



universität
wien

Diplomarbeit

Titel der Diplomarbeit

Die Überarbeitung und testtheoretische Überprüfung
eines spielbasierten Itempools zur Erfassung der
Entwicklung zweijähriger Kinder
unter besonderer Berücksichtigung der visuellen Wahrnehmung

Verfasserin

Viktoria Magdalena Kronberger

Angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag. rer. nat.)

Wien, im September 2013

Studienkennzahl: A 298

Studienrichtung: Psychologie

Betreuer: Ass.-Prof. Dr. Ursula Kastner-Koller

DANKSAGUNG

An erster Stelle möchte ich
Frau Ass.-Prof. Dr. Kastner-Koller und Frau Ass.-Prof. Dr. Deimann
meinen großen Dank für die professionelle Betreuung aussprechen!

Sigrid Punz,
meiner Diplomarbeitskollegin und nun sehr guten Freundin,
danke ich für die engagierte und tolle Zusammenarbeit!

Ein weiterer Dank gebührt
den Eltern und Kindern,
die sich die Zeit genommen haben,
um an dieser Diplomarbeit mitzuwirken.

Ein großer Dank gilt auch
meiner Mama Andrea und meinem Freund Christoph,
die die Zeit aufgebracht haben, meine Diplomarbeit Korrektur zu lesen und
mir immer mit Rat und Tat zur Seite standen.

DANKE

Inhaltsverzeichnis

Einleitung.....	11
1. Grundlagen der Entwicklungsdiagnostik.....	15
1.1. Definition des Entwicklungsbegriffs.....	15
1.2. Definition der Entwicklungsdiagnostik.....	16
1.3. Aufgaben und Ziele der Entwicklungsdiagnostik	17
1.4. Geschichte der Entwicklungsdiagnostik	18
1.5. Methoden in der Entwicklungsdiagnostik	19
1.5.1. Verhaltensbeobachtung der Kinder.....	19
1.5.2. Befragung der Bezugsperson(en).....	20
1.5.3. Testen der Kinder.....	21
1.5.3.1. Entwicklungsscreenings.....	22
1.5.3.2. Allgemeine Entwicklungstests.....	24
1.5.3.3. Spezifische Entwicklungstests	27
Zusammenfassung	28
2. Qualitätsanforderungen an Entwicklungstests.....	29
2.1. Objektivität	29
2.1.1. Durchführungsobjektivität	29
2.1.2. Auswertungsobjektivität	30
2.1.3. Interpretationsobjektivität	30
2.1.4. Objektivität in der Entwicklungsdiagnostik.....	31
2.2. Reliabilität.....	32
2.2.1. Paralleltest-Reliabilität.....	32
2.2.2. Retest-Reliabilität	33
2.2.3. Innere Konsistenz.....	33
2.2.4. Reliabilität in der Entwicklungsdiagnostik.....	33
2.3. Validität.....	34
2.3.1. Inhaltliche Gültigkeit	34
2.3.2. Kriteriumsvalidität	35

2.3.3.	Konstruktvalidität	35
2.3.4.	Validität in der Entwicklungsdiagnostik.....	36
2.4.	Normierung.....	37
2.5.	Kritische Betrachtung der allgemeinen Entwicklungstests in Hinblick auf die Gütekriterien	38
2.5.1.	ET 6-6 (Petermann, Stein & Macha, 2008)	38
2.5.2.	MFED 2-3 (Hellbrügge, 1994)	39
2.5.3.	Bayley-II (Reuner, Rosenkranz, Pietz & Horn, 2007).....	40
2.5.4.	Bayley-III (Bayley, 2006)	41
	Zusammenfassung	42
3.	Spielbasierte Diagnostik.....	43
3.1.	Geschichte der spielbasierten Diagnostik	45
3.2.	Transdisciplinary Play-Based Assessment.....	46
	Zusammenfassung	47
4.	Fähigkeitsbereich der visuellen Wahrnehmung	48
4.1.	Farbwahrnehmung.....	48
4.2.	Farbpräferenz.....	50
4.3.	Farbdifferenzierung	51
4.3.1.	Schwierigkeiten beim Erwerb der Farbbezeichnungen	51
4.3.2.	Entwicklungsschritte in der Farbdifferenzierung.....	52
4.4.	Kategorische Farbwahrnehmung	53
4.5.	Formwahrnehmung	55
4.6.	Größenwahrnehmung	56
4.7.	Diagnostik der visuellen Wahrnehmung.....	56
	Zusammenfassung	58
5.	Wie untersucht man Kinder?	59
5.1.	Räumliche Bedingungen	59
5.2.	Anforderungen an den Untersucher.....	60

5.3. Vorbereitung der Untersuchung.....	61
5.4. Durchführung der Untersuchung.....	62
5.4.1. Untersuchungszeit.....	62
5.4.2. Untersuchungsdauer.....	62
5.4.3. Anwesenheit der Bezugsperson(en) während der Untersuchung	63
5.4.4. Untersuchung	63
5.4.5. Verhaltensbeobachtung während der Untersuchung	65
5.4.6. Abschluss der Untersuchung.....	66
Zusammenfassung	66
 EMPIRISCHER TEIL	 67
6. Ausgangspunkt und Ziel dieser Arbeit	69
7. Fragestellungen	69
8. Methodik.....	70
8.1. Setting.....	71
8.2. Untersuchungsablauf	72
8.3. Untersuchungsinstrument (Kuchler et al., 2011)	73
8.3.1. Erfassung der visuellen Wahrnehmung.....	73
8.3.1.1. Vorgenommenen Veränderungen.....	74
8.3.2. Erfassung der Motorik.....	75
8.3.2.1. Erfassung der Grobmotorik	75
8.3.2.2. Erfassung der Fein- und Visumotorik (Anziehen)	75
8.3.2.3. Erfassung der Feinmotorik	76
8.3.2.4. Erfassung der Visumotorik.....	76
8.3.2.5. Vorgenommene Veränderungen.....	76
8.3.3. Erfassung der Aufmerksamkeit	77
8.3.3.1. Vorgenommene Veränderungen.....	77
8.3.4. Erfassung der sozioemotionalen Entwicklung	77
8.3.4.1. Vorgenommene Veränderungen.....	77
8.3.5. Elternfragebogen	78
8.4. Die Stichprobe	78
8.4.1. Rekrutierung der Stichprobe.....	78

8.4.2. Soziodemographische Merkmale der Stichprobe	79
9. Auswertung.....	80
10. Ergebnisse.....	83
10.1. Anzahl und Dauer der Pausen	83
10.2. Visuelle Wahrnehmung	84
10.2.1. Aktive Farbdifferenzierung	84
10.2.2. Passive Farbdifferenzierung	85
10.2.3. Stabilität des Farbkonzepts	86
10.2.4. Form- und Größendifferenzierung.....	86
10.3. Motorik.....	87
10.3.1. Grobmotorik	88
10.3.2. Fein- und Visumotorik	89
10.3.3. Feinmotorik	90
10.3.4. Visumotorik	91
10.4. Aufmerksamkeit.....	92
10.4.1. Aufmerksamkeit in Minuten.....	92
10.4.2. Aufmerksamkeit – zurückholen.....	93
10.4.3. Aufmerksamkeit – Buchseiten.....	94
10.5. Sozioemotionale Entwicklung	95
10.5.1. Selbstständigkeit.....	95
10.5.2. Lösung und Kontakt zur Bezugsperson.....	96
10.6. Zusammenfassung der Itemanalyse	98
10.7. Reliabilität.....	99
10.8. Objektivität.....	99
10.8.1. Durchführungsobjektivität.....	99
10.8.2. Auswertungsobjektivität.....	100
10.8.3. Interpretationsobjektivität.....	100
10.9. Validität.....	101
10.9.1. Inhaltliche Validität	101
10.9.2. Konstruktvalidität	101
10.9.3. Kriteriumsvalidität.....	103

DISKUSSION.....	104
ZUSAMMENFASSUNG.....	114
ABSTRACT.....	116
LITERATURVERZEICHNIS.....	117
TABELLENVERZEICHNIS.....	127
ABBILDUNGSVERZEICHNIS.....	129
ANHANG.....	133

Einleitung

Kuchler und Sapper entwickelten 2011 mit der Unterstützung von Frau Ass.-Prof. Dr. Kastner-Koller und Frau Ass.-Prof. Dr. Deimann einen Itempool zur Erfassung der Entwicklung zweijähriger Kinder.

Der Grundgedanke hierfür war, dass es im deutschsprachigen Raum einen Mangel an diagnostischen Verfahren gibt, die einerseits die wesentlichen Fähigkeitsbereiche von Kindern im dritten Lebensjahr erfassen und deren Gütekriterien andererseits als zufriedenstellend gelten.

Dieser 2011 entstandene Itempool enthält Aufgaben zur Erfassung der kognitiven Entwicklung, der visuellen Wahrnehmung, der Motorik, der Sprache, der sozioemotionalen Entwicklung, der Aufmerksamkeit und des Gedächtnisses und beruht auf einem spielbasierten Ansatz. Dieser sieht vor, die diagnostischen Intentionen in Spiele einzukleiden (Sturzbecher, 2011), um dem Kind zu ermöglichen, sein Optimum an Fähigkeiten in einer natürlichen Situation zu zeigen (Kelly-Vance, Ryalls & Glover, 2002).

Birngruber (2012) und Fuchs-Garderer (2012) sowie Putzer (2013) und Sindelar (2013) nahmen Veränderungen an diesem Itempool vor und erprobten diese an einer kleinen Stichprobe von Kindern zwischen dem 24. und 36. Lebensmonat.

Das Ziel dieser Diplomarbeit ist die erneute Überarbeitung und testtheoretische Überprüfung des Itempools.

Der theoretische Teil dieser Arbeit soll Aufschluss darüber geben, was Entwicklungsdiagnostik ausmacht, welche Methoden angewandt werden und welchen Gütekriterien die diagnostischen Verfahren entsprechen sollen. Danach folgen eine kritische Betrachtung vorhandener Entwicklungstests im deutschsprachigen Raum und die Darstellung eines neuen Ansatzes der spielbasierten Diagnostik. Anschließend wird auf den Fähigkeitsbereich der visuellen Wahrnehmung, einen Bereich des vorhin erwähnten spielbasierten Itempools, detailliert eingegangen. Um Näheres über die kognitive Entwicklung zu erfahren, wird auf die Arbeit von Punz (in Vorbereitung) verwiesen. In Kapitel 5 werden abschließend günstige Rahmenbedingungen für eine Untersuchung von Kindern beschrieben.

Im Anschluss daran werden im empirischen Teil das Ziel dieser Arbeit, die Fragestellungen und die Methode beschrieben. Danach folgt die statistische Analyse, durch die geklärt werden soll, welche Items zur Überprüfung der vorhin genannten Fähigkeitsbereiche zufriedenstellende statistische Kennwerte erzielen, ob sich Geschlechts- und Altersunterschiede in der Stichprobe der zweijährigen Kinder feststellen lassen und ob die Gütekriterien der Objektivität, Reliabilität und Validität als gegeben betrachtet werden können. Zusätzlich werden die Einschätzungen der Mütter im Fragebogen über die Fähigkeiten der Kinder mit den Erfahrungen aus der Testung verglichen. Im Hinblick auf die Fragestellungen werden alle relevanten Ergebnisse am Ende dieser Arbeit zusammengefasst.

THEORETISCHER TEIL

1. Grundlagen der Entwicklungsdiagnostik

1.1. Definition des Entwicklungsbegriffs

Seit Beginn des 19. Jahrhunderts wurden viele Versuche unternommen den Entwicklungsbegriff zu definieren, jedoch lässt sich keine einheitliche Definition in der Literatur finden. Worüber laut Ettrich (2000) jedoch Einigkeit besteht, ist, dass dieser Begriff die gesamte Zeitspanne des Lebens umfasst, also die Zeit von der Konzeption bis zum Tod.

Laut Esser und Petermann (2010) orientiert sich der engere, traditionelle Entwicklungsbegriff „an Reifungsprozessen und bezieht sich auf eine Veränderungsreihe mit mehreren, aufeinander aufbauenden Schritten in Richtung eines höherwertigen Endzustandes“ (S. 9).

Diesem Entwicklungsbegriff zu Folge ist auch Ettrich (2000) der Ansicht, dass sich die Entwicklung in einer bestimmten Abfolge von Stufen vollzieht (*Sequentialität*) und dass diese nicht umkehrbar ist (*Irreversibilität*). Zusätzlich verläuft die Entwicklung nur in eine Richtung, da sie einen End- oder Optimalzustand anstrebt (*Unidirektionalität*) und ist bei allen Menschen ähnlich. Lediglich die Geschwindigkeit, in der sie voranschreitet, variiert (*Universalität*). Außerdem ist jede Entwicklungsstufe durch neue Eigenschaften gekennzeichnet und lässt sich nicht auf frühere reduzieren (*Strukturalismus*).

Dieses Stufenkonzept wird jedoch von Petermann und Macha (2005) sehr stark in Frage gestellt, da die Entwicklung nicht linear verläuft und Sprünge nicht selten auftreten. Esser und Petermann (2010) zu Folge unterliegen Entwicklungsverläufe einer beträchtlichen Variabilität. So kann die Entwicklung von Kindern sehr unterschiedlich verlaufen, aber dennoch normal sein (Petermann & Macha, 2005).

Aktuell wird die Entwicklung als multifaktorielles Geschehen verstanden, das die Erklärung von Entwicklungsunterschieden berücksichtigt und die Wechselwirkung zwischen Person und Umwelt einbezieht. Entwicklungsprozesse werden demnach von biologischen, sozialen und psychischen Faktoren beeinflusst (Montada, 2008).

Fuiko und Wurst (2003) geben dazu eine aktuelle Definition:

Entwicklung wird als lebenslanger Prozess, als eine Sequenz von Phasen des Aufbaus, der Stabilisierung und des Abbaues verstanden; diese Phasen laufen zum Teil parallel und haben Spezialisierung und Optimierung von Fähigkeiten und Eigenschaften zum Ziel. Entwicklung ist demnach das Produkt von Reifung, Umwelteinflüssen und individueller Gestaltung des Menschen. (S. 120)

1.2. Definition der Entwicklungsdiagnostik

Die Entwicklungsdiagnostik ist ein Teilbereich der psychologischen Diagnostik, der sich vorwiegend mit der frühen Kindheit beschäftigt. Mit Hilfe standardisierter, entwicklungsdiagnostischer Verfahren werden Informationen gesammelt, um Aussagen über eine normale Entwicklung, Entwicklungsauffälligkeiten oder -abweichungen eines Kindes zu treffen oder um spezifische Störungen zu identifizieren (Rümmele, 2011).

Deimann und Kastner-Koller (2007) definieren die Entwicklungsdiagnostik folgendermaßen:

Entwicklungsdiagnostik als ein wesentlicher Bereich der Angewandten Entwicklungspsychologie beschäftigt sich mit der quantitativen und qualitativen Erfassung entwicklungsbedingter Kompetenzen im Hinblick auf die Beschreibung und Erklärung des aktuellen Kompetenzniveaus, die Prognose der zukünftigen Entwicklung sowie die Planung und Evaluation von Entwicklungsinterventionen. (S. 558)

Anhand dieser Definition lassen sich die Aufgaben der Entwicklungsdiagnostik ableiten, die nachfolgend näher beschrieben werden.

1.3. Aufgaben und Ziele der Entwicklungsdiagnostik

Deimann und Kastner-Koller (2007) listen folgende vier Aufgaben der Entwicklungsdiagnostik auf:

- *die Beschreibung des aktuellen Entwicklungsstandes*

Die Entwicklung eines Kindes wird mit einer normativ festgelegten Entwicklung verglichen. Werden Abweichungen festgestellt, werden diese je nach Richtung als Entwicklungsvorsprung (Akzeleration) oder als Entwicklungsrückstand (Retardation) interpretiert.

- *die Diagnose von Entwicklungsauffälligkeiten*

Eine Abweichung von der normativ festgelegten Entwicklung, die eine hohe Variabilität der normalen Entwicklung berücksichtigen muss, deutet auf eine Auffälligkeit bzw. im negativen Sinne auf eine Entwicklungspathologie hin.

- *die Evaluation von Maßnahmen zur Entwicklungsförderung*

Werden Auffälligkeiten aufgrund eines Abweichens der Entwicklung von der normativen Entwicklung festgestellt, so müssen entsprechende Interventionen, wie z.B. Förderprogramme, folgen, deren Behandlungserfolge evaluiert werden können.

- *die Prognose zukünftiger Entwicklung*

Mittels entwicklungspsychologischer Verfahren sollen auf Basis der aktuellen Entwicklung Aussagen bzw. Prognosen über den weiteren Entwicklungsverlauf getroffen werden.

Rümmele (2011) ergänzt diese Aufzählung von Aufgaben um *die Ermittlung von Veränderungsgeschwindigkeit und –richtung der Entwicklung*, um *den Einbezug von Entwicklungsbedingungen* und um *die Beschreibung von Entwicklungsprofilen*.

1.4. Geschichte der Entwicklungsdiagnostik

Die Anfänge der Entwicklungsdiagnostik gehen in das 19. Jahrhundert zurück und lassen sich auf Alfred Binet (1905) und Arnold Gesell (1925) zurückführen.

Binet und Simon (1905) bildeten als Erste in ihrer „Stufenleiter der Intelligenz“ intellektuelle Leistungen, die in standardisierten Situationen erhoben wurden, altersbezogen ab (Ettrich, 2000). Das Intelligenzalter eines Kindes wurde auf Basis der Anzahl der gelösten Aufgaben berechnet, das Aufschluss darüber gab, ob die Denkfähigkeit des Kindes altersgemäß, retardiert oder akzeleriert war (Deimann & Kastner-Koller, 2007).

Arnold Gesell (1925) war als Erster daran interessiert, die frühkindliche Entwicklung zu beschreiben, indem das Leistungsniveau anhand von standardisierter Aufgaben erfasst wurde. Anhand von Beobachtungen von Vorschulkindern veröffentlichte er Normen, die charakteristische Entwicklungsstufen für bestimmte Altersgruppen wiedergeben sollten (Quaiser-Pohl, 2010).

Auch Charlotte Bühler war, wie Arnold Gesell, nicht nur daran interessiert, die Intelligenz zu erfassen, sondern auch die Gesamtentwicklung zu beschreiben. Auf Basis 24-stündiger Dauerbeobachtungen von Kindern im Alter zwischen 0 bis 6 Jahren, entstand der Bühler-Hetzer-Kleinkindertest, der ein Inventar alterstypischer Verhaltensweisen darstellt. Dieser Test wurde 1932 veröffentlicht und war das bekannteste entwicklungsdiagnostische Verfahren dieser Zeit (Deimann & Kastner-Koller, 2007).

Zwei weitere einflussreiche Verfahren stellen die „Bayley Scales of Infant Development“ von Nancy Bayley dar, die ursprünglich für eine Langzeitstudie in den 30er Jahren konstruiert und im Jahre 1969 veröffentlicht wurden. Weitere Bearbeitungen folgten, bis 2007 eine deutsche Überarbeitung von Reuner, Rosenkranz, Pietz und Horn, jedoch ohne deutsche Normen, veröffentlicht wurde (Hagmann-von Arx, Meyer & Grob, 2008).

Zusätzlich kam es in England durch Ruth Griffith im Jahr 1954 zu einer Testkonstruktion mit dem Namen „Griffiths Mental Development Scales“, die bis heute in der deutschsprachigen Ausgabe in Form der „Griffiths Entwicklungsskalen“ von Brandt und Sticker (2001) Verwendung finden (Quaiser-Pohl, 2010).

Danach verschwand das Interesse an der Entwicklungsdiagnostik, bis es in den 1960er und 1970er Jahren im Zuge des amerikanischen Head-Start-Programms wieder aufflammte. Zu dieser Zeit war das Interesse für die frühzeitige Förderung von Kindern in sozial benachteiligten Familien in den Bereichen der Sprache und der kognitiven Entwicklung sehr groß, wofür geeignete Testverfahren benötigt wurden, die sich hauptsächlich auf diese beiden Fähigkeitsbereiche spezialisierten (Deimann & Kastner-Koller, 2007).

Das Interesse an Entwicklungstests nimmt seit den 1980er Jahren wieder zu und zeigt sich in Veröffentlichungen allgemeiner Entwicklungstests, wie z.B. dem „Wiener Entwicklungstest“ von Kastner-Koller und Deimann (1998, 2002, 2012) oder dem „Entwicklungstest sechs Monate bis sechs Jahre“ von Petermann, Stein und Macha (2000, 2004, 2008), deren Erstveröffentlichungen weitere Überarbeitungen folgten (Hagmann-von Arx et al., 2008).

1.5. Methoden in der Entwicklungsdiagnostik

Laut Esser und Petermann (2010) sind mit dem Begriff der Entwicklungsdiagnostik meist ausschließlich Entwicklungstests gemeint. Eine fachgerechte Diagnostik der Entwicklung beinhaltet jedoch viel mehr als die Durchführung von Tests.

Nachfolgend werden die wichtigsten Zugänge in der Entwicklungsdiagnostik beschrieben: die *Verhaltensbeobachtung* der Kinder, die *Befragung von Bezugspersonen* und das *Testen mittels standardisierten Verfahren*.

1.5.1. Verhaltensbeobachtung der Kinder

Das wichtigste Erhebungsinstrument in der Kleinkindforschung ist laut Schölmerich (2011) die Verhaltensbeobachtung, da sich Säuglinge und Kleinkinder noch nicht mitteilen können. Zusätzlich sind die Verhaltenskontrolle und die willentliche Steuerung

des Ausdrucks in diesem Alter noch deutlich eingeschränkt und somit die Verhaltens- und Ausdrucksäußerungen sehr reichhaltig und unverfälscht.

Kastner-Koller und Deimann (2009) sind der Ansicht, dass die Verhaltensbeobachtung gerade bei sehr kleinen Kindern und bei Kindern mit Entwicklungsdefiziten die einzige Möglichkeit darstellt, einen Einblick in die Fähigkeiten eines Kindes zu erlangen, solange das Sprachverständnis es nicht erlaubt, Testinstruktionen zu befolgen und die expressive Sprache bzw. auch die kognitiven Fähigkeiten ein diagnostisches Vorgehen nicht ermöglichen. Es existieren jedoch kaum standardisierte Beobachtungsinventare für Fragestellungen der Entwicklungsdiagnostik.

Ein weiteres Feld, in der die Verhaltensbeobachtung eingesetzt werden kann, ist die Interaktionsanalyse (Deimann & Kastner-Koller, 2007). Am häufigsten wird sie jedoch in der Testsituation eingesetzt, da dies eine ökonomische Form der Beobachtung darstellt. Während der Bearbeitung der Aufgaben können verschiedene Aspekte, wie z.B. die Arbeitshaltungen, die Konzentrationsfähigkeit, die Frustrationstoleranz und andere Dimensionen des Kindes, beobachtet und beurteilt werden (Kastner-Koller & Deimann, 2009).

Der Zugang mittels Verhaltensbeobachtung birgt jedoch auch Nachteile, wie z.B. können die Objektivität, Reliabilität und Validität durch eine verzerrte Personenwahrnehmung beeinträchtigt werden. Hierbei erhöhen laut Kastner-Koller und Deimann (2009) technische Hilfsmittel, wie z.B. eine Videoaufnahme oder eine computergestützte Kodierung und Auswertung, die Genauigkeit und Zuverlässigkeit der erhobenen Daten.

1.5.2. Befragung der Bezugsperson(en)

Im Rahmen einer Entwicklungsdiagnostik haben die Informationen der Eltern einen besonders hohen Stellenwert, da in den (kurzen) Untersuchungssituationen gewisse Verhaltens- und Merkmalsausprägungen der Kinder nur schwer oder gar nicht festzustellen sind. Zusätzlich kennen die Eltern das Alltagsverhalten der Kinder am besten und können am ehesten genaue Angaben zu bestehenden oder vergangenen Erkrankungen, zum bisherigen Entwicklungsverlauf, zu familiären Belastungen oder zu emotionalen und sozialen Problemen ihrer Kinder machen (Rennen-Allhoff, Allhoff, Bowi & Laaser, 1993).

Obwohl die Befragung der Bezugspersonen eine wichtige Rolle in der Praxis eingenommen hat, ist nicht außer Acht zu lassen, dass es empirische Hinweise gibt, dass es sich um eine Informationsquelle mit fraglicher Reliabilität und Validität handelt (Deimann & Kastner-Koller, 2007). Deimann, Kastner-Koller, Benka, Kainz und Schmidt (2005) dokumentieren, dass von Seiten der Eltern bei der Einschätzung von Leistungen bei Kindern mit Entwicklungsproblemen generell eine Überschätzung auftritt. Die Einschätzungen der Bezugspersonen entsprechen eher den realen Leistungen des Kindes, wenn das zu bewertende Verhalten beobachtbar ist oder Vergleichsmaßstäbe vorliegen. Zusätzlich haben laut Deimann und Kastner-Koller (2007) die Art der Befragung und das Setting einen wesentlichen Einfluss auf die Genauigkeit der Angaben.

Trotz der mangelnden Güte von Elternbefragungen ist Deegener (2009) der Ansicht, dass Informationssammlungen über Dritte umso eher erfolgen müssen, je jünger die Kinder sind. Die (Video-)Beobachtung von Kleinst- und Kleinkindern in sozialen Kontakten darf jedoch auf keinen Fall vernachlässigt werden (Deegener, 2009).

1.5.3. Testen der Kinder

Entwicklungstests zählen laut Deimann und Kastner-Koller (2007) zu den wichtigsten Erhebungsmethoden im Rahmen einer entwicklungsdiagnostischen Untersuchung und ermöglichen die bestmögliche diagnostische Annäherung an die Fähigkeiten und Fertigkeiten der untersuchten Person.

Unter dem Begriff eines Entwicklungstests lassen sich drei Verfahrenstypen differenzieren (Hagmann-von Arx et al., 2008), die nachfolgend näher beschrieben werden:

- *Entwicklungsscreenings*, die zeitökonomisch eine erste, grobe Abschätzung des Entwicklungsstandes des Kindes ermöglichen,
- *allgemeine Entwicklungstests*, die den Entwicklungsstand des Kindes in den wesentlichen Funktionsbereichen erheben und
- *spezifische Entwicklungstests*, die das Erfassen in einem bestimmten Fähigkeitsbereich, wie z.B. der Motorik oder der Sprache, erlauben.

Welche Art von Verfahren am besten geeignet ist, um eine Abklärung des Entwicklungsstandes zu ermöglichen, hängt laut Petermann und Macha (2005) von den Fragestellungen, der interessierenden Entwicklungsbereiche, der klinischen Gruppen, dem Altersbereich und der zur Verfügung stehenden Zeit ab. Zusätzlich sollte auch eine differenzielle Einschätzung der Testgütekriterien (s. Kap. 2. Qualitätsanforderungen an Entwicklungstests) vorgenommen werden (Petermann & Macha, 2005).

Zu berücksichtigen ist laut Esser und Petermann (2010) allerdings, dass alleine auf Basis von Entwicklungstests keine hinreichenden Aussagen über den Entwicklungsstand eines Kindes gemacht werden können. Eine ausführliche Diagnostik beinhaltet eine sorgfältige Anamnese, die Verhaltensbeobachtung des Kindes, die Befragung der Bezugspersonen und das Anwenden von Entwicklungstests (Sarimski, 2009; Hagmann-von Arx et al., 2008; Rümmele, 2011).

1.5.3.1. Entwicklungsscreenings

Entwicklungsscreenings sind Kurztestverfahren, haben demnach eine Durchführungsdauer von 10 bis 20 Minuten, und klassifizieren den Entwicklungsstand eines Kindes in grundlegenden Entwicklungsbereichen als auffällig oder unauffällig. Um diese Differenzierung vornehmen zu können, werden Leistungsgrenzwerte (sog. Cut-Off-Werte) festgelegt, deren Unterschreitung als auffällig bewertet wird (Petermann & Macha, 2005). Laut Tröster, Flender und Reineke (2005) ist es im Rahmen von Screenings jedoch immer sinnvoll und notwendig, weitere Informationsquellen (wie z.B. die Eltern oder Erzieher) einzubeziehen.

Das Ziel von Entwicklungsscreenings ist es, „Entwicklungsrisiken zu einem Zeitpunkt, zu dem noch keine Symptome aufgetreten sind, mit ökonomischen Methoden sensitiv und spezifisch zu erfassen“ (Deimann & Kastner-Koller, 2007, S. 564).

Unter der Sensitivität eines Screenings wird der Prozentsatz der Fälle mit Entwicklungsrisiko verstanden, der durch das Screening identifiziert wird. Die Spezifität gibt die Rate an, mit der Nicht-Risikofälle auch als solche erkannt werden (Deimann & Kastner-Koller, 2007). „Gute Screeningverfahren sollten Sensitivitätswerte und Spezifitätswerte über 80% aufweisen, damit gewährleistet ist, dass das Screening einen

hohen Anteil an Risikopersonen herausfiltert, ohne zu viele falsche Zuordnungen zu liefern.“ (Kastner-Koller & Deimann, 2011, S. 277)

Ein weiteres Gütekriterium für Entwicklungsscreenings neben der Reliabilität ist der Ratz-Index, mittels diesem die prognostische Validität eingeschätzt werden kann. Dieser Index sollte über 60% liegen, um das Verfahren als sehr gut zu bewerten. Liegt der Wert unter 34% deutet dies auf eine unzureichende Prognosefähigkeit hin (Kastner-Koller & Deimann, 2011, 278).

Eine niedrige Sensitivität hat zur Folge, dass entwicklungsauffällige Kinder übersehen und nicht frühzeitig gefördert werden, die Beeinträchtigungen zunehmen und somit das Risiko für Sekundärfolgen steigt (Fuiko, 2003; zitiert nach Esser & Petermann, 2010, S. 53). Fuiko (2003, zitiert nach Esser & Petermann, 2010, S. 53) ist deshalb der Ansicht, dass Screenings allenfalls als ersten Schritt im diagnostischen Prozess gesehen werden sollten, obwohl sie durch den zeitlichen und finanziellen geringeren Aufwand gegenüber umfangreicheren diagnostischen Untersuchungen überzeugen.

Folgende Screeningverfahren im Bereich der allgemeinen Entwicklungsdiagnostik decken den Altersbereich der zweijährigen Kinder ab:

- *Die Denver Entwicklungsskalen* (DES; Flehmig, Schloon, Uhde & Bernutz, 1973) ermöglichen eine Erfassung von Entwicklungsauffälligkeiten im Alter von 0 bis 6 Jahren in den Bereichen der Grobmotorik, der Sprache, der Feinmotorik-Adaption und dem persönlich-sozialen Kontakt.
- *Die Erweiterte Vorsorgeuntersuchung* (EVU; Melchers et al., 2003) unterstützt die Einschätzung der Entwicklung bei Kindern zwischen dem dritten und dem 64. Lebensmonat in den Bereichen der Motorik, der Sprache, der Kognition und dem Verhalten.

1.5.3.2. Allgemeine Entwicklungstests

Ein Entwicklungstest ist „ein wissenschaftliches Routineverfahren zur Untersuchung eines oder mehrerer empirisch abgrenzbarer Entwicklungsmerkmale mit dem Ziel einer möglichst quantitativen Aussage über den relativen Grad der individuellen Merkmalsausprägung“ (Lienert, 1969, S. 7; zitiert nach Quaiser-Pohl & Köhler, 2010, S. 85).

Allgemeine Entwicklungstests sind hauptsächlich für das Säuglings- und Kleinkindalter entwickelt worden, da davon ausgegangen wird, dass eine isolierte Betrachtung einzelner Funktionsbereiche noch nicht möglich bzw. zulässig ist (Fuiko & Wurst, 2003).

Die im deutschsprachigen Raum gebräuchlichen entwicklungsdiagnostischen Verfahren folgen keinem einheitlichen theoretischen Entwicklungskonzept, sondern bestehen aus einer Aufgabensammlung aus verschiedenen Fähigkeitsbereichen, die sich bewährt haben (Sarimski, 2009). Somit ermöglichen diese eine differenzierte Betrachtung der wesentlichen Fähigkeitsbereiche.

Laut Esser und Petermann (2010) greifen allgemeine Entwicklungstests auf folgende Dimensionen zurück, die Großteils historisch begründet sind und von Experten als weitgehend evident angesehen werden:

- Grobmotorik
- Auge-Hand-Koordination (Visumotorik), Handgeschicklichkeit
- Wahrnehmung
- kognitive Entwicklung
- Sprachentwicklung
- Sozialentwicklung
- emotionale Entwicklung
- Selbstständigkeit

Die Aufgaben in diesen Bereichen lassen sich einerseits in der Durchführung mit kleinen Kindern gut umsetzen und andererseits hat die Leistung der Kinder in diesen Aufgaben eine prognostische Bedeutung für die weitere Entwicklung (Sarimski, 2009).

Ettrich (2000) ist der Ansicht, dass entwicklungsdiagnostischen Verfahren, die eine globale Bewertung der allgemeinen Fähigkeiten erlauben, eine höhere Bedeutung zukommen, als solchen Verfahren, die spezifische Fähigkeiten, wie z.B. die Sprache oder die Motorik, erheben. Dies begründet Ettrich (2000) damit, dass jede entwicklungsdiagnostische Untersuchung eine Analyse des allgemeinen Leistungsstandes des Kindes einschließen muss, um auf Basis dieser Information andere Beobachtungen und Aspekte des Verhaltens sachgerecht einschätzen zu können. Hierfür ist laut Petermann und Macha (2005) eine vom Alter abhängige Durchführungsdauer zwischen 30 und 90 Minuten notwendig.

Michaelis und Niemann (2004) sehen jedoch einen Nachteil im hohen Zeitaufwand einiger Verfahren, da in vielen Tests die Motivation und Konzentration der Kinder nur wenig Berücksichtigung finden.

Einen weiteren großen Nachteil sehen Esser und Petermann (2010) in Entwicklungstests, die von einem Stufenleiterkonzept (s. Kap. 1.1. Definition des Entwicklungsbegriffes) ausgehen. Diese nehmen eine relativ starre, biologisch determinierte Abfolge von Entwicklungsschritten an, die alle Kinder in ähnlicher Weise durchlaufen. Laut Petermann und Macha (2005) sind jedoch nur für wenige Entwicklungsbereiche universelle Abfolgen beschrieben und empirisch bestätigt worden. Die Einstiegs- und Abbruchkriterien sind meist klar definiert und die Aufgaben ihrer Schwierigkeit nach in aufsteigender Reihenfolge geordnet (Petermann & Macha, 2005). „Solche eher starren Aufgabenreihungen können jedoch zu ungünstigen praktischen Konsequenzen im Sinne von Falschbeurteilungen führen.“ (Esser & Petermann, 2010, S. 69)

Als Ergebnis der Durchführung eines allgemeinen Entwicklungstests lassen sich für die verschiedenen Fähigkeitsbereiche standardisierte Werte (wie z.B. C-Werte, Prozentränge, T-Werte) bestimmen (Quaiser-Pohl & Köhler, 2010). Zusätzlich erhält man Maße für den Stand der Gesamtentwicklung, wie z.B. einen Gesamt-Entwicklungsquotienten (EQ) oder ein Gesamt-Entwicklungsalter (EA). Mittels einer anschaulichen Darstellung der errechneten Werte lässt sich ein Entwicklungsprofil erstellen, anhand dessen über Stärken und Schwächen eines Kindes berichtet werden kann.

Allgemeine Entwicklungstests für zweijährige Kinder im deutschsprachigen Raum

Nun soll eine kurze Beschreibung der allgemeinen Entwicklungstests gegeben werden, die den Altersbereich der zweijährigen Kinder abdecken und im deutschen Sprachraum weite Verbreitung gefunden haben:

- Der *Entwicklungstest 6 Monate bis 6 Jahre* (ET 6-6; Petermann, Stein & Macha, 2008) ermöglicht es, die Körpermotorik, die Handmotorik, die kognitive Entwicklung (Gedächtnis, Handlungsstrategien, Kategorisieren, Körperbewusstsein), die Sprache (Sprachverständnis, Sprachproduktion), die Sozialentwicklung und die emotionale Entwicklung von Kindern zwischen 6 Monaten und 6 Jahren zu erheben. Die durchschnittliche Durchführungsdauer beträgt zwischen 25 und 40 Minuten, kann aber im Kleinkindalter stark variieren.
- Die *Münchener Funktionelle Entwicklungsdiagnostik für das zweite und dritte Lebensjahr* (MFED 2-3; Hellbrügge, 1994) erfasst sieben Fähigkeitsbereiche. Überprüft werden die Entwicklung der Körperbeherrschung, der Feinmotorik, des Erfassens von Zusammenhängen, der Lautäußerung und des Sprechens, des Sprachverständnisses, des sozialen Verhaltens und der Selbstständigkeit. Die Durchführungsdauer wird mit 40 bis 60 Minuten angegeben.
- Die *Bayley Scales of Infant Development* (Bayley-II; Reuner, Rosenkranz, Pietz & Horn, 2007) wurden aus dem Amerikanischen für den deutschsprachigen Raum überarbeitet und erlauben eine Einschätzung der Fähigkeiten im Bereich der kognitiven und der motorischen Entwicklung bei Kindern zwischen dem 1. und dem 42. Lebensmonat. Zusätzlich wird eine Beurteilung des Verhaltens vorgenommen. Die Testung kann bis zu 60 Minuten dauern.
- Die *Bayley Scales of Infant Development* (Bayley-III; Bayley, 2006) sind eine Weiterentwicklung der Bayley-II. Hier finden sich zusätzlich zur kognitiven und motorischen Skala eine sprachliche und eine sozial-emotionale Skala. Die Einschätzung des Alltagsverhaltens der Kinder wird von den Eltern vorgenommen. Die Testung dauert bis zu 90 Minuten und ist bei Kindern zwischen dem 1. und dem 42. Lebensmonat möglich.

1.5.3.3. Spezifische Entwicklungstests

Spezifische Entwicklungstests fokussieren auf einen ausgewählten Fähigkeitsbereich (Hagmann-von Arx et al., 2008). Dazu zählen Verfahren, die der Diagnostik allgemeiner kognitiver, motorischer, sprachlicher Entwicklung sowie der Wahrnehmungsentwicklung dienlich sind (Esser & Petermann, 2010).

Sie unterscheiden sich von allgemeinen Entwicklungstests nicht nur in der Bandbreite der untersuchten Fähigkeiten und Fertigkeiten, sondern auch in der Differenziertheit, mit der die einzelnen Fähigkeitsbereiche erhoben werden (Deimann & Kastner-Koller, 2007).

Im Folgenden wird eine kurze Übersicht über Verfahren zur Einschätzung der visuellen Wahrnehmung bei Kindern im Alter zwischen 24 und 36 Monaten gegeben:

Für den Fähigkeitsbereich der Wahrnehmungsentwicklung lassen sich nur wenige Testverfahren finden, da vor allem höhere Wahrnehmungsleistungen nur schwer von kognitiven Leistungen zu trennen sind und diese somit bei kognitiven Entwicklungstests erhoben werden (Kastner-Koller & Deimann, 2011).

Dacheneder (2009) führt einige Tests zur Überprüfung der visuellen Wahrnehmung an, wobei es für Zweijährige nur die Möglichkeit gibt, diesen Fähigkeitsbereich im Rahmen von Intelligenztests zu erfassen.

- Der *Hannover-Wechsler-Intelligenztest für Kinder im Vorschulalter – III* (HAWIVA-III; Ricken, Fritz, Schuck & Preuß, 2007) bildet bei Kindern ab 2 Jahren und 6 Monaten die visuell-konstruktiven Leistungen mittels der Subtests „Mosaik-Test“ und „Figuren legen“ ab.
- Der *Nonverbale Intelligenztest für Kinder* (SON-R 2½– 7; Tellegen, Laros & Petermann, 2007) überprüft mittels der Subtests „Mosaike“, „Puzzles“ und „Zeichenmuster“ große Anteile von visuell-räumlichen und visuokonstruktiven Wahrnehmungsleistungen.
- Die *Kaufman-Assessment Battery for Children* (K-ABC; Melchers & Preuß, 1991) erhebt die Leistung in der visuellen Wahrnehmung, indem beim Subtest „Gestaltschließen“ die Fähigkeit zum visuellen Gestaltschließen und im Subtest „Dreiecke“ die Fähigkeit im visuell-räumlichen und visokonstruktiven Bereich überprüft werden.

Zusammenfassung

Entwicklung wird als multifaktorielles Geschehen über die ganze Lebensspanne hinweg verstanden, das die Erklärung von Entwicklungsunterschieden berücksichtigt und die Wechselwirkung zwischen Personen und Umwelt einbezieht. Entwicklung wird demnach von biologischen, sozialen und psychischen Faktoren beeinflusst.

Die Aufgaben der Entwicklungsdiagnostik sind die Beschreibung des aktuellen Entwicklungsstandes, die Diagnose von Entwicklungsauffälligkeiten, die Prognose zukünftiger Entwicklung und die Evaluation von Maßnahmen zur Förderung der Entwicklung. Methoden, die dafür angewandt werden, sind die Verhaltensbeobachtung, die Befragung von Bezugspersonen und das Testen mittels standardisierter Verfahren. Hinreichende Aussagen über den Entwicklungsstand eines Kindes können nur dann gewährleistet werden, wenn alle drei Zugänge gewählt werden.

Es lassen sich drei Arten entwicklungsdiagnostischer Verfahren, wie die Entwicklungsscreenings, die allgemeinen und die spezifischen Entwicklungstests, unterscheiden.

Entwicklungsscreenings sind Kurztestverfahren, die im Durchschnitt 10 bis 20 Minuten in Anspruch nehmen, und den Entwicklungsstand eines Kindes in den grundlegenden Entwicklungsbereichen als auffällig oder unauffällig klassifizieren. Allgemeine Entwicklungstests hingegen haben eine durchschnittliche Testdauer von 30 bis 90 Minuten und erfassen die Entwicklung eines Kindes in mehreren Bereichen: Grobmotorik, Auge-Hand-Koordination, Wahrnehmung, kognitive Entwicklung, Sprachentwicklung, Sozialentwicklung, emotionale Entwicklung und Selbstständigkeit. Spezifische Entwicklungstests fokussieren im Gegensatz dazu nur auf einen ausgewählten Fähigkeitsbereich, wie z.B. auf die Motorik oder die Sprache.

2. Qualitätsanforderungen an Entwicklungstests

Wie alle Testverfahren der psychologischen Diagnostik müssen auch jene, die der Erfassung der Entwicklung dienen, gewissen Gütekriterien entsprechen. Als Hauptkriterien lassen sich dafür die Objektivität, die Reliabilität und die Validität anführen, die im Folgenden näher beschrieben und in Beziehung zur Entwicklungsdiagnostik gesetzt werden. Zusätzlich wird am Ende des Kapitels auf die Normierung, einem wichtigen Nebengütekriterium, eingegangen.

2.1. Objektivität

Bortz und Döring (2006) beschreiben, dass ein Test oder Fragebogen objektiv ist, „wenn verschiedene Testanwender bei denselben Personen zu den gleichen Resultaten gelangen“ (S. 195). Lienert und Raatz (1994) sprechen dabei von interpersoneller Übereinstimmung der Untersucher. Als Maß für diese Übereinstimmung gilt der durchschnittliche Korrelationskoeffizient zwischen den Testbefunden unterschiedlicher Untersucher an verschiedenen Personen (Lienert & Raatz, 1994).

Somit lässt sich schlussfolgern, dass ein Test nicht objektiv wäre, wenn in die Durchführung oder Auswertung subjektives Wissen oder Deutungen eines Untersuchers einfließen (Bortz & Döring, 2006).

Fisseni (2004) bezeichnet die Objektivität als Maß, inwieweit eine Standardisierung des gesamten Testvorganges in der diagnostischen Situation gelingt.

Im Folgenden werden drei Formen der Objektivität näher angeführt:

die *Durchführungs-*, die *Auswertungs-* und die *Interpretationsobjektivität*. Kubinger (2009) zieht diesen Formulierungen die Begriffe Testleiterunabhängigkeit, Verrechnungssicherheit und Interpretationseindeutigkeit vor, wobei diese Begriffe den erst genannten gleichzusetzen sind.

2.1.1. Durchführungsobjektivität

Unter der Durchführungsobjektivität verstehen Lienert und Raatz (1994, S. 8) „den Grad der Unabhängigkeit der Testergebnisse von zufälligen oder systematischen Verhaltensvariationen des Untersuchers während der Testdurchführung, die ihrerseits zu Verhaltensvariationen des Pbn [Probanden] führen und dessen Ergebnis beeinflussen.“

Die Forderung nach Durchführungsobjektivität wäre verletzt, „wenn dieselbe Person die Aufgabenstellung bei dem einen Untersuchungsleiter nicht versteht, während sie bei einem anderen Untersucher problemlos arbeiten kann.“ (Bortz & Döring, 2006, S. 195)

Will der Testleiter eine maximal hohe Durchführungsobjektivität erreichen, so muss die Instruktion schriftlich so genau wie möglich festgelegt und die Untersuchungssituation so weit wie möglich standardisiert werden. Somit soll die soziale Interaktion zwischen Untersucher und Proband auf ein unumgängliches Minimum reduziert werden (Lienert & Raatz, 1994).

Nicht nur die wortwörtliche Vorgabe der Instruktionen vom Testleiter ist ein wesentlicher Faktor, sondern auch, dass die Antworten auf Rückfragen der Testpersonen schriftlich im Manual bzw. der Handanweisung festgelegt sind (Bortz & Döring, 2006). Die Durchführung lässt sich allerdings laut Fisseni (2004) nur in seltenen Fällen vollkommen standardisieren.

2.1.2. Auswertungsobjektivität

„Mit Verrechnungssicherheit ist gemeint, dass die Reglementierung (im Manual eines psychologisch-diagnostischen Verfahrens), wie die einzelnen Testleistungen bzw. –reaktionen auf Items zu numerischen oder kategorialen Testwerten zu verrechnen sind, derart exakt festgelegt ist, dass jeder Auswerter zu denselben Ergebnissen kommt.“ (Kubinger, 2009, S. 43). Die Auswertungsobjektivität wird bei festen Antwortformaten (wie z.B. Multiple-Choice) eher und leichter realisiert als bei offenen Antwortvorgaben (Lienert & Raatz, 1994). Hierfür ist eine ausführliche Auswertungsanweisung mit vielen Beispielen sehr hilfreich (Fisseni, 2004).

2.1.3. Interpretationsobjektivität

„Die Interpretationseindeutigkeit ist gegeben, wenn aus denselben Auswertungsergebnissen verschiedene „Interpreten“ zum selben Schluss gelangen.“ (Kubinger, 2009, S. 46).

Individuelle Deutungen und Interpretationen müssen vermieden werden. Anstatt dessen vergleicht man die Testwerte einer Person mit den Normen repräsentativer Stichproben (wie z.B. Alter, Geschlecht), um Aussagen darüber zu machen, ob die Ergebnisse im Über-, Unter- oder Durchschnittsbereich liegen (Bortz & Döring, 2006).

Laut Esser und Petermann (2010) fällt die Objektivität umso höher aus, „je standardisierter die Untersuchungssituation, das Testmaterial und die Instruktionen sind, je exakter die Auswertungsregeln und Bewertungskriterien definiert und je eindeutiger die Interpretationsmöglichkeiten (Normwerte, Fallbeispiele) beschrieben sind“ (S. 16). All diese Informationen werden in den Testmanualen bzw. Testhandbüchern detailliert beschrieben.

