

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

Umsetzung des Lerntestkonzepts von Guthke beim AID 3

verfasst von / submitted by

Theresa Sabo, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2018 / Vienna, 2018

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on
the student record sheet:

A 066 840

Studienrichtung lt. Studienblatt /
degree programme as it appears on
the student record sheet:

Masterstudium Psychologie

Betreut von / Supervisor:

Univ.- Prof. i. R. Mag. Dr. Klaus Kubinger

Danksagung

An erster Stelle möchte ich mich bei meinem Betreuer, Herrn Univ.-Prof. Dr. Mag. Klaus Kubinger, für sein entgegengebrachtes Vertrauen bedanken, dass ich mich mit dem ausgewählten Thema im Rahmen meiner Masterarbeit näher befassen konnte. Danken möchte ich Ihnen für das konstruktive Feedback und für die stets rasche Beantwortung jeglicher Fragen.

Für die Hilfe bei der Durchführung der Testungen und für den fachlichen Austausch möchte ich mich bei meiner Kommilitonin Theresa Renz bedanken.

Ein Dankeschön gebührt auch den Direktorinnen, sowie den LehrerInnen der beiden Schulen, die diese Untersuchung erst ermöglicht und bei der Organisation der Testungen mitgeholfen haben.

Von ganzem Herzen möchte ich mich bei meiner Familie bedanken, die mich während des Studiums nicht nur finanziell, sondern auch emotional unterstützt hat und dabei sehr viel Geduld und Verständnis für mich aufgebracht hat.

Abschließend möchte ich mich auch besonders bei meinem Freund Mogue dafür bedanken, dass du an mich glaubst, mich immer motivierst und mich dabei unterstützt meine Ziele und Träume zu erreichen.

Inhaltsverzeichnis

1. Einleitung	3
2. Theoretischer Hintergrund	4
2.1 Kritik an der herkömmlichen Intelligenzdiagnostik	4
2.2 Statusdiagnostik versus dynamische Testdiagnostik	5
2.3 Historischer Umriss der dynamischen Testdiagnostik	7
2.4 Das Lerntestkonzept nach Guthke	9
2.5 Dynamisches Testen in der Praxis	11
2.6 Vorteile des dynamischen Testens	13
3. Ziel der Untersuchung und Hypothesen	16
4. Entwicklung des Lerntests am Untertest „Formale Folgerichtigkeit“	18
4.1 Kurzbeschreibung des Untertests „Formale Folgerichtigkeit“ des AID3	18
4.2 Allgemeine Informationen zum Lerntest	18
4.3 Das Hilfesystem des Lerntests	19
5. Methode	24
5.1 Untersuchungsdesign	24
5.2 Darstellung der Stichprobe	24
5.3 Durchführung und Ablauf der Untersuchung	25
5.4 Operationalisierung und Messinstrumente	25
5.5 Beschreibung der statistischen Analyse	26
6. Ergebnisse	28
6.1 Deskriptivstatistische Beschreibung der Stichprobe	28
6.2 Explorative Datenanalyse	28
6.3 Statistische Auswertung der Hypothesen	29
7. Diskussion	35
8. Literaturverzeichnis	39
9. Abbildungsverzeichnis	44
10. Tabellenverzeichnis	45
11. Anhang	46
11.1 Anhang A- Abstract (Deutsch)	46
11.2 Anhang B- Abstract (English)	47
11.3 Anhang C- Einverständniserklärung der Eltern	48

1. Einleitung

Die vorliegende Masterarbeit beschäftigt sich damit, für den Untertest „Formale Folgerichtigkeit“ des Adaptiven Intelligenz Diagnostikums (AID 3, Kubinger & Holocher-Ertl, 2014), einen Lerntest nach dem Lerntestkonzept von Guthke (Guthke & Wiedl, 1996) zu entwickeln. Im Anschluss an die Entwicklung des Lerntests soll an einer Stichprobe, bestehend aus Schulkindern, geprüft werden, ob der Lerntest in der Lage ist die Lernfähigkeit zu erfassen und ob damit eine zusätzliche diagnostische Information erhoben werden kann, die der standardmäßige Untertest „Formale Folgerichtigkeit“ nicht erfasst.

Nachdem die Forschung im Bereich des dynamischen Testens und die Entwicklung von Lerntests in den letzten Jahrzehnten des 20. Jahrhunderts einen Aufschwung erlebt haben und die dynamische Testdiagnostik Bekanntheit in der breiteren psychologischen Fachwelt erlangt (Guthke & Wiedl, 1996; Haywood & Tzuriel, 2002), so ging diese Phase unter anderem nach dem Tod von Guthke zu Ende, weil sich niemand mehr um die Erhaltung der Lerntests kümmerte (Kubinger, 2009). Da das dynamische Testen einige Vorteile bietet und eine gute Ergänzung zur herkömmlichen Intelligenzdiagnostik darstellen kann, soll beim AID 3 das Lerntestkonzept erprobt werden. Falls sich zeigt, dass der Lerntest in der Lage ist die Lernfähigkeit zu erfassen und es sich lohnt diesen in den AID 3 zu integrieren, könnte angedacht werden den AID 3, um eine weitere attraktive Komponente für Diagnostiker zu erweitern.

Zum Aufbau dieser Arbeit ist zu sagen, dass zu Beginn des Theorieteils zunächst die grundlegenden Begrifflichkeiten des dynamischen Testens erklärt werden. Anschließend soll sowohl die historische Entwicklung der dynamischen Testdiagnostik als auch das Lerntestkonzept vorgestellt werden. Auch wird näher auf die Gründe eingegangen, warum sich die dynamische Testdiagnostik nicht ausreichend in der Praxis etablieren konnte. Im letzten Abschnitt des Theorieteils werden der Nutzen und die Vorteile des dynamischen Testens anhand von empirischen Erkenntnissen diskutiert.

Im empirischen Teil der Arbeit werden zuallererst die Fragestellungen der Untersuchung und die dazugehörigen Hypothesen vorgestellt. Danach werden die Entwicklung des Lerntests und das dafür konzipierte Hilfesystem ausführlich beschrieben. Es folgt die Methodenbeschreibung, die Präsentation der Ergebnisse und zum Schluss die Diskussion.

2. Theoretischer Hintergrund

2.1 Kritik an der herkömmlichen Intelligenzdiagnostik

Seit langer Zeit werden in der psychologischen Diagnostik wiederkehrend Diskussionen darüber geführt, ob die herkömmlichen Intelligenztests in der Lage sind, die intellektuellen Fähigkeiten von untersuchten Personen angemessen und umfassend abzubilden (Guthke & Wiedl, 1996). Traditionelle Intelligenztests messen im Allgemeinen die Produkte früheren Lernens und gründen dabei auf der Annahme, dass die Testpersonen in der Vergangenheit unter den gleichen Bedingungen die Möglichkeit hatten, ihre Fähigkeiten zu entwickeln (De Beer, 2005; Haywood, Brown, & Wingenfeld, 1990). Erkenntnissen wie, dass kognitive Fähigkeiten einer Veränderbarkeit und dem Einfluss anderer Variablen unterliegen, werden in der herkömmlichen Intelligenzdiagnostik oftmals keine große Aufmerksamkeit geschenkt (Haywood et al., 1990; Grigorenko & Sternberg, 1998).

Dass beispielsweise soziale Variablen einen Einfluss auf die gezeigten Testleistungen haben können, wie die Menge an Lernanregungen, die Testpersonen in ihrer Vergangenheit erfahren haben, wird kaum in der traditionellen Intelligenzdiagnostik mitberücksichtigt (Grigorenko & Sternberg, 1998). Problematisch wird dieser Umstand, wenn die zu untersuchenden Personen aus einem sozioökonomisch schwachen Umfeld kommen und/oder unter defizitären Lernbedingungen aufwachsen (Guthke & Wiedl, 1996). So kann es vorkommen, dass diese Personen bei den herkömmlichen Intelligenztests benachteiligt werden, da lediglich ihre bereits entwickelten Fähigkeiten gemessen werden, aber ihr möglicherweise vorhandenes Potential verschleiert bleibt (Murphy & Maree, 2006).

Auch wird immer wieder darüber diskutiert, ob zuverlässige Aussagen über psychische Eigenschaften getroffen werden können, wenn diese wie üblich nur durch einmalige Testungen erfasst werden (Guthke & Wiedl, 1996). Aus diesen Überlegungen und Diskussionen heraus, haben sich in den letzten Jahrzehnten alternative und ergänzende Ansätze zur herkömmlichen statusorientierten Intelligenzdiagnostik entwickelt, die unter dem Begriff der dynamischen Testdiagnostik zusammengefasst werden können (Guthke & Wiedl, 1996).

Anzumerken ist, dass mit der angeführten Kritik nicht behauptet werden soll, dass die herkömmliche Intelligenzdiagnostik „keinen guten Job macht“, indem was sie misst und wie sie es misst. Im Gegenteil, die traditionelle Intelligenzdiagnostik ist in der Lage die bereits entwickelten Fähigkeiten einer Person gut abzubilden, diese in Beziehung zu anderen Testleistungen zu setzen und die zukünftigen Leistungen von Personen, mit einer „normalen“ Lerngeschichte, gut vorherzusagen (Caffrey, D. Fuchs, & L. S. Fuchs, 2008; Haywood et al.,

1990). Ebenso erachten es die Vertreter der dynamischen Testdiagnostik als wichtig die bereits entwickelten Fähigkeiten zu messen, jedoch begreifen sie diese Information lediglich als einen Teil, der zu einem kompletten Abbild der kognitiven Fähigkeiten einer Person beiträgt (Lidz, 1996). Als Ersatz zur herkömmlichen Intelligenzdiagnostik wird das dynamische Testen also nicht verstanden, sondern als eine Ergänzung dazu (Grigorenko, 2009; Kirschenbaum, 1998). Optimal wäre es, wenn beide Testansätze in Kombination miteinander realisiert werden (Grigorenko, 2009; Kirschenbaum, 1998).

Im nächsten Abschnitt sollen nun die grundlegenden Unterschiede zwischen der dynamischen Testdiagnostik und der herkömmlichen statusorientierten Intelligenzdiagnostik diskutiert und die Merkmale der Dynamischen Testdiagnostik für ein tiefergehendes Verständnis skizziert werden.

2.2 Statusdiagnostik versus dynamische Testdiagnostik

Die Statusdiagnostik und die dynamische Testdiagnostik werden in der Literatur als zwei zueinander konträre Ansätze verstanden, die sich durch ein unterschiedliches Verständnis, wie kognitive Fähigkeiten erfasst und gemessen werden sollten, auszeichnen (Tzuriel, 2001). Im Allgemeinen unterscheiden sich die beiden in ihren Zielen, in ihren Prozessen, in ihren Testinstrumenten, in ihrer Testsituation und ihrer Interpretation von Ergebnissen, voneinander (Grigorenko & Sternberg, 1998; Tzuriel, 2000).

Die herkömmlichen Intelligenztests sind durch einen statusorientierten Ansatz geprägt und haben vor allem zum Ziel, Aussagen über den psychologischen Ist-Zustand hinsichtlich zeit- und situationsstabiler psychischer Merkmale zu treffen (Beckmann, 2003). Dabei liegt der Fokus darauf, zu untersuchen, wie sich die Testwerte verschiedener Personen untereinander unterscheiden und dabei die Testleistungen zu klassifizieren (Haywood et al., 1990; Schack & Guthke, 2003). Interessiert ist die statusorientierte Intelligenzdiagnostik auch daran, in welchem Ausmaß sich die kognitiven Fähigkeiten von Personen entwickelt haben und dabei die Stärken und die Schwächen einer Person zu identifizieren (Kirschenbaum, 1998). Gegenstand der Statusdiagnostik ist es aber im Allgemeinen nicht, die intraindividuelle Variabilität eines psychischen Merkmals einer Person zu erfassen, da diese intraindividuellen Unterschiede meist als Störgröße betrachtet werden (Guthke & Wiedl, 1996).

