sich diese verschiedenen Perspektiven sehr wohl voneinander unterscheiden können.

In der Literatur finden sich aber auch zahlreiche andere Ansätze, die die Entwicklung der theory of mind darstellen. Dabei stellen die Theorie-Theorie, sowie die leistungs-basierten Zugänge die beiden Hauptansätze dar.

Die *Theorie-Theorie* besagt, dass das kindliche Denken in Form von Konzepten organisiert ist, welche aus intuitiven Theorien bestehen. Jede Veränderung der Konzepte und folglich des Denkens geschieht aufgrund einer Veränderung dieser Theorien. Das kindliche Verständnis des Denkens, welches ebenfalls in Form einer Theorie gespeichert ist, unterliegt dabei gerade im Alter von 3 bis 5 Jahren einer solchen konzeptuellen Veränderung (siehe Gopnik & Meltzoff, 1997; Perner, 1991; Wellman & Gelman, 1998).

Die *leistungs-basierten Zugänge* hingegen postulieren, dass die Entwicklung der theory of mind aufgrund allgemeiner kognitiver Veränderungen geschieht. Kinder, die jünger als 4 Jahre alt sind, haben deshalb Schwierigkeiten bei false-belief-Aufgaben, weil sie über bestimmte Fähigkeiten, die zur Lösung der Aufgaben benötigt werden, noch nicht verfügen (Leslie & Roth, 1993). Die Gruppe der leistungs-basierten Zugänge kann in zwei Untergruppen eingeteilt werden, nämlich die „*nativist modular theories*“, welche besagen, dass auch schon jüngere Kinder über metarepräsentationale Konzepte zu Überzeugungen verfügen, sie jedoch in false-belief-Aufgaben durch andere kognitive Faktoren behindert werden (Leslie & Roth, 1993), und die „*executive function theories*“, welche postulieren, dass im Alter von 4 Jahren konzeptuelle Veränderungen vonstattengehen, die durch exekutive Funktionen bedingt sind und welche die Fähigkeiten bezüglich der theory of mind beeinflussen. Verschiedene Studien dazu erkannten die Impulskontrolle (z.B. Carlson & Moses, 2001) oder das Gedächtnis (z.B. Carlson et al., 2002) als Faktoren, die Einfluss darauf haben, inwieweit Aufgaben zur theory of mind gelöst werden können.

Darüber hinaus existieren aber auch noch viele andere Theorien, die die Entwicklung der theory of mind erklären zu versuchen. Zum Beispiel befassen sich etwa Hale und Tager-Flusberg (2003) mit der Rolle, die die Sprache in der Entwicklung der theory of mind spielt.

Auf die Rolle der exekutiven Funktionen sowie die Rolle der Sprache wird später noch genauer eingegangen werden.

Generelle Informationen zum Ablauf der Entstehung von Wissen um Überzeugungen fanden Wellman und Liu (2004) in einer Meta-Analyse zur theory of mind heraus. Demnach dürften Kleinkinder früher in der Lage sein, Urteile über die Wünsche anderer Personen zu bilden, als über deren Überzeugungen. Erst wenn sie dies beherrschen, sind sie in der Lage, über die unterschiedlichen Überzeugungen zweier Personen Bescheid zu wissen, und später auch über Überzeugungen, die nicht mit der Wirklichkeit übereinstimmen, also falsche Überzeugungen (siehe auch Lillard & Flavell, 1992). Nicht-Wissen wird ebenfalls früher verstanden als falsche Überzeugungen es werden. Es muss dabei allerdings unterschieden werden zwischen Situationen, in welchen das jeweilige Kind weiß, was der Wahrheit entspricht, und solchen, in denen es die Wahrheit nicht kennt. Kennt es die Wahrheit nicht, ist es früher in der Lage zu erkennen, dass zwei Personen verschiedene Überzeugungen haben und im Anschluss daran, dass aus diesen verschiedenen Überzeugungen unterschiedliche Annahmen resultieren können. Erst später kann es dies auch in Situationen, in welchen es die Wahrheit kennt. Es ist also schwieriger für das Kind, in einer solchen Situation zu erkennen, dass die Überzeugung einer Person der Wahrheit entsprechen kann, die einer anderen jedoch nicht. Erst danach ist es auch in der Lage, zu beurteilen, dass die Handlung einer Person aus einer falschen Überzeugung heraus entstanden ist. Bartsch und Wellman (1995) sprechen in diesem Zusammenhang von 3 Phasen, die in der Entwicklung der theory of mind nacheinander auftreten: „desire psychology“ (bis zum Ende des 3. Lebensjahres), „desire-belief psychology“ (ab Beginn des 4. Lebensjahres) und „belief-desire psychology“ (ab Beginn des 5. Lebensjahres).

3.1.1.3 Vorläufer der theory of mind

Die theory of mind taucht nicht einfach zu einem bestimmten Zeitpunkt in der Entwicklung eines Kleinkindes auf und ist von da an vorhanden, sondern es spielen bei ihrer Entwicklung einige Vorläuferkomponenten eine Rolle, auf welchen die theory of mind aufbaut und welche zur Qualität ihrer Entwicklung beitragen. Im Folgenden sollen das Erkennen des eigenen Spiegelbildes, das Als-ob-Spiel, die Imitation und die Intention näher auf ihren Zusammenhang mit der Entwicklung der theory of mind beleuchtet werden.

3.1.1.3.1 Erkennen des eigenen Spiegelbildes

Nach Courage und Howe (2002) besteht wissenschaftlicher Konsens darüber, dass zwei Dimensionen des eigenen Selbst existieren: das subjektive, unreflektierte „ich“, als der wissende und denkende Handelnde (nach Lewis (2003) das implizite Bewusstsein, welches ab der Geburt vorhanden ist), und das objektive, reflexive „mich“, das aus einzigartigen Charakteristiken und Besonderheiten besteht, die das eigene Selbstkonzept ausmachen (nach Lewis (2003) das explizite Bewusstsein). Die Autoren fassen die Entwicklung der objektiven Komponente des Selbst, welche vermutlich im Laufe des 2. Lebensjahres vonstattengeht, folgendermaßen zusammen: Mit etwa 3 Monaten zeigen Kinder ihrem Spiegelbild gegenüber positive Reaktionen. Im Alter von etwa 8 Monaten erkennen sie, dass sich das Spiegelbild mit ihnen gemeinsam bewegt und nutzen dies für Spiel und Imitationshandlungen. Mit 18 Monaten sind sie schließlich in der Lage, das eigene Spiegelbild als solches zu erkennen. In diesem Alter beginnen Kinder auch, Verlegenheit gegenüber ihrem Spiegelbild zu zeigen, etwa durch schüchternes Lächeln, Blickvermeidung, oder indem sie ihr Spiegelbild berühren. Es werden also Emotionen gezeigt, die bewusst auf das eigene Selbst gerichtet sind (siehe auch Lewis, 2003, welcher von „exposure embarrassment“ spricht). Etwa im Alter von 22 bis 24 Monaten können Kinder ihr Spiegelbild korrekt benennen und sprachliche Selbstbeschreibung ist möglich. Schließlich sind sie in der Lage, Emotionen zu zeigen, die das

eigene Selbst bewerten, wie Schuld, Scham, oder Stolz. Lewis (2003) spricht in diesem Zusammenhang von „evaluative embarrassment“, welche etwa ab einem Alter von 2;6 Jahren auftritt.

Lewis und Brooks-Gunn (1979) führten Untersuchungen mit dem sogenannten “mark test” durch, bei welchem dem Kind ein Farb-Punkt ins Gesicht gemalt und seine Reaktion darauf vor einem Spiegel untersucht wird. Visuelles Erkennen des eigenen Selbst, definiert als diejenigen Verhaltensweisen, die als Reaktion auf den Farb-Punkt im Gesicht gezeigt werden, fand in diesen Untersuchungen bei Kindern im 1. Lebensjahr noch nicht statt. Im Alter von 15 Monaten trat eine Reaktion das erste Mal auf und bis ins Alter von 21 Monaten zeigten so gut wie alle Kinder ein solches Verhalten, was auch mit der Beschreibung von Courage und Howe (2002) überein stimmt.

Nach Perner (1991) sind zum Erkennen des eigenen Spiegelbildes sekundäre Repräsentationen notwendig. Außerdem ist es essentiell, dass mehrere Repräsentationen gleichzeitig vorhanden sein können, um zu verstehen, dass es man selbst ist, den man im Spiegel sieht. Erstens wird nämlich eine Repräsentation von sich selbst in der Realität benötigt und zweitens eine von sich selbst im Spiegel. Dieses gleichzeitige Vorhandensein von mehreren Repräsentationen erlaubt es, eine Interpretation vorzunehmen und diese beiden Repräsentationen miteinander abzugleichen. Das Kind versteht dadurch den Unterschied zwischen dem, was es im Spiegel sieht, und dem, was außerhalb des Spiegels ist.

3.1.1.3.2 Als-ob-Spiel

Das Als-ob-Spiel stellt jene Form des Spiels dar, in der ein Kind sich reale Objekte im Geiste als andere vorstellt (zum Beispiel eine Banane als Telefon), oder sich die Objekte (oder Handlungen) gänzlich vorstellt. Für Nielsen und Dissanayake (2004) bedeutet das, dass Kleinkinder im Als-ob-Spiel beginnen, sich von ihrer Fantasie anstatt vom Stimulus, mit dem sie sich beschäftigen, leiten zu lassen. Sie verwenden ihr Wissen um in der

Vorstellung Gegenstände, sowie deren Eigenschaften und Funktionen zu verändern.

Die Grundlage für das Als-ob-Spiel bildet nach Lewis und Ramsay (2004) das Vorhandensein von Metarepräsentationen über das eigene Selbst (nach Lewis, 2003, die 2. Stufe in der Entwicklung der theory of mind; siehe auch Leslie, 1987), denn das Kind muss wissen, dass das, was es tut, nicht der Wirklichkeit entspricht. Bei der Entwicklung des Als-ob-Spiels erfolgt laut Kavanaugh, Eizenman und Harris (1997) in der Mitte des zweiten Lebensjahres ein Übergang von auf sich selbst gerichtetem Als-ob-Spiel auf ein auf andere gerichtetes Als-ob-Spiel, oder, wie Lewis und Ramsay (2004) es beschreiben, ein als Übergang von Als-ob-Spiel, das nur auf sich selbst gerichtet ist, auf ein Als-ob-Spiel, das sowohl auf andere als auch noch immer auf sich selbst gerichtet ist.

Eine weitere Voraussetzung für das Als-ob-Spiel ist nach Lillard (1993) das Vorhandensein von mentalen Repräsentationen (nach Perner, 1991: primäre Repräsentationen). Ohne Als-ob-Handlungen kann es sehr wohl auskommen, da ein Kind etwa so tun kann, als wäre ein Besen ein Pferd, ohne dass es so tut, als würde es auf dem Besen reiten, aber ohne mentale Repräsentationen ist kein Als-ob-Spiel möglich, da eine Repräsentation eines Pferdes vorhanden sein muss, um sich den Besen als Pferd vorstellen zu können. Eine Studie wurde von Lillard (1993) durchgeführt, um herauszufinden, ob auch schon Kinder über dieses Verständnis verfügen. Es zeigte sich, dass 4 und 5 Jahre alte Kinder den Stellenwert mentaler Repräsentationen beim Als-ob-Spiel wohl noch nicht begriffen haben.

Nach dem Verständnis von Perner (1991) entspricht das Als-ob-Spiel insgesamt einer sekundären Repräsentation, da etwas im Geist abgebildet wird, das nicht der Realität entspricht.

Lewis und Ramsay (2004) untersuchten den Auftritt der objektiven Selbst-Komponente, also der Metarepräsentation (im Sinne von Lewis, 2003) des eigenen Selbst in der Entwicklung von Kleinkindern. Sie erhoben dazu die Verbindung zwischen visuellem Selbst-Erkennen, der Verwendung von Personalpronomina und dem Als-ob-Spiel, da die beiden letzteren Komponenten in früheren Studien schon als Indikatoren für die

Metarepräsentation des eigenen Selbst herausgefunden worden waren. 66 Kinder wurden im Alter von 15, 18, 21 und 24 Monaten untersucht. Diejenigen Kinder, die in der Lage waren, sich selbst zu erkennen, verwendeten signifikant häufiger Personalpronomina und zeigten ebenfalls signifikant häufiger Als-ob-Spielverhalten. Ein linearer Zusammenhang zwischen der Höhe der Ausprägung des Selbst-Erkennens bzw. der Verwendung von Personalpronomina bzw. des Umschwungs von auf sich selbst gerichtetem Als-ob-Spiels in eine Mischung aus auf sich selbst und auf andere gerichtetem Als-ob-Spiel und dem Alter war außerdem vorhanden. Durch den Zusammenhang dieser drei Komponenten schon vom ersten Messzeitpunkt an sehen die Autoren belegt, dass eine Metarepräsentation des Selbst etwa ab der Mitte des zweiten Lebensjahres vorhanden ist.

3.1.1.3.3 *Imitation*

Nielsen und Dissanayake (2004) unterscheiden zwei Arten von Imitation, die synchrone Imitation und die verzögerte Imitation.

Die *synchrone Imitation* beschreiben sie als eine typische Art von Spielverhalten, welche Mitte des zweiten Lebensjahres erstmals auftritt. Dabei verwenden Kinder für ihr Spiel spontan dieselben Objekte und spielen damit zur gleichen Zeit auf eine mehr oder weniger ähnliche Art und Weise (Asendorpf, 2002). Dazu müssen die Kleinkinder die Intentionen ihrer Spielpartner bezüglich des weiteren Spielverlaufs durchschauen können, wozu ein Zusammenspiel von primären Repräsentationen, also wie das Verhalten des Spielpartners wahrgenommen wird, und sekundären Repräsentationen, also Repräsentationen über die Gedanken oder Gefühle des Spielpartners, notwendig ist.

Die *Verzögerte Imitation* hingegen beschreibt die Fähigkeit eines Individuums, das Verhalten einer Person zum Zeitpunkt der Vorführung abzuspeichern und dieses nach einer gewissen Zeit wieder abzurufen.

In einer Studie fanden Nielsen und Dissanayake (2004) heraus, dass verzögerte Imitation ab einem Alter von etwa 12 Monaten auftritt. Synchrone Imitation, das Als-ob-Spiel und das Erkennen des eigenen Spiegelbildes,

hingegen ist sogar bis ins Alter von 15 Monaten nur sehr rudimentär vorhanden, was darauf schließen lässt, dass die verzögerte Imitation im Gegensatz zu den anderen drei Fähigkeiten nicht als Indikator für Entwicklung mentaler Fertigkeiten im zweiten Lebensjahr angesehen werden kann. Dennoch dürfte sie Auswirkungen auf die Entwicklung des Als-ob-Spiels haben. So fanden sich etwa Zusammenhänge zwischen der verzögerten Imitation im Alter von 15 Monaten und dem Auftreten des Als-ob-Spiels mit 18 und 21 Monaten. Weiters korrelierte die Anzahl der Items, die bei verzögerter Imitation nachgeahmt wurden, signifikant mit der Anzahl der Als-ob-Spiel-Items mit 21 und 24 Monaten. Für die Annahme, dass sich die synchrone Imitation, das Als-ob-Spiel und das Erkennen des eigenen Spiegelbildes in der Mitte des zweiten Lebensjahres weitgehend gemeinsam entwickeln, fanden sich nur gemischte Belege. Signifikante Zusammenhänge zwischen dem Erkennen des eigenen Spiegelbilds im Alter von 15 Monaten und dem Als-ob-Spiel im Alter von 18 Monaten, darüber hinaus zwischen der synchronen Imitation und dem Als-ob-Spiel mit 24 Monaten konnten als einzige entdeckt werden. Insgesamt ließ sich lediglich sagen, dass sobald eine dieser drei Fähigkeiten gezeigt wird, etwa 3 bis 6 Monate später alle drei Fähigkeiten gezeigt werden. Keine stellt dabei jedoch die Vorläuferfähigkeit einer anderen dar. Auch Asendorpf und Baudonnière (1993) beziehungsweise Asendorpf, Warkentin und Baudonnière (1996) fanden heraus, dass die synchrone Imitation eng mit dem Erkennen des eigenen Spiegelbildes verknüpft ist und führten dies auf das Vorhandensein der Fähigkeit zu sekundären Repräsentationen zurück.

Da die synchrone Imitation, das Als-ob-Spiel und das Erkennen des eigenen Spiegelbildes, wie schon zuvor erläutert, an die Fähigkeit zu sekundären Repräsentationen geknüpft sind, können sie mit dem späteren Auftreten einer theory of mind in Verbindung gebracht werden. Denn die für die theory of mind notwendige Fähigkeit zu Metarepräsentationen baut auf ebendiesen sekundären Repräsentationen auf (siehe Perner, 1991).

3.1.1.3.4 *Intention*

Der Begriff Intention ist eng mit dem Begriff Begehren verknüpft, beschreibt jedoch nicht dasselbe. Ein Ergebnis kann begehrt, aber gleichzeitig nicht intendiert werden, oder aber kann ein Ergebnis intendiert werden, das nicht begehrt wird (Feinfield, Lee, Flavell, Green & Flavell, 1999).

Searle (1983) unterscheidet zwei Arten von Intentionen, nämlich die vorausgehenden Intentionen und die Intentionen in der Situation.

Vorausgehende Intentionen finden statt, bevor eine Handlung ausgeführt wird, und können sprachlich als „Ich werde X machen“ beschrieben werden. Sie können, müssen aber nicht zwangsläufig eine Handlung nach sich ziehen.

Intentionen in der Situation sind Intentionen, die in dem Moment, in dem eine absichtliche Bewegung oder Handlung durchgeführt wird, eine Rolle spielen. Sie werden als „Ich mache gerade X“ beschrieben und können auch ohne eine vorausgehende Intention auftreten.

Jede absichtlich durchgeführte Bewegung beinhaltet eine Intention in der Situation, jedoch nur manche Intentionen in der Situation beinhalten auch eine vorausgehende Intention.

Feinfield et al., (1999) konnte aufzeigen, dass 4-jährige Kinder sowohl das Konzept der vorausgehenden Intention, als auch das der Intention in der Situation verstehen. 3-Jährige schnitten dabei um vieles schlechter ab. Die Autoren zogen aus den Daten den Schluss, dass Kinder erst etwa zwischen 3;6 und 4 Jahren zwischen vorausgehenden Intentionen und Intentionen in der Situation unterscheiden können. Aber auch wenn sie noch nicht explizit darüber Bescheid wissen, hilft das Wissen um eine Intention Kindern unter 3 Jahren sehr wohl schon beim Problemlösen und führt zu Unterschiede in der Imitation von Handlungen. So fanden Carpenter, Call und Tomasello (2002) heraus, dass 2-Jährige besser in der Lage sind, komplexe Handlungen zu imitieren und dadurch zu einer Problemlösung zu gelangen, wenn sie die vorausgehende Intention, die hinter dieser Handlung steht, kennen, als wenn sie die Intention nicht kennen. Meltzoff (1995) zeigten, dass 18 Monate alte Kinder, die eine Person sahen, die an der vollständigen Ausführung einer

Handlung scheiterte, die vollständige Handlung nachahmten. Sie imitierten also, was das Modell zu tun versucht hatte (die Intention), nicht das, was es tatsächlich tat. In dieser Hinsicht lässt sich sagen, dass Kinder in diesem Alter schon in dem Sinne Vorgänge interpretieren, dass sie zwischen dem, was eine Person tatsächlich tut, und einem tieferen Sinn, der hinter dieser Handlung steht, unterscheiden und daraus Handlungen ableiten (siehe auch Huang, Heyes & Charman, 2002). Gergely, Bekkering und Király (2002) beziehungsweise Carpenter, Call und Tomasello (2005) zeigten, dass sogar 14 beziehungsweise 12 Monate alte Kinder schon dazu in der Lage sind, Handlungen auf ihr Ziel hin zu analysieren und sie gemäß dem Ergebnis dieser Analyse zu imitieren. Gergely et al. (2002) konnten etwa zeigen, dass 14 Monate alte Kinder eine Handlung so nachahmen, wie es ihnen am vernünftigsten erscheint. Wenn die Modellperson einen Schalter mit dem Kopf betätigte, obwohl sie auch ihre Hände frei gehabt hätte, imitierten dies 96% der Kinder (für die Kinder muss die Person demnach einen bestimmten Grund dafür gehabt haben, den Kopf anstatt der Hände zu verwenden). Waren die Hände der Modellperson allerdings bedeckt, führten 79% der Kinder die Handlung selbst nicht mit dem Kopf, sondern mit den Händen aus (es muss für sie also logischer erschienen sein, selbst die Hände zu verwenden, da diese – anders als bei der Modellperson – frei waren).

Dass die Intention mit der theory of mind zusammenhängt, liegt auf der Hand. Das Verstehen von Intentionen ist für die theory of mind wichtig, da etwa bei Maxi und der Schokolade (Wimmer & Perner, 1983) das Wissen um eine falsche Überzeugung mit der Handlung, die darauf folgt, verknüpft wird. Nicht nur Maxis falsche Überzeugung, sondern auch die Intention, die er bezüglich des Verstecks der Schokolade hat, nämlich dass er diejenige Lade öffnet, in der sich die Schokolade befindet, muss verstanden werden um eine richtige Antwort zu geben.

3.1.1.4 Einflüsse auf die theory of mind

3.1.1.4.1 Exekutivfunktionen

Nun soll auf den Zusammenhang zwischen theory of mind und bestimmten Exekutivfunktionen näher eingegangen werden, was, wie weiter oben schon erwähnt, einen Ansatz der leistungs-basierten Zugänge zur theory of mind darstellt.

Carlson und Moses (2001) stellen übersichtlich dar, was zu den Exekutivfunktionen gezählt wird. Demnach umfassen diese “those processes that serve to monitor and control thought and action, including self-regulation, planning, behavior organization, cognitive flexibility, error detection and correction, response inhibition, and resistance to interference” (S.1032).

Carlson, Mandell und Williams (2004) fanden in ihrer Langzeitstudie heraus, dass Zusammenhänge zwischen exekutiven Funktionen und theory of mind im Alter von 24 Monaten noch nicht vorhanden sind. Jedoch sagt die Leistung in Aufgaben zu den exekutiven Funktionen die Leistung in theory-of-mind-Aufgaben ein Jahr später vorher, auch wenn Variablen wie Sprachfähigkeiten oder Bildung der Mutter konstant gehalten werden. Carlson und Moses (2001) zeigten, dass der Zusammenhang zwischen den exekutiven Funktionen und der theory of mind bei 3 und 4 Jahre alten Kindern vorhanden sind.

Die Impulskontrolle scheint dabei die bedeutendste Rolle zu spielen. Darunter versteht man eine Art Selbstkontrolle, die es ermöglicht, eine vorherrschende Reaktion of einen Reiz zurückzuhalten, um eine andere Reaktion zu ermöglichen. Ein Beispiel dafür stellt etwa die Studie von Zelazo, Frye & Rapus (1996) zum Sortieren von Karten dar. Im Gegensatz zu Vierjährigen waren Dreijährige nicht in der Lage, auf eine andere Regel zum Sortieren der Karten als der gerade verwendeten zu wechseln, obwohl sie sehr wohl über Wissen bezüglich dieser Regel verfügen. Diese Impulskontrolle könnte demnach auch dafür verantwortlich sein, dass Dreijährige Probleme beim Lösen von theory-of-mind-Aufgaben haben, vor

allem bei Aufgaben zur falschen Überzeugung, wenn verlangt wird, dass sie die vorherrschende Reaktion (zu sagen, wo z.B. die Schokolade wirklich versteckt ist) unterdrücken und stattdessen eine andere Antwort geben (wo Maxi die Schokolade wohl suchen wird). In die Meta-Analyse von Wellman, Cross und Watson (2001) wurden Studien miteinbezogen, die auf dies Rücksicht nahmen und in welchen, angelehnt an Maxi und die Schokolade, die Schokolade nicht von einem Kasten in den anderen transferiert, sondern sie beispielsweise gegessen wird, um diesen Konflikt zu vermeiden. Denn wenn keine Schokolade mehr vorhanden ist, sollte auch die vorherrschende Reaktion zur Preisgabe des wahren Verstecks der Schokolade ausgeschaltet sein. Die Leistungen der Kinder stiegen dadurch zwar wie vorausgesagt an, jedoch war dieser Anstieg nicht signifikant.

Viele Studien belegen dennoch einen starken Zusammenhang zwischen Impulskontrolle, welche sich meist aus Konflikt- (eine Handlung darf ausgeführt werden, eine andere jedoch nicht) sowie Aufschieb-Aufgaben (eine Handlung darf erst nach einer gewissen Zeitspanne ausgeführt werden) zusammensetzt, und theory of mind, beziehungsweise zwischen Impulskontrolle und den einzelnen false-belief-Aufgaben (siehe etwa Carlson, Moses & Claxton, 2004). Dieser Zusammenhang bleibt auch bestehen, wenn verschiedene Faktoren, wie etwa Alter, Geschlecht oder Sprachfertigkeiten (Carlson & Moses, 2001), ja sogar das Kurzzeitgedächtnis (Carlson et al., 2002;), konstant gehalten werden. Carlson, Moses & Claxton (2004) zeigen zwar die Möglichkeit auf, dass der Zusammenhang zwischen den exekutiven Funktionen und der theory of mind auch durch das Zusammenspiel anderer exekutiver Komponenten, wie hierarchisches Planen, gesteuert werden könnte, fanden in ihrer Studie jedoch keine Belege dafür. Die Impulskontrolle scheint somit einen hinreichenden Faktor für die Erklärung des Zusammenhanges zwischen exekutiven Funktionen und theory of mind darzustellen. Carlson et al. (2002) zum Beispiel fanden jedoch heraus, dass gerade die falsche Überzeugung am besten von Konflikt-Aufgaben, nicht aber von Aufschieb-Aufgaben vorhergesagt wird. Der Unterschied zwischen diesen beiden Arten von Aufgaben ist, dass die Konflikt-Aufgaben neben der Impulskontrolle auch Anforderungen an das Kurzzeitgedächtnis stellen. Dies stellt die Vermutung nahe, dass eine

Kombination von Impulskontrolle und Kurzzeitgedächtnis notwendig ist, um Aufgaben zur falschen Überzeugung lösen zu können. Diese Ergebnisse gelten allerdings nur für die false-belief-Aufgaben, nicht für Aufgaben zur Unterscheidung von Erscheinungsbild und Wirklichkeit.

3.1.1.4.2 Sprachfähigkeiten

Auch auf die Rolle der Sprache für die Entwicklung der theory of mind soll anschließend noch eingegangen werden.

Um die Auswirkung von Sprache auf die theory of mind zu untersuchen, werden nach Lohmann und Tomasello (2003) meist Korrelationsstudien oder Trainingsstudien verwendet. Das Problem bei Korrelationsstudien ist, dass sie Raum für viele Interpretationsmöglichkeiten lassen. Trainingsstudien hingegen sind in der Lage kausale Zusammenhänge aufzuzeigen.

In ihrer Trainingsstudie untersuchten die Autoren den Zusammenhang zwischen Sprache und Verständnis von falscher Überzeugung, wie auch die Beschaffenheit dieses Zusammenhanges. Es zeigte sich, dass sowohl durch die sprachliche Strukturierung der eigenen Erfahrungen durch andere, als auch durch Training des Satzbaus und der Satzergänzungen die Leistung von Kindern im Alter von 3;3 Jahren bis 3;10 Jahren in Aufgaben zur falschen Überzeugung verbessert werden konnten. Am besten wirkte sich ein Training aus, das beide Komponenten beinhaltet.

Hale und Tager-Flusberg (2003) führten ebenfalls eine Trainingsstudie durch um herauszufinden, ob das Training von Satzergänzungen einen Einfluss auf die theory of mind hat. Es stellte sich heraus, dass ein Satz-Ergänzungs-Training bei 36 bis 58 Monate alten Kindern die gleichen Verbesserungen bei false-belief-Aufgaben erzielt wie ein gezieltes theory-of-mind-Training.

Korrelationsstudien hingegen wurden etwa von Astington und Jenkins (1999) oder Farrar und Maag (2002) durchgeführt. Erstere fanden heraus, dass die Sprache bei 3;4 Jahre alten Kindern mit der theory-of-mind-Leistung 7 Monate später in Zusammenhang steht und dass vor allem – wie auch schon oben erwähnt – der Satzbau dabei eine große Rolle spielt. Farrar und Maag

(2002) konnten den Wortschatz und die grammatikalischen Fähigkeiten im Alter von 2 Jahren mit Leistungen bezüglich theory of mind im Alter von 4 Jahren in Verbindung bringen. Vorsichtig kann gesagt werden, dass frühe grammatikalische Fähigkeiten mit der theory of mind auch dann zusammenhängen, wenn der Effekt, den der Wortschatz hat, berücksichtigt wird.

Ruffman, Slade und Crowe (2002) untersuchten schließlich den Einfluss des Sprachverhaltens der Mütter, nämlich Äußerungen über mentale Zustände, auf die theory-of-mind-Leistung der Kinder. Es zeigte sich tatsächlich ein solcher Zusammenhang, sogar wenn Variablen wie die Sprachfertigkeiten des Kindes oder der sozioökonomische Status der Mutter konstant gehalten wurden.