2.1.4. Objektivität in der Entwicklungsdiagnostik

Esser und Petermann (2010) erwähnen, dass sich die Objektivität bei älteren Kindern und Erwachsenen weitgehend erreichen lässt, die idealen Bedingungen bei jüngeren Kindern jedoch nicht immer realisiert werden können. Margraf-Stiksrud (2003, zitiert nach Esser & Petermann, 2010, S. 16) weist darauf hin, dass „ein völlig schematisches Vorgehen bei der Durchführung eines Entwicklungstests bei Kleinkindern vermutlich noch unzuverlässigere Werte ergeben würde als ein flexibles, der Situation angepasstes Vorgehen.“

Petermann und Macha (2005) sind der Ansicht, dass der Interpretationsobjektivität von Testresultaten eine besondere Bedeutung zukommt. Zwar liefern Entwicklungstests einen Leistungsstatus, der objektiv eingeordnet werden kann, jedoch ist eine entwicklungsbezogene Interpretation häufig nur bei Integration verschiedener Quellen möglich. Dies kann bei gleichen Entwicklungsprofilen zu unterschiedlichen Bewertungen führen (Petermann & Macha, 2005).

Laut Fuiko (2003, zitiert nach Esser & Petermann, 2010, S. 53) müssen sehr viele variierende Faktoren, wie z.B. Interesse, Motivation, Aufmerksamkeitsspanne, Aktivitätsniveau und die körperliche Befindlichkeit des Kindes berücksichtigt werden. Eine intensive Einarbeitung in das Verfahren sowie viel Erfahrung im Umgang mit Kindern sind unumgängliche Faktoren, die beachtet werden müssen, um ein möglichst objektives Vorgehen zu garantieren (Esser & Petermann, 2010). Laut Petermann und Macha (2005) werden jedoch nur wenige empirische Studien zur Überprüfung der Durchführungs- und Auswertungsobjektivität im Bereich der Entwicklungsdiagnostik vorgelegt.

2.2. Reliabilität

„Die Reliabilität eines Tests kennzeichnet den Grad der Genauigkeit, mit dem das geprüfte Merkmal gemessen wird.“ (Bortz & Döring, 2006, S. 196)

Eine grundlegende Voraussetzung für die Messgenauigkeit ist laut Esser und Petermann (2010) die präzise Durchführung und Auswertung der Testaufgaben (s. Kap. 2.1. Objektivität).

Anders als bei dem Gütekriterium der Objektivität, lässt sich bei diesem Hauptkriterium der Reliabilität ein Koeffizient errechnen, der zwischen 0 und 1 liegt, wobei ein Wert von 0 bedeutet, dass dieses Ergebnis ausschließlich aus Messfehlern besteht und ein Wert von 1 darauf hindeutet, dass kein Messfehler vorliegt und somit der Test das geprüfte Merkmal genau misst. In der Praxis sollten die Testverfahren einen Reliabilitätskoeffizienten von 0,8 aufwärts aufweisen, wobei Werte zwischen 0,8 und 0,9 als mittelmäßig und Reliabilitäten über 0,9 als hoch gelten (Weise, 1975, S. 219; zitiert nach Bortz & Döring, 2006, S. 199).

Im Folgenden werden drei Methoden zur Überprüfung der Reliabilität vorgestellt: die *Paralleltest-Reliabilität*, die *Retest-Reliabilität* und die *Innere Konsistenz*.

2.2.1. Paralleltest-Reliabilität

Hier wird davon ausgegangen, „dass zwar nicht eine Testwiederholung mit dem ursprünglichen Test möglich ist, jedoch mit einer völlig gleichwertigen Nachahmung des Tests (sog. „Paralleltest“).“ (Kubinger, 2009, S. 50)

Es werden also den Testpersonen zwei miteinander streng vergleichbare Tests vorgelegt und die Ergebnisse miteinander korreliert (Lienert & Raatz, 1994). Diese Korrelationen sind laut Kubinger (2009) jedoch „nur ein Kriterium dafür, wie gut die Konstruktion des Paralleltests gelungen ist, und weniger ein Kriterium für die Exaktheit der Messungen“ (S. 51).

Bei dieser Form der Reliabilitätsberechnung ist jedoch mit Lerneffekten zu rechnen, da die Lösungsprinzipien des einen Tests auf den Paralleltest übertragen werden können (Fisseni, 2004).

2.2.2. Retest-Reliabilität

Bei der Retest-Methode wiederholt die Testperson den gleichen Test nach einem gewissen Zeitabstand, in der Regel nach 2 bis 3 Monaten (Esser & Petermann, 2010). Die Korrelation der Ergebnisse zwischen dem Test und dem Retest werden berechnet und ergeben den Reliabilitätskoeffizienten.

Laut Bortz und Döring (2006) besteht bei dieser Methode die Gefahr der Überschätzung der Reliabilität, da die Lösungen der Testaufgaben vor allem bei kurzen Tests mit interessanten Items erinnert werden können. Dieser Erinnerungseffekt nimmt jedoch ab, je länger der zeitliche Abstand zwischen den Testvorgaben ist.

2.2.3. Innere Konsistenz

Eine weitere Möglichkeit die Reliabilität bzw. die Messgenauigkeit zu berechnen, ist die Beurteilung, wie sehr die einzelnen Teile bzw. Items eines Tests dasselbe messen.

Die „Split-Half“-Methode (Testhalbierungsmethode) ist laut Esser und Petermann (2010) die bekannteste Methode zur Reliabilitätsberechnung, wonach es nach der Bearbeitung des Tests zu einer Teilung dessen in zwei Teile kommt. Für jede Testhälfte werden die Ergebnisse einzeln berechnet, danach die Korrelation dieser beiden Testteile analysiert und auf die Gesamtlänge hin korrigiert.

Eine andere Möglichkeit zur Bestimmung der Reliabilität ist die sog. Konsistenzanalyse, wobei der Test nicht nur in zwei Teile zerlegt wird, sondern in so viele Teile wie er Items hat (Kubinger, 2009).

2.2.4. Reliabilität in der Entwicklungsdiagnostik

Die Möglichkeit der Reliabilitätsberechnung mittels Paralleltests ist laut Esser und Petermann (2010, S. 17) für den Bereich der Entwicklungsdiagnostik nicht geeignet, „da meist keine Paralleltests vorliegen“, weil es sich als sehr schwierig gestaltet, „tatsächlich indikationsidentische Items zu finden“ (vgl. Petermann & Macha, 2005).

Auch die Retest-Methode ist nicht anzuraten, da bei den Kindern „rasche Veränderungen in Wochen- oder Monatsschritten stattfinden, wobei die Entwicklungsgeschwindigkeit und die Verlaufsformen variieren.“ (Esser & Petermann, 2010, S. 17)

Wiederholt eine Person einen Test in kurzer Zeit erneut, so lassen sich zusätzlich weder Übungs-, noch Lerneffekte ausschließen (vgl. Petermann & Macha, 2005).

Die Methode der Konsistenzanalyse lässt sich im Bereich der Entwicklungsdiagnostik auch nur bedingt einsetzen, da meist nur wenige Items pro Altersstufe zur Verfügung stehen und die Skalen, als Voraussetzung der Konsistenzanalyse, meist nicht homogen sind (Esser & Petermann, 2010).

Am ehesten für die Überprüfung der Reliabilität von Entwicklungstests empfiehlt Ettrich (2000) die Split-Half bzw. Testhalbierungsmethode oder die Konsistenzanalyse.

2.3. Validität

„Unter Validität eines Tests ist zu verstehen, dass er tatsächlich jenes psychische Merkmal misst, welches er zu messen beansprucht.“ (Kubinger, 2009, S. 55)

Laut Lienert und Raatz (1994, S. 10) ist ein Test demnach „vollkommen valide, wenn seine Ergebnisse einen unmittelbaren und fehlerfreien Rückschluss auf den Ausprägungsgrad des zu erfassenden Persönlichkeits- oder Verhaltensmerkmals zulassen, wenn also der individuelle Testpunktwert eines Pb [Probanden] diesen auf der Merkmalsskala eindeutig lokalisiert.“

Esser und Petermann (2010), Kubinger (2009), Bortz und Döring (2006) und Fisseni (2004) sind sich einig, dass die Validität das wichtigste Gütekriterium darstellt, es jedoch am schwierigsten bzw. aufwendigsten zu überprüfen ist.

Wie bei der Reliabilität lassen sich auch bei der Validität Korrelationskoeffizienten errechnen. Validitäten zwischen 0,4 und 0,6 gelten als mittelmäßig und Koeffizienten über 0,6 als hoch (Weise, 1975, S. 219; zitiert nach Bortz & Döring, 2006, S. 202).

Im Folgenden werden die drei Hauptarten der Validität näher beschrieben: die *Inhaltliche Gültigkeit*, die *Konstruktvalidität* und die *Kriteriumsvalidität*.

2.3.1. Inhaltliche Gültigkeit

Von inhaltlicher Gültigkeit bzw. Inhaltsvalidität spricht Fisseni (2004, S. 63), „wenn der Inhalt der Test-Items das Zielmerkmal hinreichend genau definiert.“ Laut Lienert und Raatz (1994) und Kubinger (2009) soll der Test das optimale Kriterium des interessierenden Merkmals darstellen.

Berechnen lässt sich die inhaltliche Gültigkeit nicht, sie beruht allein auf subjektiven Einschätzungen (Bortz & Döring, 2006). Aus diesem Grund ist die Inhaltsvalidität laut

Bortz und Döring (2006) kein Testgütekriterium, sondern nur ein Ziel, das bei der Konstruktion eines Tests berücksichtigt werden sollte.

Inhaltliche Gültigkeit wird laut Kubinger (2009) am ehesten durch ein Experten-Rating erreicht, bei dem jedes Item dahingehend geprüft wird, „ob es in Bezug auf die gegebene operationale Definition dessen, was der Test messen soll, passt“ (S. 56).

2.3.2. Kriteriumsvalidität

Werden die Testergebnisse mit einem Kriterium, das unabhängig vom Test erhoben wird, korreliert, so nennt man diese Möglichkeit der Validitätsberechnung „Kriteriumsvalidität“. Der Vorteil dieser Methode ist, dass sie eine statistische Kennzahl liefert, die das Ausmaß der Validität eines Tests bestimmt (Kubinger, 2009; Lienert & Raatz, 1994). Jedoch erweist es sich laut Kubinger (2009) und Bortz und Döring (2006) als schwierig, ein geeignetes Außenkriterium zu finden, das zusätzlich noch messgenau ist.

Meist wird die Validität daran gemessen, ob der Testwert das spätere Verhalten korrekt voraussagt. Diese Form der Validität wird prognostische Validität genannt, während die Übereinstimmungsvalidität eine Form ist, bei der der Testwert und der Kriteriumswert zum selben Messzeitpunkt erhoben werden (Bortz & Döring, 2006).

Lienert und Raatz (1994) berichten, dass die kriteriumsbezogene Validität eines Tests umso größer ist, je größer die Gemeinsamkeit des von Test und Kriterium erfassten Merkmalsanteils. Sie ist andererseits umso geringer, je geringer die Reliabilitätskoeffizienten ausfallen.

2.3.3. Konstruktvalidität

„Erfüllt ein Test (nicht nur einen pragmatischen Anspruch, sondern) gewisse theoretische bzw. theoriegeleitete Vorstellungen in Bezug auf irgendein „Konstrukt“, dann ist ihm Konstruktvalidität zuzusprechen.“ (Kubinger, 2009, S. 57)

Es lassen sich zwei Formen der Konstruktvalidität unterscheiden:

Konvergente Konstruktvalidität entsteht, wenn der Test hoch mit konstruktnahen Variablen korreliert und diskriminante Konstruktvalidität entsteht, wenn sich negative oder nur gering positive Korrelationen des Tests mit den Variablen zeigen, mit denen er keine theoretische Übereinstimmung aufweist (Esser & Petermann, 2010).

Zur Bestimmung wird häufig die Faktorenanalyse herangezogen (Esser & Petermann, 2010), Bortz und Döring (2006) empfehlen die Multi-Trait-Multi-Method-Methode und Kubinger (2009) führt weitere nicht korrelative Ansätze an.

Laut Fisseni (2004) beinhaltet die Konstruktvalidität die inhaltliche und die kriteriumsbezogenen Validität.

2.3.4. Validität in der Entwicklungsdiagnostik

Laut Esser und Petermann (2010) und Hagmann-von Arx et al. (2008) lässt sich die inhaltliche Gültigkeit häufig bei Entwicklungstests finden. Jedoch wird sie meist nur mit der Tatsache begründet, dass mit zunehmendem Alter komplexere sowie eine höhere Anzahl von Aufgaben gelöst werden können (Esser & Petermann, 2010; Sarimski, 2009).

Als Beleg für die Kriteriumsvalidität dürften laut Sarimski (2009) bereits mittlere Korrelationskoeffizienten angesehen werden. Als Begründung gibt er an, dass sich im Säuglings- und Kleinkindalter eher selten Untersuchungen finden lassen, bei denen die Resultate eines Entwicklungstests mit denen eines anderen verglichen werden. Vollständige Übereinstimmung ist dabei nicht zu erwarten, weil sich die Konzepte der verschiedenen Verfahren unterscheiden.

Vor allem die prognostische Validität ist im Rahmen der Früherkennung von besonderem Interesse, jedoch erfüllen nicht alle entwicklungsdiagnostischen Untersuchungen dieses Kriterium (Esser & Petermann, 2010). Der Grund dafür ist laut Sarimski (2009), dass die prognostische Aussagekraft von entwicklungsdiagnostischen Verfahren „durch die besonders schnellen entwicklungsbedingten Veränderungen sowie durch die Schwierigkeit begrenzt ist, Vorläufer von komplexeren Fähigkeiten in den Bereichen Wahrnehmung, Kognition und Sprache zuverlässig abzubilden“ (S. 131).

Laut Esser und Petermann (2010) fällt die prognostische Validität vor allem dann gering aus, wenn Kinder getestet werden, die jünger als 2 Jahre alt sind und die Abstände zwischen den Untersuchungen sehr lange ausfallen.

Skranes (1998, zitiert nach Sarimski, 2009, S. 131) berichtet, dass sich „erst bei Kindern, die jenseits des zweiten Lebensjahres mit den Bayley-Entwicklungsskalen untersucht wurden, recht hohe Korrelationen mit später erhobenen Intelligenztestwerten“ ergaben.

Renziehausen und Petermann (2007) sind der Ansicht, dass Entwicklungsprognosen umso zuverlässiger sind, je deutlicher die Entwicklungsabweichungen sind. Deshalb sollte der Untersucher eher von einem gegenwärtigen Entwicklungsstand sprechen, ohne damit eine Prognose für die Leistungsentwicklung im Vor- und Grundschulalter zu stellen (Sarimski, 2009).

Im Bezug auf die Konstruktvalidität bei Entwicklungstests beschreiben Esser und Petermann (2010, S. 20), dass immer wieder die Frage aufkam, „ob der allgemeine Entwicklungsstand überhaupt als latente Variable angenommen werden kann, da es sich bei diesem Merkmal um ein sehr heterogenes Merkmal handelt, das sich zudem im Verlauf der Entwicklung eines Kindes qualitativ verändert.“ Laut Macha, Proske und Petermann (2005) und Hagmann-von Arx et al. (2008) gehen alle Entwicklungstests von stetig zunehmenden Testleitungen mit dem Alter aus. Weisen also benachbarte Altersbereiche ähnliche Testaufgaben auf, so lässt sich dieser Alterstrend meist unmittelbar aus den Normen ablesen und muss somit nicht gesondert überprüft werden.

2.4. Normierung

Zusätzlich zu den drei oben beschriebenen Hauptkriterien, soll hier auch näher auf die Normierung bzw. die Eichung, dessen Begriff Kubinger (2009) vorzieht, eingegangen werden.

„Ein Test erfüllt das Gütekriterium *Eichung*, wenn für sein Bezugssystem zur Relativierung des individuellen Testergebnisses (die sog. „Eichtabellen“) folgende Bedingungen gegeben sind:

- Die Eichtabellen sind gültig, d.h. nicht veraltet,
- die Population, für die die Eichtabellen gelten, ist definiert,
- die für die Erstellung der Eichtabellen herangezogene Stichprobe ist repräsentativ.“ (Kubinger, 2009, S. 68)

Zusätzlich stellt für Esser und Petermann (2010) auch die Größe der Stichprobe ein wesentliches Qualitätskriterium für die Normierung dar. Als gering gelten laut Fisseni

(2004) weniger als 150 Personen pro Gruppe, als mittel 150-300 Personen und darüber liegende Stichprobengrößen werden als hoch beurteilt.

Die Aktualität der Normen haben deshalb einen sehr wichtigen Stellenwert, da bei Flynn (1987, zitiert nach Esser & Petermann, 2010, S. 21) Normvergleiche gezeigt haben, dass bei Intelligenztests die Testnormen schon innerhalb von zwei Jahrzehnten zu einer Überschätzung der Intelligenz um zehn Punkte führen können. Diese Überschätzung der Fertigkeiten lassen sich jedoch nicht nur bei Intelligenztests, sondern auch bei Entwicklungstests beobachten (Esser & Petermann, 2010). Deshalb sollten laut Macha et al. (2005) spätestens alle zehn Jahre Normierungen der Tests vorgenommen werden.

Normierung in der Entwicklungsdiagnostik

Um Aussagen darüber treffen zu können, über welche individuellen Stärken und Schwächen ein Kind verfügt, sind Vergleiche mit der entsprechenden Referenzgruppe mittels der Normtabellen erforderlich. Damit es zu keinen Überschätzungen der Fertigkeiten und Fähigkeiten eines Kindes kommt und damit niedrige und sehr niedrige Leistungen präzise erfasst werden können, sind aktuelle Normierungen mit großen Stichproben äußerst wichtig, die diesen Leistungsbereich hinreichend repräsentieren (Esser & Petermann, 2010).

2.5. Kritische Betrachtung der allgemeinen Entwicklungstests in Hinblick auf die Gütekriterien

Im Folgenden sollen nun die im deutschsprachigen Raum weit verbreiteten allgemeinen Entwicklungstests für Zweijährige in Hinblick auf die soeben beschriebenen Gütekriterien kritisch betrachtet werden.

2.5.1. ET 6-6 (Petermann, Stein & Macha, 2008)

Die Autoren machen im Manual keine Angaben zur Objektivität, jedoch ist davon auszugehen, dass sie in ausreichender Höhe gewährleistet ist (Sarimski, 2001; Lücking & Scheithauer, 2006; Esser & Petermann, 2010).

Im Handbuch des ET 6-6 wird darauf hingewiesen, dass eine Überprüfung der Reliabilität mittels Konsistenzmaße und Splithalf-Reliabilitäten aufgrund der inhaltlich heterogenen Skalen keine angemessenen Zugänge sind. Auch wurde auf die Berechnungen der Retest-

Reliabilitäten mit der Begründung verzichtet, dass aufgrund der zu erwartenden Lern- und Erinnerungseffekte auch diese nicht angemessen seien (Esser & Petermann, 2010).

Die Konstruktvalidität ist laut Macha et al. (2005) gegeben, da geschlechtsspezifische Normen vorliegen und sowohl Alterstrends als auch Entwicklungsverläufe berücksichtigt wurden. Auch dokumentiert der ET 6-6 Zusammenhänge mit einem etablierten Leistungstest (Kaufman Assessment Battery for Children; K-ABC; Melchers & Preuß, 2003) und lässt somit auf eine kriteriengebundene Validität schließen. Inhaltliche Gültigkeit kann diesem Verfahren, wie allen Verfahren in der Entwicklungsdiagnostik, zugesprochen werden (Petermann & Macha, 2005).

Im Weiteren ist der ET 6-6 laut Macha et al. (2005) in der Lage, frühgeborene von reifgeborenen Kindern zu unterscheiden. Dieses Verfahren differenziert also zwischen Gruppen, jedoch nicht innerhalb von Gruppen.

Die Normwerte aus den Jahren 1999 und 2000 (Esser und Petermann, 2010) sind nun älter als zehn Jahre und somit bedarf es einer Neunormierung. Da die herangezogene Stichprobe nur aus drei deutschen Regionen stammt, ist auch eine höhere Repräsentativität der Normstichprobe wünschenswert (Lücking & Scheithauer, 2006).

Macha et al. (2005) sehen im ET 6-6 ein empfehlenswertes Verfahren, das eine erste Orientierung über ein breites Leistungsspektrum im Rahmen einer Eingangsdiagnostik erlaubt. Auch Hagmann-von Arx et al. (2008) können dieses Verfahren aufgrund der richtungsweisenden Validität empfehlen. Lücking und Scheithauer (2006) wünschen sich jedoch weitere und differenzierende Befunde zur Reliabilität und Validität sowie zu den Normwerten.

Im März 2013 wurde der ET 6-6-R („Entwicklungstest für Kinder im Alter von sechs Monaten bis sechs Jahren – Revision“) von Petermann und Macha veröffentlicht, der wesentliche Veränderungen zur letzten Version des ET 6-6 von 2008, wie z.B. eine Neunormierung und eine umfassende Dokumentation der Reliabilität, enthält.

2.5.2. MFED 2-3 (Hellbrügge, 1994)

Das Handbuch der MFED 2-3 enthält Aussagen über die Durchführungs- und Auswertungsobjektivität, die durch den Aufbau des Verfahrens als gewährleistet angesehen werden können (Esser & Petermann, 2010). Rennen-Allhoff (1990, zitiert nach

Esser & Petermann, 2010, S. 73) und Berth und Kramp (1990, zitiert nach Esser & Petermann, 2010, S. 73) bezeichnen die Objektivität der Untertests als hoch.

Ergebnisse zur Retest-Reliabilität und interne Konsistenz werden von den Autoren nicht mitgeteilt (Esser & Petermann, 2010). Auch werden keine Angaben zur Validierung (Hagmann-von Arx et al., 2008; Macha et al., 2005) und zur Differenzierungsfähigkeit der auffälligen von nicht auffälligen Kindern gemacht (Macha et al., 2005).

Dem Kriterium der Normierung kann die MFED 2-3 nicht gerecht werden (Hagmann-von Arx et al., 2008), da das Verfahren einerseits sehr veraltet ist (ca. 1970), die Normen keine empirisch ermittelten Daten sind (Macha et al., 2005) und die Zusammensetzung dieser Werte nicht näher erläutert wird (Esser & Petermann, 2010).

Laut Hagmann- von Arx et al. (2008) ist dieses Verfahren nur bedingt zu empfehlen, da die Testgütekriterien nicht zufriedenstellend und die Normen nicht hinreichend sind. Zusätzlich führen Macha et al. (2005) als Schwäche dieses Verfahrens das veraltete Stufenleiterkonzept an (s. Kap. 1.1. Definition des Entwicklungsbegriffes). Jedoch hat sich die MFED aufgrund der Übersichtlichkeit und dem hohen Maß an Praktikabilität in den kinderärztlichen Praxen etabliert (Esser & Petermann, 2010).

2.5.3. Bayley-II (Reuner, Rosenkranz, Pietz & Horn, 2007)

Laut Esser und Petermann (2010) weisen die Aufgaben der Bayley-II ein hohes Maß an Standardisierung auf. Die Auswertung und die Interpretation erfolgen regelgeleitet und somit lässt sich die Objektivität dieses Verfahrens als gewährleistet ansehen.

Sowohl zur Reliabilität als auch zur Validität liegen umfangreiche Ergebnisse vor.

Laut Esser und Petermann (2010) weist die Bayley-II Retest-Reliabilitäten und interne Konsistenzen im zufriedenstellenden Bereich auf, jedoch muss von einer Überschätzung der internen Konsistenzen ausgegangen werden, da nicht getestete Aufgaben als erfüllt gelten. Angaben zu Korrelationen mit anderen Entwicklungs- und Leistungstests lassen auf eine gute Validität schließen (Esser & Petermann, 2010). Auch die Differenzierungsfähigkeit wird anhand spezifischer Risikogruppen überprüft und fällt tendenziell hypothesenkonform aus (Esser & Petermann, 2010). Bezüglich der Konstruktvalidität zeigen sich laut Hagmann-von Arx et al. (2008) geringe bis mittlere Skalenkorrelationen. Die Inhaltsvalidität ist laut Petermann und Macha (2005) gegeben.

Die Normen der deutschen Version bestehen laut Esser und Petermann (2010) aus den Originalnormen der amerikanischen Version (BSID II) aus den Jahren 1991 und 1992. Deutsche Normen liegen nicht vor (Hagmann-von Arx et al., 2008) und lassen deshalb verbale Leistungen für deutschsprachige Kinder nur eingeschränkt interpretieren. Außerdem liegt im englischsprachigen Raum eine weiter entwickelte Form mit dem BSID-III (siehe 2.5.4. Bayley-III) vor, die inhaltlich deutlich ergänzt und differenziert wurde (Esser & Petermann, 2010).

Esser und Petermann (2010) können die deutsche Bayley-II dort empfehlen, „wo entwicklungspsychologisch und testdiagnostisch besonders kenntnisreiche Anwender mit einem ausreichenden Zeitbudget umfangreiche Beobachtungen unter standardisierten Bedingungen vornehmen wollen“ (S. 90). Jedoch lässt sich der Einsatz dieses Verfahrens bei Verwendung der Originalnormen nur mit großer Umsicht anraten (Esser & Petermann, 2010).

2.5.4. Bayley-III (Bayley, 2006)

Aufgrund der genauen Anleitung in der Durchführung, Auswertung und Interpretation dieses Verfahrens, lässt sich die Objektivität als gegeben ansehen.

Die interne Konsistenz kann für alle Skalen und Altersgruppen mit Werten zwischen .86 und .93 angenommen werden (Bayley, 2006). Für die „Adaptive Behavior Scale“ fallen die internen Konsistenzen für Kinder unter drei Monaten mit Werten $< .70$ jedoch gering aus (Kastner-Koller & Deimann, 2011). Die Retest-Reliabilitäten für die Skalen der kognitiven, sprachlichen und motorischen Entwicklung und der Einschätzung des Alltagsverhaltens durch die Eltern fallen zufriedenstellend aus (Albers & Grieve, 2007). Für die Skala der sozioemotionalen Entwicklung werden keine Retest-Reliabilitäten angegeben (Bayley, 2006; Albers & Grieve, 2007).

Laut Petermann und Macha (2005) ist die inhaltliche Validität, wie bei jedem Entwicklungstests, als gegeben anzusehen. Bayley (2006) macht weitere, sehr ausführliche Angaben zur zufriedenstellenden Validität.

Die Normierung wurde im Jahre 2004 vorgenommen, an der 1700 Kinder teilnahmen (Bayley, 2006), weshalb die Stichprobe laut Fisseni (2004) als sehr groß beurteilt werden kann. Die Bayley-III kann dem Gütekriterium der Normierung im deutschsprachigen Raum jedoch nicht entsprechen, da diese nur für den amerikanischen Raum vorliegt. Jene

Aufgaben, die verbale Leistungen der Kinder abverlangen, können somit, wie bei der Bayley-II, nur eingeschränkt interpretiert werden. Zusätzlich ist anzumerken, dass eine baldige Neunormierung angedacht werden sollte, da die letzte nun neun Jahre zurück liegt.

Sehr positiv anzumerken ist allerdings, dass eine Einschätzung der Fähigkeiten der Kinder bis zu einem Alter von 5 Monaten und 15 Tagen in 10-Tages-Abschnitten vorgenommen werden kann (Bayley, 2006).

Obwohl die Bayley-III sehr nützliche und notwendige Verbesserungen zur Bayley-II enthält (Albers & Grieve, 2007), ist die Verwendung der originalen, amerikanischen Normen nur eingeschränkt zu empfehlen (Esser & Petermann, 2010). Auch zu bedenken ist, dass die Aufgaben nicht in übersetzter Form vorliegen (Kastner-Koller & Deimann, 2011).

Zusammenfassung

Alle Verfahren, die in der Diagnostik eingesetzt werden, müssen gewissen Qualitätsanforderungen entsprechen. Um ein Verfahren als qualitativ gut oder schlecht beurteilen zu können, ist die Betrachtung der Hauptgütekriterien, nämlich der Objektivität, der Reliabilität und der Validität, unumgänglich.

Die Objektivität unterteilt sich in die Durchführungs-, Auswertungs-, und Interpretationsobjektivität. Bei der Durchführungsobjektivität handelt es sich um die Unabhängigkeit des Verhaltens und der Testleistungen der Testperson vom Untersucher. Die Auswertungsobjektivität meint, dass unterschiedliche Testleiter bei der Auswertung desselben Tests zu denselben Ergebnissen kommen und auch die gleichen Schlüsse daraus ziehen (Interpretationsobjektivität).

Unter dem Gütekriterium der Reliabilität versteht man die Genauigkeit, mit der ein Test ein Merkmal misst. Hierfür gibt es mehrere Möglichkeiten, diese zu berechnen: die Retest-Reliabilität, die Paralleltestreliabilität und die innere Konsistenz. Der errechnete Koeffizient liegt zwischen 0 und 1, im günstigsten Fall über 0,9. Auch ein Wert über 0,8 ist zufriedenstellend.

Die Validität ist das wichtigste, jedoch am schwierigsten zu überprüfende Gütekriterium. Sie beschreibt, inwiefern ein Test das Merkmal misst, das er auch zu messen beansprucht.

Um dieses Gütekriterium zu überprüfen, lassen sich die folgenden Möglichkeiten anführen: die inhaltliche Gültigkeit, die Konstrukt- und die Kriteriumsvalidität.

Im Weiteren ist in der Diagnostik auch das Nebenkriterium der Normierung sehr wichtig. Hierfür ist es vor allem wesentlich, dass die Normen aktuell sind und einer großen Stichprobe entstammen, um zu verhindern, dass Fehleinschätzungen gemacht werden.

Die kritische Betrachtung der allgemeinen Entwicklungstests für zweijährige Kinder im deutschsprachigen Raum hat ergeben, dass alle vier Verfahren (ET 6-6, MFED 2-3, Bayley-II und -III) methodische Mängel aufweisen, die die Anwendung dieser nur bedingt empfehlen.

3. Spielbasierte Diagnostik

Da soeben gezeigt werden konnte, dass die Verfahren zur Erfassung der Entwicklung zweijähriger Kinder den Gütekriterien nicht zufriedenstellend entsprechen, wird nun auf einen weiteren Zugang in der Entwicklungsdiagnostik, der spielbasierten Diagnostik, eingegangen.

Der Itempool, der von Kuchler et al. (2011) erstellt wurde und für diese Arbeit zum Einsatz kam, folgt diesem Ansatz.

Was ist spielbasierte Diagnostik?

Unter einem spielbasierten diagnostischen Verfahren versteht Sturzbecher (2001) die Einkleidung der diagnostischen Intentionen in Spiele als Handlungsrahmen, wobei die Fähigkeiten der Kinder beobachtet werden (Kelly-Vance & Ryalls, 2005).

Die Beschäftigung mit Spielsachen und Gegenständen ist für jüngere Kinder eine äußerst wichtige (Vig, 2007) und darüber hinaus sehr natürliche und vertraute Aktivität (Lowenthal, 1997) und kann somit als die dominierende und beliebteste Tätigkeit eines Vorschulkindes bezeichnet werden (Witzlack, 2001).

Das Testen mit standardisierten Verfahren stellt eine prüfungsähnliche, unnatürliche Situation für die Kinder dar, in der komplexe Anforderungen gestellt werden (Kelly-Vance et al., 2002). Im Gegensatz dazu können Beobachtungen von Kindern in natürlichen Situationen Informationen beinhalten, die sehr bedeutsam sein können (Linder, 1993; zitiert nach Lowenthal, 1997, S. 43). Auch Vig (2007) ist der Ansicht, dass

aus der Verhaltensbeobachtung von Kindern in Spielsituationen wertvolle Schlüsse bezüglich der Entwicklung gezogen werden können, weshalb Irblich und Renner (2009) bei Kindern unter vier Jahren einen spielerischen Zugang zur Testung empfehlen.

Das Kind soll sich freiwillig in die Spielsituation begeben (Sturzbecher, 2001), die im Inhalt und in der Struktur flexibel an die Bedürfnisse und die Persönlichkeit des Kindes angepasst wird (Lowenthal, 1997; Cherney, Kelly-Vance, Glover, Ruane & Ryalls, 2003). Im Gegensatz dazu werden bei standardisierten Verfahren die Aufgaben, meist mit ansteigendem Schwierigkeitsgrad, strikt der Reihenfolge nach vorgegeben (Witzlack, 2001).

Aufgrund der Flexibilität in der Abfolge hat das Spiel eine stark motivierende Komponente für die Kinder (Bronfenbrenner, 1979; zitiert nach Kelly-Vance et al., 2002, S. 170) und wirkt im Gegensatz zu traditionellen Tests nicht bedrohlich (Lowenthal, 1997). Das Spielen bewirkt ein kooperativeres Verhalten von Seiten der Kinder, das auch bei Bagnato und Neisworth (1994) bestätigt werden konnte. In dieser groß angelegten Studie wurde bei standardisierten Tests festgestellt, dass 43% der Stichprobe als untestbar gelten.

Weitere entscheidende Vorteile sind die Anwendung spielbasierter Verfahren bei entwicklungsverzögerten Kindern (Vig, 2007), die adaptiv zu gestaltende Durchführung und das Vermeiden von Misserfolgserlebnissen (Witzlack, 2001). Dies ermöglicht den Kindern sowohl ihr Optimum an Fähigkeiten (Vygotsky, 1967; zitiert nach Kelly-Vance et al., 2002, S. 170; Cherney et al., 2003) als auch ihr Wesen und ihre Individualität, ihre Bedürfnisse und Wünsche, ihre Fähigkeiten und Fertigkeiten, ihr Temperament und ihre Ausdauer zu zeigen (Witzlack, 2001).

Farmer-Dougan und Kaszuba (1999), Kelly-Vance et al. (2002) und Dykeman (2008) sind der Ansicht, dass durch die spielbasierte Diagnostik mehr relevante Aussagen gewonnen werden und somit genauere Einschätzungen der Fähigkeiten und Fertigkeiten der Kinder möglich sind. Traditionelle standardisierte Verfahren hingegen unterschätzen den wahren Entwicklungsstand eines Kindes (Kelly-Vance, Needelman, Troia & Ryalls, 1999; zitiert nach Cherney et al., 2003, S. 95) und nehmen zudem mehr Zeit in Anspruch (Kelly-Vance et al., 2002). Zusätzlich resultieren bei standardisierten Tests lediglich Werte, während bei spielbasierten Verfahren die Beschreibung der Stärken und

Schwächen eines Kindes ausführlich möglich sind (Kelly-Vance et al., 2002; Dykeman, 2008). Spielbasierte Verfahren haben sich laut Dykeman (2008) also als zumindest sinnvolle Ergänzung in der Untersuchung von Kindern herausgestellt und auch Vig (2007) ist der Ansicht, dass dieser Zugang ein sehr wertvoller ist.

Der Bekanntheitsgrad von spielbasierten Verfahren steigt, obwohl empirische Belege zu den Gütekriterien meist fehlen (Farmer-Dougan & Kaszuba, 1999; Kelly-Vance et al., 2002; Kelly-Vance & Ryalls, 2005).

Um die Durchführung zu standardisieren, listet Witzlack (2001) fünf Komponenten auf, die kontrolliert werden müssen, um objektives Untersuchen zu ermöglichen: der Spielinhalt, die Spielrolle des Kindes und des Spielpartners, die Spielgegenstände und die Abfolge des Angebots der Spielgegenstände.

Laut Kelly-Vance et al. (2002) können bei spielbasierten Verfahren sowohl die Auswertungsobjektivität als auch die Retest-Reliabilität (Kelly-Vance et al., 2005) erreicht werden.

Die Kriteriumsvalidität wurde von Kelly-Vance et al. (1999, zitiert nach Kelly-Vance et al., 2002, S. 172) untersucht. Dabei wurden die Ergebnisse eines spielbasierten Verfahrens, dem Transdisciplinary Play-Based Assessment (*TPBA*; s. Kap. 3.2. Transdisciplinary Play-Based Assessment), und der Bayley-II verglichen und eine hohe Korrelation festgestellt. Somit kann diese Art der Validität angenommen werden.

Auch die soziale Validität kann laut Myers et al. (1996, zitiert nach Kelly-Vance et al., 2002, S. 172) für zumindest das TPBA angenommen werden, da sowohl die Anwender als auch die Eltern aufgrund der kürzeren Inanspruchnahme der Zeit und genaueren Einschätzungen der Entwicklung diesem Verfahren sehr positiv entgegengetreten (Kelly-Vance et al., 2002).

3.1. Geschichte der spielbasierten Diagnostik

Witzlack gibt 2001 einen Überblick über das Entstehen des Ansatzes der spielbasierten Diagnostik:

Melanie Klein gilt mit ihrer Kinderspielanalyse (1919; 1932; „Analytische Spieltechnik“, 1976) als Begründerin der Spieldiagnostik, da sie als Erste Spiele und Spielsachen bei Kindern ab zwei Jahren für therapeutische und diagnostische Zwecke nutzte.

Fast gleichzeitig zu Klein hat auch Margaret Lowenfeld mit ihrem Weltspiel (1925; 1935; „Welt“-Technik, 1976) den spieldiagnostischen Ansatz gefunden, der die Entwicklung der spielbasierten Diagnostik am nachhaltigsten beeinflusste.

Beide Methoden sind grundverschieden, da Klein jede Spielaktivität des Kindes psychoanalytisch deutete, während sich Lowenfeld ausdrücklich von psychoanalytischen Interpretationen distanzierte.

Charlotte Bühler beschäftigte sich ab 1935 intensiv mit dem Weltspiel von Lowenfeld, standardisierte neben der Instruktion auch die Auswertung und nannte das Verfahren „Welt-Test“ (1936; 1952; 1955).

Alle danach folgenden Ansätze, wie z.B. der Dorftest von Arthus (1939; 1949), das Handpuppenspiel von Rambert („Kasperle-Test“, 1969), der Scenotest von v. Staabs (1938; 1943; 1951; 1985) oder die Ericamethode von Harding (1945; 1959; „Dreiweltentest“, 1972) zeigten kaum Fortschritte oder neue Ansätze.

In den 1980er Jahren war es der Arbeitsgruppe um Witzlack wichtig, das methodologische Prinzip zu verfolgen, für die Entscheidung über die Schulfähigkeit eines Kindes alle verfügbaren diagnostischen Quellen mit einzubeziehen. Aufgrund dessen entstanden der spielbasierte „Familien-Interaktionstest für Kinder“ (FIT-K, Sturzbecher, 1993) und der „Spielbeobachtungsbogen zur Analyse des psychischen Entwicklungsstandes drei- bis vierjähriger Vorschulkinder“ (SBB, Großmann, 1987). 1988 veröffentlichte Witzlack den Einschulungstest „Kaufmannsladen“, der ein Kernstück seiner spielorientierten Entwicklungsdiagnostik war.

Auf ein weiteres spielbasiertes Verfahren soll im Folgenden eingegangen werden.

3.2. Transdisciplinary Play-Based Assessment

Linder publizierte 1993 erstmals ihr mittlerweile sehr bekanntes spielbasiertes Verfahren „Transdisciplinary Play-Based Assessment“, dessen Überarbeitung 2008 folgte, und das bei Kindern zwischen Null und sechs Jahren angewandt werden kann.

Begonnen wird mit einer unstrukturierten Spielsituation, in der das Kind die Kontrolle übernimmt. Nach 30 Minuten wird auf Spielaktivitäten hingewiesen, die das Kind noch nicht gezeigt hat. Während der Spielsituationen wird das Verhalten des Kindes beobachtet und laut Anleitungen des Manuals kodiert. Es werden die Fähigkeiten der Kinder in den Fähigkeitsbereichen der kognitiven Entwicklung (Imitation, Problemlösen,

Diskriminierungs- und Klassifizierungskompetenz, Eins- zu- eins Zuordnung, Zeichnen und das Verstehen von Abfolgen), der Sprache, des Sozialverhaltens, der Fein- und der Grobmotorik erhoben.

Das Ziel des TPBA ist es sowohl Stärken als auch Schwächen zu erfassen und diese detailliert zu beschreiben. Aufgrund dessen können fundierte Entscheidungen getroffen werden, welche Interventionen für welche Fähigkeitsbereiche eingeleitet werden müssen und später, inwieweit das Kind Fortschritte gemacht hat.

Zusammenfassung

Ein weiterer Zugang der Entwicklungsdiagnostik ist die spielbasierte Diagnostik, worunter die Einkleidung diagnostischer Intentionen in Spielsituationen verstanden wird. Melanie Klein gilt als Begründerin dieses Ansatzes, wenn gleich Margret Lowenfeld ihn am nachhaltigsten beeinflusste.

Da die Beschäftigung mit Spielsachen eine sehr natürliche und vertraute Aktivität für junge Kinder darstellt und sich die Kinder freiwillig in die Befragungssituation begeben, ist es ihnen möglich, aufgrund ihrer Motivation, ihr Optimum an Fähigkeiten zu zeigen. Der Inhalt und die Struktur des Spiels werden den Bedürfnissen der Kinder angepasst. Dies bedeutet eine adaptiv zu gestaltende Durchführung und das Vermeiden von Misserfolgserlebnissen.

Aufgrund der spielbasierten Diagnostik können mehr relevante Aussagen, genauere Einschätzungen der Fähigkeiten und Fertigkeiten und somit detailliertere Beschreibungen des Entwicklungsstandes des Kindes gemacht werden. Als Beispiele für spielbasierte Verfahren können der *Kaufmannsladen* und das *Transdisciplinary Play-Based Assessment* in der zweiten Version angeführt werden.

Negativ anzumerken ist, dass empirische Belege zu den Gütekriterien dieser Verfahren meist fehlen.

4. Fähigkeitsbereich der visuellen Wahrnehmung

Ein weiteres spielbasiertes Verfahren, das sich noch in der Entwicklung befindet, ist der spielbasierte Itempool zur Erfassung der Entwicklung Zweijähriger von Kuchler et al. (2011). Dieses soll im Rahmen dieser Arbeit überarbeitet und testtheoretisch überprüft werden. Fähigkeitsbereiche, die damit erfasst werden, sind die der kognitiven Entwicklung, der visuellen Wahrnehmung, der Motorik, der Sprache, der sozioemotionalen Entwicklung, der Aufmerksamkeit und des Gedächtnisses.

In dieser Arbeit wird besonderes Augenmerk auf den Fähigkeitsbereich der visuellen Wahrnehmung gelegt, weshalb im Folgenden detailliert auf die Farbwahrnehmung, die Farbpräferenz, die aktive und passive Farbdifferenzierung, die kategorische Farbwahrnehmung, die Form- und die Größenwahrnehmung eingegangen wird. Um zu erfahren, wie diese bei Kuchler et al. (2011) erhoben werden, wird auf das Kap. 8.3.1. Erfassung der visuellen Wahrnehmung verwiesen. Zusätzlich werden am Ende dieses Kapitels die Möglichkeiten der Diagnostik beschrieben.

Näheres über den Fähigkeitsbereich der kognitiven Entwicklung ist bei Punz (in Vorbereitung) zu lesen.

Berk (2005) ist der Ansicht, dass Menschen beim Erkunden der Umwelt vom Sehsinn mehr als von jedem anderen Sinn abhängig sind. Leiden Kinder an einer minimalen oder fehlenden Sehkraft, so scheint dies die Erfahrung in der kognitiven und motorischen Entwicklung zu verändern, so dass die Meilensteine der Entwicklung erst später als bei Gleichaltrigen erlangt werden (Hatton, Bailey, Burchinal & Ferrell, 1997; zitiert nach Berk, 2005, S. 183).

4.1. Farbwahrnehmung

In der Wahrnehmung unserer Umwelt ist die Farbe omnipräsent. Wir haben alltäglich mit ihr zu tun, wenn wir unsere Kleidung farblich abstimmen oder uns an der Ampel orientieren. Wir ordnen den Farben Gefühle zu (z.B. steht Rot für Wut) und bringen sie mit einzelnen Bedeutungen in Verbindung, wie z.B. wird mit der Farbe Rot eine drohende Gefahr signalisiert (Goldstein, 2008).

Welche Farbe wir wahrnehmen, hängt zum Großteil von den Wellenlängen des Lichts ab, die ins Auge reflektiert werden (Pinel & Pauli, 2007). Diese Wellenlängen von etwa 380nm bis 750nm lösen im Auge Sinnesempfindungen aus (Fischer, 1995, S. 61). In welcher Beziehung diese Wellenlängen mit der Aktivität des Nervensystems stehen, behandeln zwei verschiedene Theorien des Farbsehens (Pinel & Pauli, 2007):

- *Die Dreifarbentheorie* wurde 1802 von Thomas Young aufgestellt und 1852 von Hermann von Helmholtz verbessert. Nach dieser Theorie gibt es drei verschiedene Arten von Farbrezeptoren (Zapfen), die alle eine unterschiedliche spektrale Sensitivität besitzen. Die Farbe eines bestimmten Reizes wird durch das Verhältnis der Aktivität aller drei Rezeptorarten bestimmt.
- *Die Gegenfarbentheorie* wurde 1878 von Ewald Hering vorgeschlagen. Er ging davon aus, dass es zwei verschiedene Arten von Zellen zur Farbkodierung gibt und eine weitere zur Helligkeitskodierung. Jede der drei Arten von Zellen kodiert zwei komplementäre Farbwahrnehmungen (Rot-Grün, Blau-Gelb, Schwarz-Weiß).

Für beide Theorien konnten physiologische Belege gefunden werden (Goldstein, 2008).

Laut Pinel und Pauli (2007) können achromatische Farben (Schwarz, Weiß, Grau) und chromatische Farben (Blau, Grün, Gelb, Rot) unterschieden werden. Bei Abwesenheit von Licht wird Schwarz wahrgenommen, die Wahrnehmung von Weiß entsteht durch eine intensive Mischung einer großen Bandbreite an Wellenlängen in einem annähernd gleichen Anteil und Grau wird durch dieselbe Mischung bei niedrigeren Intensitäten wahrgenommen.

Laut Goldstein (2008) werden Rot, Gelb, Grün und Blau als die vier Grundfarben betrachtet, wobei sich Blau am kurzwelligen Ende, Grün im mittleren Bereich und Gelb und Rot am langwelligen Ende des Spektrums befinden.

Insgesamt können wir Menschen etwa 200 verschiedene Farben erkennen und unterscheiden (Goldstein, 2008). Die Wichtigkeit der Farbe wird laut Fischer (1995) meist jedoch erst dann erkannt, wenn diesbezüglich eine Störung bzw. ein Unvermögen auftritt, sämtliche Farben zu erkennen bzw. zu unterscheiden.

4.2. Farbpräferenz

In der Literatur zeigen sich unterschiedliche Ergebnisse in Bezug darauf, welche Farben von Kindern präferiert werden.