Im Gegensatz dazu, hat sich die dynamische Testdiagnostik genau dies zum Ziel gemacht, nämlich Veränderungen im Testprozess hervorzurufen, um so zu validen Aussagen über die kognitiven Fähigkeiten und deren Veränderbarkeit zu gelangen (Schreiber, 2002). Dabei versucht die dynamische Testdiagnostik nicht nur den aktuellen Zustand eines psychischen Merkmals festzustellen, sondern ist vor allem an dem Niveau interessiert, dass

gegenwärtig noch nicht sichtbar ist, aber möglicherweise vorhanden (Schack & Guthke, 2003). Um dieses Niveau sichtbar zu machen, bedienen sich dynamische Testansätze einer Art Prozessdiagnostik, mit welcher Veränderungsmessung möglich gemacht wird und so der Lernprozess selbst untersucht wird, anstatt wie bei der Statusdiagnostik nur das Lernprodukt (Beckmann, 2003; Caffrey et al., 2008; Ruijsenaars, Castelijns, & Hamers, 1993). An dieser Stelle ist zu erwähnen, dass die Voraussetzung für die Prozessdiagnostik, die Statusdiagnostik selbst ist, da eine Veränderung eines psychischen Merkmales nur über die Erfassung des Status zu zwei verschiedenen Messzeitpunkten realisiert werden kann (Beckmann, 2003).

Passend zu den bereits genannten Punkten, schlagen Guthke und Wiedl (1996, S.8) folgende Definition für den Begriff des Dynamischen Testens vor: „Dynamische Testdiagnostik ist ein Sammelbegriff für testdiagnostische Ansätze, die über die gezielte Evozierung und Erfassung der intraindividuellen Variabilität im Testprozess entweder auf eine validere Erfassung des aktuellen Standes eines psychischen Merkmals und/oder seiner Veränderbarkeit abzielen.“

Aus der Definition von Guthke und Wiedl (1996) lassen sich zwei Zielbereiche des dynamischen Testens ableiten. Der Zielbereich Eins bezieht sich darauf ein psychisches Merkmal valider zu erfassen und dabei das Höchstmaß der aktuellen geistigen Leistungsfähigkeit zu ermitteln (Guthke & Wiedl, 1996). Hingegen wird vom Zielbereich Zwei gesprochen, wenn es darum geht die Veränderbarkeit eines Merkmals zu erfassen und diese entsprechend einer „neuen“ Eigenschaft definiert wird. Im Rahmen der dynamischen Intelligenzdiagnostik wird diese „neue“ Eigenschaft vielfach als Lernfähigkeit oder Lernpotential bezeichnet (Guthke & Wiedl, 1996). Es geht also entweder darum Intelligenz mithilfe eines dynamischen Zugangs umfassender zu untersuchen oder die Messung einer neuen Eigenschaft anzustreben (Kubinger, 2009). Anzumerken ist, dass es zwischen den beiden Zielbereichen einen fließenden Übergang gibt und diese praktisch nicht voneinander zu trennen sind (Guthke & Wiedl, 1996).

Wenn man die statusorientierte Testdiagnostik und die dynamische Testdiagnostik bezüglich ihrer Verfahren miteinander vergleicht, werden sehr konkrete Unterschiede zwischen diesen beiden sichtbar. Bei statusorientierten Verfahren hat der/die TestleiterIn die Funktion, den Testpersonen die Aufgaben nacheinander zu präsentieren als auch die Antworten zu dokumentieren, ohne dabei in irgendeiner Art und Weise während der Testung zu intervenieren oder einen Einfluss auf die Testsituation zu nehmen (Tzuriel, 2000). Anders gestaltet sich dies bei den dynamischen Verfahren, bei welchen der/die TestleiterIn den Testpersonen gezielt Anregungen in Form von Feedback, Hinweisen und Hilfestellungen gibt,

um entweder das Optimum der aktuellen geistigen Leistungsfähigkeit oder die Lernfähigkeit/das Lernpotential einer Person zu erfassen (Tzuriel, 2000). Hierbei nimmt der/die TestleiterIn bei der Durchführung eine eher aktive Rolle ein, anstatt einer passiven Rolle, wie es bei statusorientierten Verfahren der Fall ist (Lidz, 1995). Des Weiteren zeichnet sich die Beziehung zwischen dem/der TestleiterIn und der Testperson beim dynamischen Testen nicht so stark durch Neutralität aus, wie bei der Statusdiagnostik, sondern durch einen eher interaktiven, lehrhaften und helfenden Charakter (Grigorenko & Sternberg, 1998). Bei der Durchführung kann es, je nach dynamischem Verfahren, variieren, ob die Testpersonen individualisiertes Feedback oder standardisiertes Feedback, welches aus einer Hierarchie von Hinweisen und Hilfestellungen besteht, bekommen (Kirschenbaum, 1998).

2.3 Historischer Umriss der dynamischen Testdiagnostik

Obwohl das dynamische Testen erst in den letzten Jahrzehnten des 20. Jahrhunderts verstärkt untersucht wurde, reicht die theoretische Grundlage historisch viel weiter zurück (Grigorenko & Sternberg, 1998). Um die Idee hinter dem dynamischen Testen besser zu verstehen, sollen in diesem Abschnitt einige der wichtigsten Wegbereiter und Vertreter der dynamischen Testdiagnostik vorgestellt werden. Anzumerken ist an dieser Stelle, dass die historische Entwicklung nur selektiv dargestellt werden kann, da es ansonsten den Rahmen dieser Arbeit sprengen würde.

Binet, der bis heute als Pionier für die Entwicklung von Intelligenztests gilt, war einer der ersten der betonte, dass es wichtig sei auch die Lernprozesse zu erfassen und nicht nur wie bisher die Lernprodukte, die auf früheren Möglichkeiten zu lernen beruhen (Binet, 1911, zitiert nach Haywood & Tzuriel, 2002, S. 43; Haywood & Tzuriel, 2002). Mit Simon gemeinsam ging Binet (1908, zitiert nach Guthke & Wiedl, 1996, S. 13) einen Schritt weiter und argumentierte, dass zwischen der angeborenen Intelligenz und Merkmalen, wie sie durch leistungsrelevante Charaktereigenschaften vermittelt werden, eine Unterscheidung getroffen werden muss. Erkannt wurde dabei erstmals, dass kognitive Merkmale einer Veränderbarkeit durch andere Einflussgrößen unterliegen. Methodisch haben die beiden Autoren den Versuch unternommen, innerhalb einer ihrer Verfahren, eine Aufgabengruppe zusammenzustellen, die unbeeinflusst von den Charaktermerkmalen der Testperson, die Intelligenz messen sollte (Binet, 1908, zitiert nach Guthke & Wiedl, 1996, S. 13). Meumann (1922) der sich näher mit den Arbeiten von Simon und Binet befasste, kritisierte daran, dass die Milieufaktoren durch deren Verfahren noch nicht ausreichend kontrolliert werden. Neben der Begabungskomponente und Entwicklungskomponente hielt er es für unabdingbar, auch die Umweltkomponente bei der Intelligenztestung nicht außer Acht zu lassen.

Die eben vorgestellten Überlegungen der Autoren zeigen, dass schon relativ früh Kritik an der herkömmlichen Intelligenzdiagnostik geübt und das Problem der Vernachlässigung des Umwelteinflusses in der Testdiagnostik erkannt wurde. Neben Binet (1909, zitiert nach Grigorenko & Sternberg 1998, S. 77), der eine prozessorientierte Testdiagnostik befürwortete, argumentierte auch Thorndike, dass es notwendig sei, die „Fähigkeit zum Lernen“ als Teil der Intelligenz zu erfassen (Thorndike, 1922).

Theoretisch begründet wurde das dynamische Testen durch Vygotski mit seiner Theorie „der Zone der nächsten Entwicklung“ (Cole, John-Steiner, Scribner, & Souberman, 1978). Schon Anfang der 30er Jahre stellte er fest, dass man noch mehr über die kognitive Entwicklung eines Kindes erfahren kann, wenn man erfasst was das Kind erreicht, wenn es Hilfestellungen während der Testung bekommt, anstatt wie üblich darauf zu verzichten (Lauchlan & Elliot, 2001). Dabei maß Vygotski der Zusammenarbeit des Erwachsenen mit dem Kind als auch der Zusammenarbeit unter Gleichaltrigen eine besondere Bedeutung, als potentiellen Weg zur Erfassung der intellektuellen Fähigkeiten, zu (Haywood et al., 1990). Seiner Theorie nach, können sich Kinder einerseits dadurch auszeichnen, was sie bereits eigenständig leisten können und andererseits dadurch, was sie leisten können, wenn sie dabei Hilfe durch eine andere Person bekommen (Lidz, 1995). Des Weiteren ist festzuhalten, dass die „Zone der nächsten Entwicklung“ nicht als eine Eigenschaft einer Person verstanden werden kann, da sie erst durch die Interaktion zwischen dem Hilfesteller und dem Kind selbst kreiert wird (Lidz, 1995). Folglich sind es die Effekte der Interaktion auf die mentalen Prozesse des Kindes, die gemessen werden können und die „Zone der nächsten Entwicklung“ sichtbar machen (Lidz, 1995).

Vygotski legte die Relevanz der Auseinandersetzung mit der „Zone der nächsten Entwicklung“ mit folgendem hypothetischen Beispiel dar: Zwei gleichaltrige Kinder erreichen in einem traditionellen Intelligenztest dasselbe Intelligenzalter von acht Jahren (Cole et al., 1978). Die herkömmliche Statusdiagnostik würde bei diesem Ergebnis davon ausgehen, dass sich die Kinder weder in ihrer kognitiven Entwicklung noch in ihrer zukünftigen Leistung unterscheiden (Cole et al., 1978). Bleibt man an diesem Punkt nicht stehen, indem man den Kindern zeigt, wie sie mit verschiedenen Problemstellungen umgehen können, stellt sich die Frage, ob die Kinder noch als „gleich gut“ bezüglich ihrer kognitiven Fähigkeiten eingestuft werden können. Gibt man also den beiden Kindern bei der Testdurchführung Anregungen in Form von Hilfestellungen, erreicht das eine Kind ein Intelligenzalter von 9 Jahren und das andere ein Intelligenzalter von 12 Jahren. Die Frage, die sich Vygotski an diesem Punkt stellt, ist, „are these children now mentally the same?“ (Cole

et al., 1978, S.86). Mit ziemlicher Sicherheit lässt sich diese Frage mit einem „Nein“ beantworten.

Das vorgestellte Beispiel zeigt, dass es nicht nur wichtig ist, die bereits herangereiften Fähigkeiten eines Kindes zu erfassen, sondern auch die Heranreifenden. In Vygotskis theoretischem Konzept hält er fest, dass es wichtig sei, eine Unterscheidung zwischen der „Zone der aktuellen Entwicklung“, den bereits herangereiften Fähigkeiten und der „Zone der nächsten Entwicklung“, den Heranreifenden zu treffen. Definiert wird die „Zone der nächsten Entwicklung“ als die Differenz zwischen dem momentanen Entwicklungsstand eines Kindes und dem Niveau, das das Kind mithilfe eines Erwachsenen erreicht (Cole et al., 1978). Durch seine Erkenntnisse hat Vygotski zu seiner Zeit den dynamischen Testansatz am weitesten ausgebaut und wird daher oft als der wesentliche Begründer der dynamischen Testdiagnostik angesehen (Elliott, 2010; Grigorenko & Sternberg, 1998; Guthke & Wiedl, 1996). Mit seinem theoretischen Konzept stellte Vygotski eine theoretische Grundlage bereit, welche viele Autoren, darunter auch Guthke, aufgriffen, weiterentwickelten und auf dieser Basis ihre Messinstrumente begründeten (Schack & Guthke, 2003).