3.1.2 Numerisches Wissen

3.1.2.1 Zählen

Das Erlernen des Zählens stellt einen langwierigen Prozess dar. Beginnen Kinder in einem Alter von 2 Jahren zu zählen, so brauchen sie doch etwa 4 Jahre, bis sie die Abfolge der Zahlen von 1 bis 100 beherrschen (Fuson & Hall, 1983). Dabei ist es den meisten 2-Jährigen recht bald möglich, das Zählen mit „1, 2, 3“ zu beginnen, wenngleich sie diese Abfolge in unterschiedlichster Art und Weise fortsetzen (Gelman & Gallistel, 1978). Sehr oft ist der darauffolgende Zahlenblock von Kind zu Kind unterschiedlich, jedoch auch in mehreren Durchgängen konsistent (zum Beispiel „5, 7, 4, 9“). Schließlich folgt ein Block von Zahlen, der auch in sich nicht mehr konsistent ist (etwa „10, 15, 12, 19, 14“) (Fuson, Richards & Briars, 1982). Die Studie von Fuson, Richards und Briars (1982) legt den Schluss nahe, dass im Alter von 3;6 bis 4;6 Jahren eine Zahlenfolge mit der höchsten richtigen Ziffer zwischen 10 und 14 beherrscht wird, von 4;6 bis 5 Jahren zwischen 14 und 20 und von 5 bis 6 Jahren mit der höchsten richtigen Ziffer zwischen 30 und 72.

Aunio und Niemivirta (2010) beschreiben nach Fuson (1988) und Van de Rijt (1996, zitiert nach Aunio & Niemivirta, 2010) 6 aufeinander folgenden Phasen, die das Erlernen des Zählens darstellen:

1. *Erstes Verstehen von Mengen*: Diese Phase tritt im Alter von 2 Jahren auf. Dabei können Kinder rudimentär zwischen Mengen unterscheiden und wissen, dass für unterschiedliche Mengen unterschiedliche Zahlwörter verwendet werden.
2. *Phase des akustischen Zählens*: Im Alter von etwa 3 Jahren haben Kinder schon verschiedene Zahlwörter gelernt, können diese aber noch nicht in der richtigen Reihenfolge aufsagen.
3. *Asynchrone Phase*: In dieser Phase, welche mit rund 4 Jahren beginnt, können Zahlwörter schon in der korrekten Reihenfolge aufgesagt und mit Objekten in Verbindung gebracht werden, indem beim Zählen auf die zu zählenden Objekte gezeigt wird. Jedoch stimmt das Zeigen mit den Zahlwörtern meist noch nicht überein.
4. *Synchrone Phase*: Diese Phase beginnt etwa im Alter von 4;6 Jahren. Hier stimmt das Zählen mit dem Zeigen auf Objekte schließlich überein.
5. *Resultative Phase des Zählens*: Mit rund 5 Jahren verstehen Kinder, dass jedes Objekt beim Zählen nur einmal mit einer Zahl bezeichnet wird und dass die zuletzt genannte Zahl der Menge der Objekte entspricht.
6. *Gekürzte Phase des Zählens*: In dieser Phase, im Alter von etwa 5;6 Jahren, ist es schließlich möglich, bei einer anderen Zahl als 1 mit dem Zählen zu beginnen.

Damit diese einzelnen Schritte vollzogen werden können, müssen aber bestimmte Regeln des Zählens gelernt, angewendet und eingehalten werden. In der Frage, warum und wie diese Regeln erlernt werden, gehen die Meinungen allerdings auseinander. Gelman und Gallistel (1978) etwa erachten die folgenden 5 Prinzipien, welche sie schon vor dem Erlernen des Zählens als implizit vorhanden sehen, als grundlegend für das Erlernen des Zählens:

1. Das *Prinzip der Eins-zu-Eins-Korrespondenz* (jedem zu zählenden Objekt wird genau ein Zahlwort zugeordnet)

2. Das *Prinzip der stabilen Reihenfolge* (die Zahlwörter haben immer die gleiche Reihenfolge)
3. Das *Kardinalzahlprinzip* (die beim Zählen zuletzt genannte Zahl entspricht der Anzahl der Objekte einer Menge)
4. Das *Prinzip der Irrelevanz der Reihenfolge* (die zu zählenden Items können in jeder beliebigen Reihenfolge gezählt werden)
5. Das *Prinzip der Abstraktion* (jede Menge diskreter Objekte ist zählbar)

Vor allem das Kardinalprinzip ist dabei entscheidend dafür, dass ein vollständiges Verständnis des Zählens von Objekten erlangt wird.

Dem gegenüber stehen Theorien, die besagen, dass Kinder das Zählen zuerst als Aktivität ohne besondere Bedeutung erlernen, so wie das Aufsagen eines Reims, bevor sie die Prinzipien, die dahinter stehen, verstehen und anwenden können (siehe etwa Briars & Siegler, 1984; Fuson & Hall, 1983). Betrachtet man die dazu vorhandene Literatur, so finden sich vermehrt Belege, die gegen das implizite Vorhandensein der 5 Prinzipien sprechen, und es wird gezeigt, dass diese von Kindern erst verstanden werden, wenn sie mit dem Erlernen des Zählens bereits begonnen haben (z.B. Frye, Braisby, Lowe, Maroudas & Nicholls, 1989; Wynn, 1990, 1992). Frye et al. (1989) fanden auch Variationen in der Entwicklung des Wissens über diese 5 Prinzipien. Studien, die einen Unterschied im Wissen um die Prinzipien je nach Fortschritt des numerischen Wissens oder je nach Alter aufzeigen, führten etwa Le Corre, Van de Walle, Brannon und Carey (2006) oder Wynn (1990) durch.

3.1.2.2 Erfassung von Mengen

Im Alter von 2 Jahren ist das Erfassen von Mengen durch eine große Variation gekennzeichnet. In der Studie von Gelman und Gallistel (1978) waren 50% aller teilnehmenden 2-Jährigen mit einem Altersdurchschnitt von 31 Monaten in der Lage eine Menge von 2 Items richtig zu benennen. Die anderen 50%, die dies nicht schafften, hatten ein durchschnittliches Alter von 27 Monaten. 25% aller Kinder konnten die Anzahl bei einer Menge von 3 Items wiedergeben, diese waren mindestens 32 Monate alt.

Neuere Ergebnisse zur Mengenerfassung bei 3-jährigen Kleinkindern liefern Condry und Spelke (2008). Obwohl diese die Bedeutung eines Zahlworts als Bezeichnung für eine Menge bei niedrigen Zahlwörtern schon verstehen, können sie dies bei größeren Zahlwörtern noch nicht. Die Unterscheidung zwischen einer Menge von 10 und einer Menge von 2 oder 1 ist möglich, jedoch nicht die Unterscheidung zwischen einer Menge von 10 und einer von 5. Obwohl die meisten Kinder in diesem Alter bis 10 zählen können, verfügen sie über kein Verständnis dafür, dass Wörter, die in der Zählabfolge später an der Reihe sind, auch größere Mengen repräsentierten. Ebenfalls dürften sie auch nicht verstehen, dass ein noch nicht bekanntes Zahlwort eine Menge auch dann noch bezeichnet, wenn die Objekte neu angeordnet werden.

Auch Wynn (1990) führte Studien zur Erfassung von Mengen bei Kindern zwischen 2;6 und 3;6 Jahren durch. Hierbei mussten die Kinder verschiedene Mengen von Objekten weitergeben. Je nach Antwortverhalten wurden die Kinder kategorisiert in Zähler und Greifer. Die Zähler zählten die Items beim Übergeben laut, gaben ein Item nach dem anderen, bis die richtige Anzahl erreicht wurde, oder zählten die Items nach und veränderten, wenn notwendig, deren Anzahl. Die Greifer hingegen gaben einfach eine Handvoll Items auf einmal, ohne sie zu zählen. Wenn sie aufgefordert wurden, nachzuprüfen, ob sie die richtige Anzahl von Items gegeben hatten, zählten nur die Zähler diese nach. Die Greifer wendeten unterschiedliche Strategien an. Sie benannten entweder einfach das letzte Item mit der gefragten Zahl (z.B. eins, zwei, sechs), leugneten (z.B. eins, zwei, drei; das sind sechs), veränderten die Anzahl der gegebenen Items, oder argumentierten einfach gar nicht (zu ähnlichen Ergebnissen kamen hierzu auch Le Corre et al., 2006). Die Autorin schloss aus den Ergebnissen, dass im Alter von 3 bis 3;6 Jahren ein Sprung von Greifer zu Zähler stattfindet, der mit dem Auftreten eines Verständnisses für das Kardinalprinzip einhergehen dürfte. Dennoch dürften die meisten Greifer zumindest auch schon ein Verständnis für die Menge von 1 besitzen. Bei mehr als einem Item können sie allerdings nicht einmal zwischen größeren und kleineren Mengen unterscheiden. Die Ergebnisse bestätigten weiters auch die zuvor dargestellten Ergebnisse von Condry und Spelke (2008), dass Kinder zuerst

über ein Verständnis für kleine Mengen verfügen, was sich später auf größere Mengen ausweitet.

Viele Autoren vertreten die Meinung, dass das Erfassen kleiner und großer Mengen durch unterschiedliche Mechanismen bedingt ist. Nach Watson, Maylor und Bruce (2007) liefern viele in der Vergangenheit durchgeführte Studien Belege dafür, dass kleine Mengen im Gegensatz zu größeren Mengen nicht gezählt werden müssen um die Anzahl ihrer Items zu erfassen. Mengen, die aus bis zu 4 Einheiten bestehen, können schnell und präzise erfasst werden, größere Mengen dagegen wesentlich langsamer. Kaufman, Lord, Reese und Volkmann (1949) nennen die Fähigkeit zur schnellen Erfassung kleiner Mengen „subitization“, also spontane Mengenerfassung. Andere Autoren jedoch (zum Beispiel Gelman & Gallistel, 1978, oder Silverman & Rose, 1980) halten dies nicht für stimmig. Die Hypothesen, dass Kinder kleine Mengen nicht zählen, sondern spontan erfassen, beziehungsweise sie erst zu zählen beginnen, wenn sie auch gelernt haben größere Mengen zu zählen, lehnen sie mit Berufung auf zahlreiche Studien kategorisch ab. Sie vertreten die Meinung, dass auch bei kleinen Mengen das Zählen gegenüber der spontanen Erfassung den Vorrang hat. Eine andere Theorie vertreten zum Beispiel Feigenson, Dehaene und Spelke (2004). Sie nehmen an, dass es zwei Kernsysteme gibt, welche einerseits eine annähernde Repräsentation von Mengen und andererseits eine genaue Repräsentation von einzelnen Objekten ermöglichen. Diese beiden Systeme werden je nach Mengengröße aktiv. Bei größeren Mengen schaltet sich das Kernsystem zur ungefähren Repräsentation ein, kleine Mengen werden durch das Kernsystem zur genauen Repräsentation exakt erfasst.

3.1.3 Schlussfolgerndes Denken

3.1.3.1 Induktives Denken

Beim induktiven Denken geht es darum, von Eigenschaften einzelner Kategoriemitglieder auf die Eigenschaften anderer Mitglieder derselben

Kategorie zu schließen, also vom Einzelfall auf die Allgemeinheit. Ihr Aufbau sieht folgendermaßen aus: *Prämissen* werden vorgegeben (Individuum X weist die Eigenschaft Y auf; Individuum X gehört der Kategorie Z an) und aus diesen erfolgt eine *Schlussfolgerung* (auch andere Mitglieder der Kategorie Z weisen die Eigenschaft Y auf). Da Induktionen also Informationen zu den Prämissen hinzufügen, die nicht auf ihren Wahrheitsgehalt überprüft werden können (alle Mitglieder einer Kategorie weisen dieselben Eigenschaften auf), sind sie bis zu einem gewissen Grad unsicher (Evans, 2010).

Das Ziehen von induktiven Schlussfolgerungen gestaltet sich umso einfacher, je ähnlicher sich die einzelnen Mitglieder untereinander schon in Bezug auf perzeptuelle, also äußerlich wahrnehmbare Merkmale sind. Bei natürlichen, aber auch bei künstlichen Kategorien, verbinden die Mitglieder jedoch nicht immer äußerliche Ähnlichkeiten, was die Gefahr birgt, sich beim induktiven Schließen von der äußeren Erscheinung fehlleiten zu lassen. Trotzdem sind 4-Jährige (Gelman & Markman, 1986; Gelman, Collman & Maccoby, 1986), aber auch schon 3-Jährige (Gelman & Markman, 1987), und sogar 2;6 Jahre alte Kinder (Gelman & Coley, 1990) in der Lage, bei der Generalisierung neuer Informationen über ein Individuum äußere Ähnlichkeiten zu ignorieren und stattdessen aufgrund der Kategoriezugehörigkeit auf Gleichheit zwischen diesem und anderen Individuen zu schließen. Kinder unter 3 Jahren sind dabei allerdings auf die zusätzliche Bereitstellung von sprachlichen Informationen angewiesen, da sie sonst eher dem Fehler unterliegen, die Zuschreibung von Ähnlichkeiten auf der äußeren Erscheinung aufzubauen (Gelman & Coley, 1990).

3.1.3.2 Deduktives Denken

Bei deduktiven Schlussfolgerungen wird ebenfalls aus Prämissen eine logische Antwort abgeleitet, jedoch erfolgt die Schlussfolgerung hier von der Allgemeinheit auf ein Individuum, daher stellt die Schlussfolgerung, sofern sie richtig gezogen wurde, die einzige wahre und gültige Antwort dar. Zur Prüfung dieser Art des Denkens werden meist Syllogismen verwendet, die folgenden Aufbau aufweisen: Zunächst wird eine *allgemeine Prämisse* dargeboten („Alle Schafe fahren mit dem Fahrrad.“), danach eine *auf einen*

Einzelfall bezogene Prämisse („Bill ist ein Schaf.“), und anschließend erfolgt die *Frage* („Fährt Bill mit dem Fahrrad?“). Aus den vorgegebenen Prämissen ist nur eine einzige *Schlussfolgerung* möglich („Ja, Bill fährt mit dem Fahrrad.“). Dabei spielt es keine Rolle, ob die Prämissen mit der Realität übereinstimmen oder nicht (Goswami, 1998/2001). Da deduktive Schlüsse also nur auf den Prämissen aufbauen und kein weiteres Wissen hinzugefügt wird, können sie mit einer größeren Sicherheit gezogen werden als induktive Schlüsse (Evans, 2010).

Das Problem beim deduktiven Denken von Kleinkindern ist jedoch oft, dass sie gerade mit denjenigen Prämissen Probleme haben, die ihr Wissen um die Realität widerlegen. Richards und Sanderson (1999) fanden heraus, dass 2-, 3- und 4-Jährige in der Lage waren, trotzdem logisch korrekte Schlussfolgerungen zu treffen, und zwar wenn sie ihre Fantasie verwendeten. Die Aufgaben waren in der oben schon dargestellten Form von Syllogismen dargeboten. Lediglich bei der gestellten Frage wurde eine zweite Alternative gegeben, die die tatsächliche Realität widerspiegelte (z.B. „Fährt Bill mit dem Fahrrad oder geht er zu Fuß?“). Ob die richtige oder die falsche Alternative zuerst dargeboten wurde, machte allerdings keinen Unterschied. Am besten gelang den Kindern das Treffen von korrekten Schlussfolgerungen in den Untersuchungsbedingungen, in denen sie entweder einfach dazu angehalten wurden, ihre Fantasie zu verwenden, oder sie sich einen Fantasieplaneten vorstellen sollten, auf dem das Erzählte geschieht, da somit die Dinge, die die Wirklichkeit widerlegten, ohne Probleme neben der Realität existieren konnten. In diesen Bedingungen gaben Kinder auch häufiger richtige theoretische Begründungen für ihre Antworten. Auch Harris, German und Mills (1996) fanden Belege dafür, dass 3-Jährige sehr wohl schon Schlussfolgerungen treffen können, wenn die Prämissen die Realität widerlegen, dies aber leichter ist, wenn es sich dabei um Aufgaben handelt, die die Realität der Vergangenheit widerlegen, als um hypothetische Aufgaben über die Zukunft. Die Kinder können beurteilen, ob der Realität widersprechende Prämissen ein Ergebnis verhindert hätten. Ebenso können sie zwischen solchen Prämissen unterscheiden, die das Ergebnis herbei geführt hätten, und solchen, die es nicht zugelassen hätten.

Eine andere Art, deduktives Denken zu erfassen, sind sogenannte Selektionsaufgaben, die nach dem Schema „wenn p, dann q“ aufgebaut sind (siehe Wason, 1966). Bei Kindern werden diese Selektionsaufgaben in Form von Erlaubnisregeln dargeboten und die Kinder müssen bewerten, ob eine Regel verletzt wird oder nicht. Harris und Núñez (1996) konnten zeigen, dass schon 3- und 4-Jährige Aufgaben zu solchen Erlaubnisregeln lösen können. Die Regeln (zum Beispiel: Wenn Sally draußen spielt, muss sie einen Mantel anziehen) wurden in eine Geschichte eingebaut und den Kindern wurden Fragen darüber gestellt, um zu sehen, ob sie daraus darauf schließen können, ob eine Situation (Sally spielt mit / ohne Mantel draußen) die Regel verletzt oder nicht. Beide Altersgruppen schnitten in dieser Aufgabe gut ab, auch wenn zusätzlich berücksichtigt werden musste, ob die in Frage stehende Handlung überhaupt ausgeführt wurde (Sally spielt draußen), oder nicht (Sally spielt im Haus). Kinder in diesem Alter dürften also schon wissen, dass eine Regel nur beachtet werden muss, wenn die in Frage stehende Handlung ausgeführt wird.

3.1.3.3 Zusammenhänge / Unterschiede zwischen induktivem und deduktivem Denken

Nach Rotello und Heit (2009) gibt es zwei Ansätze, die den Zusammenhang zwischen deduktivem und induktivem Denken beschreiben, den Aufgaben-Zugang und den Prozess-Zugang. Beim *Aufgaben-Zugang* geht es allein um die Aufgaben, bei welchen Induktionen oder Deduktionen angewendet werden, und die Begründung für die eine oder die andere Art der logischen Schlussfolgerung, da davon ausgegangen wird, dass sich die Aufgaben zum induktiven beziehungsweise zum deduktiven Denken voneinander unterscheiden. Die dahinterstehenden kognitiven Prozesse sind dabei nicht von Interesse. Beim *Prozess-Zugang* hingegen steht die Frage im Zentrum, von welchen kognitiven Prozessen Induktionen beziehungsweise Deduktionen gesteuert werden und welche Ähnlichkeiten oder Unterschiede es dabei gibt. Der Prozess-Zugang kann unterteilt werden in einen Ein-Prozess-Zugang und einen Zwei-Prozess-Zugang. Aus Sichtweise des *Ein-Prozess-Zuganges* bauen Induktionen und Deduktion auf denselben

kognitiven Mechanismen auf. Im Gegensatz dazu wird beim *Zwei-Prozess-Zugang* (siehe etwa Sloman, 1996) davon ausgegangen, dass sowohl heuristische als auch analytische Prozesse auf das schlussfolgernde Denken einwirken, jedoch auf unterschiedliche Art und Weise, je nachdem, ob es sich um eine Induktion oder eine Deduktion handelt. Induktionen werden nach diesem Ansatz durch schnelle heuristische Prozesse gesteuert, wobei es darum geht, Schlussfolgerungen aufgrund von Informationen über Kontext und Ähnlichkeit zu ziehen. Deduktionen werden durch langsamere analytische Prozesse beeinflusst, was zu abwägenderen, meist genaueren Schlussfolgerungen führt.

Heit und Rotello (2010) beziehungsweise Rips (2001) führten Studien zum Vergleich von Ein-Prozess- und Zwei-Prozess-Theorien durch. Die Ergebnisse beider Studien sprachen gegen den Ein-Prozess-Zugang und für die Gültigkeit des Zwei-Prozess-Zugangs. Induktives und deduktives Denken dürften demnach durch verschiedene kognitive Prozesse gesteuert werden.

3.1.3.4 Analoges Denken

Beim analogen Denken geht es darum, sich frühere Probleme und deren erfolgreiche Lösung ins Gedächtnis zu rufen, um ein neues Problem auf dieselbe Weise zu lösen. Dazu muss auf die Gemeinsamkeiten des aktuellen Problems mit dem früheren Problem geachtet werden, da dies die Grundlage für einen erfolgreichen Transfer des Lösungsweges auf das neue Problem darstellt. Konkret muss der Zusammenhang zwischen den am früheren Problem beteiligten Objekten dem zwischen den am jetzigen Problem beteiligten Objekten entsprechen (Goswami, 1998/2001). Die meisten Aufgaben zum analogen Denken sind im „a zu b ist wie c zu d“ - Format gestaltet, wobei a, b und c bereits vorgegeben sind und ein passendes d eingesetzt werden muss.

Nach Goswami und Brown (1990) sind schon 4-Jährige in der Lage, solche Aufgaben zu lösen, aber auch im Alter von 3 Jahren verfügen Kinder bereits über die Fähigkeit zum analogen Denken, wenn sie die Relationen, die in den Aufgaben vorkommen, verstehen, was eher der Fall ist, wenn diese

kausaler Natur sind (zum Beispiel: Apfel zu geschnittenem Apfel ist wie Knetmasse zu geschnittener Knetmasse; Goswami & Brown, 1989). Generell gilt wohl: Je vertrauter die Kinder mit den verwendeten Relationen sind, desto einfacher sind die Analogien zu lösen. In diesem Zusammenhang zeigten Singer-Freeman und Goswami (2001), dass die vorhergehende Erfahrung mit dem Untersuchungsmaterial eine große Rolle dabei spielt, wie gut Aufgaben zu den Analogien gelöst werden. Für Kinder unter 3 Jahren dürfte diese Art von Aufgaben hingegen um Vieles schwieriger zu lösen sein, da sie zu abstrakt ist (Goswami, 1998/2001).

3.2 Visuelle Wahrnehmung

3.2.1 Objekterkennung

Das Erkennen eines Objektes gliedert sich in 2 Schritte auf: die Objektwahrnehmung und die Mustererkennung (Anderson, 2005/2007).

3.2.1.1 Objektwahrnehmung

Folgende Gestaltprinzipien, wie sie von verschiedenen Gestaltpsychologen formuliert wurden, bestimmen die Wahrnehmung von Szenen. Sie sind dafür da, die Szene in Objekte zu gliedern:

Prinzip der Nähe: Nahe beieinander liegende Einheiten werden als zusammengehörig wahrgenommen.

Prinzip der Ähnlichkeit: Ähnlich aussehende Objekte werden zu Gruppen zusammengefasst.

Prinzip des glatten Verlaufs: Eine Linie wird eher als Linie betrachtet, wenn sie einen glatten Verlauf hat, als wenn sie abgelenkt ist.

Prinzip der Geschlossenheit / der guten Gestalt: Objekte werden als solche wahrgenommen, auch wenn sie teilweise verdeckt sind.

3.2.1.2 Mustererkennung

Die Mustererkennung dient der Identifikation der Objekte, in die eine wahrgenommene Szene zuvor zergliedert wurde. Eine Theorie dazu beruht auf der Annahme eines *Schablonenabgleichs*. Dabei wird jedes Bild eines Objektes mit allen im Gedächtnis gespeicherten Mustern (Schablonen) verglichen, um zu einer Übereinstimmung und einer Erkennung des Objektes zu gelangen.

Da die Annahme einer solchen Objektidentifikation allerdings auch mit Schwierigkeiten verbunden ist, wurde eine andere Form der Objekterkennung vorgeschlagen, die *Merkmalsanalyse*. Hierbei wird ein Reiz als eine Kombination unterschiedlicher Grund-Merkmale (zum Beispiel Striche) gesehen, bestehend aus einer bestimmten Anzahl dieser Merkmale, die in einer bestimmten Position relativ zueinander stehen. Zuerst werden die einzelnen Merkmale erkannt, anschließend die Kombination aus diesen.

Trotzdem ist es nur schwer vorstellbar, wie komplexere Objekte nur durch einige wenige Merkmale erkannt werden können. Biederman (1987) schlägt in seiner *Theorie der komponentialen Erkennung* daher eine Objekterkennung in 3 Stufen vor:

1. Ein wahrzunehmendes Objekt wird in Teilobjekte zergliedert.
2. Jedes Teilobjekt kann einer Kategorie zugeordnet werden. Biederman (1987) spricht von 36 grundlegenden Kategorien, den sogenannten Geons (geometric icons). Das Erkennen eines Geons ist vergleichbar mit dem Erkennen eines Buchstabens.
3. Aus diesen Geons lassen sich alle denkbaren Objekte aufbauen, so wie aus Buchstaben alle denkbaren Wörter aufgebaut werden können. Somit wird ein Objekt als das Muster erkannt, das aus diesen bestimmten Geons zusammengesetzt ist.

Aber auch der Kontext spielt bei der Wahrnehmung von Objekten eine große Rolle. Biederman, Mezzanotte und Rabinowitz (1982) definierten 5 Regeln, welchen Bildern allgemein unterliegen und bei deren Verletzung das Wahrnehmen eines Bildes erschwert wird:

1. *Unterlagenregel*: Gegenstände liegen auf einer Unterlage auf.
2. *Verdeckungsregel*: Normalerweise verdecken Gegenstände ihren Hintergrund.
3. *Wahrscheinlichkeitsregel*: Gegenstände sind nur in gewissen Kontexten anzutreffen.
4. *Ortsregel*: Ein in einem Kontext wahrscheinlicher Gegenstand tritt dort nur an bestimmten Orten auf.
5. *Größenregel*: Bekannte Gegenstände werden nur in bestimmten Größen vorgefunden.

Je mehr Regeln verletzt werden, desto höher sind die Reaktionszeit und die Fehlerrate bei der Wahrnehmung des entsprechenden Gegenstandes. Die Wahrnehmung der ganzen Szene wird davon allerdings nicht beeinträchtigt.

3.2.2 Farbwahrnehmung / -differenzierung

Nach Boothe (2002) wird bei der Wahrnehmung von Farben allgemein zwischen drei Dimensionen unterschieden: Farbton, Sättigung und Helligkeit. Der *Farbton* ist das, was im normalen Sprachgebrauch als „Farbe“ wahrgenommen wird. Ein Mensch kann rund 200 Farbtöne unterscheiden, je nach Wellenlänge wird eine andere Farbe erkannt. Die *Sättigung* bestimmt, wie „rein“ eine Farbe wahrgenommen wird, und die *Helligkeit* gibt schließlich an, wie hell oder dunkel ein bestimmter Farbton ist.

Im 1. Lebensjahr ist die Farbe rot diejenige, die allen anderen vorgezogen wird. Auch benannt wird rot gemeinsam mit gelb als erste Farbe, blau und grün können erst um einiges später unterschieden werden. Mit 3 bis 4 Jahren kann ein Kind alle wichtigen Farben benennen, nur grün und blau werden noch öfters verwechselt (Holle, 1984/2000).

Schon Johnson (1977) fand heraus, dass vor dem Alter von 3;6 Jahren ein großer Sprung in der Entwicklung der Farbwahrnehmung stattfindet, wobei sich die Leistung der Jungen im Vergleich zu der von Mädchen langsamer entwickelt. Rot, blau, gelb und grün sind die ersten Farben, die von Kindern benannt werden können. Dabei können 38% der Kinder im Alter von 30 bis

33 Monaten die vier Grundfarben richtig benennen, mit 34 bis 37 Monaten sind 50% dazu in der Lage und im Alter von 38 bis 42 Monaten 56%. Hingegen können dies schon 71% der 42 bis 45 Monate alten Kinder, 72% der Kinder im Alter von 46 bis 49 Monaten und 79% der 50 bis 53 Monate alten Kinder.

3.2.3 Formwahrnehmung / -differenzierung

Nach Bertenthal, Campos und Haith (1980) beginnt sich die Fähigkeit zur Formwahrnehmung zwischen 5 und 7 Monaten zu entwickeln. Unter gewissen Bedingungen sind jedoch auch schon 4 Monate alte Kinder zur Wahrnehmung von Formen in der Lage (Kavsek, 2002). Im Alter von einem Jahr ist es einem Kind nach Holle (1984/2000) möglich, einen Gegenstand zu erkennen, auch wenn er aus unterschiedlichen Perspektiven, von verschiedenen Abständen aus, oder vor verschiedenen Hintergründen zu sehen ist (Formkonstanz). Ist es 2 Jahre alt, ist das Kind in der Lage, geometrische Formen in entsprechende Lücken zu platzieren. Bis 2;6 Jahre beginnt es, sich Gegenstände auch vorstellen zu können und es kann außerdem schon große Bausteine zu Paaren sortieren. Eine Einschränkung, die in diesem Alter allerdings gegeben ist, ist, dass ein Kind einen Gegenstand noch nicht als Ganzes wahrnehmen kann, sondern immer nur eine Eigenschaft auf einmal.

Gruber (1992) führte in einer Vorarbeit zum WET (Kastner-Koller & Deimann, 2002) eine Studie zur visuellen Wahrnehmung 3- bis 6-jähriger Kinder durch. Bezüglich der Formwahrnehmung kam er zu dem Ergebnis, dass das Differenzieren zwischen verschiedenen Formen mit 3 Jahren bereits beherrscht wird, obwohl er bei seiner Literaturrecherche herausgefunden hatte, dass sich diese Fähigkeit erst zwischen 3 und 5 Jahren entwickeln würde. Daher wird angenommen, dass die Entwicklung der Fähigkeit zur Formdifferenzierung genau in dem hier interessierenden Altersbereich von 2;0 bis 2;11 Jahren relevant ist, nicht zuletzt, da die empirische Erhebung der Formdifferenzierung in der vorliegenden Studie an die Arbeit von Gruber (1992) angelehnt ist.

3.2.4 Größenwahrnehmung / -differenzierung

Was die Größendifferenzierung betrifft, fand Gruber (1992) ähnlich wie bei der Formdifferenzierung heraus, dass sich die Fähigkeit zur Unterscheidung verschiedener Größen eines Gegenstandes ab dem Alter von 3 Jahren herausbildet. In seiner empirischen Studie zeigte sich jedoch, dass die Aufgaben selbst für 3-Jährige schon zu leicht waren. Daher wird hier davon ausgegangen, dass die Unterscheidung verschiedener Größen desselben Objektes eine für die hier interessierende Altersgruppe passende Aufgabe darstellt.