Mehrfach wird bestätigt, dass Kinder chromatische Stimuli und Farben mit einer längeren Wellenlänge (Rot und Gelb) achromatischen Stimuli und Farben mit einer kürzeren Wellenlänge (Blau und Grün) vorziehen. Bei Adams (1987) betrachteten drei Monate alte Säuglinge die Farben Rot und Gelb länger als Stimuli in den Farben Blau und Grün und auch Holle war 1988 der Ansicht, dass die beiden erstgenannten Farben von Kindern im ersten Lebensjahr präferiert werden. Als Erklärung gab sie an, dass Kinder den roten Anteil des Farbspektrums früher erkennen können und den blauen und violetten Anteil zuletzt differenzieren. Teller, Civan und Bronson-Castain (2004) konnten in ihrer Studie mit 12-Wochen alten Babies zeigen, dass chromatische Stimuli schon von klein an präferiert werden und dies von der Helligkeit der Farben unabhängig ist. Diese Präferenz ist auch bei dreijährigen Kindern zu finden, die von den chromatischen Farben Rot am meisten und Braun am wenigsten präferieren (Zentner, 2001). Auch Pitchford und Mullen konnten 2005 bei zweijährigen Kindern zeigen, dass diese die Farben Braun und Grau signifikant weniger präferierten als alle anderen Farben (Rot, Gelb, Grün, Blau, Orange, Rosa, Lila, Schwarz und Weiß). Bornstein (1976, zitiert nach Franklin, Bevis, Ling & Hurlbert, 2010, S. 346) zeigte in seiner Studie mit viermonatigen Säuglingen jedoch, dass diese Rot und Blau am längsten fixierten, während sie Grün und Gelb am wenigsten ansahen. Auch Zemach, Chang und Teller (2007) fanden heraus, dass dreimonatige Säuglinge in ihrer Stichprobe die Farben Blau und Lila bzw. Violett am längsten fixierten und Gelb und Grün am wenigsten. Rot wurde in dieser Studie nur mittelmäßig präferiert.

Als Grund für diese widersprüchlichen Studienergebnisse geben Zemach et al. (2007) und Franklin et al. (2010) die subjektiven Farbbezeichnungen in den Studien und die verschiedenen Stimuli an, die unterschiedliche Mengen an chromatischem Kontrast beinhalten. All diese genannten Studien stimmen nur in der Aussage überein, dass chromatische Farben den achromatischen vorgezogen werden und dass die Farben Braun und Grau am wenigsten präferiert werden. Von den vier Grundfarben werden Rot, Gelb und Blau der Farbe Grün vorgezogen.

Laut Li (2007) verändern sich die Farbpräferenzen mit ansteigendem Alter und sind von Kultur zu Kultur unterschiedlich (Taylor, Clifford & Franklin, 2012).

4.3. Farbdifferenzierung

4.3.1. Schwierigkeiten beim Erwerb der Farbbezeichnungen

Zahlwörter werden laut Carey (1978, zitiert nach O'Hanlon & Roberson, 2006, S. 275) oftmals schon nach erstmaligem Hören von den Kindern gelernt, während ihnen das Erlernen der Farben weitaus schwerer fällt (Soja, 1994; Sandhofer & Smith, 2001; Kowalski & Zimiles, 2006; O'Hanlon & Roberson, 2006; Sandhofer & Thom, 2006). Soja (1994) ist der Ansicht, dass es Phasen gibt, in denen es nahezu unmöglich ist, Kindern Farben zu erlernen, so dass früher davon ausgegangen wurde, dass Kinder farbenblind sind. Sobald sie jedoch die erste Farbbezeichnung erworben haben, fällt es ihnen leichter, sich weitere anzueignen (Franklin, 2006).

Pitchford und Mullen (2002) bestätigen in ihrer Studie, dass ein jahrelanges Üben der Farben notwendig ist, um diese bei den Kindern zu festigen. Rice (1980, zitiert nach Sandhofer & Smith, 2001, S. 600) berichtet in ihrer Studie darüber, dass die Kinder im Durchschnitt 1.080 Versuche benötigten, um die drei Farben Rot, Grün, und Gelb zu erlernen und dass explizite Instruktionen notwendig sind. Aus diesem Grund ist erst relativ spät in der Sprachentwicklung davon auszugehen, dass Kinder die Farbbezeichnungen vollständig erwerben (Franklin, 2006; Pitchford, 2006). Zusätzlich sind zu Beginn die erlernten Farben sehr fehleranfällig, da dies das Verständnis für diese voraussetzt (Pitchford, 2006).

Obwohl das Alter beim Erwerb der Farben gesunken ist und das Erlernen der Farbbezeichnungen nicht mehr sehr viel schwieriger zu sein scheint als bei anderen abstrakten Eigenschaften (Franklin, 2006), können dennoch einige Faktoren aufgezeigt werden, die den Erwerb schwieriger gestalten.

Kindern fällt es laut Kowalski und Zimiles (2006) schwer, die Farbe eines Gegenstandes als eigenständige Eigenschaft zu sehen. Die Autoren geben an, dass erst Dreijährige diese Fähigkeit erlangen. O'Hanlon und Roberson (2006) sind der Ansicht, dass die Schwierigkeiten entstehen, wenn Mütter Farben signifikant seltener benennen als andere Wörter bzw. Eigenschaften und somit die Aufmerksamkeit der Kinder nur selten auf die Farbeigenschaften eines Gegenstandes gelenkt werden. Soja (1994) sieht den Grund darin, dass die Farbe eine kategorielle Eigenschaft mit vielen Abstufungen ist und die

Differenzierung somit schwerer fällt, als bei einer Eigenschaft, die nur zwei Abstufungen hat, wie z.B. „hell“ und „dunkel“.

Laut Franklin (2006) bestehen die Schwierigkeiten nicht darin, die Bedeutung einer Farbe oder die eigentliche Bezeichnung einer Farbe zu erlernen, sondern darin, die Wörter den Farben richtig zuzuordnen, sodass manchmal eine Farbbezeichnung für mehrere Farben verwendet wird. Dies lässt wiederum darauf schließen, dass Kinder Schwierigkeiten haben, die unterschiedlichen Farben einzelnen Kategorien zuzuteilen.

4.3.2. Entwicklungsschritte in der Farbdifferenzierung

Kinder beginnen im Alter zwischen zwei und drei Jahren die Bezeichnungen für Farben zu erlernen (Pitchford & Mullen, 2002). Die Studienergebnisse von Pitchford und Mullen (2005) bestätigen diese Annahme und zeigen, dass zweijährige Kinder signifikant weniger Farben richtig benennen können als dreijährige Kinder, und Dreijährige signifikant weniger Farben richtig benennen können als Vierjährige.

Carey (1982, zitiert nach Sandhofer & Smith, 2001, S. 600) ist der Ansicht, dass das Verknüpfen der Frage „Welche Farbe ist das?“ mit einer Farbbezeichnung die erste erworbene Kompetenz im Erwerb des Wissens über Farben ist. Sandhofer und Smith (1999) zeigten in ihrer Studie, dass Kinder mit durchschnittlich 27 Monaten wissen, dass auf diese Frage mit einer Farbbezeichnung geantwortet werden muss, auch wenn die Antwort zum Großteil noch nicht richtig ausfällt. Richtige Antworten werden laut Sandhofer und Smith (1999) erst gegeben, wenn die Kinder fünf oder mehr Farben benennen können.

Die zweite Stufe wird laut Sandhofer und Smith (1999) mit durchschnittlich 28,5 Monaten erlangt. Die Kinder sind in der Lage, Farben passiv zu differenzieren und auf die Frage „Wo ist der grüne Stift?“ den richtigen Stift zu zeigen.

Die dritte Stufe wird mit durchschnittlich 30,5 Monaten erreicht, wenn das Farbkonzept der Kinder stabil ist und sie Objekte verschiedenen Farben nach sortieren können (Sandhofer & Smith, 1999). Das Sortieren von Objekten stellt eine wesentliche Voraussetzung für den Erwerb der Farbbezeichnungen dar. In einer Studie von Rice (1980, zitiert nach Kowalski & Zimiles, 2006, S. 302) konnte gezeigt werden, dass zwei- bis dreijährige Kinder, die nicht in der Lage waren, Objekte den Farben nach zu sortieren,

signifikant mehr Schwierigkeiten hatten Farbbezeichnungen zu erlernen, als Kinder, die die Farbeigenschaften von Objekten abstrahieren konnten.

Holle (1988) ist der Ansicht, dass Kinder mit etwa drei bis vier Jahren alle wichtigen Farben benennen können.

Dabei treten laut Pitchford und Mullen (2005) zwischen dem 36. und 40. Lebensmonat in beliebiger Reihenfolge die Farben Rot, Grün, Gelb, Blau, Schwarz, Weiß, Orange, Rosa und Lila auf. Sechs bis neun Monate später können die Kinder die Farben Braun und Grau richtig benennen. Zusätzlich werden die erst genannten Farben nicht nur signifikant häufiger aktiv richtig wiedergegeben, sondern auch signifikant häufiger passiv wiedererkannt als die Farben Braun und Grau (Pitchford & Mullen, 2005).

Bornstein (1985, zitiert nach Franklin et al., 2008, S. 18224) gibt an, dass das Alter, in dem Kinder Farben benennen können, kulturabhängig ist.

Als wichtiger Einflussfaktor auf den Erwerb der Benennung der Farben bei zweijährigen Kindern hat sich die Häufigkeit der verwendeten Farbbezeichnungen von Seiten der Mütter herausgestellt. Kinder können jene Farben, die in der Interaktion mit ihren Müttern häufiger benannt werden, selbst signifikant häufiger richtig benennen (Pitchford & Mullen, 2005).

Zusätzlich geht Zentner (2001) davon aus, dass die Farbpräferenz das richtige Benennen der Farben beeinflusst. Pitchford, Davis und Scerif (2009) griffen diese Vermutung auf und konnten feststellen, dass die präferierten Farben signifikant häufiger richtig benannt werden können.

Johnson zeigte 1977, dass Mädchen bessere Leistungen beim Benennen der Farben erzielen als Buben und auch Holle (1988) ist der Ansicht, dass Mädchen für Farben sensibler zu sein scheinen und sich in diesem Bereich früher entwickeln.

4.4. Kategorische Farbwahrnehmung

Kategorische Farbwahrnehmung kann laut Bornstein und Korda (1984, zitiert nach Franklin, Wright & Davies, 2009, S. 239) als gegeben angenommen werden, wenn die Unterscheidung zwischen Farben unterschiedlicher Kategorien schneller geschieht, als

die Unterscheidung zwischen Farben derselben Kategorie. Franklin et al. (2008) konnten diese Fähigkeit bei Kindern zwischen zwei und fünf Jahren beobachten und herausfinden, dass sie unabhängig davon ist, wie viele bzw. ob das Kind bereits Farben benennen kann (vgl. Franklin, Clifford, Williamson & Davies, 2005). Auch Ozturk, Shayan, Liszkowski und Majid (2013) bestätigten diese Fähigkeit bei acht Monate alten Kindern, Sandhofer & Smith (2001) bei Zweijährigen und O'Hanlon und Roberson (2007) bei Dreijährigen.

Im Gegensatz dazu schlussfolgerten Gilbert, Regier, Kay und Ivry (2006) und Roberson, Davies und Davidoff (2000) aus ihren Studienergebnissen, dass die kategorielle Farbwahrnehmung von der Sprache bzw. den bereits erworbenen Farbbezeichnungen abhängig ist.

Franklin et al. (2009) analysierten viele Studien und kamen zu dem Schluss, dass sich kein Konsensus bezüglich des Einflusses der Sprache auf die kategorielle Farbwahrnehmung finden lässt. Ein Ansatz für weitere Überlegungen hierfür wäre, dass Kleinkinder, wie schon in Kap. 4.3.2. Entwicklungsschritte in der Farbdifferenzierung beschrieben, vor dem aktiven Benennen von Farben sprachliche Assoziationen bilden und somit nicht davon ausgegangen werden kann, dass die kategorielle Farbwahrnehmung von der Sprache völlig unabhängig ist (Davidoff, Goldstein & Roberson, 2009). Wovon jedoch mit Sicherheit ausgegangen werden kann, ist, dass sie vor der Entwicklung der Sprache in der rechten Hemisphäre lokalisiert ist, dies jedoch mit Erlernen der Sprache zur linken Hemisphäre wechselt, wo sie auch bei Erwachsenen zu finden ist (Franklin et al., 2008; Roberson, Pak & Hanley, 2008).

Stabilität des Farbkonzepts

Der nächste Entwicklungsschritt in der kategoriellen Farbwahrnehmung, der ebenso vor dem vollständigen Erwerb der Farbbezeichnungen eintritt, ist das Abstrahieren der Farbe von verschiedenen Objekten, um gleichfarbige Gegenstände zusammenzufinden (Sandhofer & Smith, 1999). Laut Soja (1994) fällt dies den Kindern schwer, wenn die zu sortierenden Objekte sehr verschieden sind, wie z.B. ein rotes Glas und ein roter Ball. Leichter fällt es, wenn die Objekte ident oder sehr ähnlich sind. Soja (1992) fand heraus, dass zweijährige Mädchen bessere Leistungen im schlussfolgernden Denken bezüglich der Eigenschaften (wie z.B. Farben) von Objekten zeigen, als gleichaltrige Buben.

Das Sortieren verschiedener Objekte nach Farben können Kinder mit ungefähr drei Jahren (Pitchford & Mullen, 2001; Kowalski & Zimiles, 2006). In der Studie von

Sandhofer und Smith (1999) war dies schon mit durchschnittlich 30,5 Monaten möglich und Linder (1993, zitiert nach Dykeman, 2008, S. 407) geht davon aus, dass diese Fähigkeit zwischen dem 33. und 36. Lebensmonat ausgebildet ist. Daraus lässt sich schlussfolgern, dass Kinder ab dem dritten Lebensjahr das Sortieren von Objekten nach Farben erlernen und dies noch vor dem vollständig richtigen Benennen der Farben beherrschen (Sandhofer & Smith, 1999).

4.5. Formwahrnehmung

Slater, Morison und Rose (1983, zitiert nach Goswami, 2001, S. 41) kommen zu dem Schluss, dass Neugeborene bereits verschiedene Formen wahrnehmen können und zwischen einem Kreuz und einem Kreis differenzieren.

Da die Kinder ab dem neunten Lebensmonat hauptsächlich mit den Augen und Fingern Figuren wahrnehmen, können sie ab diesem Alter laut Holle (1988) einzelne Gegenstände wiedererkennen. Mit einem Jahr ist es ihnen möglich, den gleichen Gegenstand aus verschiedenen Blickwinkeln und Abständen und vor unterschiedlichem Hintergrund zu identifizieren (Holle, 1988). Zwischen dem 15. und 18. Lebensmonat sind Kinder in der Lage, zwischen einem Kreis und einem Quadrat zu unterscheiden (Linder, 1993; zitiert nach Dykeman, 2008, S. 407). Einfache geometrische Figuren in entsprechende Lücken platzieren, können sie mit ungefähr zwei Jahren, und mit zweieinhalb Jahren können die Kinder große Blauklötze zu Paaren sortieren (Holle, 1988).

Beim Erwerb der Sprache bei 18 bis 24 Monate alten Kinder spielt der „shape bias“ laut Yee, Jones und Smith (2012) eine entscheidende Rolle. Dieser besagt, dass Kinder Gegenstände und Objekte aufgrund ihrer Form kategorisieren und so neue Wörter erwerben. Die fokussierte Aufmerksamkeit auf die Form von Objekten ist also ein entscheidender Prädiktor für den schnellen Wortschatzanstieg in diesem Alter (Yee et al., 2012). Bei sprachverzögerten Kindern konnte der „shape bias“ nicht gefunden werden (Jones, 2003).

Sollen Gegenstände nach einer beliebigen Eigenschaft (Farbe, Form, Größe), ohne Einfluss der Sprache, sortiert werden, so tendieren Kleinkinder dazu, Objekte der Farbe nach zusammenzufinden. Mit zunehmendem Alter sortieren Kinder die Gegenstände

immer mehr nach der Form (Prevor & Diamond, 2005; La Heij & Boelens, 2010). Prevor und Diamond (2005) kommen zu dem Schluss, dass bei Kindern zwischen drei und dreieinhalb Jahren ein Wechsel des Interesses von der Farbe zur Form stattfindet. Auch Pitchford und Mullen (2001) sind der Ansicht, dass sich die Kinder, nach abgeschlossenem Erwerb der Farbbezeichnungen mit etwa vier Jahren, mehr für die Formeigenschaften eines Gegenstandes als für dessen Farbe interessieren.

4.6. Größenwahrnehmung

Während Sandhofer und Smith (1999) der Ansicht sind, dass nicht nur der Erwerb der Farbbezeichnungen langsamer als bei anderen Wörtern verläuft, sondern auch der Erwerb der Größenbezeichnungen, konnten Pitchford und Mullen (2001) keinen Unterschied in der Schwierigkeit des Benennens der Wörter finden.

Ein Unterschied existiert jedoch zwischen dem Erwerb der Farb- und der Größenbezeichnungen, der einerseits in der Anzahl der unterschiedlichen Kategorien und andererseits in der unterschiedlichen Stufenabfolge (Sandhofer & Smith, 1999, 2001) besteht.

Kinder erwerben mit ca. 26,5 Monaten die Fähigkeit, Objekte der Größe nach zu sortieren, mit ca. 27 Monaten können sie das größere Auto zeigen, wenn die Frage lautet „Zeige mir das große Auto!“ und mit ca. 28 Monaten ist es den Kindern möglich, die Frage nach der Größe eines Gegenstandes mit einer Größenbezeichnung zu beantworten.

Carey (1982, zitiert nach Sandhofer & Smith, 2001, S. 601) fand heraus, dass zweijährige Kinder die Größenbezeichnungen „groß“ und „klein“ schon sehr gut beherrschen. Trotzdem können Schwierigkeiten im Verständnis der Größenbezeichnungen auftreten, da ein Auto, obwohl es groß ist, im Vergleich zu anderen Autos klein sein kann (Sera & Smith, 1987; zitiert nach Sandhofer & Smith, 2001, S. 601).

4.7. Diagnostik der visuellen Wahrnehmung

In der Literatur der 1970-er Jahre galten visuelle Wahrnehmungsstörungen als wesentliche Vorläuferfunktionen für das Lesen und Rechtschreiben, heute werden sie jedoch zunehmend für Störungen des rechnerischen Denkens diskutiert (Dacheneder, 2009). Da deutlich erkennbare Schwächen in diesen Bereichen von großem Nachteil für

einen erfolgreichen Bildungsweg, das anschließende Berufsleben und vor allem auch für die persönliche Befindlichkeit sind (Rindermann & Kwiatkowski, 2010), ist eine ausführliche und sorgfältige Diagnostik unumgänglich.

Im Säuglings- und Kleinkindalter werden im Zuge der Mutter-Kind-Pass Untersuchungen Augenuntersuchungen vorgeschrieben. Diese können zwischen dem 10. und 14. Lebensmonat bei einem Kinderarzt oder einer Kinderärztin durchgeführt werden. Zwischen dem 22. und 26. Lebensmonat muss ein Facharzt oder eine Fachärztin für Augenheilkunde aufgesucht werden.

Da für solch junge Kinder keine Wahrnehmungsuntersuchungen im engeren Sinne stattfinden, werden im deutschen Sprachraum die Lea-Tests von Lea Hyvärinen (2004) angewandt, die häufig von Orthoptisten und Orthoptistinnen oder Mitarbeitern und Mitarbeiterinnen der Frühförderung für Sehbehinderte und Blinde eingesetzt werden (Dacheneder, 2009).

Manche Aspekte der visuellen Wahrnehmung bei Zweijährigen können in verschiedenen Intelligenztests erfasst werden (Dacheneder, 2009):

- Mit dem *Hannover-Wechsler-Intelligenztest für Kinder im Vorschulalter – III* (HAWIVA-III; Ricken et al., 2007) können im Alter zwischen 2;6 bis 3;11 Jahren mittels der Untertests „Mosaik-Test“ und „Figuren legen“ die visuell-konstruktiven Leistungen abgebildet werden.
- Die Subtests „Mosaik“, „Puzzles“ und „Zeichenmuster“ des „*Nonverbalen Intelligenztests für Kinder*“ (SON-R 2½-7; Tellegen et al., 2007) erfassen visuell-räumliche und visuokonstruktive Wahrnehmungsleistungen.
- In der „*Kaufman-Assessment-Battery for Children*“ (K-ABC; Melchers & Preuß, 2007) bildet der Subtest „Gestaltschließen“ die Fähigkeit zum visuellen Gestaltschließen ab.

Verdacht bezüglich einer visuellen Wahrnehmungsstörung ist laut Dacheneder (2009) angebracht, wenn im Spielverhalten beobachtet werden kann, dass ein Kind bestimmte Spiele, wie z.B. das konstruktive Bauen oder das Puzzle meidet, ablehnt oder nicht beherrscht, es selten gegenständlich malt, beim Ausmalen die Begrenzungen nicht einhält und es in seiner Malentwicklung nicht mit Gleichaltrigen mithalten kann.

Zusammenfassung

Wir Menschen können in etwa 200 verschiedene Farben wahrnehmen, wobei sich chromatische (wie z.B. Blau, Grün, Rot und Gelb) von achromatischen Farben (Schwarz, Weiß und Grau) unterscheiden lassen. Als die vier Grundfarben werden Rot, Gelb, Grün und Blau bezeichnet.

Bezüglich der Präferenz für Farben konnten unterschiedliche Ergebnisse gefunden werden, wobei mit Sicherheit davon ausgegangen werden kann, dass Kinder chromatische Farben den achromatischen vorziehen. Außerdem präferieren Kinder die Farben Braun und Grau am wenigsten und von den vier Grundfarben werden Rot, Gelb und Blau der Farbe Grün vorgezogen. Die Farbpräferenzen verändern sich mit ansteigendem Alter und sind kulturabhängig.

Im Alter zwischen zwei und drei Jahren beginnen Kinder die Bezeichnungen für Farben zu erlernen. Mit durchschnittlich 27 Monaten erwerben die Kinder das Wissen, dass auf die Frage „Welche Farbe ist das?“ eine Farbbezeichnung folgen muss. Kurze Zeit später, mit ca. 28,5 Monaten, können die Kinder die Farben passiv differenzieren und mit etwa drei bis vier Jahren alle wichtigen Farben aktiv benennen. Zwischen dem 36. und dem 40. Lebensmonat erwerben die Kinder in beliebiger Reihenfolge die Bezeichnungen für die Farben Rot, Grün, Gelb, Blau, Schwarz, Weiß, Orange, Rosa und Lila. Erst sechs bis neun Monate später können die Kinder die Farben Braun und Grau benennen.

Ab dem dritten Lebensjahr erlernen die Kinder das Sortieren von Gegenständen nach Farben, das sich als wichtige Vorläuferfähigkeit für das Benennen der Farben herausgestellt hat.

Die fokussierte Aufmerksamkeit auf die Form von Objekten ist ein entscheidender Prädiktor für den schnellen Wortschatzanstieg im Alter zwischen 18 und 24 Monaten und steigt nach abgeschlossenem Erwerb der Farbbezeichnungen mit etwa vier Jahren.

Objekte ihrer Größe nach zu sortieren gelingt Kindern mit ca. 26,5 Monaten, mit durchschnittlich 27 Monaten können sie den größeren vom kleineren Gegenstand unterscheiden und mit ca. 28 Monaten ist das aktive Benennen einer Größenbezeichnung möglich.

Die Diagnostik von visuellen Wahrnehmungsstörungen passiert hauptsächlich im Zuge der Mutter-Kind-Pass Untersuchungen zum ersten und zweiten Geburtstag. Zusätzlich kann dieser Fähigkeitsbereich in verschiedenen Intelligenztests, wie z.B. dem HAWIVA-III, dem SON-R 2½-7 und der K-ABC, erfasst werden.

5. Wie untersucht man Kinder?

Wie auch immer die Untersuchung bei Kindern gestaltet wird, ob mit standardisierten Tests oder spielbasierten Verfahren, auf gewisse Rahmenbedingungen sollte geachtet werden. Diese werden nachfolgend näher beschrieben.

Kinder, vor allem kleine Kinder, zu testen ist nicht immer einfach und erfolgt oft nicht völlig vorschriftsgemäß. Denn gerade sehr kleine Kinder halten sich nicht unbedingt an die formalen Vorgaben einer Testdurchführung. Sie kennen noch kaum soziale Normen und Regeln und neigen deshalb zu eigenwilligem Handeln. (Quaiser-Pohl, 2010, S. 57)

Aus diesen Gründen sind laut Esser und Petermann (2010) ein hohes Maß an Erfahrung im Umgang mit Kindern in der jeweiligen Altersgruppe sowie die Beherrschung des Untersuchungsmaterials und des –ablaufs unabdingbare Voraussetzungen für eine objektive und kindgerechte Untersuchung.

Weitere günstige Faktoren für die psychologische Untersuchung von Kindern werden im Folgenden beschrieben:

5.1. Räumliche Bedingungen

Der Raum, in der eine psychologische Untersuchung stattfindet, löst laut Preston (2005) und Quaiser-Pohl (2010) emotionale Reaktionen bei einem Kind aus. Deshalb ist es umso wichtiger, diesen Raum angenehm und freundlich zu gestalten, um es dem Kind, der begleitenden Bezugsperson und auch dem Untersucher oder der Untersucherin zu ermöglichen, sich wohl zu fühlen (Esser & Petermann, 2010; Irblich & Renner, 2009).

Empfohlen wird von Esser und Petermann (2010) ein heller, warmer Raum und auch Irblich und Renner (2009) weisen auf eine den Bedürfnissen des Kindes entsprechende Raumtemperatur sowie blendfreie Beleuchtung hin. Der Raum sollte ausgestattet sein mit kindgerechten Möbeln, wie z.B. höhenverstellbaren Stühlen und Tischen (Preston, 2005; Esser & Petermann, 2010; Irblich & Renner, 2009) sowie mit Teppichboden, Matten und Decken (Esser & Petermann, 2010; Irblich & Renner, 2009).

Im Weiteren sollen störende Außengeräusche vermieden werden, um dem Kind möglichst wenig Ablenkung zu bieten (Esser & Petermann, 2010; Irblich & Renner, 2009).

Irblich und Renner (2009) empfehlen, das Material vor Beginn der Untersuchung für die Untersuchungsperson übersichtlich, jedoch für das Kind außerhalb der Reichweite zu platzieren. Nicht benötigtes Untersuchungsmaterial soll im Koffer gelassen werden, um dem Kind keinen Anlass für Ablenkung zu bieten (Esser & Petermann, 2010). Preston (2005), Quaiser-Pohl (2010) und Irblich und Renner (2009) raten zu einem überschaubaren Angebot altersentsprechender Spielmöglichkeiten und zu keiner leicht zu erreichenden Wanddekoration, wie z.B. Postern. Da Kinder diese gerne explorieren, dafür jedoch keine Zeit bleibt, kann dies zu einer Frustration des Kindes führen. Die Folgen davon wären laut Preston (2005) und Quaiser-Pohl (2010), dass sich die Verarbeitungskapazität des Kindes reduziert, dies die Genauigkeit der Ergebnisse gefährdet und zu einer Verlängerung der Untersuchung führen kann.

Auch eine Überfüllung des Testraumes mit fremden, körperlich überragenden Erwachsenen soll vermieden werden, da dies zu einer Verunsicherung des Kindes führen kann. Kinder flüchten auf Grund dessen zu ihren Eltern, akzeptieren nur einfachste Aufgaben, geben zögerliche Antworten oder verweigern die weitere Mitarbeit (Preston, 2005; Quaiser-Pohl, 2010).

Um auf unangemessenes Verhalten eines Kindes schnell reagieren zu können, empfehlen Irblich und Renner (2009) sowie Esser und Petermann (2010) die Sitzposition von Kind und Untersucher entweder vis-à-vis oder über das Eck.

5.2. Anforderungen an den Untersucher

Esser und Petermann (2010) sind sich einig, dass eine entwicklungsdiagnostische Untersuchung im Säuglings- und Kleinkindalter deutlich höhere Anforderungen an die Untersucher stellt als die Leistungsdiagnostik im Schul- und Erwachsenenalter.

Die Arbeitshaltungen bei kleinen Kindern sind noch nicht ausgebildet, „was zur Folge hat, dass Motivation, Konzentration und Ausdauer sehr stark durch Außenreize gesteuert werden“ (Kastner-Koller & Deimann, 2012, S. 57). Deshalb ist ein hohes Maß an

Flexibilität und Erfahrung mit der jeweiligen Altersgruppe sowie in der Durchführung eine wichtige Voraussetzung (Kastner-Koller & Deimann, 2012). Denn bei der Untersuchung von Kindern gilt ganz besonders, dass die Geübtheit der Untersucher die Güte eines Ergebnisses maßgeblich beeinflusst (Quaiser-Pohl, 2010).

Laut Esser und Ballaschk (2008, zitiert nach Esser & Petermann, 2010, S. 29) ist es die Aufgabe der Untersuchungsperson, die Kinder bestmöglich auf die Untersuchungen vorzubereiten, die Reihenfolge der Items dem Gefühlszustand der Kinder anzupassen und damit auch schwierige Untersuchungssituationen mit schwankender Kooperationsbereitschaft der Kinder erfolgreich zu meistern. Das Bemühen um Standardisierung der Untersuchungssituation soll nicht dazu führen, dass das flexible Handeln der Untersuchungsperson eingeschränkt ist (Quaiser-Pohl, 2010).

Da Kinder sehr sensibel auf eine gleichgültige oder ablehnende Haltung des Untersuchers oder der Untersucherin reagieren, sollte die Arbeit mit Kindern Freude bereiten (Esser & Petermann, 2010; Irblich & Renner, 2009) und dies auch in einer kindgemäßen Ausdrucksweise zum Vorschein kommen. Eine freundliche Stimme signalisiert dem Kind Wertschätzung und Interesse (Irlich & Renner, 2009), zusätzlich sollte Kritik vermieden werden und ein angemessenes Blickkontaktangebot berücksichtigt werden (Kastner-Koller & Deimann, 2012).

Im Weiteren benötigt der Untersucher eine breite Palette an Möglichkeiten, um mit Kindern in Kontakt zu kommen und mit ihnen zu interagieren (Esser & Petermann, 2010; Irlich & Renner, 2009). Laut Bracken (2007, zitiert nach Irlich & Renner, 2009, S. 23) ist ein Untersucherverhalten dann optimal, wenn es die Balance zwischen freundlicher Aufgeschlossenheit und aufgabenbezogener Sachlichkeit hält.

5.3. Vorbereitung der Untersuchung

Preston (2005) erwähnt, dass es wichtig ist, Fakten, wie z.B. das Alter des Kindes oder den Grund für die Untersuchung, schon bei der Terminvereinbarung von der Bezugsperson zu erfahren, um die geeigneten Verfahren vorbereiten, bzw. damit einhergehende Schwierigkeiten zu antizipieren und sich somit bestmöglich darauf vorbereiten zu können.

Kinder, vor allem sehr kleine, sind in ihrem Verhalten wenig vorhersehbar und in ihrer Kooperationsbereitschaft sprunghaft (Irblich & Renner, 2009). Umso wichtiger ist es, sich als Untersucher oder Untersucherin vor einer Testung mit dem Material besonders vertraut zu machen und sich einzuarbeiten, um auf alle Eventualitäten vorbereitet zu sein und flexibel reagieren zu können.

Die Untersuchungsperson muss mit den Testinstruktionen gut vertraut sein, diese sollten auf keinen Fall vorgelesen, aber doch wortgetreu wiedergegeben werden (Kastner-Koller & Deimann, 2012). Wichtig ist es außerdem, die Aufgabenmaterialien nach einer vorangegangenen Testung wieder in die richtige Reihenfolge und an den richtigen Platz zu bringen (Preston, 2005), da Unterbrechungen dieser Art den Fluss der Untersuchung stören und dadurch möglicherweise das Interesse des Kindes verloren geht (Quaiser-Pohl, 2010).

Reuner et al. (2008) sehen außerdem die Kenntnis alterstypischer Verhaltensweisen als großen Vorteil, Preston (2005) die Kenntnisse über den Ablauf der psychischen Entwicklung und relevanten Entwicklungsprozesse.

5.4. Durchführung der Untersuchung

5.4.1. Untersuchungszeit

Als günstige Tageszeit für die Untersuchung von Kindern erweist sich laut Esser und Petermann (2010) der Vormittag, da die Kinder üblicherweise zu dieser Zeit besonders leistungsfähig und interessiert sind. Bei der Terminvereinbarung sollten die Wach-Schlaf-Rhythmen der Kinder berücksichtigt werden (Irblich & Renner, 2009) und laut Quaiser-Pohl (2010) Zeitpunkte vermieden werden, zu denen die Kinder gewöhnlich schlafen.

5.4.2. Untersuchungsdauer

Quaiser-Pohl (2010) ist der Ansicht, dass die Untersuchungsdauer von Kindern eine Zeitspanne von 30 bis 45 Minuten nicht überschreiten sollte, bei Esser und Petermann (2010) ist eine Dauer bis zu 50 Minuten im annehmbaren Bereich. Die Eltern sollten im Vorhinein über die Dauer und den Ablauf der Untersuchung in Kenntnis gesetzt werden (Irblich & Renner, 2009).

5.4.3. Anwesenheit der Bezugsperson(en) während der Untersuchung

Abhängig vom Alter des Kindes ist die Bezugsperson während der Untersuchung anwesend oder nicht. Auf jeden Fall ist dies im Vorhinein mit dieser abzuklären (Irblich & Renner, 2009). Empfehlenswert ist die Anwesenheit von zumindest einem Elternteil bei Kindern bis zu vier Jahren (Irblich & Renner, 2009). Andere Autoren, wie z.B. Kamphaus (2005, zitiert nach Irblich & Renner, 2009, S. 27), empfehlen jedoch schon ab einem Alter von drei Jahren, die Kinder in Abwesenheit der Bezugsperson zu untersuchen. Ziel ist es laut Irblich und Renner (2009), möglichst optimale Untersuchungsbedingungen herzustellen, deshalb ist dies im Einzelfall abzuwägen und zu entscheiden.

Oftmals können Eltern für ihre Kinder in der Untersuchungssituation eine Ablenkung darstellen, die es ihnen nicht ermöglicht, sich auf das Aufgabenmaterial zu konzentrieren. In diesen Fällen ist es besser, die Kinder in Abwesenheit der Eltern zu untersuchen. Ist dies der Fall, muss dem Kind mitgeteilt werden, wo sich diese während der Untersuchung aufhalten, und wann es sie wiedersieht (Irblich & Renner, 2009).

Andere Kinder können sich wiederum nur schwer von ihren Eltern lösen und haben Trennungsängste. Hier wäre es laut Irblich und Renner (2009) kontraproduktiv und keineswegs sinnvoll, das Kind von der Mutter bzw. dem Vater zu trennen. Somit kann es sich auch als durchaus günstig erweisen, Eltern bei der Untersuchung zuschauen zu lassen. Laut Ford und Dahinten (2005, zitiert nach Irblich & Renner, 2009, S. 27) ist es in diesem Fall jedoch notwendig, die Eltern darauf hinzuweisen, dass die Kinder nicht alle Aufgaben lösen können und dass dies auch nicht erforderlich ist, um ein gutes Ergebnis zu erzielen. Außerdem sollen sie sich während der Untersuchungssituation zurückhalten und möglichst im Hintergrund, ohne direkten Blickkontakt zum Kind, sitzen (Quaiser-Pohl, 2010).

5.4.4. Untersuchung

Petermann et al. (2004) empfehlen bei der Kontaktaufnahme mit Säuglingen und Kleinkindern, das Kind nur kurz auf gleicher Augenhöhe zu begrüßen und sich vorzustellen, sich dann zunächst jedoch der begleitenden Bezugsperson zuzuwenden. So wird dem Kind die Möglichkeit gegeben, sich als scheinbar passiver Beobachter in Ruhe

ein Bild von der fremden Person zu machen. Auch sollte es den Kindern möglich sein, den Untersuchungsraum frei zu erkunden, wobei ihnen eventuell beiläufig etwas vom Material angeboten werden kann.

Da viele Kinder, vor allem im Kleinkindalter, anfänglich scheu und zögerlich gegenüber fremden Personen sind, ist es hilfreich, zum Einstieg ein interessantes Spielzeug anzubieten, das nicht zum eigentlichen Untersuchungsmaterial gehört (Reuner et. al., 2008). Mit Beginn der Untersuchung, sollte dies jedoch zur Seite gelegt werden. Dem Kind sollte in jedem Fall ausreichend Zeit zur Verfügung stehen, damit es mit der Untersuchungsperson vertraut werden kann (Reuner et. al., 2008).

Lässt sich die Bereitschaft zur Kooperation von Seiten des Kindes erkennen, indem es sich z.B. von seiner Bezugsperson löst, im Raum spielt und auf (verbale) Kontaktaufnahme des Untersuchers oder der Untersucherin positiv reagiert, kann laut Irblich und Renner (2009) mit der eigentlichen Untersuchung begonnen werden.

Die primären Bedürfnisse des Kindes (Hunger, Durst, Schlaf, trockene Kleidung) sollten bei Beginn und auch während der Untersuchung befriedigt sein. Zeigt das Kind Anzeichen von Müdigkeit, sollte abgebrochen und ein neuer Untersuchungszeitpunkt vereinbart werden (Kastner-Koller & Deimann, 2012).

Als Einstieg in die Untersuchungssituation eignen sich laut Esser und Petermann (2010) einfache Aufgaben mit einem hohen Aufforderungscharakter. Schwierige Aufgaben, die sprachliche Reaktionen des Kindes erfordern, sollten erst im weiteren Verlauf der Untersuchung eingesetzt werden.

Manche Kinder legen einen langsamen Arbeitsstil an den Tag. In dieser Situation ist es Preston (2005) zufolge wichtig, den Kindern ausreichend Zeit zu geben, sie nicht zu stressen und in der Rolle des Untersuchers oder der Untersucherin ruhig und geduldig zu bleiben. Andere Kinder wiederum benötigen eine straffere Testdurchführung. Hier ist es die Aufgabe der Untersuchungsperson, die Vorgabe der Aufgaben flexibel an das Kind abzustimmen.

Beobachtet der Untersucher oder die Untersucherin motorische Unruhe, ablenkende Fragen, abschweifende Blicke und zunehmend längere Reaktionszeiten, kann dies auf

eine zunehmende Belastung oder Ermüdung des Kindes hinweisen. Ist dies der Fall, so sind rechtzeitig Pausen anzubieten (Irblich & Renner, 2009).

Für viele Kinder ist es wichtig und auch effektiv, häufig Bestätigung zu erfahren, indem sie für ihre Mühe bzw. ihre Leistung gelobt werden. Dies ist auch ein wichtiger Faktor für diese Kinder, um weiterhin kooperativ zu bleiben (Irblich & Renner, 2009).

Die Motivation, die Kooperationsbereitschaft und das Interesse des Kindes müssen während der gesamten Untersuchung aufrecht erhalten werden. Irblich und Renner (2009, S. 28) führen folgende unterstützende Faktoren an:

- freundliche Zugewandtheit bei gleichzeitiger Aufgabenorientierung
- klare Strukturierung der Situation und Verhaltenssteuerung durch den Untersucher bzw. die Untersucherin
- Erkennen der Bedürfnisse des Kindes und deren Berücksichtigung
- Abwechslung der Aktivitäten während der Testung
- Bestätigung und Lob, aber auch Aufforderungen bei geringer Anstrengungsbereitschaft
- Trost und Ermutigung bei Misserfolgserlebnissen
- auflockernde Bemerkungen oder eine kleine Unterhaltung über nicht testbezogene Themen bei ängstlich angespannten Kindern

5.4.5. Verhaltensbeobachtung während der Untersuchung

Die Verhaltensbeobachtung des Kindes während der Untersuchung spielt eine wesentliche Rolle. So können laut Irblich & Renner (2009) nicht nur körperliche Merkmale, äußeres Erscheinungsbild, hygienischer Zustand, Sprechweise und motorische Geschicklichkeit beobachtet werden, sondern auch diagnostische Hinweise zu den Testwerten eingeholt werden. Weitere Fähigkeiten bzw. Eigenschaften, die beobachtet werden können, sind z.B. die Aufmerksamkeit, Ausdauer, Frustrationstoleranz, das Interesse, und die Lenkbarkeit bei fremdgestellten Aufgaben. Testergebnisse sollten nach Ansicht von Irblich und Renner (2009, S. 30) nie interpretiert werden, „ohne die Motivation des Kindes, seine verbalen Äußerungen und sein Verhalten während der Aufgabenbearbeitung zu berücksichtigen.“

5.4.6. Abschluss der Untersuchung

Am Ende einer Untersuchung empfehlen Irblich und Renner (2009) sich beim Kind für seine Mitarbeit zu bedanken und danach zu fragen, wie es die Untersuchung empfand. Zusätzlich können bei aufgetretenen Problemen diese in kindgemäßer Form angesprochen werden, um einen versöhnlichen Abschluss mit dem Kind zu finden.

Zusammenfassung

Um eine entwicklungsdiagnostische Untersuchung objektiv und kindgerecht zu gestalten, müssen viele begünstigende Faktoren berücksichtigt werden.

Der Raum, in der die Untersuchung stattfindet, sollte hell, warm und mit kindgerechten Möbeln ausgestattet sein. Es soll darauf geachtet werden, mögliche Außengeräusche zu vermeiden, bzw. ablenkende Spielsachen oder Wanddekorationen außerhalb der Reichweite des Kindes zu platzieren. Von der untersuchenden Person ist genügend Erfahrung im Umgang mit der jeweiligen Altersgruppe bzw. in der Durchführung der psychologischen Testverfahren zu erwarten. Die Arbeit mit Kindern sollte Freude bereiten. Auf eine kindgemäße Ausdrucksweise, freundliche Stimme und ein angemessenes Blickkontaktangebot sollte geachtet werden. Die Vorbereitung auf die Testsituation bzw. die Einarbeitung in bestimmte Verfahren sollte als äußerst wichtig erachtet werden, um unnötige Unterbrechungen in der Untersuchungssituation zu vermeiden. Als günstige Tageszeit für die Untersuchung von Kindern erweist sich der Vormittag, wobei die Untersuchungsdauer 50 Minuten nicht überschreiten sollte.

Ob die Bezugsperson während der Untersuchung anwesend ist oder nicht, ist abhängig vom Alter und der Persönlichkeit des Kindes.

Zu Beginn der Untersuchung ist es wichtig, dem Kind genügend Zeit zu geben, um sich mit der Untersuchungsperson vertraut zu machen, es soll mit einfachen Aufgaben, die keine Sprache verlangen, begonnen werden. Eine flexible Anpassung der untersuchenden Person an das Kind und auch Lob für seine Leistungen sind wichtig, um die Kooperation des Kindes nicht zu gefährden. Die Verhaltensbeobachtung während der Untersuchung wird als äußerst sinnvoll erachtet, um die Testergebnisse mit Hinweisen aus der Beobachtung in Bezug setzen zu können.

Am Ende der Untersuchung ist es wichtig, sich beim Kind für seine Mitarbeit zu bedanken.

EMPIRISCHER TEIL

6. Ausgangspunkt und Ziel dieser Arbeit

Im Rahmen zweier Diplomarbeiten wurde im Jahr 2011 von Kuchler und Sapper mit der Unterstützung von Ass.-Prof. Dr. Kastner-Koller und Ass.-Prof. Dr. Deimann ein Itempool zur spielbasierten Erfassung der Entwicklung Zweijähriger erstellt. Dieser umfasst die Fähigkeitsbereiche der motorischen, sprachlichen, sozioemotionalen und kognitiven Entwicklung, der visuellen Wahrnehmung, des Gedächtnisses und der Aufmerksamkeit.

Birngruber (2012), Fuchs-Gaderer (2012), Putzer (2013) und Sindelar (2013) überarbeiteten und entwickelten diesen Itempool weiter und erprobten ihre Veränderungen jeweils an einer Stichprobe von Kindern im dritten Lebensjahr.

Das Ziel dieser Diplomarbeit ist es, aufgrund der statistischen Ergebnisse und Kritikpunkte der soeben genannten Vorgängerdiploarbeiten, weitere Modifikationen an den Items vorzunehmen, den Aufgabenpool sowohl im Inhalt als auch in der Vorgabe dem Alter der zweijährigen Kinder noch besser anzupassen und die Ergebnisse am Ende testtheoretisch zu überprüfen.

Mit Hilfe der statistischen Auswertungen sollen die nachfolgend beschriebenen Fragestellungen beantwortet werden.

7. Fragestellungen

Die erste Fragestellung beschäftigt sich damit, welche Items geeignet sind, den Entwicklungsstand der Kinder in den jeweiligen Bereichen adäquat abzubilden und welche nicht.

Hierfür werden die Trennschärfen und Schwierigkeitsindizes der einzelnen Items berechnet, um zu entscheiden, welche Aufgaben genügend Aussagekraft haben, um im Itempool zu bleiben, und welche ausgeschlossen bzw. überarbeitet werden müssen.

Im Zuge der zweiten Fragestellung wird der Frage nach den Hauptgütekriterien nachgegangen. Es wird untersucht, ob der vorliegende Itempool die Qualitätsanforderungen in Bezug auf die Reliabilität, Objektivität und Validität erfüllt.

Um zu überprüfen, ob die Items *reliabel* genug messen, wird die innere Konsistenz mittels Cronbach-Alpha (α) berechnet, das laut Field (2009, S. 675) ab einem Wert von .70 zufriedenstellend und ab .90 laut Weise (1975, S. 219, zitiert nach Bortz & Döring, 2009, S. 199) hoch ausfällt.

Für die *Objektivität* können keine statistischen Kennwerte berechnet werden. Die Beurteilung dieses Kriteriums wird aufgrund der Beobachtungen aus den Untersuchungen vorgenommen.

Für den Nachweis des Gütekriteriums der *Validität* werden Korrelationen der Testleistung mit dem Alter der Kinder berechnet, da laut Macha, Proske und Petermann (2005) und Hagmann-von Arx et al. (2008) alle Entwicklungstests von stetig zunehmenden Testleistungen mit dem Alter ausgehen (s. Kap. 2.3.4. Validität in der Entwicklungsdiagnostik). Dieser Zugang dient dem Beleg der Konstruktvalidität.

Eine weitere Fragestellung bezieht sich darauf, ob sich signifikante Geschlechts- und Altersunterschiede in den Testergebnissen feststellen lassen. Es wird davon ausgegangen, dass sich die Mädchen nicht signifikant von den Buben in ihren Leistungen unterscheiden (zweiseitiges Testen) und dass ältere Kinder bessere Testergebnisse erbringen als Jüngere (einseitiges Testen). Diese Analysen dienen zum weiteren Nachweis der Konstruktvalidität.

Zusätzlich interessiert, inwieweit die Aussagen der Bezugspersonen im Elternfragebogen über den Fähigkeitsbereich der Selbstständigkeit der Kinder mit den Beobachtungen und Ergebnissen der Testung übereinstimmen. Die Testergebnisse der Kinder werden also mit einem Außenkriterium, nämlich der Einschätzungen der Bezugspersonen im Fragebogen, korreliert, um zu belegen, ob der Skala „Selbstständigkeit“ Kriteriumsvalidität zuzusprechen ist.

8. Methodik

Nach der Darstellung der Fragestellungen werden nun das Setting, der Ablauf der Untersuchung, das Untersuchungsinstrument und die Stichprobe näher beschrieben.

Um detaillierte Informationen zur Vorgabe und Auswertung der einzelnen Items zu erhalten, wird auf das Manual zum Itempool zur spielbasierten Erfassung der Entwicklung Zweijähriger von Kuchler, Sapper, Deimann und Kastner-Koller (2011), überarbeitet von Punz und Kronberger (2013), verwiesen.

8.1. Setting

Das Ziel der Erstellung dieses Itempools zur Erfassung des Entwicklungsstandes zweijähriger Kinder war, möglichst spielerisch die wichtigsten Fähigkeitsbereiche der motorischen, kognitiven, sprachlichen, sozioemotionalen Entwicklung, der visuellen Wahrnehmung, des Gedächtnisses und der Aufmerksamkeit zu erfassen.

Die insgesamt 20 Testungen fanden vom 05.02. bis zum 07.03.2013 im Kleinkindertestraum der Test- und Beratungsstelle der Fakultät für Psychologie statt.