Bevor nun in weiterer Folge auf Guthkes Beiträge zum dynamischen Testen eingegangen wird, sollen nun noch kurz die Zubin Axiome vorgestellt werden, die zwar weitgehend unbeachtet von anderen dynamischen Vertretern blieben, jedoch für das Lerntestkonzept nach Guthke eine essentielle Grundlage darstellen (Guthke, Beckmann, Wiedl, 2003). Zubin (1950) argumentierte, dass die Analyse der Gesetzmäßigkeiten eines Individuums im Vordergrund stehen müsse und die „Gruppenstatistik“ in den Hintergrund gerückt werden soll. Was genau darunter zu verstehen ist, formulierte Zubin (1950) in den folgenden Axiomen: Wichtig sei es, dass jedes Individuum als ein unabhängiges Universum betrachtet wird, dass erst mit anderen verglichen oder mit ähnlichen Gruppen zusammengefasst werden darf, wenn man sich mit dem individuellen Universum bereits eingehend beschäftigt hat. Ebenso soll jedes Individuum ein spezifisches Leistungsniveau haben und ein beobachteter Testwert nur eine Zufallsstichprobe darstellen. Eine weitere Feststellung besagt, dass jedes Individuum und jedes Merkmal durch einen gewissen Grad an Variabilität gekennzeichnet ist, deren Spielbreite von Person zu Person unterschiedlich ist (Zubin, 1950).

2.4 Das Lerntestkonzept nach Guthke

Das Lerntestkonzept stellt eine von vielen Varianten des dynamischen Testens dar und wurde von Guthke formuliert und entwickelt (Guthke, 1972; Guthke, 1992; Guthke et al., 2003). Dabei reicht die theoretische Grundlage des Lerntestkonzepts historisch weit zurück

und stützt sich überwiegend auf Vygotskis Theorie der „Zone der nächsten Entwicklung“, als auch auf die Zubin Axiome (Guthke & Wiedl, 1996). Nachdem in früherer Vergangenheit der Fokus der dynamischen Testdiagnostik darauf lag, die Schwächen und Unzulänglichkeiten der Statusdiagnostik aufzudecken und die theoretische Grundlage zu spezifizieren, wurde mit Beginn der 70er Jahre eine neue Phase eingeleitet. Der neuen Phase zugehörig, ging es Guthke, wie auch anderen Vertretern, um die systematische Erstellung von wissenschaftlich fundierten Untersuchungsverfahren, die sich in der Praxis bewähren sollten (Guthke & Wiedl, 1996; Haywood & Tzuriel, 2002). Dabei wurde einerseits besonderer Wert auf die Standardisierung der Testsituation gelegt und andererseits darauf, dass die Testverfahren nach psychometrischen Standards abgesichert werden (Guthke & Wiedl, 1996).

In seinem Lerntestkonzept unterschied Guthke zwischen einem Intelligenzstatus und der „intellektuellen Lernfähigkeit“ (Guthke & Wiedl, 1996). Die im Lerntest erzeugte „Fähigkeitssteigerung“ wird dabei als intellektuelle Lernfähigkeit interpretiert und als über die Intelligenz hinaus gehende zusätzliche diagnostische Information aufgefasst, die aber nicht im Sinne einer komplett neuen Eigenschaft verstanden wird (Guthke & Wiedl, 1996). Operationalisiert wurde die Lernfähigkeit über das Ausmaß von Leistungsveränderungen unter leistungsoptimierenden Bedingungen (Guthke & Wiedl, 1996). Es ist bis heute noch nicht ausreichend geklärt, ob mit den entwickelten Lerntests auch wirklich die intellektuelle Lernfähigkeit erfasst wird und nicht „nur“ die Intelligenz im Sinne des Zubin Axioms mit ihrem „Veränderungsspielraum“ (Guthke et al., 2003)

Beim Lerntest handelt es sich um eine diagnostische Methode, bei welcher ein/e TestleiterIn das Ausmaß erfasst, inwieweit eine Testperson in der Lage ist seine/ihre Testleistung zu verbessern, wenn diese Hilfestellungen und/oder Hinweise in einer Testsituation bekommt oder Trainingsprogramme zwischen einem Prä- und Posttest absolviert (Tzuriel, 2001). Ziel der Lerntests ist es, zuverlässigere Aussagen über die zukünftige Entwicklung von Personen zu treffen und pädagogische als auch therapeutische Vorschläge aus der Testung abzuleiten (Guthke, 1992). Dabei wurden die Lerntests zunächst als Langzeitlerntests konstruiert und später auch als Kurzzeitlerntests.

Die Langzeitlerntests bestehen aus insgesamt drei Phasen: einem Prätest, einem Posttest und einer zwischengeschalteten Trainingsphase. Der Prätest wird wie ein konventioneller Intelligenztest durchgeführt und es werden damit die bereits entwickelten kognitiven Fähigkeiten erfasst (Guthke & Wiedl, 1996). Im Anschluss an den Prätest findet die Trainingsphase statt, die je nach Verfahren unterschiedlich gestaltet ist. Dabei haben alle Verfahren gemeinsam, dass sie den Testpersonen in der Trainingsphase Lösungsstrategien

vermitteln, die diese dazu befähigen sollen, die Anforderungen des Tests besser zu bewältigen (Guthke & Wiedl, 1996; Schack & Guthke, 2003). Anschließend folgt der Posttest, der im Sinne „der Zone der nächsten Entwicklung“ die Lernfähigkeit der Testpersonen erfasst. Durch diese Art von Testprozedur wird angestrebt die intraindividuelle Variabilität der Intelligenz messbar zu machen (Guthke & Wiedl, 1996). Dabei betont Guthke, dass mit den Langzeitleerntests keine völlig neue Eigenschaft erfasst wird, sondern eine zusätzliche diagnostische Information, die als „der Grad der minimalen im Lerntest erzeugten Fähigkeitssteigerung“ definiert wird (Guthke & Wiedl, 1996, S.81).

Bei den Kurzzeitleerntests wird auf die Trennung zwischen der Test- und Trainingsphase verzichtet und die Pädagogisierungsphase direkt in den Testprozess hinein verlagert (Guthke & Wiedl, 1996; Schack & Guthke, 2003). Dabei erhält die Testperson während der Testung entweder einfache Feedbacks in Form von Richtig-Falsch-Informationen, allgemeine Lösungshinweise und/oder auf den Fehler der Testperson speziell abgestimmte Denkhilfen (Guthke & Wiedl, 1996). Die Entwicklung und der Einsatz von Kurzzeitleerntests wurde aus testökonomischen Gründen angestrebt, da die Langzeitleerntests in der Praxis für Individualtestungen zu aufwendig erschienen (Guthke & Wiedl, 1996).

In den 1980er Jahren wurde schließlich eine neue Variante des Kurzzeitleerntests entwickelt, die unter dem Begriff „Diagnostische Programme“ bekannt wurde (Guthke, 1978; Schack & Guthke, 2003). Durchgeführt in nur einer Testphase, zeichnen sich die diagnostischen Programme dadurch aus, dass sie sich stärker um eine grundlagenpsychologisch fundierte Testkonstruktion und um eine objektive Schwierigkeitsbestimmung der Testitems bemühen (Guthke & Wiedl, 1996). Die Diagnostischen Programme unterscheiden sich von anderen Kurzzeitleerntests dadurch, dass diese bei der Testauswertung nicht nur die Richtig-/Falsch Antworten summieren, sondern eine genauere Lernprozessanalyse anstreben, bei der nicht nur der Lerngewinn, sondern auch der Zeitverbrauch, die Anzahl, die Art und der „Einsatzort“ der benötigten Hilfen analysiert wird (Guthke & Wiedl, 1996; Schack & Guthke, 2003). Als ein Beispiel für ein solches Diagnostisches Programm kann die Adaptive Computergestützte Intelligenzlerntestbatterie (ACIL, Guthke, 1995) genannt werden, die sich durch einen adaptiven Charakter im Sinne eines individuell verzweigten Lernens auszeichnet.

2.5 Dynamisches Testen in der Praxis

Nachdem ab den 70er Jahren das Konzept des dynamischen Testens auch in der breiteren psychologischen Fachwelt bekannt wurde, ging die dynamische Testdiagnostik, wie zuvor schon beschrieben, in eine neue Phase über, in der sich verschiedene Arbeitsgruppen,

um die systematische Erstellung von wissenschaftlich fundierten Untersuchungsverfahren bemühten (Guthke & Wiedl, 1996; Haywood & Tzuriel, 2002). Das Learning Potential Assessment Device (LPAD, von Feuerstein, Rand, & Hoffmann, 1979), die Lerntestbatterie „Schlussfolgerndes Denken“ (LTS, Guthke, Jäger, & Schmidt, 1983) und die ACIL von Guthke (1995) sind hier als nennenswerte Verfahren, die in den letzten Jahrzehnten des 20. Jahrhunderts entwickelt wurden, anzuführen.

Obwohl viele ForscherInnen und PraktikerInnen von der Idee des dynamischen Testens überzeugt waren und bis heute sind, konnten sich die dynamischen Verfahren und Lerntests als Alternative oder Ergänzung zu den statusorientierten Intelligenztests in der Praxis nicht wirklich durchsetzen. Die Frage, warum es die Lerntests nicht in die Routinepraxis geschafft haben, wird vielfach damit beantwortet, dass die meisten Verfahren, wie es auch auf den als Langzeitlerntest konzipierten LTS (Guthke et al., 1983) zutrifft, zu zeitintensiv sind und der Durchführungsaufwand dadurch zu groß ist (Guthke & Stein, 1996; Guthke et al., 2003; Haywood et al., 1990; Murphy & Maree, 2006). Da für das Problem der langen Durchführungsdauer bereits Lösungen in Form von Kurzzeitlerntests existieren, kann dies vermutlich nicht mehr als einer der Hauptgründe ausgemacht werden, warum dynamische Verfahren kein essentieller Bestandteil in der Praxis sind (Guthke et al., 2003). Dem Vorwurf, dass die Testkonstruktion oftmals nicht theoriegeleitet ist, wurde durch die Entwicklung von Diagnostischen Programmen, die eben dies forcierten, ebenfalls bereits entgegengewirkt (Guthke & Stein, 1996).

Nicht abzustreiten ist jedoch, dass dynamische Verfahren oft nicht ausreichend nach psychometrischen Standards abgesichert sind und die Umsetzung bzw. die Überprüfung dieser, eine große Herausforderung für die dynamische Testdiagnostik darstellt (Murphy & Maree, 2006; Tzuriel, 2001). Dabei wird darüber diskutiert, dass die messmethodischen und psychometrischen Probleme noch viel größer sind, als die bei den herkömmlichen Tests (Guthke & Wiedl, 1996). Als problematisch werden unter anderem die Zuverlässigkeit von Differenzwerten, die Zweifelhaftigkeit der jeweils gewählten Lerntestauswertungsparameter und die Außenkriterien bei der Validierung angesehen (Guthke & Wiedl, 1996). Bei Langzeitlerntests muss bezüglich der Differenzwerte beispielsweise beachtet werden, dass zwar die Prä- und Posttestwerte zufriedenstellende Reliabilitätswerte aufweisen können, dieser Umstand aber nicht garantiert, dass dies auch auf die Differenzwerte zutrifft (Murphy & Maree, 2006). Bei Kurzzeitlerntests und den Diagnostischen Programmen ist die Überprüfung der Reliabilität noch schwieriger zu lösen, da weder die

Testhalbierungsmethode, die Konsistenzanalyse noch Cronbachs Alpha dafür geeignet erscheinen (Guthke & Wiedl, 1996).