3.3 Gedächtnis

Die gängigste Unterteilung des Gedächtnisses ist wohl die in ein Ultrakurzzeit- (oder sensorisches), ein Kurzzeit- und ein Langzeitgedächtnis. Das sensorische Gedächtnis ist für die erste Wahrnehmung eines Stimulus zuständig, das Kurzzeitgedächtnis beinhaltet Informationen, die den momentanen Aufmerksamkeitsfokus betreffen, und das Langzeitgedächtnis umfasst all jene Informationen, die aus dem Kurzzeitgedächtnis dorthin transferiert und für längere Zeit behalten werden. Diese Informationen können über Jahre hinweg gespeichert werden und sind über das Kurzzeitgedächtnis immer wieder ins Bewusstsein rücktransportierbar. Im Folgenden werden diese drei Speicher genauer vorgestellt.

3.3.1 Ultrakurzzeitgedächtnis

Das Ultrakurzzeitgedächtnis umfasst die Zeitspanne von ein paar hundert Millisekunden und dient der ersten Registrierung eines Stimulus, damit dieser vom Kurzzeitgedächtnis weiterverarbeitet werden kann.

Zur Erfassung des visuellen sensorischen Gedächtnisses dienen nach Anderson (2005/2007) meist Untersuchungen, bei welchen für ungefähr 50

Millisekunden ein Reiz (etwa ein paar Reihen mit jeweils einigen Buchstaben) in das Blickfeld der Testperson projiziert wird. Die Testperson muss anschließend angeben, wie viele Items sie sich gemerkt hat. Meist sind dies 3 bis 6 Buchstaben. Sperling (1960) veränderte diese Aufgaben, indem er den Probanden nach jeder Darbietung einen Hinweis gab, welche Buchstabenreihe abgefragt werden würde. So konnten meist alle oder fast alle Buchstaben aus einer Viererreihe wiedergegeben werden.

Um das phonologische sensorische Gedächtnis zu erfassen, werden die Probanden meist mit phonologischen Stimuli beschallt, welche sie anschließend wiedergeben müssen.

3.3.2 Kurzzeit- bzw. Arbeitsgedächtnis

Die Speicherkapazität des Kurzzeitgedächtnisses ist begrenzt, sie umfasst durchschnittlich 7 Einheiten. Miller (1956) führte dazu den Begriff „Chunks“ ein. Ein „Chunk“ kann einer Einheit entsprechen, aber durch sogenanntes „Chunking“ können auch mehrere einzelne Einheiten zu einem „Chunk“ zusammengefasst werden und nehmen dadurch weniger Speicherplatz im Kurzzeitgedächtnis ein. Denn es ist nach Miller nicht die Anzahl der Einheiten, die die Beschränkung des Kurzzeitgedächtnisses bestimmt, sondern die Anzahl der „Chunks“. Ein „Chunk“ kann somit einen einzelnen Buchstaben repräsentieren, eine Silbe, einen Wortteil, oder gar ein ganzes Wort oder eine Phrase.

Als Alternative zum Kurzzeitgedächtnis schufen Baddeley und Hitch (1974, zitiert nach Baddeley, 2000) den Terminus „Arbeitsgedächtnis“. Sie unterteilten das Arbeitsgedächtnis in 3 Einheiten: die phonologische Schleife, den visuell-räumlichen Notizblock und die zentrale Exekutive. Später fügte Baddeley diesem Modell noch den episodischen Puffer hinzu (Baddeley, 2000). Die phonologische Schleife dient dabei der temporären Speicherung von akustischer Information, der visuell-räumliche Notizblock der kurzfristigen Speicherung von visuellen und räumlichen Informationen und der episodische Puffer ist dazu da, komplexe Informationen sowohl aus den beiden Speichern, als auch aus dem Langzeitgedächtnis zu integrieren.

Er ist somit eine Art Schnittpunkt der phonologischen Schleife, dem visuell-räumlichen Notizblock und dem Langzeitgedächtnis. Die zentrale Exekutive stellt die Kontrollinstanz dar, die auch in der Lage ist, verschiedenartige Informationen in „Episoden“ zusammenzufassen um so die Speicherung verschiedenartiger Informationen im episodischen Puffer zu erleichtern (Baddeley, 2000).

Zur Erfassung der Kapazität des Kurzzeitgedächtnisses wird der Testperson üblicherweise eine Reihe von Ziffern vorgesagt, die sie reproduzieren muss. Es wird gemessen, wie viele Ziffern die Person in ihrem Kurzzeitgedächtnis zu behalten imstande ist. Nach Dempster (1981) bezeichnet die Gedächtnisspanne „the maximum length of series that can be reproduced 50% of the time“ (S.65). Im Alter von 2 Jahren beträgt diese Gedächtnisspanne 2 Ziffern, mit 3 Jahren 3 Ziffern, und mit etwa 5 Jahren 4 Ziffern, bis sie bis zum Erwachsenenalter schließlich auf 7 Ziffern ansteigt. Bei Wörtern ist die Spanne geringer. Zwar schaffen es sowohl Zwei- als auch Dreijährige, 3 Wörter im Gedächtnis zu behalten, und Fünfjährige kommen auf 4 Wörter, jedoch steigt die durchschnittliche Gedächtnisspanne im Erwachsenenalter nur auf 5 Wörter.

Um das Kurzzeitgedächtnis bei vorsprachlichen Kindern zu erfassen, eignet sich die oben genannte Methode klarerweise nur bedingt. Meist wird bei Kleinkindern ein sogenannter delayed-response task angewendet, also eine Aufgabenstellung mit verzögerter Lösung der Aufgabe, wie ihn Hunter (1913, zitiert nach Reznick, 2009) erstmals einsetzte. Dabei wird vor den Augen des Kindes ein Gegenstand versteckt, das Kind kurzzeitig abgelenkt wird und es soll diesen Gegenstand schließlich suchen.

3.3.3 Langzeitgedächtnis

Zur Unterteilung des Langzeitgedächtnisses existieren verschiedene Ansätze. Sehr oft erfolgt in der Literatur eine Trennung des Gedächtnisses in eine „deklarative“ und eine „nicht deklarative“ Komponente (siehe Cohen & Squire, 1980, zitiert nach Squire, 1992), beziehungsweise in eine „explizite“

und eine „implizite“ Komponente (siehe Graf & Schacter, 1985; Schacter, 1987). Der Begriff „explizit“ ist dabei gleichbedeutend mit „deklarativ“ und „implizit“ entspricht dem Begriff „nicht deklarativ“.

Das *deklarative* oder *explizite* Gedächtnis, welches schnell und flexibel ist, beinhaltet alle bewussten Gedächtnisleistungen und bezieht sich auf das Wissen von Fakten und Ereignissen, die in der Vergangenheit liegen. Unterteilt werden können die Inhalte des deklarativen Gedächtnisses in episodisches und semantisches Wissen (Tulving, 1972, war der erste, der diese Unterscheidung traf). Das *episodische* Wissen umfasst das autobiographische Wissen, also das, was man selbst erlebt hat, das *semantische* Wissen hingegen beinhaltet Weltwissen sowie das Wissen um Fakten, Regeln und Konzepte.

Das *implizite* oder *nicht deklarative* Gedächtnis hingegen beschreibt alle Gedächtnisleistungen, die nicht durch das Bewusstsein gesteuert sind, wie etwa die Anwendung einmal erlernter Fertigkeiten, oder Gewohnheiten (Graf & Schacter, 1985). Es läuft unbewusst und ist weniger flexibel (Squire, 1994). Je nachdem, wie der Abruf aus dem expliziten Gedächtnis erfolgt, unterscheidet Squire (1987) zwischen „recognition“, also dem Wiedererkennen und „recall“, dem freien Erinnern. Der Unterschied zwischen diesen beiden Formen des Abrufs liegt in der Menge, mit der Wahrnehmungsunterstützungen verwendet werden, um eine innere Repräsentation des zu Erinnernden hervorzurufen. Beim Wiedererkennen sind Hinweisreize vorhanden, die ein Erinnern ermöglichen. Beim freien Erinnern hingegen gibt es eine solche Unterstützung nicht. Hier muss die innere Repräsentation des zu Erinnernden aus der Vergangenheit hervorgerufen werden um sie ins Bewusstsein zu bringen.

Nach Bauer (2002) galt lange Zeit die Meinung, dass die Fähigkeit zum freien Erinnern etwa im Alter von drei Jahren beginnt. Hinweise darauf lieferten die Tatsache, dass die sogenannte infantile Amnesie, also die Zeit in der Kindheit, an die wir uns als Erwachsene nicht erinnern können, mit etwa 3 Jahren und 6 Monaten endet und dass es Kindern bis 3 Jahren schwer fällt, Auskunft über die Vergangenheit zu geben. Das Problem war, dass lange Zeit kein Instrument vorhanden war, um die freie Erinnerungsleistung

auch bei vorsprachlichen Kindern zu erheben. Als aktuelle Methode der Wahl beschreibt Bauer (2002) hierzu die Imitation. Mit dieser Methode ließ sich zeigen, dass die Fähigkeit zum freien Erinnern schon mit neun Monaten besteht, dass sich diese Art von Erhebung vor allem ab dem 2. Lebensjahr als verlässliche Methode zur Gedächtnisüberprüfung erweist (Bauer, 2002). Meltzoff (1988) war der erste, der dies für die Erfassung des Gedächtnisses bei Kleinkindern einsetzte. Unterschieden wird hierbei zwischen der verzögerten und der hervorgerufenen Imitation.

Bei der *verzögerten Imitation* wird vom Untersuchungsleiter eine Handlung durchgeführt, zu der bestimmte, für das Kind neuartige, Objekte verwendet werden, und das Kind soll diese Handlung nach einer gewissen Zeitspanne imitieren. Der Vorteil dieser Art der Gedächtniserfassung liegt darin, dass vorab kein Training notwendig ist, da Kinder von sich aus die Handlungen anderer imitieren (z.B. Meltzoff, 1988). Außerdem sind keine Sprachfertigkeiten nötig, was auch die Anwendung an vorsprachlichen Kindern, beziehungsweise solchen, die gerade zu sprechen begonnen haben, erlaubt (Hayne & Simcock, 2009).

Bei der *hervorgerufenen Imitation* handelt es sich um eine Abwandlung der verzögerten Imitation. Auch hierbei zeigt der Untersuchungsleiter eine Handlung mit noch nicht bekannten Objekten vor, welche das Kind später imitieren muss. Jedoch beschreibt der Untersuchungsleiter dabei verbal seine Handlung und was er damit bezweckt, das Kind hat weiters die Möglichkeit, vor dem Testdurchgang zu üben, und außerdem werden verbale Hinweise gegeben (Hayne & Simcock, 2009). Gerade diese verbalen Hinweise sind es, die die Erinnerungsleistung deutlich steigern dürften (Hayne & Herbert, 2004).

Studien belegen, dass im Alter von 20 Monaten etwa 70% der Kinder eine freie Erinnerungsleistung durch Imitation noch nach 12 Monaten schaffen (Bauer, Wenner, Dropik, Wewerka & Howe, 2000). Generell gilt, je älter ein Kind ist, desto länger können die Intervalle sein, nach welchen das Kind die Handlung noch immer vollständig imitieren kann (Bauer 2004). Auch die Flexibilität des Abrufens, also das Anwenden früherer Erfahrungen auf neue Situationen, wird, wie Hayne und Simcock (2009) zusammenfassen, mit dem Alter immer besser.

3.4 Arbeitshaltungen

3.4.1 Aufmerksamkeit

Der Terminus „Aufmerksamkeit“ beschreibt nach Wagner-Menghin (2003) die Fähigkeit, eine Handlung präzise und beständig auszuführen, wobei Dingen, die für diese Handlung nicht wichtig sind, keine Beachtung geschenkt wird.

Im Allgemeinen wird meist eine Unterscheidung zwischen fokussierter und geteilter Aufmerksamkeit getroffen. Unter *fokussierter* (oder auch *selektiver*) *Aufmerksamkeit* wird diejenige Art von Aufmerksamkeit verstanden, bei welcher ein einziger Stimulus oder eine einzige Handlung im Zentrum des Interesses steht und alles Übrige mehr oder weniger effektiv ausgeblendet wird. *Geteilte Aufmerksamkeit* hingegen beschreibt eine Art von Aufmerksamkeit, bei der mehrere Stimuli oder Handlungen gleichzeitig im Fokus stehen.

Neumann (1992) erachtet 5 Komponenten als wichtig für die Aufmerksamkeit:

1. *Verhaltenshemmung*: Diese ist dazu da, eine von zwei unvereinbaren Handlungen zu hemmen, da die beiden nicht gleichzeitig ausgeführt werden können.
2. *Regulation des psychophysiologischen Erregungsniveaus*
3. *Informationsselektion zur Handlungssteuerung*: Nachdem entschieden ist, dass eine Handlung ausgeführt wird, wird durch die Selektion der Information spezifiziert, wie diese ausgeführt ist.
4. *Handlungsplanung*: Mit dieser werden Handlungen koordiniert. Sie ermöglicht es, Operationen ohne vorherige Übung in einer bestimmten Abfolge auszuführen.
5. *Hemmung beim Einsatz von Fertigkeiten*: Durch diese Hemmung wird verhindert, dass eine Fähigkeit gleichzeitig in verschiedenen Handlungen eingesetzt wird.

Zur Entstehung und Entwicklung der Aufmerksamkeit im Kleinkindalter existieren derzeit nur wenige, vor allem aktuelle, Untersuchungen, wenngleich es nach Ruff und Capozzoli (2002) sehr wohl Gründe dafür gibt, anzunehmen, dass eine Veränderung der Aufmerksamkeit vom Kleinkind- bis ins Vorschulalter stattfindet. Zum einen schreitet die Gewöhnung an Ablenkungen von außen voran, was bewirken müsste, dass diese mit zunehmendem Alter nicht mehr so schnell in das Aufmerksamkeitsfeld der Kinder rücken. Andererseits dürfte auch die Entwicklung der Impulskontrolle, die eine mit dem Alter immer stärker werdende Ausprägung aufweist (Kochanska, Murray & Harlan, 2000), dazu beitragen, dass sich die Kinder immer weniger ablenken lassen, je älter sie werden, und mit der Aufmerksamkeit bei der momentan ausgeführten Handlung oder dem Stimulus, der sich gerade im Aufmerksamkeitsfeld befindet, bleiben.

Diejenigen Studien, die sich mit dem hier in Frage stehenden Altersbereich befassten, legen den Schluss nahe, dass die Aufmerksamkeit, wie vermutet, mit zunehmendem Alter immer besser wird. Etwa fanden im Jahr 1990 Ruff und Lawson in einer Studie an Kindern im Alter von 1, 2 und 3;6 Jahren heraus, dass die fokussierte Aufmerksamkeit, die von Kindern im freien Spiel gezeigt wird, mit dem Alter zunimmt. Auch was die Art der Aufmerksamkeit beziehungsweise der Unaufmerksamkeit betrifft, liefern die Ergebnisse Hinweise darauf, dass sich diese verändert, je älter ein Kind wird. Auf diesen Ergebnissen bauten Ruff und Capozzoli (2002) ihre Studie auf. Sie untersuchten Kinder im Alter von 10, 26 und 42 Monaten und konnten die Ergebnisse von Ruff und Lawson (1990) replizieren. Denn auch hier veränderte sich die Aufmerksamkeit über die drei Alterszeitpunkte hinweg. Konkret nahm die Dauer der beiläufigen Aufmerksamkeit zwischen 26 und 42 Monaten ab, die beständige Aufmerksamkeit stieg immer weiter an, und auch die fokussierte Aufmerksamkeit steigerte sich zwischen 26 und 42 Monaten erheblich. Die Autorinnen nehmen an, dass die Entwicklung selbstregulativer Fähigkeiten der Grund dafür ist, dass sich die Ablenkbarkeit in dieser Zeitspanne immer mehr vermindert, je älter ein Kind wird, und die Fähigkeit zur Aufmerksamkeit immer weiter zunimmt.

Unter Selbstregulation wird die Fähigkeit verstanden, das eigene Verhalten insofern an eine Situation anzupassen, dass es den kognitiven, emotionalen

und sozialen Anforderungen dieser gerecht wird (Posner & Rothbart, 2000). Dabei spielt nach Posner und Rothbart (1998) vor allem Interaktion von Erregung und Hemmung eine wichtige Rolle. Im 2. Lebensjahr ist das Ziel der aufmerksamkeitsregulierenden Prozesse, fokussierte Aufmerksamkeit und zielgerichtetes Verhalten immer länger aufrecht erhalten zu können (Edelman, 2004). Dies bedeutet auch, dass mit steigender Regulationsfähigkeit die Ausdauer eines Kindes bei Aufgaben immer weiter zunimmt (Deater-Deckard, Petrill, Thompson & DeThorne, 2005). Ebenso wird es ab dieser Zeit möglich, die Aufmerksamkeit hierarchisch nach Wichtigkeit zu organisieren (Posner & Rothbart, 1998; 2000).

Bei der Ausbildung der Aufmerksamkeitsregulierung werden nacheinander 3 Aufmerksamkeitssysteme aktiv. Im ersten Lebensjahr eines Kindes beginnt sich das *retikuläre Aktivierungssystem* zu entwickeln, das es ermöglicht, die Aufmerksamkeit auf bestimmte Aspekte der Umwelt zu fokussieren und vor Ablenkungen zu schützen (Derryberry & Rothbart, 1997). Am Ende des ersten Lebensjahres bildet sich das *posteriore Aufmerksamkeitssystem* heraus, durch welches die Aufmerksamkeit von einem Ort an einen anderen gelenkt werden kann, indem die Aufmerksamkeit auf etwas gerichtet und davon wieder gelöst wird (Rothbart, Derryberry & Posner, 1994). Schließlich entwickelt sich im Kleinkindalter (Derryberry & Rothbart, 1997) das *anteriore Aufmerksamkeitssystem*, welches ein ausführendes System ist, das sensorische Informationen reguliert (Rothbart, Derryberry & Posner, 1994). Dieses System wird bewusst aktiviert und unterliegt willentlicher Kontrolle (Posner & Rothbart, 1992).

3.4.2 Leistungsmotivation

Folgt man der Definition von Heckhausen (1974), wird von leistungsmotiviertem Handeln dann gesprochen, wenn die Handlungen einer Person, sowie deren Ergebnisse durch einen Tüchtigkeitsmaßstab von der Person selbst als Erfolge oder Misserfolge klassifiziert werden.

Als Vorläufer des leistungsmotivierten Handelns wird die sogenannte Effekt- oder Wirksamkeitsmotivation gesehen (z.B. Barrett, 1998; Heckhausen,

1985). Im Kleinkindalter orientiert sich ein Kind noch daran, ob als Folge einer selbst durchgeführten Handlung ein – folglich selbst verursachter – Effekt eintritt oder nicht (Wirksamkeitsmotivation). Ungefähr im Übergang vom zum Vorschulalter erfolgt allerdings ein Wechsel weg davon zu einer selbstständigen Bewertung dieses Effekts (Leistungsmotivation). Reagiert ein wirksamkeitsmotiviertes Kind auf Erfolge noch mit Freude und auf Misserfolge mit Ärger, so ändern sich diese Reaktionen bei vorhandener Leistungsmotivation in Stolz beziehungsweise Scham um.

Holodynski (1992) entwickelte ein Phasenmodell, das die Entstehung des leistungsmotivierten Handelns verdeutlicht. Der Übergang zu einer Bewertung der eigenen Leistung wird nach diesem Modell aufgrund von sozialen Bewertungen durch die Bezugsperson des Kindes ermöglicht. Im Vergleich zu anderen vorhandenen Modellen zur Leistungsmotivation meint Holodynski, dass leistungsmotiviertes Handeln im Vorschulalter noch davon abhängig ist, ob eine erwachsene Person anwesend ist oder nicht. Die Ergebnisse aus der Studie von Holodynski (2006) sprechen für diese Annahme.

3.5 Zusammenfassung

Dieser Abschnitt befasste sich mit denjenigen Entwicklungsbereichen, auf die im empirischen Teil der Arbeit das Augenmerk gelegt wird: die kognitive Entwicklung, die visuelle Wahrnehmung, das Gedächtnis, sowie die Arbeitshaltungen. Unter den Bereich der kognitiven Entwicklung wurden die Bereiche „theory of mind“, „numerisches Wissen“ und „schlussfolgerndes Denken“ gefasst.

Bei der theory of mind wurde auf deren Erfassung, nämlich mittels Aufgaben zur falschen Überzeugung oder mittels Aufgaben zur Unterscheidung von Erscheinungsbild und Wirklichkeit, eingegangen. Da die theory of mind selbst aber in dem hier interessierenden Altersbereich noch nicht ausgereift ist, wurde außerdem auf ihre Vorläuferfähigkeiten, das Erkennen des eigenen Spiegelbildes, das Als-ob-Spiel, die Imitation und die Intention, eingegangen.

Auch die Einflüsse, die etwa die Exekutivfunktionen oder die Sprachfähigkeiten auf die theory of mind haben, wurden dargestellt.

Bezüglich des numerischen Wissens wurde die Entwicklung des Zählens und der Mengenerfassung geschildert. Kinder beginnen mit 2 Jahren zu zählen, im selben Alter dürfte sich die Fähigkeit zum Erfassen von kleinen Mengen herausbilden.

Zum schlussfolgernden Denken wurden dessen drei Kategorien, das induktive, das deduktive und das analoge Denken, beschrieben, deren Unterschiede und Zusammenhänge erläutert, sowie deren Erhebungsmethoden dargestellt.

Was die visuelle Wahrnehmung betrifft, wurde zuerst erklärt, wie die allgemeine Objekterkennung abläuft, nämlich über Objektwahrnehmung und Mustererkennung. Kurz wurde auch auf den Einfluss, den der Kontext auf die Wahrnehmung eines Objektes hat, eingegangen. Danach wurde die Wahrnehmung beziehungsweise Differenzierung von Formen, Größen und Farben beschrieben.

Hinsichtlich des Gedächtnisses wurde die Unterteilung in Ultrakurzzeit-, Kurzzeit- und Langzeitgedächtnis vorgestellt, sowie die Möglichkeiten um diese einzelnen Komponenten zu erheben. Auch eine Möglichkeit, das Langzeitgedächtnis vorsprachlicher Kinder zu erheben, nämlich mittels Imitation, wurde dargestellt.

Was die Arbeitshaltungen betrifft, wurde auf die Aufmerksamkeit und die Leistungsmotivation eingegangen. Es zeigte sich, dass beide Komponenten bei Kleinkindern schwierig zu erheben sind, da bei diesen die fokussierte Aufmerksamkeit noch nicht sehr ausgeprägt ist und auch die Leistungsmotivation sich erst herausbildet.

II. EMPIRISCHER TEIL

4 Zielsetzung und Fragestellungen

4.1 Zielsetzungen

Die Literaturrecherche zur Entwicklungsdiagnostik 2-jähriger Kinder brachte das Ergebnis, dass aktuell kein Verfahren existiert, das theoretisch und empirisch fundiert zur Diagnostik in diesem Altersbereich herangezogen werden kann. Zusätzlich brachte die Recherche über die spielbasierte Diagnostik die Erkenntnis, dass die Verbindung von spielbasierten Elementen mit denen der klassischen Diagnostik eine gute Voraussetzung für eine möglichst standardisierte und vergleichbare, gleichzeitig aber auch kindgerechte Erhebung der Fähigkeiten eines Kindes darstellt.

Das Ziel dieser Untersuchung ist es nun, Aufgaben zur Erfassung der allgemeinen Entwicklung im Kleinkindalter zu erproben, die den klassischen und den spielbasierten diagnostischen Zugang miteinander verbinden. Ein spielerischer Aufbau der gesamten Testsituation wurde gewählt, da im Zentrum des Interesses vor allem der Gedanke stand, für das zu untersuchende Kind eine Umgebung zu schaffen, die frei von Druck ist, sodass der Prozentsatz der verweigten Items so gering wie möglich gehalten wird, da sich das Kind in der Situation möglichst wohl fühlen und ihm das „Spielen“ Spaß machen soll. Durch den Einbau von Testitems in die Spielsituation, die standardisiert erfasst werden, soll die allgemeine Entwicklung des jeweiligen Kindes erhoben und mit anderen Kindern vergleichbar gemacht werden. Das Kind kann jedoch während der gesamten diagnostischen Untersuchung frei entscheiden, was es tun möchte und auch Vorschläge des Testleiters / der Testleiterin ablehnen.

Die ersten Schritte dieses Vorhabens wurden bereits umgesetzt. Theoretisch wurden im ersten Teil dieser Arbeit, sowie bei Kuchler (in Druck), verschiedene Entwicklungsbereiche auf ihre Relevanz im 3. Lebensjahr hin beleuchtet und daraus Aufgaben abgeleitet um diese empirisch zu erfassen. Dieser Itempool erfasst die Sprache, die Motorik, die kognitive Entwicklung, die visuelle Wahrnehmung, das Gedächtnis, die Arbeitshaltungen und die sozial-emotionale Entwicklung (zu Sprache, Motorik und sozial-emotionaler Entwicklung siehe Kuchler, in Druck).

Ziel ist es herauszufinden, ob ein spielbasierter Zugang praktikabel und eine Anwendung bei Kindern von 24 bis 36 Monaten überhaupt möglich ist.

An 20 Kindern wurden die Items bereits vorgegeben, die daraus gewonnenen Ergebnisse sollen im Folgenden dargestellt werden.

4.2 Fragestellungen

Die vorrangige Fragestellung dieser Studie lautet: *Ist die Vorgabe des Itempools zur Erfassung der kognitiven Entwicklung, der visuellen Wahrnehmung, des Gedächtnisses und der Arbeitshaltungen von Kindern im Alter von 2;0 bis 2;11 Jahren geeignet und praktikabel?*

Dabei steht im Zentrum des Interesses, ob die einzelnen Items jedes Entwicklungsbereiches geeignet sind, um die jeweiligen Entwicklungsbereiche zu erfassen. Die weiteren Fragestellungen lauten also:

- *Sind die Items des Bereichs „kognitive Entwicklung“ geeignet, um die kognitive Entwicklung bei Kindern im Alter von 2;0 bis 2;11 Jahren zu erfassen?*
- *Sind die Items des Bereichs „visuelle Wahrnehmung“ geeignet, um die visuelle Wahrnehmung bei Kindern im Alter von 2;0 bis 2;11 Jahren zu erfassen?*
- *Sind die Items des Bereichs „Gedächtnis“ geeignet, um das Gedächtnis bei Kindern im Alter von 2;0 bis 2;11 Jahren zu erfassen?*
- *Sind die Items des Bereichs „Arbeitshaltungen“ geeignet, um die Arbeitshaltungen bei Kindern im Alter von 2;0 bis 2;11 Jahren zu erfassen?*

5 Methode

5.1 Beschreibung des Itempools

Zur Entwicklung der Items wurden aus der Theorie diejenigen Bereiche herausgesucht, welche im Altersbereich von 2;0 bis 2;11 Jahren von Bedeutung sind, und welche sich auch als empirisch gut zu erheben erwiesen. Diese sind:

- Sprache (rezeptive und expressive Sprache)
- Motorik (Grob- und Feinmotorik)

-
- kognitive Entwicklung (theory of mind, numerisches Wissen, schlussfolgerndes Denken)
 - Gedächtnis (phonologisches und visuelles Gedächtnis)
 - visuelle Wahrnehmung (Farb-, Form-, und Größendifferenzierung)
 - Arbeitshaltungen (Aufmerksamkeit)
 - sozial-emotionale Entwicklung

Diese Arbeit spezialisiert sich dabei auf die Bereiche kognitive Entwicklung, Gedächtnis, visuelle Wahrnehmung und Arbeitshaltungen. Zu Motorik, Sprache und sozial-emotionaler Entwicklung siehe Kuchler (in Druck).

Das Testmaterial, das zur Erhebung dieser Fähigkeiten verwendet wird, besteht aus einer Puppenküche mit Kochutensilien (Töpfe, Pfannen, Besteck, Teller, Becher, usw.), einem Kaufmannsladen mit verschiedenem Zubehör (einem Einkaufswagen, verschiedenfarbigem Obst und Gemüse, Produktpäckchen) und einer Spielebox. Diese beinhaltet Papier und Farbstifte, Bausteine, ein Memory-Spiel bestehend aus 4 Paaren, ein Bilderset, einen großen und einen kleinen Ball, sowie 3 Bücher. Das erste Buch, das Buch zum deduktiven Denken (dD-Buch), ist ein selbst entworfenes und gestaltetes Buch aus Moosgummi (siehe Kuchler, Sapper, Deimann & Kastner-Koller, 2011). Beim zweiten Buch, dem Theory-of-Mind-Buch (ToM-Buch), handelt es sich um eine adaptierte Form von „In unserem Haus gibt's viel zu sehen“ (Capdevila, 1984/1993; siehe Kuchler et al., 2011). Das letzte Buch, das Magnetbuch („Was gehört wohin?: Mein großes Spielbuch mit 22 Magnetteilen“, Eberhard, 2007) wurde in seiner ursprünglichen Form eingesetzt (siehe Kuchler et al., 2011). Schließlich gibt es noch eine Reihe von Päckchen mit Bildern von geometrischen Figuren (siehe Kuchler et al., 2011), die in das Kaufmannsladen- und Puppenküche-Spiel mit einbezogen werden. Außerdem ist am Boden des Testraumes mit Klebeband ein 2 Meter langer Streifen geklebt.

Aus dem Bereich der kognitiven Entwicklung werden das numerische Wissen und ein Teil der theory of mind, nämlich das Als-ob-Spiel, im Spiel mit Puppenküche und Kaufmannsladen erfasst. Der zweite Teil der theory of mind, das Hineinversetzen in andere, wird über das ToM-Buch abgefragt, das deduktive Denken über das eigens gestaltete dD-Buch. Zur visuellen Wahrnehmung wird die Farbdifferenzierung einerseits in der Spielsituation mit Puppenküche und

Kaufmannsladen erfasst, andererseits über das Zeichnen mit den Buntstiften. Die Form- und Größendifferenzierung wird mithilfe der Päckchen mit den geometrischen Formen abgefragt. Schließlich wird das visuelle Gedächtnis mit dem Memory-Set und den Kärtchen erhoben, das phonologische Gedächtnis in der Spielsituation. Die Aufmerksamkeit wird über die 3 Bilderbücher, das Zeichnen und das Bauen mit den Bausteinen quantifiziert.