Dieser Raum war ausgestattet mit einer Puppenküche, die das Hauptsetting dieses Verfahrens ist, da sie einen hohen Aufforderungscharakter für die Kinder besitzt. Zusätzlich waren ein Kaufmannsladen, eine Stiege und eine Spielebox vorhanden, in der sich jene Materialien befanden, die sich in das Spiel mit der Puppenküche nicht einbauen ließen. Darüber hinaus gab es auch Items, die beiläufig beobachtet wurden.

Um dem Kind eine Sitzmöglichkeit zu bieten, befand sich ein Kindertisch mit vier Sesseln im Raum, auf einem davon saß ein Teddybär, für den gekocht werden konnte. Zusätzlich standen auch der Bezugsperson und der Beobachterin ein Tisch und zwei Sessel zur Verfügung.

Die Bezugsperson des Kindes war bei der Testung anwesend und konnte auch in das Spiel mit einbezogen werden. Zu empfehlen war dies, wenn das Kind von der Testleiterin nicht zur Lösung der Aufgaben motiviert werden konnte.

Während die Testleiterin mit dem Kind spielte und versuchte, die vorzugebenden Items durchzuführen, war zusätzlich eine Beobachterin anwesend, die die gelösten und nicht gelösten Items auf einem Protokollbogen (s. Anhang B) notierte. Zudem war es auch ihre Aufgabe, die Videokamera zu bedienen, um die gesamte Spielsituation aufzuzeichnen.

Beide Personen, sowohl die Bezugsperson als auch die Beobachterin, saßen eher abseits und griffen so wenig wie möglich in das Spiel ein, außer wenn es das Kind wünschte oder es für die Motivation des Kindes förderlich war.

8.2. Untersuchungsablauf

Das Kind wurde von der Testleiterin und der Beobachterin meist als Erstes begrüßt, indem ihm auf Augenhöhe die Hand gereicht und es beim Namen angesprochen wurde. Erst danach wurde(n) die begleitende(n) Bezugsperson(en) und gegebenenfalls Geschwisterkinder begrüßt. Jedoch war es nur vier Mal der Fall, dass ein Geschwister mitkam. Bei zwei Untersuchungen waren beide Elternteile anwesend. Sobald der Testraum, in dem zuvor alles vorbereitet wurde, gemeinsam betreten wurde, wurde das mündliche Einverständnis für die Aufzeichnung eingeholt, um sofort mit der Videoaufnahme beginnen zu können. Dies erleichterte u.a. das Protokollieren der Gesamtzeit und der Zeit, die das Kind benötigte, bis es mit der Testleiterin in der Puppenküche interagierte.

Während die Testleiterin versuchte, mit dem Kind in Kontakt zu treten und gemeinsam die Puppenküche zu explorieren, wurde die Bezugsperson von der Beobachterin aufgeklärt, dass das Spielen ca. zwei Stunden in Anspruch nehmen würde und sie zu jeder Zeit mitspielen dürfe, wenn es das Kind möchte. Sie wurde jedoch darauf hingewiesen, keine Gegenstände oder Farben zu benennen, da dies Teil des Verfahrens sei. Im Weiteren wurde sie angeleitet, vom Kind unbeantwortete Fragen zu wiederholen, um eine verwertbare Antwort zu bekommen. Auch wurde die Bezugsperson darüber informiert, dass sehr viele Spiele vorbereitet worden wären, die für das Kind einerseits leicht bzw. altersadäquat sein werden, manche Spiele dem Kind jedoch schwerfallen werden. Es sei jedoch das Ziel dieser Diplomarbeit, jene Spiele herauszufinden bzw. zu überarbeiten, die für das Alter der Zweijährigen noch zu schwierig bzw. nicht interessant sind. Im Weiteren teilte die Beobachterin der Bezugsperson mit, dass Pausen zu jeder Zeit willkommen sind, wenn das Kind etwas essen bzw. trinken möchte oder andere Bedürfnisse hat.

Während diesen anfänglichen Informationen beschäftigte sich das Kind meist schon mit der Puppenküche und interagierte mit der Testleiterin. War dem so, so wurde die Bezugsperson zum Tisch gebeten, um den vorbereiteten Elternfragebogen (Anhang C) auszufüllen und die Einverständniserklärung für die Videoaufzeichnung (Anhang D) zu unterschreiben. Fand eine gemeinsame Exploration der Puppenküche mit der Bezugsperson statt, so dauerte diese meist nur kurz, da sich die Kinder sehr schnell von

den Bezugspersonen lösen konnten (s. Kap. 10.5.2. Lösung und Kontakt zur Bezugsperson).

Bei der Vorgabe der Items wurde den Angaben aus dem Manual (Kuchler et al., 2011) gefolgt.

Da es keine festgelegte Reihenfolge der Items gibt, konnte flexibel entschieden werden, welche Aufgabe als nächste vorgegeben wurde. Dies ermöglichte eine individuelle Anpassung an die Bedürfnisse des Kindes, da die Testleiterin aufgrund des Verhaltens des Kindes entschied, womit gespielt wurde. Im Weiteren wurde das höchst gezeigte Verhalten des Kindes kodiert. Ziel war es, den Kindern alle Items vorzugeben, um ihnen zu ermöglichen, diese zu probieren und zu lösen, um so ihr Optimum an Fähigkeiten zeigen zu können.

Benötigte das Kind Pausen, so wurden diese so lange gemacht, bis sich das Kind von selbst dem Spielmaterial wieder zuwandte.

Waren sich die Testleiterin und die Beobachterin einig, dass dem Kind alle Items zumindest einmal vorgegeben wurden oder ein weiteres Testen aufgrund der Müdigkeit oder des Desinteresses des Kindes nicht mehr möglich gewesen wäre, so wurde aus der Spielebox eine Belohnung geholt. Diese wurde dem Kind von der Testleiterin übergeben, die sich ausdrücklich bei ihm für das tolle Spielen bedankte. Auch der Bezugsperson wurde ein kleines Dankeschön für das Aufbringen der Zeit überreicht.

8.3. Untersuchungsinstrument (Kuchler et al., 2011)

Nun erfolgt die nähere Beschreibung der Fähigkeitsbereiche der visuellen Wahrnehmung, der Motorik, der Aufmerksamkeit und der sozioemotionalen Entwicklung. Um zu erfahren, wie die Sprache, die kognitive Entwicklung und das Gedächtnis erhoben wurden, wird auf Punz (in Vorbereitung) verwiesen. Zusätzlich wird auch kurz auf den Elternfragebogen eingegangen.

8.3.1. Erfassung der visuellen Wahrnehmung

Der Fähigkeitsbereich der visuellen Wahrnehmung wird unterteilt in die aktive und passive Farbdifferenzierung, in die Stabilität des Farbkonzepts und in die Form- und Größendifferenzierung.

Die Erhebung der aktiven und passiven Farbdifferenzierung wird während dem Spielen in der Puppenküche und bei anderen Aufgaben, wie z.B. dem Zeichnen, eingebracht. Insgesamt soll das Kind sieben Farben („rot“, „grün“, „gelb“, „blau“, „orange“, „lila“, „weiß“) aktiv benennen, indem es auf die Frage „Welche Farbe ist das?“ antwortet, und passiv differenzieren, indem es der Frage „Kannst du mir bitte den grünen Stift geben?“ zufolge richtig handelt. Zusätzlich wird erfasst, ob das Farbkonzept des Kindes schon stabil ist, indem es das Spielobst nach Farben („rot“, „gelb“, „grün“ und „orange“) sortieren soll.

Um die Fähigkeiten in der Form- und Größenwahrnehmung zu erfassen, wird überprüft, ob das Kind in der Lage ist, unterschiedliche geometrische Formen und Vierecke und unterschiedlich große Rechtecke und Kreise zu differenzieren.

8.3.1.1. Vorgenommenen Veränderungen

Da sich die Vorgabe der Aufgaben zur Form- und Größendifferenzierung und zur Stabilität des Farbkonzepts in den Vorgängerdiplomarbeiten von Birngruber (2012), Fuchs-Gaderer (2012), Putzer (2013) und Sindelar (2013) als schwierig herausgestellt hatte, wurde diese vor der Erhebung verändert.

Um die Vorgabe zur Erfassung der Stabilität des Farbkonzepts für die Kinder interessanter zu gestalten, wurden vier Stofftiere mit unterschiedlich farbigen Maschen („rot“, „grün“, „gelb“, „orange“) mit einbezogen. Die Kinder wurden gebeten, der Testleiterin beim Füttern der Tiere behilflich zu sein, da die Tiere schon sehr großen Hunger hätten. Daraufhin wurde ihnen erklärt, dass die Tiere nur jenes Futter fressen würden, das dieselbe Farbe wie deren Masche habe.

Die Fähigkeit zur Form- und Größendifferenzierung wurde mittels extra dafür angefertigten Bodenmatten erfasst. Für jede dieser Aufgaben standen jeweils fünf Matten zur Verfügung: das Zielitem wurde vor der Testleiterin auf den Boden gelegt. Die Aufgabe des Kindes war es, eine dem Zielitem idente Form aus den vier Antwortmatten auf dem Boden zu finden („Hast du einen Teppich, der genauso aussieht wie meiner?“). Die Kinder sollten unterschiedliche Formen und Vierecke differenzieren und verschieden große Rechtecke und Kreise, anstatt wie bei Putzer (2013) und Sindelar (2013) verschieden große Dreiecke, unterscheiden.

Zusätzlich wurden die Farben „weiß“, „blau“ und „lila“ nicht mehr bei den Aufgaben zur Form- und Größendifferenzierung passiv abgefragt, sondern beim Auffädeln der Perlen. Die Kinder sollten der Testleiterin die „weiße“, „blaue“ und „lila“ Perle zeigen bzw. auffädeln.

8.3.2. Erfassung der Motorik

Die Erfassung der Motorik gliedert sich in die Bereiche der Grob-, Fein- und Visumotorik. Jeder dieser Unterbereiche besteht wiederum aus mehreren Items. Zusätzlich dienen Aufgaben zum Anziehen der Erhebung der Fein- und Visumotorik und werden zu einer eigenen, kombinierten Skala dieser Fähigkeitsbereiche zusammengefasst.

8.3.2.1. Erfassung der Grobmotorik

Die Erfassung der Grobmotorik besteht aus insgesamt elf Items und unterteilt sich in die Bereiche Gleichgewicht und Ballspiel.

Um das Gleichgewicht zu erfassen, wurden mit den Kindern fünf Turnübungen durchgeführt, die ihnen die Testleiterin zuvor vorgezeigt hatte. Ein weiterer Bereich der Grobmotorik, der im Laufe der Untersuchung erfasst wurde, war das Stiegen Steigen. Es wurde auf einer dreistufigen Stiege, die an einer Wand platziert war, beobachtet, ob die Kinder im Wechsel- oder Nachstellschritt bzw. mit oder ohne anhalten diese hinauf bzw. hinunter gingen und ob sie von der letzten Stufe hüpfen konnten. Um die Kinder für diese Aufgabe zu motivieren, wurde ein Luftballon an der Wand befestigt, den sie herunterholen sollten. Zusätzlich wurde erfasst, ob es den Kindern gelang, den Ball zu werfen, zu fangen und mit dem Fuß weg zu kicken.

8.3.2.2. Erfassung der Fein- und Visumotorik (Anziehen)

Im Weiteren wurde beobachtet, ob sich die Kinder alleine anziehen können (z.B. Schuhe anziehen und Hut aufsetzen) bzw. die einzelnen Schritte beim Anziehen einer Hose und eines T-Shirts alleine meistern können. Der Einschätzung dieses Bereiches dienen insgesamt sieben Items.

8.3.2.3. Erfassung der Feinmotorik

Die Feinmotorik wurde während dem Spielen in der Puppenküche und auch beim Zeichnen erhoben. Auch die Items bzgl. des Anziehens wurden noch einmal gesondert diesem Fähigkeitsbereich zugeteilt. Insgesamt dienten 14 Items der Erhebung dieses der Feinmotorik.

8.3.2.4. Erfassung der Visumotorik

Um den Funktionsbereich der Visumotorik zu erfassen, wurden den Kindern 11 Items vorgegeben. Diese wurden in das Spiel in der Puppenküche und beim Zeichnen eingebaut. Zusätzlich diente das Bauen mit Bausteinen, das Auffädeln von Perlen und das Puzzlebauen der Erhebung dieses Fähigkeitsbereiches.

8.3.2.5. Vorgenommene Veränderungen

In den Untersuchungen bei Putzer (2013) und Sindelar (2013) kamen zwei Bälle in unterschiedlichen Größen zum Einsatz. Für diese Untersuchung wurde den Kindern nur mehr ein Ball mit einem Durchmesser von 12cm zur Verfügung gestellt.

Um die Fähigkeiten der Kinder beim Anziehen spielerischer und standardisierter zu erfassen, konnten sie sich in einen Clown verkleiden, indem sie in ein Kostüm, bestehend aus einem Hut, einem T-Shirt und einer Hose, schlüpften. Die Kodierung dieser Aufgaben wurde differenzierter als bei den Vorgängerinnen vorgenommen.

Eine weitere Veränderung in der Vorgabe war es, die Kinder nicht mehr von einer Vorlage die gewünschten Formen abzeichnen zu lassen, sondern gemeinsam mit ihnen ein Bild zu malen. Die Testleiterin malte einen Wurm, der aus einigen Kreisen bestand. Das Kind wurde aufgefordert, den Wurm länger zu malen, indem es ein paar Kreise hinzu malt. Danach wurde eine Wiese, bestehend aus einem horizontalen Strich für den Boden und aus vielen vertikalen Strichen für die Grashalme, gezeichnet. Welches Bild dargestellt wurde, hing vom Interesse des Kindes ab und wurde beliebig angepasst. Interessierte sich z.B. ein Kind sehr für Autos, so wurden die Reifen mit Kreisen abgebildet, die Straße mit einem horizontalen Strich und der Stamm eines Baumes mit vertikalen Linien.

8.3.3. Erfassung der Aufmerksamkeit

Für die Erhebung der Aufmerksamkeit wurden keine zusätzlichen Materialien benötigt. Während der Durchführung bestimmter Spiele, wie beim Zeichnen, beim Bauen mit den Bausteinen und des Puzzles, beim Auffädeln der Perlen und bei der Vorgabe des Buches, wurde erfasst, wie viele Sekunden lang das Kind der Aufgabe aufmerksam folgen konnte. Zusätzlich wurde protokolliert, wenn es unaufmerksam geworden war, ob es sich wieder zurückführen ließ und wie viele Seiten es sich konzentriert im Buch ansah.

8.3.3.1. Vorgenommene Veränderungen

Putzer (2013) und Sindelar (2013) gaben bei ihren Untersuchungen insgesamt drei Bücher vor, wobei während der Durchführung jedes Buches protokolliert wurde, wie viele Sekunden das Kind aufmerksam war. Bei den Testungen für diese Arbeit wurde auf die Durchführung von zwei Büchern verzichtet.

8.3.4. Erfassung der sozioemotionalen Entwicklung

Der sozioemotionale Entwicklungsstand des Kindes wurde erhoben, indem die Lösung des Kindes von der Bezugsperson am Beginn der Testung und der Kontakt zur Bezugsperson während der Testung beobachtet wurden. Diese beiden Items wurden jeweils auf einer fünfstufigen Ratingskala von der Testleiterin und der Beobachterin eingeschätzt. Zusätzlich war von Interesse, inwieweit das Kind versuchte, sich selbstständig zu verkleiden und ob es dies auch schaffte. Beobachtet wurde hier, ob das Kind beim Anziehen sofort um Hilfe bat oder ob es das Verkleiden zuerst selbst versuchte. In weiterer Folge wurde kodiert, ob das Kind sich selbst anziehen konnte.

8.3.4.1. Vorgenommene Veränderungen

Putzer (2013) und Sindelar (2013) stellten den Kindern u. a. ein Kleid zum Anziehen zur Verfügung. Da jedoch jedes Kind, egal welchem Geschlecht zugehörig, dieselben Bedingungen haben sollte, wurde dieses aus dem Itempool genommen, um die Vorgabe zu standardisieren. Die Schuhe (Crocs), die von Putzer (2013) und Sindelar (2013) eingeführt wurden, wurden beibehalten, eine Hose, ein T-Shirt und ein Hut eines Clownkostüms wurden neu in den Itempool aufgenommen.

8.3.5. Elternfragebogen

Der Elternfragebogen (s. Anhang C), der von Birngruber (2012) und Fuchs-Gaderer (2012) in Anlehnung an den Elternfragebogen zur Früherkennung von Risikokindern (ELFRA) von Grimm und Doil (2006) und den Elternfragebogen der Diplomarbeit von Metzler (2003) erfasst worden war, wurde zu Beginn der Untersuchungen der anwesenden Bezugsperson vorgelegt, mit der Bitte darum, ihn auszufüllen. Darin wurden soziodemographische Daten des Kindes erfasst und auf einer vierstufigen Ratingskala Aussagen zur Selbstständigkeit, zum Sozial- und Spielverhalten der Kinder gemacht. Zusätzlich wurden Informationen zum Wortschatz und zu den grammatikalischen Kompetenzen der Kinder eingeholt.

8.4. Die Stichprobe

8.4.1. Rekrutierung der Stichprobe

Der Großteil der 20 Kinder wurde im Dezember 2012 und Jänner 2013 über eine Kinderarztpraxis in Wien akquiriert. Es wurden sowohl Elternbriefe (s. Anhang E) in der Ordination ausgehängt, als auch die Mütter persönlich darauf angesprochen, ob sie Interesse hätten an einer Diplomarbeit mitzuwirken. Im Zuge dieses ersten Kontaktes wurden diese über den Ablauf, die Dauer und das Ziel dieser Untersuchung informiert. Viele Mütter waren sehr interessiert, stellten Fragen und teilten sofort deren Zusage mit. Daraufhin wurden deren Telefonnummern und E-Mail-Adressen notiert, um im gegebenen Fall genauere Informationen senden zu können. Zusätzlich wurden auch in einigen Kindergärten und Kinderkrippen die Elternbriefe ausgehängt und im privaten Umfeld nach interessierten Eltern mit zweijährigen Kindern gefragt.

Noch während der laufenden Akquirierung der Kinder wurden im Jänner 2013, meist telefonisch, aber auch per E-Mail, Termine mit den Müttern fixiert, die bereits im persönlichen Gespräch der Teilnahme an der Untersuchung zugesagt hatten. Die Terminvereinbarung gestaltete sich als sehr flexibel, da der Kleinkindertestraum des Institutes für Psychologie im Februar 2013 an drei ganzen Tagen der Woche für diese Untersuchung zur Verfügung stand und somit die Termine größtenteils zu Zeiten festgelegt werden konnten, zu denen die Kinder zum Spielen motiviert waren.

8.4.2. Soziodemographische Merkmale der Stichprobe

Es gehen 20 Kinder in die Stichprobe ein, wobei das Geschlechterverhältnis mit zehn Buben und zehn Mädchen ausgewogen ist. Das durchschnittliche Testalter der Kinder betrug 28.5 Monate ($SD = 3.154$), wobei das jüngste Kind zum Zeitpunkt der Untersuchung 24 und das älteste Kind 35 Monate alt war.

In Tabelle 1 wird das Alter der Kinder für Mädchen und Buben getrennt aufgelistet.

Tabelle 1: Darstellung des Alters der Kinder in Monaten, getrennt nach Geschlecht

Anzahl		Alter in Monaten											Gesamt
		24	25	26	27	28	29	30	31	33	34	35	
Geschlecht	männlich	1	1	3	0	2	1	0	1	0	0	1	10
	weiblich	0	1	1	2	1	0	1	2	1	1	0	10
Gesamt		1	2	4	2	3	1	1	3	1	1	1	20

Aus den vorgegebenen Elternfragebögen (s. Anhang C) ließen sich folgende Informationen einholen: Die Väter der Kinder waren im Durchschnitt 34.95 Jahre ($SD = 7.082$), die Mütter 32 Jahre ($SD = 5.181$) alt, wobei die jüngsten Väter und Mütter jeweils 23, die ältesten Väter 46 und die ältesten Mütter 42 Jahre alt waren.

Gleich viele Mütter wie Väter, nämlich jeweils sieben (35%), wiesen einen Universitäts- bzw. Fachhochschulabschluss vor, sechs Mütter (30%) und vier Väter (20%) absolvierten die Matura. Eine Lehre bzw. Fachschule beendeten sieben Mütter (35%) und acht Väter (40%) und ein Vater (5%) vollendete alleine die Pflichtschule.

Alle 20 Mütter und 18 Väter (90%) leben mit ihren Kindern in einem gemeinsamen Haushalt. Neun Kinder (45%) haben ein Geschwister, von denen drei Kinder ein älteres und sechs Kinder ein jüngeres Geschwister haben. Mehrsprachig wachsen sieben Kinder (35%) auf, sechs davon erlernen zwei Sprachen und ein Kind drei Sprachen.

Den Kindergarten besuchen 17 Kinder (85%) und zusätzliche Sport- oder Freizeitkurse acht Kinder (40%). Nur eine Mutter berichtet über eine Risikoschwangerschaft (5%). Unter gesundheitlichen Problemen leidet laut den Angaben aus den Fragebögen kein Kind.

9. Auswertung

Es wird nun auf die Methoden der statistischen Auswertung eingegangen, um die nachfolgend aufgelisteten Ergebnisse zu den einzelnen Fähigkeitsbereichen verständlich zu machen.

Um die erste Fragestellung bezüglich der Aussagekraft der einzelnen Items beantworten zu können, werden die Schwierigkeiten und die Trennschärfen der Items berechnet.

Der Schwierigkeitsindex gibt die prozentuelle Häufigkeit an, mit der das Item von der Stichprobe gelöst wurde (Lienert & Raatz, 1994). Liegt dieser Wert zwischen .20 und .80 spricht man von einer adäquaten Schwierigkeit (Bortz & Döring, 2006, S. 219). Liegt der Wert unter .20 bedeutet dies, dass die Lösung dieses Items für die Stichprobe sehr schwierig war und das Item somit nur selten gelöst wurde. Ein Wert von über .80 zeigt, dass das Item der Stichprobe sehr leicht gefallen ist und somit oft gelöst wurde.

Die Trennschärfe eines Items gibt an, wie gut ein Item den Gesamtwert einer Skala vorhersagen kann (Bortz & Döring, 2006). Eine mittlere Trennschärfe liegt laut Weise (1975, S. 219, zitiert nach Bortz & Döring, 2006, S. 220) zwischen .30 und .50, von einer hohen Trennschärfe spricht man, wenn der Wert höher als .50 ist. Liegt der Wert eines Items unter .30 bedeutet dies, dass dieses Item nicht genügend dazu beiträgt, den Gesamtwert einer Skala voraus zu sagen und somit von der Skala ausgeschlossen werden soll (Bortz & Döring, 2006). Zu beachten ist laut Bortz und Döring (2006), dass die Trennschärfe eines Items von seiner Schwierigkeit abhängt. Items mit mittleren Schwierigkeiten besitzen die höchsten Trennschärfen, während bei sehr leichten und sehr schwierigen Items die Trennschärfen niedrig sind.

Oftmals erhöht sich durch Vernachlässigung jener Items mit geringen Trennschärfen das Cronbach Alpha (α) wesentlich. Dieser Wert ist eine Kennzahl für die innere Konsistenz bzw. der **Reliabilität** (siehe Kapitel 2.2.3. Innere Konsistenz). Lässt sich ein Cronbach α von über .70 berechnen, so kann man laut Field (2009, S. 675) schlussfolgern, dass die Skala zufriedenstellend genau misst und das Gütekriterium der Reliabilität als gegeben angesehen werden kann. Ein Wert von über .90 entspricht einer hohen Reliabilität (Weise, 1975, S. 210, zitiert nach Bortz & Döring, 2009, S. 199).

Ob die **Objektivität** als gegeben angenommen werden kann, wird im Anschluss der Ergebnisse zur statistischen Auswertung erläutert, da für diese keine Kennwerte berechnet werden können. Die Beurteilung wird aufgrund der Beobachtungen aus den Untersuchungen vorgenommen.

Bezüglich der **Validität** werden Alterseffekte analysiert. Sind die Voraussetzungen pro Skala gegeben, werden Interkorrelationen des Alters der Kinder in Monaten und der Testleistung mittels der Pearson-Korrelation berechnet. Bei einer Verletzung der Voraussetzungen wird das Kendall-tau herangezogen, das laut Field (2009) eine bessere Einschätzung als die Spearman-Korrelation erlaubt. Dieser Korrelationskoeffizient dient der Einschätzung der Konstruktvalidität für die einzelnen Skalen. Können signifikante Korrelationen beobachtet werden, so bedeutet dies, dass die Kinder mit ansteigendem Alter bessere Leistungen erbringen und das Gütekriterium der Konstruktvalidität angenommen werden kann.

Weitere Erkenntnisse zur Konstruktvalidität liefern Geschlechts- bzw. Altersunterschiede. Bei gegebenen Voraussetzungen der Normalverteilung und der Homogenität der Varianzen der abhängigen Variablen werden diese mittels einer ANOVA berechnet. Hierfür wird das Alter der Kinder in zwei Gruppen aufgeteilt. Die Kinder, die zwischen 24 und 29 Monaten alt sind, bilden die jüngere Altersgruppe, und die Kinder zwischen dem 30. und dem 36. Lebensmonat die ältere Gruppe.

In der jüngeren Altersgruppe sind insgesamt 13 Kinder (acht Burschen und fünf Mädchen) und weitere sieben Kinder gehören der älteren Altersgruppe an (zwei Burschen und fünf Mädchen).

In der Tabelle 2 ist diese Aufteilung dargestellt.

Tabelle 2: Aufteilung der Stichprobe in die jüngere und ältere Altersgruppe

Anzahl		Alter in Monaten											Gesamt
		24	25	26	27	28	29	30	31	33	34	35	
Geschlecht	männlich	1	1	3	0	2	1	0	1	0	0	1	10
	weiblich	0	1	1	2	1	0	1	2	1	1	0	10
Gesamt		1	2	4	2	3	1	1	3	1	1	1	20
Altersgruppen	männlich	8						2					10
	weiblich	5						5					10
Gesamt		13						7					20

Die Normalverteilungen der Daten der abhängigen Variablen werden hauptsächlich mittels der Histogramme und den Normalverteilungskurven und zusätzlich mit den Kennwerten zur Schiefe und Kurtosis überprüft (s. Anhang A). Der Levene-Test weist auf eine Homogenität der Varianzen hin, wenn $p > .05$ und über die Heterogenität der Varianzen wenn $p \leq .05$ ist (s. Anhang A).

Ist die Normalverteilung gegeben, die Voraussetzung der Homogenität der Varianzen jedoch nicht, wird ein T-Test für unabhängige Stichproben herangezogen.

Sind beide Voraussetzungen nicht gegeben bzw. ist nur die Voraussetzung der Normalverteilung verletzt, so wird auf den parameterfreien Mann-Whitney-U-Test zurückgegriffen.

Die Geschlechtsunterschiede werden zweiseitig mit einem Signifikanzniveau α von $p = .05$ getestet, da kein Unterschied in den Leistungen der Mädchen und Buben angenommen wird. Jedoch wird davon ausgegangen, dass ältere Kinder höhere Werte erreichen als jüngere Kinder. Somit werden die Alterunterschiede einseitig mit einem Signifikanzniveau α von $p = .025$ getestet. Da die Stichprobe sehr klein ist, wurden für die Mann-Whitney-U-Tests die exakten Signifikanzen für eine einseitige bzw. zweiseitige Testung angegeben, um detailliertere Ergebnisse zu erhalten.

Im Zuge der Frage, ob es einen Zusammenhang zwischen den Einschätzungen der Bezugspersonen im Fragebogen und den Beobachtungen und Erfahrungen aus der Testsituation zur Selbstständigkeit gibt (Kriteriumsvalidität), werden die Voraussetzungen für eine Pearson-Korrelation und bei Verletzung dieser ein Kendall-tau berechnet, da diese Korrelation laut Field (2009) der Spearman-Korrelation vorzuziehen

ist. Eine signifikante Korrelation ($p \leq .05$) bedeutet, dass dieser Zusammenhang signifikant von 0 (keine Korrelation) abweicht und somit anzunehmen ist. Fällt eine Korrelation nicht signifikant aus ($p > .05$), wird davon ausgegangen, dass kein Zusammenhang besteht (Field, 2009) und die Kriteriumsvalidität für die Skala „Selbstständigkeit“ als nicht gegeben anzusehen ist.

10. Ergebnisse

Im Folgenden wird eine detaillierte Auflistung der deskriptiven Analyse zur Gesamtdauer, zur Anzahl und der Dauer der Pausen gemacht. Anschließend werden die statistischen Ergebnisse inklusive der Reliabilitätskennwerte der Fähigkeitsbereiche „Visuelle Wahrnehmung“, „Motorik“, „Aufmerksamkeit“ und „Sozioemotionale Entwicklung“ ausführlich beschrieben. Zusätzlich werden im Anschluss Stellungnahmen zu den Gütekriterien der Objektivität und der Validität gemacht.

10.1. Anzahl und Dauer der Pausen

Aufgrund der Videoaufzeichnungen der Spielsituationen können folgende Zeitangaben gemacht werden:

Die Gesamttestdauer setzt sich zusammen aus der Zeit vom Betreten bis zum Verlassen des Testraums durch das Kind und seine Bezugsperson. Die durchschnittliche Gesamttestdauer betrug $MW = 120.39$ Minuten ($SD = 16.747$), wobei die kürzeste Testung 96 Minuten und die längste 160 Minuten dauerte. Die durchschnittliche Beschäftigungsdauer mit dem Material, also Gesamttestdauer minus der Anlaufzeit ($MW = 4.08$; $SD = 3.778$) und der Verabschiedung ($MW = 12.20$; $SD = 6.321$), dauerte $MW = 99.42$ Minuten ($SD = 11.426$), wobei die kürzeste Beschäftigung 80 Minuten und die längste 124 Minuten dauerte.

Die Kinder benötigten zwischen keiner und sieben Pausen, wobei zwei Kinder (10%) keine, 12 Kinder (60%) eine und zwei Kinder (10%) sieben Pausen benötigten. Die restlichen 20% benötigten zwischen zwei und drei Pausen. Die Länge der Pausen variierte zwischen einer und 20 Minuten. Die durchschnittliche Pausendauer betrug $MW = 4.64$ Minuten ($SD = 5.127$). Die Hälfte der Kinder benötigte drei Minuten Pause während des Spielens.

10.2. Visuelle Wahrnehmung

Für den Fähigkeitsbereich der visuellen Wahrnehmung mit den einzelnen Skalen der „Aktiven Farbdifferenzierung“, der „Passiven Farbdifferenzierung“, der „Stabilität des Farbkonzepts“ und der „Form- und Größendifferenzierung“ konnte vor der Selektion der nicht trennscharfen Items ein Cronbach Alpha von .902 und nach der Selektion des Items „rot“ (aktive Farbdifferenzierung) ein zufriedenstellendes Cronbach Alpha von .901 berechnet werden.

Nun werden die Ergebnisse der einzelnen Skalen detailliert beschrieben.

10.2.1. Aktive Farbdifferenzierung

Für die Skala der aktiven Farbdifferenzierung, in die die Werte von 19 Kindern mit einfließen, konnte für die sieben Items ein Cronbach Alpha von .826 berechnet werden, welches als sehr zufriedenstellend gewertet werden kann. Die Schwierigkeiten aller Items liegen im mittleren Bereich zwischen .20 und .80. Da die Werte der Trennschärfen für die Items „grün“, „orange“, „blau“ und „lila“ hoch ausfallen, können diese Items das Gesamtergebnis der Skala sehr gut repräsentieren. Auch die Items „gelb“ und „weiß“ weisen zufriedenstellende Werte auf. Die Trennschärfe des Items „rot“ liegt jedoch mit einem Wert von .13 im zu niedrigen Bereich. Dieses Item wird für die weiteren Analysen ausgeschlossen und so erhöht sich das Cronbach Alpha auf .870. Die Skala enthält nun sechs Items.

In Tabelle 3 sind alle Werte der Itemanalyse ersichtlich.

Tabelle 3: Ergebnisse der Itemanalyse zur aktiven Farbdifferenzierung

N = 19	Schwierigkeit	Trennschärfe	Trennschärfe nach der Selektion
rot	.58	.13	-
grün	.47	.82	.83
gelb	.58	.49	.60
orange	.32	.74	.62
weiß	.47	.47	.51
blau	.47	.82	.83
lila	.53	.61	.64

Zusätzlich wurde eine Skalenanalyse gemacht, um mögliche Geschlechts- und Altersunterschiede feststellen zu können.

Die Mann-Whitney-U-Tests liefern keine signifikanten Ergebnisse bzgl. des Geschlechts und des Alters. Mädchen und Buben zeigen gleich gute Leistungen in der aktiven Farbdifferenzierung aller vorgegebenen Items ($U = 25.00$, $z = -1.664$, $p = .105$, $r = -.37$) und auch die älteren Kinder unterscheiden sich von den jüngeren Kindern nicht signifikant ($U = 23.00$, $z = -1.636$, $p = .056$, $r = -.37$). Auch lässt sich kein Zusammenhang erkennen, dass die Kinder mit ansteigendem Alter bessere Leistungen in der aktiven Farbdifferenzierung erzielen ($\tau = .243$, $p = .091$).

10.2.2. Passive Farbdifferenzierung

Diese Skala, die die Werte von 19 Kindern beinhaltet, zeigt ein α von .766, das zufriedenstellend ausfällt. Die Schwierigkeiten von sechs Items liegen im angemessenen Bereich, außer die des Items „gelb“ (.89). Die passive Farbdifferenzierung dieser Farbe fiel der untersuchten Stichprobe sehr leicht. Die Trennschärfen aller Items liegen zumindest im mittelmäßigen Bereich. Die Items „grün“, „gelb“ und „orange“ weisen sogar hohe Trennschärfen von über .50 auf. Um das Cronbach Alpha nicht zu verringern, darf kein Item selektiert werden.

In Tabelle 4 wird eine übersichtliche Darstellung aller Ergebnisse der Itemanalyse zur passiven Farbdifferenzierung gegeben.

Tabelle 4: Ergebnisse der Itemanalyse zur passiven Farbdifferenzierung

N = 19	Schwierigkeit	Trennschärfe
rot	.68	.42
grün	.53	.62
gelb	.89	.51
orange	.42	.58
weiß	.63	.43
blau	.68	.42
lila	.79	.45

Hinsichtlich der Skalenanalyse ergeben sich beim Mann-Whitney-U-Test keine signifikanten Geschlechtsunterschiede ($U = 37.00$, $z = -.996$, $p = .338$, $r = -.22$). Auch die älteren Kinder unterscheiden sich von den jüngeren nicht signifikant ($U = 36.50$, $z = -.723$, $p = .249$, $r = -.16$) und es kann nicht von einem Zusammenhang zwischen Alter und Testleistung ausgegangen werden ($\tau = .163$, $p = .176$).

10.2.3. Stabilität des Farbkonzepts

Das Cronbach α für diese Skala liegt bei .857, das einen hohen Wert darstellt. Die Schwierigkeiten aller vier Items liegen im mittleren, angemessenen Bereich. Die Trennschärfen sind alle höher als .50 und somit als trennscharf einzuschätzen. Auch bei dieser Skala darf kein Item ausgeschlossen werden, um eine Verringerung der Reliabilität zu vermeiden. In diese Skala fließen die Werte von allen 20 Kindern mit ein.

Alle Ergebnisse der Itemanalyse sind in der Tabelle 5 aufgelistet.

Tabelle 5: Ergebnisse der Itemanalyse zur Stabilität des Farbkonzepts

N = 20	Schwierigkeit	Trennschärfe
rot	.55	.67
grün	.55	.77
gelb	.40	.70
orange	.55	.67

Hinsichtlich der Leistungsunterschiede zwischen Mädchen und Buben zeigen sich beim Mann-Whitney-U-Test keine signifikanten Unterschiede ($U = 36.50$, $z = -1.063$, $p = .308$, $r = -.24$).

Auch für die Analyse des Altersunterschiedes wurde ein Mann-Whitney-U-Test angewandt. Ältere Kinder ($MW = 3.14$, $SD = 1.46$) unterscheiden sich signifikant ($U = 19.00$, $z = -2.188$, $p = .017$, $r = -.49$) von den jüngeren Kindern ($MW = 1.46$, $SD = 1.56$). Zusätzlich konnte gezeigt werden, dass je älter die Kinder sind, desto stabiler ist das Farbkonzept ($\tau = .595$, $p = .001$).

10.2.4. Form- und Größendifferenzierung

Hier wurde ein Cronbach Alpha von .710 berechnet, das gerade noch im zufriedenstellenden Bereich liegt. Da jedoch in diese Skala die Werte von nur 17 Kindern mit einfließen, wurden alle fehlenden Werte auf 0 gesetzt und somit ein erhöhtes von .735 berechnet. Die Itemanalyse mit den auf 0 gesetzten Werten zeigt, dass die Schwierigkeiten aller vier Items im akzeptablen Bereich liegen. Im Weiteren ist zu sehen, dass drei von vier Items hohe Trennschärfen von über .50 aufweisen und auch das vierte Item („Größendifferenzierung II“) einen Wert im annehmbaren Bereich zeigt.

Aus diesem Grund wird auch für diese Skala keine Selektion der Items vorgenommen.

Alle weiteren Werte zu dieser Itemanalyse sind der Tabelle 6 zu entnehmen.

Tabelle 6: Ergebnisse der Itemanalyse zur Form- und Größendifferenzierung

N = 20	Schwierigkeit	Trennschärfe
Formdifferenzierung I (versch. Formen)	.65	.55
Formdifferenzierung II (versch. Vierecke)	.50	.62
Größendifferenzierung I (Rechtecke)	.35	.61
Größendifferenzierung II (Kreise)	.25	.34

Um Aussagen darüber treffen zu können, ob sich Mädchen und Buben in den Ergebnissen der Form- und Größendifferenzierung unterscheiden, wurde ein Mann-Whitney-U-Test herangezogen. Hier zeigen sich keine signifikanten Geschlechtsunterschiede ($U = 35.50$, $z = -1.121$, $p = .271$, $r = -.25$).

Der Mann-Whitney-U-Test liefert jedoch für den Altersunterschied ein signifikantes Ergebnis ($U = 11.00$, $z = -2.796$, $p = .003$, $r = -.63$). Demnach zeigen ältere Kinder ($MW = 3.00$, $SD = 1.15$) signifikant bessere Leistungen in der Form- und Größendifferenzierung als jüngere Kinder ($MW = 1.08$, $SD = 1.12$). Daraus kann die Schlussfolgerung gezogen werden, dass die Fähigkeit in der Form- und Größendifferenzierung mit steigendem Alter zunimmt ($\tau = .450$, $p = .006$).

10.3. Motorik

Für den Fähigkeitsbereich der Motorik, der sich in die Grob-, Fein- und Visumotorik unterteilen lässt und insgesamt 33 Items beinhaltet, konnte vor der Selektion ein akzeptables Cronbach α von .773 berechnet werden. Nach der Selektion der nicht trennscharfen Items („von Stiege hüpfen“, „Ball werfen“, „Ball fangen“, „Fußkick“, „Zehengang“, „beidbeinig hüpfen“, „stehen bleiben“, „Kochlöffel aufhängen“, „Glas öffnen“, „auffädeln vorgezeigt“, „aufkehren“, „umfüllen“ und „daneben schütten“) konnte es auf .814 erhöht werden und fällt somit zufriedenstellend aus.

Im Folgenden werden nun im Detail die Itemanalysen der einzelnen Bereiche näher dargestellt.

10.3.1. Grobmotorik

Für die Skala der Grobmotorik, bestehend aus elf Items, wurde ein nicht zufriedenstellendes Cronbach α von .490 berechnet. Sie enthält sehr leichte Items, wie z.B. „Ball werfen“, „Fußkick“ und „beidbeinig hüpfen“ und sehr schwierige Items, wie z.B. „Ball fangen“, „Seitwärtsgang“ und „stehen bleiben“. Das Item „Ball werfen“ wird von der Analyse ausgeschlossen, da es alle Kinder gelöst haben, somit eine Varianz von 0 entstand und es keine Aussagekraft hat. Viele Items zeigen sehr niedrige Trennschärfen, wie z.B. „von Stiege hüpfen“, „Ball fangen“, „Fußkick“, „Zehengang“, „beidbeinig hüpfen“ und „stehen bleiben“. Werden diese 6 Items ausgeschlossen, so erhöht sich das Cronbach α auf .641, das dennoch nicht im zufriedenstellenden Bereich liegt. Sowohl vor als auch nach der Itemselektion fließen die Werte von 18 Kindern mit ein.

In Tabelle 7 sind alle Werte der Itemanalyse ersichtlich.

Tabelle 7: Ergebnisse der Itemanalyse zur Grobmotorik

N = 18	Schwierigkeit	Trennschärfe	Trennschärfe nach der Selektion
Stiege hinauf	.69	.45	.43
Stiege herab	.54	.34	.64
von Stiege hüpfen	.72	.25	-
Ball werfen	1.00	-	-
Ball fangen	.11	.003	-
Fußkick	.89	-.01	-
balancieren	.28	.33	.40
Seitwärtsgang	.11	.40	.35
Zehengang	.56	.22	-
beidbeinig hüpfen	.83	.14	-
stehen bleiben	.11	.005	-

Die Mann-Whitney-U-Tests zeigen signifikante Geschlechtsunterschiede ($U = 18.00$, $z = -1.994$, $p = .049$, $r = -.47$) und Altersunterschiede ($U = 19.50$, $z = -1.756$, $p = .043$, $r = -.41$). Demnach erzielten Mädchen ($MW = 2.05$, $SD = 0.32$) bessere Leistungen in der Grobmotorik als Buben ($MW = 1.09$, $SD = 0.16$) und ältere Kinder ($MW = 2.07$, $SD = 0.36$) bessere Leistungen als jüngere Kinder ($MW = 1.34$, $SD = 0.25$). Auch hier ist erwiesen, dass die Kinder in dieser Stichprobe mit steigendem Alter höhere Leistungen in der Grobmotorik erzielten ($\tau = .320$, $p = .043$).

10.3.2. Fein- und Visumotorik

Die Skala Fein- und Visumotorik besteht aus sieben Anzieh-Items und erreicht ein α von .860, das sehr zufriedenstellend ausfällt. Die Analyse beinhaltet die Werte von allen 20 Kindern. Alle Items, bis auf eines, zeigen eine adäquate Schwierigkeit zwischen .20 und .80. Das Item „in Ärmel schlüpfen“ fiel der Stichprobe jedoch sehr schwer.

Die Trennschärfen können als mittelmäßig bis hoch beurteilt werden und somit kommt es zu keiner Selektion der Items.

In Tabelle 8 sind alle Werte zur Schwierigkeit und zur Trennschärfe ersichtlich.

Tabelle 8: Ergebnisse der Itemanalyse zur Fein-/Visumotorik

N = 20	Schwierigkeit	Trennschärfe
Schuhe anziehen	.70	.48
Hosenbein anziehen	.33	.76
Hose hinauf ziehen	.20	.85
Hut aufsetzen	.80	.32
in Ärmel schlüpfen	.10	.54
mit Kopf in T-Shirt schlüpfen	.20	.65
T-Shirt hinunter ziehen	.20	.85

Mittels Mann-Whitney-U-Tests konnte kein signifikanter Altersunterschied ($U = 29.00$, $z = -1.343$, $p = .097$, $r = -.30$) und auch kein Zusammenhang, dass die Leistungen in diesem Bereich mit steigendem Alter zunehmen, festgestellt werden ($\tau = .072$, $p = .343$). Jedoch zeigt der signifikante Geschlechtsunterschied ($U = 18.50$, $z = -2.445$, $p = .012$, $r = -.55$), dass Mädchen ($MW = 3.70$, $SD = 0.75$) bessere Leistungen beim Anziehen von Kleidung erzielen als die Buben ($MW = 1.35$, $SD = 0.28$).

10.3.3. Feinmotorik

Für die Skala der Feinmotorik, bestehend aus 14 Items, wurde ein α von .850 berechnet, das im zufriedenstellenden Bereich liegt. Diese Skala weist sehr leichte Items („Teller einordnen“, „Kochlöffel aufhängen“, „Glas öffnen“, „Stifthaltung“ und „Blatt halten“) auf. Ein Item („in Ärmel schlüpfen“) fiel der Stichprobe sehr schwer. Alle weiteren Items weisen adäquate Schwierigkeiten zwischen .20 und .80 auf. Die Trennschärfen können als mittelmäßig bis hoch eingestuft werden, die Items „Kochlöffel aufhängen“, „Glas öffnen“ erreichen jedoch sehr niedrige Trennschärfen, sodass das Ausschließen dieser zwei Items das Cronbach α auf .864 erhöht. Sowohl vor, als auch nach der Selektion können die Werte von allen 20 Kindern in die Analyse mit hineingenommen werden.

Die weiteren Ergebnisse dieser Itemanalyse sind der Tabelle 9 zu entnehmen.

Tabelle 9: Ergebnisse der Itemanalyse zur Feinmotorik

N = 20	Schwierigkeit	Trennschärfe	Trennschärfe nach der Selektion
Teller einordnen	.90	.55	.54
Kochlöffel aufhängen	.90	.25	-
Glas öffnen	.90	.13	-
Glas schließen	.75	.50	.48
Stifthaltung	.93	.43	.48
Unterarm am Tisch	.70	.31	.38
Blatt halten	.95	.44	.48
Schuhe anziehen	.70	.61	.61
Hosenbein anziehen	.33	.75	.73
Hose hinauf ziehen	.20	.72	.72
Hut aufsetzen	.80	.45	.46
in Ärmel schlüpfen	.10	.47	.47
mit Kopf in T-Shirt schlüpfen	.20	.60	.59
T-Shirt hinunter ziehen	.20	.72	.72

Beim Mann-Whitney-U-Test ergibt sich kein signifikanter Unterschied zwischen den jüngeren und älteren Kindern ($U = 32.00$, $z = -1.078$, $p = .151$, $r = -.24$) und auch konnte kein signifikanter Zusammenhang zwischen dem Alter und Testleistung festgestellt werden ($\tau = .028$, $p = .434$). Jedoch konnten signifikante Unterschiede zwischen Mädchen und Buben berechnet werden ($U = 24.00$, $z = -1.980$, $p = .047$, $r = -.44$). Demzufolge erzielen Mädchen ($MW = 8.25$, $SD = 0.90$) bessere Leistungen im Bereich der Feinmotorik als Burschen ($MW = 5.25$, $SD = 0.73$).

10.3.4. Visumotorik

Hier wurde für elf Items ein α von .572 berechnet, das im niedrigen Bereich liegt. Die Schwierigkeiten liegen im adäquaten bis hohen Bereich. Die Items „umfüllen“ und „Perlen auffädeln“ fielen der Stichprobe sehr leicht. Die Trennschärfen sind alle im mittelmäßigen bis hohen Bereich angesiedelt.

Allerdings kommt es zu einer Selektion von vier Items („auffädeln vorgezeigt“, „aufkehren“, „umfüllen“ und „daneben schütten“), da diese niedrige bzw. negative Trennschärfen aufweisen. Das Cronbach α erhöht sich aufgrund dieser Selektion auf .796, das nun als akzeptabel einzustufen ist. In diese Analyse gehen 19 Werte ein.

Eine detaillierte Auflistung der Ergebnisse zur Itemanalyse ist in Tabelle 10 zu sehen.