Kritisiert wird auch, dass die Gültigkeit der dynamischen Verfahren vorwiegend an Außenkriterien, wie der Schulleistung oder der Schätzung der Intelligenz durch Lehrerurteile, geprüft wird. Dies liegt einerseits daran, dass man keine besseren Kriterien für die Lernfähigkeit fand und andererseits daran, dass der Vergleich zwischen den herkömmlichen Intelligenztests und den Lerntests bezüglich ihrer Gültigkeit erst dann angemessen erschien, wenn man die gleichen Außenkriterien zur Validierung nutzte (Guthke & Wiedl, 1996). Der Vorwurf, dass zur Validierung von Intelligenztests Lehrerschätzungen und Schulnoten herangezogen werden, ist also keineswegs nur an die dynamischen Testdiagnostik zu richten, sondern auch an die herkömmlichen Intelligenzdiagnostik. Die Kritik in Bezug auf die messmethodischen und die psychometrischen Probleme der dynamischen Testdiagnostik muss in Zukunft auf jeden Fall mehr Beachtung geschenkt und Lösungen dafür gesucht werden. Gleichzeitig ist es aber wichtig, dass diese Probleme nicht jegliche Weiterentwicklung und Innovation der dynamischen Testdiagnostik hemmen (Guthke & Wiedl, 1996).

Ein weiterer wesentlicher Grund, weshalb im deutschsprachigen Raum die dynamische Testdiagnostik weitgehend aus der Praxis und Forschung verschwunden ist, hängt vermutlich auch damit zusammen, dass nach dem Ableben von Jürgen Guthke im Jahr 2004 (Beckmann, 2004), der die Weiterentwicklung und Forschung im Bereich des dynamischen Testens maßgeblich vorantrieb, es niemanden mehr gab, der dessen Arbeit weiterführte. Dabei scheint es lohnenswert die dynamische Testdiagnostik weiterzuentwickeln und zu erforschen, wenn man den Nutzen und die vielen Vorteile des dynamischen Testens gegenüber den herkömmlichen Intelligenztests in Betracht zieht.

2.6 Vorteile des dynamischen Testens

Wie schon zu Beginn dieser Arbeit berichtet wurde, lag die Vermutung schon immer nahe, dass Kinder aus sozioökonomisch schwachen Verhältnissen bei herkömmlichen Tests benachteiligt werden, da bei diesen nur der momentane kognitive Status erfasst wird. Forschungsergebnisse bestätigen diese Annahme und stellten fest, dass herkömmliche Intelligenztests die kognitiven Fähigkeiten von Kindern aus sozioökonomisch schwachen Verhältnissen unterschätzen (Hamers, Hessels, & Pennings, 1996; Hessels, 1997). In einer Studie von Tzuriel (1989) wurde beispielsweise herausgefunden, dass sich Kindergartenkinder aus der Mittelschicht im Vergleich zu Kindergartenkindern aus der Unterschicht hinsichtlich ihrer Prätestleistungen unterscheiden. Die Posttestwerte der beiden

Gruppen glichen sich aber nach dem Training weitgehend aneinander an. Auch zeigte sich, dass die Prä- und Posttestwerte der Kinder aus der Mittelschicht viel höher miteinander übereinstimmten, als dies bei den Unterschichtskindern der Fall war (Tzuriel, 1989). Die Ergebnisse dieser Studie deuten darauf hin, dass das dynamische Testen in einem geringeren Ausmaß von Milieufaktoren beeinflusst wird und dass dies, gerade bei Kindern aus einem sozioökonomisch schwachen Umfeld, eine validere Erfassung ihrer kognitiven Fähigkeiten zulässt, als dies bei den herkömmlichen Intelligenztests der Fall ist.

Eine andere Annahme der dynamischen Testdiagnostik, dass Kinder aus anderen Kulturen, durch mögliche Sprachschwierigkeiten oder durch mangelnde Lernmöglichkeiten weniger gute Leistungen in den herkömmlichen Intelligenztests erreichen, konnte ebenfalls bestätigt werden (Tzuriel, 2001). Es wurde herausgefunden, dass dynamische Testverfahren, im Vergleich zu herkömmlichen Intelligenztests, die kognitiven Fähigkeiten dieser Kinder besser abbilden (Hessels & Hamers, 1993; Tzuriel & Kaufman, 1999). So wurde in einer niederländischen Studie festgestellt, dass die Kinder kultureller Minderheiten im Vergleich zu den aus den Niederlanden stammenden Kindern zwar signifikant geringere Leistungen in den Lerntests erreichten, diese Unterschiede zwischen den beiden Gruppen, jedoch viel kleiner waren als bei einem gewöhnlichen Intelligenztest (Hessels & Hamers, 1993). Des Weiteren wurde in einer in Israel durchgeführten Studie festgestellt, dass die in Israel geborenen Kinder in zwei verschiedenen Lerntests bessere Ergebnisse erzielten als die äthiopischen Immigrantenkinder (Tzuriel & Kaufman, 1999). Es wurde dabei herausgefunden, dass die Immigrantenkinder generell niedrigere Testwerte in den Prätests der beiden Lerntests erreichten, als die in Israel geborenen Kinder. Allerdings zeigten die Ergebnisse der Posttests, dass sich die Immigrantenkinder in ihrer Leistung stärker steigern konnten, als die in Israel geborenen Kinder. Im Posttest näherten sich die beiden Gruppen in ihrer Leistung weitgehend aneinander an (Tzuriel & Kaufman, 1999). Die Ergebnisse weisen darauf hin, dass es bei Kindern aus anderen Kulturkreisen, wie es auch bei Kindern aus einem sozioökonomisch schwachen Umfeld festgestellt wurde, akkurater ist, dynamische Verfahren zur Messung der kognitiven Fähigkeiten zu verwenden, da diese bei den herkömmlichen Intelligenztests benachteiligt erscheinen.

Betreffend die prognostische Validität zeigt sich, dass dynamische Verfahren und herkömmliche Intelligenztests in einem ähnlichen Ausmaß mit zukünftigen Erfolgsmaßen korrelieren (Caffrey et al., 2008). In einer Studie von Gitter (1988, zitiert nach Guthke, 1992, S. 141), bei der insgesamt mehr als 400 Kinder mit einem herkömmlichen Intelligenztest und mit einem Lerntest getestet wurden, zeigte sich, dass sich die Validitätskoeffizienten der

beiden Verfahrensvarianten in Bezug auf die gesamte Stichprobe nur geringfügig unterschieden. Als jedoch eine Subgruppe von Kindern näher untersucht wurde, die vor dem Schuleintritt von KindergärtnerInnen als besonders leistungsschwach beschrieben wurde, zeigte sich, dass der Lerntest bezüglich seiner prognostischen Validität dem Statustest klar überlegen war (Gitter, 1988, zitiert nach, Guthke 1992, S. 141). Generell wurde in mehreren Studien festgestellt, dass sich für Probanden aus „irregulären Lernbedingungen“ beim dynamischen Testen die größeren Validitätsvorteile im Vergleich zu den Statustests zeigen (Guthke et al., 2003). Festzuhalten ist, dass Lerntests aber auch für leistungsstarke Testpersonen mit einer „normalen“ Lerngeschichte geeignet sind (Guthke et al., 2003).

Neben der höheren prognostischen Validität von Lerntests bei Testanden aus „irregulären Lernbedingungen“ (Guthke et al., 2003), sollen dynamische Verfahren auch in der Lage sein, festzustellen, ob es einer Person entweder an Wissen mangelt oder an den Fähigkeiten (Haywood et al., 1990). Auch sollen LehrerInnen durch das dynamische Testen mehr Informationen über eine Person und ihre Fähigkeiten erhalten, sodass sie den Unterricht danach besser ausrichten können, als dies bei den Standardintelligenztests der Fall ist (Haywood & Tzuriel, 2002).

Betrachtet man die aus den empirischen Ergebnissen abgeleiteten Vorteile, so sprechen diese dafür, der dynamischen Testdiagnostik in Zukunft wieder mehr Aufmerksamkeit zu schenken. So sollte in diesem Bereich verstärkt Forschung betrieben werden und neue Verfahren in Bezug auf das dynamische Testen entwickelt werden. Mit der folgenden empirischen Untersuchung soll ein erster Schritt in diese Richtung unternommen werden, indem für den Untertest 12 - „Formale Folgerichtigkeit“ des Adaptiven Intelligenz Diagnostikums (AID 3, Kubinger & Holoher-Ertl, 2014) ein Lerntest in Anlehnung an das Lerntestkonzept von Guthke entwickelt wird.

3. Ziel der Untersuchung und Hypothesen

Das vorrangige Ziel der Untersuchung war die Entwicklung eines Lerntests, angelehnt an Guthkes Lerntestkonzept für den AID 3, der die Lernfähigkeit von Testpersonen erfassen soll. Dabei wurde entschieden die Umsetzung des Lerntestkonzepts beim Untertest „Formale Folgerichtigkeit“ des AID 3 zu erproben. Um festzustellen, ob mit dem entwickelten Lerntest die Lernfähigkeit erfasst werden kann, wird geprüft, ob mit dem Lerntest ein signifikanter Lerngewinn zu erzielen ist, der deutlich über den reinen Testwiederholungsgewinn hinausgeht. Des Weiteren soll untersucht werden, ob sich die Posttestwerte der Testpersonen, die mit dem Lerntest getestet wurden, nicht mit genügend hoher Wahrscheinlichkeit aus den Prätestwerten voraussagen lassen. Dabei ist es für die Lerntestidee substantiell, dass sich die Posttestwerte nicht mit genügend hoher Wahrscheinlichkeit aus den Prätestwerten voraussagen lassen, da nur so geklärt werden kann, ob der Posttest dem Prätest auch wirklich eine neue diagnostische Information hinzufügt (Guthke & Wiedl, 1996). Schließlich soll auch untersucht werden, ob sich die Testpersonen mit einem niedrigen/mittleren/hohen Lerngewinn hinsichtlich ihrer Prätestleistung, ihrem Alter, ihrer Nationalität (Österreich versus alle anderen Nationalitäten) und ihrem Geschlecht signifikant voneinander unterscheiden. Aus diesen Überlegungen lassen sich nun die folgenden Hypothesen für die Studie ableiten.

Hypothese 1: Mit dem Lerntest ist ein signifikanter Lerngewinn zu erzielen, der deutlich über den reinen Testwiederholungsgewinn hinausgeht.

Hypothese 2: Die Posttestwerte des Lerntests lassen sich nicht mit genügend hoher Wahrscheinlichkeit aus den Prätestwerten voraussagen.

Hypothese 3: Die Testpersonen mit einem niedrigen Lerngewinn/ mit einem mittleren Lerngewinn/ mit einem hohen Lerngewinn unterscheiden sich bezüglich ihrer Prätestleistung signifikant voneinander.

Hypothese 4: Die Testpersonen mit einem niedrigen Lerngewinn/ mit einem mittleren Lerngewinn/ mit einem hohen Lerngewinn unterscheiden sich bezüglich ihres Alters signifikant voneinander.

Hypothese 5: Die Testpersonen mit einem niedrigen Lerngewinn/ mit einem mittleren Lerngewinn/ mit einem hohen Lerngewinn unterscheiden sich bezüglich ihres Geschlechts signifikant voneinander.

Hypothese 6: Die Testpersonen mit einem niedrigen Lerngewinn/ mit einem mittleren Lerngewinn/ mit einem hohen Lerngewinn unterscheiden sich bezüglich ihrer Nationalität (Österreich versus alle anderen Nationalitäten) signifikant voneinander.

4. Entwicklung des Lerntests am Untertest „Formale Folgerichtigkeit“

4.1 Kurzbeschreibung des Untertests „Formale Folgerichtigkeit“ des AID3

Wie bereits erwähnt, wurde der Lerntest für den Untertest „Formale Folgerichtigkeit“ des AID 3 entwickelt. Mit dem manuell-visuellen Untertest soll die Fähigkeit erfasst werden, „Gesetzmäßigkeiten und logisch zwingende Zusammenhänge erkennen und zweckentsprechend verwerten zu können“ (Kubinger & Holocher-Ertl, 2014, S. 13). Dabei wird mithilfe von figuralem Aufgabenmaterial logisch-schlussfolgerndes Denken gemessen (Hagenmüller, 2011). Die Aufgabe der Testpersonen besteht darin, aus einer Reihe von Plättchen, die sich in ihrer Farbe, Form und Größe unterscheiden, dasjenige Plättchen auszuwählen, dass eine bereits bestehende Abfolge entsprechender Plättchen, logisch fortsetzt (Kubinger & Holocher-Ertl, 2014).