Für eine Beschreibung der Erhebung der Sprachfertigkeiten, der Motorik und der sozial-emotionalen Entwicklung siehe Kuchler (in Druck).

Im Folgenden wird auf die hier relevanten Entwicklungsbereiche mit den jeweiligen Items genauer eingegangen.

5.1.1 Kognitive Entwicklung

5.1.1.1 Theory of mind

Da sich die theory of mind selbst erst ab einem Alter von etwa 3 Jahren zu entwickeln beginnt (siehe zum Beispiel Carlson, Moses & Breton, 2002; Carlson, Moses & Claxton, 2004; Wellman, Cross & Watson, 2001), werden lediglich einzelne Vorläuferkomponenten der theory of mind erhoben. Die Entscheidung fiel aufgrund der vorhandenen Möglichkeiten einer empirischen Erhebung auf das Als-ob-Spiel, sowie auf die Fähigkeit, Urteile über Wünsche anderer Personen zu bilden, indem man sich in andere hineinversetzt. Bezüglich des Vorhandenseins des Als-ob-Spiels wird durch Beobachtung in der Spielsituation in der Puppenküche eingeschätzt, wie weit dieses bereits entwickelt ist. Zu den einzelnen Ausprägungskategorien siehe Kuchler, et al. (2011). Für das Bilden von Urteilen über Wünsche anderer werden die Kinder dazu angehalten, die Handlungen von in einem Bilderbuch abgebildeten Personen vorherzusagen. Die einzelnen Items sind in Kuchler et al. (2011) dargestellt.

5.1.1.2 Numerisches Wissen

Zur Erhebung des numerischen Wissens wird die Spielsituation in der Puppenküche herangezogen. Der erste Punkt, der erhoben wird, betrifft das Erfassen von Mengen. Zum einen wird erhoben, wie groß die maximale Anzahl von Objekten ist, die ein Kind richtig benennen kann, da dies bis zum Alter von 3 Jahren bei kleinen Mengen schon möglich sein sollte (Condry & Spelke, 2008; Wynn, 1990). Andererseits wird auch das implizite Wissen über Mengen erfasst, also ob das Kind schon eine geforderte Anzahl an Objekten geben kann. Weiters wird hierbei darauf geachtet, wie das Kind die Objekte hergibt, also ob es eine Handvoll gibt, oder eines nach dem anderen, und ob es dabei zählt. Für eine genaue Beschreibung aller Items siehe Kuchler et al. (2011).

5.1.1.3 Schlussfolgerndes Denken

Zur Untersuchung des schlussfolgernden Denkens wurde die Überprüfung des deduktiven Denkens ausgewählt. Die Art der Überprüfung dieser Fähigkeit ist an die Untersuchung von Richards und Sanderson (1999) angelehnt, da diese herausfanden, dass mithilfe von Fantasie-Aufgaben schon 2- bis 4-Jährige in der Lage sind, deduktive Schlüsse zu ziehen. Anhand des selbst gestalteten dD-Buches werden in Form von Syllogismen verschiedene Items zum deduktiven Denken abgefragt. Zu den einzelnen Items siehe Kuchler et al. (2011).

5.1.2 Visuelle Wahrnehmung

Im hier interessierenden Altersbereich sollte sich vor allem das Wissen um die vier Grundfarben entwickeln (Holle, 1984/2000; Johnson, 1977). Um den Erhebungsbereich der Farbdifferenzierung jedoch etwas weiter zu fassen, werden folgende 7 Farben sowohl aktiv als auch passiv abgefragt: rot, grün, blau, gelb, orange, lila und weiß.

Für die Erhebung, welche Farben ein Kind aktiv schon beherrscht, werden in der Puppenküchensituation verschiedene Obstsorten in den Farben rot, gelb, grün und orange abgefragt. Da die Möglichkeiten durch das vorhandene Spielzeug-

Obst allerdings relativ eingeschränkt sind, wird das Zeichnen ebenfalls dazu genutzt, das Wissen um die Farben zu erheben.

Um auch das passive Beherrschen dieser Farben zu überprüfen, werden die Aufgaben zur Form- und zur Größendifferenzierung verwendet. Die hierbei benötigten Päckchen sind unter Tüchern in verschiedenen Farben versteckt, zu welchen das Kind bei jedem Item hingehen muss. Außerdem wird wieder die Zeichensituation dazu benutzt, die restlichen Farben abzufragen.

Zusätzlich ist von Interesse, ob das Farbkonzept eines Kindes schon so stabil ist, dass verschiedene Objekte einer Farbe beim Sortieren als zusammengehörig gesehen werden, auch wenn es sich um verschiedene Abstufungen ein- und derselben Farbe handelt (zum Beispiel hellgrün und dunkelgrün). Dazu wird das verschiedenfarbige Spielzeug-Obst aus der Puppenküchensituation verwendet, wobei das Kind verschiedenfarbige Obst- und Gemüseutensilien in Körbe sortieren soll. Eine genaue Beschreibung aller Items findet sich in Kuchler et al. (2011).

Die Fähigkeit zur Formdifferenzierung wird in Anlehnung an Gruber (1992) erhoben, der in einer Vorarbeit zum WET (Kastner-Koller & Deimann, 2002) eine Studie zur visuellen Wahrnehmung 3- bis 6-jähriger Kinder durchführte. Er kam zu dem Ergebnis, dass die Aufgaben, die er seinen Testpersonen stellte, selbst für die Altersgruppe der 3-Jährigen zu leicht waren. Aufgrund dessen wird angenommen, dass sie in dem hier relevanten Altersbereich von 2;0 bis 2;11 Jahren differenzieren und daher wurden sie in nur leicht abgeänderter Form übernommen. Bei den Items geht es einerseits um die Differenzierung verschiedener geometrischer Formen (Kreis, Quadrat, Dreieck, Rechteck), andererseits um die Differenzierung verschieden aussehender Vierecke. Eine genauere Beschreibung findet sich in Kuchler et al. (2011).

Auch die Erhebung der Größendifferenzierung ist an Gruber (1992) angelehnt. Denn ähnlich wie bei der Formdifferenzierung zeigte sich in seiner empirischen Studie, dass die Aufgaben selbst für 3-Jährige zu leicht waren. Daher wird davon ausgegangen, dass die Unterscheidung von Größen, wie sie bei Gruber (1992) durchgeführt wurde, für die hier interessierende Altersgruppe eine passende Aufgabe darstellt, weshalb sie ebenfalls in abgeänderter Form übernommen

wurde. Hier geht es um die Differenzierung verschieden großer Rechtecke beziehungsweise Dreiecke. Auch hierzu siehe Kuchler et al. (2011).

5.1.3 Gedächtnis

Beim phonologischen Gedächtnis wird auf die Ergebnisse von Dempster (1981) aufgebaut, nach welchen die Ziffernspanne für Zweijährige etwa bei 2 Ziffern und die Wortspanne bei 3 Wörtern liegt. Obwohl die Leistungsdifferenzierung bei Ziffern besser möglich ist als bei Wörtern (während mit 3 Jahren schon 3 Ziffern und mit 5 Jahren 4 Ziffern gemerkt werden können, ändert sich die Wortspanne kaum und steigt erst mit 5 Jahren auf 4 Wörter an), konnte keine passende Möglichkeit gefunden werden, das Wiederholen einer Zahlenfolge in die Spielsituation mit der Puppenküche einzubauen. Daher wird die Wortspanne der Kinder in Form von „Gibst du mir bitte ein/eine A (z.B. Banane) und ein/eine B (z.B. Apfel)?“ abgefragt.

Um das visuelle Gedächtnis zu erheben, wird, ähnlich wie bei den delayed-response tasks (siehe Reznick, 2009) mit einem Bilderset sowie Memory-Karten überprüft, ob das Kind ein kurz zuvor gesehenes Bild wieder erkennen beziehungsweise wieder finden kann. Sowohl zu den Items des phonologischen als auch zu den Items des visuellen Gedächtnisses siehe Kuchler et al. (2011).

5.1.4 Arbeitshaltungen

Bezüglich der Arbeitshaltungen wird lediglich die fokussierte Aufmerksamkeit des Kindes beobachtet. Eine Überprüfung der Leistungsmotivation wäre nicht sinnvoll, da diese im Kleinkindalter noch nicht entwickelt ist (Barett, 1998; Heckhausen, 1985). Dazu werden einzelne der zuvor vorgestellten Aufgaben herangezogen (Bauen mit den Bausteinen, Zeichnen, Durchführung des Magnetbuches, des ToM-Buches und des dD-Buches) und es wird notiert, wie viele Minuten lang das Kind bei der Aufgabe aufmerksam ist, wie vielen Seiten dies bei einem Bilderbuch entspricht, ob sich das Kind wieder zur Aufgabe zurückführen lässt, wenn es während der Bearbeitung unaufmerksam wird, und wie viele Items das Kind vollständig aufmerksam bearbeitet. Eine genaue Beschreibung ist in Kuchler et al. (2011) zu finden.

5.2 Stichprobe

Die Stichprobe, welche durch ein Schneeballsystem zustande kam, bestand aus 20 Kindern, 11 (55%) davon weiblich, 9 (45%) männlich. Das Durchschnittsalter lag bei 29,1 Monaten, mit einer Spannweite von 24 bis 35 Monaten. Unter den teilnehmenden Kindern befanden sich 2 Frühgeburten, jedoch gaben deren Bezugspersonen an, dass bei den Kindern keine Entwicklungsauffälligkeiten vorliegen würden. Von den übrigen Kindern gab eine Mutter an, dass bei ihrem Sohn ein Entwicklungsvorsprung gegenüber Gleichaltrigen bestehe, jedoch sei dies nur ihre persönliche Einschätzung. 2 der teilnehmenden Kinder wachsen zweisprachig auf. Ein 2;10 Jahre alter Junge spricht finnisch und deutsch. Ein 2;11 Jahre altes Mädchen tschechisch und deutsch. Bei dem Mädchen kam es teilweise vor, dass sie tschechisch antwortete. Hier übersetzte die Mutter ins Deutsche. Eine Aufgabe galt bei diesem Mädchen als gelöst, wenn die Antwort richtig war, egal ob sie in der deutschen oder in der tschechischen Sprache gegeben wurde.

5.3 Durchführung der Untersuchung

Die Testungen wurden von April 2010 bis Oktober 2010 im Kleinkindertestraum der Test- und Beratungsstelle des Instituts für Entwicklungspsychologie und Psychologische Diagnostik durchgeführt.

Zuerst wurde der Itempool an 3 Kindern erprobt, um zu überprüfen, ob die Durchführung der Items auch so funktionierte, wie sie in der Theorie angedacht war. Aufgrund der Erfahrungen, die bei der Durchführung an diesen 3 Kindern gemacht wurden, wurden noch einige Dinge adaptiert und verändert, so wurden beispielsweise die Päckchen zur Form- und Größendifferenzierung umgestaltet. Nach diesen Änderungen blieb der Itempool jedoch in seiner nun aktuellen Form bestehen und wurde während der Durchführung an den restlichen 17 Kindern nicht mehr verändert.

Bei 2 Testungen waren beide Elternteile anwesend, ansonsten war ausschließlich die Mutter dabei. Die Instruktion an die Bezugsperson lautete, dass sich diese im

Hintergrund halten sollte, jedoch jederzeit mitspielen könne, wenn das Kind danach verlange. Allerdings sollte sie hierbei keine Handlungen aktiv initiieren, sondern sich an die Handlungen des Kindes anpassen. Wichtig dabei war, ihr zu erklären, dass sie dem Kind etwa keine Fragen stellen, oder Dinge für das Kind benennen sollte.

Ebenfalls bei 2 Testungen war zusätzlich ein Geschwisterkind anwesend, einmal handelte es sich um einen etwa 10 Monate alten Bruder, einmal um eine etwa 5 Monate alte Schwester des Testkindes. Im ersten Fall verhielt es sich so, dass die Bezugsperson die meiste Zeit mit dem Geschwisterkind spielte, während sich die Testleiterin mit dem Testkind beschäftigte, im anderen Fall schlief das Geschwisterkind fast die ganze Zeit über im Kinderwagen.

Die gesamte Testung wurde auf Video aufgezeichnet und bei jeder bis auf eine Testung war eine Beobachterin anwesend, die, wie die Testleiterin, die Antworten der Kinder mitprotokollierte. Die Beobachterin hielt sich während der Testung im Hintergrund, konnte jedoch, wenn die Situation danach verlangte, ebenfalls motivierend eingreifen. Wenn das Testkind auf die Beobachterin zukam, reagierte diese selbstverständlich darauf und ließ sich ins Spiel mit einbinden.

Aufgrund des flexiblen Eingehens auf das jeweilige Testkind gestaltete sich jede Testsituation unterschiedlich, jedoch war der grobe Ablauf immer der, dass die Testung mit einer kurzen Begrüßung und Erklärung des Ablaufes für die Bezugsperson begonnen wurde, währenddessen sich das Kind im Testraum umsehen konnte. Zu diesem Zeitpunkt unterschrieb die Bezugsperson auch die Einverständniserklärung für die Videoaufzeichnung. Danach folgte eine Exploration der Puppenküche und/oder des Kaufmannsladens mit der Bezugsperson oder alleine mit der Testleiterin, je nach Verlangen des Kindes. Als erstes wurde dem Kind ein Teddybär gezeigt, der an einem Tisch saß, und gesagt, dass dieser Bär hungrig sei, weswegen nun für ihn gekocht werden müsse. Anschließend wurde meist damit begonnen, gemeinsam die Kaufmannsladensutensilien in den Kaufmannsladen zu räumen. Diese Handlung war dazu da, Kontakt mit dem Kind herzustellen. Manchmal wurde auch sofort zu den ersten Items übergegangen, wenn das Kind offensichtlich keine Aufwärmphase benötigte. Dass mit Kaufmannsladen und Puppenküche begonnen und später zu den Spielen aus der Spielebox gewechselt wurde, stellte den

Normalfall dar, wenngleich diese Reihenfolge nicht zwingend eingehalten werden musste. Hier orientierte sich die Testleiterin ganz nach den Wünschen des Kindes. Zwischenzeitliche Pausen, etwa zum Essen, waren möglich, wenn das Kind von selbst danach verlangte, die Mutter dies wünschte, oder die Aufmerksamkeit des Kindes sichtlich nachließ.

Wurden alle Items vorgegeben, oder war absehbar, dass das Kind nicht mehr bereit war, weiterzumachen, wurde der Bezugsperson mitgeteilt, dass die Testung beendet sei. Nach einer Verabschiedung war die gesamte Testdurchführung zu Ende.

Die durchschnittliche Testdauer lag bei 115,474 Minuten, mit einer Spannweite von 89,5 bis 141,5 Minuten. Davon nahm die Anlaufzeit, also die Zeit, die das Kind bis zur Exploration des Testmaterials mit der Testleiterin benötigte, inklusive Vorstellung und Erklären des Ablaufes 0,5 bis 21 Minuten (durchschnittlich 4,125 Minuten) ein. Die Verabschiedung dauerte im Schnitt 8,605 Minuten, mit einer Spannweite von 4,5 bis 17,5 Minuten. Bei 6 Testungen wurde zwischendurch keine einzige Pause gemacht, bei 8 Testungen eine Pause, bei 5 Testungen 2 Pausen und bei einer Testung wurden 3 Pausen gemacht. Hauptsächlich aßen die Kinder in den Pausen, tranken, gingen mit der Bezugsperson auf die Toilette beziehungsweise wurden gewickelt, oder saßen lediglich am Schoß der Bezugsperson. Insgesamt dauerten die Pausen pro Kind 0 bis 18 Minuten, mit einem Durchschnitt von 3,775 Minuten.

6 Ergebnisse

Zur Beantwortung der Fragestellungen, also um zu sehen, ob die einzelnen Aufgabengruppen dazu geeignet sind, die zugehörigen Entwicklungsbereiche zu erfassen, wird zunächst eine Einschätzung der Praktikabilität der Aufgabenbereiche in der Testsituation gegeben. In einer Schilderung der Handhabung der einzelnen Itemgruppen soll kurz ein persönlicher Eindruck von der Durchführung der Aufgaben an 2-Jährigen vermittelt werden.

Danach sind die Lösungshäufigkeiten der einzelnen Items für jeden Aufgabenbereich dargestellt, um einen Überblick darüber zu gewinnen, wie diese verteilt sind. Angegeben ist, wie viele gültige und wie viele fehlende Werte es pro

Item gibt. Ein Wert wurde dann als fehlend gewertet, wenn das Antwortverhalten eines Kindes aus irgendeinem Grund nicht beobachtbar oder kodierbar war, wenn das Item aufgrund von Unwillen oder Müdigkeit des Kindes nicht vorgegeben wurde oder das Kind es nicht durchführen wollte, wenn das Kind zu unaufmerksam war, etwas anderes tat als verlangt, oder weglief, oder wenn es sich dabei um ein Kind aus den Probetestungen handelte, bei welchem die Durchführung dieses Items noch anders abgelaufen war. Bei den gültigen Werten wurde meist eine Unterscheidung zwischen 0 und 1, also zwischen „nicht gelöst“ und „gelöst“ getroffen.

Anschließend sind die Itemtrennschärfen für alle Items pro Itemgruppe angegeben. Je nach Ergebnis wurden unter Umständen einzelne Items aus den anschließenden Reliabilitätsberechnungen ausgeschossen. Die Reliabilität wurde sowohl für die einzelnen Itemgruppen, als auch für die gesamten Entwicklungsbereiche berechnet.

Außerdem ist die Korrelation eines jeden Entwicklungsbereichs mit dem Alter dargestellt. Dahinter steht die Hypothese, dass die jeweiligen Aufgaben umso besser beherrscht werden, je älter ein Kind ist. Da diese Annahme gerichtet ist, wurde einseitig getestet. Aufgrund der nicht vorhandenen Voraussetzungen für eine Korrelation nach Pearson bot es sich an, ein parameterfreies Verfahren zu verwenden, nämlich eine Korrelation nach Spearman.

Schließlich wird noch auf das zentrale Problem der vielen fehlenden Werte eingegangen.

6.1 Ergebnisse zur kognitiven Entwicklung

6.1.1 Ergebnisse zur theory of mind

Der Aufgabenbereich „theory of mind“ setzt sich aus den Itemgruppen „in andere hineinversetzen“ und „Als-ob-Spiel“ zusammen.

Ob die Itemgruppe „in andere hineinversetzen“ von einem Kind bearbeitet wurde oder nicht, war sehr oft davon abhängig, wie aufmerksam das Kind bei dem dafür verwendeten ToM-Buch war. Der Beginn des Buches gestaltete sich meist noch relativ einfach, außer ein Kind lehnte von Anfang an das gesamte Bilderbuch ab. Doch da dieses Buch zusätzlich auch Aufgaben zum Wortschatz beinhaltete

(siehe Kuchler, in Druck) und deshalb noch mehr Zeit beanspruchte, war es oft nicht einfach, die Aufmerksamkeit des Kindes die ganze Zeit über aufrecht zu halten. Vor allem für die jüngeren Kinder schien es schwierig zu sein, ihre Aufmerksamkeit nicht abdriften zu lassen. Was sich auch zeigte, war ein Zusammenhang mit den sprachlichen Fähigkeiten. Vor allem diejenigen Kinder, deren Sprachfähigkeiten noch nicht weit ausgeprägt zu sein schienen (diese Kinder zeigten noch keine Dreiwortäußerungen; siehe Kuchler, in Druck), schwiegen bei den Items meist, oder reagierten gar nicht, was als falsche Antwort gewertet wurde.

Die Beobachtung des Als-ob-Spiels im freien Spiel mit der Puppenküche und dem Kaufmannsladen war gut zu handhaben. Da das freie Spiel meist viel Zeit in Anspruch nahm, ergaben sich zahlreiche Möglichkeiten für die Kinder, Als-ob-Spielhandlungen zu zeigen. Auch die Zuordnung der Verhaltensweisen zu den einzelnen Kategorien erwies sich als praxistauglich.

In Tabelle 1 sind die Lösungshäufigkeiten der einzelnen Items der beiden Itemgruppen der theory of mind angegeben. Auffällig ist hier, dass in der Aufgabengruppe „in andere hineinversetzen“ das Item „hungrig“ von der Hälfte der teilnehmenden Kinder nicht bearbeitet wurde. Am schwierigsten dürfte das Item „schlafen“ mit 11 falschen Lösungen gewesen sein, am einfachsten das Item „Pflaster“ mit 9 richtigen Lösungen.

Tabelle 1. Lösungshäufigkeiten der Items des Aufgabenbereichs „theory of mind“

[THEORY OF MIND]								
IN ANDERE HINEINVERSETZEN			ALS-OB-SPIEL					
	gültig		fehlend		gültig			fehlend
	0	1			kein Als-ob-Spiel vorhanden	Als-ob-Spiel mit realen Dingen	Als-ob-Spiel gänzlich vorgestellt	
schmutzige Hände	9	5	6	Als-ob-Spiel	2	11	5	2
hungrig	5	5	10					
schlafen	11	4	5					
Pflaster	4	9	7					
durstig	7	5	8					

Sämtliche Items der Aufgabengruppe „in andere hineinversetzen“ weisen eine ausreichend hohe Itemtrennschärfe auf. Die Trennschärfe des Items „schmutzige Hände“ ist mit 0.345 zwar eher gering, jedoch noch annehmbar (siehe Tabelle B.1 im Anhang).

Die Reliabilitätsberechnung, die mittels Cronbach Alpha durchgeführt wurde, ergibt für die Itemgruppe „in andere hineinversetzen“ eine Reliabilität von $\alpha=0.791$ (siehe ebenfalls Tabelle B.1 im Anhang). Eine Berechnung der Reliabilität des gesamten Bereichs „theory of mind“ erweist sich als nicht sinnvoll, da das Als-ob-Spiel negativ mit dem Gesamtscore korreliert (siehe Tabelle B.2 im Anhang). Da es also anscheinend nicht dasselbe misst wie die Itemgruppe „in andere hineinversetzen“, werden die beiden im Folgenden getrennt voneinander auf ihre Korrelation mit dem Alter hin untersucht.

Die Korrelationsberechnung ergibt weder eine signifikante Korrelation des Als-ob-Spiels mit dem Alter, noch eine signifikante Korrelation der Itemgruppe „in andere hineinversetzen“ mit dem Alter. Die genauen Werte sind in Tabelle 2 dargestellt.

Tabelle 2. Korrelationen des Aufgabenbereichs „theory of mind“

		Alter des Kindes (in Monaten)	Theory of mind - Hineinversetzen - MW	Theory of mind - Als-ob-Spiel
Alter des Kindes (in Monaten)	r	1,000	-,045	,301
	Sig. (1-seitig)	.	,437	,113
	N	20	15	18
Theory of mind - Hineinversetzen - MW	r	-,045	1,000	-,195
	Sig. (1-seitig)	,437	.	,252
	N	15	15	14
Theory of mind - Als-ob-Spiel	r	,301	-,195	1,000
	Sig. (1-seitig)	,113	,252	.
	N	18	14	18

6.1.2 Ergebnisse zum numerischen Wissen

Das numerische Wissen wird über die Itemgruppen „aktive Mengenerfassung“ und „passive Mengenerfassung“ erhoben, weiters über die Beobachtung, wie das Kind die einzelnen Mengen reicht, und ob es dabei zählt.

Die Aufgaben zur aktiven und passiven Mengenerfassung ließen sich gut in die Spielsituation mit dem Kaufmannsladen und der Puppenküche einbauen, da Situationen, in welchen nach etwas gefragt wurde, einfach hergestellt werden konnten, ohne dass dies unnatürlich gewirkt hätte. Selbst wenn ein Kind nicht reagierte, war es ohne Probleme möglich, die Aufgabe einfach später noch einmal zu wiederholen.

Die Beobachtung, wie ein Kind der Testleiterin eine bestimmte Menge reicht, wäre in der Testsituation selbst nicht allzu schwierig zu machen gewesen, doch da die Entscheidung, diese Itemgruppe hinzuzufügen, erst nach Beendigung der Testdurchläufe getroffen wurde, war es schwierig, diese Punkte aufgrund des Videos zu beurteilen. Genauso verhielt es sich bei der Itemgruppe „Kind zählt beim Benennen und Geben einer Menge“.

In den Tabellen 3.a und 3.b sind die Lösungshäufigkeiten der einzelnen Items zum Aufgabenbereich „numerisches Wissen“ angegeben. Augenscheinlich ist hier, dass das Item „Geben von 1“ von allen Kindern gelöst werden konnte und dass das Item „Geben von 4“ mit 13 fehlenden Werten 65 Prozent der Kinder nicht vorgegeben wurde. Auch die Items „Benennen von 4“ und „Benennen von 5“ waren bei demselben Prozentsatz an Kindern nicht abfragbar. Die Itemgruppen „wie hergegeben“ und „zählt dabei“ weisen ebenfalls viele fehlende Werte auf, was jedoch daran liegt, dass diese Itemgruppen erst nach Beendigung der Testdurchläufe hinzugefügt wurden und im Nachhinein über die Videos nur schwer zu kodieren waren.

Tabelle 3.a. Lösungshäufigkeiten der Items des Aufgabenbereichs „numerisches Wissen“

NUMERISCHES WISSEN							
GEBEN				BENENNEN			
	gültig		fehlend		gültig		fehlend
	0	1			0	1	
...von 1	0	20	0	...von 2	11	7	2
...von 2	6	13	1	...von 3	10	5	5
...von 3	12	4	4	...von 4	5	2	13
...von 4	6	1	13	...von 5	6	1	13

Tabelle 3.b. Lösungshäufigkeiten der Items des Aufgabenbereichs „numerisches Wissen“ (Fortsetzung)

NUMERISCHES WISSEN								
WIE HERGEBEN					ZÄHLT DABEI			
	gültig			fehlend		gültig		
	nur eines	eine Handvoll	eines nach dem anderen			0		fehlend
						1	1	
Geben von 2	2	4	12	2	Geben von 2	18	0	2
Geben von 3	1	2	6	11	Geben von 3	7	1	12
Geben von 4	1	1	4	14	Geben von 4	5	0	15
					Benennen von 2	9	2	9
					Benennen von 3	6	4	10
					Benennen von 4	3	3	14
					Benennen von 5	3	3	14

Beim Geben und Benennen von Mengen weisen die Items „Geben von 1“ und „Geben von 4“ eine Itemtrennschärfe unter 0.3 auf (siehe Tabelle B.3 im Anhang). In den Aufgabengruppen „hergegeben“ und „zählt dabei“ sind es 4 Items („hergeben: Geben von 2“, „zählt dabei: Geben von 2“, „zählt dabei: Geben von 4“, „zählt dabei: Benennen von 2“), welche über eine zu geringe Itemtrennschärfe, im letzten Fall sogar eine negative Korrelation mit dem Gesamtscore verfügen (siehe Tabelle B.5 im Anhang).

Unter Ausschluss dieser Items ergibt sich für die Reliabilität der aktiven und passiven Erfassung von Mengen (Geben und Benennen einer Menge) eine Wert

von $\alpha=0.738$ (siehe Tabelle B.4 im Anhang). Die Itemgruppen „hergeben“ und „zählt dabei“ weisen eine Reliabilität von $\alpha=0.829$ auf (siehe Tabelle B.6 im Anhang). Berechnet man nun die Reliabilität für den gesamten Bereich „numerisches Wissen“, ergibt sich, unter Ausschluss aller oben angeführten Items, eine Reliabilität von $\alpha=0.872$ (siehe Tabelle B.7 im Anhang).

Was die Korrelationen mit dem Alter betrifft, so zeigt sich unter Ausschluss aller genannten Items mit zu geringer Itemtrennschärfe eine signifikante Korrelation ($p<0.05$) der Itemgruppe „Kind zählt beim Benennen und Geben einer Menge“ mit dem Alter. Die Itemgruppe „Benennen einer Menge“ korreliert mit $p=0.057$ knapp nicht signifikant mit dem Alter. Die genauen Werte sind in den Tabellen 4.a und 4.b dargestellt.

Tabelle 4.a. Korrelationen des Aufgabenbereichs „numerisches Wissen“

		Alter des Kindes (in Monaten)	Num. Wissen - Geben - MW (ohne „1“, „4“)	Num. Wissen - Benennen - MW	Num. Wissen - Geben + Benennen - MW (ohne „Geben v. 1“, „Geben v. 4“)
Alter des Kindes (in Monaten)	r	1,000	,053	,385	,276
	Sig. (1-seitig)	.	,415	,057	,119
	N	20	19	18	20
Num. Wissen - Geben - MW (ohne „1“, „4“)	r	,053	1,000	,026	,783(**)
	Sig. (1-seitig)	,415	.	,461	,000
	N	19	19	17	19
Num. Wissen - Benennen - MW	r	,385	,026	1,000	,625(**)
	Sig. (1-seitig)	,057	,461	.	,003
	N	18	17	18	18
Num. Wissen - Geben + Benennen - MW (ohne „Geben v. 1“, „Geben v. 4“)	r	,276	,783(**)	,625(**)	1,000
	Sig. (1-seitig)	,119	,000	,003	.
	N	20	19	18	20
Num. Wissen - Geben - wie hergeben - MW (ohne „Geben v. 2“)	r	-,150	-,044	-,290	-,163
	Sig. (1-seitig)	,339	,451	,208	,326
	N	10	10	10	10
Num. Wissen - Geben + Benennen - zählt dabei - MW (ohne „Geben v. 2“, „Geben v. 4“, „Benennen v. 2“)	r	,560(*)	-,032	,658(**)	,518(*)
	Sig. (1-seitig)	,029	,460	,010	,042
	N	12	12	12	12
Num. Wissen - wie hergegeben + zählt dabei - MW (ohne hergeben: „Geben v. 2“, zählt dabei: „Geben v. 2“, „Geben v. 4“, „Benennen v. 2“)	r	,352	-,021	,466	,441
	Sig. (1-seitig)	,131	,474	,064	,076
	N	12	12	12	12
Num. Wissen - MW (ohne „Geben v. 1“, „Geben v. 4“, wie hergeben: „Geben v. 2“, zählt dabei: „Geben v. 2“, „Geben v. 4“, „Benennen v. 2“)	r	,314	,690(**)	,661(**)	,948(**)
	Sig. (1-seitig)	,089	,001	,001	,000
	N	20	19	18	20

* Die Korrelation ist auf dem 0,05 Niveau signifikant (einseitig).