Tabelle 10: Ergebnisse der Itemanalyse zur Visumotorik

N = 19	Schwierigkeit	Trennschärfe	Trennschärfe nach der Selektion
horizontale Linie zeichnen	.68	.55	.73
vertikale Linie zeichnen	.58	.47	.66
Kreis zeichnen	.74	.62	.50
mit Schere schneiden	.74	.30	.41
Turm bauen	.45	.35	.60
Perlen auffädeln	.84	.52	.40
auffädeln vorgezeigt	.58	-.19	-
Puzzle bauen	.79	.31	.45
aufkehren	.68	.19	-
umfüllen	.84	.24	-
daneben schütten	.74	-.31	-

Es lassen sich mittels Mann-Whitney-U-Test keine signifikanten Geschlechtsunterschiede feststellen. Mädchen zeigen demnach gleich gute Leistungen im Bereich der Visumotorik wie Burschen ($U = 26.00$, $z = -1.558$, $p = .126$, $r = -.35$). Bezüglich des Alters konnte der Mann-Whitney-U-Test zeigen ($U = 20.00$, $z = -1.868$, $p = .032$, $r = -.42$), dass die älteren Kinder ($MW = 5.11$, $SD = 1.39$) signifikant bessere Leistungen erbringen als die jüngeren Kinder ($MW = 3.48$, $SD = 1.77$). Demnach lässt sich schlussfolgern, dass sich die Leistungen der Kinder mit ansteigendem Alter verbessern ($\tau = .435$, $p = .007$).

10.4. Aufmerksamkeit

Es wurden drei verschiedene Aspekte der Aufmerksamkeit in der Testsituation beobachtet. Protokolliert wurde, wie viele Sekunden lang das Kind eine Aufgabe bis zur ersten Unaufmerksamkeit aufmerksam bearbeiten konnte. Dies wurde bei den Items „Zeichnen“, „Perlen auffädeln“, „Bausteine bauen“, „Puzzle bauen“ und „Buch lesen“ beobachtet und danach in Minuten umgewandelt. Wenn es zu einer Unaufmerksamkeit kam, wurde notiert, ob sich das Kind wieder zur eigentlichen Aufgabe zurückführen ließ. Zusätzlich wurde beim Ansehen des Buches beobachtet, wie viele Seiten das Kind aufmerksam betrachtete.

10.4.1. Aufmerksamkeit in Minuten

Die Kinder waren im Durchschnitt $MW = 17.27$ Minuten ($SD = 7.67$) lang über alle fünf Items aufmerksam. Die Mädchen ($MW = 18.68$, $SD = 7.50$) und die älteren Kinder ($MW = 20.98$, $SD = 8.10$) waren dabei etwas länger als die Buben ($MW = 15.85$, $SD = 7.97$) und die jüngeren Kinder ($MW = 15.27$, $SD = 6.92$) konzentriert. Das Aufmerksamkeitsniveau der Mädchen und der älteren Kinder lag bei mindestens 10.40 und bei maximal 31.58 Minuten, das der Buben und der jüngeren Kinder bei mindestens 3.53 und bei höchstens 24.17 Minuten.

In Tabelle 11 sind alle Werte ersichtlich.

Tabelle 11: Darstellung der Aufmerksamkeit in Minuten getrennt nach Geschlecht und Alter				
in Minuten	Mädchen	Buben	jung	alt
MW	18.68	15.85	15.27	20.98
SD	7.50	7.97	6.92	8.10
Minimum	10.40	3.53	3.53	10.40
Maximum	31.58	24.17	24.17	31.58

Geschlechts- und Altersunterschiede konnten mit den Mann-Whitney-U-Tests nicht festgestellt werden. Sowohl Mädchen und Buben ($U = 43.00$, $z = -.529$, $p = .616$, $r = -.12$), als auch die jüngere und die ältere Altersgruppe ($U = 28.50$, $z = -1.348$, $p = .095$, $r = -.30$) unterscheiden sich nicht signifikant von einander. Die Hypothese, dass je älter die Kinder sind, umso bessere Leistungen werden erreicht, erweist sich als nicht signifikant ($\tau = .181$, $p = .140$).

10.4.2. Aufmerksamkeit – zurückholen

Wurden die Kinder bei den fünf oben erwähnten Aufgaben unaufmerksam, so wurde versucht, sie wieder zum eigentlichen Item zurückzuführen. Gelingt dies, so wurde dies mit 0.5 Punkten kodiert, gelang dies nicht, wurde dieses Verhalten mit 0 Punkten gewertet. War das Kind durchgängig aufmerksam, bekam es 1 Punkt. Diese Punkte wurden über alle fünf Aufgaben addiert. Je höher diese Summe ausfällt, umso mehr Aufgaben haben die Kinder durchgängig aufmerksam bearbeitet.

Insgesamt erreichten sieben Kinder (35%), davon fünf Mädchen und zwei Buben, bzw. vier ältere und drei jüngere Kinder, einen Summenscore von 4.0, welcher bedeutet, dass sie zumindest drei Aufgaben durchgängig aufmerksam bearbeitet haben. Bei zwei Items wurden sie zwar unaufmerksam, jedoch ließen sie sich wieder zurückführen. Vier Kinder (20%), davon ein Mädchen und drei Buben bzw. zwei jüngere und zwei ältere Kinder, erreichten einen Summenscore von 4.5, der bedeutet, dass diese Kinder vier Aufgaben durchgängig aufmerksam bearbeitet haben, bei einer Aufgabe unaufmerksam wurden, sich jedoch wieder zurückführen ließen. Zwei Mädchen aus der älteren Altersgruppe waren bei allen fünf Aufgaben durchwegs aufmerksam.

Daraus lässt sich schlussfolgern, dass 13 Kinder (65%) eine sehr gute Aufmerksamkeit mit einem Summenscore von 4.0 aufwärts zeigen konnten, und dass alle sieben Kinder der älteren Altersgruppe sehr konzentriert waren.

In Tabelle 12 sind alle Häufigkeiten ersichtlich.

Tabelle 12: Darstellung der Häufigkeiten bzgl. der Anzahl der aufmerksam bearbeiteten Items getrennt nach Geschlecht und Alter

	Mädchen	Buben	jung	alt
2.5	0	2	2	0
3.0	0	2	2	0
3.5	2	1	3	0
4.0	5	2	4	3
4.5	1	3	2	2
5.0	2	0	0	2
gesamt	10	10	13	7

Den Ergebnissen des Mann-Whitney-U-Tests zufolge lassen sich keine signifikanten Geschlechtsunterschiede feststellen ($U = 32.50$, $z = -1.361$, $p = .191$, $r = -.30$), jedoch unterscheidet sich die ältere ($MW = 4.43$, $SD = 0.45$) von der jüngeren Altersgruppe (MW

= 3.58, $SD = 0.67$) signifikant ($U = 14.00$, $z = -2.568$, $p = .006$, $r = -.57$). Daraus lässt sich schlussfolgern, dass je älter die Kinder sind, umso bessere Leistungen werden erzielt ($\tau = .469$, $p = .004$).

10.4.3. Aufmerksamkeit – Buchseiten

Hier wurde beobachtet, wie viele Buchseiten sich die Kinder aufmerksam ansahen.

Ein Mädchen betrachtete nur eine Seite des Buches aufmerksam, drei Mädchen zwei Seiten und sechs Mädchen alle fünf Seiten. Von den Buben sah sich keiner alle fünf Seiten aufmerksam an, und nur einer betrachtete vier Seiten konzentriert. Drei Seiten sahen sich zwei Buben aufmerksam an, zwei Seiten drei Buben, und vier Burschen zeigten kaum Interesse beim Buchansetzen und sahen sich somit keine einzige Seite aufmerksam an. Bezüglich des Alters lassen sich aus der Häufigkeitstabelle keine deutlichen Unterschiede erkennen, sodass sowohl ältere als auch jüngere Kinder sich keine einzige Seite bzw. alle fünf Seiten aufmerksam angesehen haben.

In Tabelle 13 sind alle Werte aufgelistet.

Tabelle 13: Häufigkeitsverteilung getrennt nach Geschlecht und Alter bzgl. der Anzahl der angesehenen Buchseiten

	Mädchen	Buben	jung	alt
0 Seiten	0	4	3	1
1 Seite	1	0	1	0
2 Seiten	3	3	4	2
3 Seiten	0	2	1	1
4 Seiten	0	1	1	0
5 Seiten	6	0	3	3
gesamt	10	10	13	7

Dem Mann-Whitney-U-Test zufolge unterscheiden sich die Mädchen ($MW = 0.74$, $SD = 0.34$) von den Buben ($MW = 0.32$, $SD = 0.30$) signifikant in der Anzahl der angesehenen Buchseiten ($U = 19.50$, $z = -2.379$, $p = .015$, $r = -.53$). Die ältere Altersgruppe unterscheidet sich von der jüngeren laut Mann-Whitney-U-Test jedoch nicht signifikant ($U = 34.50$, $z = -.899$, $p = .207$, $r = -.20$). Auch der Zusammenhang zwischen Alter und Testleistung zeigt sich als nicht signifikant ($\tau = .213$, $p = .117$).

10.5. Sozioemotionale Entwicklung

Der Fähigkeitsbereich der sozioemotionalen Entwicklung setzt sich zusammen aus der Selbstständigkeit, die das Kind beim Anziehen eines Kostüms zeigt, der Lösung von der Bezugsperson zu Beginn der Testung und des Kontaktes zur Bezugsperson während des Spieles.

10.5.1. Selbstständigkeit

Für die Skala der Selbstständigkeit, bestehend aus elf Items, konnte ein Cronbach Alpha von .898 berechnet werden, das sehr zufriedenstellend ausfällt. Die Schwierigkeiten von neun Items liegen im adäquaten Bereich. Das Item „Hut aufsetzen“ ist der Stichprobe sehr leicht gefallen, das Item „in Ärmel schlüpfen“ sehr schwer. Die Trennschärfen liegen im mittelmäßigen, zum Großteil jedoch im hohen Bereich. Daher kommt es zu keiner Selektion der Items. In diese Auswertung fließen die Werte von allen 20 Kindern ein.

In Tabelle 14 sind alle Ergebnisse ersichtlich.

Tabelle 14: Ergebnisse der Itemanalyse zur Selbstständigkeit

N = 20	Schwierigkeit	Trennschärfe
Schuhe anziehen: versucht	.75	.63
Hut aufsetzen: versucht	.85	.38
Hose anziehen: versucht	.50	.69
T-Shirt anziehen: versucht	.35	.74
Schuhe anziehen: geschafft	.70	.58
Hosenbein anziehen: geschafft	.33	.74
Hose hinauf ziehen: geschafft	.20	.77
Hut aufsetzen: geschafft	.80	.49
in Ärmel schlüpfen: geschafft	.10	.50
mit Kopf in T-Shirt schlüpfen: geschafft	.20	.65
T-Shirt hinunter ziehen: geschafft	.20	.77

Mittels Mann-Whitney-U-Tests konnten weder signifikante Unterschiede zwischen Mädchen und Buben ($U = 24.50$, $z = -1.959$, $p = .052$, $r = -.44$) noch zwischen den jüngeren und den älteren Kindern festgestellt werden ($U = 32.00$, $z = -1.087$, $p = .146$, $r = -.24$). Die Hypothese, dass je älter die Kinder sind, umso bessere Leistungen erzielen sie, kann nicht bestätigt werden ($\tau = .099$, $p = .285$).

10.5.2. Lösung und Kontakt zur Bezugsperson

Während der Testsituation wurden von der Testleiterin und der Beobachterin zwei Items zur sozioemotionalen Entwicklung beobachtet und auf einer fünfstufigen Ratingskala eingeschätzt.

Lösung von der Bezugsperson

Ein Kind (5%) konnte sich vor der Testsituation sehr lange nicht von der Mutter lösen und bei einem weiteren Kind (5%) musste die Bezugsperson eine Zeit lang mitspielen. Lediglich ein kurzes gemeinsames Explorieren der Spielmaterialien mit der Bezugsperson war bei zwei Kindern (10%) nötig. 11 Kinder (55%) verabschiedeten sich von der Bezugsperson und beschäftigen sich mit den Spielmaterialien. Weitere fünf Kinder (25%) suchten keinen Kontakt. Somit lässt sich schlussfolgern, dass sich die Kinder sehr schnell auf die unbekannte Situation im Testraum einließen und mit der Testleiterin interagierten, was bei 16 Kindern (80%) der Fall war.

In Tabelle 15 sind alle Häufigkeiten ersichtlich.

Tabelle 15: Häufigkeitsverteilung im Bezug auf die Lösung des Kindes von der Bezugsperson getrennt nach Geschlecht und Alter

Lösung	Mädchen	Buben	jung	alt
Körperkontakt	0	1	1	0
mitspielen	1	0	0	1
Exploration	2	0	1	1
Verabschiedung	4	7	7	4
kein Kontakt	3	2	4	1
gesamt	10	10	13	7

Laut Mann-Whitney-U-Test unterscheiden sich weder Mädchen von Buben ($U = 48.00$, $z = -.167$, $p = 1.00$, $r = -.04$), noch die jüngere von der älteren Altersgruppe ($U = 35.50$, $z = -.876$, $p = .247$, $r = -.20$) signifikant. Der Zusammenhang zwischen Alter und Testleistung ist nicht signifikant ($\tau = -.230$, $p = .108$).

Ein Kind (5%) suchte während der Testung keinen **Kontakt zur Bezugsperson** und ein weiteres Kind (5%) versicherte sich lediglich über Blickkontakt, dass die Bezugsperson noch im Raum ist. Acht Kinder (40%) gingen zur Bezugsperson, wollten Lob für ihre Handlungen, wobei jedoch kein gemeinsames Spiel entstand. 10 Kinder (50%) sprachen die Bezugsperson direkt an und wollten sie in ihr Spiel mit einbinden. Von keinem Kind wurde ständiger Körperkontakt während der Testung aufgesucht. Hier ist deutlich zu sehen, dass 18 Kinder (90%) ein sehr starkes Bedürfnis nach Kontakt zur Bezugsperson während der Spielsituation hatten.

In Tabelle 16 sind alle Häufigkeiten dargestellt.

Tabelle 16: Häufigkeitsverteilung in Bezug auf den Kontakt zur Bezugsperson getrennt nach Geschlecht und Alter

Kontakt	Mädchen	Buben	jung	alt
Körperkontakt	0	0	0	0
gem. Spiel	5	5	6	4
kein gem. Spiel	3	5	6	2
Blickkontakt	1	0	1	0
kein Kontakt	1	0	0	1
gesamt	10	10	13	7

Auch hier ließen sich laut Mann-Whitney-U-Tests keine signifikanten Geschlechts- ($U = 45.00$, $z = -.419$, $p = .745$, $r = -.09$) und Altersunterschiede ($U = 43.00$, $z = -.220$, $p = -.478$, $r = -.05$) feststellen. Der Zusammenhang zwischen Alter und Testleistung erweist sich auch hier als nicht signifikant ($\tau = -.153$, $p = .211$).

10.6. Zusammenfassung der Itemanalyse

Die Itemanalyse gibt Antwort auf die erste Fragestellung, welche Items zu wenig Aussagekraft haben, um im Itempool zu bleiben und aus diesem Grund für weitere Analysen ausgeschlossen werden.

Die Tabelle 17 enthält alle Items, für die eine Trennschärfe von $< .3$ berechnet wurde.

Tabelle 17: Auflistung aller nicht trennscharfen Items aus der Itemanalyse

Fähigkeitsbereich	Item	Trennschärfe
Visuelle Wahrnehmung: aktive Farbdifferenzierung	rot	.13
	von Stiege hüpfen	.25
Grobmotorik	Zehengang	.22
	beidbeinig hüpfen	.14
	stehen bleiben	.01
	Ball fangen	.00
	Ball werfen	-
	Fußkick	-.01
	Kochlöffel aufhängen	.25
Feinmotorik	Glas öffnen	.13
	aufkehren	.19
Visumotorik	auffädeln vorgezeigt	-.19
	umfüllen	.24
	daneben schütten	-.31

Diese oben aufgelisteten 14 Items erreichten aufgrund der zu hohen bzw. zu niedrigen Schwierigkeitsindizes zu geringe Werte in der Trennschärfe. Da sie aufgrund dessen zu wenig Aussagekraft haben, um die jeweiligen Skalen vorherzusagen, wurde sie ausgeschlossen und bei den weiteren Analysen vernachlässigt.

10.7. Reliabilität

Obwohl die Kennwerte zu den internen Konsistenzen mittels dem Cronbach Alpha bereits für jede Skala angeführt wurden, sollen sie nun noch einmal übersichtlich in der Tabelle 18 aufgelistet werden. Fett gedruckt sind all jene Werte, die auf eine zufriedenstellende Reliabilität hinweisen. Zusätzlich wird mit „n“ angegeben, wie viele Items jede Skala beinhaltet.

Tabelle 18: Auflistung der internen Konsistenzen inklusive der Anzahl der Items pro Skala

Fähigkeitsbereich	Skala	Cronbach α	n (Anzahl der Items)
Visuelle Wahrnehmung ($\alpha = .901$)	Aktive Farbdifferenzierung	.870	6
	Passive Farbdifferenzierung	.766	7
	Farbstabilität	.857	4
	Form- und Größendifferenzierung	.735	4
	Grobmotorik	.641	11
Motorik ($\alpha = .814$)	Fein- und Visumotorik	.860	7
	Feinmotorik	.864	14
	Visumotorik	.796	11
Sozioemotionale Entwicklung	Selbstständigkeit	.898	11

Das Gütekriterium der Reliabilität kann für alle hier aufgelisteten Skalen, bis auf den Bereich der Grobmotorik, als gegeben angesehen werden, da das Cronbach α immer einen Wert von über .70 erreicht und somit zufriedenstellend ausfällt. Über eine hohe Reliabilität ($\alpha > .90$) verfügt jedoch keine Skala.

10.8. Objektivität

Nun soll anhand der Erfahrungen und der Beobachtungen aus den Testsituationen beurteilt werden, ob das Gütekriterium der Objektivität als gegeben angenommen werden kann.

10.8.1. Durchführungsobjektivität

Das Manual zum spielbasierten Itempool von Kuchler, Sapper, Deimann und Kastner-Koller (2011) bzw. die überarbeitete Version von Punz und Kronberger (2013) geben Aufschluss darüber, mit welchen Materialien der Testraum ausgestattet sein soll, welche Instruktionen den Kindern zu den jeweiligen Items gegeben und wie diese Aufgaben

angeleitet und durchgeführt werden sollen. Diese Rahmenbedingungen sprechen für die Annahme der Durchführungsobjektivität.

Da die soziale Interaktion bei zweijährigen Kindern nicht, wie bei Lienert und Raatz (1994) empfohlen (siehe Kapitel 2 – Objektivität), auf ein unumgängliches Minimum reduziert werden kann, sondern der Testleiter bzw. die Testleiterin flexibel auf die individuellen Bedürfnisse des Kindes eingehen muss, kann die Durchführungsobjektivität nur begrenzt angenommen werden. Aufgrund der Berücksichtigung der Bedürfnisse und der Motivation des Kindes, wird die Reihenfolge der Items individuell dem Kind angepasst, lediglich die oben genannten Rahmenbedingungen sprechen für ein standardisiertes Vorgehen.

10.8.2. Auswertungsobjektivität

Die Auswertungsobjektivität kann als gegeben angesehen werden, da im Manual zum spielbasierten Itempool zur Erfassung der Entwicklung Zweijähriger von Kuchler, Sapper, Deimann und Kastner-Koller (2011), bzw. in der Überarbeitung von Punz und Kronberger (2013) genaue Anweisungen gegeben werden, wie die Reaktionen und die Verhaltensweisen der Kinder zu verrechnen sind. Da im Manual zusätzlich Beispiele für das Verhalten der Kinder in den jeweiligen Verrechnungskategorien beschrieben werden, und das Antwortformat keine offene, sondern nur geschlossene Kategorien enthält, sollten alle Testleiter und Testleiterinnen zu denselben Ergebnissen kommen.

Zusätzlich dient die Videoaufzeichnung der Testsituation dafür, Unsicherheiten bezüglich der Verrechnung entgegenzuwirken, um im gegebenen Fall die fragliche Situation noch einmal anzusehen.

10.8.3. Interpretationsobjektivität

Da dieses zur Anwendung gekommene, spielbasierte Verfahren aktuell noch über keine Normen verfügt, ist es demnach nicht möglich, die Testwerte einer Person mit den Normwerten einer repräsentativen Stichprobe zu vergleichen. Da es jedoch das Ziel ist, dieses Verfahren zu publizieren, werden in Zukunft Normierungen gemacht werden, um dem Entstehen individueller Deutungen und Interpretationen entgegen zu wirken.

10.9. Validität

10.9.1. Inhaltliche Validität

Bei der Erstellung des spielbasierten Itempools zur Erfassung der Entwicklung Zweijähriger von Kuchler, Sapper, Deimann und Kastner-Koller (2011), wurde das Ziel verfolgt, eine inhaltliche Gültigkeit der Items zu erreichen. Dies bedeutet, wie schon in Kapitel 2 – Validität beschrieben, dass der Inhalt der Items das Zielmerkmal hinreichend genau definiert (Fisseni, 2004) und der Test das optimale Kriterium des interessierenden Merkmals darstellt (Lienert & Raatz, 1994; Kubinger, 2009).

Da die Items dieses Itempools aufgrund des Ablaufes der kindlichen Entwicklung und des Alters aufgenommen wurden, wird die inhaltliche Validität angenommen.

10.9.2. Konstruktvalidität

Zusätzlich werden Alterseffekte als ein weiterer Zugang gesehen, das Gütekriterium der Validität, genauer der Konstruktvalidität, zu belegen. Hierfür wurden Interkorrelationen des Alters in Monaten und der Testleistung in den einzelnen Skalen berechnet.

Es zeigte sich für diese Stichprobe, dass die Kinder in den Bereichen der Farbstabilität, der Form- und Größendifferenzierung, der Grobmotorik, der Visumotorik und der Aufmerksamkeit: zurückholen mit steigendem Alter bessere Leistungen erzielten und dass für diese Skalen die Konstruktvalidität als gegeben angesehen werden kann.

In Tabelle 19 sind die Korrelationen (Kendall's tau) mit den entsprechenden Signifikanzwerten ersichtlich. Fett gedruckt sind all jene Werte, die auf einen signifikanten Alterseffekt bzw. auf Konstruktvalidität hinweisen.

Tabelle 19: Interkorrelationen des Alters in Monaten mit der Testleistung

	Kendalls- tau	p
Aktive Farbdifferenzierung	.243	.091
Passive Farbdifferenzierung	.163	.176
Farbstabilität	.595	.001
Form- und Größendifferenzierung	.450	.006
Grobmotorik	.320	.043
Fein- und Visumotorik	.072	.343
Feinmotorik	.028	.434
Visumotorik	.435	.007
Aufmerksamkeit: in Sekunden	.181	.140
Aufmerksamkeit: zurückholen	.469	.004
Aufmerksamkeit: Buchseiten	.213	.117
Selbstständigkeit	.099	.285
Lösung zur Bezugsperson	-.230	.108
Kontakt zur Bezugsperson	-.153	.211

Auch die berechneten Geschlechts- und Altersunterschiede können dem Nachweis der Konstruktvalidität dienen.

Signifikante Altersunterschiede konnten in den Bereichen der Farbstabilität, der Form- und Größendifferenzierung und der Visumotorik gefunden werden, wobei ältere Kinder hier bessere Leistungen erbrachten als die jüngeren Kinder. Auch hat in dieser Stichprobe die ältere Altersgruppe gezeigt, dass sie signifikant mehr Aufgaben durchgängig aufmerksam bearbeiten konnte als die jüngere Altersgruppe.

Signifikante Geschlechtsunterschiede wurden in den Bereichen der Grob-, der Feinmotorik und beim selbstständigen Anziehen von Kleidung (Fein- und Visumotorik) gefunden. Mädchen erzielten hier signifikant höhere Werte als die Buben und zusätzlich sahen sie sich auch signifikant mehr Seiten im Buch aufmerksam an als die Jungen.

In Tabelle 20 sind die Mittelwertsunterschiede und die Signifikanzwerte, getrennt nach Alter und Geschlecht, ersichtlich. Fett gedruckt sind all jene Werte, die auf einen signifikanten Unterschied hinweisen.

Tabelle 20: Mittelwertsunterschiede für das Alter und das Geschlecht, inklusive Signifikanzwerte

	ALTER			GESCHLECHT		
	jung	alt	p	Mädchen	Buben	p
Aktive Farbdifferenzierung	2.167	4.000	.056	3.600	2.000	.105
Passive Farbdifferenzierung	4.077	5.000	.249	5.000	3.800	.338
Farbstabilität	1.462	3.143	.017	2.500	1.600	.308
Form- und Größendifferenzierung	1.077	3.000	.003	2.100	1.400	.271
Grobmotorik	1.341	2.071	.043	2.050	1.094	.049
Fein- und Visumotorik	2.154	3.214	.097	3.700	1.350	.012
Feinmotorik	6.192	7.786	.151	8.250	5.250	.047
Visumotorik	3.479	5.107	.032	4.775	3.306	.126
Aufmerksamkeit: in Sekunden	915.923	1259.143	.095	6.600	3.350	.616
Aufmerksamkeit: zurückholen	3.577	4.429	.006	3.90	3.90	.191
Aufmerksamkeit: Buchseiten	.477	.629	.207	2.80	2.50	.015
Selbstständigkeit	4.462	5.929	.146	1121.40	950.70	.052
Lösung zur Bezugsperson	4.00	3.71	.247	4.150	3.600	1.00
Kontakt zur Bezugsperson	2.62	2.71	.478	.740	.320	.745

10.9.3. Kriteriumsvalidität

Nun wird näher auf den Zusammenhang zwischen den Einschätzungen der Bezugspersonen im Fragebogen zur Selbstständigkeit ihrer Kinder und den Beobachtungen dazu in der Testsituation eingegangen, da dies ein Zugang ist, die Kriteriumsvalidität zu belegen. Näheres über den Zusammenhang des aktiven Wortschatzes und der Grammatik aus dem Fragebogen und der Testsituation ist bei Punz (in Vorbereitung) nachzulesen.

Die Angaben zur Selbstständigkeit im Fragebogen werden anhand einer vierstufigen Ratingskala, mit den Kategorien „nie“, „manchmal“, „oft“ und „immer“, von der Bezugsperson eingeschätzt. Diese Skala beinhaltet 16 Items, wie z.B. „Mein Kind braucht tagsüber Windeln.“, oder „Mein Kind kann sich alleine anziehen.“, wofür eine Reliabilität von .414 berechnet werden konnte, die sehr gering ausfällt.

Der Zusammenhang zwischen den Einschätzungen der Bezugspersonen und der Beobachtungen in den Testsituationen zur Selbstständigkeit fiel nicht signifikant aus, $\tau = -.188$, $p > .05$. Somit kann nicht von einem Zusammenhang bzw. einer Kriteriumsvalidität dieser Skala ausgegangen werden.

DISKUSSION

Das Ziel dieser Diplomarbeit war die Überarbeitung und testtheoretische Überprüfung des von Kuchler et al. 2011 entwickelten spielbasierten Itempools zur Erfassung der Entwicklung zweijähriger Kinder.

Die folgenden Anmerkungen zur Durchführung dieses Verfahrens und dessen statistischen Ergebnissen sollen der Weiterentwicklung dieses Itempools dienen.

Allgemeine Anmerkungen zum spielbasierten Itempool von Kuchler et al. (2011)

Die Puppenküche als Setting dieses Verfahrens besitzt einen sehr großen Aufforderungscharakter und hat sich für alle Kinder als sehr interessant erwiesen. Die erste Kontaktaufnahme mit dem Kind fiel sehr leicht, da es sich meist sehr schnell in das Spiel mit der Puppenküche einbinden ließ. Sehr vorteilhaft ist, dass die Items, die in der Puppenküche vorgesehen sind, keine sprachlichen Reaktionen von Seiten des Kindes verlangen und sich somit als Einstieg in die Untersuchungssituation eignen (Esser & Petermann, 2010). Zu empfehlen ist allerdings, das Angebot an Töpfen, Besteck, etc. zu reduzieren, da manche Kinder alle Kochutensilien verwenden wollen und somit nur schwer für die anderen vorbereiteten Aufgaben zu begeistern sind. Auch Preston (2005), Quaiser-Pohl (2010) und Irblich und Renner (2009) raten, wie in Kapitel 5 (Wie untersucht man Kinder?) beschrieben, zu einem überschaubaren Angebot altersentsprechender Spielmöglichkeiten.

Da die Videoaufzeichnung der gesamten Spielsituation ein wichtiger Faktor ist, um laut Kastner-Koller und Deimann (2009) die Objektivität der Auswertung (s. Kap. 2.1. Objektivität) zu gewährleisten und diese auch zu erleichtern, ist es empfehlenswert, das zuerst mündliche Einverständnis für die Videoaufnahme sofort nach Betreten des Testraumes einzuholen, damit sogleich damit begonnen werden kann. Zudem wird auch das Protokollieren der Anlaufzeit, der Gesamtzeit, etc., das im Nachhinein stattfindet, vereinfacht.

Quaiser-Pohl (2010) und Esser und Petermann (2010) empfehlen eine maximale Untersuchungsdauer von 45 bis 50 Minuten (s. Kap. 5.4.2 Untersuchungsdauer). Obwohl die durchschnittliche Gesamttestdauer bei $MW = 120.39$ Minuten ($SD = 16.747$) lag, die kürzeste Testung 96 Minuten und die längste 160 Minuten dauerte, erwies sie sich als

angemessen. Der Grund für diese lange Durchführungsdauer liegt zum Großteil darin, dass versucht würde, das Optimum an Leistung aus den Kindern herauszuholen und somit die Items mehrmals vorgegeben wurden, wenn das Kind einmal eine falsche Antwort gab. Natürlich benötigte es ein hohes Maß an Flexibilität und Erfahrung mit der Altersgruppe der zweijährigen Kinder sowie in der Durchführung der Spielsituation, da die Arbeitshaltungen bei kleinen Kindern noch nicht ausgebildet sind (Kastner-Koller & Deimann, 2012). Auch bei den Testungen für diese Diplomarbeit wurde deutlich, dass eine Untersuchung im Kleinkindalter höhere Anforderungen an die Testleiterin stellt als die Leistungsdiagnostik im Schul- und Erwachsenenalter (Esser & Petermann, 2010).

Eine große Hilfe stellte jedoch die Anwesenheit von zumindest einer Bezugsperson dar. Diese wurde zu Beginn der Untersuchung angeleitet, unbeantwortete Frage zu wiederholen, um eine verwertbare Antwort zu bekommen. Auch wurde sie darauf hingewiesen, dass sie zu jeder Zeit mitspielen dürfe, wenn es das Kind möchte, sie jedoch keine Gegenstände und Farben benennen solle. Waren die Kinder in das Spiel mit der Testleiterin vertieft, so hielt sich die Bezugsperson meist zurück, wie von Quaiser-Pohl (2010) empfohlen (s. Kap. 5.4.3. Anwesenheit der Bezugsperson(en) während der Untersuchung).

Die Anwesenheit der Beobachterin war bei den ersten Testungen eine große Erleichterung, da diese hinweisen konnte, wenn sich z.B. bestimmte Items aufgrund des Verhaltens der Kinder anboten. Aufgrund der zunehmenden Erfahrung im Umgang mit den zweijährigen Kindern und mit dem Untersuchungsmaterial, kann auf die Anwesenheit eines Beobachters bzw. einer Beobachterin aus ökonomischen Gründen verzichtet werden, sobald sich der Testleiter bzw. die Testleiterin in der Durchführung dieses Verfahrens sicher fühlt und eine Videoaufzeichnung des ganzen Raumes möglich ist.

Diskussion der Durchführung und der Ergebnisse des Funktionsbereiches „Visuelle Wahrnehmung“

Bei der Durchführung der Skalen der **aktiven und passiven Farbdifferenzierung** wurden keine Änderungen vorgenommen, da sie sich bei den Vorgängerinnen (Kuchler, 2011; Sapper, 2011; Birngruber, 2012; Fuchs-Garderer, 2012; Putzer, 2013; Sindelar, 2013) als leicht durchführbar erwiesen hatte. Die Erhebung findet während dem Spielen in der Puppenküche und bei anderen Aufgaben statt. Das Kind soll sieben Farben aktiv

benennen, indem es auf die Frage „Welche Farbe ist das?“ antwortet, und passiv differenzieren, indem es der Frage „Kannst du mir bitte den grünen Stift geben?“ zufolge richtig handelt. Aufgrund der vorgegebenen Fragen und der eindeutig richtig zu kodierenden Antworten kann die Durchführungs- und Auswertungsobjektivität als gegeben angesehen werden.

Für den Bereich der **aktiven Farbdifferenzierung** konnten Schwierigkeiten im angemessenen Bereich errechnet werden (.32 - .58). Auch die Trennschärfen zeigten sich für sechs der sieben Items im angemessenen bis hohen Bereich (.47 - .82). Das Item „rot“ wurde aufgrund der niedrigen Itemtrennschärfe von .13 für die weiteren Analysen ausgeschlossen. Zu beachten ist allerdings, dass die hier getestete, kleine Stichprobe von 20 Kindern keine Rückschlüsse auf die Allgemeinheit zulässt und somit dieses Item auf jeden Fall weiterhin im Itempool behalten werden soll. Das Gütekriterium der Reliabilität fällt mit einem Cronbach Alpha von .826 als sehr zufriedenstellend aus und kann somit als gegeben angesehen werden. Geschlechts- ($p = .105$) und Altersunterschiede ($p = .056$) konnten nicht festgestellt werden, wobei sich eine Tendenz beim Alter erkennen lässt. Das nicht Vorhandensein von Alterseffekten ($p = .176$) deutet darauf hin, dass die Konstruktvalidität nicht nachgewiesen werden kann.

Auch bei der **passiven Farbdifferenzierung** konnten Schwierigkeiten (.42 - .89) und Trennschärfen (.42 - .62) im adäquaten Bereich berechnet werden und es kam daher zu keiner Selektion der Items. Die Reliabilität deutet mit einem Cronbach Alpha von .766 auf eine genau messende Skala hin. Die Analyse auf Geschlechts- ($p = .338$) und Altersunterschiede ($p = .249$) fiel nicht signifikant aus. Auch von einem Alterstrend kann nicht ausgegangen werden ($p = .176$) und somit auch nicht von einer Konstruktvalidität.

Da sich bei den Vorgängerdiplomarbeiten herausgestellt hatte, dass die Vorgabe zur **Erfassung der Farbstabilität** zu lange und die Durchführung für die Kinder langweilig war, wurde im Zuge dieser Diplomarbeit versucht, diese Aufgabe interessanter zu gestalten. Hierfür wurden vier Stofftiere mit unterschiedlich farbigen Maschen mit einbezogen. Die Kinder wurden gebeten, der Testleiterin beim Füttern der Tiere behilflich zu sein, da die Tiere schon sehr großen Hunger hätten. Jedes dieser Tiere würde nur jenes Futter fressen, das dieselbe Farbe wie deren Masche habe.

Auch mit Zuhilfenahme der Stofftiere war die Durchführung teilweise schwierig, da das Füttern der Tiere den Kindern zu lange dauerte. Zu empfehlen ist, in die Körbchen der Tiere den Maschen farblich gleiche Servietten zu legen, um die Zugehörigkeit des Spielobjektes deutlicher zu machen. Zu erwähnen ist jedoch, dass sich die Kuscheltiere neben der Puppenküche sehr gut zur ersten Kontaktaufnahme mit den Kindern eignen.

Schon bei Sindelar (2013) zeigten sich die Ergebnisse zu dieser Skala als sehr zufriedenstellend, welche sich auch bei dieser Stichprobe wiederfinden lassen. Die Schwierigkeiten liegen im adäquaten Bereich (.40 bis .55) und die Trennschärfen lassen sich als hoch einstufen (.67-.77). Pitchford und Mullen (2001) und Kowalski und Zimiles (2006) gehen davon aus, dass das Sortieren von Objekten nach Farben erst mit ungefähr drei Jahren möglich ist. Diese Stichprobe zeigt, dass auch jüngere Kinder, wie auch bei Sandhofer und Smith (1999), diese Fähigkeit bereits erworben haben (s. Kap. 4.4. Kategorische Farbwahrnehmung). Die Reliabilität liegt bei $\alpha = .857$ und zeigt somit, dass die Skala genau misst und das Gütekriterium als gegeben anzusehen ist.

Während Mädchen und Buben sich in der Leistung nicht signifikant unterscheiden ($p = .308$), zeigen die Ergebnisse, dass das Farbkonzept umso stabiler ist, je älter die Kinder sind ($p = .017$). Dieser Alterseffekt weist auf das Vorhandensein der Konstruktvalidität hin. Auch hier lässt sich die Durchführungs- und Auswertungsobjektivität als gegeben ansehen, da die Instruktion, die Vorgabe und die Kodierung eindeutig festgelegt sind und keine Variabilität zulassen.

Auch bei der Skala der **Form- und Größendifferenzierung** wurde eine Änderung in der Vorgabe vorgenommen, da auch sie sich bei den Vorgängerinnen als schwierig erwiesen hatte.

Auffällig ist, dass diese Skala die einzige der im Zuge dieser Diplomarbeit bearbeiteten Funktionsbereiche ist, bei der mehrere fehlende Werte zu verzeichnen sind. Verwunderlich ist dies jedoch nicht, da Prevor und Diamond (2005) der Ansicht sind, dass Kinder erst nach dem dritten Geburtstag das Interesse für die Form eines Gegenstandes entwickeln. Laut Pitchford und Mullen (2001) entwickelt sie sich sogar erst mit etwa vier Jahren (s. Kap. 4.5. Formwahrnehmung). Bei drei Kindern konnte diese Aufgabe aufgrund fortschreitender Müdigkeit nicht mehr durchgeführt werden, wobei erwähnt werden muss, dass diese Aufgabe meist erst am Ende der Testung vorgegeben wurde. Die fehlenden Werte wurden auf 0 gesetzt und somit als falsch gelöst gewertet.

Dies kann zu einer Überschätzung der internen Konsistenz führen. Die Reliabilität zeigt ein α von .735, das im zufriedenstellenden Bereich liegt. Auch die Itemschwierigkeiten (.25 bis .65) und die Trennschärfen (.34 bis .62) fallen zufriedenstellend aus. Hier konnten zwar keine signifikanten Geschlechtsunterschiede ($p = .271$) festgestellt werden, jedoch zeigen die Auswertungen, dass die Fähigkeit der Form- und Größendifferenzierung mit steigendem Alter zunimmt ($p = .006$). Auch hier kann die Konstruktvalidität, die Durchführungs- und Auswertungsobjektivität als gegeben angesehen werden.

Diskussion der Durchführung und der Ergebnisse des Funktionsbereiches „Motorik“

Dieser Funktionsbereich besteht aus vier Skalen, der Grobmotorik, der Fein-/Visumotorik, der Feinmotorik und der Visumotorik.

Der Bereich „**Grobmotorik**“ wird anhand von elf Items erfasst. Diese werden vom Testleiter bzw. der Testleiterin vorgezeigt und das Kind hat die Aufgabe, diese dementsprechend nachzumachen bzw. nachzuturnen. Auch hier finden sich sehr leichte Items und sehr schwierige Items (.11 – 1.00). Da das Item „Ball fangen“ eine Trennschärfe von nur .003 erreicht und daher selektiert wurde, ist zu empfehlen, bei den folgenden Testungen einen etwas größeren Schaumstoffball zu verwenden, damit den Kindern das Fangen etwas erleichtert wird. Im Gegensatz dazu lösten alle Kinder das Item „Ball werfen“. Beim „Fußkick“ könnte man die Kinder in ein Tor, z.B. beim Kaufmannsladen, oder unter den Tisch, schießen lassen, um dies interessanter zu gestalten.

Die Aufgaben „balancieren“, „Zehenspitzenengang“ und „beidbeinig hüpfen“ haben gut funktioniert, wobei etwas Kreativität bei der Vorgabe gefragt war, da dies die Kinder nicht immer sofort machen wollten. Häufig war es hilfreich, wenn diese Aufgaben mit der Mama gemeinsam durchgeführt werden konnten. Obwohl der Klebestreifen, auf dem balanciert und seitwärts gegangen werden sollte, statt den im Manual vorgeschriebenen 8cm nur 5cm breit war, erreichten diese beiden Items dennoch eine Trennschärfe $> .30$. Trotzdem ist zu empfehlen, den Klebestreifen wieder auf 8cm zu verbreitern.

Auch die Durchführung des Items „stehen bleiben“ bereitete so manche Probleme, da der Kleinkindertestraum nicht allzu groß war und die Kinder meist zur Bezugsperson liefen. Sagte die Testleiterin „Stopp!“ so schien es, als dass die Kinder nicht verstanden haben,

worum es geht und der Bezugsperson in die Arme liefen. Obwohl die Testleiterin und die Beobachterin jedem Kind vorzeigten, worum es ging, wurde dieses Item nur sehr selten gelöst. Fraglich ist, ob die Kinder mit den Worten „Stopp!“ oder „Halt!“ vertraut sind und die Bedeutung kennen.

Beim Hinauf- bzw. Hinuntergehen der Stiege sollte differenzierter erfasst werden, wie die Kinder dies bewerkstelligen. Manche Kinder sind mit Händen und Füßen, also auf allen Vieren, hinauf gekrabbelt, manche Kinder haben sich beim Testleiter angehalten. Derzeit wird nur protokolliert, dass sich das Kind angehalten hat. Es macht aber einen Unterschied, wo es sich anhält.

Nach der Selektion von sechs Items erhöhte sich die Reliabilität zwar von $\alpha = .490$ auf $\alpha = .641$, dennoch ist dieses Gütekriterium als nicht gegeben anzusehen.

Die Skalenanalysen zeigten, dass Mädchen in der Grobmotorik bessere Leistungen erzielten als die Buben ($p = .049$) und dass die Testleistung signifikant mit dem Alter der Kinder korreliert ($p = .043$). Aufgrund dieses Alterseffekts kann von der Konstruktvalidität ausgegangen werden. Zusätzlich lässt sich auch die Durchführungs- und Auswertungsobjektivität als gegeben einschätzen.

Die Skala ***Fein- und Visumotorik*** besteht aus sieben Anzieh-Items, wobei erfasst wird, welche Teilschritte die Kinder beim Anziehen von Kleidung selbstständig schaffen. Hierfür konnten Schwierigkeiten im Bereich von .10 bis .80 und Trennschärfen von .32 bis .85 berechnet werden, weshalb keine Selektion der Items vorgenommen wurde. Die Reliabilität lässt sich mit einem Cronbach Alpha von .898 als gegeben ansehen. Mittels der Skalenanalyse konnte festgestellt werden, dass Mädchen signifikant bessere Leistungen beim Anziehen erbrachten als Buben ($p = .012$). Ein Alterseffekt, der als Nachweis der Konstruktvalidität dient, konnte nicht festgestellt werden ($p = .343$). Die Durchführungs- und Auswertungsobjektivität ist gegeben.

Die Skala „***Feinmotorik***“ besteht aus 14 Items, die einerseits in der Puppenküche vorgegeben, beim Zeichnen beobachtet und beim Anziehen von Kleidung durchgeführt werden. Die Vorgabe bereitete keine Schwierigkeiten, da die Kinder sehr interessiert in der Puppenküche spielten und somit gerne mit den Utensilien hantierten. Die Anziehitens erforderten etwas Verhandlungsgeschick von Seiten der Testleiterin (s. Diskussion der

Durchführung und der Ergebnisse des Funktionsbereiches „Sozioemotionale Entwicklung“). Die Skala enthält Items mit Schwierigkeiten zwischen .10 und .90 und Trennschärfen zwischen .13 und .75. Selektiert werden mussten zwei Items, aufgrund dessen sich das Cronbach Alpha von .850 auf .864 erhöhte und die Reliabilität somit als gegeben anzusehen ist. Mädchen erzielten signifikant bessere Leistungen als Buben ($p = .047$), aber es konnte kein Alterstrend festgestellt werden ($p = .434$), weshalb die Konstruktvalidität als nicht gegeben anzusehen ist. Die Durchführungs- und Auswertungsobjektivität sind gegeben.

Die Skala „*Visumotorik*“ besteht aus elf Items. Diese weisen Schwierigkeiten im Bereich von .45 - .84 und Trennschärfen im Bereich von -.31 - .62 auf. Aufgrund letzterer Werte kam es zu einer Selektion von vier Items, wodurch sich das Cronbach Alpha von .572 auf .796 erhöht. Dieser Wert liegt nun im angemessenen Bereich und deutet auf das Gütekriterium der Reliabilität hin. Mädchen und Buben zeigten gleich gute Leistungen ($p = .126$) und es konnte ein Alterstrend festgestellt werden ($p = .007$), der die Konstruktvalidität belegt. Zusätzlich lassen sich die inhaltliche Validität, die Durchführungs- und die Auswertungsobjektivität als gegeben ansehen.

Für die Durchführung der Items zum Zeichnen werden die laminierten Kärtchen mit dem vorgezeichneten Kreis, vertikalen und horizontalen Strich nicht benötigt. Mit den Kindern kann gemeinsam ein Wurm gezeichnet werden, der viele Kugeln benötigt, um groß zu werden. Die Kinder sollen animiert werden, dem Wurm eine Kugel zu zeichnen, damit er länger wird. Danach kann eine Wiese (horizontaler Strich) mit den Grashalmen (vertikale Striche) gezeichnet werden. Diese Vorgabe wurde als spielerischer empfunden und hat sich als sehr effektiv erwiesen. Die Schere soll den Kindern erst nach dem Beenden der Aufgaben zum Nachzeichnen und zur aktiven und passiven Farbdifferenzierung gegeben werden, da sie meist sehr lange in Gebrauch ist. Falls sich die Kinder nicht mehr davon ablenken lassen, ist es sinnvoll, dass der Beobachter bzw. die Beobachterin das Kind um die Schere bittet, da er bzw. sie sie auch benötigt.

Das Spielen mit den Bausteinen war für einige Kinder nicht sehr interessant. Schwierig war die Instruktion, da im Manual vorgegeben wurde, dass nicht erwähnt werden darf, was das Kind bauen könnte, wie z.B. „Kannst du mir einen Turm bauen? Oder ein Haus?“

Dem Kind wird das Umfüllen der Bohnen von einem in den anderen Becher, und wieder zurück vorgeführt. Danach soll das Kind selbst versuchen, die Bohnen jeweils mit der linken und der rechten Hand, ohne daneben zu schütten, umzufüllen. Diese Aufgabe eignet sich besonders als Überleitung für die Aufgabe des Aufkehrens. Denn entweder verschüttet das Kind beim Umfüllen einige Bohnen oder der Testleiter bzw. die Testleiterin ist ungeschickt, sodass einige Bohnen zu Boden fallen.

Das Perlenschnüraufhängen wurde von den meisten Kindern sehr bemüht durchgeführt. Falls sich ein Kind nicht motivieren lässt, hat es sich als effektiv erwiesen, ihm ein Geheimnis ins Ohr zu flüstern, nämlich dass das Kind für die Mama eine Kette aufhängen könnte. Hier hat der Testleiter bzw. die Testleiterin alle Perlen in beiden Händen und das Kind wird gefragt, ob es die blaue, weiße und lilafarbene Perle aufhängen kann. Dafür werden die Perlen in den Händen versteckt, damit das Kind noch keine Perle nehmen kann. Erst nachdem die Frage nach einer bestimmten Perle gestellt wurde, werden dem Kind alle Perlen gezeigt. Nach dem Aufhängen aller Perlen lässt man das Kind zählen, wie viele es aufgehängt hat und erfasst somit, ob bzw. wie weit das Kind zählen kann und ob eine Eins-zu-Eins-Zuordnung vorhanden ist (siehe Punz, 2013).