4.2 Allgemeine Informationen zum Lerntest

Da der Lerntest in den Untertest „Formale Folgerichtigkeit“ des AID 3 integriert wurde, musste für diese Arbeit kein neues Aufgabenmaterial entwickelt werden. So wurde das Aufgabenmaterial des Untertests „Formale Folgerichtigkeit“, wie es standardgemäß beim AID 3 vorgegeben wird, verwendet. Neben der Aufgabenstellung, die für die Testpersonen beim Lerntest gleich blieb, wurde die Instruktionsanweisung als auch die Durchführung der Warming-Up Aufgabe nicht verändert.

Auch wurde bei dem Lerntest das branched testing Design beibehalten, bei der die Kinder, je nach Altersgruppe, bei einer bestimmten Aufgabengruppe starten und diese anschließend, je nach Richtig oder Falschlösung der Aufgaben, gemäß dem Verzweigungssystem zur nächsten Aufgabengruppe weitergeleitet werden. Bei der Verrechnung der Antworten wird beim Lerntest der erste Lösungsversuch der Testpersonen gewertet. Das ist der Versuch, bei dem die Testpersonen noch keine Hilfestellungen bekommen haben. Löst also eine Testperson die Aufgabe beim ersten Versuch, wird die Aufgabe als „richtig“ gewertet. Ist die Testperson nicht in der Lage die entsprechende Aufgabe beim ersten Versuch zu lösen, wird die Aufgabe als „falsch“ gewertet und das Hilfesystem des Lerntests kommt ab diesem Punkt zum Tragen. Bevor also bei einer falschen Antwort zur nächsten Aufgabe, gemäß dem Verzweigungssystem, übergegangen wird, wird dem Kind durch Hinweise und Hilfen die Möglichkeit gegeben, die eigenen Fehler zu erkennen und zur richtigen Lösung zu gelangen. Wie das Hilfesystem genau aufgebaut ist, wird im nächsten Schritt ausführlich beschrieben. Die Fähigkeitsparameter sind beim

entwickelten Lerntest, wie es auch bei der standardmäßigen Vorgabe des Untertests „Formale Folgerichtigkeit“ der Fall ist, dem Manual zu entnehmen.

4.3 Das Hilfesystem des Lerntests

Zum entwickelten Hilfesystem ist allgemein festzuhalten, dass die Testpersonen, neben einem Feedback über die Lösung/Nicht-Lösung einer Aufgabe, spezifische Hinweise als auch Hilfestellungen in Form einer intensiveren Auseinandersetzung mit der Art des gemachten Fehlers bekommen können. Festzuhalten ist, dass es sich dabei um standardisierte Hilfen handelt, die immer auf den Typ des gemachten Fehlers Bezug nehmen. Der Typ des gemachten Fehlers kann sich bei dem Untertest „Formale Folgerichtigkeit“ auf die Dimensionen Farbe, Form und Größe der Plättchen beziehen. Das Hilfesystem besteht insgesamt maximal aus zwei Hilfsschritten. Ist die Testperson nicht in der Lage beim ersten Versuch die Aufgabe zu lösen, wird der erste Hilfsschritt dargeboten, der einen Hinweis darüber gibt, welche Dimension näher beachtet werden muss. Löst die Testperson die Aufgabe mithilfe des Hinweises wird zur nächsten Aufgabe übergegangen und der zweite Hilfsschritt ausgelassen. Löst die Testperson die Aufgabe mithilfe des Hinweises nicht, wird der zweite Hilfsschritt dargeboten, der eine intensivere Auseinandersetzung mit dem Typ des gemachten Fehlers beinhaltet. Anzumerken ist, dass sich der entwickelte Lerntest an dem Konzept der Hilfestellungen des Subtests „Adaptiver Figurenfolgenlerntest“ (ADAFI) der ACIL orientiert (Guthke, 1995).

Bevor die Hilfestellungen zum Einsatz kommen, geht es darum, dass der/die TestleiterIn zuallererst den Fehlertyp bestimmt, der von der Testperson gemacht wurde. Der/die TestleiterIn identifiziert dabei welche Dimension(en) (Farbe, Form und/oder Größe) von der Testperson nicht richtig erkannt wurde(n). Es wird also bestimmt, ob die Farbabfolge, die Formabfolge und/oder die Größenabfolge nicht richtig erkannt wurden. Die Hilfestellungen beziehen sich dann konkret auf den Typ des gemachten Fehlers.



Abb. 1: Beispielaufgabe aus dem Untertest „Formale Folgerichtigkeit“ des AID 3

Sieht man sich die Beispielaufgabe aus Abbildung 1 (Abb. 1) an, wäre die Lösung für die Aufgabe „ein kleines grünes Quadrat“. Wenn die Testperson stattdessen „ein kleines

gelbes Quadrat“ als Antwort wählt, ist die Aufgabe als falsch zu werten. Es ist dann davon auszugehen, dass die Testperson zwar die logische Abfolge der Dimensionen „Form“ und „Größe“ richtig erkannt hat, jedoch nicht die der Dimension „Farbe“.

Anschließend erfolgt der erste Hilfsschritt: Dabei wird bei dem gegebenen Beispiel ein standardisierter Hinweis gegeben, der sich auf die Dimension Farbe bezieht: „Beachte wie die Farbe der Figuren wechselt.“. Würde sich der Fehler in dem Beispiel auf die Dimension Größe beziehen, würde die Testperson den Hinweis: „Beachte wie die Größe der Figuren wechselt.“ bekommen. Würde sich der Fehler auf die Dimension der Form beziehen, würde das Kind die Hilfestellung: „Beachte wie die Form der Figuren wechselt.“ erhalten.

Natürlich kann es vorkommen, dass sich der Fehler nicht nur auf eine Dimension beschränkt. Es ist möglich, dass sich mehrere Fehlertypen in einem Beispiel ergeben, also auf mehreren Dimensionen gleichzeitig Fehler vorliegen. Wählt die Testperson zum Beispiel ein großes gelbes Quadrat als Antwort, so wurde die Abfolge der Dimension Farbe, als auch die Abfolge der Dimension Größe nicht richtig erkannt. Hierbei kann der/die TestleiterIn der Testperson den Hinweis: „Beachte wie die Farbe und die Größe der Figuren wechselt.“ geben. Der Hinweis soll sich immer auf die nicht richtig erkannte(n) Dimension(en) beziehen. Wählt die Testperson ein großes gelbes Dreieck, liegen hier Fehler auf allen drei Dimensionen vor. Der Hinweis bezieht sich nun auf alle drei Dimensionen: „Beachte wie die Farbe, die Größe und die Form der Figuren wechselt“. Nachdem man der Testperson den auf den Fehler bezogenen verbalen Hinweis gegeben hat, fordert man die Person auf: „Probiere es nochmal.“. Löst die Testperson die Aufgabe mithilfe des Hinweises, gibt man der Person die Rückmeldung: „Das ist richtig.“ und es wird zur nächsten Aufgabe übergegangen, ohne dass der zweite Hilfsschritt noch zur Anwendung kommt.

Ist die Testperson nicht in der Lage die Aufgabe mithilfe des Hinweises zu lösen, wird die Aufgabe um einen zweiten Hilfsschritt erweitert: Es wird nun mit dem Kind jede einzelne Dimension gesondert betrachtet, die beim zweiten Versuch, also nach dem gegebenen ersten Hilfsschritt, nicht richtig erkannt wurde. Wenn die Testperson beispielsweise die Abfolge der Dimension „Farbe“ beim zweiten Versuch nach gegebenem Hinweis nicht richtig erkannt hat, betrachtet man die Dimension Farbe mit der Testperson gesondert. Dies geschieht indem man auf jede einzelne Figur von links nach rechts tippt (aus der Sicht der Testperson) und dabei die Farbe jeder einzelnen Figur benennt. In Anlehnung an das Beispiel würde der Ablauf dafür so aussehen: Zuerst wird der Testperson gesagt: „Sehen wir uns noch einmal die Farben der Figuren genauer an.“. Anschließend werden von dem/der TestleiterIn die Figuren in der Reihenfolge mit dem Zeigefinger angetippt und deren Farbe benannt. Bezogen auf das

Beispiel in Abb.1 würde man sagen: „Wir haben hier Grün - Grün - Gelb - Gelb - Grün“. Anschließend wird auf das Kästchen mit dem Fragezeichen getippt und die Testperson gefragt: „Welche Farbe hat die Figur an dieser Stelle?“. Ist das Kind nicht in der Lage, die richtige Antwort zu erkennen, wird die Aufgabe aufgelöst, indem man dem Kind zeigt, welches Plättchen an die Stelle gehört und man beim Ablegen des Plättchens sagt: „So habe ich das gemeint.“.

Erkennt das Kind, dass die richtige Antwort „grün“ ist, fordert man das Kind auf erneut ein Plättchen zu wählen und auf das Kästchen mit dem Fragezeichen zu legen: „Nimm ein Plättchen, von welchem du denkst, dass es an diese Stelle gehört und lege es dort hin.“. Liegt das Kind mit seiner Wahl richtig, wird dem Kind die Rückmeldung: „Das ist richtig.“ gegeben und man geht zur nächsten Aufgabe über. Erkennt die Testperson die richtige Antwort nicht, wird die Aufgabe aufgelöst und es wird dem Kind gezeigt, welche Figur an die Stelle des Kästchens mit dem Fragezeichen gehört. Beim Ablegen des richtigen Plättchens auf die Stelle mit dem Fragezeichen sagt man: „So habe ich das gemeint.“.

Wenn sich der Fehler nach dem zweiten Versuch, also nach gegebenen Hinweis, nicht auf die Farbe wie im Beispiel bezieht, sondern auf eine andere Dimension, wird trotzdem nach dem gleichen Schema vorgegangen. Es wird dann die Anleitung, an die jeweilige Dimension, die nicht erkannt wurde, angepasst. Wenn sich der Fehler auf die Dimension Form bezieht, sagt man: „Sehen wir uns noch einmal die Form der Figuren genauer an. Wir haben hier...“. Es werden wieder die Formen in ihrer Reihenfolge nacheinander benannt und am Ende auf das Kästchen mit dem Fragezeichen getippt und gefragt: „Welche Form hat die Figur an dieser Stelle?“. Gleiches gilt für die Dimension Größe, bei der gesagt wird „Sehen wir uns noch einmal die Größe der Figuren genauer an. Wir haben hier...“. Alles andere läuft bei der Dimension Größe genauso ab, wie beim zweiten Hilfsschritt für die Dimensionen Farbe und Form.

Falls die Testperson beim zweiten Versuch nach dem ersten Hilfsschritt Fehler auf mehreren Dimensionen gemacht hat, wird jede einzelne Dimension, bei der die Abfolge nicht richtig erkannt wurde, einzeln betrachtet. Es wird nach dem gleichen Prinzip vorgegangen, wie das bei nur einer Dimension geschieht. Wenn beispielsweise die Form als auch die Farbe beim zweiten Versuch, also nach dem gegebenen Hinweis, nicht richtig erkannt wurde, wird die Form als auch die Farbe nacheinander gesondert betrachtet, wie zuvor ausführlich beschrieben. In welcher Reihenfolge die Dimensionen einzeln betrachtet werden, ist nicht festgelegt und kann von dem/der TestleiterIn frei entschieden werden. Scheitert das Kind nach dem Abtippen der Figuren bei einer der Fragen „Welche Farbe/Form/Größe hat die Figur an

dieser Stelle?“, wird die gesonderte Betrachtung der Dimensionen abgebrochen und die Aufgabe sofort aufgelöst, indem der Testperson gezeigt wird, welches Plättchen auf das Kästchen mit dem Fragezeichen gehört. Dabei wird wieder gesagt: „So habe ich das gemeint.“. Kann die Testperson bei der gesonderten Betrachtung der einzelnen Dimensionen, die jeweilige Frage „Welche Farbe/Form/Größe hat die Figur an dieser Stelle?“ richtig beantworten, wird das Kind anschließend aufgefordert ein Plättchen zu wählen und auf das Kästchen mit dem Fragezeichen zu legen. Dies geschieht mit der gleichen Aufforderung, wie dies auch im Beispiel zuvor dargelegt wurde. Legt das Kind das richtige Plättchen auf die Stelle, wird dem Kind die Rückmeldung „Das ist richtig.“ gegeben und man geht zur nächsten Aufgabe über. Liegt das Kind mit seiner Wahl des Plättchens falsch, löst man die Aufgabe, wie zuvor bereits beschrieben, auf. Nachdem man dem Kind die Lösung gezeigt hat, geht man zur nächsten Aufgabe über.