** Die Korrelation ist auf dem 0,01 Niveau signifikant (einseitig).

Tabelle 4.b. Korrelationen des Aufgabenbereichs „numerisches Wissen“
(Fortsetzung)

		Num. Wissen - Geben - wie hergeben - MW (ohne „Geben v. 2“)	Num. Wissen - Geben + Benennen - zählt dabei - MW (ohne „Geben v. 2“, „Geben v. 4“, „Benennen v. 2“)	Num. Wissen - wie hergegeben + zählt dabei - MW (ohne: wie hergeben: „Geben von 2“ + zählt dabei: „Geben von 2“, „Geben von 4“, „Benenne von 2“)	Num. Wissen - MW (ohne „Geben v. 1“, „Geben v. 4“, wie hergeben: „Geben v. 2, zählt dabei: „Geben v. 2“, „Geben v. 4“, „Benennen v. 2“)
Alter des Kindes (in Monaten)	r	-,150	,560(*)	,352	,314
	Sig. (1-seitig)	,339	,029	,131	,089
	N	10	12	12	20
Num. Wissen - Geben - MW (ohne „1“, „4“)	r	-,044	-,032	-,021	,690(**)
	Sig. (1-seitig)	,451	,460	,474	,001
	N	10	12	12	19
Num. Wissen - Benennen - MW	r	-,290	,658(**)	,466	,661(**)
	Sig. (1-seitig)	,208	,010	,064	,001
	N	10	12	12	18
Num. Wissen - Geben + Benennen - MW (ohne „Geben v. 1“, „Geben v. 4“)	r	-,163	,518(*)	,441	,948(**)
	Sig. (1-seitig)	,326	,042	,076	,000
	N	10	12	12	20
Num. Wissen - Geben - wie hergeben - MW (ohne „Geben v. 2“)	r	1,000	-,233	,319	-,101
	Sig. (1-seitig)	.	,258	,185	,391
	N	10	10	10	10
Num. Wissen - Geben + Benennen - zählt dabei - MW (ohne „Geben v. 2“, „Geben v. 4“, „Benennen v. 2“)	r	-,233	1,000	,798(**)	,763(**)
	Sig. (1-seitig)	,258	.	,001	,002
	N	10	12	12	12
Num. Wissen - wie hergegeben + zählt dabei - MW (ohne hergeben: „Geben v. 2“, zählt dabei: „Geben v. 2“, „Geben v. 4“, „Benennen v. 2“)	r	,319	,798(**)	1,000	,703(**)
	Sig. (1-seitig)	,185	,001	.	,005
	N	10	12	12	12
Num. Wissen - MW (ohne „Geben v. 1“, „Geben v. 4“, wie hergeben: „Geben v. 2“, zählt dabei: „Geben v. 2“, „Geben v. 4“, „Benennen v. 2“)	r	-,101	,763(**)	,703(**)	1,000
	Sig. (1-seitig)	,391	,002	,005	.
	N	10	12	12	20

* Die Korrelation ist auf dem 0,05 Niveau signifikant (einseitig).

** Die Korrelation ist auf dem 0,01 Niveau signifikant (einseitig).

6.1.3 Ergebnisse zum schlussfolgernden Denken

Die Vorgabe des Aufgabenbereichs „schlussfolgerndes Denken“ über die Items zum deduktiven Denken erwies sich als gut handhabbar. Da das dD-Buch sehr bunt und ansprechend gestaltet ist, sprachen die meisten Kinder darauf an. Jedoch zeigte sich hier wieder das Problem der Aufmerksamkeit, da es für einige Kinder, meist für die jüngeren, nicht möglich war, das Bilderbuch von Anfang bis Ende vollständig zu bearbeiten. Allerdings ist anzumerken, dass sich die Durchführung dieses Buches noch am einfachsten gestaltete, was nicht zuletzt sicherlich dadurch bedingt ist, dass es sich hierbei um das kürzeste der drei Bilderbücher handelt.

Wie in Tabelle 5, in der die Lösungshäufigkeiten aller Items angegeben sind, ersichtlich ist, konnte das Item „Elefant mit Flügeln“ viel öfters als alle anderen nicht gelöst werden.

Tabelle 5. Lösungshäufigkeiten der Items des Aufgabenbereichs „schlussfolgerndes Denken“

SCHLUSSFOLGERNDES DENKEN			
DEDUKTIVES DENKEN			
	gültig		fehlend
	0	1	
Katze mit Besteck	6	9	5
Schaf auf Fahrrad	5	10	5
Elefant mit Flügeln	11	4	5
Hase mit Eis	4	11	5
Hund mit Auto	3	12	5

Genau dieses Item ist es auch, welches als einziges eine Itemtrennschärfe unterhalb des Cut-off-Werts von 0.3 aufweist (0.053; siehe Tabelle B.8 im Anhang), weshalb es aus der weiteren Berechnung ausgeschlossen wird.

Die Reliabilitätsberechnung des Aufgabenbereichs „schlussfolgerndes Denken“ ergibt unter Ausschluss des Items „Elefant mit Flügeln“ eine Reliabilität von $\alpha=0.673$ (siehe Tabelle B.9 im Anhang).

In die Korrelationsberechnung wurde dieses Item ebenfalls nicht mit einbezogen. Insgesamt weist der Aufgabenbereich „schlussfolgerndes Denken“ eine signifikante Korrelation mit dem Alter auf (siehe Tabelle 6).

Tabelle 6. Korrelationen des Aufgabenbereichs „schlussfolgerndes Denken“

		Alter des Kindes (in Monaten)	dD-Buch - MW (ohne „Elefant“)
Alter des Kindes (in Monaten)	r	1,000	,646(**)
	Sig. (1-seitig)	.	,005
	N	20	15
dD-Buch - MW (ohne „Elefant“)	r	,646(**)	1,000
	Sig. (1-seitig)	,005	.
	N	15	15

** Die Korrelation ist auf dem 0,01 Niveau signifikant (einseitig).

6.2 Ergebnisse zur visuellen Wahrnehmung

Der Aufgabenbereich „visuelle Wahrnehmung“ wird über Items zur aktiven und passiven Farbdifferenzierung, über die Itemgruppe „Sortieren nach Farben“, sowie über Aufgaben zur Form- und Größendifferenzierung erhoben.

Das Abfragen der Itemgruppen „Farben aktiv“ und „Farben passiv“ ließ sich gut in die Spielsituation mit der Puppenküche und dem Kaufmannsladen sowie beim Zeichnen einbauen. Wenn ein Kind nicht antwortete, konnte das jeweilige Item auf einfache Art und Weise wiederholt werden. Lediglich die Farben, die über die verdeckten Päckchen mit den geometrischen Formen abgefragt wurden, erwiesen sich als etwas schwieriger vorzugeben, da sich einige Kinder schlichtweg nicht für die Päckchen begeistern ließen und nicht bereit waren, zu den Tüchern mit den richtigen (oder falschen) Farben zu gehen.

Weiters war die Itemgruppe „Sortieren nach Farben“ nicht einfach zu handhaben. Da diese Aufgabe im Vergleich zu anderen Aufgaben relativ viel Zeit in Anspruch nahm und vor allem bei der Instruktion viel Aufmerksamkeit von Seiten der Kinder erforderte, brachen etliche Kinder schon zu Beginn ab und wollten die Itemgruppe auch zu keinem anderen Zeitpunkt mehr fortsetzen. Von den übrigen Kindern schafften es viele nicht, ihre Aufmerksamkeit die ganze Aufgabe über aufrecht zu halten und brachen während der Vorgabe ab.

Schließlich war auch die Durchführung der Aufgaben zur Form- und Größendifferenzierung, wie schon erwähnt, nicht immer leicht. Machten beim ersten Item der Formdifferenzierung noch die meisten Kinder mit, so weigerten sich in weiterer Folge viele, das richtige Päckchen zu suchen, passten nicht auf,

liefen weg, oder machten einfach etwas anderes. Auch war es oft der Fall, dass die Kinder einfach ein Päckchen nahmen, ohne die Instruktion überhaupt vollständig gehört zu haben.

Die Lösungshäufigkeiten der einzelnen Items des Aufgabenbereichs sind in den Tabellen 7.a und 7.b angegeben. Auffällig ist hier, dass bei der Itemgruppe „Sortieren nach Farben“ mehr als die Hälfte der Werte fehlen. Auch bei den Itemgruppen „Formdifferenzierung“ und „Größendifferenzierung“ fehlen, bis auf das Item „Formdifferenzierung 1“, über 50 Prozent der Werte.

Tabelle 7.a. Lösungshäufigkeiten der Items des Aufgabenbereichs „visuelle Wahrnehmung“

VISUELLE WAHRNEHMUNG							
FARBEN - aktiv				FARBEN – passiv			
	gültig		fehlend		gültig		fehlend
	0	1			0	1	
rot	9	10	1	rot	5	11	4
grün	10	9	1	grün	9	8	3
gelb	10	9	1	gelb	9	7	4
orange	12	7	1	orange	6	7	7
blau	8	8	4	blau	6	9	5
weiß	5	7	8	weiß	2	12	6
lila	8	7	5	lila	2	11	7

Tabelle 7.b. Lösungshäufigkeiten der Items des Aufgabenbereichs „visuelle Wahrnehmung“ (Fortsetzung)

VISUELLE WAHRNEHMUNG							
FARBEN SORTIEREN				FORM- und GRÖSSENDIFFERENZIERUNG			
	gültig		fehlend		gültig		fehlend
	0	1			0	1	
rot	4	5	11	Formdifferenzierung 1	7	9	4
grün	3	6	11	Formdifferenzierung 2	1	9	10
gelb	2	7	11	Größendifferenzierung 1	5	5	10
orange	6	3	11	Größendifferenzierung 2	4	4	12

Die Itemtrennschärfen aller Items sind ausreichend hoch (siehe Tabellen B.10 bis B.12 im Anhang).

Die Berechnung der Reliabilität mittels Cronbach Alpha für die Items zur Erfassung der visuellen Wahrnehmung ergibt für die Itemgruppe „Farbdifferenzierung“ eine Reliabilität von $\alpha=0.954$. Die Itemgruppe „Formdifferenzierung“ weist eine Reliabilität von $\alpha=0.888$ auf und die Itemgruppe „Größendifferenzierung“ eine Reliabilität von $\alpha=0.927$. Für den Aufgabenbereich insgesamt ergibt sich eine Reliabilität von $\alpha=0.962$ (siehe die Tabelle B.10 bis B.13 im Anhang).

Die Korrelationen der Itemgruppen mit dem Alter brachten das Ergebnis, dass sowohl das aktive Benennen als auch das passive Erkennen von Farben signifikant positiv mit dem Alter korrelieren ($p<0.05$ beziehungsweise $p<0.01$). Fasst man das aktive Benennen, das passive Erkennen und das Sortieren nach Farben zusammen, ergibt sich ebenfalls eine signifikante Korrelation mit dem Alter. Ebenso sind die Items zur Formdifferenzierung und zur Größendifferenzierung, wenn sie zusammengefasst werden, mit $p<0.05$ positiv mit dem Alter korreliert. Auch der Gesamtwert „visuelle Wahrnehmung - gesamt“ steht in einem positiven Zusammenhang mit dem Alter, welcher mit $p<0.05$ signifikant ist. Betrachtet man die Korrelationen innerhalb der einzelnen Itemgruppen, wird deutlich, dass mit Ausnahme der Itemgruppe „Sortieren nach Farben“ fast alle Aufgabenbereiche miteinander hoch signifikant korreliert sind. Die genauen Werte sind in den Tabellen 8.a und 8.b zu sehen.

Tabelle 8.a. Korrelationen des Aufgabenbereichs „visuelle Wahrnehmung“

		Alter des Kindes (in Monaten)	visuelle Wahrnehm. - Farbe aktiv - MW	visuelle Wahrnehm. - Farbe passiv - MW	visuelle Wahrnehm. - Farbe aktiv + passiv - MW	visuelle Wahrnehm. - Sortieren nach Farben - MW
Alter des Kindes (in Monaten)	r	1,000	,434(*)	,553(**)	,517(**)	-,393
	Sig. (1-seitig)	.	,028	,006	,010	,148
	N	20	20	20	20	9
visuelle Wahrnehm. - Farbe aktiv - MW	r	,434(*)	1,000	,852(**)	,909(**)	,317
	Sig. (1-seitig)	,028	.	,000	,000	,203
	N	20	20	20	20	9
visuelle Wahrnehm. - Farbe passiv - MW	r	,553(**)	,852(**)	1,000	,982(**)	,230
	Sig. (1-seitig)	,006	,000	.	,000	,276
	N	20	20	20	20	9
visuelle Wahrnehm. - Farbe aktiv + passiv - MW	r	,517(**)	,909(**)	,982(**)	1,000	,310
	Sig. (1-seitig)	,010	,000	,000	.	,208
	N	20	20	20	20	9
visuelle Wahrnehm. - Sortieren nach Farben - MW	r	-,393	,317	,230	,310	1,000
	Sig. (1-seitig)	,148	,203	,276	,208	.
	N	9	9	9	9	9
visuelle Wahrnehm. - Farbe aktiv + passiv + Sortieren - MW	r	,456(*)	,891(**)	,948(**)	,971(**)	,753(**)
	Sig. (1-seitig)	,022	,000	,000	,000	,010
	N	20	20	20	20	9
Visuelle Wahrnehm. - Formdiff. - MW	r	,357	,712(**)	,691(**)	,722(**)	,525
	Sig. (1-seitig)	,087	,001	,002	,001	,113
	N	16	16	16	16	7
Visuelle Wahrnehm. - Größendiff. - MW	r	,483	,679(*)	,724(**)	,703(*)	-,426
	Sig. (1-seitig)	,078	,015	,009	,012	,200
	N	10	10	10	10	6
Visuelle Wahrnehm. - Form- und Größendiff. - MW	r	,448(*)	,743(**)	,734(**)	,756(**)	,092
	Sig. (1-seitig)	,041	,000	,001	,000	,422
	N	16	16	16	16	7
Visuelle Wahrnehm. - MW	r	,507(*)	,866(**)	,925(**)	,941(**)	,319
	Sig. (1-seitig)	,011	,000	,000	,000	,201
	N	20	20	20	20	9

* Die Korrelation ist auf dem 0,05 Niveau signifikant (einseitig).

** Die Korrelation ist auf dem 0,01 Niveau signifikant (einseitig).

Tabelle 8.b. Korrelationen des Aufgabenbereichs „visuelle Wahrnehmung“
(Fortsetzung)

		visuelle Wahrnehm. - Farbe aktiv + passiv + Sortieren - MW	visuelle Wahrnehm. - Formdiff. - MW	visuelle Wahrnehm. - Größendiff. - MW	Visuelle Wahrnehm. - Form- und Größendiff. - MW	Visuelle Wahrnehm. - MW
Alter des Kindes (in Monaten)	r	,456(*)	,357	,483	,448(*)	,507(*)
	Sig. (1-seitig)	,022	,087	,078	,041	,011
	N	20	16	10	16	20
visuelle Wahrnehm. - Farbe aktiv - MW	r	,891(**)	,712(**)	,679(*)	,743(**)	,866(**)
	Sig. (1-seitig)	,000	,001	,015	,000	,000
	N	20	16	10	16	20
visuelle Wahrnehm. - Farbe passiv - MW	r	,948(**)	,691(**)	,724(**)	,734(**)	,925(**)
	Sig. (1-seitig)	,000	,002	,009	,001	,000
	N	20	16	10	16	20
visuelle Wahrnehm. - Farbe aktiv + passiv - MW	r	,971(**)	,722(**)	,703(*)	,756(**)	,941(**)
	Sig. (1-seitig)	,000	,001	,012	,000	,000
	N	20	16	10	16	20
visuelle Wahrnehm. - Sortieren nach Farben - MW	r	,753(**)	,525	-,426	,092	,319
	Sig. (1-seitig)	,010	,113	,200	,422	,201
	N	9	7	6	7	9
visuelle Wahrnehm. - Farbe aktiv + passiv + Sortieren - MW	r	1,000	,661(**)	,475	,653(**)	,913(**)
	Sig. (1-seitig)	.	,003	,083	,003	,000
	N	20	16	10	16	20
Visuelle Wahrnehm. - Formdiff. - MW	r	,661(**)	1,000	,481	,943(**)	,869(**)
	Sig. (1-seitig)	,003	.	,080	,000	,000
	N	16	16	10	16	16
Visuelle Wahrnehm. - Größendiff. - MW	r	,475	,481	1,000	,949(**)	,887(**)
	Sig. (1-seitig)	,083	,080	.	,000	,000
	N	10	10	10	10	10
Visuelle Wahrnehm. - Form- und Größendiff. - MW	r	,653(**)	,943(**)	,949(**)	1,000	,927(**)
	Sig. (1-seitig)	,003	,000	,000	.	,000
	N	16	16	10	16	16
Visuelle Wahrnehm. - MW	r	,913(**)	,869(**)	,887(**)	,927(**)	1,000
	Sig. (1-seitig)	,000	,000	,000	,000	.
	N	20	16	10	16	20

* Die Korrelation ist auf dem 0,05 Niveau signifikant (einseitig).

** Die Korrelation ist auf dem 0,01 Niveau signifikant (einseitig).

6.3 Ergebnisse zum Gedächtnis

Der Erhebung des Aufgabenbereichs „Gedächtnis“ dienen die Itemgruppen „phonologisches Gedächtnis“ und „visuelles Gedächtnis“.

Die Aufgaben zum phonologischen Gedächtnis ließen sich gut in die Spielsituation einbauen. Jedoch war es schwierig, die Kinder daran zu hindern, sich bei der Vorgabe eines Items von anderen Dingen ablenken zu lassen. Oft passierte es, dass ein Kind, während es die geforderten Objekte in die Hand nahm, plötzlich andere Objekte entdeckte, die interessanter zu sein schienen, und einfach begann, etwas anderes zu tun.

Auch bei der Vorgabe der Items zum visuellen Gedächtnis war es schwierig, die Kinder überhaupt zum Mitmachen zu bewegen, da viele von Anfang an nicht bereit waren, sich die Kärtchen auch nur anzusehen. Bei den Memory-Karten ergab sich zusätzlich das Problem, dass viele Kinder nicht warteten, bis die Kärtchen von der Testleiterin aufgedeckt wurden, sondern die Instruktion ignorierten, selbst hingingriffen und die Kärtchen umdrehten, bevor die Testleiterin dies verhindern konnte.

In Tabelle 9 ist auffällig, dass das Item „Gib mir A“ des phonologischen Gedächtnisses 19 richtige Antworten enthält. Die Items „Gib mir A+B+C“ und „Gib mir A+B+C+D“ weisen jeweils 14, also fast 75 Prozent fehlende Werte auf. Die fehlenden Werte beim Memory-Spiel des Aufgabenbereichs „visuelles Gedächtnis“ sind mit mehr als der Hälfte ebenfalls hoch ausgeprägt.

Tabelle 9. Lösungshäufigkeiten der Items des Aufgabenbereichs „Gedächtnis“

GEDÄCHTNIS							
PHONOLOGISCHES GEDÄCHTNIS				VISUELLES GEDÄCHTNIS			
	gültig		fehlend		gültig		fehlend
	0	1			0	1	
Gib mir A	1	19	0	Memory 1	2	8	10
Gib mir A + B	9	4	7	Memory 2	3	6	11
Gib mir A + B + C	5	1	14	Memory 3	2	7	11
Gib mir A + B + C + D	6	0	14	Memory 4	2	7	11
				Teddybär	12	4	4
				Haus	8	7	5

Die Trennschärpen zweier Items („Gib mir A“ und „Gib mir A+B+C+D“) sind beim phonologischen Gedächtnis nur unzureichend, weshalb sie aus der weiteren Berechnung ausgeschlossen werden. In der Itemgruppe „visuelles Gedächtnis“ ist lediglich die Trennschärpe des Items „Teddybär“ gering, jedoch mit 0.302 gerade noch akzeptabel. Zu den Itemtrennschärpen siehe die Tabellen B.14 und B.16 im Anhang.

Bei der Reliabilitätsberechnung der Itemgruppe „phonologisches Gedächtnis“ zeigt sich nach Entfernung der beiden Items mit zu geringer Itemtrennschärfe eine Reliabilität von $\alpha=0.557$. Für die Itemgruppe „visuelles Gedächtnis“ ergibt sich eine

Reliabilität von $\alpha=0.887$. Insgesamt erreicht die Reliabilität für den Bereich „Gedächtnis“ unter Ausschluss der beiden genannten Items einen Wert von $\alpha=0.882$ (siehe die Tabellen B.15 bis B.17 im Anhang). Das Item „Gib mir A+B+C“ weist nach der Zusammenfassung der beiden Itemgruppen allerdings ebenfalls eine zu geringe Itemtrennschärfe von 0.262 auf. Sein Ausschluss aus der Berechnung erhöht die Gesamt-Reliabilität dieses Aufgabenbereichs auf $\alpha=0.895$ (siehe Tabelle B.18 im Anhang).

Was die Korrelationen betrifft, so gibt es keine signifikanten Korrelationen einzelner Itemgruppen mit dem Alter. Auch die beiden Bereiche „phonologisches Gedächtnis“ und „visuelles Gedächtnis“ sind nicht signifikant miteinander korreliert (siehe Tabelle 10).

Tabelle 10. Korrelationen des Aufgabenbereichs „Gedächtnis“

		Alter des Kindes (in Monaten)	Phonologisches Gedächtnis - MW (ohne „Gib mir A“, „Gib mir A+B+C+D“)	Visuelles Gedächtnis - MW	Gedächtnis - MW (ohne „Gib mir A“, „Gib mir A+B+C+D“)
Alter des Kindes (in Monaten)	r	1,000	,061	,263	,323
	Sig. (1-seitig)	.	,421	,163	,096
	N	20	13	16	18
Phonologisches Gedächtnis – MW (ohne „Gib mir A“, „Gib mir A+B+C+D“)	r	,061	1,000	,375	,840(**)
	Sig. (1-seitig)	,421	.	,128	,000
	N	13	13	11	13
Visuelles Gedächtnis - MW	r	,263	,375	1,000	,880(**)
	Sig. (1-seitig)	,163	,128	.	,000
	N	16	11	16	16
Gedächtnis - MW (ohne „Gib mir A“, „Gib mir A+B+C+D“)	r	,323	,840(**)	,880(**)	1,000
	Sig. (1-seitig)	,096	,000	,000	.
	N	18	13	16	18

** Die Korrelation ist auf dem 0,01 Niveau signifikant (einseitig).

6.4 Ergebnisse zu den Arbeitshaltungen

Da hierzu keine eigenen Items vorhanden sind, sondern die teilweise zuvor schon ausgeführten Aufgaben herangezogen werden („Bausteine bauen“, „Zeichnen“, Durchführung des Magnetbuches, des ToM-Buches und des dD-Buches), wird

eine Beurteilung der Items zum Bereich „Arbeitshaltungen“ schon in den vorangegangenen Ergebnisdarstellungen gegeben.

Die Itemhäufigkeiten der Itemgruppen „lässt sich zurückholen“ und „komplett aufmerksam bearbeitet“ sind in Tabelle 11 angegeben. Hier sticht heraus, dass die Hälfte der Kinder das Item „Zeichnen“ und fast die Hälfte der Kinder das Item „dD-Buch“ komplett aufmerksam bearbeitet hat.

Tabelle 11. Itemhäufigkeiten des Aufgabenbereichs „Arbeitshaltungen“

AUFMERKSAMKEIT								
LÄSST SICH ZURÜCKHOLEN					KOMPLETT AUFM. BEARBEITET			
	gültig			fehlend		gültig		
	lässt sich nicht zurückholen	lässt sich zurückholen	alles aufmerksam bearbeitet			0	1	fehlend
Bausteine	11	2	3	4	Bausteine	13	3	4
Zeichnen	5	2	10	3	Zeichnen	7	10	3
Magnetbuch	11	2	2	5	Magnetbuch	13	2	5
ToM-Buch	9	0	3	8	ToM-Buch	9	3	8
dD-Buch	5	4	9	2	dD-Buch	9	9	2

Was die Itemtrennschärfen betrifft, so weist lediglich das Item „Zeichnen“ aus der Itemgruppe „längste Aufmerksamkeitsspanne (Min.)“ eine Trennschärfe von unter 0.3 auf (0.106; siehe Tabelle B.19 im Anhang).

Die Reliabilitätsberechnung mittels Cronbach Alpha ergibt für die Itemgruppen „längste Aufmerksamkeitsspanne (Min.)“ und „längste Aufmerksamkeitsspanne (S.)“ unter Ausschluss des Items „längste Aufmerksamkeitsspanne (Min.) - Zeichnen“ eine Reliabilität von $\alpha=0.845$ (siehe Tabelle B.20 im Anhang). Die Itemgruppen „lässt sich zurückholen“ und „komplett aufmerksam bearbeitet“ weisen gemeinsam eine Reliabilität von $\alpha=0.819$ auf (siehe Tabelle B.21 im Anhang). Insgesamt ergibt sich für den Bereich „Aufmerksamkeit“ unter Ausschluss des Items „längste Aufmerksamkeitsspanne (Min.) - Zeichnen“ eine Reliabilität von $\alpha=0.845$ (siehe Tabelle B.22 im Anhang).

Bei der Berechnung der Korrelationen unter Ausschluss des Items „Aufmerksamkeit (Minuten) - Zeichnen“ ergeben sich keine signifikanten positiven Korrelationen einzelner Itemgruppen mit dem Alter. Immerhin sind aber alle Itemgruppen untereinander mit $p < 0.01$ positiv korreliert (siehe Tabelle 12).

Tabelle 12. Korrelationen zum Aufgabenbereich „Arbeitshaltungen“

		Alter des Kindes (in Monaten)	längste Aufm.spanne - MW (Minuten) (ohne „Zeichnen“)	längste Aufm.spanne - MW (Seiten)	lässt sich wieder zurückholen - MW	komplett aufmerksam bearbeitet - MW
Alter des Kindes (in Monaten)	r	1,000	,292	,310	,254	,184
	Sig. (1-seitig)	.	,105	,091	,140	,219
	N	20	20	20	20	20
längste Aufm.spanne - MW (Minuten) (ohne „Zeichnen“)	r	,292	1,000	,786(**)	,696(**)	,566(**)
	Sig. (1-seitig)	,105	.	,000	,000	,005
	N	20	20	20	20	20
längste Aufm.spanne - MW (Seiten)	r	,310	,786(**)	1,000	,841(**)	,765(**)
	Sig. (1-seitig)	,091	,000	.	,000	,000
	N	20	20	20	20	20
lässt sich wieder zurückholen - MW	r	,254	,696(**)	,841(**)	1,000	,963(**)
	Sig. (1-seitig)	,140	,000	,000	.	,000
	N	20	20	20	20	20
komplett aufmerksam bearbeitet - MW	r	,184	,566(**)	,765(**)	,963(**)	1,000
	Sig. (1-seitig)	,219	,005	,000	,000	.
	N	20	20	20	20	20

** Die Korrelation ist auf dem 0,01 Niveau signifikant (einseitig).

6.5 Problem der fehlenden Werte

Das Hauptproblem bei der Durchführung der Testungen bestand darin, dass die Werte vieler Items als fehlend kategorisiert werden mussten, da sie nicht kodierbar waren, nicht durchgeführt wurden, die Kinder sie verweigerten, oder deren Aufmerksamkeit nicht hoch genug war. Durchschnittlich ergaben sich pro Kind 26,85 fehlende Werte von insgesamt 79 Items, also 33,5% der Items wurden nicht bearbeitet. Davon wurden durchschnittlich 16,2 Items vom Kind aktiv verweigert, 3,75 entfielen auf mangelnde Aufmerksamkeit. Die Probetestungen, wo sich einige Items noch vom endgültigen Itemkatalog unterschieden, weswegen diese nicht in die Berechnung mit einbezogen werden konnten, nahmen pro Kind durchschnittlich 4,5 der fehlenden Werte ein. 2,4 Items je Kind waren schließlich nicht beobachtbar oder kodierbar.

In den Tabellen 13 bis 18 sind die fehlenden Werte der einzelnen Aufgabengruppen angeführt.