Das Puzzle empfand der Großteil der Kinder als sehr interessant und somit waren sie in der Durchführung sehr bemüht.

Diskussion der Durchführung und der Ergebnisse des Funktionsbereiches „Aufmerksamkeit“

Für die Erfassung der Aufmerksamkeit der Kinder werden keine zusätzlichen Spielmaterialien benötigt, weshalb sich diese als sehr einfach zu erheben darstellt. Jedoch ist die Videoaufzeichnung der Spielsituation unerlässlich. Mithilfe dieser wird im Nachhinein evaluiert, wie viele Sekunden lang das Kind insgesamt bei bestimmten Aufgaben aufmerksam war. Hier kann es zu einer unterschiedlichen Anzahl an Sekunden kommen, je nachdem wann angefangen bzw. aufgehört wird zu protokollieren. Allerdings sollte sich diese Differenz in Grenzen halten. Zusätzlich wird erfasst, ob sich das Kind wieder zurückführen ließ, wenn es unaufmerksam geworden war, und wie viele Seiten im Buch es sich konzentriert ansah. Bei den Vorgängerinnen wurden insgesamt drei Bücher vorgegeben, für diese Arbeit wurde jedoch auf die Durchführung von zwei Büchern aufgrund mangelnder statistischer Kennwerte verzichtet.

Die Auswertung für den Funktionsbereich *Aufmerksamkeit in Sekunden* wurde deskriptiv vorgenommen und zeigt, dass die Mädchen und die älteren Kinder etwas länger als die Buben und die jüngeren Kinder aufmerksam waren, jedoch erweist sich dieser Geschlechts- ($p = .616$) und Altersunterschied ($p = .095$) als nicht signifikant. Die Ergebnisse zeigen im Weiteren, dass sich die Mädchen signifikant mehr Buchseiten aufmerksam ansahen ($p = .015$) als die Buben (*Aufmerksamkeit: Buchseiten*) und dass die ältere Altersgruppe mehr Aufgaben durchgängig aufmerksam bearbeiten konnte ($p = .006$) als die jüngere (*Aufmerksamkeit: zurückholen*).

Diskussion der Durchführung und der Ergebnisse des Funktionsbereiches „Sozioemotionale Entwicklung“

Die Erfassung der sozioemotionalen Entwicklung gliedert sich in zwei Teile.

Den einen Bereich bildet die Erhebung der *Selbstständigkeit* des Kindes beim Anziehen von verschiedenen Kleidungsstücken. Hier ist von Interesse, ob das Kind selbstständig versucht, sich zu verkleiden, ob es um Hilfe bittet oder ob es Desinteresse zeigt. Zusätzlich wird erfasst, wie viele Teilschritte die Kinder beim Anziehen alleine schaffen (Items der Skala Fein- und Visumotorik).

Für einen Großteil der Kinder waren die Hose und das T-Shirt des Clownkostüms nicht interessant. Viele Kinder argumentierten, dass sie schon eine Hose bzw. ein T-Shirt an hätten und sie deshalb diese Kleidungsstücke nicht mehr anziehen können bzw. wollen. Manchmal konnten die Kinder überredet werden, indem ihnen erzählt wurde, dass dies eine ganz besondere Kleidung sei, da man diese über die normale Kleidung anziehen könne. Auch war es sehr hilfreich, den Kindern zu sagen und zu zeigen, dass diese Hose für den Testleiter bzw. der Testleiterin viel zu klein ist. Die Kinder wurden animiert, indem sie gefragt wurden: „Kannst du die Hose mal anziehen? Vielleicht passt sie dir!“ Das Anprobieren der Schuhe und des Hutes wurde von den Kindern meist ohne Anforderung gemacht.

Diese Skala enthält elf Items, von denen neun Items Schwierigkeiten (.20 - .80) und Trennschärfen (.49 - .77) im angemessenen Bereich aufweisen. Ein Item war zu schwer (.10), ein Item war der Stichprobe zu leicht (.85). Jedoch kam es zu keiner Selektion von Items. Das Gütekriterium der Reliabilität kann als gegeben angesehen werden, da das Cronbach Alpha einen Wert von .898 erreicht.

Geschlechts- ($p = .052$) und Altersunterschiede ($p = .146$) konnten nicht festgestellt werden, weshalb neben der gültigen inhaltlichen Validität die Konstruktvalidität als nicht gegeben gilt. Die Durchführungs- und Auswertungsobjektivität kann angenommen werden, wobei jedoch genau zu protokollieren ist, wie viel dem Kind beim Anziehen der Kleidung geholfen wurde, d.h. ob z.B. der Arm selbst in den Ärmel fand oder ob dieser etwas gehoben wurde, um es dem Kind zu erleichtern.

Der zweite Teil beruht auf der Verhaltensbeobachtung des Kindes zu Beginn und während der Testung. Hier wird von der Testleiterin bzw. dem Testleiter und der Beobachterin bzw. dem Beobachter getrennt voneinander eine Einschätzung auf einer fünfstufigen Ratingskala vorgenommen. Diese Bewertungen bzgl. der **Lösung des Kindes zu Beginn der Testung** und des **Kontaktes des Kindes während der Testung** werden im Nachhinein verglichen, besprochen und gegebenenfalls angeglichen. Da sich gezeigt hat, dass die Definitionen der fünfstufigen Ratingskala nicht alle gezeigten Verhaltensweisen der Kinder abdecken, wurde von Punz und Kronberger (2013) in der Überarbeitung des Manuals von Kuchler et al. (2011) versucht, diese zu präzisieren, um zu gewährleisten, dass die Auswertungsobjektivität gegeben ist und nicht mehr auf subjektiven Einschätzungen beruht.

Aufgrund der deskriptiven Auswertung zur Skala „Lösung des Kindes von der Bezugsperson“ wird der Eindruck bestätigt, dass die Kinder sehr schnell mit der Testleiterin interagierten. 16 Kinder (80%) konnten sich sehr schnell auf die unbekannte Situation im Testraum einlassen und begannen mit der Testleiterin zu spielen. Hierbei zeigten sich keine Geschlechts- ($p = 1.00$) und Altersunterschiede ($p = .247$).

18 Kinder (90%) hatten ein sehr starkes Bedürfnis nach Kontakt zur Bezugsperson während der Testung. Sie holten Lob für ihre Handlungen ein, sprachen die Bezugsperson direkt an und wollten sie in ihr Spiel mit einbinden. Dies wurde gewährleistet, wobei darauf hingewiesen wurde, dass keine Gegenstände oder Farben benannt werden sollen. Das Miteinbeziehen der Bezugsperson ist vor allem dann zu empfehlen, wenn die Kinder auf die Frage des Testleiters bzw. der Testleiterin nicht antworten bzw. reagieren.

ZUSAMMENFASSUNG

Das Ziel dieser Diplomarbeit ist die Überarbeitung und testtheoretische Überprüfung des spielbasierten Itempools zur Erfassung der Entwicklung zweijähriger Kinder von Kuchler et al. (2011).

Hierfür kam dieser Itempool an insgesamt 20 Kindern (zehn Mädchen und zehn Buben) im Alter zwischen 24 und 36 Monaten zur Anwendung. Für alle Items der Skalen „Visuelle Wahrnehmung“, „Motorik“ und „Selbstständigkeit“ wurden die Schwierigkeiten und Trennschärfen berechnet, um Items mit niedrigen Werten in der Trennschärfe für die weiteren Analysen ausschließen zu können. Zusätzlich wurden die Reliabilitäten mittels Cronbach Alpha α berechnet, das auf alle, bis auf die Skala „Grobmotorik“ ($\alpha = .641$), einen Wert $> .700$ erreicht und somit zufriedenstellend ausfällt.

Die Objektivität kann als gegeben angesehen werden, da die Vorgabe, die Ausführung, die Instruktion und die Auswertung zu den einzelnen Items im Manual von Kuchler et al. (2011) genau beschrieben werden und aufgrund dessen standardisiert durchgeführt werden können. Allerdings verfügt dieses spielbasierte Verfahren derzeit noch über keine Normierungen, weshalb die Interpretationsobjektivität noch nicht als gegeben anzusehen ist.

Bezüglich des Gütekriteriums der Validität ist zu erwähnen, dass die inhaltliche Validität vorhanden ist, da die Items aufgrund des Ablaufes der kindlichen Entwicklung und des Alters in den Itempool aufgenommen wurden.

Um die Konstruktvalidität nachweisen zu können, wurden Korrelationen zwischen dem Alter der Kinder in Monaten und den Testleistungen berechnet. Es konnte festgestellt werden, dass die Kinder in den Bereichen der Farbstabilität, der Form- und Größendifferenzierung, der Grobmotorik und der Visumotorik mit zunehmendem Alter bessere Leistungen erbringen. Auch hat in dieser Stichprobe die ältere Altersgruppe gezeigt, dass sie signifikant mehr Aufgaben durchgängig aufmerksam bearbeiten konnten als die jüngere Altersgruppe.

Im Weiteren wurden Skalenanalysen mittels Mann-Whitney-U-Tests gemacht, um mögliche Geschlechtsunterschiede festzustellen. Die Ergebnisse zeigen, dass die Mädchen in dieser Stichprobe signifikant bessere Leistungen in den Fähigkeitsbereichen der Grobmotorik, der Feinmotorik und im Anziehen von Kleidung (Fein- und

Visumotorik) erzielten und dass sie sich auch signifikant mehr Buchseiten aufmerksam angesehen hatten.

Zusätzlich von Interesse war, ob sich ein Zusammenhang zwischen den Einschätzungen der Bezugsperson zur Selbstständigkeit ihres Kindes mit den Beobachtungen aus der Testsituation finden lässt. Ein solcher konnte jedoch nicht gefunden werden, weshalb auch die Kriteriumsvalidität als nicht gegeben anzusehen ist.

ABSTRACT

The aim of this thesis is the revision and theoretical examination of the play-based itempool for investigating the development of two-year-olds by Kuchler et al. (2011).

This itempool was used on twenty children (ten girls and ten boys) at the age of 24 months up to 36 months. The difficulty indices and the statistical power were calculated for all items of the following scales: „visual perception“, „motor function“ and „independence“ to eliminate bad data from further analysis. In addition the reliability with Cronbach Alpha was calculated, which is satisfactory ($\alpha > .700$) for the most scales apart from the gross motor skills ($\alpha = .641$).

The objectivity can be seen as given since the guidelines, the realization, the instruction and the analysis are not only well described in the manual of Kuchler et al. (2011), but also should be performed standardized. The objectivity of the interpretation cannot be taken for granted because there are no norms available yet.

The content validity is given because of the fact that the items are produced in terms of the developmental process and the age of the children. To find out if the construct validity can be taken for granted correlations of the children's age (in months) and the results were calculated. These results show that children perform better in the concept of colour, the form and size differentiation, the gross motor skills and the hand-eye-coordination when they get older. Furthermore the older group was able to deal with items more attentively than the younger group did. The construct validity is given for those scales, but not for all.

There were also Mann-Whitney-U-Tests used to find out whether there are statistical significant differences between boys and girls. The results show that girls perform significantly better in gross-motor skills, fine-motor skills and in getting dressed on their own. Also girls looked at significant more pages of the book more attentively than boys did.

In addition it was interesting to see if there is a reference between the evaluation of the role model to the independence of the child and the observation during the testing. According to the statistical analysis there was no connection found and because of that the criterion validity cannot be seen as given.

LITERATURVERZEICHNIS

- Adams, R. J. (1987). An evaluation of color preference in early infancy. *Infant Behavior & Development*, 10, 143-150.
- 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.
- Arthus, H. (1949). *Le village: Test d'activité créatrice*. Paris: Hartmann.
- Berk, L. E. (2005). *Entwicklungspsychologie* (3., aktualisierte Aufl.). München: Pearson.
- Birngruber, A. (2012). *Die Überarbeitung und testtheoretische Überprüfung eines spielbasierten Itempools zur Erfassung der Entwicklung im dritten Lebensjahr unter besonderer Berücksichtigung der kognitiven Entwicklung, der visuellen Wahrnehmung, des Gedächtnisses und der Aufmerksamkeit*. Unveröffentlichte Diplomarbeit, Universität, Wien.
- 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.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4., überarbeitete Aufl.). Heidelberg: Springer.
- Brandt, I. & Sticker, E. J. (2001). *Griffiths-Entwicklungsskalen*. Göttingen: Beltz.
- Bühler, C. (1955). *Der Welt-Test*. Göttingen: Hogrefe.
- Cherney, I. D., Kelly-Vance, L., Glover, K. G., Ruane, A. & Ryalls, B. O. (2003). The Effects of Stereotyped Toys and Gender on Play Assessment in Children Aged 18-47 Months. *Educational Psychology*, 23, 95-105.
- Dacheneder, W. (2009). Diagnostik der visuellen Wahrnehmung. In D. Irblich & G. Renner (Hrsg.), *Diagnostik in der Klinischen Kinderpsychologie. Die ersten sieben Lebensjahre* (S. 179-194). Göttingen: Hogrefe.
- Davidoff, J., Goldstein, J. & Roberson, D. (2009). Nature versus nurture: The simple contrast. *Journal of Experimental Child Psychology*, 102, 246-250.
- Deegener, G. (2009). Eltern und Erzieher als Informationsquellen. In D. Irblich & G. Renner (Hrsg.), *Diagnostik in der Klinischen Kinderpsychologie. Die ersten sieben Lebensjahre* (S. 86-96). Göttingen: Hogrefe.

- Deimann, P. & Kastner-Koller, U. (2007). Entwicklungsdiagnostik. In M. Hasselhorn & W. Schneider (Hrsg.), *Handbuch der Entwicklungspsychologie* (S. 558-569). Göttingen: Hogrefe.
- Deimann, P., Kastner-Koller, U., Benka, M., Kainz, S. & Schmidt, H. (2005). Mütter als Entwicklungsdiagnostikerinnen. Der Entwicklungsstand von Kindergartenkindern im Urteil ihrer Mütter. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 37, 122-134.
- Dykeman, B. F. (2008). Play-Based Neuropsychological Assessment of Toddlers. *Journal of Instructional Psychology*, 35, 405-408.
- Esser, G. & Petermann, F. (2010). *Entwicklungsdiagnostik*. Göttingen: Hogrefe.
- Ettrich, K. U. (2000). *Entwicklungsdiagnostik im Vorschulalter: Grundlagen – Verfahren – Neuentwicklungen – Screenings*. Göttingen: Vandenhoeck & Ruprecht.
- Farmer-Dougan, V. & Kaszuba, T. (1999). Reliability and Validity of Play-based Observations: relationship between the PLAY behaviour observation system and standardised measures of cognitive and social skills. *Educational Psychology*, 19, 429-440.
- Fischer, H. (1995). *Entwicklung der visuellen Wahrnehmung*. Weinheim: Beltz.
- Fisseni, H.-J. (2004). *Lehrbuch der psychologischen Diagnostik*. Göttingen: Hogrefe.
- Flehmg, I., Schloon, M., Uhde, J. & Bernuth, H. v. (1973). *Denver Entwicklungsskalen*. Hamburg: Hamburger Spastikerverein.
- Franklin, A. (2006). Constraints on children's color term acquisition. *Journal of Experimental Child Psychology*, 94, 322-327.
- Franklin, A., Bevis, L., Ling, Y. & Hurlbert, A. (2010). Biological components of colour preference in infancy. *Developmental Science*, 13, 346-354.
- Franklin, A., Clifford, A., Williamson, E. & Davies, I. (2005). Color term knowledge does not affect categorical perception of color in toddlers. *Journal of Experimental Child Psychology*, 90, 114-141.
- Franklin, A., Drivonikou, G. V., Clifford, A., Kay, P., Regier, T. & Davies, I. R. L. (2008). Lateralization of categorical perception of color changes with color term acquisition. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 105, 18221-18225.

- Franklin, A., Wright, O. & Davies, I. R. L. (2009). What can we learn from toddlers about categorical perception of color? Comments on Goldstein, Davidoff, and Roberson. *Journal of Experimental Child Psychology*, 102, 239-245.
- Fuchs-Gaderer, M. (2012). *Die Überarbeitung und testtheoretische Überprüfung eines spielbasierten Itempools zur Erfassung der Entwicklung im dritten Lebensjahr unter besonderer Berücksichtigung der Bereiche Motorik, Sprache und sozioemotionale Entwicklung*. Unveröffentlichte Diplomarbeit, Universität, Wien.
- Fuiko, R. & Wurst, E. (2003). Entwicklungsdiagnostik. In K. D. Kubinger & R. S. Jäger (Hrsg.), *Schlüsselbegriffe der Psychologischen Diagnostik* (S. 119-123). Weinheim: Beltz.
- Gilbert, A. L., Regier, T., Kay, P., & Ivry, R. B. (2006). Whorf hypothesis is supported in the right visual field but not the left. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 103, 489–494.
- Goldstein, E. B. (2008). *Wahrnehmungspsychologie. Der Grundkurs*. (7. Aufl.). Berlin: Springer.
- Goswami, U. (2001). *So denken Kinder. Einführung in die Psychologie der kognitiven Entwicklung*. Bern: Huber.
- Hagmann-von Arx, P., Meyer, C. S. & Grob, A. (2008). Intelligenz- und Entwicklungsdiagnostik im deutschen Sprachraum. *Kindheit und Entwicklung*, 17, 232-242.
- Harding, G. (1972). *Spieldiagnostik. Das Spiel als diagnostisches Mittel in der Kinderpsychiatrie*. Weinheim: Beltz.
- Hellbrügge, T. (Hrsg.). (1994). *Münchener Funktionelle Entwicklungsdiagnostik (zweites und drittes Lebensjahr)* (4., korrigierte und erweiterte Aufl.). München: Universität München, Institut für Soziale Pädiatrie und Jugendmedizin.
- Holle, B. (1988). *Die motorische und perzeptuelle Entwicklung des Kindes. Ein praktisches Lehrbuch für die Arbeit mit normalen und retardierten Kindern*. Weinheim: Beltz.
- Hyvärinen, L. (2004). *Das funktionale Sehen in der Frühbetreuung und im Spezialunterricht sehgeschädigter Kinder*. Zugriff am 07.05.2013, <http://www.lea-test.fi/de/sehuber/waldkirc.html>.

- Irblich, D. & Renner, G. (2009). Wie untersucht man Kinder? In D. Irblich & G. Renner (Hrsg.), *Diagnostik in der Klinischen Kinderpsychologie. Die ersten sieben Lebensjahre* (S. 21-32). Göttingen: Hogrefe.
- Johnson, E. G. (1977). The Development of Color Knowledge in Preschool Children. *Child Development*, 48, 308-311.
- Jones, S. S. (2003). Late talkers show no shape bias in a novel name extension task. *Developmental Science*, 6, 477-483.
- Kastner-Koller, U. & Deimann, P. (2009). Beobachtung und Befragung von Kindern. In D. Irblich & G. Renner (Hrsg.), *Diagnostik in der Klinischen Kinderpsychologie. Die ersten sieben Lebensjahre* (S. 97-107). Göttingen: Hogrefe.
- Kastner-Koller, U. & Deimann, P. (2011). Entwicklungstests. In L. F. Hornke, M. Amelang & M. Kersting (Hrsg.), *Leistungs-, Intelligenz- und Verhaltensdiagnostik* (Enzyklopädie der Psychologie, Serie II Psychologische Diagnostik, Bd. 3, S. 275-304). Göttingen: Hogrefe.
- Kastner-Koller, U. & Deimann, P. (2012). *Der Wiener Entwicklungstest* (3., überarbeitete und erweiterte Aufl.). Göttingen: Hogrefe.
- 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.
- Klein, M. (1976). Die psychoanalytische Spieltechnik, ihre Geschichte und Bedeutung. In: G. Biermann (Hrsg.). *Handbuch der Kinderpsychotherapie* (4. Auflage, Bd 1, S. 151-168). München: Reinhardt.
- Kowalski, K. & Zimiles, H. (2006). The relation between children's conceptual functioning with color and color term acquisition. *Journal of Experimental Child Psychology*, 94, 301-321.
- Kubinger, K. D. (2009). *Psychologische Diagnostik. Theorie und Praxis psychologischen Diagnostizierens* (2., überarbeitete und erweiterte Aufl.). Göttingen: Hogrefe.
- Kuchler, M. (2011). *Die Entwicklung und Erprobung eines Itempools zur spielbasierten Erfassung der Entwicklung Zweijähriger unter besonderer Berücksichtigung der Motorik, der Sprache und der sozialemotionalen Kompetenzen*. Unveröffentlichte Diplomarbeit, Universität, Wien.

- Kuchler, M., Sapper, E., Deimann, P. & Kastner-Koller, U. (2011). *Manual zum Itempool zur spielbasierten Erfassung der Entwicklung Zweijähriger*. Unveröffentlichtes Manuskript, Universität, Wien.
- La Heij, W. & Boelens, H. (2010). Object interference in children's colour and position naming: Lexical interference or task-set competition? *Language and Cognitive Processes*, 25, 568-588.
- Li, J. (2007). The development of color preferences. *Psychological Science*, 30, 891-894.
- Lienert, G. A. & Raatz, U. (1994). *Testaufbau und Testanalyse* (5. Aufl.). Weinheim: Beltz.
- Linder, T.W. (2008). *The Transdisciplinary Play Based Assessment-2 (TPBA)*. Baltimore: Brookes.
- Lissmann, I., Domsch, H. & Lohaus, A. (2006). Zur Stabilität und Validität von Entwicklungstestergebnissen im Alter von sechs Monaten bis zwei Jahren. Eine Analyse am Beispiel des ET 6-6. *Kindheit und Entwicklung*, 15, 35-44.
- Lowenfeld, M. (1976). Die „Welt“-Technik in der Kinderpsychotherapie. In: Biermann, G. (Hrsg.). *Handbuch der Kinderpsychotherapie* (4. Auflage, Bd 1, S. 442-451). München: Reinhardt.
- Lowenthal, B. (1997). Useful Early Childhood Assessment: Play-Based, Interviews and Multiple Intelligences. *Early Child Development and Care*, 129, 43-49.
- Lücking, A. & Scheithauer, H. (2006). Entwicklungstest sechs Monate bis sechs Jahre (ET 6-6) (2. Auflage, 2004). *Diagnostica*, 52, 100-103.
- Macha, T., Proske, A. & Petermann, F. (2005). Validität von Entwicklungstests. *Kindheit und Entwicklung*, 14, 150-162.
- Melchers, P., Floß, S., Brandt, I., Eßer, K.-J., Lehmkuhl, G., Rauh, H. & Sticker, E. (2003). *EVU - Erweiterte Vorsorgeuntersuchung*. Leiden: PITS Testverlag.
- Melchers, P. & Preuß, U. (1991). *Kaufman-Assessment Battery for Children, Deutschsprachige Fassung (K-ABC)*. Amsterdam: Swets & Zeitlinger.
- Michaelis, R. & Niemann, G. (2004). *Entwicklungsneurologie und Neuropädiatrie. Grundlagen diagnostischer Strategien* (3. Aufl.). Stuttgart: Thieme.
- Montada, L. (2008). Fragen, Konzepte, Perspektiven. In R. Oerter & L. Montada (Hrsg.), *Entwicklungspsychologie* (6., überarbeitete Aufl., S. 3-48). Weinheim: Beltz.

- O'Hanlon, C. G. & Roberson, D. (2006). Learning in context: Linguistic and attentional constraints on children's color term learning. *Journal of Experimental Child Psychology*, 94, 275-300.
- O'Hanlon, C. G. & Roberson, D. (2007). What constrains children's learning of novel shape terms? *Journal of Experimental Child Psychology*, 97, 138-148.
- Ozturk, O., Shayan, S., Liszkowski, U. & Majid, A. (2013). Language is not necessary for color categories. *Developmental Science*, 16, 111-115.
- Petermann, F. & Macha, T. (2005). Entwicklungsdiagnostik. *Kindheit und Entwicklung*, 14, 131-139.
- Petermann, F. & Renziehausen, A. (2005). *Neuropsychologisches Entwicklungsscreening (NES)*. Bern: Huber.
- Petermann, F., Stein, I. A. & Macha, T. (2008). *Entwicklungstest sechs Monate bis sechs Jahre (ET 6-6)* (3. Aufl.). Frankfurt: Pearson Assessment.
- Pinel, J. P. J. & Pauli, P. (Hrsg.). (2007). *Biopsychologie* (6. aktualisierte Aufl.). München: Pearson.
- Pitchford, N. J. (2006). Reflections on how color term acquisition is constrained. *Journal of Experimental Child Psychology*, 94, 328-333.
- Pitchford, N. J., Davis, E. E. & Scerif, G. (2009). Does colour preference have a role in colour term acquisition? *British Journal of Developmental Psychology*, 27, 993-1012.
- Pitchford, N. J. & Mullen, K. T. (2001). Conceptualization of perceptual attributes: A special case for color? *Journal of Experimental Child Psychology*, 80, 289-314.
- Pitchford, N. J. & Mullen, K. T. (2002). Is the development of basic colour terms in young children constrained? *Perception*, 31, 1349-1370.
- Pitchford, N. J. & Mullen, K. T. (2005). The role of perception, language, and preference in the developmental acquisition of basic color terms. *Journal of Experimental Child Psychology*, 90, 275-302.
- Preston, P. (2005). *Testing Children. A Practitioner's Guide to the Assessment of Mental Development in Infants and Young Children*. Göttingen: Hogrefe.
- Prevor, M. B. & Diamond, A. (2005). Color-object interference in young children: A Stroop effect in children 3½–6½ years old. *Cognitive Development*, 20, 256-278.

- Punz, S. (in Vorbereitung). *Die Überarbeitung und testtheoretische Überprüfung eines Itempools zur spielbasierten Erfassung der Entwicklung Zweijähriger unter besonderer Berücksichtigung der kognitiven Entwicklung*. Universität, Wien.
- Putzer, S. (2013). *Die Entwicklung und Erprobung eines Itempools zur spielbasierten Erfassung der Entwicklung Zweijähriger unter besonderer Berücksichtigung der Grobmotorik*. Universität, Wien.
- Quaiser-Pohl, C. (2010). Wie teste ich Kinder? Zur Gestaltung der diagnostischen Situation. In C. Quaiser-Pohl & H. Rindermann (Hrsg.), *Entwicklungsdiagnostik* (S. 57-68). München: Reinhardt.
- Quaiser-Pohl, C. & Köhler, A. (2010). Allgemeine Entwicklungstests. In C. Quaiser-Pohl & H. Rindermann (Hrsg.), *Entwicklungsdiagnostik* (S. 85-100). München: Reinhardt.
- Rambert, M. L. (1969). *Das Puppenspiel in der Kinderpsychotherapie*. München: Reinhardt.
- Rennen-Allhoff, B., Allhoff, P., Bowi, U. & Laaser, U. (1993). *Elternbeteiligung bei Entwicklungsdiagnostik und Vorsorge. Verfahren zur Früherkennung von Krankheiten und Entwicklungsstörungen im Säuglings- und Kindesalter durch Eltern*. Weinheim: Juventa.
- Renziehausen, A. & Petermann, F. (2007). Zur prädiktiven Validität des neuropsychologischen Entwicklungs-Screenings NES. *Kindheit und Entwicklung*, 16, 62-72.
- Reuner, G., Rosenkranz, J., Pietz, J. & Horn, R. (2008). *Bayley II. Bayley Scales of Infant Development* (2nd edition) – *Deutsche Fassung* (2., korrigierte Aufl.). Frankfurt a. M.: Pearson Assessment & Information.
- Ricken, G., Fritz, A., Schuck, K. D. & Preuß, U. (2007). *HAWIVA-III. Hannover-Wechsler-Intelligenztest für das Vorschulalter – III*. Bern: Huber.
- Rindermann, H. & Kwiatkowski, V. (2010). Diagnostik von Teilleistungsstörungen. In: C. Quaiser-Pohl (Hrsg.), *Entwicklungsdiagnostik* (S. 295-318). München: Reinhardt.
- Roberson, D., Davies, I. R. L., Davidoff, J. (2000). Color categories are not universal: Replications and new evidence from a stone age culture. *Journal of Experimental Psychology: General*, 129, 369-398.

- Roberson, D., Pak, H. S. & Hanley, J. R. (2008). Categorical perception of colour in the left and right visual field is verbally mediated: Evidence from Korean. *Cognition*, 107, 752-762.
- Rümmele, A. (2011). Inhalt und Ziele einer entwicklungsorientierten Diagnostik. In H. Keller (Hrsg.), *Handbuch der Kleinkindforschung* (4., überarbeitete Aufl., S. 928-959). Bern: Huber.
- Sandhofer, C. M. & Smith, L. B. (1999). Learning color words involves learning a system of mappings. *Developmental Psychology*, 35, 668-679.
- Sandhofer, C. M. & Smith, L. B. (2001). Why Children Learn Color and Size Words So Differently: Evidence From Adults' Learning of Artificial Terms. *Journal of Experimental Psychology: General*, 130, 600-620.
- Sandhofer, C. M. & Thom, E. E. (2006). Taking the task seriously: Reflections on measures of color acquisition. *Journal of Experimental Child Psychology*, 94, 344-348.
- Sapper, E. (2011). *Die Entwicklung und Erprobung eines spielbasierten Itemspools zur Erfassung der Entwicklung Zweijähriger unter besonderer Berücksichtigung der kognitiven Entwicklung, der visuellen Wahrnehmung, des Gedächtnisses und der Arbeitshaltungen*. Unveröffentlichte Diplomarbeit, Universität, Wien.
- Sarimski, K. (2001). Testinformation Entwicklungstest von 6 Monaten bis 6 Jahren (ET 6-6). *Diagnostica*, 47, 107-109.
- Sarimski, K. (2009). Entwicklungsdiagnostik. In D. Irblich & G. Renner (Hrsg.), *Diagnostik in der Klinischen Kinderpsychologie. Die ersten sieben Lebensjahre* (S. 123-135). Göttingen: Hogrefe.
- Schölmerich, A. (2011). Verhaltensbeobachtung. In H. Keller (Hrsg.), *Handbuch der Kleinkindforschung* (4. Aufl., S. 768-789). Bern: Huber.
- Sindelar, I. (2013). *Die Entwicklung und Erprobung eines Itemspools zur spielbasierten Erfassung der Entwicklung Zweijähriger unter besonderer Berücksichtigung der Feinmotorik und der Visumotorik*. Universität, Wien.
- Soja, N. N. (1992). Inferences about the meanings of nouns: The relationship between perception and syntax. *Cognitive Development*, 7, 29-45.
- Soja, N. N. (1994). Young children's concept of color and its relation to the acquisition of color words. *Child Development*, 65, 918-937.

- Staabs, G. v. (1985). Der Sceno-Test. *Ein Beitrag zur Erfassung unbewußter charakterologischer Struktur in Diagnostik und Therapie* (6. Auflage). Bern: Huber.
- Sturzbecher, D. (1993). Der Familien-Interaktions-Test für Vorschulkinder (FIT-K) und erste Ergebnisse seiner Erprobung. *Psychologie in Erziehung und Unterricht*, 40, 249-262.
- Sturzbecher, D. (2001). Methodische Lösungsansätze zur Befragung jüngerer Kinder. In D. Sturzbecher (Hrsg.), *Spielbasierte Befragungstechniken. Interaktionsdiagnostische Verfahren für Begutachtung, Beratung und Forschung* (S. 51-63). Göttingen: Hogrefe.
- Taylor, C., Clifford, A. & Franklin, A. (2012). Color Preferences Are Not Universal. *Journal of Experimental Psychology: General*, Advance online publication. Doi: 10.1037/a0030273.
- Tellegen, P. J., Laros, J. A. & Petermann, F. (2007). *SON-R 21/2-7. Non-verbaler Intelligenztest*. Göttingen: Hogrefe.
- Teller, D. Y., Civan, A. & Bronson-Castain, K. (2004). Infants' spontaneous color preferences are not due to adult-like brightness variations. *Visual Neuroscience*, 21, 397-401.
- Tröster, H., Flender, J. & Reinecke, D. (2005). Dortmunder Entwicklungsscreening für den Kindergarten (DESK 3-6). *Kindheit und Entwicklung*, 14, 140-149.
- Vig, S. (2007). Young Children's Object Play: A Window on Development. *Journal of developmental and physical disabilities*, 19, 201-215.
- 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). Göttingen: Hogrefe.
- Yee, M., Jones, S. S. & Smith, L. B. (2012). Changes in visual object recognition precede the shape bias in early noun learning. *Frontiers in Psychology*, 3, 1-13.
- Zemach, I., Chang, S. & Teller, D. Y. (2007). Infant color vision: prediction of infants' spontaneous color preferences. *Vision Research*, 47, 1368-1381.
- Zentner, M. R. (2001). Preferences for colours and colour emotion combinations in early childhood. *Developmental Science*, 4, 389-398.

TABELLENVERZEICHNIS

Tabelle 1: Darstellung des Alters der Kinder in Monaten, getrennt nach Geschlecht.....	79
Tabelle 2: Aufteilung der Stichprobe in die jüngere und ältere Altersgruppe.....	82
Tabelle 3: Ergebnisse der Itemanalyse zur aktiven Farbdifferenzierung.....	84
Tabelle 4: Ergebnisse der Itemanalyse zur passiven Farbdifferenzierung.....	85
Tabelle 5: Ergebnisse der Itemanalyse zur Stabilität des Farbkonzepts.....	86
Tabelle 6: Ergebnisse der Itemanalyse zur Form- und Größendifferenzierung.....	87
Tabelle 7: Ergebnisse der Itemanalyse zur Grobmotorik.....	88
Tabelle 8: Ergebnisse der Itemanalyse zur Fein-/Visumotorik.....	89
Tabelle 9: Ergebnisse der Itemanalyse zur Feinmotorik.....	90
Tabelle 10: Ergebnisse der Itemanalyse zur Visumotorik.....	91
Tabelle 11: Darstellung der Aufmerksamkeit in Minuten getrennt nach Geschlecht und Alter.....	92
Tabelle 12: Darstellung der Häufigkeiten bzgl. der Anzahl der aufmerksam bearbeiteten Items getrennt nach Geschlecht und Alter.....	93
Tabelle 13: Häufigkeitsverteilung getrennt nach Geschlecht und Alter bzgl. der Anzahl der angesehen Buchseiten.....	94
Tabelle 14: Ergebnisse der Itemanalyse zur Selbstständigkeit.....	95
Tabelle 15: Häufigkeitsverteilung im Bezug auf die Lösung des Kindes von der Bezugsperson getrennt nach Geschlecht und Alter.....	96
Tabelle 16: Häufigkeitsverteilung in Bezug auf den Kontakt zur Bezugsperson getrennt nach Geschlecht und Alter.....	97
Tabelle 17: Auflistung aller nicht trennscharfen Items aus der Itemanalyse.....	98
Tabelle 18: Auflistung der internen Konsistenzen inklusive der Anzahl der Items pro Skala.....	99
Tabelle 19: Interkorrelationen des Alters in Monaten mit der Testleistung.....	102
Tabelle 20: Mittelwertsunterschiede für das Alter und das Geschlecht, inklusive Signifikanzwert.....	103
Tabelle 21: Überprüfung der Normalverteilung mittels Schiefe und Kurtosis für die männliche Stichprobe.....	148

Tabelle 22: Überprüfung der Normalverteilung mittels Schiefe und Kurtosis für die weibliche Stichprobe.....	149
Tabelle 23: Überprüfung der Normalverteilung mittels Schiefe und Kurtosis für die jüngere Altersgruppe.....	150
Tabelle 24: Überprüfung der Normalverteilung mittels Schiefe und Kurtosis für die ältere Altersgruppe.....	151
Tabelle 25: Überprüfung der Varianzhomogenität, getrennt nach Geschlecht und Alter	152

ABBILDUNGSVERZEICHNIS

Abb. A.1.1. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „aktive Farbdifferenzierung“ nach der Selektion, getrennt nach Geschlecht und Alter.....	134
Abb. A.1.2. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „passive Farbdifferenzierung“, getrennt nach Geschlecht und Alter.....	135
Abb. A.1.3. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „Farbstabilität“, getrennt nach Geschlecht und Alter.....	136
Abb. A.1.4. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „Form- und Größendifferenzierung“, getrennt nach Geschlecht und Alter.	137
Abb. A.1.5. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „Grobmotorik“, getrennt nach Geschlecht und Alter.....	138
Abb. A.1.6. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „Fein- und Visumotorik“, getrennt nach Geschlecht und Alter.....	139
Abb. A.1.7. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „Feinmotorik“, getrennt nach Geschlecht und Alter.....	140
Abb. A.1.8. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „Visumotorik“, getrennt nach Geschlecht und Alter.....	141
Abb. A.1.9. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „Selbstständigkeit“, getrennt nach Geschlecht und Alter.....	142
Abb. A.1.10. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „Lösung von der Bezugsperson“, getrennt nach Geschlecht und Alter...	143
Abb. A.1.11. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „Kontakt zur Bezugsperson“, getrennt nach Geschlecht und Alter.....	144
Abb. A.1.12. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „Aufmerksamkeit in Sekunden“, getrennt nach Geschlecht und Alter.....	145
Abb. A.1.13. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „Aufmerksamkeit: zurückholen“, getrennt nach Geschlecht und Alter.....	146
Abb. A.1.14. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „Aufmerksamkeit - Buchseiten“, getrennt nach Geschlecht und Alter.....	147

Abb. A.4.1. Histogramm mit der Normalverteilungskurve für das Alter in Monaten und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis....	153
Abb. A.4.2. Histogramm mit der Normalverteilungskurve für die Skala „Farbdifferenzierung aktiv“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis.....	153
Abb. A.4.3. Histogramm mit der Normalverteilungskurve für die Skala „Farbdifferenzierung passiv“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis.....	154
Abb. A.4.4. Histogramm mit der Normalverteilungskurve für die Skala „Farbstabilität“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis	154
Abb. A.4.5. Histogramm mit der Normalverteilungskurve für die Skala „Form- und Größendifferenzierung“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis.....	155
Abb. A.4.6. Histogramm mit der Normalverteilungskurve für die Skala „Grobmotorik“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis	155
Abb. A.4.7. Histogramm mit der Normalverteilungskurve für die Skala „Fein- und Visumotorik“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis.....	156
Abb. A.4.8. Histogramm mit der Normalverteilungskurve für die Skala „Feinmotorik“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis	156
Abb. A.4.9. Histogramm mit der Normalverteilungskurve für die Skala „Visumotorik“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis	157
Abb. A.4.10. Histogramm mit der Normalverteilungskurve für die Skala „Aufmerksamkeit - Sekunden“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis.....	157
Abb. A.4.11. Histogramm mit der Normalverteilungskurve für die Skala „Aufmerksamkeit - zurückholen“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis.....	158

Abb. A.4.12.	Histogramm mit der Normalverteilungskurve für die Skala „Aufmerksamkeit - Buchseiten“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis.....	158
Abb. A.4.13.	Histogramm mit der Normalverteilungskurve für die Skala „Selbstständigkeit“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis.....	159
Abb. A.4.14.	Histogramm mit der Normalverteilungskurve für die Skala „Lösung von der Bezugsperson“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis.....	159
Abb. A.4.15.	Histogramm mit der Normalverteilungskurve für die Skala „Kontakt zur Bezugsperson“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis.....	160

ANHANG

A ÜBERPRÜFUNG DER VORAUSSETZUNGEN

- A.1. Überprüfung der Normalverteilung mittels Histogramme für eine ANOVA
- A.2. Überprüfung der Normalverteilung mittels Schiefe und Kurtosis für eine ANOVA
- A.3. Überprüfung der Varianzhomogenität mittels Levene-Test für eine ANOVA
- A.4. Überprüfung der Normalverteilung mittels Histogramme für eine Pearson-Korrelation

B PROTOKOLLBOGEN

C ELTERNFRAGEBOGEN

D EINVERSTÄNDNISERKLÄRUNG FÜR DIE VIDEOAUFZEICHNUNG

E ELTERNBRIEF

F CURRICULUM VITAE

A ÜBERPRÜFUNG DER VORAUSSETZUNGEN FÜR EINE ANOVA

A.1. Überprüfung der Normalverteilung mittels Histogramme

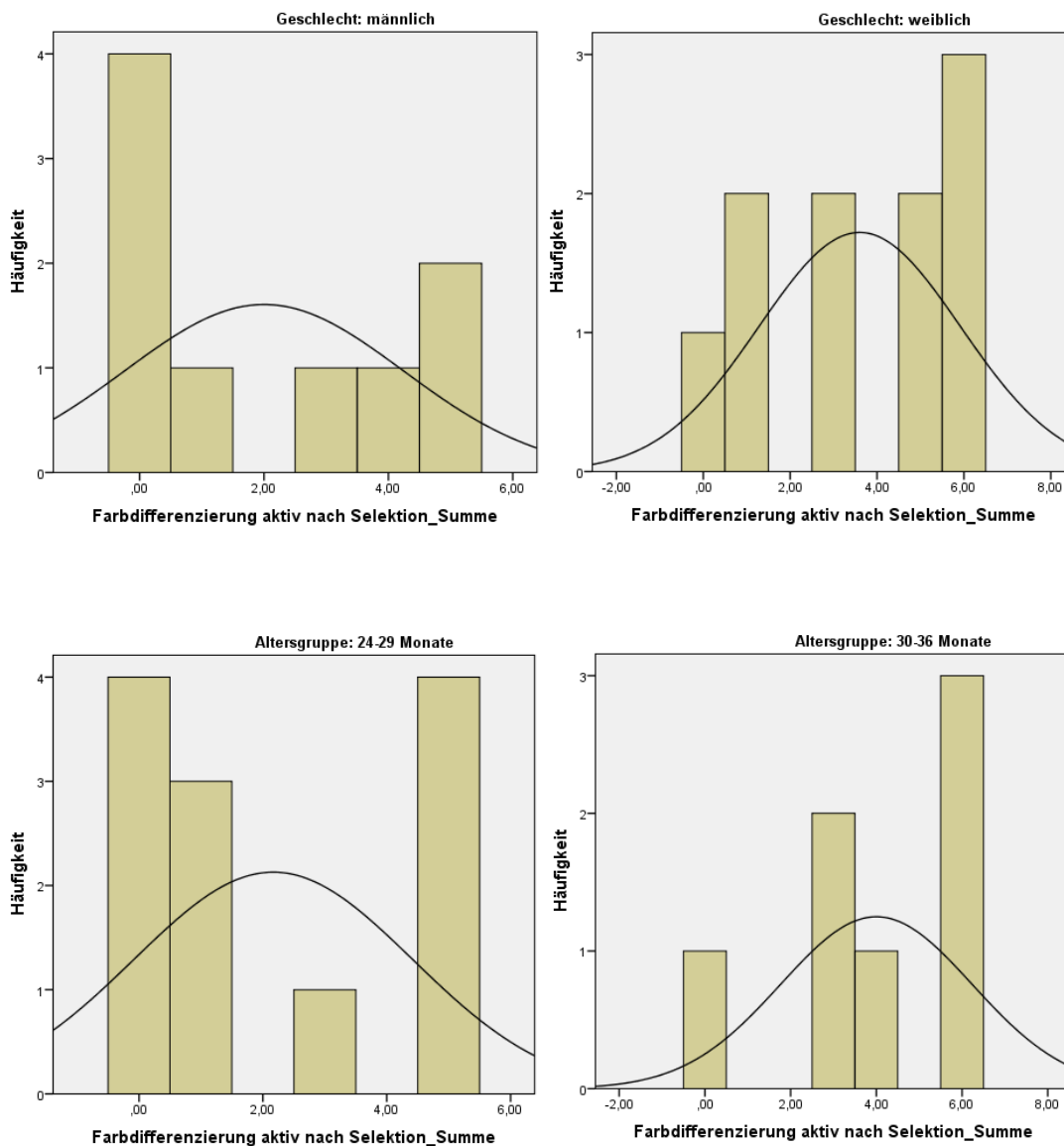


Abb. A.1.1. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „aktive Farbdifferenzierung“ nach der Selektion, getrennt nach Geschlecht und Alter

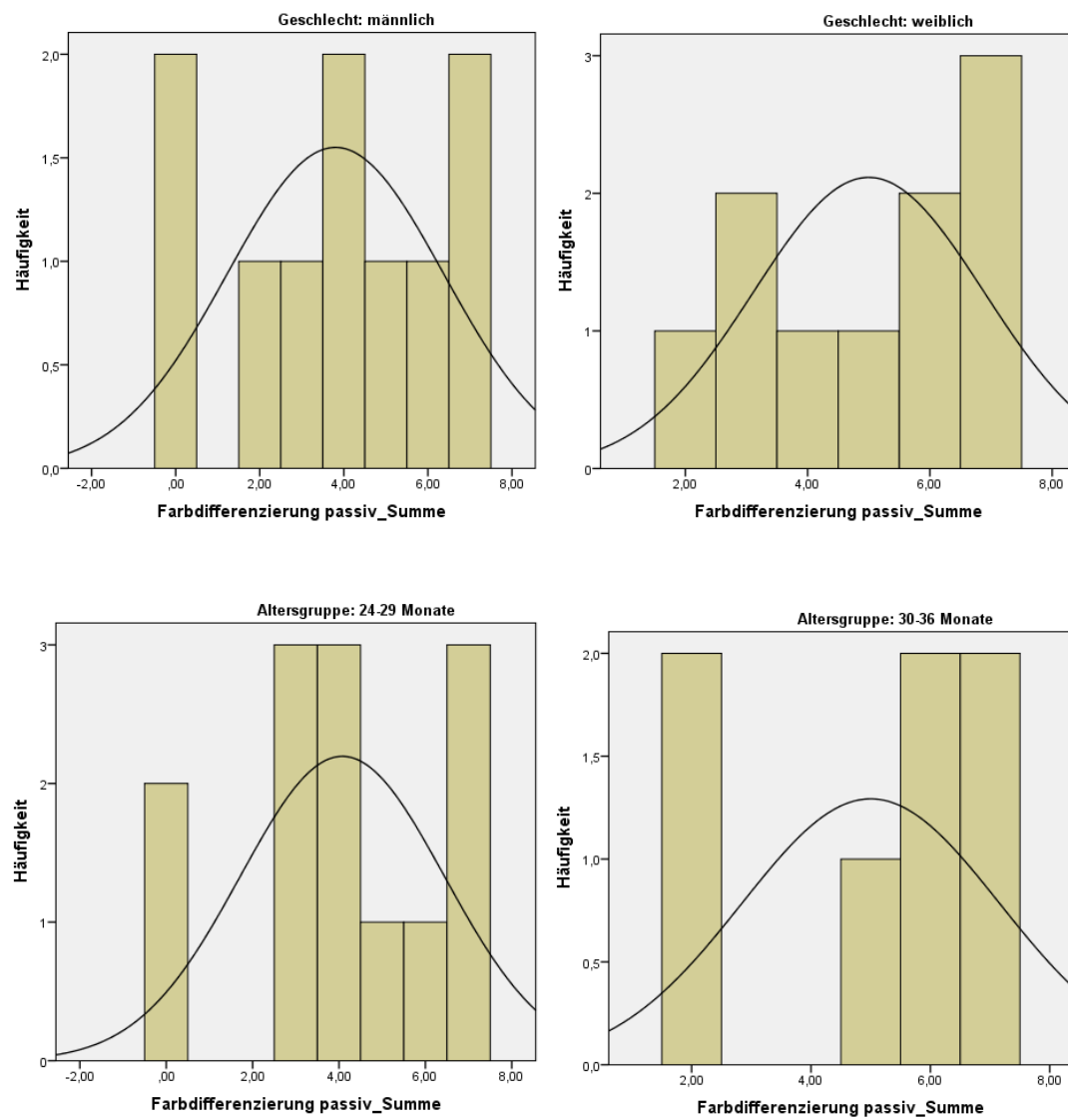


Abb. A.1.2. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „passive Farbdifferenzierung“, getrennt nach Geschlecht und Alter