Neben den eben genannten Dimensionen, gibt es beim Untertest „Formale Folgerichtigkeit“ bei ein paar Aufgaben noch „die Richtung des Dreiecks“ als zusätzliche Dimension, die beachtet werden muss. Bezüglich der Verrechnung der Antworten und der Durchführungsanleitung gibt es beim standardmäßigen Untertest für diese zusätzliche Dimension eine etwas andere Handhabung als bei den anderen Dimensionen. Dies musste bei der Entwicklung des Lerntests mitberücksichtigt werden und deshalb leiten sich auch für das Hilfesystem bezogen auf die Dimension „Richtung des Dreiecks“ etwas andere Regeln ab, als das bei den anderen Dimensionen der Fall ist.

Tritt der Fall ein, dass die Testperson bei ihrem ersten Versuch, die Richtung des Dreiecks nicht sofort erkennt und keine Fehler auf anderen Dimensionen gemacht wurden, fragt man die Testperson, wie das auch beim standardmäßigen Untertest der Fall ist: „Wohin zeigt das Dreieck?“. Ist das Kind in der Lage die Richtung zu korrigieren, wird die Aufgabe als „Richtig“ gewertet und es wird zur nächsten Aufgabe übergegangen.

Wenn das Kind nach dem ersten fehlgeschlagenen Versuch bezüglich der Richtung des Dreiecks, die Frage „Wohin zeigt das Dreieck?“ nicht richtig beantworten kann und beim ersten Versuch auf keinen anderen Dimensionen Fehler gemacht wurden, wird die Aufgabe als falsch gewertet und zum zweiten Hilfsschritt übergegangen. Dieser zweite Hilfsschritt gestaltet sich folgendermaßen: Der Testperson wird gesagt: „Sehen wir uns nochmal die Richtung des Dreiecks genauer an.“. Anschließend werden die Figuren in ihrer Abfolge nacheinander angetippt und dabei wird beispielsweise gesagt: „Dieses Dreieck zeigt nach oben, dieses Dreieck zeigt nach unten, der Kreis hat keine Richtung, das Dreieck zeigt nach rechts...“. Im nächsten Schritt tippt man auf die Stelle mit dem Fragezeichen und man fragt

das Kind: „In welche Richtung zeigt das Dreieck?“. Kann das Kind die Richtung nicht richtig benennen, löst man die Aufgabe, wie zuvor schon bei den anderen Dimensionen beschrieben, auf. Kann das Kind die Richtung des Dreiecks richtig benennen, wird das Kind aufgefordert erneut ein Plättchen zu wählen. Liegt das Kind dann mit seiner Wahl richtig, sagt man „Das ist richtig.“. Liegt das Kind nach der Aufforderung ein Plättchen zu wählen, mit seiner Wahl falsch wird die Aufgabe aufgelöst.

Wenn neben der Dimension der Richtung des Dreiecks, die Dimension(en) Farbe und/oder Größe nicht richtig erkannt wurde(n), hält man sich an die Regeln, die für diese Dimensionen vorhin schon genau beschrieben wurden. Verrechnet wird die Aufgabe sofort als falsch, da auf den anderen Dimensionen bereits beim ersten Versuch Fehler gemacht wurden. Anschließend wendet man für die Dimensionen Farbe und/oder Größe den ersten Hilsschritt, wie gewohnt, an. Im zweiten Hilsschritt werden, wie zuvor schon beschrieben, alle Dimensionen einzeln betrachtet inklusive der Richtung des Dreiecks. Die Regeln und die Durchführung des zweiten Hilsschritts bleiben in diesem Fall gleich.

Tritt der Fall ein, dass bei einer Aufgabe die Lösung ein Dreieck mit einer bestimmten Richtung ist und das Kind ein Plättchen wählt, dass eine andere Form als ein Dreieck hat, kann man an dieser Stelle selbstverständlich nicht fragen: „Wohin zeigt das Dreieck?“. Hier ist folgendermaßen vorzugehen: Zuerst wird dem Kind der Hinweis gegeben „Beachte wie die Formen wechseln.“. Anschließend wird das Kind aufgefordert: „Probiere es nochmal.“. Wenn das Kind nach dem ersten Hilsschritt die richtige Form erkannt hat und die Richtung noch nicht, fragt man die Testperson „Wohin zeigt das Dreieck?“. Ist das Kind nicht in der Lage die Richtung zu erschließen, wird für diesen Fall, wie gewohnt der zweite Hilsschritt durchgeführt. Tritt dieser Sonderfall ein, ist es unabdingbar, dass vor der Dimension Richtung, die Frage nach der Form geklärt wird. Wenn die Form nach dem ersten Hilsschritt von dem Kind nicht erkannt wird, kann kein Hilsschritt zur Richtung gegeben werden. Das bedeutet, dass die zweite Hilfestellung nicht mehr vorgegeben werden kann. Der Testperson wird dann die Lösung der Aufgabe gezeigt und gesagt: „So habe ich das gemeint.“.

5. Methode

5.1 Untersuchungsdesign

Zur Überprüfung der zuvor beschriebenen Hypothesen wurde eine experimentelle Untersuchung durchgeführt. Jede Testperson wurde vor Beginn der Untersuchung zufällig auf eine von zwei Bedingungen, entweder der Experimentalgruppe oder der Kontrollgruppe, zugeteilt. Die Randomisierung erfolgte dabei mithilfe eines Würfels. Beim Würfeln einer geraden Zahl wurde die Testperson der Experimentalgruppe zugeteilt. Wurde eine ungerade Zahl gewürfelt, erfolgte die Zuteilung der Testperson zur Kontrollgruppe. Für die Untersuchung wurden alle Kinder zwei Mal getestet, um deren Prä- und Posttestleistungen zu erfassen. Beim ersten Testdurchgang wurden alle StudienteilnehmerInnen mit dem standardmäßigen Untertest „Formale Folgerichtigkeit“ des AID 3 getestet. Nach einer kurzen Pause folgte dann der zweite Testdurchgang, bei dem die Testpersonen, die der Experimentalgruppe zugeteilt wurden, mit dem entwickelten Lerntest getestet wurden und die Leistungen der Testpersonen der Kontrollgruppe mit dem standardmäßigen Untertest „Formale Folgerichtigkeit“ wiederholt erfasst wurden.

5.2 Darstellung der Stichprobe

Zur Rekrutierung der StudienteilnehmerInnen wurden zunächst E-Mails an insgesamt 70 Wiener Volksschulen ausgeschickt, in denen die DirektorInnen über das Forschungsvorhaben informiert wurden. Dabei wurde angefragt, ob die Durchführung der Studie an den Schulen möglich ist und dem E-Mail wurde eine detaillierte Beschreibung über den Ablauf, den Zweck und dem Ziel der Studie beigelegt. Auf diese E-Mail meldeten sich sechs DirektorInnen zurück, die an der Studie interessiert waren. Von den sechs interessierten Volksschulen, haben schlussendlich zwei Direktorinnen zugestimmt die Studie unterstützen zu wollen. Drei Direktorinnen erteilten nach einer Abstimmung im Schulforum eine Absage und eine weitere Direktorin entschied sich nach anfänglichem Interesse, ohne nähere Angaben gegen die Unterstützung der Untersuchung. Nach einem persönlichen Gespräch mit den beiden Direktorinnen der verbleibenden Volksschulen, teilten die KlassenlehrerInnen Einverständniserklärungen (siehe Anhang C) für die Eltern in den Klassen aus, die zwei Wochen später wieder eingesammelt wurden. Für die Studie wurden ausschließlich Kinder im Alter zwischen sechs und elf Jahren rekrutiert. Insgesamt waren 110 Eltern damit einverstanden, ihre Kinder an der Untersuchung teilnehmen zu lassen. Vier Kinder mussten wegen mangelndem Sprachverständnisses ausgeschlossen werden. Nicht an der Testung teilnehmen konnten ebenfalls zwei weitere Kinder, die zum Zeitpunkt der Untersuchung die

Schule gewechselt hatten und zwei Kinder, die im Zeitraum der Testung aufgrund längerer Krankheit nicht in der Schule waren. Ein Fall wurde nachträglich ausgeschlossen, weil bei der Testung ein Durchführungsfehler unterlaufen ist.

5.3 Durchführung und Ablauf der Untersuchung

Die Untersuchungen fanden im Zeitraum von zirka Mitte April 2018 bis Mitte Mai 2018 in der Volksschule Pannaschgasse 6 und in der Volksschule Klausenburger Straße 25 statt. Durchgeführt wurden die Untersuchungen immer vormittags zwischen acht und zwölf Uhr. Neben dieser Studie wurde noch eine weitere Untersuchung für eine Masterarbeit (Renz, 2018) in den Schulen durchgeführt. Dies gestaltete sich so, dass beide Testleiterinnen neben der eigenen Untersuchung, noch für die jeweils andere Masterarbeit die Daten miterhoben haben. Da zu den Terminen immer beide Testleiterinnen anwesend waren, konnten so parallel zwei Kinder getestet werden. Getestet wurden die Kinder in der Volksschule Klausenburger Straße 25 in einem separaten Raum und in der Volksschule Pannaschgasse 6 am Gang vor den Klassenzimmern. An beiden Testorten war eine ungestörte Testung ohne Lärmbelästigung oder sonstiger Ablenkung möglich.

Zum Ablauf der Untersuchung ist zu erwähnen, dass die Kinder entweder von den KlassenlehrerInnen nacheinander zur Testung geschickt oder von den Testleiterinnen selbst aus der Klasse abgeholt wurden. Zu Beginn der Untersuchung wurde jedem Kind der standardmäßige Untertest „Formale Folgerichtigkeit“ vorgegeben. Anschließend wurde die Testung für die Masterarbeit der Kollegin durchgeführt, da eine Pause zwischen der ersten und zweiten Testung nötig war, damit sich die Kinder nicht mehr im Detail an die Aufgaben des Untertests „Formale Folgerichtigkeit“ erinnern konnten. Im Anschluss an die Erhebung der Daten für die andere Masterarbeit, wurde je nach Zuteilung zur Experimentalgruppe oder Kontrollgruppe, entweder der Lerntest oder wiederholt der standardmäßige Untertest „Formale Folgerichtigkeit“ vorgegeben. Die Durchführungsdauer der Testung lag für beide Studien zusammen bei zirka 40 bis 60 Minuten pro Testperson. Den Kindern wurde nach der Testung noch die Möglichkeit gegeben, Fragen zur Testung zu stellen und sie erhielten als Belohnung für ihre Teilnahme jeweils zwei Bonbons.

5.4 Operationalisierung und Messinstrumente

Für die Erhebung der Daten wurde der Untertest 12- „Formale Folgerichtigkeit“ des AID 3 (Kubinger & Holocher-Ertl, 2014) eingesetzt. Ebenfalls kam als Messinstrument, der für diese Studie entwickelte Lerntest zum Einsatz. Da der Lerntest im vierten Kapitel schon ausführlich beschrieben wurde, wird an dieser Stelle auf eine wiederholte Darstellung

verzichtet. Für eine vollständige Beschreibung des Untertests „Formale Folgerichtigkeit“ des AID 3 wird auf dessen Manual (Kubinger & Holocher-Ertl, 2014) verwiesen.