Tabelle 13. Übersicht fehlender Werte „theory of mind“

THEORY OF MIND				
	Probekind	nicht beobachtbar / kodierbar	nicht vorgegeben / Kind will nicht	unaufmerksam / tut etwas anderes als verlangt / läuft weg
IN ANDERE HINEINVERSETZEN				
schmutzige Hände	3		4	2
hungrig			7	
schlafen			5	
Pflaster			5	2
durstig			7	1
ALS-OB-SPIEL				
Als-ob-Spiel		2		

Tabelle 14. Übersicht fehlender Werte „numerisches Wissen“

NUMERISCHES WISSEN				
	Probekind	nicht beobachtbar / kodierbar	nicht vorgegeben / Kind will nicht	unaufmerksam / tut etwas anderes als verlangt / läuft weg
GEBEN				
...von 1				
...von 2				1
...von 3			1	3
...von 4			13	
BENENNEN				
...von 2			2	
...von 3			5	
...von 4			13	
...von 5			13	
HERGEBEN				
Geben von 2		2		
Geben von 3		10	1	
Geben von 4		1	13	
ZÄHLT DABEI				
Geben von 2		2		
Geben von 3		11	1	
Geben von 4		2	13	
Benennen von 2		7	2	
Benennen von 3		4	6	
Benennen von 4		1	13	
Benennen von 5		1	13	

Tabelle 15. Übersicht fehlender Werte „schlussfolgerndes Denken“

SCHLUSSFOLGERNDES DENKEN				
	Probekind	nicht beobachtbar / kodierbar	nicht vorgegeben / Kind will nicht	unaufmerksam / tut etwas anderes als verlangt / läuft weg
DEDUKTIVES DENKEN				
Katze mit Besteck			2	3
Schaf auf Fahrrad			5	
Elefant mit Flügeln			5	
Hase mit Eis			5	
Hund mit Auto			5	

Tabelle 16. Übersicht fehlender Werte „visuelle Wahrnehmung“

VISUELLE WAHRNEHMUNG				
	Probekind	nicht beobachtbar / kodierbar	nicht vorgegeben / Kind will nicht	unaufmerksam / tut etwas anderes als verlangt / läuft weg
FORM- UND GRÖSSEN-DIFFERENZIERUNG				
Formdiff. 1	3			1
Formdiff. 2	3		6	1
Größendiff. 1	3		7	
Größendiff. 2	3		7	2
FARBDIFFERENZIERUNG - aktiv				
rot				1
grün				1
gelb				1
orange				1
blau	3			1
weiß	3			5
lila	3			2
FARBDIFFERENZIERUNG - passiv				
rot				
grün	3			
gelb	3		1	
orange	3		3	1
blau			5	
weiß			5	1
lila			7	
FARBEN SORTIEREN				
rot			1	10
grün			1	10
gelb			1	10
orange			1	10

Tabelle 17. Übersicht fehlender Werte „Gedächtnis“

GEDÄCHTNIS				
	Probekind	nicht beobachtbar / kodierbar	nicht vorgegeben / Kind will nicht	unaufmerksam / tut etwas anderes als verlangt / läuft weg
PHONOLOGISCHES GEDÄCHTNIS				
Gib mir A				
Gib mir A + B			5	2
Gib mir A + B + C			11	3
Gib mir A + B + C + D			13	1
VISUELLES GEDÄCHTNIS				
Memory 1	3		4	3
Memory 2	3		5	3
Memory 3	3		7	1
Memory 4	3		8	
Teddybär	3			1
Haus	3		1	1

Tabelle 18. Übersicht fehlender Werte „Arbeitshaltungen“

ARBEITSHALTUNGEN				
	Probekind	nicht beobachtbar / kodierbar	nicht vorgegeben / Kind will nicht	unaufmerksam / tut etwas anderes als verlangt / läuft weg
AUFMERKSAMKEIT				
Bausteine			4	
Zeichnen	3			
Magnetbuch	3		2	
ToM-Buch	3	1	4	
dD-Buch			2	

Die Hauptursache für das Auftreten fehlender Werte stellt, wie in den Tabellen 13 bis 18 ersichtlich ist, die Verweigerung von Items durch die Kinder dar. Es gestaltete sich teilweise überaus schwierig, oft leider auch unmöglich, ein Kind zur Durchführung eines Items zu motivieren, wenn es dies nicht wollte. Dass sich die Vorgabe der Items stark am Kind orientierte und die einzelnen Itemgruppen immer nur zu dem Zeitpunkt durchgeführt wurden, zu dem das Kind sich dazu bereit erklärte, hätte dem entgegenwirken sollen. In der Realität stellte sich allerdings heraus, dass viele Kinder bestimmte Aufgaben nicht nur ausschließlich zu einem von ihnen gewählten Zeitpunkt, sondern einige Items schlichtweg gar nicht

durchführen wollten. Hier konnte auch durch Überredungskünste und Motivation, sogar durch Zureden durch die Mutter, oft nichts erreicht werden.

Neben der Verweigerung war aber auch das Nachlassen der Aufmerksamkeit ein wichtiger Grund für die nicht mögliche Durchführung vieler Items. Während der gesamten Testung war ersichtlich, dass die Leistung, die ein Kind in den einzelnen Aufgabenbereichen zeigte, sehr stark von seiner Aufmerksamkeit abhängig war. Entweder musste ein Item während der Durchführung abgebrochen werden, weil das Kind von etwas anderem abgelenkt wurde, oder es konnte von Anfang an nicht durchgeführt werden, weil das Kind nicht auf die Frage der Testleiterin reagierte und seine Aufmerksamkeit auf etwas anderes richtete.

In Tabelle 19 ist schließlich dargestellt, dass die Vorgabe der Bilderbücher insgesamt sehr viel Zeit in Anspruch nahm und die Durchführungsdauer der einzelnen Bücher teilweise sehr stark variierte.

Tabelle 19. Dauer der einzelnen Bilderbücher

	N	Minimum	Maximum	Mittelwert	Standardabweichung
Magnetbuch - Dauer insgesamt (Minuten)	15	0,5	18,0	8,800	4,8763
ToM-Buch - Dauer insgesamt (Minuten)	12	2,5	10,5	6,958	1,9005
dD-Buch - Dauer insgesamt (Minuten)	18	0,5	8,0	3,944	1,5424

7 Diskussion

Um die Frage zu beantworten, ob der entwickelte Itempool zur Entwicklungsdiagnostik bei Kindern im Alter von 2;0 bis 2;11 Jahren praktikabel ist, und ob sich die Aufgaben zur Erhebung der einzelnen Entwicklungsbereiche eignen, soll zunächst auf jeden Entwicklungsbereich separat eingegangen werden, danach wird die allgemeine Vorgabe der Items besprochen und schließlich wird eine zusammenfassende Bewertung des gesamten Itempools gegeben.

7.1 Aufgabenbereiche

7.1.1 Kognitive Entwicklung

Was die kognitive Entwicklung betrifft, so werden die einzelnen Unterkategorien dieses Entwicklungsbereichs im Folgenden separat voneinander betrachtet.

7.1.1.1 Theory of mind

Da sich die theory of mind erst ab einem Alter von 3 Jahren zu entwickeln beginnt (siehe zum Beispiel Carlson, Moses & Breton, 2002; Carlson, Moses & Claxton, 2004; Wellman, Cross & Watson, 2001), wurde das Augenmerk in dieser Untersuchung auf ihre Vorläufer gelegt. Aufgrund praktischer Überlegungen wurde einerseits die Fähigkeit, Urteile über die Wünsche anderer Personen zu bilden, andererseits das Vorhandensein des Als-ob-Spiels ausgewählt. Das Bilden von Urteilen über die Wünsche anderer wurde mithilfe eines Bilderbuches, des ToM-Buches, erhoben, in welchem verschiedene Personen abgebildet sind, in die sich das Kind hineinversetzen musste, um deren weitere Handlungsschritte vorherzusagen. Das Als-ob-Spiel wurde in der Spielsituation mit Puppenküche und Kaufmannsladen beobachtet und kategorisiert.

Für das Lösen oder Nicht-Lösen der Aufgabengruppe „Hineinversetzen in andere“ spielten die Aufmerksamkeit und die Verweigerung durch die Kinder eine große Rolle, weswegen hier eventuell eine andere Form der Item-Vorgabe als mithilfe eines Buches angedacht werden sollte, da das Kind hierbei ruhig am Tisch zu sitzen hat und von ihm verlangt wird, seine Aufmerksamkeit vollständig auf das Buch zu richten. Dass im Zuge der Vorgabe des ToM-Buches auch gleichzeitig Items zum Wortschatz abgefragt wurden (siehe Kuchler, in Druck), trug dazu bei, dass sich die ohnehin schon lange Dauer der Vorgabe noch weiter erhöhte. Möglicherweise reicht es also aus, die Items zur Erfassung des Wortschatzes aus dem ToM-Buch herauszunehmen, damit die Durchführungsdauer dieses Buches verringert und die Anzahl der bearbeiteten Items erhöht wird. Das Als-ob-Spiel war in der Testsituation gut beobachtbar und auch die Kategorisierung erwies sich als gut handhabbar.

Was die Itemschwierigkeiten betrifft, so sollten die Items „schlafen“ und „Pflaster“ noch einmal genauer betrachtet werden, da das eine für die Kinder auf den ersten Blick sehr schwierig und das andere sehr leicht gewesen sein dürfte. Insgesamt weisen alle Items zur Aufgabengruppe „in andere hineinversetzen“ ausreichend hohe Itemtrennschärfen auf und auch die Reliabilität der Aufgabengruppe ist akzeptabel.

Weder die Items zum Hineinversetzen in andere noch das Als-ob-Spiel bilden jedoch den Altersverlauf des Entwicklungsbereichs „theory of mind“ in der vorliegenden Altersgruppe ab, es bestehen also keine Zusammenhänge mit dem Alter, was eine Überarbeitung beider Itemgruppen sinnvoll erscheinen lässt.

Insgesamt lässt sich also sagen, dass dem nicht vorhandenen Zusammenhang der Itemgruppen mit dem Alter auf den Grund gegangen werden sollte, wobei vor allem die Items „schlafen“ und „Pflaster“ bezüglich ihrer Itemschwierigkeit einer genaueren Betrachtung bedürfen. Auch eine Trennung der Items zum Wortschatz und zur theory of mind im ToM-Buch erscheint sinnvoll, da die Vorgabe des gesamten Buches insgesamt viel Zeit in Anspruch nimmt, was der durchgehenden Aufmerksamkeit der Kinder abträglich sein dürfte.

7.1.1.2 Numerisches Wissen

Die Erhebung des numerischen Wissens erfolgte durch die Überprüfung der aktiven und passiven Mengenerfassung des Kindes, da dies bis zum Alter von 3 Jahren bei kleinen Mengen schon möglich sein sollte (Condry & Spelke, 2008; Wynn, 1990). Weiters wurde beobachtet, auf welche Weise das Kind die geforderte Menge reicht und ob es dabei, sowie beim Benennen der Mengen, zählt oder nicht, da nach Wynn (1990) auch dies einen Hinweis auf den Entwicklungsstand des numerischen Wissens liefert.

Für die praktische Anwendung in der Testsituation eignete sich die Aufgabengruppe prinzipiell gut. Die Items wurden von den Kindern angenommen, konnten leicht beobachtet und bei Bedarf wiederholt werden. Dennoch zeigten sich einige fehlende Werte, aufgrund derer es notwendig sein wird, zu überlegen, ob die Vorgabe der Items eventuell anders gestaltet werden könnte, damit nicht so viele Kinder verweigern.

Die Tatsache, dass das Item „Geben von 1“ zumindest in dieser Untersuchung von allen Kindern gelöst werden konnte, zeigt zusätzlich, dass die Vorgabe dieses Items offensichtlich keinen allzu großen Gewinn bringt.

Die Reliabilitäten sind für die einzelnen Itemgruppen und den gesamten Bereich „numerisches Wissen“ akzeptabel bis gut, wenn man die Items „Geben von 1“, „Geben von 4“, „hergeben: Geben von 2“, „zählt dabei: Geben von 2“, „zählt dabei: Geben von 4“ und „zählt dabei: Benennen von 2“ aufgrund ihrer zu geringen Itemtrennschärfen aus der Berechnung eliminiert.

Sowohl die Itemgruppen „Kind zählt beim Benennen und Geben einer Menge“ und „Benennen einer Menge“ dürften den Entwicklungsverlauf gut darstellen, da diese beiden Gruppen signifikant beziehungsweise nur knapp nicht signifikant mit dem Alter korrelieren. Bei den übrigen beiden Itemgruppen „Geben einer Menge“ und „Hergeben“ stellt sich die Frage, inwieweit eine Beibehaltung im Itempool mit einer verringerten Itemanzahl (2 beziehungsweise 3 Items) aufgrund der nicht vorhandenen Darstellung eines Alterstrends sinnvoll ist. Zumindest für eine Beibehaltung der Itemgruppe „Geben einer Menge“ spricht aber, dass die Items für die Itemgruppe „Zählt beim Geben und Benennen einer Menge“ benötigt werden.

Zusammenfassend lässt sich zu diesem Bereich sagen, dass zu überlegen ist, ob die Erfassung des numerischen Wissens nicht auf die Itemgruppen „Benennen einer Menge“, „Geben einer Menge“ und „Zählt beim Benennen und Geben einer Menge“ beschränkt werden sollte, wobei die Gruppe „Geben einer Menge“ nur aus den Items „2 Objekte“ und „3 Objekte“ bestehen würde. Allerdings ist diese Überlegung aufgrund der vielen fehlenden Werte bei einer Stichprobe von nur 20 Kindern nur mit Vorsicht zu genießen. Eventuell sollte der gesamte Aufgabenbereich in einer Folgestudie unverändert übernommen werden, um zusätzliche Daten zu sammeln und damit ein aussagekräftigeres Ergebnis zu gewinnen, welches auf eine größere Anzahl von Daten gestützt ist.

7.1.1.3 Schlussfolgerndes Denken

Das schlussfolgernde Denken wurde über Items zum deduktiven Denken erhoben. Da mithilfe von Fantasie-Aufgaben schon 2- bis 4-jährige Kinder in der Lage sind,

deduktive Schlüsse zu ziehen (Richards & Sanderson, 1999), wurde ein Bilderbuch (dD-Buch) gestaltet, mit welchem in Form von Syllogismen Items zum deduktiven Denken abgefragt werden.

Die Durchführung der Items in der Praxis funktionierte gut, wenngleich hier, wie auch schon in den zuvor diskutierten Bereichen, wieder der Aufmerksamkeit und dem Willen der Kinder eine wichtige Rolle zukam.

Das Item „Elefant mit Flügeln“ konnte im Vergleich zu allen anderen am häufigsten nicht gelöst werden. Es ist möglich, dass dies daher rührt, dass es das einzige Item ist, in der die zweite Auswahlmöglichkeit die Verneinung der ersten ist („Hat Cäsar Flügel oder nicht?“; zum Vergleich das Item „Katze mit Besteck“: „Isst Julia ihr Essen aus dem Napf oder mit Messer und Gabel?“). Daher sollte es für zukünftige Erhebungen nicht mehr berücksichtigt, oder zumindest abgewandelt und an die übrigen Items angepasst werden.

Die Aufgabengruppe zum deduktiven Denken weist insgesamt, selbst wenn das Item „Elefant mit Flügeln“ aufgrund seiner zu geringen Itemtrennschärfe aus der Berechnung ausgeschlossen wird, nur eine geringe Reliabilität auf. Aufgrund dessen ist für eine weitere Bearbeitung zu überlegen, wie diese Aufgabengruppe umgestaltet und verbessert werden könnte, damit die Reliabilität steigt.

Die Korrelation mit dem Alter ist hingegen signifikant, was darauf schließen lässt, dass dieser Aufgabenbereich den Altersverlauf gut abbildet.

Alles in allem lässt sich sagen, dass der Bereich „deduktives Denken“ eine gute Basis für weiterführende Untersuchungen darstellt, jedoch sind Veränderungen, die der Verbesserung der Reliabilität dienen, unbedingt notwendig.

7.1.2 Visuelle Wahrnehmung

Die Bereich „visuelle Wahrnehmung“ wurde über Items zur aktiven und passiven Farbdifferenzierung, zur Stabilität des Farbkonzeptes, zur Formdifferenzierung und zur Größendifferenzierung erhoben, da sich in dem hier interessierenden Altersbereich zum einen gerade das Wissen um die vier Grundfarben entwickelt (Holle, 1984/2000; Johnson, 1977). Zum anderen kam Gruber (1992) zu dem Ergebnis, dass die Items zur Form- und Größendifferenzierung in seiner Studie für

die Gruppe der 3- bis 6-Jährigen zu einfach waren. Aufgrund dessen wurden diese für die vorliegende Studie in leicht abgeänderter Form übernommen.

Bei den Aufgaben zur Form- und Größendifferenzierung sowie dem Sortieren nach Farben, kam es zu vielen fehlenden Werten, was deren Eignung für den Einsatz in der Praxis einschränkt. Es zeigte sich in der Testsituation, dass gerade diese Items sehr oft durch aufmerksamkeitsbedingte Abbrüche gekennzeichnet waren, oder aufgrund von Verweigerung gar nicht erst probiert wurden. Hier erscheint es sinnvoll, an der Art der Durchführung der Items zu arbeiten, da vor allem die Instruktionen viel Aufmerksamkeit von den Kindern fordern. Die Items zum Benennen und Erkennen von Farben erwiesen sich hingegen als praxistauglich, da sie sich gut ins Spiel einbauen ließen.

Die Itemtrennschärfen der Items, sowie die Reliabilität der einzelnen Itemgruppen weisen allesamt hohe Werte auf.

Bei der Berechnung der Korrelationen zeigt sich, dass der Gesamtwert der Aufgabengruppe einen positiven Zusammenhang mit dem Alter aufweist. Betrachtet man die einzelnen Itemgruppen gesondert, stellen das Erkennen und das Benennen von Farben den Entwicklungsverlauf der visuellen Wahrnehmung von 2;0 bis 2;11 Jahren gut dar. Fasst man das aktive Benennen, das passive Erkennen und das Sortieren von Farben zusammen, ergibt sich ebenfalls ein signifikanter Zusammenhang mit dem Alter. Dasselbe trifft auf die Itemgruppen zur Form- und Größendifferenzierung zu, wenn man sie zusammengefasst betrachtet. Untereinander korrelieren bis auf das „Sortieren nach Farben“ alle Itemgruppen hoch miteinander.

Insgesamt ist also zu überlegen, ob die Itemgruppe „Sortieren nach Farben“ nicht weggelassen werden sollte. Da sie nicht signifikant mit den übrigen Itemgruppen dieses Entwicklungsbereichs korreliert, viele fehlende Werte beinhaltet und auch sonst nicht unbedingt notwendig für ein aussagekräftiges Ergebnis ist, wäre ihr Weglassen kein Verlust. Hingegen würde es zu einer Verkürzung der gesamten Testdauer beitragen. Die Items zur Form- und zur Größendifferenzierung sollten noch einmal überarbeitet werden, um sie praxistauglicher zu gestalten, damit die große Anzahl an fehlenden Werten reduziert werden kann.

7.1.3 Gedächtnis

Der Erhebung des Gedächtnisses dienten Aufgaben zum phonologischen sowie zum visuellen Gedächtnis. Das phonologische Gedächtnis wurde erhoben, indem die Gedächtnisspanne für Wörter abgefragt wurde, welche im Alter von 2 Jahren bei 3 Wörtern liegt (Dempster, 1981). Das visuelle Gedächtnis wurde in einer ähnlichen Form wie bei den delayed-response tasks bei Reznick (2009) erhoben, nämlich indem überprüft wurde, ob das Kind ein kurz zuvor gesehenes Bild wieder erkennen beziehungsweise wieder finden kann.

Sowohl die Ergebnisse zum phonologischen, als auch die Ergebnisse zum visuellen Gedächtnis weisen viele fehlende Werte auf. Ein Großteil dieser entstand durch Verweigerung der Kinder. Eine Änderung der Vorgabe der Items wäre aufgrund dessen in jedem Fall ratsam, um die Eignung für eine praktische Anwendung zu erhöhen.

Die Items zum phonologischen Gedächtnis weisen auch nach Ausschluss zweier Items mit zu geringer Itemtrennschärfe („Gib mir A“, „Gib mir A+B+C+D“) eine nicht ausreichend hohe Reliabilität auf, weswegen eine Überarbeitung angebracht erscheint. Die Reliabilität der Items zum visuellen Gedächtnis ist annehmbar. Insgesamt ist die Reliabilität der gesamten Aufgabengruppe ausreichend hoch, jedoch weist bei der Berechnung der Gesamt-Reliabilität noch ein weiteres Item des phonologischen Gedächtnisses eine zu geringe Itemtrennschärfe auf, was die Überarbeitungsbedürftigkeit dieser Itemgruppe noch weiter unterstreicht.

Der Altersverlauf dürfte durch keine der Aufgabengruppen gut abgebildet werden, da es keine signifikanten Korrelationen mit dem Alter gibt. Es empfiehlt sich daher, den gesamten Aufgabenbereich noch einmal zu überarbeiten und eventuell andere Wege zu finden, um das Gedächtnis in der Altersgruppe der 2- bis 3-Jährigen adäquat zu erheben.

Zusammenfassend kann gesagt werden, dass sowohl die Aufgaben zum phonologischen, als auch die zum visuellen Gedächtnis einer Überarbeitung bedürfen und eventuell Alternativen für die Erhebung dieser Bereiche gefunden werden sollten.

7.1.4 Arbeitshaltungen

Was die Aufgabengruppe „Arbeitshaltungen“ betrifft, so wurden für die Erhebung lediglich Aufgaben zur fokussierten Aufmerksamkeit verwendet. Eine Überprüfung der Leistungsmotivation wäre nicht sinnvoll gewesen, da diese im Kleinkindalter noch nicht entwickelt ist (Barett, 1998; Heckhausen, 1985). Die Items zum Spielen mit den Bausteinen, zum Zeichnen, zum Magnetbuch, zum ToM-Buch und zum dD-Buch wurden dazu herangezogen.

Die Diskussion der Handhabung der Items sowie der Lösungshäufigkeiten ist bei den jeweils zu den Materialien zugehörigen Aufgabengruppen nachzulesen.

Die Reliabilitäten der einzelnen Itemgruppen sind relativ hoch, lediglich das Item „Aufmerksamkeit (Minuten), Zeichnen“ wurde aufgrund einer zu geringen Itemtrennschärfe eliminiert. Dies ist aber wohl darauf zurückzuführen, dass einige Itemgruppen dasselbe in nur abgeänderter Form erheben. So bringt es lediglich einen irrelevanten Informationsgewinn, zusätzlich zu den Minuten, die ein Kind aufmerksam war, auch die Seitenanzahl bei den Bilderbüchern zu erheben. Auch die Itemgruppe „komplett aufmerksam bearbeitet“ ist eigentlich in der Gruppe „lässt sich zurückholen“ enthalten und bringt daher keinen Mehrgewinn.

Entgegen den Vermutungen aus der Literatur (zum Beispiel Ruff & Capozzoli, 2002; Ruff & Lawson, 1990) zeigte sich kein signifikanter positiver Zusammenhang mit dem Alter.

Hier stellt sich also die Frage, inwieweit die Erfassung der Aufmerksamkeit sinnvoll ist, um Aussagen über den Entwicklungsstand eines Kindes ableiten zu können. Neue Möglichkeiten der Erhebung sollten eventuell angedacht werden.

7.2 Testsituation

Für den Testleiter / die Testleiterin verlangt die Durchführung der Items im Vorfeld viel Übung und eine intensive Auseinandersetzung mit dem Material. Obwohl in dieser Studie die Testleiterinnen auch gleichzeitig die Entwicklerinnen des Itempools waren und daher die Vertrautheit mit dem Material schon gegeben war, erwies sich der Einbau der Items in die Spielsituation gerade in den ersten Testungen noch als schwierig. Dies schlug sich vor allem in der Testdauer nieder,

die zu Beginn länger war als in späteren Durchgängen. Testleiter / Testleiterinnen, die zum ersten Mal mit den Aufgaben in Berührung kommen, benötigen daher sicherlich viel Vorbereitungszeit. Die Testvorgabe, die je nach Kind etwa 1,5 bis 2 Stunden dauert, verlangt vom Testleiter / von der Testleiterin von Anfang bis Ende hohe Konzentration und Engagement ab. Je öfter die Vorgabe jedoch durchgeführt wird, desto einfacher wird sie.

Die Anwesenheit eines Beobachters / einer Beobachterin ist nach der Erfahrung, die durch die 20 Testungen gewonnen wurde, nicht zwingend notwendig, wenn man mit den Items vertraut ist. Zumindest was die hier relevanten Entwicklungsbereiche betrifft ist es für einen Testleiter / eine Testleiterin möglich, die Untersuchung allein durchzuführen und trotzdem alle Antworten des Kindes mitzuprotokollieren. Jedoch bedarf dies sicherlich einiger Übung. Für ungeübte Testleiter / Testleiterinnen empfiehlt sich daher ein zusätzlicher Beobachter / eine zusätzliche Beobachterin oder ein Videomitschnitt, um keine relevanten Äußerungen oder Reaktionen des Kindes zu verpassen.

Die Anwesenheit einer Bezugsperson, die bei Kindern im Alter von 2;0 bis 2;11 Jahren in jedem Fall notwendig ist, war nie störend. Zu Beginn der Testung wurde der jeweiligen Bezugsperson erklärt, wie sie sich verhalten und während der Testung auf das Kind reagieren solle. Diese Anweisungen wurden im Normalfall auch eingehalten. Die Bezugspersonen hielten sich in den allermeisten Fällen im Hintergrund, waren jedoch für ihre Kinder da, wenn diese nach ihnen verlangten, genau so, wie es vorgesehen war.

Was die Testdauer betrifft, so ist diese mit 1,5 bis 2 Stunden sehr lange. Dem könnte mit einer Aufteilung auf zwei oder mehrere Testungstermine entgegengewirkt werden. In der vorliegenden Studie dauerte eine Testung immer so lange, bis alle Entwicklungsbereiche fertig abgefragt waren, oder das Kind sichtlich nicht mehr weiter machen wollte oder konnte. Würde man die Testung auf zumindest zwei Termine aufteilen, würde die Dauer auf ein akzeptables Maß gesenkt werden. Allerdings stellt sich hier die Frage, ob eine akzeptable Länge nicht auch einfacher durch eine Verringerung der Itemanzahl erreicht werden könnte.

Generell war die Aufmerksamkeit ein zentrales Thema bei der Durchführung des Itempools. Etliche Itemgruppen konnten nicht vollständig durchgeführt werden,

weil die Kinder nach kurzer Zeit nicht mehr aufmerksam waren oder sich mit etwas anderem beschäftigten. Vor allem bei den Instruktionen fiel auf, dass diese teilweise wohl zu lange dauerten, da die Kinder oft nicht bis zum Ende warteten, sondern sich schon vorher mit dem Material beschäftigen wollten. Auch bei den Bilderbüchern ist auffällig, dass deren Dauer jeweils eine hohe Spannweite aufweist. Dies schlug sich in der Aufmerksamkeit der Kinder nieder und wirft die Forderung nach einer Kürzung der Bücher auf, um dadurch vielleicht zu einer vollständigeren Bearbeitung durch die Kinder zu gelangen. Auch die Kürzung anderer Itemgruppen oder Instruktionen ist anzuraten.

Neben der mangelnden Aufmerksamkeit stellte auch die Verweigerung von Items durch die Kinder ein zentrales Problem bei der Vorgabe des Itempools dar. Viele Kinder ließen sich weder durch die Testleiterin, noch durch die Bezugsperson von der Durchführung einzelner Items überzeugen. Um diesem Problem entgegenzuwirken, bietet sich hier wieder das Ansetzen eines zweiten Testungstermins an, bei welchem das Kind unter Umständen doch Gefallen an den zuvor verweigerten Items findet. Außerdem könnten bei einigen Aufgabenbereichen andere Möglichkeiten der Vorgabe überlegt werden.

7.3 Zusammenfassung

Die Idee, einen klassischen diagnostischen Ansatz mit einem spielbasierten zu verbinden, ließ sich gut in die Realität umsetzen. Es funktionierte, die verschiedenen Items in die Testsituation einzubauen, ohne den Charakter eines freien Spiels zu verlieren. Das Setting mit Puppenküche und Kaufmannsladen erwies sich für die Erhebung als geeignet, da alle Kinder darauf ansprachen und mit den vorgegebenen Utensilien spielten. Die Spielebox mit den verschiedenen Zusatz-Spielen bot eine gute Ergänzung, da auf diese zurückgegriffen werden konnte, wenn sich ein Kind nicht mehr mit der Puppenküche oder dem Kaufmannsladen beschäftigen wollte.

Allerdings stellten die oft nicht genug vorhandene Aufmerksamkeit und die Verweigerung von Items durch die Kinder zentrale Probleme dar. Dem könnte durch die Umsetzung eines zweiten Testungstermins entgegengewirkt werden. Abgesehen davon sollten alle Instruktionen so kurz wie möglich gehalten werden,

da sich herausstellte, dass einige für die Kinder offensichtlich zu lange waren, und es sollte überlegt werden, wie die Itemanzahl insgesamt sinnvoll gekürzt werden könnte.

Die einzelnen Aufgabengruppen stellen schon eine gute Basis für die Erhebung der Entwicklungsbereiche dar, jedoch müssen manche Wege der Vorgabe, einzelne Items, oder ganze Itemgruppen noch einmal überarbeitet, verbessert, oder teilweise durch neue Ideen ersetzt werden.