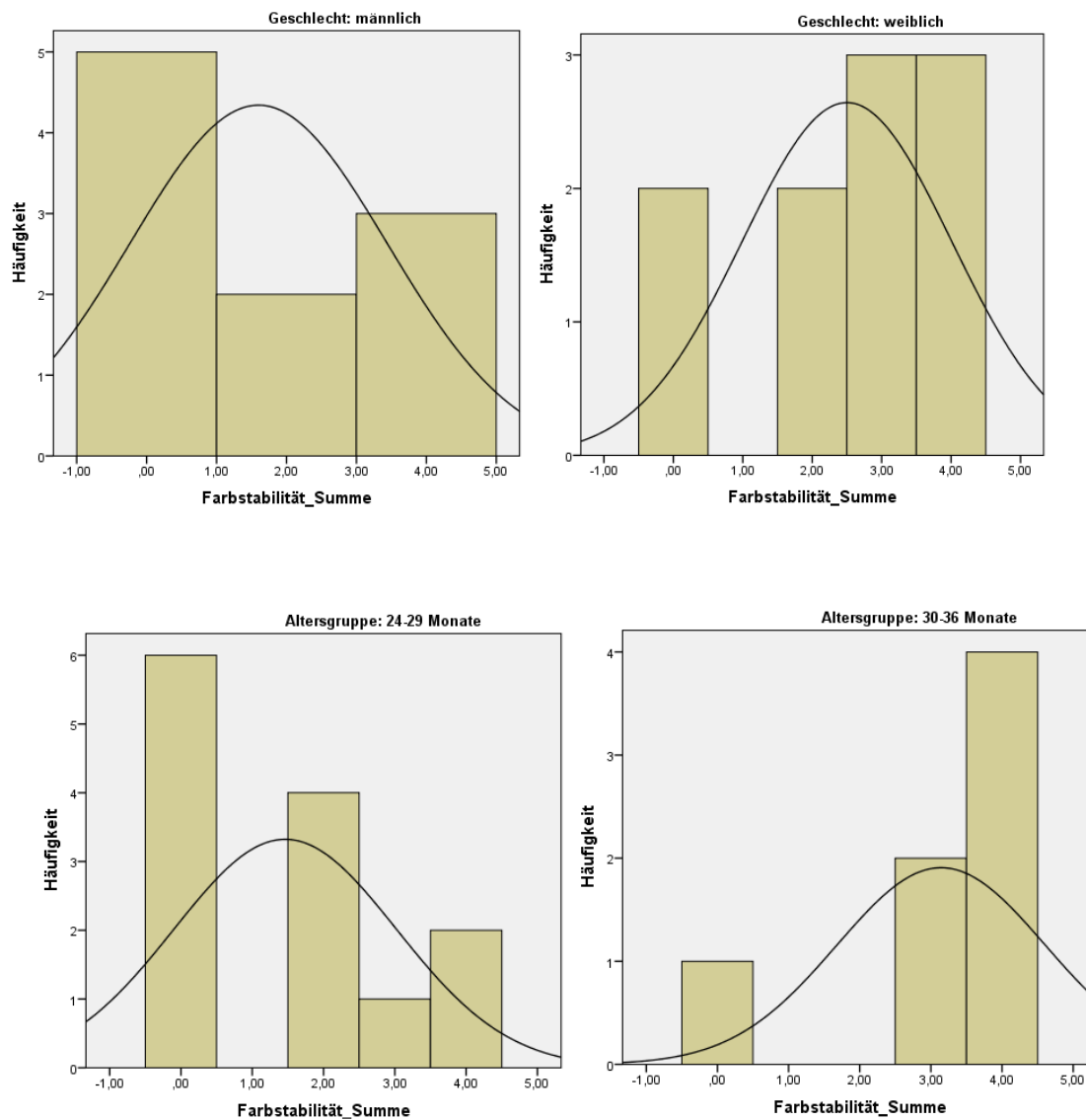


Abb. A.1.3. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „Farbstabilität“ nach der Selektion, getrennt nach Geschlecht und Alter

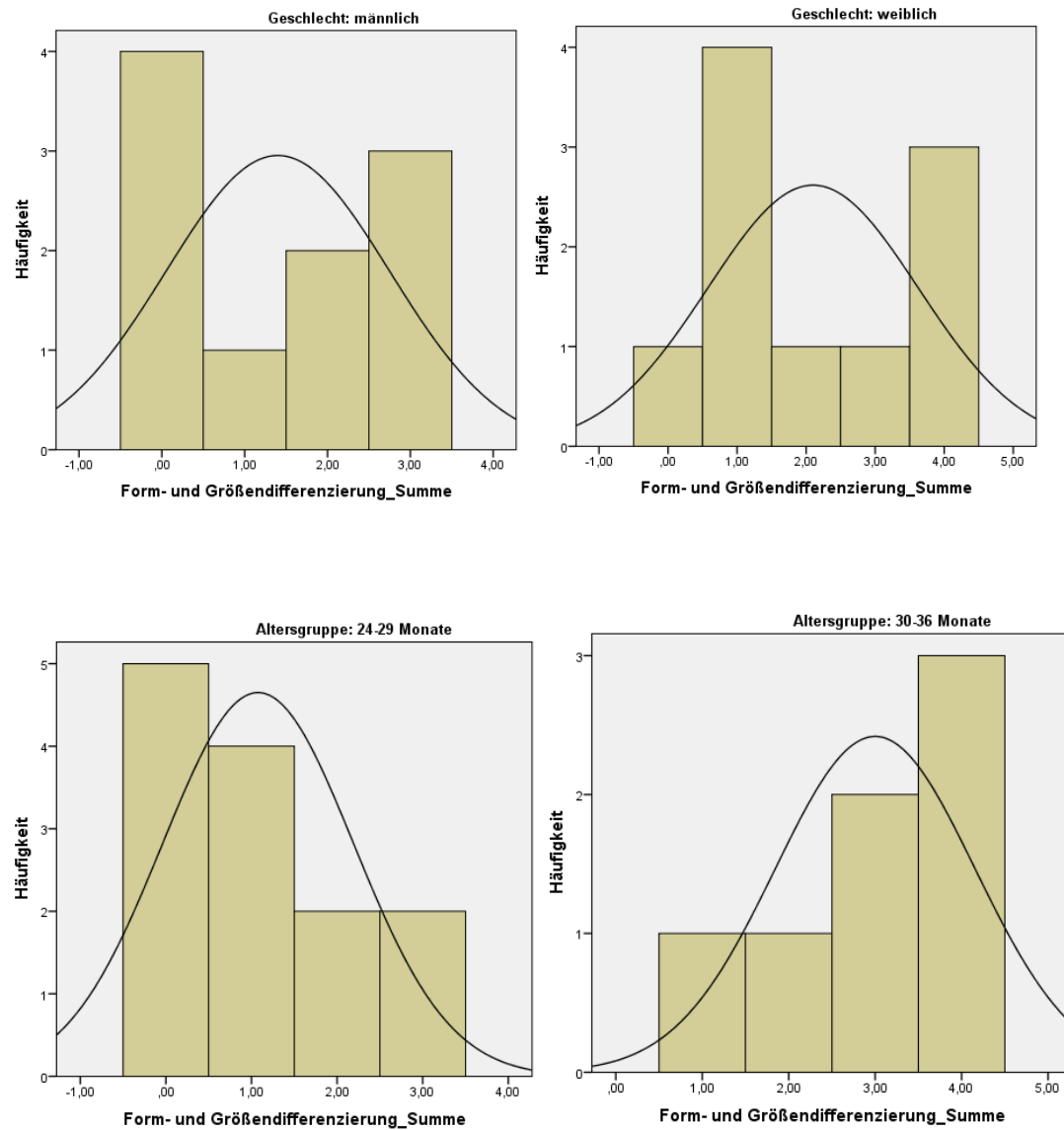


Abb. A.1.4. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „Form- und Größendifferenzierung“, getrennt nach Geschlecht und Alter

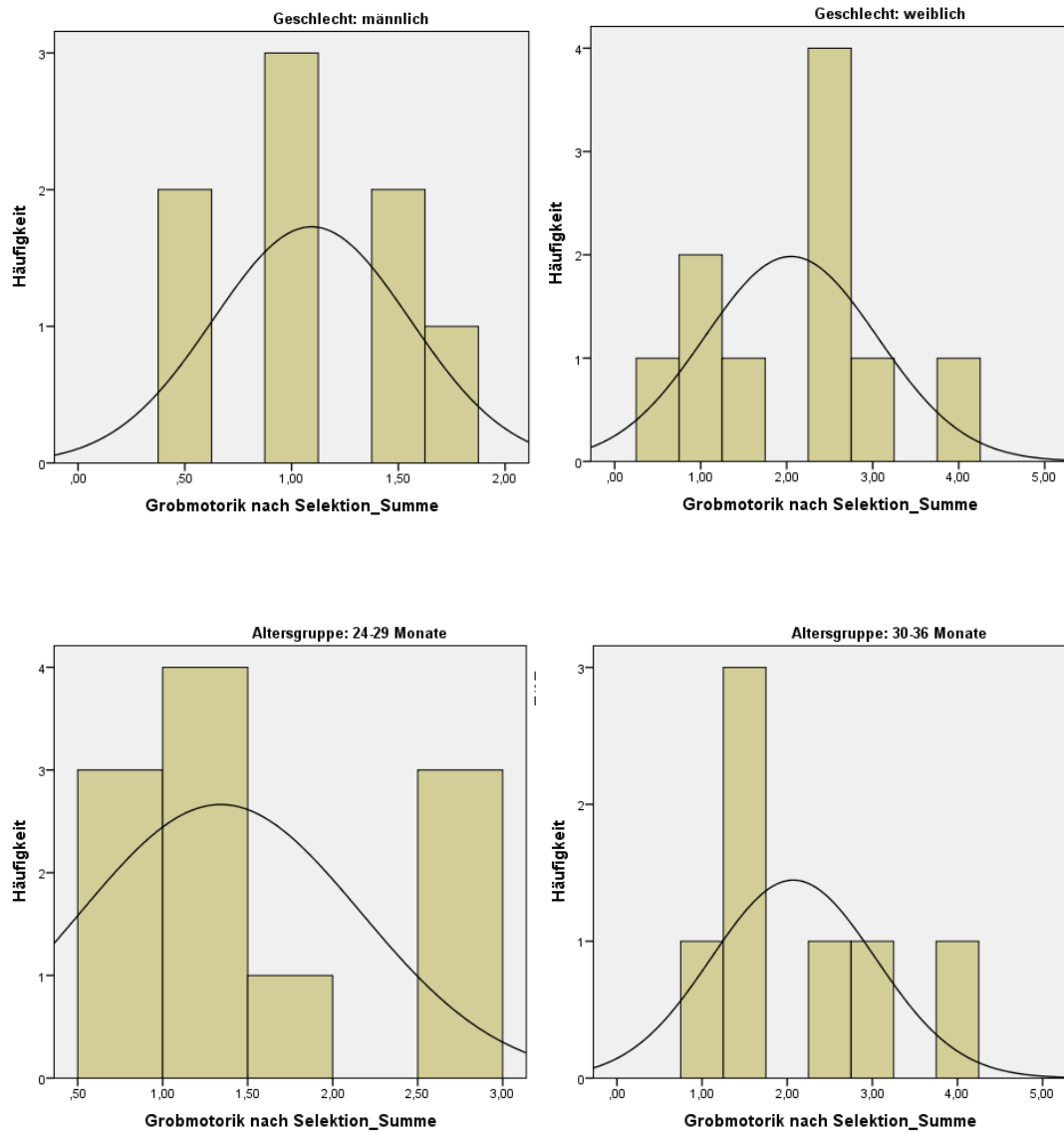


Abb. A.1.5. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „**Grobmotorik**“ nach der Selektion, getrennt nach Geschlecht und Alter

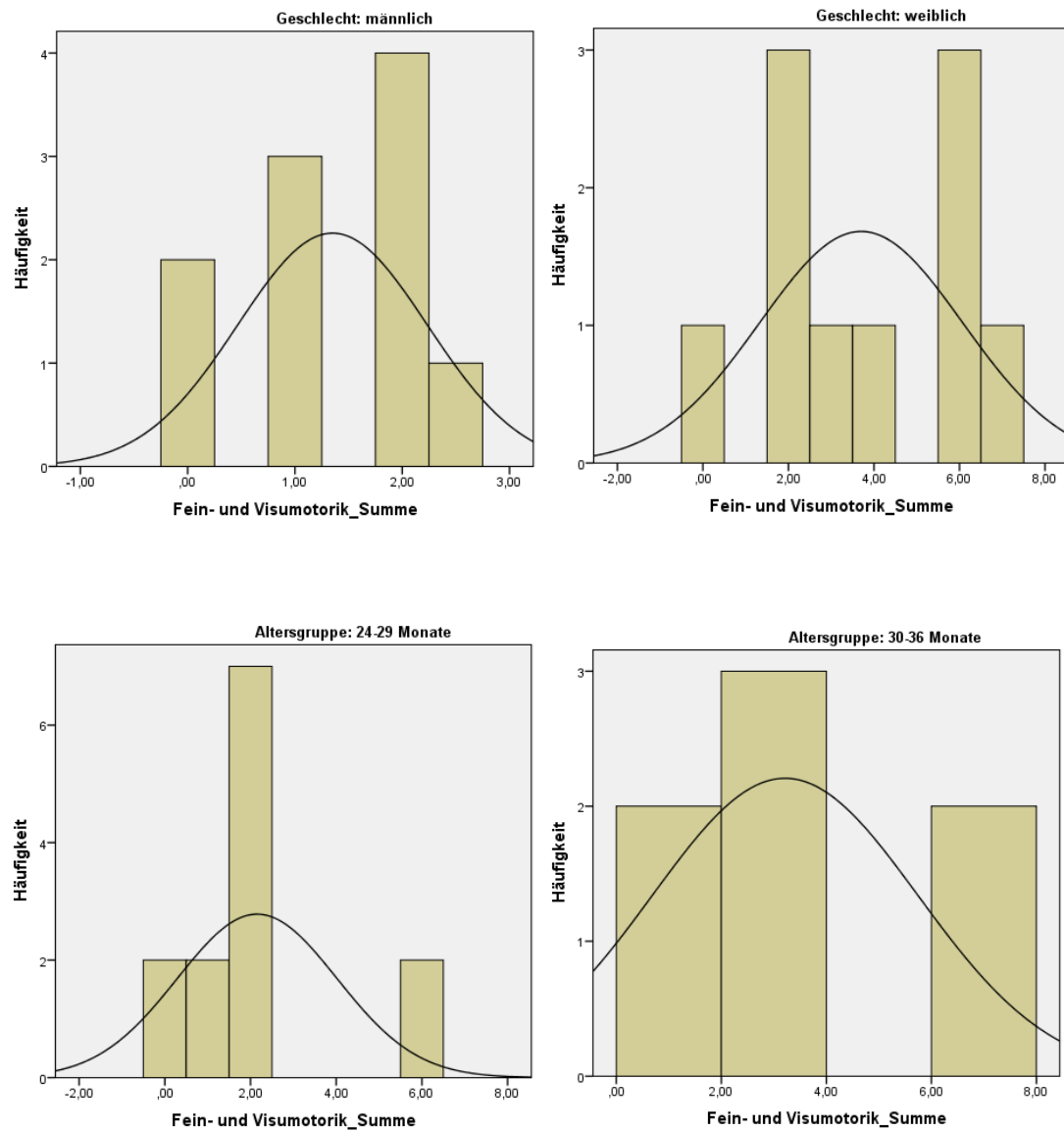


Abb. A.1.6. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „Fein- und Visumotorik“ nach der Selektion, getrennt nach Geschlecht und Alter

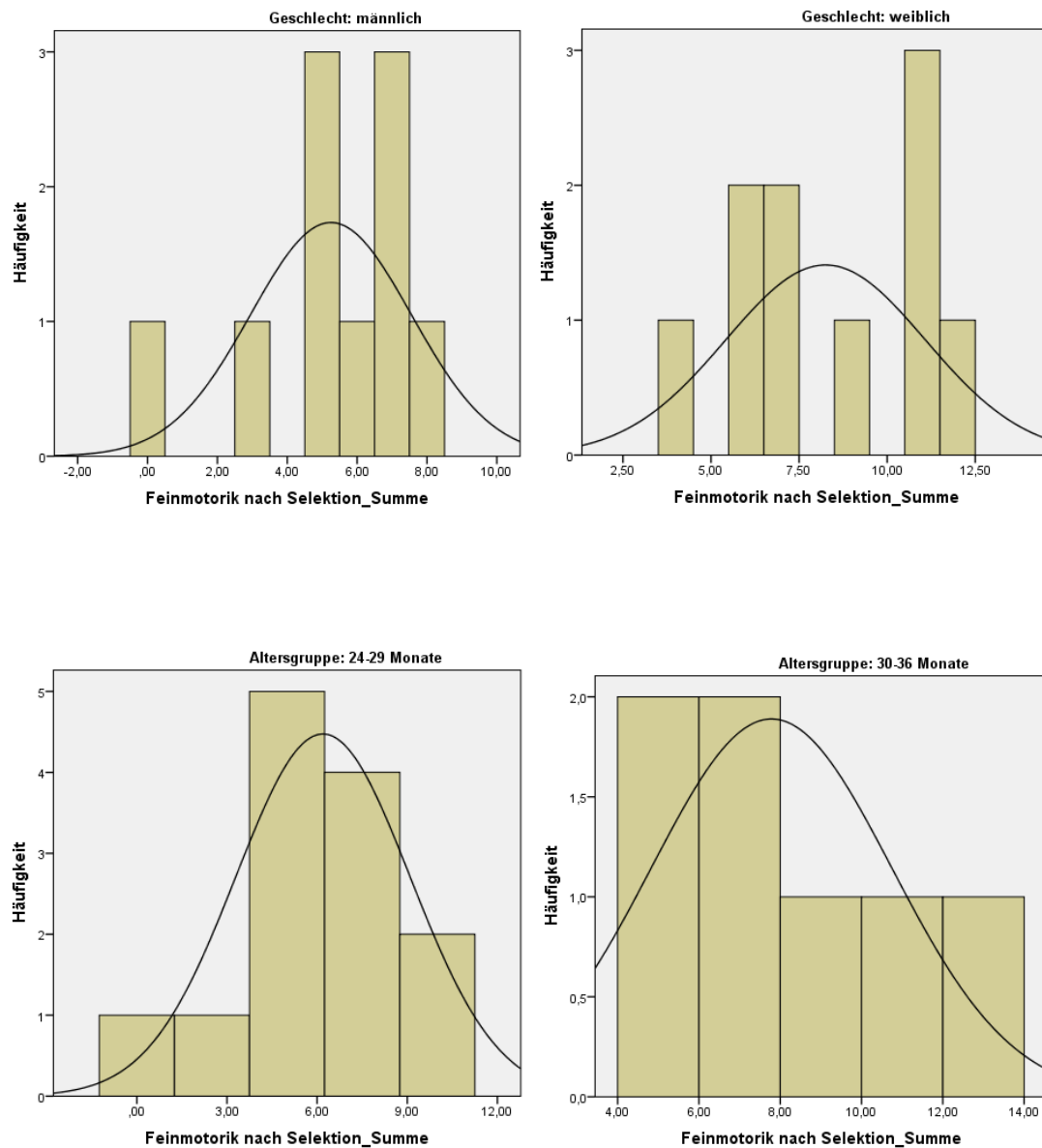


Abb. A.1.7. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „Feinmotorik“ nach der Selektion, getrennt nach Geschlecht und Alter

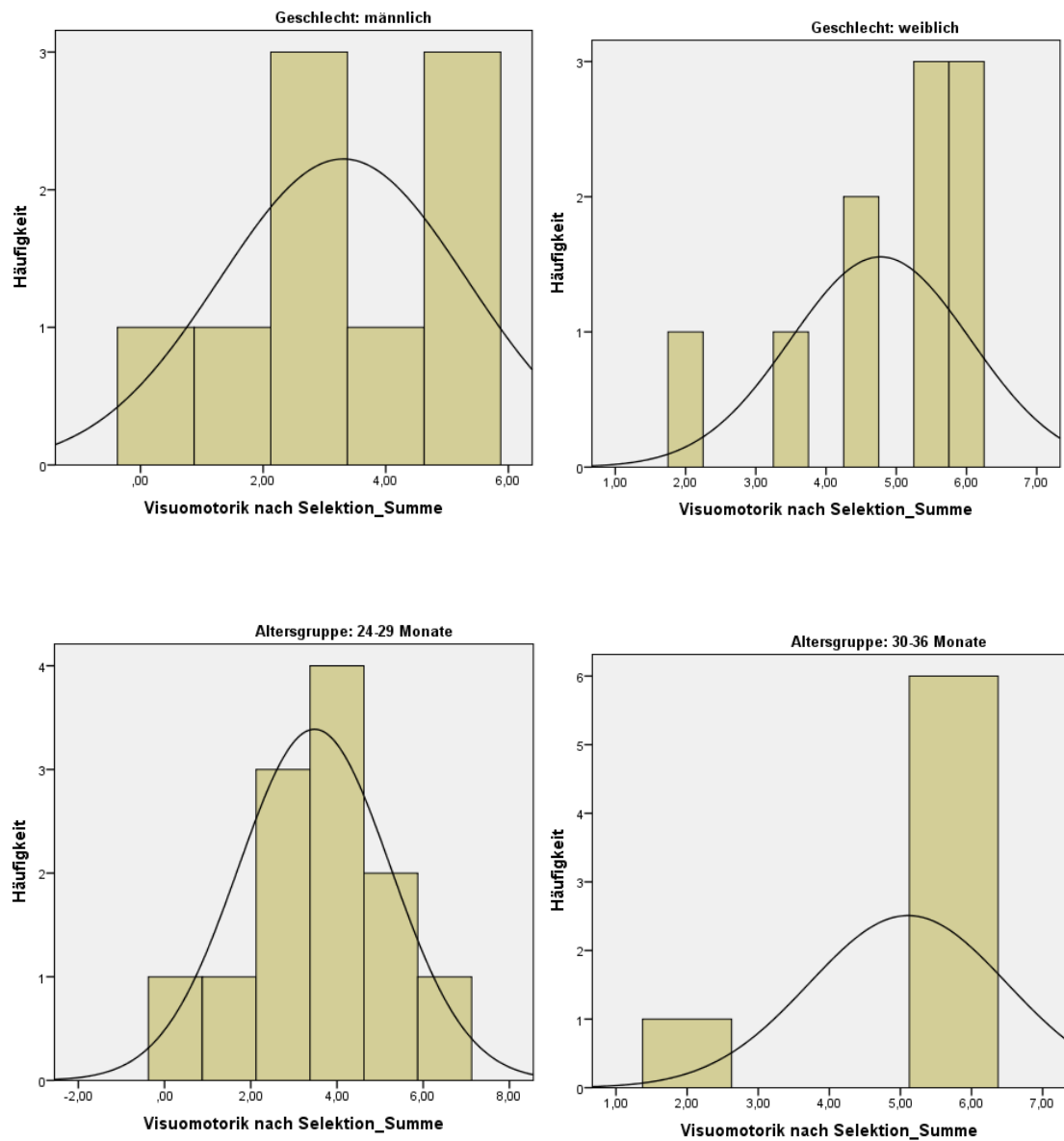


Abb. A.1.8. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „Visumotorik“ nach der Selektion, getrennt nach Geschlecht und Alter

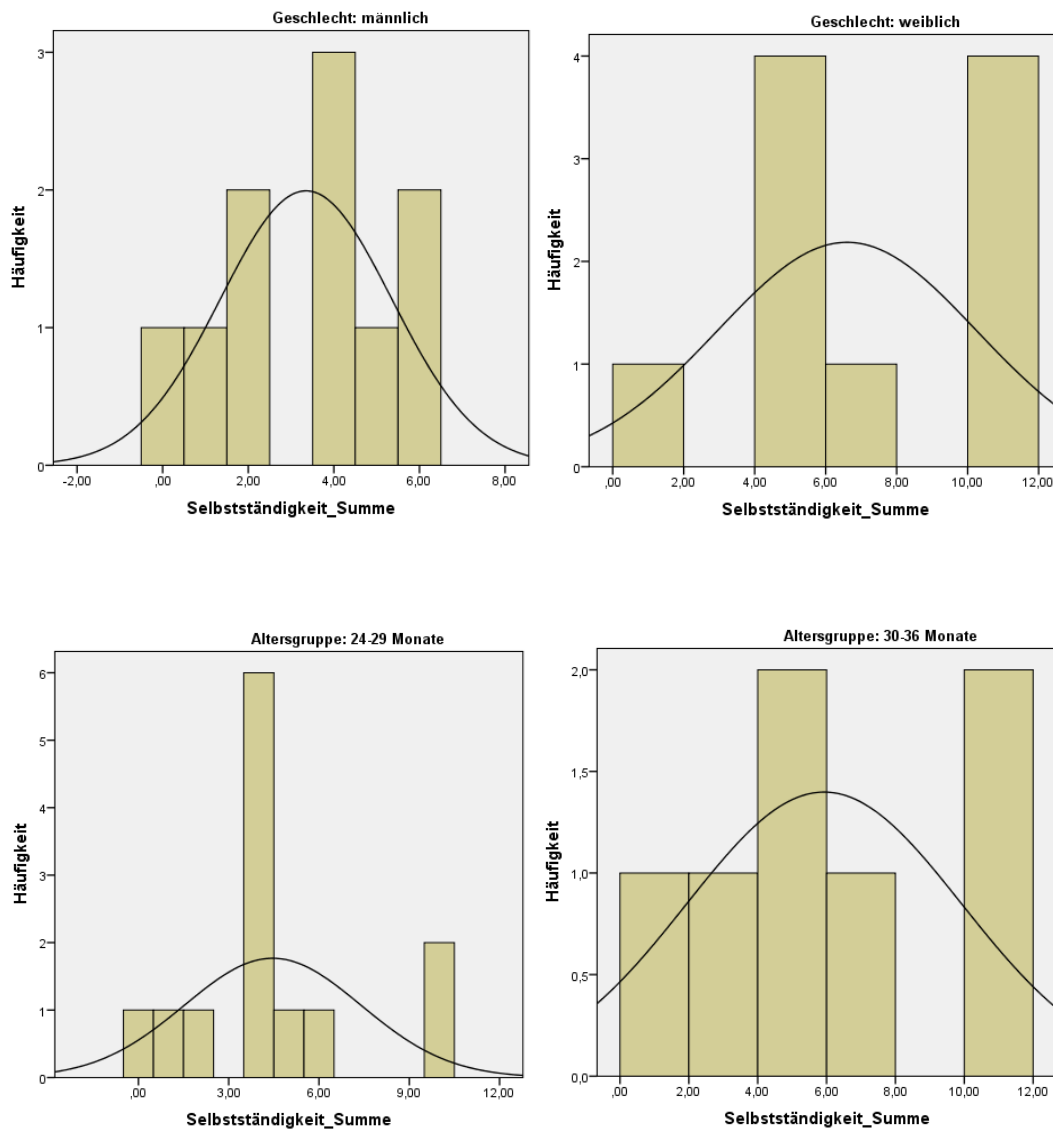


Abb. A.1.9. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „Selbstständigkeit“, getrennt nach Geschlecht und Alter

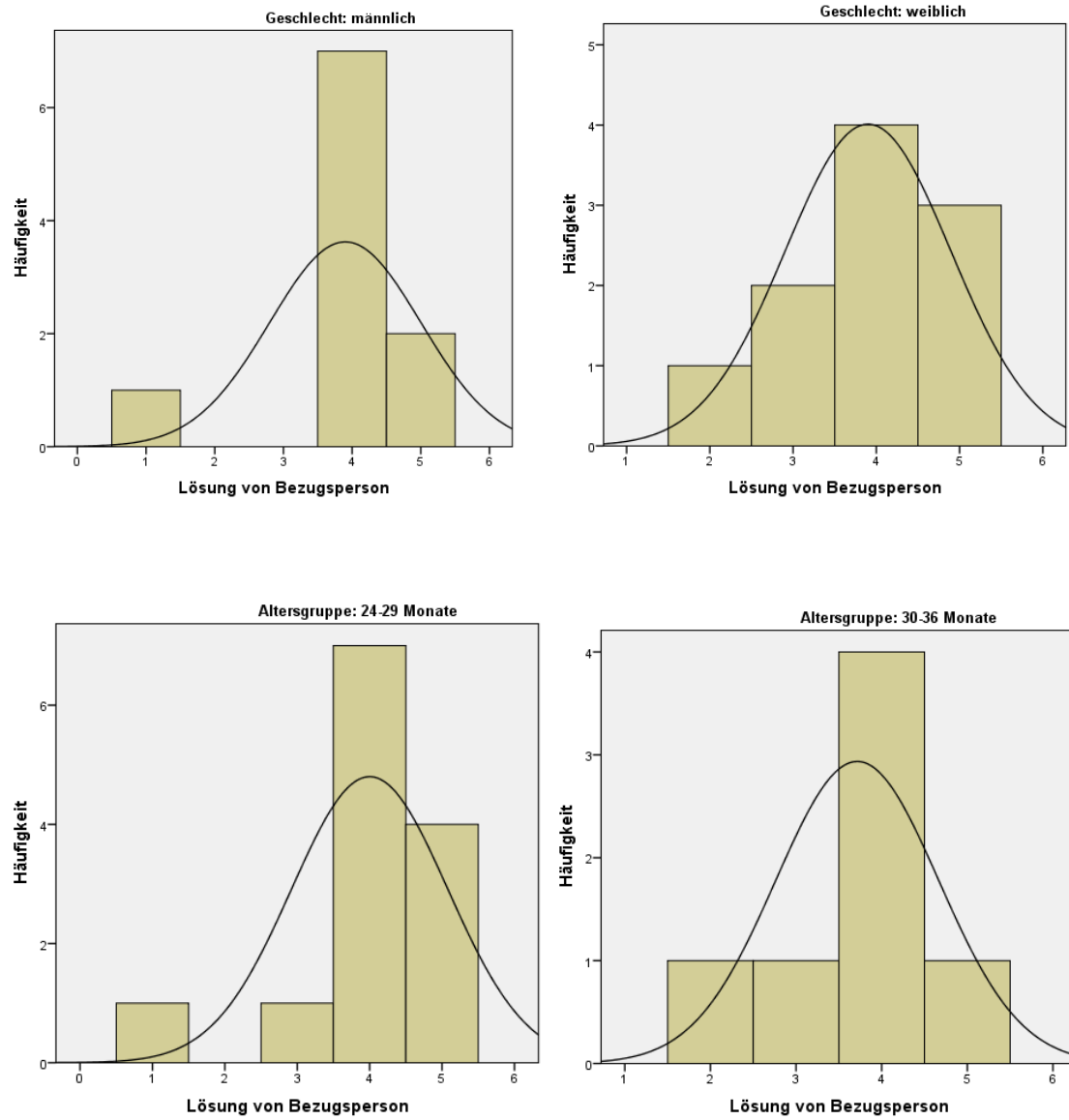


Abb. A.1.10. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „Lösung von der Bezugsperson“, getrennt nach Geschlecht und Alter

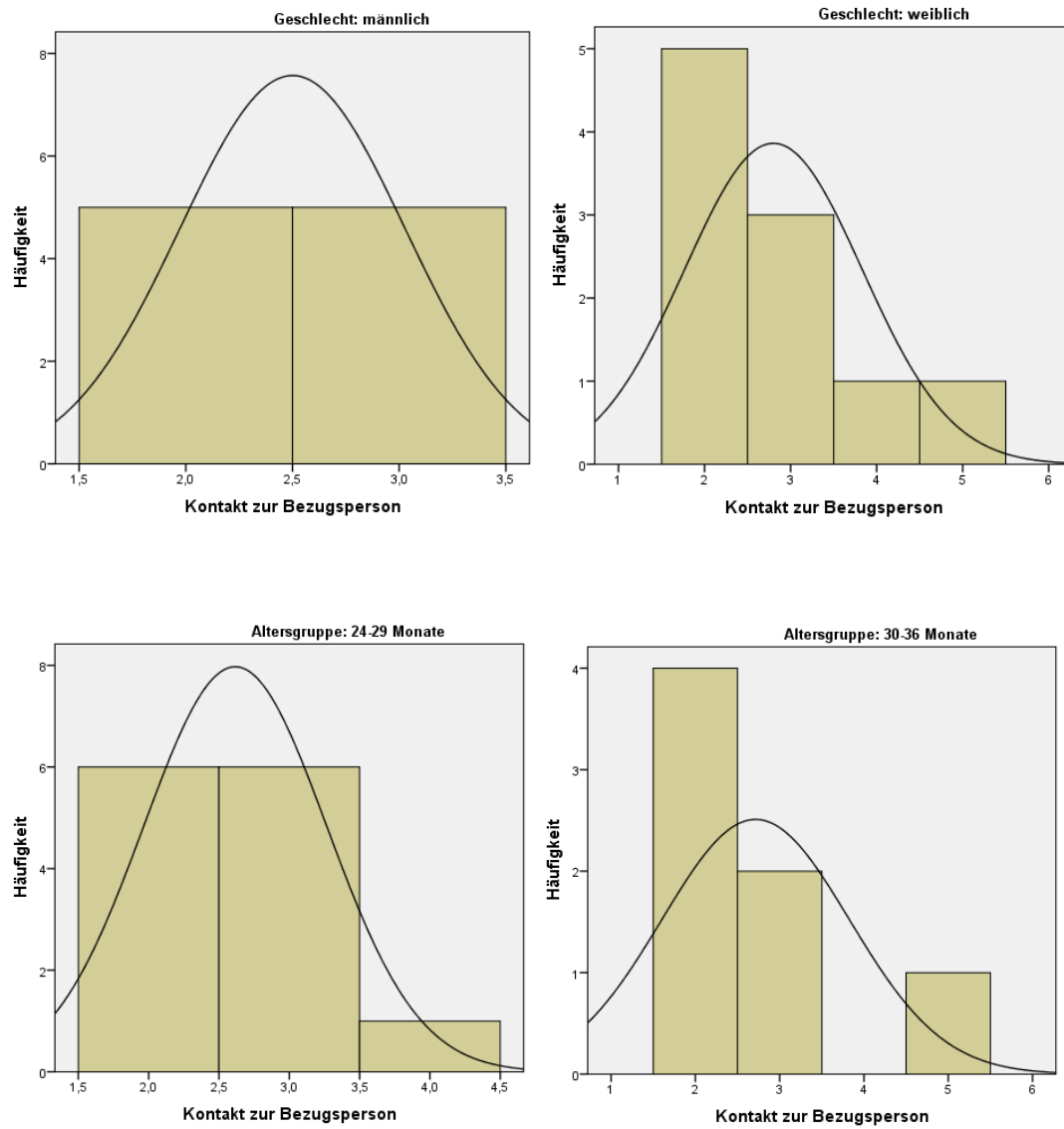


Abb. A.1.11. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „Kontakt zur Bezugsperson“, getrennt nach Geschlecht und Alter

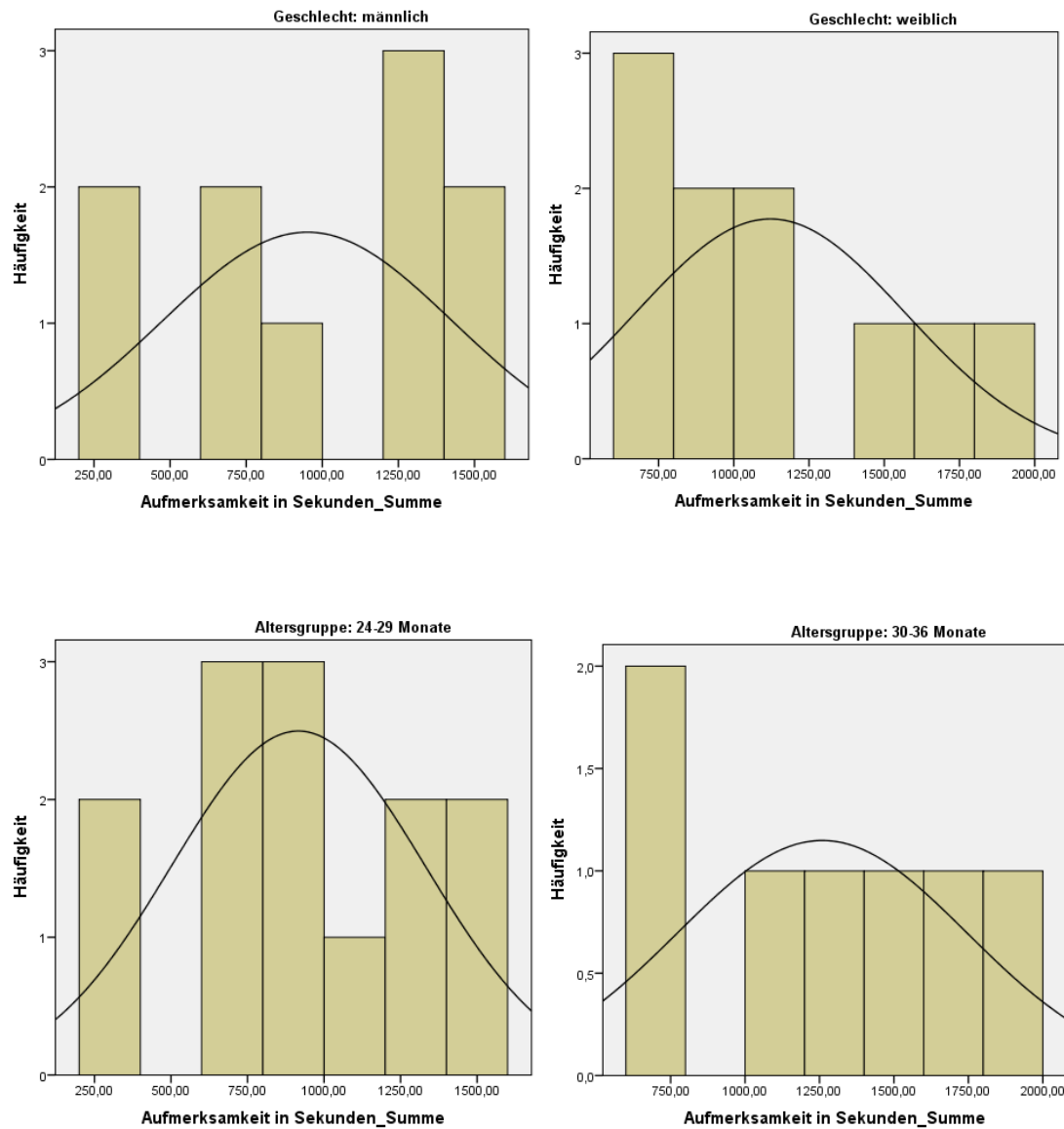


Abb. A.1.12. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „Aufmerksamkeit - in Sekunden“, getrennt nach Geschlecht und Alter

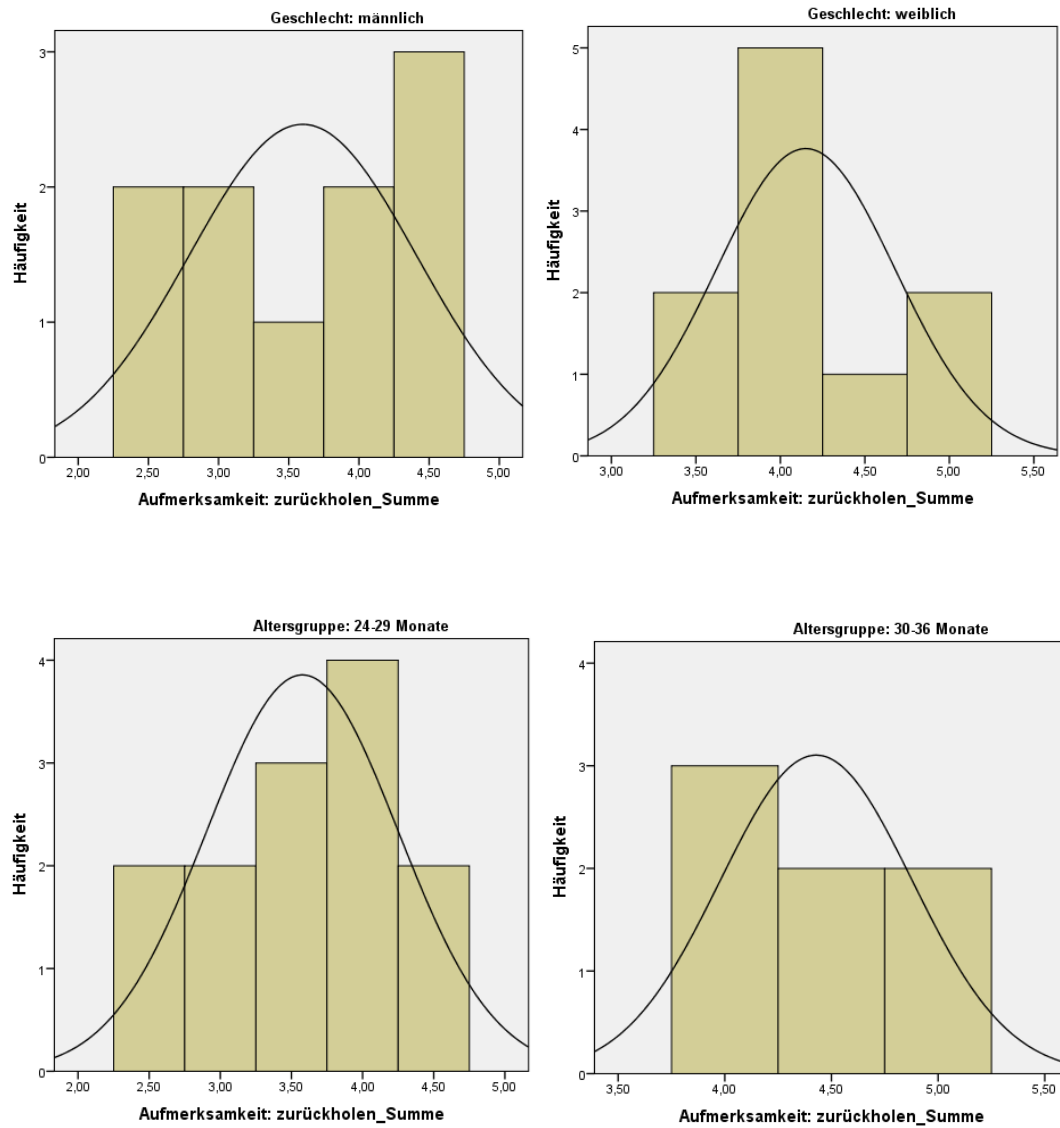


Abb. A.1.13. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „Aufmerksamkeit - zurückholen“, getrennt nach Geschlecht und Alter

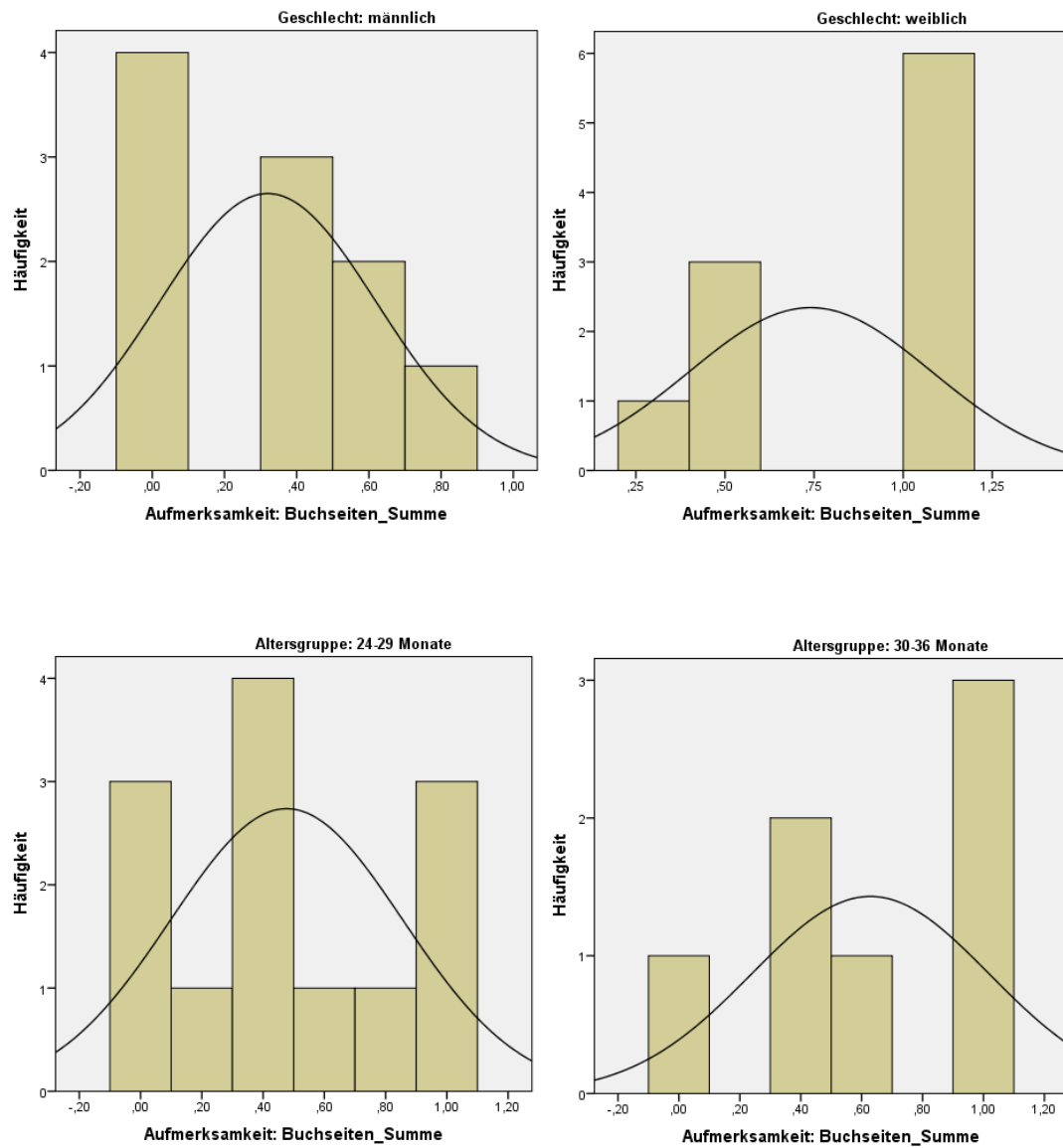


Abb. A.1.14. Histogramme mit den Normalverteilungskurven für die Summenscores der Skala „Aufmerksamkeit - Buchseiten“, getrennt nach Geschlecht und Alter

A.2. Überprüfung der Normalverteilung mittels Schiefe und Kurtosis

Tabelle 21: Überprüfung der Normalverteilung mittels Schiefe und Kurtosis für die männliche Stichprobe