Um die tatsächliche Leistung der StudienteilnehmerInnen zu erfassen, wurden bei der standardmäßigen Vorgabe des Untertests „Formale Folgerichtigkeit“ für jede einzelne Person die Fähigkeitsparameter ermittelt, die dem Manual des AID 3 zu entnehmen sind. Da die Verrechnung der Testantworten beim Lerntest gleich blieb, wurden auch bei diesem die Fähigkeitsparameter aus dem Manual des AID 3 entnommen. Um zu prüfen, ob mit dem Lerntest ein signifikanter Lerngewinn zu erzielen ist, der über den reinen Testwiederholungsgewinn hinausgeht, musste der Lerngewinn operationalisiert werden. Da für jedes Kind die Prätestleistung und die Posttestleistung erfasst wurde, errechnete sich der Lerngewinn einer jeden Testperson aus der Differenz der beiden erfassten Fähigkeitsparameter: $\text{Lerngewinn} = \text{Fähigkeitsparameter}_{\text{Posttest}} - \text{Fähigkeitsparameter}_{\text{Prätest}}$

5.5 Beschreibung der statistischen Analyse

Zur statistischen Analyse der ersten Fragestellung, ob ein signifikanter Lerngewinn zu erzielen ist, der deutlich über den reinen Testwiederholungsgewinn hinausgeht, wurde entschieden einen Zwei-Stichproben-*t*-Test durchzuführen. Die Analyse wird mit dem Welch-Test vorgenommen, da neuere Forschungsergebnisse wie die von Rasch, Kubinger und Moder (2011) zeigen, dass der Welch-Test grundsätzlich dem *t*-Test vorzuziehen ist. Um zu untersuchen, ob sich die Posttestwerte der Experimentalgruppe nicht mit genügend hoher Wahrscheinlichkeit aus den Prätestwerten voraussagen lassen, wurde eine einfache lineare Regressionsanalyse durchgeführt. Dabei fungierte als abhängige Variable der Fähigkeitsparameter der zweiten Testung (Lerntest) und als Prädiktor der Fähigkeitsparameter der ersten Testung (Standardmäßiger Untertest „Formale Folgerichtigkeit“). Es wurden für diese Analyse nur die Daten der Personen herangezogen, die der Experimentalgruppe zugeteilt waren. Für die Untersuchung der Hypothesen drei bis sechs wurde vor der statistischen Analyse jeder Studienteilnehmer/jede Studienteilnehmerin der Experimentalgruppe abhängig von der Höhe ihres tatsächlichen Lerngewinns, einer von drei Gruppen zugeteilt: niedriger Lerngewinn/ mittlerer Lerngewinn/ hoher Lerngewinn. Zur Gruppe „niedriger Lerngewinn“ wurden die Testpersonen zugeteilt, die sich in dem Drittel mit dem niedrigsten Lerngewinn aller Testpersonen der Experimentalgruppe befanden und der Lerngewinn einem Wert zwischen -0,7 und 0,3 entsprach. Zur Gruppe „mittlerer Lerngewinn“ wurden die StudienteilnehmerInnen zugeteilt, deren Lerngewinn dem mittleren Drittel aller Personen der Experimentalgruppe zuzuordnen war und der Wert des Lerngewinns zwischen 0,5 und 1,1 lag. Der Gruppe „hoher Lerngewinn“ wurde das Drittel aller

Testpersonen mit dem höchsten Lerngewinn zugeteilt. Für diese Gruppe lag der Wert des Lerngewinns zwischen 1,2 und 2,9. Für die Untersuchung, ob es zwischen den Gruppen signifikante Unterschiede bezüglich der Prätestleistung gibt, wurde eine Varianzanalyse durchgeführt. Um zu klären, ob es zwischen den Gruppen signifikante Unterschiede bezüglich des Alters gibt, wurde der Kruskal-Wallis-Test bevorzugt, da die Voraussetzung der Varianzhomogenität für die Durchführung der Varianzanalyse als verletzt galt. Mit Chi-Quadrat-Tests wurde schließlich überprüft, ob es zwischen den Gruppen signifikante Unterschiede hinsichtlich des Geschlechts und der Nationalität (Österreich versus alle anderen Nationen) gibt. Die statistische Analyse der Hypothesen wurde mit dem Programm SPSS durchgeführt.

6. Ergebnisse

6.1 Deskriptivstatistische Beschreibung der Stichprobe

Die Stichprobe umfasste insgesamt $N = 101$ SchülerInnen. Davon besuchten 25 Kinder die Volksschule Klausenburger Straße und 76 Kinder die Volksschule Pannaschgasse. Die Stichprobe setzte sich aus 48 Mädchen (47,5%) und 53 Jungen (52,5%) zusammen. Die Altersspanne der StudienteilnehmerInnen reichte von 6-11 Jahren ($M = 8.66$; $SD = 1.032$).

Die Frage „Aus welchem Land kommst du?“, beantworteten, 41 Kinder mit Österreich, 12 Kinder mit Serbien, 9 Kinder mit Syrien, 7 Kinder mit Türkei, 5 Kinder mit Ägypten, 3 Kinder mit Kroatien, 3 Kinder mit Bulgarien, 2 Kinder mit Albanien. Jeweils ein Kind gab an aus Rumänien, Palästina, Tunesien, Tschechien, Japan, Libyen, Mazedonien, Tschetschenien, Italien, Griechenland, Bosnien, China, Philippinen, Polen, Kosovo, Ungarn, Litauen, Indien und Vietnam, zu stammen. Insgesamt stammen die Kinder aus dieser Stichprobe aus 27 unterschiedlichen Ländern.

Wie sich die Anzahl der StudienteilnehmerInnen, das Geschlecht und das Alter auf die Versuchsbedingungen verteilen, wird in der Tabelle (Tab. 1) unterhalb dargestellt. So wurden durch die Randomisierung etwas mehr Personen der Experimentalgruppe zugeteilt als der Kontrollgruppe. In einem sehr geringen Ausmaß war in beiden Versuchsbedingungen der männliche Anteil größer als der weibliche Anteil. Die beiden Gruppen unterscheiden sich unwesentlich in ihrem Altersdurchschnitt voneinander.

Tabelle 1

Deskriptivstatistik der beiden Versuchsbedingungen

<i>Versuchsbedingung</i>	n_G	n_M	n_W	M_A	SD_A
<i>EXG-Lerntest</i>	55	29	26	8.76	0.942
<i>KG-Standard</i>	46	24	22	8.54	1.130

Anmerkung. n_G = Anzahl an StudienteilnehmerInnen in der jeweiligen Gruppe, n_M =Anzahl männlicher Studienteilnehmer, n_W = Anzahl weiblicher Studienteilnehmer, M_A =Mittelwert des Alters, SD_A = Standardabweichung des Alters

6.2 Explorative Datenanalyse

Nach einer genauen Analyse der Daten fiel auf, dass sich drei Testpersonen in der Kontrollgruppe in ihren Leistungen vom Prätest auf den Posttest erheblich gesteigert haben. Diese drei Testpersonen haben sich alle mindestens um einen Fähigkeitsparameterwert von

3.0 gesteigert. Dies scheint hinsichtlich der einfachen Testwiederholung des standardmäßigen Untertests „Formale Folgerichtigkeit“ nicht nachvollziehbar. Möglicherweise ist eine solche Leistungssteigerung auf einen Durchführungsfehler zurückzuführen. Beispielsweise wäre es möglich, dass die Testpersonen bei der ersten Testung die Instruktion oder die Aufgabenstellung nicht vollständig verstanden haben und bei der zweiten Testung anschließend schon. Über den Grund kann allerdings nur spekuliert werden. Trotzdem scheint es sicherer zu sein, diese Fälle aus der Untersuchung nachträglich auszuschließen, um die Ergebnisse der Studie nicht zu verfälschen. Da entschieden wurde die Fälle aus der Kontrollgruppe mit einer Leistungssteigerung von einem Fähigkeitsparameter ≥ 3.0 auszuschließen, wurde dies auch bei der Experimentalgruppe gemacht. So wurden in der Experimentalgruppe ebenfalls drei Fälle ausgeschlossen, die das Ausschlusskriterium einer Leistungssteigerung von einem Fähigkeitsparameter von ≥ 3.0 aufwiesen. Insgesamt wurden damit vor der Durchführung der statistischen Analyse der Hypothesen sechs Fälle aus der Untersuchung ausgeschlossen.

6.3 Statistische Auswertung der Hypothesen

Hypothese 1. Im Durchschnitt zeigt sich bei den Testpersonen die mit Lerntest getestet wurden ein größerer Lerngewinn ($M = 0.85$, $SD = 0.88$), als bei den Testpersonen die wiederholt mit dem standardmäßigen Untertest getestet wurden ($M = 0.46$, $SD = 1.17$). Dabei konnte ein signifikanter Unterschied bezüglich des Lerngewinns $t_{\text{Welch}}(76,74) = -1.84$, p (einseitig) $< .05$ zwischen der Experimentalgruppe, die mit dem Lerntest getestet wurde und der Kontrollgruppe, die mit dem standardmäßigen Untertest „Formale Folgerichtigkeit“ getestet wurde, festgestellt werden. Die Hypothese, dass mit dem Lerntest ein signifikanter Lerngewinn zu erzielen ist, der deutlich über den reinen Testwiederholungsgewinn hinausgeht, kann somit als bestätigt angesehen werden.

Zur Veranschaulichung wurden die Unterschiede bezüglich des Lerngewinns zwischen der Experimentalgruppe und der Kontrollgruppe in der folgenden Abbildung (Abb. 2) dargestellt. Dabei wurden die Mittelwerte der Fähigkeitsparameter beider Gruppen des Prä- und Posttests abgebildet. Bei der Vorgabe des standardmäßigen Untertests im Prätest erreicht die Experimentalgruppe im Durchschnitt einen Fähigkeitsparameter von 0.48 und die Kontrollgruppe einen Fähigkeitsparameter von -0.04. Im Posttest zeigt sich bei der Experimentalgruppe, bei der der Lerntest durchgeführt wurde, ein durchschnittlicher Fähigkeitsparameter von 1.33 und bei der Kontrollgruppe, mit welcher der standardmäßige Untertest wiederholt wurde, ein durchschnittlicher Fähigkeitsparameter von 0.42. Wie aus der statistischen Analyse mit dem Welch-Test hervorgeht, wird auch aus der Abbildung, der

höhere Lerngewinn der Experimentalgruppe gegenüber der Kontrollgruppe ersichtlich. Da der Lerngewinn der Experimentalgruppe deutlich über den reinen Testwiederholungsgewinn hinausgeht, kann davon ausgegangen werden, dass der Lerntest in der Lage ist die Lernfähigkeit von Testpersonen zu erfassen.