III. LITERATURVERZEICHNIS

-
- Albers, C. A. & Grieve, A. J. (2007). Test review: Bayley, N. (2006). Bayley Scales of Infant and Toddler Development – Third Edition. San Antonio, TX: Harcourt Assessment. *Journal of Psychoeducational Assessment*, 25, 180-190.
- Anderson, J. R. (2007). *Kognitive Psychologie* (6. Aufl.) (G. Plata, Übers.). Berlin: Springer. (Original erschienen 2005: Cognitive psychology and its implications, Sixth Edition)
- Appl, D. J. (2000). Clarifying the preschool assessment process: Traditional practices and alternative approaches. *Early Childhood Education Journal*, 27, 219-225.
- Asendorpf, J. B. (2002). Self-awareness, other-awareness, and secondary representation. In A. N. Meltzoff & W. Prinz (Eds.), *The imitative mind: Development, evolution, and brain bases. Cambridge studies in cognitive perceptual development* (pp. 63-73). New York: Cambridge University Press.
- Asendorpf, J. B. & Baudonnière, P.-M. (1993). Self-awareness and other-awareness: Mirror self-recognition and synchronic imitation among unfamiliar peers. *Developmental Psychology*, 29, 88-95.
- Asendorpf, J. B., Warkentin, V. & Baudonnière, P.-M. (1996). Self-awareness and other-awareness II: Mirror self-recognition, social contingency awareness, and synchronic imitation. *Developmental Psychology*, 32, 313-321.
- Astington, J. W. & Jenkins J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology*, 35, 1311-1320.
- Athanasidou, M. S. (2007). Play-based approaches to preschool assessment. In B. A. Bracken & R. J. Nagle (Eds.), *Psychoeducational assessment of preschool children* (4th ed.) (pp. 219-238). Mahwah: Lawrence Erlbaum Associates Publishers.
- Aunio, P. & Niemivirta, M. (2010). Predicting children's mathematical performance in grade one by early numeracy. *Learning and Individual Differences*, 20, 427-435.
- Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4, 417-423.
- Bagnato, S. J. (2005). The authentic alternative for assessment in early intervention: An emerging evidence-based practice. *Journal of Early Intervention*, 28, 17-22.
- Bagnato, S. J. & Neisworth, J. T. (1994). A national study of the social and treatment "invalidity" of intelligence testing for early intervention. *School Psychology Quarterly*, 9, 81-102.
- Barrett, K. C. (1998). A functionalist perspective to the development of emotions. In M. F. Mascolo & S. Griffin (Eds.), *What develops in emotional development?* (pp.109-133). New York: Plenum Press.
- Bartsch, K. & Wellman, H. M. (1995). *Children talk about the mind*. New York: Oxford University Press.
- Bauer, P. J. (2002). Long-term recall memory: Behavioral and neuro-developmental changes in the first 2 years of life. *Current directions in psychological science*, 11, 137-141.
- Bauer, P. J. (2004). Getting explicit memory off the ground: Steps toward construction of a neurodevelopmental account of changes in the first two years of life. *Developmental Review*, 24, 347-373.

- Bauer, P. J., Wenner, J. A., Dropik, P. L., Wewerka, S. S. & Howe, M. L. (2000). Parameters of remembering and forgetting in the transition from infancy to early childhood. *Monographs of the Society for Research in Child Development*, 65, i-213.
- Bayley, N. (1969). *Bayley Scales of Infant Development*. New York: Psychological Corp.
- Bayley, N. (2006). *Bayley Scales of Infant and Toddler Development – Third edition*. San Antonio: Harcourt Assessment.
- Bayley, N. (2007). *Bayley-II: Bayley Scales of Infant Development* (2nd ed.) (G. Reuner, J. Rosenkranz, J. Pietz & R. Horn, Übers.). Frankfurt: Harcourt Test Services. (Original erschienen 1993: Bayley Scales of Infant Development, Second Edition)-
- Bertenthal, B. I., Campos, J. J. & Haith, M. M. (1980). Development of visual organization: The perception of subjective contours. *Child Development*, 51, 1072-1080.
- Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, 94, 115-147.
- Biederman, I., Mezzanotte, R. J. & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, 14, 143-177.
- Boothe, R. G. (2002). *Perception of the visual environment*. New York: Springer.
- Briars, D. & Siegler, R. S. (1984). A featural analysis of preschoolers' counting knowledge. *Developmental Psychology*, 20, 607-618.
- Capdevila, R. (1993). *In unserem Haus gibt's viel zu sehen*. Wien: Tosa. (Original erschienen 1984: Mirem la casa)
- Carlson, S. M., Mandell, D. J. & Williams, L. (2004). Executive function and theory of mind: Stability and prediction from ages 2 to 3. *Developmental Psychology*, 40, 1105-1122.
- Carlson, S. M. & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, 72, 1032-1053.
- Carlson, S. M., Moses, L. J. & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development*, 11, 73-92.
- Carlson, S. M., Moses, L. J. & Claxton, L. J. (2004). Individual differences in executive functioning and theory of mind: An investigation of inhibitory control and planning ability. *Journal of Experimental Child Psychology*, 87, 299-319.
- Carpenter, M., Call, J. & Tomasello, M. (2002). Understanding "prior intentions" enables two-year-olds to imitatively learn a complex task. *Child Development*, 73, 1431-1441.
- Carpenter, M., Call, J. & Tomasello, M. (2005). Twelve- and 18-month-olds copy actions in terms of goals. *Developmental Science*, 8, F13-F20.
- Cicchetti, D. & Wagner, S. (1990). Alternative assessment strategies for the evaluation of infants and toddlers: An organizational perspective. In S. J. Meisels (Ed.), *Handbook of early childhood intervention* (pp. 246-253). Cambridge: Cambridge Univ. Press.
- Condry, K. F. & Spelke, E. S. (2008). The development of language and abstract concepts: The case of natural number. *Journal of Experimental Psychology: General*, 137, 22-38.
- Courage, M. L. & Howe, M. L. (2002). From infant to child: The dynamics of cognitive change in the second year of life. *Psychological Bulletin*, 128, 250-277.

-
- Deater-Deckard, K., Petrill, S. A., Thompson, L. A. & DeThorne, L. S. (2005). A cross-sectional behavioral genetic analysis of task persistence in the transition to middle childhood. *Developmental Science*, 8, F21-F26.
- Deimann, P. & Kastner-Koller, U. (2007). Entwicklungsdiagnostik. In M. Hasselhorn & W. Schneider (Hrsg.), *Handbuch der Entwicklungspsychologie* (S. 558-569). Göttingen: Hogrefe.
- Dempster, F. N. (1981). Memory span: Sources of individual and developmental differences. *Psychological Bulletin*, 89, 63-100.
- Derryberry, D. & Rothbart, M. K. (1997). Reactive and effortful processes in the organization of temperament. *Development and Psychopathology*, 9, 633-652.
- Eberhard, I. (2007). *Was gehört wohin?: Mein großes Spielbuch mit 22 Magnetteilen*. Ravensburg: Ravensburger.
- Edelman, G. M. (2004). *Das Licht des Geistes: Wie Bewusstsein entsteht* (Ch. Trunk, Übers.). Düsseldorf: Walter. (Original erschienen 2004: *Wider than the sky: The phenomenal gift of consciousness*)
- Eisert, D. & Lamorey, S. (1996). Play as a window on child development: The relationship between play and other developmental domains. *Early Education and Development*, 7, 221-235.
- Evans, J. St. B. T. (2010). On the nature of inductive inference. *American Journal of Psychology*, 123, 118-121.
- Farrar, M. J. & Maag, L. (2002). Early language development and the emergence of a theory of mind. *First Language*, 22, 197-213.
- Feigenson, L., Dehaene, S. & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, 8, 307-314.
- Feinfield, K. A., Lee, P. P., Flavell, E. R., Green, F. L. & Flavell, J. H. (1999). Young children's understanding of intention. *Cognitive Development*, 14, 463-486.
- Flavell, J. H. (1986). The development of children's knowledge about the appearance-reality distinction. *American Psychologist*, 41, 418-425.
- Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, 101, 171-191.
- Folio, R. M. & Fewell, R. R. (2000). *Peabody Developmental Motor Scales – Second Edition*. Austin, TX: Pro-Ed.
- Frye, D., Braisby, N., Lowe, J., Maroudas, C. & Nicholls, J. (1989). Young children's understanding of counting and cardinality. *Child Development*, 60, 1158-1171.
- Fuson, K. C. (1988). *Children's counting and concepts of number*. New York: Springer.
- Fuson, K. C. & Hall, J. (1983). The acquisition of early number word meanings: A conceptual analysis of review. In H. P. Ginsburg (Ed.), *The development of mathematical thinking* (pp. 49-107). New York: Academic Press.
- Fuson, K. C., Richards, J. & Briars, D. J. (1982). The acquisition and elaboration of the number word sequence. In Ch. J. Brainerd (Ed.), *Children's logical and mathematical cognition: Progress in cognitive development research* (pp. 33-92). New York: Springer.
- Gelman, R. & Gallistel, C. R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.

- Gelman, S. A. & Coley, J. D. (1990). The importance of knowing a dodo is a bird: Categories and inferences in 2-year-old children. *Developmental Psychology*, 26, 796-804.
- Gelman, S. A., Collman, P. & Maccoby, E. E. (1986). Inferring properties from categories versus inferring categories from properties: The case of gender. *Child Development*, 57, 396-404.
- Gelman, S. A. & Markman, E. M. (1986). Categories and induction in young children. *Cognition*, 23, 183-209.
- Gelman, S. A. & Markman, E. M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development*, 58, 1532-1541.
- Gergely, G., Bekkering, H. & Király, I. (2002). Developmental psychology: Relational imitation in preverbal infants. *Nature*, 415, 755.
- Gopnik, A. & Meltzoff, A. N. (1997). *Words, thoughts, and theories*. Cambridge, MA: MIT Press.
- Goswami, U. (2001). *So denken Kinder: Einführung in die Psychologie der kognitiven Entwicklung* (1. Aufl.) (M. Wengenroth, Übers.). Bern: Huber. (Original erschienen 1998: Cognition in children)
- Goswami, U. & Brown, A. L. (1989). Melting chocolate and melting snowmen: Analogical reasoning and causal relations. *Cognition*, 35, 69-95.
- Goswami, U. & Brown, A. L. (1990). Higher-order structure and relational reasoning: Contrasting analogical and thematic relations. *Cognition*, 36, 207-226.
- Graf, P. & Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 501-518.
- Grimm, H. & Aktas, M. (2002). Entwicklungstests im Vorschulalter: Beurteilung ihrer Nützlichkeit durch praktisch tätige Psychologen. *Frühförderung interdisziplinär*, 21, 163-177.
- Gruber, K. (1992). *Konstruktion eines Itempools für die Dimensionen "Motorik", "Wahrnehmung" und "Gedächtnis" für 3- bis 6-jährige Kinder*. Unveröff. Dipl. Arbeit, Universität, Wien.
- Hagmann von Arx, P., Meyer, Ch. S. & Grob, A. (2008). Intelligenz- und Entwicklungsdiagnostik im deutschen Sprachraum. *Kindheit und Entwicklung*, 17, 232-242.
- Hale, C. M. & Tager-Flusberg, H. (2003). The influence of language on theory of mind: A training study. *Developmental Science*, 6, 346-359.
- Hansen, M. B. & Markman, E. M. (2005). Appearance questions can be misleading: A discourse-based account of the appearance-reality problem. *Cognitive Psychology*, 50, 233-263.
- Harris, P. L., German, T. & Mills, P. (1996). Children's use of counterfactual thinking in causal reasoning. *Cognition*, 61, 233-259.
- Harris, P. L. & Núñez, M. (1996). Understanding of permission rules by preschool children. *Child Development*, 67, 1572-1591.
- Harrison, P. L. (2009). Preschool assessment. In T. B. Gutkin & C. R. Reynolds (Eds.), *The handbook of school psychology* (4th ed.) (pp. 247-268). Hoboken: Wiley.
- Hayne, H. & Herbert, J. (2004). Verbal cues facilitate memory retrieval during infancy. *Journal of Experimental Child Psychology*, 89, 127-139.

-
- Hayne, H. & Simcock, G. (2009). Memory development in toddlers. In M. L. Courage & N. Cowan (Eds.), *The development of memory in infancy and childhood* (pp. 43-68). Hove: Psychology Press.
- Heckhausen, H. (1974). *Leistung und Chancengleichheit*. Göttingen: Hogrefe.
- Heckhausen, H. (1985). Emotionen im Leistungsverhalten aus ontogenetischer Sicht. In C. Eggers (Hrsg.), *Emotionalität und Motivation im Kindes- und Jugendalter* (S. 95-131). Frankfurt am Main: Fachbuchh. für Psychologie, Verl.-Abt.
- Heit, E. & Rotello, C. M. (2010). Relations between inductive reasoning and deductive reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36, 805-812.
- Hellbrügge, T. (Hrsg.) (1994). Münchener funktionelle Entwicklungsdiagnostik: Zweites und drittes Lebensjahr [Durchführungs-, Beurteilungs- und Interpretationshinweise] (4., korr. u. erw. Aufl.). München: Deutsche Akademie für Entwicklungs-Rehabilitation.
- Hellbrügge, T. & Pechstein, J. (1968). Entwicklungsphysiologische Tabellen für das Säuglingsalter. *Fortschritte der Medizin*, 86, 608-609.
- Holle, B. (2000). Die motorische und perzeptuelle Entwicklung des Kindes: Ein praktisches Lehrbuch für die Arbeit mit normalen und retardierten Kindern (4. Aufl.) (R. Heine, A. Schulze., Übers.). Weinheim: Beltz. (Original erschienen 1984: Normale og retarderede børns motoriske udvikling)
- Holodynski, M. (1992). *Leistungstätigkeit und soziale Interaktion: Ein tätigkeitstheoretisches Modell zur Entstehung der Leistungsmotivation*. Heidelberg: Asanger.
- Holodynski, M. (2006). Die Entwicklung der Leistungsmotivation im Vorschulalter: Soziale Bewertungen und ihre Auswirkung auf Stolz-, Scham- und Ausdauerreaktionen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 38, 2-17.
- Huang, C.-T., Heyes, C. & Charman, T. (2002). Infant's behavioural reenactment of "failed attempts": Exploring the roles of emulation learning, stimulus enhancement, and understanding of intentions. *Developmental Psychology*, 38, 840-855.
- Johnson, E. G. (1977). The development of color knowledge in preschool children. *Child Development*, 48, 308-311.
- Johnson, S. & Marlow, N. (2006). Developmental screen or developmental testing? *Early Human Development*, 82, 173-183.
- Kastner-Koller, U. & Deimann, P. (in Druck). Entwicklungstests. In: M. Amelang & L. F. Hornke (Hrsg.), *Enzyklopädie Psychologische Diagnostik*. Göttingen: Hogrefe.
- Kastner-Koller, U. & Deimann, P. (2002). *Wiener Entwicklungstest. Ein Verfahren zur Erfassung des allgemeinen Entwicklungsstandes bei Kindern von 3 bis 6 Jahren*. (2., überarb. u. neu norm. Aufl.). Göttingen: Hogrefe.
- Kaufman, E. L., Lord, M. W., Reese, T. W. & Volkman, J. (1949). The discrimination of visual number. *The American Journal of Psychology*, 62, 498-525.
- Kavanaugh, R. D., Eizenman, D. R. & Harris, P. L. (1997). Young children's understanding of pretense expressions of independent agency. *Developmental Psychology*, 33, 764-770.
- Kavsek, M. J. (2002). The perception of static subjective contours in infancy. *Child Development*, 73, 331-344.

- Kelly-Vance, L. & Ryalls, B. O. (2005). A systematic, reliable approach to play assessment in preschoolers. *School Psychology International*, 26, 398-412.
- Kelly-Vance, L., Ryalls, B. O. & Glover, K. G. (2002). The use of play assessment to evaluate the cognitive skills of two- and three-year-old children. *School Psychology International*, 23, 169-185.
- Kochanska, G., Murray, K. T. & Harlan, E. T. (2000). Effortful control in early childhood: Continuity and change, antecedents, and implications for social development. *Developmental Psychology*, 36, 220-232.
- Kubinger, K. D. (2009). *Psychologische Diagnostik: Theorie und Praxis psychologischen Diagnostizierens* (2., überarb. u. erw. Aufl.). Göttingen: Hogrefe.
- Kuchler, M. (in Druck). Die Entwicklung und Erprobung eines Itempools zur spielbasierten Erfassung der Entwicklung Zweijährige unter besonderer Berücksichtigung der Sprache, der Motorik und der sozialemotionalen Kompetenzen. Unveröff. Dipl. Arbeit, Universität, Wien.
- Kuchler, M., Sapper, E., Deimann, P. & Kastner-Koller, U. (2011). Manual zum Itempool zur spielbasierten Erfassung der Entwicklung Zweijähriger. Unveröff. Manuskript.
- Le Corre, M., Van de Walle, G., Brannon, E. M. & Carey, S. (2006). Re-visiting the competence/performance debate in the acquisition of the counting principles. *Cognitive Psychology*, 52, 130-169.
- Leslie, A. M. (1987). Pretense and representation: The origins of "theory of mind". *Psychological Review*, 94, 412-426.
- Leslie, A. M. & Roth, D. (1993). What autism teaches us about representation. In S. Baron-Cohen, H. Tager-Flusberg & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from autism* (pp. 335-366). New York: Oxford University Press.
- Lewis, M. (2003). The emergence of consciousness and its role in human development. *Annals of the New York Academy of Sciences*, 1001, 104-133.
- Lewis, M. & Brooks-Gunn, J. (1979). *Social cognition and the acquisition of self*. New York: Plenum Press.
- Lewis, M. & Ramsay, D. (2004). Development of self-recognition, personal pronoun use, and pretend play during the 2nd year. *Child Development*, 75, 1821-1831.
- Lienert, G. A. & Raatz, U. (1998). *Testaufbau und Testanalyse* (6. Aufl.). Weinheim: PVU.
- Lifter, K. (2000). Linking assessment to intervention for children with developmental disabilities or at-risk for developmental delay: The Developmental Play Assessment (DPA) instrument. In K. Gitlin-Weiner, A. Sandgrund & Ch. Schaefer (Eds.), *Play diagnosis and assessment* (2nd ed.) (pp. 228-261). Hoboken: John Wiley & Sons Inc.
- Lillard, A. S. (1993). Young children's conceptualization of pretense: Action or mental representational state? *Child Development*, 64, 372-386.
- Lillard, A. S. & Flavell, J. H. (1992). Young children's understanding of different mental states. *Developmental Psychology*, 28, 626-634.
- Linder, T. W. (1990). *Transdisciplinary play-based assessment: A functional approach to working with young children*. Baltimore: Brookes.

-
- Lissmann, I., Domsch, H. & Lohaus, A. (2006). Zur Stabilität und Validität von Entwicklungstestergebnissen im Alter von sechs Monaten bis zwei Jahren. *Kindheit und Entwicklung, 15*, 35-44.
- Lissmann, I., Korntheuer, P. & Lohaus, A. (2007). Assoziation zwischen der Kompetenz zur Kontingenzerkennung im Alter von drei und sechs Monaten und dem Entwicklungsstand im Alter von ein und zwei Jahren. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 39*, 187-195.
- Lohmann, H. & Tomasello, M. (2003). The role of language in the development of false belief understanding: A training study. *Child Development, 74*, 1130-1144.
- Lücking, A. & Scheithauer, H. (2006). Entwicklungstest sechs Monate bis sechs Jahre (ET 6-6). Franz Petermann, Iris A. Stein, Thorsten Macha (2., veränderte Auflage, 2004). [Frankfurt am Main: Harcourt Test Services, Preis des Gesamtsatzes: Euro 1.016,50]. *Diagnostica, 52*, 100-103.
- Macha, T., Proske, A. & Petermann, F. (2005). Validität von Entwicklungstests. *Kindheit und Entwicklung, 14*, 150-162.
- Melchers, P. & Preuß, U. (1994). *Kaufman assessment battery for children* (2. korr. Aufl.) (dt.-sprachige Fassung). Amsterdam: Swets & Zeitlinger.
- Meltzoff, A. N. (1988). Infant imitation after a 1-week delay: Long-term memory for novel acts and multiple stimuli. *Developmental Psychology, 24*, 470-476.
- Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology, 31*, 838-850.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81-97.
- Mogford-Bevan, K. (2000). The Play Observation Kit (POKIT): An observational assessment technique for young children. In K. Gitlin-Weiner, A. Sandgrund & Ch. Schaefer (Eds.), *Play diagnosis and assessment* (2nd ed.) (pp. 228-261). Hoboken: John Wiley & Sons Inc.
- Naggl, M. (2007). Entwicklungstest 6-6: Zweite Auflage. *Frühförderung interdisziplinär, 26*, 41-43.
- Neisworth, J. T. & Bagnato, S.J. (2004). The mismeasure of young children: The authentic assessment alternative. *Infants and Young Children, 7*, 198-212.
- Neumann, O. (1992). Theorien der Aufmerksamkeit: von Metaphern zu Mechanismen. *Psychologische Rundschau, 34*, 83-101.
- Nielsen, M. & Dissanayake, C. (2004). Pretend play, mirror self-recognition and imitation: A longitudinal investigation through the second year. *Infant Behavior and Development, 27*, 342-365.
- Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: MIT Press.
- Petermann, F. & Macha, T. (2003). Strategien in der testgestützten allgemeinen Entwicklungsdiagnostik. *Monatsschrift Kinderheilkunde, 151*, 6-13.
- Petermann, F. & Macha, T. (2005). Entwicklungsdiagnostik. *Kindheit und Entwicklung, 14*, 131-139.
- Petermann, F. & Macha, T. (2006). Trends in der modernen Entwicklungsdiagnostik. *Praxis Ergotherapie, 19*, 4-8.

- Petermann, F., Stein, I. A. & Macha, T. (2004). *Entwicklungsdiagnostik mit dem ET 6-6* (2. veränderte Aufl.). Frankfurt/Main: Harcourt Test Services.
- Posner, M. I. & Rothbart, M. K. (1992). Attentional mechanisms and conscious experience. In D. Milner & M. Ruggs (Eds.), *The neuropsychology of consciousness* (pp. 91-111). San Diego, CA: Academic Press.
- Posner, M. I. & Rothbart, M. K. (1998). Attention, self-regulation and consciousness. *Philosophical Transactions of the Royal Society, B*, 353, 1915-1927.
- Posner, M. I. & Rothbart, M. K. (2000). Developing mechanisms of self-regulation. *Development and Psychopathology*, 12, 427-441.
- Preston, P. (2005). *Testing children: A practitioner's guide to the assessment of mental development in infants and young children*. Göttingen: Hogrefe.
- Reuner, G. & Pietz, J. (2006). Entwicklungsdiagnostik im Säuglings- und Kleinkindalter. *Monatsschrift Kinderheilkunde*, 154, 305-313.
- Reznick, J. S. (2009). Working memory in infants and toddlers. In M. L. Courage & N. Cowan (Eds.), *The development of memory in infancy and childhood* (pp. 343-365). Hove: Psychology Press.
- Richards, C. A. & Sanderson, J. A. (1999). The role of imagination in facilitating deductive reasoning in 2-, 3- and 4-year-olds. *Cognition*, 72, B1-B9.
- Rips, L. J. (2001). Two kinds of reasoning. *Psychological Science*, 12, 129-134.
- Rotello, C. M. & Heit, E. (2009). Modeling the effects of argument length and validity on inductive and deductive reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 1317-1330.
- Rothbart, M. K., Derryberry, D. & Posner, M. I. (1994). A psychobiological approach to the development of temperament. In J. E. Bates & T. D. Wachs (Eds.), *Temperament: Individual differences at the interface of biology and behavior* (pp. 83-116). Washington, DC: American Psychological Association.
- Ruff, H. A. & Capozzoli, M. C. (2002). Development of attention and distractibility in the first 4 years of life. *Developmental Psychology*, 39, 877-890.
- Ruff, H. A. & Lawson, K. R. (1990). Development of sustained, focused attention in young children during free play. *Developmental Psychology*, 26, 85-93.
- Ruffman, T., Slade, L. & Crowe, E. (2002). The relation between children's and mothers' mental state language and theory-of-mind understanding. *Child Development*, 73, 734-751.
- Sapp, F., Lee, K. & Muir, D. (2000). Three-year-olds' difficulty with the appearance-reality distinction: Is it real or is it apparent? *Developmental Psychology*, 36, 547-560.
- Sarimski, K. (2001). Entwicklungstest von 6 Monaten bis 6 Jahren (ET 6-6). Franz Petermann, Iris A. Stein (2000). [Frankfurt: Swets Test Services, 120 Seiten, Preis des Handbuchs: DM 3,-, Preis des kompletten Tests: DM 1800,-]. *Diagnostica*, 47, 107-109.
- Sarimski, K. (2002). Entwicklungstest 6-6: Ein Gewinn für die Arbeit in der Frühförderung? *Frühförderung interdisziplinär*, 21, 178-181.
- Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 13, 501-518.

-
- Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind*. New York: Cambridge University Press.
- Singer-Freeman, K. E. & Goswami, U. (2001). Does half a pizza equal half a box of chocolates? Proportional matching in an analogy task. *Cognitive Development*, 16, 811-829.
- Silverman, I. W. & Rose, A. P. (1980). Subitizing and counting skills in 3-year-olds. *Developmental Psychology*, 16, 539-540.
- Slovan, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, 119, 3-22.
- Sparrow, S. S., Balla, D. A. & Cicchetti, D. V. (1984). *Vineland Adaptive Behavior Scale – Interview edition*. Circle Pines, MN: American Guidance Service.
- Sperling, G. (1960). The information available in brief visual presentation. *Psychological Monographs: General and Applied*, 74, 1-29.
- Squire, L. R. (1987). *Memory and brain*. New York: Oxford Univ. Press.
- Squire, L. R. (1992). Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. *Journal of Cognitive Neuroscience*, 4, 232-243.
- Squire, L. R. (1994). Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. In D. L. Schacter & E. Tulving (Eds.), *Memory systems* (pp. 203-232). Cambridge: MIT Press.
- Taylor, M. & Flavell, J. H. (1984). Seeing and believing: Children's understanding of distinction between appearance and reality. *Child Development*, 55, 1710-1720.
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381-403). New York: Academic Press.
- Tzuriel, D. (2000). Dynamic assessment of young children: Educational and intervention perspectives. *Educational Psychology Review*, 12, 385-435.
- Vig, S. (2007). Young children's object play: A window on development. *Journal of Developmental and Physical Disabilities*, 19, 201-215.
- Vig, S. & Sanders, M. (2007). Cognitive Assessment. In M. R. Brassard & A. E. Boehm, *Preschool assessment: Principles and practices* (pp. 383-420). New York: Guilford Press.
- Wagner-Menghin, M. M. (2003). Konzentrationstest. In K. D. Kubinger & R. S. Jäger (Hrsg.), *Schlüsselbegriffe der Psychologischen Diagnostik* (S. 282-286). Weinheim: Beltz.
- Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology* (pp. 135-151). Harmondsworth, UK: Penguin Books.
- Watson, D. G., Maylor, E. A. & Bruce, L. A. M. (2007). The role of eye movements in subitizing and counting. *Journal of Experimental Psychology: Human Perception and Performance*, 33, 1389-1399.
- Wechsler, D. (2002). *Wechsler Preschool and Primary Scales of Intelligence – Third Edition*. San Antonio, TX: The Psychological Corporation.
- Wellman, H. M., Cross, D. & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, 72, 655-684.

- Wellman, H. M. & Gelman, S.A. (1998). Knowledge acquisition in foundational domains. In D. Kuhn & R. S. Siegler (Eds.), *Handbook of child psychology* (5th ed.), Volume 2: Cognition, perception and language (pp. 523-573). New York: Wiley.
- Wellman, H. M. & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, 75, 523-541.
- Wimmer, H. & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13, 103-128.
- Witzlack, G. (2001). Spielanalytische Entwicklungsdiagnostik – Historische Wurzeln und Ansätze in der DDR. In D. Sturzbecher (Hrsg.), *Spielbasierte Befragungstechniken. Interaktionsdiagnostische Verfahren für Begutachtung, Beratung und Forschung* (S. 218-227). Goettingen: Hogrefe.
- Wynn, K. (1990). Children's understanding of counting. *Cognition*, 36, 155-193.
- Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology*, 24, 220-251.
- Zelazo, P. D., Frye, D. & Rapus T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development*, 11, 37-63.
- Zimmermann, I. L., Steiner, V. G. & Pond, R. E. (2002). *Preschool Language Scale – Fourth Edition*. San Antonio, TX: The Psychological Corporation.
- Ziv, M. & Frye, D. (2003). The relation between desire and false belief in children's theory of mind: No satisfaction? *Developmental Psychology*, 39, 859-876.

IV. ANHANG

A) Elternbrief und Einverständniserklärung

UNIVERSITÄT  WIEN

Fakultät für Psychologie

**Institut für Entwicklungspsychologie
und Psychologische Diagnostik**

Liebe Eltern!

Wir sind Studentinnen der Psychologie an der Universität Wien und schreiben derzeit unsere Diplomarbeiten im Arbeitsbereich Entwicklungspsychologie bei Frau Dr. Pia Deimann und Frau Dr. Ursula Kastner-Koller.

Unsere Diplomarbeiten beschäftigen sich mit der Entwicklung von Kindern im Alter von 24 bis 36 Monaten. Wir wollen einen spielerischen Zugang zur Feststellung der Entwicklung so junger Kinder erproben. Ziel unserer Arbeit ist es, ein genaues Bild über die Fähigkeiten normal entwickelter Kinder in dieser Altersgruppe zu bekommen.

Die Spielsituation findet in einem großen Spielzimmer des Instituts für Entwicklungspsychologie und Psychologische Diagnostik der Universität Wien statt (1010 Wien, Liebiggasse 5/1. Stock). Die Begleitperson bleibt selbstverständlich anwesend. Wir spielen mit dem Kind bekannte und neue Spiele, wobei mit einer Dauer von 1 bis 2 Stunden zu rechnen ist. Das gesamte Spiel wird auf Video aufgezeichnet und anonymisiert im Rahmen der Studie ausgewertet.

Wenn Sie möchten, können Sie gerne eine Kopie der Aufnahme erhalten.

Wir würden uns freuen, wenn Sie und Ihr Kind an der Untersuchung teilnehmen! Falls Sie Interesse haben, kontaktieren Sie uns bitte so bald wie möglich entweder telefonisch oder per E-Mail, damit genügend Zeit zur Planung und Koordination der Termine bleibt:

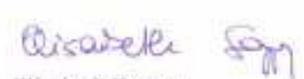
Mara Kuchler: 0699/10085664 mara.kuchler@univie.ac.at
oder
Elisabeth Sapper: 0699/12387533 elisabeth.sapper@gmx.at

Sie können sich natürlich auch gerne bei uns melden, wenn Sie noch Fragen haben oder sich unverbindlich über das Projekt informieren möchten.

Mit freundlichen Grüßen



Mara Kuchler



Elisabeth Sapper

Abbildung A.1. Elternbrief

Liebe Eltern,

da es für uns nicht möglich ist, während des Spiels mit Ihrem Kind alles zu erfassen, was für die Entwicklung unseres Verfahrens relevant ist, bitten wir Sie, die Spielsituation mit dem Kind auf Video aufzeichnen zu dürfen.