BUBEN	N	MW	SD	Schiefe	Kurtosis
Aktive Farbdifferenzierung	9	2.000	2.236	.431	-1.957
Passive Farbdifferenzierung	10	3.800	2.573	-.311	-1.039
Stabilität des Farbkonzepts	10	1.600	1.838	.473	-1.807
Form- und Größendifferenzierung	10	1.400	1.349	.095	-2.018
Grobmotorik	8	1.094	.462	-.009	-1.212
Fein-/Visumotorik	10	1.350	.883	-.489	-1.061
Feinmotorik	10	5.250	2.300	-1.463	2.203
Visumotorik	9	3.306	2.019	.004	-1.311
Selbstständigkeit	10	3.350	2.001	-.375	-1.047
Lösung von der Bezugsperson	10	3.90	1.101	-2.263	6.621
Kontakt zur Bezugsperson	10	2.50	.527	.000	-2.571
Aufmerksamkeit in Sekunden	10	950.70	478.34	-.551	-1.362
Aufmerksamkeit: zurückholen	10	3.600	.810	-.204	-1.695
Aufmerksamkeit: Buchseiten	10	.320	.301	.127	-1.487

Tabelle 22: Überprüfung der Normalverteilung mittels Schiefe und Kurtosis für die weibliche Stichprobe

MÄDCHEN	N	MW	SD	Schiefe	Kurtosis
Aktive Farbdifferenzierung	10	3.600	2.319	-.393	-1.571
Passive Farbdifferenzierung	10	5.000	1.886	-.373	-1.508
Stabilität des Farbkonzepts	10	2.500	1.509	-.848	-.401
Form- und Größendifferenzierung	10	2.100	1.524	.264	-1.706
Grobmotorik	10	2.050	1.006	-.080	-.699
Fein-/Visumotorik	10	3.700	2.371	.055	-1.459
Feinmotorik	10	8.250	2.831	-.007	-1.654
Visumotorik	10	4.775	1.283	-1.220	1.135
Selbstständigkeit	10	6.600	3.650	-.393	-.881
Lösung von der Bezugsperson	10	3.90	.994	-.610	-.157
Kontakt zur Bezugsperson	10	2.80	1.033	1.241	.946
Aufmerksamkeit in Sekunden	10	1121.40	449.81	.692	-.711
Aufmerksamkeit: zurückholen	10	4.150	.530	.659	-.406
Aufmerksamkeit: Buchseiten	10	.740	.341	-.597	-1.865

Tabelle 23: Überprüfung der Normalverteilung mittels Schiefe und Kurtosis für die jüngere Altersgruppe

JÜNGERE ALTERSGRUPPE	N	MW	SD	Schiefe	Kurtosis
Aktive Farbdifferenzierung	12	2.167	2.250	.442	-1.868
Passive Farbdifferenzierung	13	4.077	2.362	-.422	-.425
Stabilität des Farbkonzepts	13	1.462	1.561	.473	-1.223
Form- und Größendifferenzierung	13	1.077	1.115	.678	-.760
Grobmotorik	11	1.341	.824	.612	-1.379
Fein-/Visumotorik	13	2.154	1.864	1.379	1.604
Feinmotorik	13	6.192	2.898	-.200	1.252
Visumotorik	12	3.479	1.766	-.283	-.542
Selbstständigkeit	13	4.462	2.933	.772	.664
Lösung von der Bezugsperson	13	4.00	1.080	-1.876	4.784
Kontakt zur Bezugsperson	13	2.62	.659	.572	-.332
Aufmerksamkeit in Sekunden	13	915.92	415.41	-.376	-.867
Aufmerksamkeit: zurückholen	13	3.577	.672	-.327	-.895
Aufmerksamkeit: Buchseiten	13	.477	.378	.211	-1.273

Tabelle 24: Überprüfung der Normalverteilung mittels Schiefe und Kurtosis für die ältere Altersgruppe

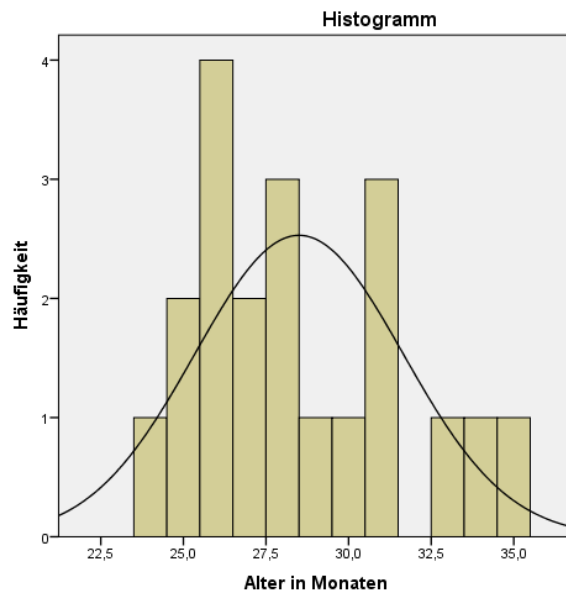
ÄLTERE ALTERSGRUPPE	N	MW	SD	Schiefe	Kurtosis
Aktive Farbdifferenzierung	7	4.000	2.236	-.877	.312
Passive Farbdifferenzierung	7	5.000	2.160	-.833	-1.200
Stabilität des Farbkonzepts	7	3.143	1.464	-2.122	4.735
Form- und Größendifferenzierung	7	3.000	1.155	-.909	-.150
Grobmotorik	7	2.071	.965	.864	-.118
Fein-/Visumotorik	7	3.214	2.531	.448	-.884
Feinmotorik	7	7.786	2.956	.275	-1.178
Visumotorik	7	5.107	1.391	-2.483	6.350
Selbstständigkeit	7	5.929	3.994	-.247	-.969
Lösung von der Bezugsperson	7	3.71	.951	-.863	1.245
Kontakt zur Bezugsperson	7	2.71	1.113	1.784	3.231
Aufmerksamkeit in Sekunden	7	1259.14	486.34	-.054	-1.426
Aufmerksamkeit: zurückholen	7	4.429	.450	.353	-1.817
Aufmerksamkeit: Buchseiten	7	.629	.390	-.465	-1.005

A.3. Überprüfung der Varianzhomogenität mittels Levene-Test

Tabelle 25: Überprüfung der Varianzhomogenität, getrennt nach Geschlecht und Alter

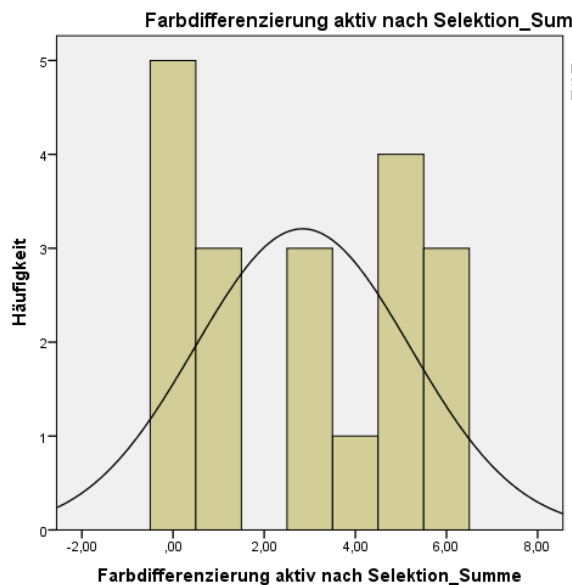
	Geschlecht		Alter		df
	F	p	F	p	
Aktive Farbdifferenzierung	.000	1.000	.469	.503	17
Passive Farbdifferenzierung	.715	.409	.013	.910	18
Stabilität des Farbkonzepts	1.321	.265	.955	.341	18
Form- und Größendifferenzierung	.237	.632	.004	.953	18
Grobmotorik	6.469	.022	.223	.643	16
Fein-/Visumotorik	14.423	.001	1.406	.251	18
Feinmotorik	1.772	.200	.156	.698	18
Visumotorik	2.134	.162	.908	.354	17
Selbstständigkeit	5.743	.028	1.101	.308	18
Lösung von der Bezugsperson	.210	.652	.054	.834	18
Kontakt zur Bezugsperson	2.531	.129	1.385	.255	18
Aufmerksamkeit in Sekunden	.235	.633	.254	.620	18
Aufmerksamkeit: zurückholen	4.105	.058	1.395	.253	18
Aufmerksamkeit: Buchseiten	1.221	.284	.008	.928	18

A.4. Überprüfung der Normalverteilung mittels Histogramme für eine Pearson-Korrelation



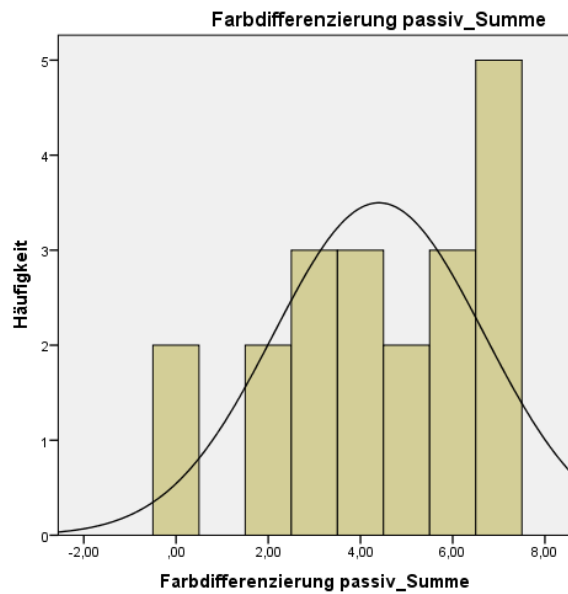
N = 20	Alter in Monaten
MW	28.50
SD	3.154
Schiefe	0.626
Kurtosis	-0.539

Abb. A.4.1. Histogramm mit der Normalverteilungskurve für das **Alter in Monaten** und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis



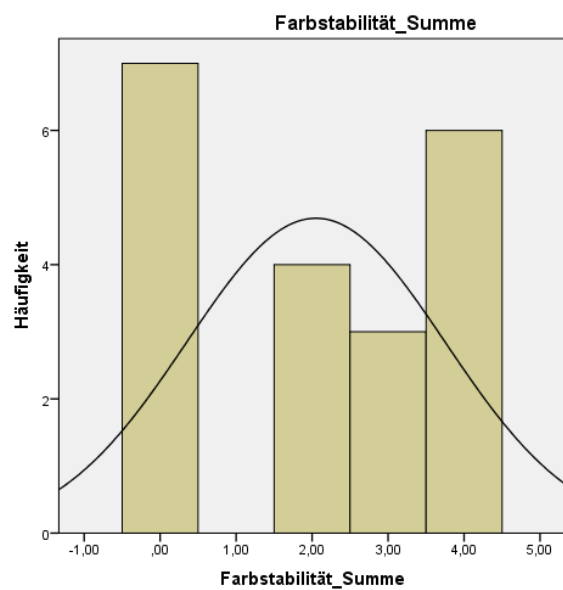
N = 20	Aktive Farbdifferenzierung
MW	2.842
SD	2.363
Schiefe	0.013
Kurtosis	-1.713

Abb. A.4.2. Histogramm mit der Normalverteilungskurve für die Skala „**Farbdifferenzierung aktiv**“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis



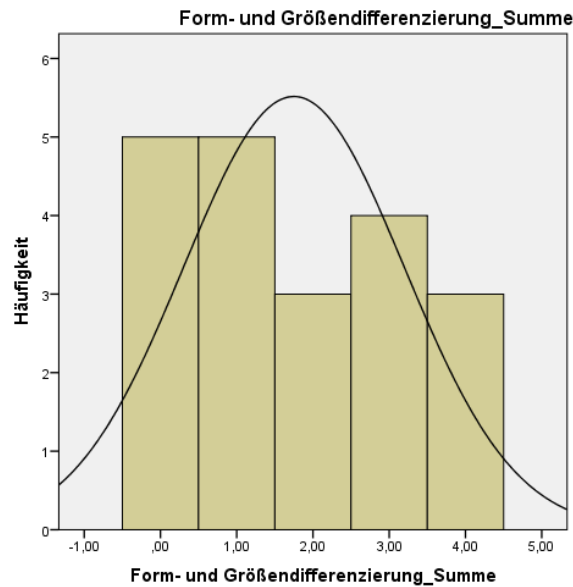
N = 20	Passive Farbdifferenzierung
MW	4.400
SD	2.280
Schiefe	-0.522
Kurtosis	-0.665

Abb. A.4.3. Histogramm mit der Normalverteilungskurve für die Skala „Farbdifferenzierung passiv“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis



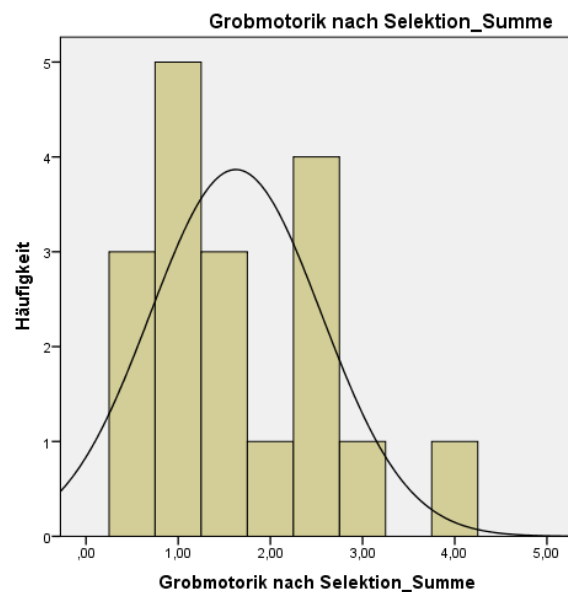
N = 20	Farbstabilität
MW	2.050
SD	1.701
Schiefe	-0.157
Kurtosis	-1.701

Abb. A.4.4. Histogramm mit der Normalverteilungskurve für die Skala „Farbstabilität“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis



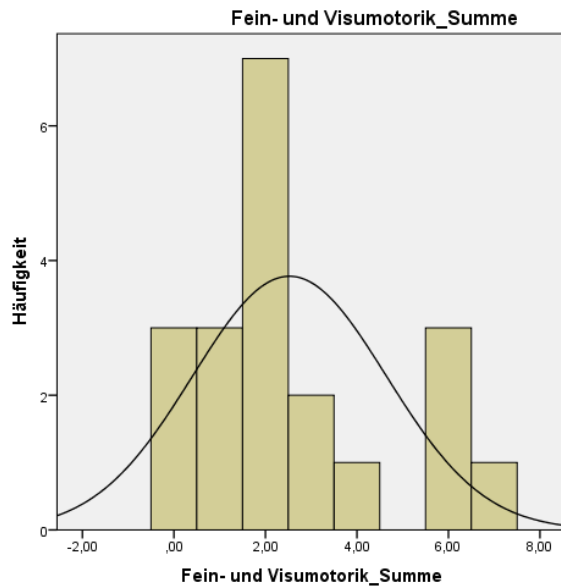
N = 20	Form- und Größendifferenzierung
MW	1.750
SD	1.446
Schiefe	0.254
Kurtosis	-1.309

Abb. A.4.5. Histogramm mit der Normalverteilungskurve für die Skala „Form- und Größendifferenzierung“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis



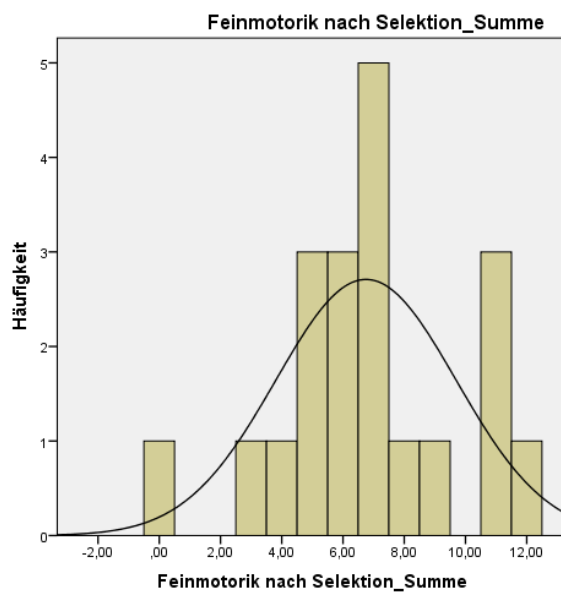
N = 20	Grobmotorik
MW	1.625
SD	0.929
Schiefe	0.678
Kurtosis	-0.240

Abb. A.4.6. Histogramm mit der Normalverteilungskurve für die Skala „Grobmotorik“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis



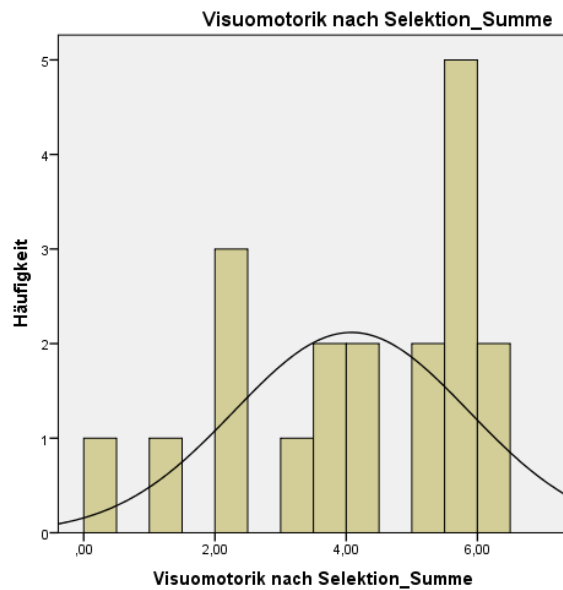
N = 20	Fein- und Visumotorik
MW	2.525
SD	2.118
Schiefe	0.963
Kurtosis	-0.028

Abb. A.4.7. Histogramm mit der Normalverteilungskurve für die Skala „**Fein- und Visumotorik**“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis



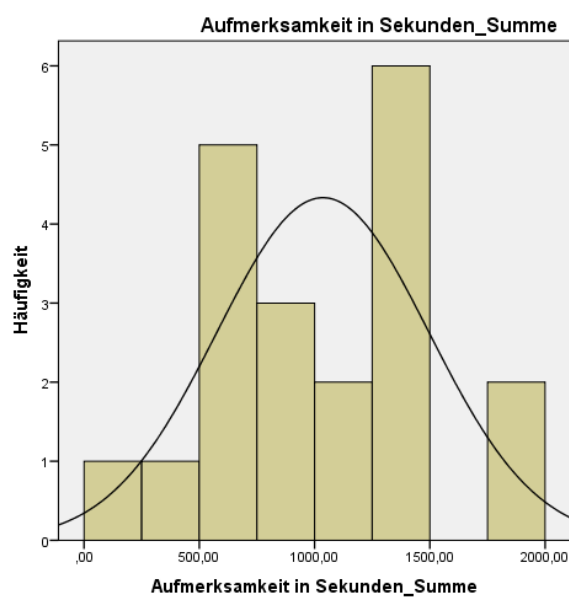
N = 20	Feinmotorik
MW	6.750
SD	2.945
Schiefe	-0.044
Kurtosis	0.410

Abb. A.4.8. Histogramm mit der Normalverteilungskurve für die Skala „**Feinmotorik**“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis



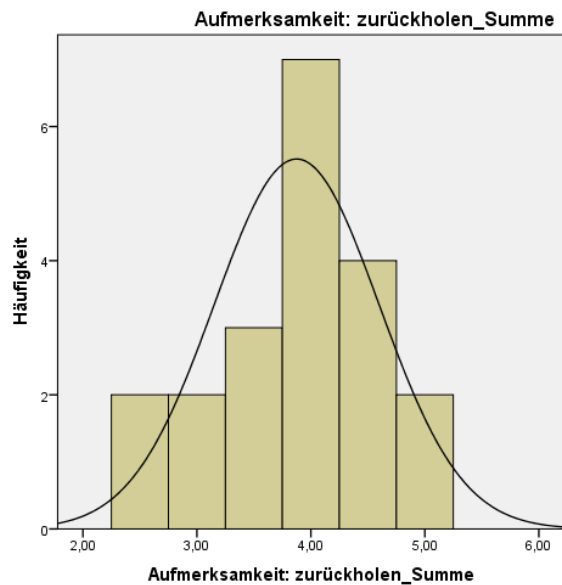
N = 20	Visumotorik
MW	4.079
SD	1.789
Schiefe	-0.692
Kurtosis	-0.672

Abb. A.4.9. Histogramm mit der Normalverteilungskurve für die Skala „**Visumotorik**“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis



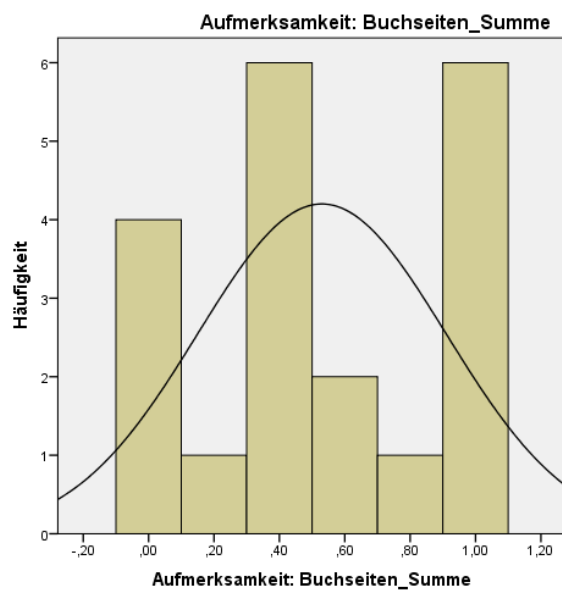
N = 20	Aufmerksamkeit: in Sekunden
MW	1036.050
SD	460.316
Schiefe	-0.025
Kurtosis	-0.533

Abb. A.4.10. Histogramm mit der Normalverteilungskurve für die Skala „**Aufmerksamkeit - Sekunden**“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis



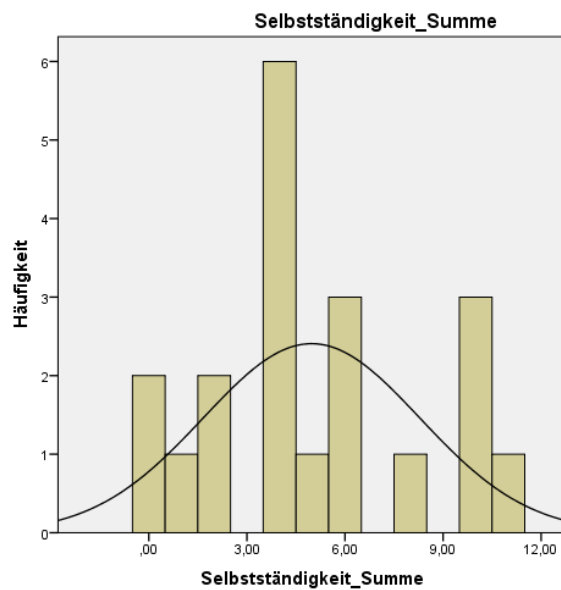
N = 20	Aufmerksamkeit: zurückholen
MW	3.875
SD	0.7232
Schiefe	-0.442
Kurtosis	-0.318

Abb. A.4.11. Histogramm mit der Normalverteilungskurve für die Skala „Aufmerksamkeit - zurückholen“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis



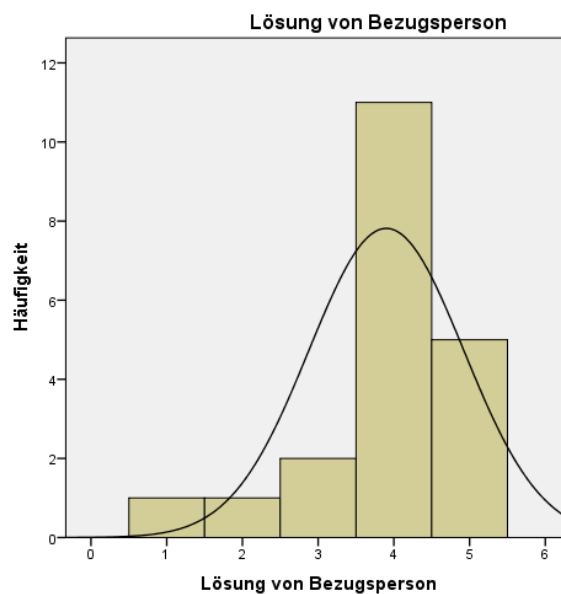
N = 20	Aufmerksamkeit: Buchseiten
MW	0.530
SD	0.380
Schiefe	-0.001
Kurtosis	-1.363

Abb. A.4.12. Histogramm mit der Normalverteilungskurve für die Skala „Aufmerksamkeit - Buchseiten“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis



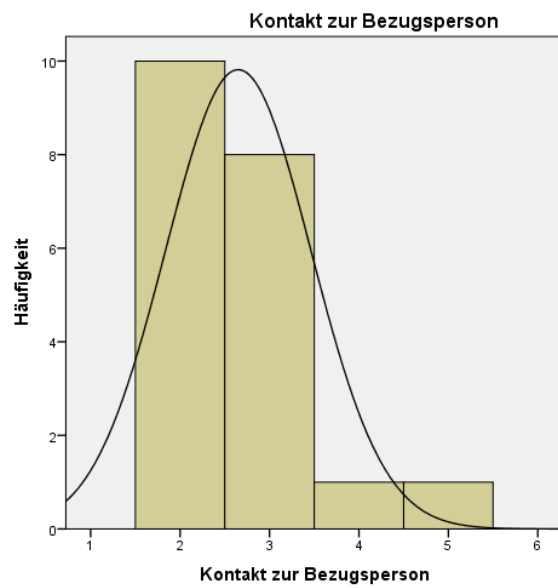
N = 20	Selbstständigkeit
MW	4.975
SD	3.315
Schiefe	0.401
Kurtosis	-0.601

Abb. A.4.13. Histogramm mit der Normalverteilungskurve für die Skala „Selbstständigkeit“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis



N = 20	Lösung von der Bezugsperson
MW	3.90
SD	1.021
Schiefe	-1.431
Kurtosis	2.607

Abb. A.4.14. Histogramm mit der Normalverteilungskurve für die Skala „Lösung von der Bezugsperson“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis



N = 20	Kontakt zur Bezugsperson
MW	2.65
SD	0.813
Schiefe	1.420
Kurtosis	2.376

Abb. A.4.15. Histogramm mit der Normalverteilungskurve für die Skala „**Kontakt zur Bezugsperson**“ und der Tabelle inklusive Mittelwert, Standardabweichung, Schiefe und Kurtosis

B PROTOKOLLBOGEN

Name des Testkinds: _____

Name des Testleiters: _____

Datum: _____

Uhrzeit (Anfang/Ende): _____

Geburtsdatum: _____

Alter: _____

Anwesende Personen: _____

Kontaktaufnahme zum Kind: _____

✓ Gelöst	X nicht gelöst
v. verweigert	- nicht vorgegeben

KIND-BEZUGSPERSON-BEZIEHUNG WORTSCHATZ PUPPENKÜCHE (OBST)

Ratingkala	Lösung von Bezugsperson (1- lange Zeit nicht möglich, 5- gleichgültig)				
	1	2	3	4	5
	Körperkontakt	mitspielen	Exploration	Verabschiedung	gleichgültig
	Kontakt zur Bezugsperson während Testung (1 – dauernd, 5 – gleichgültig)				
	1	2	3	4	5
	Körperkontakt	spielen	hingehen	Blickkontakt	gleichgültig

DESKRIPTIVE ANGABEN

Dauer	• Gesamt		
	• Anlaufzeit		
	• Beschäftigung mit Materialien		
	• Pause(n) – Anzahl		
	• Pause(n) – Minuten gesamt		
	• Verabschiedung		

Visuelle Wahrnehmung PUPPENKÜCHE (Obst sortieren nach Farbe)

Stabilität Farbkonzept Min 3	Rot		
	Grün		
	Gelb		
	orange		

Feinmotorik PUPPENKUCHE (0/1 + Anzahl aufschreiben)

Feinmotorik	Tellereinordnen (mind. 3)		
	Anzahl		
	Kochlöffel aufhängen (mind. 1)		
	Anzahl		
	Schraubverschlussglas öffnen		
	Schraubverschlussglas schließen		
	Aufkehren		
	Umfüllaufgabe		
	links & rechts umgefüllt (0/1)		
	nur mit einer Hand umgefüllt (0,5)		
	ohne daneben schütten		

Phonologisches Gedächtnis PUPPENKUCHE (0/1)

Gedächtnis	Gib mir A		
	Gib mir A + B		
	Gib mir A + B + C		

WORTSCHATZ (aktiv) PUPPENKUCHE (OBST) + Farbe benennen (0/1)

Wortschatz aktiv & Farbdifferenzierung aktiv	Obst benennen		
	Apfel		
	Birne		
	Banane		
	Zitrone		
	Karotte		
	Orange		
	Farbe benennen		
	rot		
	grün		
	gelb		
	orange		

Als ob Spiel - PUPPENKÜCHE

	Als-Ob-Spiel (1/0,5/0)		
	Kein Als-Ob-Spiel vorhanden (0)		
	Als-Ob-Spiel mit realen Dingen (0,5)		
	Als-Ob-Spiel ganzlich vorgestellt		
	Objekt anders verwendet (1)		

Numerisches Wissen - PUPPENKUCHE

Kognitive Entwicklung	Numerisches Wissen		
	Passive Mengenerfassung		
	Verständnis von 1: Gib mir 1...		
	Verständnis von 2: Gib mir 2...		
	Gibt eine Handvoll		
	gibt eines nach dem anderen		
	zählt offensichtlich		
	Verständnis von 3: gib mir 3		
	Gibt eine Handvoll		
	gibt eines nach dem anderen		
	zählt offensichtlich		
	Aktive Mengenerfassung		
	Wie viele sind das: Menge von 2		
	zählt offensichtlich		
	Wie viele sind das: Menge von 3		
	zählt offensichtlich		
	Zählen (0/1)		
	Anzahl Zählen		
	Zählen richtig (0/1)		
	Anzahl Zählen richtig		
	Eins zu eins Zuordnung (0/1)		
	Anzahl Eins zu eins Zuordnung		
	Stabile Zahlenfolge (0/1)		
	Anzahl Stabile Zahlenfolge		
	Richtige Reihenfolge (0/1)		
	Anzahl Richtige Reihenfolge		
	Kardinalsprinzip (0/1)		
	Anzahl Kardinalsprinzip		

VERKLEIDEN – Selbstständigkeit und Feinmotorische Fähigkeiten (0/1)

Selbstständigkeit	Schuhe anziehen		
	Versucht selbstständig		
	Sucht Hilfe		
	Desinteresse		
	Schuhe anziehen geschafft		
	Hose anziehen		
	Versucht selbstständig		
	Sucht Hilfe		
	Desinteresse		
	In Hosenbein reinschlüpfen		
	Hose hochziehen		
	Hut aufsetzen		
	Versucht selbstständig		
	Sucht Hilfe		
	Desinteresse		
	Aufsetzen geschafft		
	T-Shirt anziehen		
	Versucht selbstständig		
	Sucht Hilfe		
	Desinteresse		
	In Ärmel hineinschlüpfen		
	Mit Kopf durchschlüpfen		
	Shirt hinunterziehen		

TOM-BUCH (0/1 verbal oder nonverbal)

TOM-Buch	Aufmerksamkeit (0/0,5/1)		
	Dauer bis zur 1. Unaufmerksamkeit (s)		
	Lässt sich zurückholen		
	Seiten aufmerksam angesehen (gesamt)	S.	
	Küche benennen		
	schmutzige Hände		
	Torte/Käse (aktiv)		
	Besen passiv		
	Wohnzimmer benennen		
	Hungry		
	Lampe (aktiv)		
	Fernseher (passiv)		
	Tür (passiv)		
	Kinderzimmer benennen		
	Schlafen		
	Fliesen		
	Schuh (aktiv)		
	Stuhl/Sessel (passiv)		
	Gitterbett (passiv)		
	Schlafzimmer benennen		
	durstig		
	Kasten (aktiv)		
	Flasche (aktiv)		
	Fenster (passiv)		
	Bild (passiv)		
	Badezimmer (benennen)		
	Wasser (aktiv)		
	Badewanne (passiv)		

GROBMOTORIK STIEGE

Stiege	Hinaufgehen		
	mit anhalten (0,5)		
	ohne anhalten (1)		
	Nachstellschritt (0,5)		
	Wechselschritt (1)		
	Hinuntergehen		
	mit anhalten (0,5)		
	ohne anhalten (1)		
	Nachstellschritt (0,5)		
	Wechselschritt (1)		
	von letzter Stufe hüpfen (0/1)		

GROBMOTORIK BALL (MUSIK! 2 Meter 0/1)

Ballspielen (2 von 3)	Werfen		
	Anzahl geworfen		
	Fangen		
	Anzahl gefangen		
	Fußkick		

GROBMOTORIK STEHEN, HÜPFEN, LAUFEN (1/0)

	Schmaler Weg		
	Vorwärts Balancieren-Linie		
	Seitwärts Gehen-Linie		
	Zehenspitzen-Gang-(Linie egal)		
	Beidbeiniges Hüpfen (3 Mal)		
	Stehen bleiben aus vollem Lauf		
	1. Versuch:		
	Sofort		
	2 Schritte		
	fällt um		
	2. Versuch		
	Sofort		
	2 Schritte		
	fällt um		
	3. Versuch		
	Sofort		
	2 Schritte		
	fällt um		

FEINMOTORIK & FARBDIFFERENZIERUNG & VISUMOTORIK (1/0)

Zeichnen	Aufmerksamkeit (0/0,5/1)		
	Dauer bis 1. Unterbrechung		
	lässt sich wieder zurückführen		
	Stifthaltung:		
	Primitive Formen (0,5)		
	Erwachsene Stifthaltung (1)		
	Unterarm ruht auf Tisch		
	Hand zum Blatt halten verwendet		
	Linien zeichnen:		
	Horizontal		
	Vertikal		
	Kreis		
	Farben erkennen passiv (Stifte)		
	Grün		
	Gelb		
	Rot		
	Orange		
	Farben benennen aktiv (Stifte):		
	Blau		
	Weiß		
	Lila/violett		
	Blatt mit Schere schneiden (mind 2)		

VISUMOTORIK – Formen

Turm	Aufmerksamkeit (0/0,5/1)		
	Dauer bis 1. Unterbrechung		
	lässt sich wieder zurückführen		
	Vertikal		
	Anzahl höchster Turm		
	Horizontal		
	Mauer		
	3D		
	Anzahl Bausteine gesamt		

Skizze von Turm

VISUMOTORIK - TISCHAUFGABEN

	Perlen fädeln (min 2)		
	Anzahl Perlen		
	Vorgezeigt		
	Aufmerksamkeit (0/0,5/1)		
	Dauer bis zur ersten Unterbrechung		
	Ließ sich zurückholen		
	Perlen Farben passiv – GIB MIR		
	blau		
	lila		
	weiß		
	Puzzle		
	Anzahl Puzzleteile (max 6)		
	Aufmerksamkeit (0/0,5/1)		
	Dauer bis zur ersten Unterbrechung		
	Ließ sich zurückholen		

VISUELLES GEDÄCHTNIS - TISCHAUFGABEN

Visuelles Ged.	Memory Sonne		
	Memory Haus		
	Memory Baum		
	Teddybär		
	Haus		

FARB-, FORM- UND GRÖßENDIFF. (BODENMATTEN)

	Form- und Größendifferenzierung		
	Formdiff. 1 (Kreis, Quadrat, Dreieck, Rechteck)		
	Formdiff. 2 (Vierecke)		
	Größendiff. 1 (Rechtecke)		
	Größendiff. 2 (Kreise)		

SPRACHENTWICKLUNG

Satzlänge	Einwortäußerungen		
	Zweiwortäußerungen		
	Dreiwortäußerungen		
	Mehrwortäußerungen		
	Satz korrekt (S, P, O)		
Haupt- und Nebensätze	Vorkommen von Verbindungen		
	Verbindung mit und/aber		
	Verbindung mit weil		
	Verbindungen mit dass		
	Verbindungen mit wenn		
	Verbindungen mit ob		
	Relativsätze		
	- mit korrektem Relativpronomen - mit wo - ohne Relativpronomen		
Verben	Stellung im Aussagesatz		
	Endstellung		
	Verbstellung korrekt		
	Partizip Perfekt (kommt vor)		
	ohne ge- gebildet		
	mit ge- gebildet		
	Korrekte Partizipendung		
	mit Hilfsverben (haben+sein)		
Verneinung	Verneinung		
	Korrekte Wortstellung		
	Bildung		
	mit nein		

	mit nicht		
Fragen	Fragenintonation		
	Ja/Nein-Fragen		
	Inversion von Subjekt und Prädikat		
	W-Fragen		
	Inversion von Subjekt und Prädikat		
	Fragewort vorhanden (außer wo?)		
	Wo? Vorhanden		
	Alternativfragen		
	Wortstellung korrekt		

C ELTERNFRAGEBOGEN

Datum: _____

Elternfragebogen

Angaben zum Kind:

Vor- und Zuname Ihres Kindes: _____

Geschlecht: ☐ männlich ☐ weiblich Geburtsdatum: _____

Besucht Ihr Kind zurzeit eine/n Kindergarten/Krippe? ☐ Ja ☐ Nein

Wenn ja: Seit wann: _____ Stunden pro
Woche: _____

Besucht Ihr Kind Sport- oder Freizeitkurse? ☐ Ja ☐ Nein

Wenn ja, welche: _____

Angaben zur Familie des Kindes:

Eltern:

	Name	höchste abgeschlossene Ausbildung	Alter	gemeinsamer Haushalt
Vater:				☐ Ja ☐ Nein
Mutter:				☐ Ja ☐ Nein

Geschwister:

Name	Alter	Geschlecht	gemeinsamer Haushalt
		☐ M ☐ W	☐ Ja ☐ Nein
		☐ M ☐ W	☐ Ja ☐ Nein
		☐ M ☐ W	☐ Ja ☐ Nein
		☐ M ☐ W	☐ Ja ☐ Nein
		☐ M ☐ W	☐ Ja ☐ Nein

Wächst Ihr Kind mehrsprachig auf? ☐ Ja ☐ Nein

Wenn ja, welche Sprachen: _____

Muttersprache: _____

Angaben zu Geburt/Schwangerschaft/Gesundheit:

Risikoschwangerschaft: ☐ Ja ☐ Nein

Geburt in wievielter Woche: _____

Leidet Ihr Kind unter gesundheitlichen Beeinträchtigungen? ☐ Ja ☐ Nein

Wenn ja, welche: _____

Angaben zur Selbständigkeit:

Bitte beantworten Sie die folgenden Fragen nach Ihrer geschätzten Häufigkeit des Auftretens anhand der Zahlenkategorien 1 – 4.

	nie	manchmal	oft	immer
Mein Kind braucht tagsüber Windeln.	1	2	3	4
Mein Kind braucht in der Nacht Windeln.	1	2	3	4
Mein Kind sagt mir Bescheid, wenn es auf die Toilette gehen will.	1	2	3	4
Mein Kind kann sich 15 Minuten alleine beschäftigen	1	2	3	4
Mein Kind kann alleine ein Glas halten und daraus trinken.	1	2	3	4
Mein Kind teilt mir mit, wenn es Hunger oder Durst hat.	1	2	3	4
Mein Kind kann alleine mit einem Löffel essen	1	2	3	4
Mein Kind kann alleine mit einer Gabel essen.	1	2	3	4
Mein Kind geht alleine die Stiegen hinauf				
mit Festhalten	1	2	3	4
ohne Festhalten	1	2	3	4
Mein Kind geht alleine die Stiegen hinunter				
mit Festhalten	1	2	3	4
ohne Festhalten	1	2	3	4
Mein Kind sagt, wenn es etwas möchte.	1	2	3	4
Mein Kind sagt, wenn es etwas nicht möchte.	1	2	3	4
Mein Kind kann sich alleine anziehen.	1	2	3	4
Mein Kind kann den rechten vom linken Schuh unterscheiden.	1	2	3	4

Angaben zum Sozial- und Spielverhalten:

	nie	manchmal	oft	immer
Mein Kind spielt gerne mit anderen Kindern.	1	2	3	4
Mein Kind bevorzugt seine Eltern/Bezugsperson als Spielpartner.	1	2	3	4
Mein Kind kann einen Ball fangen.	1	2	3	4
Mein Kind kann einen Ball werfen.	1	2	3	4
Mein Kind turnt/klettert gerne.	1	2	3	4
Mein Kind kann im Spiel einfache Regeln befolgen.	1	2	3	4
Mein Kind spielt mit Konstruktionsspielzeug. (z.B.: Bauklötze, Duplo, Lego, ...)	1	2	3	4
Meinem Kind fallen viele Dinge ein, die es gerne spielen möchte.	1	2	3	4
Mein Kind imitiert im Spiel alltägliche Handlungen von Erwachsenen.	1	2	3	4
Mein Kind spielt Rollenspiele.	1	2	3	4

Angaben zur Sprache:

Hier finden Sie eine Wortliste. Bitte kreuzen Sie jene Wörter an, die Sie schon öfters von Ihrem Kind gehört haben. Bitte beachten Sie, dass Sie nur Wörter ankreuzen, die Ihr Kind selbst verwendet. Dazu zählen auch Wörter, die es etwas anders ausspricht (z.B.: „Nie“ statt „Knie“). Falls Ihr Kind für etwas ein anderes Wort benutzt, schreiben Sie es bitte daneben (z.B.: „heihei“ statt „schlafen“).

- | | | |
|------------------------------------|----------------------------------|-------------------------------------|
| ☐ schmutzig | ☐ Auto | ☐ Fenster |
| ☐ Torte | ☐ Apfel | ☐ Bild |
| ☐ Käse | ☐ Birne | ☐ Badezimmer |
| ☐ Besen | ☐ Banane | ☐ Wasser |
| ☐ hungrig | ☐ Zitrone | ☐ Badewanne |
| ☐ Lampe | ☐ Karotte | ☐ Katze |
| ☐ Fernseher | ☐ Orange | ☐ Besteck |
| ☐ Tür | ☐ Sessel | ☐ Schlaf |

- | | | |
|-----------------------------------|---------------------------------------|----------------------------------|
| ☐ schlafen | ☐ Gitterbett | ☐ Fahrrad |
| ☐ Pflaster | ☐ Schlafzimmer | ☐ fliegen |
| ☐ Bub | ☐ durstig | ☐ Elefant |
| ☐ Knie | ☐ Mädchen | ☐ Hase |
| ☐ Mann | ☐ Kasten | ☐ Eis |
| ☐ Schuh | ☐ Tasche | ☐ Hund |

Mein Kind hat bereits begonnen, zwei Wörter miteinander zu verbinden.	☐ Ja ☐ Nein
Mein Kind hat bereits begonnen, drei oder mehrere Wörter miteinander zu verbinden.	☐ Ja ☐ Nein
Mein Kind verwendet andere Fragewörter außer „Wo“.	☐ Ja ☐ Nein
Mein Kind verwendet bereits die Vergangenheitsform mit den Hilfsverben „haben“ und „sein“.	☐ Ja ☐ Nein
Mein Kind verwendet die Vergangenheitsform schon richtig (z.B.: gegessen, weh getan, ...).	☐ Ja ☐ Nein
Mein Kind verwendet bei der Verneinung das Wort „nicht“. (z.B.: nicht schlafen, nicht Zähneputzen, ...).	☐ Ja ☐ Nein

Vielen Dank für Ihre Mitarbeit!

D EINVERSTÄNDNISERKLÄRUNG FÜR DIE VIDEOAUFZEICHNUNG

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 Diplomarbeitsbetreuerinnen, Ass.-Prof. Dr. Pia Deimann und Ass.-Prof. Dr. Ursula Kastner-Koller, niemand das Video sehen wird und dieses spätestens nach Beendigung unserer Diplomarbeiten gelöscht werden.

Sigrid Punz und Viktoria Kronberger

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

Datum

Unterschrift

LIEBE ELTERN ALLER ZWEI- BIS DREIJÄHRIGEN KINDER!

Im Rahmen unserer Diplomarbeiten sind wir, zwei Studentinnen der Psychologie, auf der Suche nach Eltern mit Kindern, die uns bei unserer Studie unterstützen möchten.

Worum es geht:

Am Institut für Entwicklungspsychologie und Psychologische Diagnostik der Universität Wien ist es Ziel, ein spielbasiertes Verfahren zur Erfassung der Entwicklung zwei- bis dreijähriger Kinder zu entwickeln. Wir wollen Informationen sammeln, welches Verhalten Zweijährige im Spiel üblicherweise zeigen.

Was passiert?

Wir spielen, lesen, turnen und lachen circa zwei Stunden lang mit Ihrem Kind in einem Testraum der Uni Wien. Sie können in Ruhe entspannen und zusehen.

Wann?

Wir würden gerne mit Ihnen und Ihrem Kind für Februar 2013 einen Termin vereinbaren.

Welchen Vorteil haben Sie davon:

Sie haben die Möglichkeit zu erfahren, welche Fähigkeiten eines zweijährigen Kindes als besonders wichtig gelten und wie diese Fähigkeiten erfasst werden.



Wenn Sie Interesse an einem spannenden Vormittag oder Nachmittag haben und zusätzlich noch einen Beitrag zu unserem Projekt leisten wollen, dann freuen wir uns sehr Sie und Ihr Kind kennenzulernen!

Für alle Fragen stehen wir Ihnen jederzeit sehr gerne telefonisch oder per E-Mail zur Verfügung!

SIGRID PUNZ

a0705400@unet.univie.ac.at
0664/6523819



VIKTORIA KRONBERGER

a0703387@unet.univie.ac.at
0699/10106691



F CURRICULUM VITAE

Viktoria Magdalena Kronberger

Geburtsdatum: 24.05.1989,

Geburtsort: Zwettl

Staatsangehörigkeit: Österreich

Schulausbildung

<i>Juni 2007</i>	Matura mit gutem Erfolg
<i>1999-2007</i>	Bundesgymnasium Zwettl (Sprachlicher/Humanistischer Zweig)
<i>1995-1999</i>	Volksschule Ottenschlag

Berufsausbildung

<i>seit März 2008</i>	Diplomstudium Psychologie an der Universität Wien mit Schwerpunkt der Kinder- und Jugendpsychologie
<i>Oktober 2007 - Jänner 2008</i>	Bachelorstudium Bildungswissenschaften an der Universität Wien

Studienrelevante Tätigkeiten

<i>Juli 2013</i>	Praktikum im Ambulatorium für ganzheitliche Förderung und Entwicklungsdiagnostik in Zwettl (NÖ)
<i>Juni 2013</i>	Mitarbeit als Testleiterin an der Normierung der AID-Gruppe, Wien
<i>Jänner 2013</i>	Teilnahme am Symposium „Aufmerksamkeit bei Kindern und Jugendlichen“, Wien

<i>Oktober 2012 – Jänner 2013</i>	Praktikum im Ambulatorium für ganzheitliche Förderung und Entwicklungsdiagnostik in Zwettl (NÖ)
<i>September 2012</i>	Praktikum in der Österreichischen Autistenhilfe, Wien
<i>August 2012</i>	Praktikum in der Rehabilitationsklinik Gars am Kamp (NÖ)
<i>April 2012</i>	Teilnahme als Testleiterin am Projekt „Bewegte Klasse“ der Donauuniversität Krems an der Donau (NÖ)
<i>September 2011 – Juni 2012</i>	Praktikum in der Österreichischen Autistenhilfe, Wien
<i>2007-2012</i>	Nachhilfe für Kinder zwischen 11 und 14 Jahren

Berufserfahrung

<i>seit April 2012</i>	Ordinationsassistentin in einer Gruppenarztpraxis für Kinder- und Jugendheilkunde im 22.Bezirk (DDr. Peter Voitl)
<i>Februar 2009 - März 2012</i>	Ordinationsassistentin in einer Praxis für Innere Medizin im 7.Bezirk (Dr. Uta Kopetzki-Kogler)