Wir versichern Ihnen, dass außer uns und unseren Diplomarbeitsbetreuern, Ass.-Prof. Dr. Pia Deimann und Ass.-Prof. Dr. Ursula Kastner-Koller, niemand die Videos sehen wird und diese spätestens nach Beendigung unserer Diplomarbeiten gelöscht werden.

Mara Kuchler und Elisabeth Sapper

Ich, _____, erkläre mich damit einverstanden, dass meine Tochter / mein Sohn _____, geboren am _____ im Rahmen der Mitwirkung an der Diplomarbeitsstudie von Frau Kuchler und Frau Sapper auf Video aufgezeichnet werden darf.

Datum

Unterschrift

Abbildung A.2. Einverständniserklärung für die Videoaufzeichnung während der Testung

B) Reliabilitätstabellen

Tabelle B.1. Reliabilitätsberechnung Subtest „Theory of mind“, „in andere hineinversetzen“

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala- Korrelation	Cronbachs Alpha, wenn Item weggelassen
Theory of mind - Hineinversetzen - Schmutzige Hände	1,15	2,134	,345	,819
Theory of mind - Hineinversetzen - hungrig	1,15	1,818	,637	,729
Theory of mind - Hineinversetzen - schlafen	1,20	1,958	,568	,753
Theory of mind - Hineinversetzen - Pflaster	,95	1,734	,583	,749
Theory of mind - Hineinversetzen - durstig	1,15	1,713	,747	,692

Anmerkung: Cronbach Alpha = .791

Tabelle B.2. Reliabilitätsberechnung Subtest „Theory of mind“ gesamt

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala- Korrelation	Cronbachs Alpha, wenn Item weggelassen
Theory of mind - Hineinversetzen - Schmutzige Hände	1,675	2,191	,370	,713
Theory of mind - Hineinversetzen - hungrig	1,675	1,928	,608	,640
Theory of mind - Hineinversetzen - schlafen	1,725	2,065	,544	,663
Theory of mind - Hineinversetzen - Pflaster	1,475	1,855	,547	,658
Theory of mind - Hineinversetzen - durstig	1,675	1,823	,713	,604
Theory of mind - Als-ob-Spiel	1,400	2,779	-,018	,791

Anmerkung: Cronbach Alpha = .725

Tabelle B.3. Reliabilitätsberechnung Subtest „numerisches Wissen“, aktive und passive Mengenerfassung

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala- Korrelation	Cronbachs Alpha, wenn Item weggelassen
Numerisches Wissen - Geben von 1	1,65	2,661	,000	,711
Numerisches Wissen - Geben von 2	2,00	1,789	,482	,645
Numerisches Wissen - Geben von 3	2,45	1,945	,478	,644
Numerisches Wissen - Geben von 4	2,60	2,568	,059	,717
Numerisches Wissen - Benennen von 2	2,30	1,695	,570	,616
Numerisches Wissen - Benennen von 3	2,40	1,937	,426	,659
Numerisches Wissen - Benennen von 4	2,55	2,155	,454	,655
Numerisches Wissen - Benennen von 5	2,60	2,253	,533	,654

Anmerkung: Cronbach Alpha = .696

Tabelle B.4. Reliabilitätsberechnung Subtest „numerisches Wissen“, aktive und passive Mengenerfassung, unter Ausschluss der Items „Geben von 1“ und „Geben von 4“

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala- Korrelation	Cronbachs Alpha, wenn Item weggelassen
Numerisches Wissen - Geben von 2	,95	1,734	,461	,709
Numerisches Wissen - Geben von 3	1,40	1,937	,405	,720
Numerisches Wissen - Benennen von 2	1,25	1,566	,623	,651
Numerisches Wissen - Benennen von 3	1,35	1,818	,461	,705
Numerisches Wissen - Benennen von 4	1,50	2,053	,477	,705
Numerisches Wissen - Benennen von 5	1,55	2,155	,553	,705

Anmerkung: Cronbach Alpha = .738

Tabelle B.5. Reliabilitätsberechnung Subtest „numerisches Wissen“, „hergeben“ und „zählt dabei“

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala- Korrelation	Cronbachs Alpha, wenn Item weggelassen
Numerisches Wissen - Geben von 2 - wie hergegeben	1,225	2,670	,263	,718
Numerisches Wissen - Geben von 3 - wie hergegeben	1,575	2,086	,669	,629
Numerisches Wissen - Geben von 4 - wie hergegeben	1,700	2,353	,528	,664
Numerisches Wissen - Geben von 2 - zählt dabei	1,925	3,191	,000	,725
Numerisches Wissen - Geben von 3 - zählt dabei	1,875	2,760	,514	,683
Numerisches Wissen - Geben von 4 - zählt dabei	1,925	3,191	,000	,725
Numerisches Wissen - Benennen von 2 - zählt dabei	1,825	3,270	-,156	,768
Numerisches Wissen - Benennen von 3 - zählt dabei	1,725	2,539	,370	,697
Numerisches Wissen - Benennen von 4 - zählt dabei	1,775	2,302	,679	,637
Numerisches Wissen - Benennen von 5 - zählt dabei	1,775	2,302	,679	,637

Anmerkung: Cronbach Alpha = .716

Tabelle B.6. Reliabilitätsberechnung Subtest „numerisches Wissen“, „hergeben“ und „zählt dabei“, unter Ausschluss der Items „hergeben: Geben von 2“, „zählt dabei: Geben von 2“, „zählt dabei: Geben von 4“ und „zählt dabei: Benennen von 2“

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala- Korrelation	Cronbachs Alpha, wenn Item weggelassen
Numerisches Wissen - Geben von 3 - wie hergegeben	,775	1,749	,696	,780
Numerisches Wissen - Geben von 4 - wie hergegeben	,900	2,016	,534	,816
Numerisches Wissen - Geben von 3 - zählt dabei	1,075	2,349	,603	,813
Numerisches Wissen - Benennen von 3 - zählt dabei	,925	2,139	,421	,840
Numerisches Wissen - Benennen von 4 - zählt dabei	,975	1,934	,731	,774
Numerisches Wissen - Benennen von 5 - zählt dabei	,975	1,934	,731	,774

Anmerkung: Cronbach Alpha = .829

Tabelle B.7. Reliabilitätsberechnung Subtest „numerisches Wissen“ gesamt, unter Ausschluss der Items „Geben von 1“, „Geben von 4“, „hergeben: Geben von 2“, „zählt dabei: Geben von 2“, „zählt dabei: Geben von 4“ und „zählt dabei: Benennen von 2“

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala- Korrelation	Cronbachs Alpha, wenn Item weggelassen
Numerisches Wissen - Geben von 2	2,075	7,718	,485	,869
Numerisches Wissen - Geben von 3	2,525	8,118	,423	,871
Numerisches Wissen - Benennen von 2	2,375	7,207	,696	,853
Numerisches Wissen - Benennen von 3	2,475	7,749	,537	,864
Numerisches Wissen - Benennen von 4	2,625	8,260	,521	,865
Numerisches Wissen - Benennen von 5	2,675	8,560	,509	,867
Numerisches Wissen - Geben von 3 - wie hergegeben	2,375	7,286	,713	,851
Numerisches Wissen - Geben von 4 - wie hergegeben	2,500	7,868	,534	,864
Numerisches Wissen - Geben von 3 - zählt dabei	2,675	8,665	,426	,870
Numerisches Wissen - Benennen von 3 - zählt dabei	2,525	7,802	,569	,862
Numerisches Wissen - Benennen von 4 - zählt dabei	2,575	7,639	,742	,852
Numerisches Wissen - Benennen von 5 - zählt dabei	2,575	7,744	,685	,855

Anmerkung: Cronbach Alpha = .872

Tabelle B.8. Reliabilitätsberechnung Subtest „deduktives Denken“

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala- Korrelation	Cronbachs Alpha, wenn Item weggelassen
dD-Buch - Katze mit Besteck	1,85	1,608	,354	,550
dD-Buch - Schaf auf Fahrrad	1,80	1,642	,320	,568
dD-Buch - Elefant mit Flügeln	2,10	2,095	,053	,673
dD-Buch - Hase mit Eis	1,75	1,566	,391	,529
dD-Buch - Hund im Auto	1,70	1,274	,705	,336

Anmerkung: Cronbach Alpha = .602

Tabelle B.9. Reliabilitätsberechnung Subtest „deduktives Denken“, unter Ausschluss des Items „Elefant mit Flügeln“

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala- Korrelation	Cronbachs Alpha, wenn Item weggelassen
dD-Buch - Katze mit Besteck	1,65	1,292	,467	,599
dD-Buch - Schaf auf Fahrrad	1,60	1,411	,346	,677
dD-Buch - Hase mit Eis	1,55	1,418	,342	,679
dD-Buch - Hund im Auto	1,50	1,105	,697	,436

Anmerkung: Cronbach Alpha = .673

Tabelle B.10. Reliabilitätsberechnung Subtest „visuelle Wahrnehmung“, „Farbdifferenzierung“

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala- Korrelation	Cronbachs Alpha, wenn Item weggelassen
visuelle Wahrnehmung - Farbe - aktiv - rot	6,65	38,345	,770	,950
visuelle Wahrnehmung - Farbe - aktiv - grün	6,70	38,642	,725	,951
visuelle Wahrnehmung - Farbe - aktiv - gelb	6,70	37,800	,867	,949
visuelle Wahrnehmung - Farbe - aktiv - orange	6,80	38,484	,787	,950
visuelle Wahrnehmung - Farbe - aktiv - blau	6,75	39,145	,653	,952
visuelle Wahrnehmung - Farbe - aktiv - weiß	6,80	37,853	,899	,948
visuelle Wahrnehmung - Farbe - aktiv - lila	6,80	38,589	,769	,950
visuelle Wahrnehmung - Farbe - passiv - rot	6,60	39,726	,546	,954
visuelle Wahrnehmung - Farbe - passiv - grün	6,75	38,408	,777	,950
visuelle Wahrnehmung - Farbe - passiv - gelb	6,80	37,853	,899	,948
visuelle Wahrnehmung - Farbe - passiv - orange	6,80	39,116	,678	,952
visuelle Wahrnehmung - Farbe - passiv - blau	6,70	38,221	,796	,950
visuelle Wahrnehmung - Farbe - passiv - weiß	6,55	39,418	,607	,953
visuelle Wahrnehmung - Farbe - passiv - lila	6,60	39,095	,650	,952
visuelle Wahrnehmung - Sortieren nach Farben - rot	6,90	40,095	,571	,953
visuelle Wahrnehmung - Sortieren nach Farben - grün	6,85	40,134	,528	,954
visuelle Wahrnehmung - Sortieren nach Farben - gelb	6,80	39,116	,678	,952
visuelle Wahrnehmung - Sortieren nach Farben - orange	7,00	40,316	,656	,952

Anmerkung: Cronbach Alpha = .954

Tabelle B.11. Reliabilitätsberechnung Subtest „visuelle Wahrnehmung“, „Formdifferenzierung“

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala- Korrelation	Cronbachs Alpha, wenn Item weggelassen
Visuelle Wahrnehmung - Formdifferenzierung 1	,45	,261	,798	
Visuelle Wahrnehmung - Formdifferenzierung 2	,45	,261	,798	

Anmerkung: Cronbach Alpha = .888

Tabelle B.12. Reliabilitätsberechnung Subtest „visuelle Wahrnehmung“, „Größendifferenzierung“

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala- Korrelation	Cronbachs Alpha, wenn Item weggelassen
Visuelle Wahrnehmung - Größendifferenzierung 1	,20	,168	,866	
Visuelle Wahrnehmung - Größendifferenzierung 2	,25	,197	,866	

Anmerkung: Cronbach Alpha = .927

Tabelle B.13. Reliabilitätsberechnung Subtest „visuelle Wahrnehmung“ gesamt

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala- Korrelation	Cronbachs Alpha, wenn Item weggelassen
visuelle Wahrnehmung - Farbe - aktiv - rot	8,00	57,579	,744	,960
visuelle Wahrnehmung - Farbe - aktiv - grün	8,05	57,734	,727	,960
visuelle Wahrnehmung - Farbe - aktiv - gelb	8,05	56,787	,856	,959
visuelle Wahrnehmung - Farbe - aktiv - orange	8,15	57,397	,808	,959
visuelle Wahrnehmung - Farbe - aktiv - blau	8,10	58,095	,690	,961
visuelle Wahrnehmung - Farbe - aktiv - weiß	8,15	56,871	,884	,958
visuelle Wahrnehmung - Farbe - aktiv - lila	8,15	57,608	,779	,960
visuelle Wahrnehmung - Farbe - passiv - rot	7,95	58,892	,572	,962
visuelle Wahrnehmung - Farbe - passiv - grün	8,10	57,568	,762	,960
visuelle Wahrnehmung - Farbe - passiv - gelb	8,15	56,871	,884	,958
visuelle Wahrnehmung - Farbe - passiv - orange	8,15	58,345	,675	,961
visuelle Wahrnehmung - Farbe - passiv - blau	8,05	56,997	,827	,959
visuelle Wahrnehmung - Farbe - passiv - weiß	7,90	58,516	,632	,961
visuelle Wahrnehmung - Farbe - passiv - lila	7,95	58,050	,684	,961
visuelle Wahrnehmung - Sortieren nach Farben - rot	8,25	59,671	,548	,962
visuelle Wahrnehmung - Sortieren nach Farben - grün	8,20	59,853	,489	,963
visuelle Wahrnehmung - Sortieren nach Farben - gelb	8,15	58,555	,646	,961
visuelle Wahrnehmung - Sortieren nach Farben - orange	8,35	59,818	,649	,961
Visuelle Wahrnehmung - Formdifferenzierung 1	8,05	57,418	,770	,960
Visuelle Wahrnehmung - Formdifferenzierung 2	8,05	57,629	,741	,960
Visuelle Wahrnehmung - Größendifferenzierung 1	8,25	58,724	,692	,961
Visuelle Wahrnehmung - Größendifferenzierung 2	8,30	58,853	,732	,960

Anmerkung: Cronbach Alpha = .962

Tabelle B.14. Reliabilitätsberechnung Subtest „Gedächtnis“, „phonologisches Gedächtnis“

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala- Korrelation	Cronbachs Alpha, wenn Item weggelassen
Phonologisches Gedächtnis - Gib mir A	,25	,303	,107	,417
Phonologisches Gedächtnis - Gib mir A + B	1,00	,105	,395	,075
Phonologisches Gedächtnis - Gib mir A + B + C	1,15	,239	,409	,132
Phonologisches Gedächtnis - Gib mir A + B + C + D	1,20	,379	,000	,438

Anmerkung: Cronbach Alpha = .389

Tabelle B.15. Reliabilitätsberechnung Subtest „Gedächtnis“, „phonologisches Gedächtnis“, unter Ausschluss der Items „Gib mir A“ und „Gib mir A+B+C+D“

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item- Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen
Phonologisches Gedächtnis - Gib mir A + B	,05	,050	,459	
Phonologisches Gedächtnis - Gib mir A + B + C	,20	,168	,459	

Anmerkung: Cronbach Alpha = .557

Tabelle B.16. Reliabilitätsberechnung Subtest „Gedächtnis“, „visuelles Gedächtnis“

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala- Korrelation	Cronbachs Alpha, wenn Item weggelassen
Visuelles Gedächtnis - Memory 1	1,55	3,418	,827	,845
Visuelles Gedächtnis - Memory 2	1,65	3,713	,703	,866
Visuelles Gedächtnis - Memory 3	1,60	3,516	,792	,851
Visuelles Gedächtnis - Memory 4	1,60	3,411	,862	,839
Visuelles Gedächtnis - Teddybär	1,75	4,513	,302	,920
Visuelles Gedächtnis - Haus	1,60	3,621	,723	,863

Anmerkung: Cronbach Alpha = .887

Tabelle B.17. Reliabilitätsberechnung Subtest „Gedächtnis“ gesamt, unter Ausschluss der Items „Gib mir A“ und „Gib mir A+B+C+D“

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala- Korrelation	Cronbachs Alpha, wenn Item weggelassen
Phonologisches Gedächtnis - Gib mir A + B	2,00	5,474	,658	,866
Phonologisches Gedächtnis - Gib mir A + B + C	2,15	6,555	,262	,895
Visuelles Gedächtnis - Memory 1	1,80	5,011	,730	,858
Visuelles Gedächtnis - Memory 2	1,90	5,253	,664	,865
Visuelles Gedächtnis - Memory 3	1,85	4,871	,831	,846
Visuelles Gedächtnis - Memory 4	1,85	4,766	,889	,839
Visuelles Gedächtnis - Teddybär	2,00	6,000	,366	,892
Visuelles Gedächtnis - Haus	1,85	5,082	,718	,859

Anmerkung: Cronbach Alpha = .882

Tabelle B.18. Reliabilitätsberechnung Subtest „Gedächtnis“ gesamt, unter Ausschluss der Items „Gib mir A“, „Gib mir A+B+C“ und „Gib mir A+B+C+D“

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala- Korrelation	Cronbachs Alpha, wenn Item weggelassen
Phonologisches Gedächtnis - Gib mir A + B	1,95	5,208	,629	,887
Visuelles Gedächtnis - Memory 1	1,75	4,618	,780	,868
Visuelles Gedächtnis - Memory 2	1,85	4,871	,705	,878
Visuelles Gedächtnis - Memory 3	1,80	4,589	,823	,863
Visuelles Gedächtnis - Memory 4	1,80	4,484	,884	,855
Visuelles Gedächtnis - Teddybär	1,95	5,734	,332	,915
Visuelles Gedächtnis - Haus	1,80	4,800	,707	,878

Anmerkung: Cronbach Alpha = .895

Tabelle B.19. Reliabilitätsberechnung Subtest „Aufmerksamkeit“, „längste Aufmerksamkeitsspanne (Min.)“ und „längste Aufmerksamkeitsspanne (S.)“

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala- Korrelation	Cronbachs Alpha, wenn Item weggelassen
längste Aufmerksamkeitsspanne (Minuten) - Bausteine	18,575	149,191	,369	,787
längste Aufmerksamkeitsspanne (Minuten) - Zeichnen	15,700	160,168	,106	,845
längste Aufmerksamkeitsspanne (Minuten) - Magnetbuch	17,600	121,095	,761	,715
längste Aufmerksamkeitsspanne (Minuten) - ToM-Buch	18,500	131,026	,726	,728
längste Aufmerksamkeitsspanne (Minuten) - dD-Buch	17,775	153,907	,430	,776
längste Aufmerksamkeitsspanne (Seiten) - Magnetbuch	18,500	126,000	,826	,711
längste Aufmerksamkeitsspanne (Seiten) - ToM-Buch	19,450	147,629	,817	,743
längste Aufmerksamkeitsspanne (Seiten) - dD-Buch	16,700	154,905	,304	,795

Anmerkung: Cronbach Alpha = .789

Tabelle B.20. Reliabilitätsberechnung Subtest „Aufmerksamkeit“, „längste Aufmerksamkeitsspanne (Min.)“ und „längste Aufmerksamkeitsspanne (S.)“ unter Ausschluss des Items „längste Aufmerksamkeitsspanne (Min.) - Zeichnen“

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item- Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen
längste Aufmerksamkeitsspanne (Minuten) - Bausteine	13,875	128,839	,384	,858
längste Aufmerksamkeitsspanne (Minuten) - Magnetbuch	12,900	108,068	,678	,812
längste Aufmerksamkeitsspanne (Minuten) - ToM-Buch	13,800	115,589	,672	,812
längste Aufmerksamkeitsspanne (Minuten) - dD-Buch	13,075	127,797	,575	,828
längste Aufmerksamkeitsspanne (Seiten) - Magnetbuch	13,800	109,063	,811	,789
längste Aufmerksamkeitsspanne (Seiten) - ToM-Buch	14,750	127,724	,844	,808
längste Aufmerksamkeitsspanne (Seiten) - dD-Buch	12,000	127,947	,433	,849

Anmerkung: Cronbach Alpha = .845

Tabelle B.21. Reliabilitätsberechnung Subtest „Aufmerksamkeit“, „lässt sich zurückholen“ und „alles aufmerksam bearbeitet“

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala- Korrelation	Cronbachs Alpha, wenn Item weggelassen
lässt sich wieder zurückholen - Bausteine	2,750	5,329	,605	,793
lässt sich wieder zurückholen - Zeichnen	2,400	5,411	,390	,817
lässt sich wieder zurückholen - Magnetbuch	2,800	5,484	,606	,796
lässt sich wieder zurückholen - ToM-Buch	2,800	5,168	,733	,781
lässt sich wieder zurückholen - dD-Buch	2,400	5,411	,427	,812
komplett aufmerksam bearbeitet - Bausteine	2,800	5,432	,561	,798
komplett aufmerksam bearbeitet - Zeichnen	2,450	5,418	,353	,824
komplett aufmerksam bearbeitet - Magnetbuch	2,850	5,661	,525	,803
komplett aufmerksam bearbeitet - ToM-Buch	2,800	5,168	,733	,781
komplett aufmerksam bearbeitet - dD-Buch	2,500	5,421	,354	,824

Anmerkung: Cronbach Alpha = .819

Tabelle B.22. Reliabilitätsberechnung Subtest „Aufmerksamkeit“ gesamt, unter Ausschluss des Items „längste Aufmerksamkeitsspanne (Min.), Zeichnen“

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala- Korrelation	Cronbachs Alpha, wenn Item weggelassen
längste Aufmerksamkeitsspanne (Minuten) - Bausteine	16,825	188,534	,373	,849
längste Aufmerksamkeitsspanne (Minuten) - Magnetbuch	15,850	161,029	,696	,824
längste Aufmerksamkeitsspanne (Minuten) - ToM-Buch	16,750	168,408	,719	,819
längste Aufmerksamkeitsspanne (Minuten) - dD-Buch	16,025	184,802	,596	,828
längste Aufmerksamkeitsspanne (Seiten) - Magnetbuch	16,750	162,934	,813	,811
längste Aufmerksamkeitsspanne (Seiten) - ToM-Buch	17,700	184,326	,872	,815
längste Aufmerksamkeitsspanne (Seiten) - dD-Buch	14,950	184,682	,462	,841
lässt sich wieder zurückholen - Bausteine	18,450	216,261	,688	,841
lässt sich wieder zurückholen - Zeichnen	18,100	218,305	,384	,844
lässt sich wieder zurückholen - Magnetbuch	18,500	216,789	,737	,841
lässt sich wieder zurückholen - ToM-Buch	18,500	216,474	,688	,841
lässt sich wieder zurückholen - dD-Buch	18,100	216,121	,574	,841
komplett aufmerksam bearbeitet - Bausteine	18,500	218,053	,540	,843
komplett aufmerksam bearbeitet - Zeichnen	18,150	218,555	,344	,844
komplett aufmerksam bearbeitet - Magnetbuch	18,550	217,839	,671	,842
komplett aufmerksam bearbeitet - ToM-Buch	18,500	216,474	,688	,841
komplett aufmerksam bearbeitet - dD-Buch	18,200	216,747	,467	,842

Anmerkung: Cronbach Alpha = .845

C.1) Abstract deutsch

Das Ziel der vorliegenden Studie war es, einen Itempool für die Erfassung der allgemeinen Entwicklung Zweijähriger zu schaffen. Da die Anwendung klassischer diagnostischer Verfahren in dieser Altersgruppe durch viele Nachteile gekennzeichnet ist (siehe zum Beispiel Bagnato, 2005; Kelly-Vance et al., 2002) und sich auch die Durchführung spielbasierter Ansätze als nicht optimal erweist (Athanasίου, 2007; Linder, 1990), wurde versucht, die beiden miteinander zu verbinden, um einerseits ihre Vorteile zu nutzen und andererseits ihre Nachteile möglichst gering zu halten. In dieser Studie geht es dabei um die Entwicklungsbereiche „kognitive Entwicklung“, „visuelle Wahrnehmung“, „Gedächtnis“ und „Arbeitshaltungen“. Für weitere Bereiche siehe Kuchler (in Druck). Der entwickelte Itempool (für eine genaue Beschreibung siehe Kuchler et al., 2011) wurde im Zuge der Studie an 20 Kindern erprobt. Insgesamt erwiesen sich die einzelnen Itemgruppen als gut handhabbar, wenngleich die oft sehr geringe Aufmerksamkeit, Ausdauer und Kooperationsbereitschaft der Kinder zentrale Probleme bei der Durchführung der einzelnen Aufgaben darstellten, weswegen es viele fehlende Werte gab. Daher sollte es das Ziel weiterführender Untersuchungen sein, die Art der Erfassung der Entwicklungsbereiche in einigen Fällen so zu überarbeiten und zu verbessern, dass sie den Möglichkeiten und Grenzen der Aufmerksamkeit Zweijähriger gerecht wird. Auch sollten einige wenige Items aufgrund ihrer Itemschwierigkeiten noch einmal überarbeitet werden. Die Reliabilitäten der meisten Itemgruppen erwiesen sich als gut, wenngleich einige Items über zu geringe Itemtrennschärfen verfügen und daher noch einer genaueren Betrachtung bedürfen. Ein Teil der Aufgabengruppen ist in der Lage, den Altersverlauf des ihnen zugehörigen Entwicklungsbereichs darzustellen, bei manchen hingegen besteht keine signifikante Korrelation der Ergebnisse mit dem Alter. Alles in allem stellt der vorliegende Itempool jedoch eine annehmbare Grundlage dar, um das langfristige Ziel der Entwicklung eines Testverfahrens für die Erfassung der Entwicklung Zweijähriger zu erreichen.

C.1) Abstract english

This investigation was conducted in order to create an item pool for the acquisition of the general development of two-year-olds. Because of the fact that usually conducted norm-referenced testing in this age-range is afflicted with many disadvantages (see Bagnato, 2005; Kelly-Vance et al., 2002) and play-based assessments are subject to handicaps too (Athanasidou, 2007; Linder, 1990) the goal was to connect these two approaches in order to use their advantages on the one hand and to decrease their disadvantages on the other hand. This study contains the measurement of the cognitive development, the visual perception, the memory and the working attitudes. For other developmental domains see Kuchler (in press). The created item pool (for a description see Kuchler et al., 2011) was tested on 20 children. All in all the items were good to handle, although the attention and cooperation of the children emerged to be a huge problem in the implementation of the tasks. That's why there is many data missing. Because of that the mode of measurement should be reconsidered and improved in some cases, so that it meets the possibilities and limitations of two-year-olds. Also, some items should be reconsidered because of their item-difficulty. The reliability of most of the items is sufficient or good, but the item-total correlations sometimes are too low, therefore a few items need to be overworked. Several tasks are able to display the age-trend of the measured developmental domain, but some do not show a significant correlation of results and age. All in all the presented item pool provides an acceptable fundament in order to reach the long-ranging goal of developing a norm-reverenced but play-based assessment to assess the general development of two-year-olds.

D) Curriculum Vitae

Elisabeth Sapper

geboren am in	11.07.1987 Wien
Staatsbürgerschaft	Österreich
Familienstand	ledig
Lenkerberechtigung	B
Hobbys	Lesen, Snowboarden, Malen

Schulausbildung	1993-1997 Volksschule Loosdorf 1997-2005 Stiftsgymnasium Melk, humanistischer Zweig (Maturaabschluss mit ausgezeichnetem Erfolg)
-----------------	---

derzeitige Tätigkeit(en) Oktober 2005 – voraussichtlich Oktober 2011:

Diplomstudium Psychologie
(Schwerpunkt: Angewandte Kinder- und
Jugendpsychologie)

seit Juli 2011:

Rezeptionistin im *A & O Hotel und Hostel Wien
Stadthalle*

März 2007 – Juni 2011:

Praktikantin in der *Allianz Investmentbank AG,
Abteilung Kontoführung und Zahlungsverkehr:*

- Verwaltung von Kundenkonten
 - Bearbeitung von Kundenanfragen über Telefon /
E-Mail
 - Tätigkeiten im Bereich Zahlungsverkehr
 - Postbearbeitung
 - Archivierung
-

Studienrelevante Praxis Oktober 2009 – Juni 2010:

Absolvierung der Lehrveranstaltung „*Kinder- und Jugendpsychologische Interventionen*“:
Lerntherapeutische Betreuung eines Jugendlichen mit Lese-Rechtschreib-Störung, 2x pro Woche über ein gesamtes Schuljahr hinweg

März 2009 – Februar 2010:

Absolvierung der Lehrveranstaltungen „*Kinder- und Jugend-psychologische Beratungspraxis – Schwerpunkt: Schulkinder, Jugendliche*“ und „*Kinder- und Jugendpsychologische Beratungspraxis – Schwerpunkt: Klein-, Vorschulkinder*“:
Durchführung bzw. Beobachtung der Durchführung von Statusdiagnostik und Erstellung von Gutachten in vier Fällen

März – Juni 2009:

Absolvierung der Lehrveranstaltung „*Hochbegabendiagnostik*“:
Durchführung von Hochbegabendiagnostik und Erstellung von Gutachten bei Volksschulkindern in zwei Fällen

September – Oktober 2008:

6-Wochen-Pflichtpraktikum in der *Universitätsklinik für Psychiatrie des Kindes- und Jugendalters* des AKH Wien:

- Durchführung und Auswertung testpsychologischer Untersuchungen
- Verhaltensbeobachtungen im ambulanten und stationären Bereich
- Anwendung von Intelligenz- und Leistungstests, sowie neuropsychologischer Testverfahren zur differenzierten Teilleistungsanalyse und Persönlichkeitsdiagnostik
- Erstellung von klinisch-psychologischen Vorbefunden
- Durchführung von Anamnesegesprächen
- Teilnahme an Testbesprechungs- und Beratungsgesprächen

Besondere Kenntnisse gute Anwenderkenntnisse der Microsoft-Office-Programme,
10-Finger-Tastaturschreibsystem, Statistikprogramm SPSS, Führerschein B

Fremdsprachkenntnisse sehr gute Englischkenntnisse in Wort und Schrift