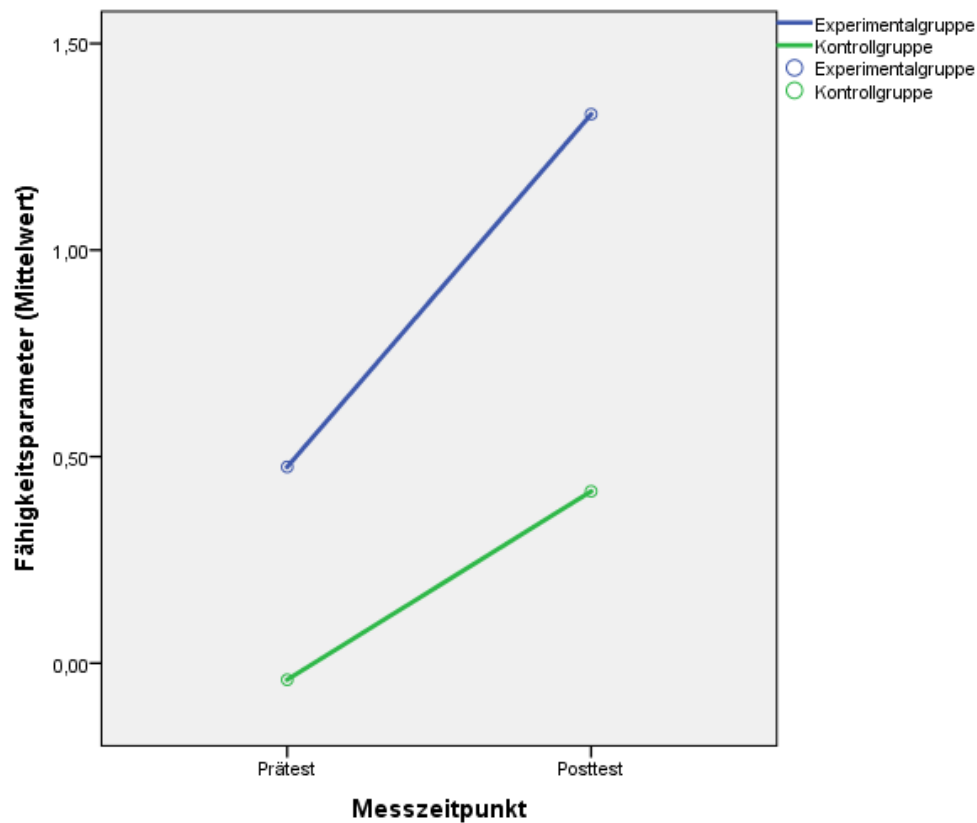


Abb. 2: Prä- und Posttestmittelwerte der Experimentalgruppe und der Kontrollgruppe

Die nächste Abbildung (Abb. 3) illustriert die Verteilung aller Merkmalsausprägungen des Lerngewinns der Kontrollgruppe als auch der Experimentalgruppe. Es ist zu erkennen, dass die Merkmalsausprägungen der Experimentalgruppe eher rechts verteilt sind im Vergleich zu den Merkmalsausprägungen der Kontrollgruppe. Die Grafik bestätigt die bereits gewonnene Einsicht, dass die Testpersonen, die mit dem Lerntest getestet wurden, einen größeren Lerngewinn erzielen, als die Testpersonen der Kontrollgruppe. Zwar zeigt die Abbildung, dass sich auch in der Kontrollgruppe einige Personen stark in ihrer Leistung verbessert haben, allerdings haben sich in der Kontrollgruppe manche Personen auch stark in ihrer Leistung verschlechtert, was auf die Experimentalgruppe in diesem Ausmaß nicht zutrifft. Selbst wenn man den Ausreißer der Kontrollgruppe mit einer Fähigkeitsparameterdifferenz von -4,4 außer

Acht lässt, deutet die Grafik noch immer darauf hin, dass sich die Merkmalsausprägungen der Experimentalgruppe stärker rechts verteilen als die Merkmalsausprägungen der Kontrollgruppe.

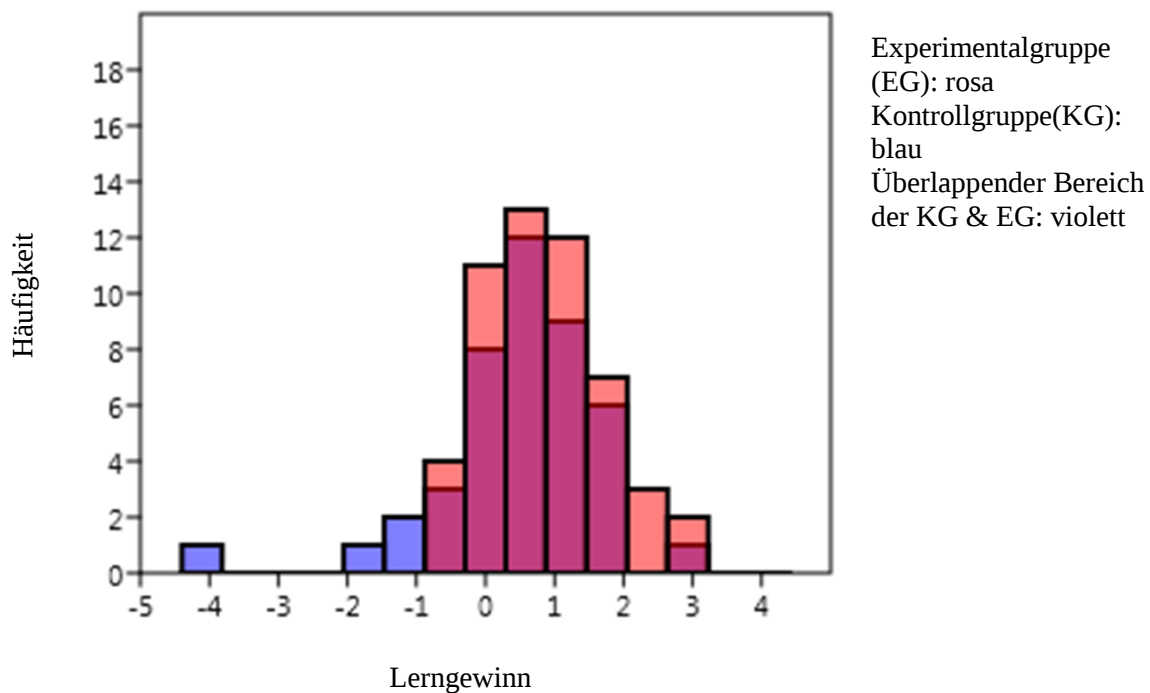


Abb. 3: Verteilung des Lerngewinns für die Experimentalgruppe und für die Kontrollgruppe

Zum Lerngewinn der Testpersonen der Experimentalgruppe ist zudem festzuhalten, dass es einerseits Kinder gab, die von dem Lerntest profitierten und andererseits Kinder, die sich entweder in ihrer Leistung verschlechterten, gleich blieben oder sich nicht wesentlich verbesserten. Die Annahme, dass ein Kind von dem Lerntest profitiert hat und somit von einem wesentlichen Lerngewinn gesprochen werden kann, wurde so definiert, dass die Leistungssteigerung eines Kindes über dem Mittelwert des Lerngewinns der Kontrollgruppe ($M = 0.46$) liegen muss. Unter dieser Annahme zeigte sich, dass bei 36 Kindern davon ausgegangen werden kann, dass sie von dem Lerntest profitiert haben und bei 16 Kindern, dies nicht der Fall war. Von den 16 Kindern haben sich fünf Kinder in ihrer Leistung verschlechtert, neun Kinder haben sich weder verbessert noch verschlechtert und zwei Kinder erfüllten das Kriterium eines wesentlichen Lerngewinns nicht. In der nächsten Abbildung (Abb. 4) wird die Verteilung des Lerngewinns der Experimentalgruppe noch einmal gesondert ohne die Merkmalsausprägungen der Kontrollgruppe veranschaulicht.

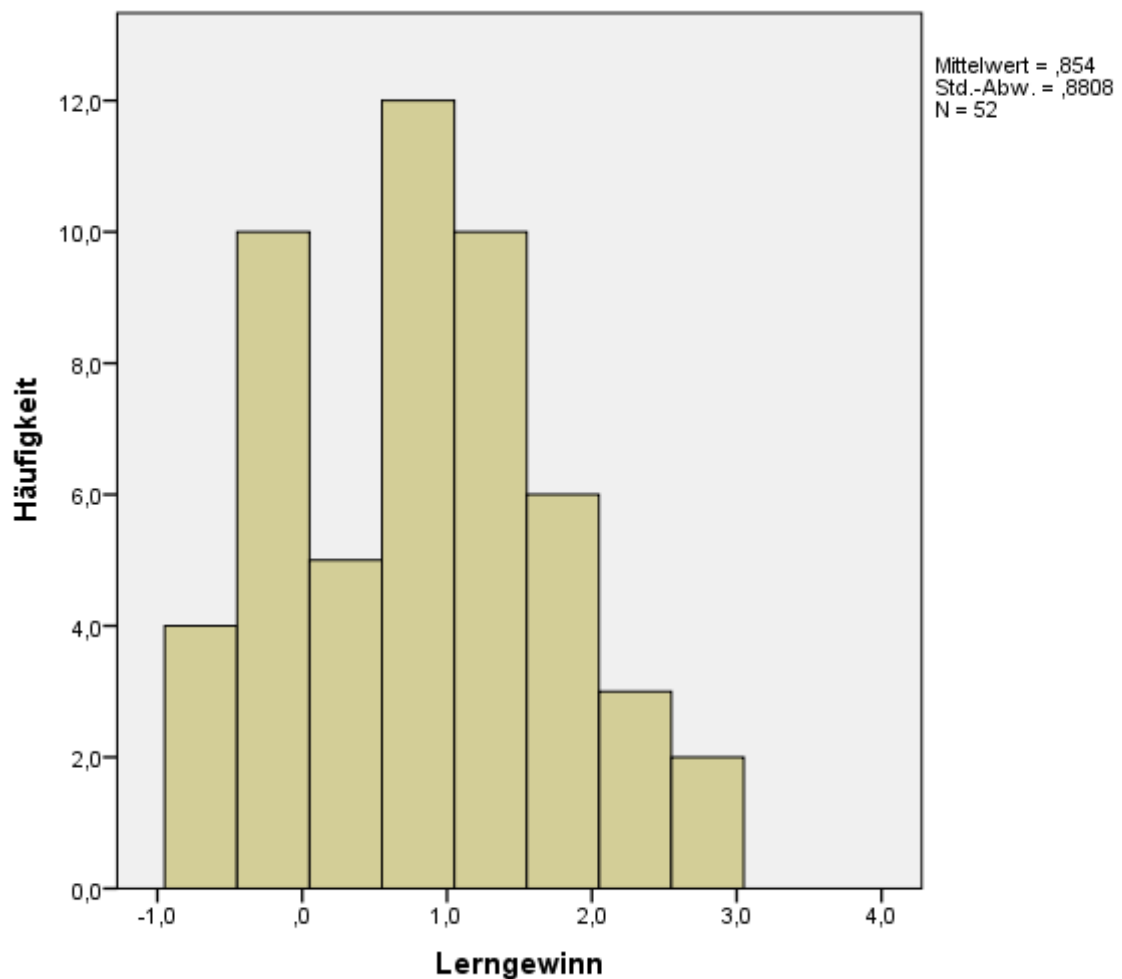


Abb. 4: Verteilung des Lerngewinns der Experimentalgruppe

Hypothese 2. Die zweite Hypothese besagt, dass sich die Posttestwerte des Lerntests nicht mit genügend hoher Wahrscheinlichkeit aus den Prätestwerten voraussagen lassen. Aus der Analyse geht hervor, dass die Prätestwerte und Posttestwerte in einem relativ hohen Ausmaß miteinander korrelieren ($\beta=.81$). Die Prätestwerte erklären 65,9% der Varianz der Posttestwerte ($R^2=.659$, $F(1,50)=96.525$, $p<.001$). Das bedeutet, dass ein erheblicher Teil der Varianz der Posttestwerte, also 34,1 % nicht durch die Prätestwerte erklärt werden kann. Aus diesem Grund kann die Hypothese Zwei ebenfalls als bestätigt angesehen werden.

Hypothese 3. Es konnte nicht gezeigt werden, dass sich die Testpersonen mit einem niedrigen/mittleren/hohen Lerngewinn bezüglich ihrer Prätestleistungen signifikant voneinander unterscheiden, $F(2,49)=2.471$, $p=0.095$ (siehe Mittelwerte und Standardabweichung der Gruppen in Tab. 2). Es ist jedoch eine Tendenz zu erkennen, dass die Gruppe mit einem niedrigen Lerngewinn im Mittel einer besseren Prätestleistung zeigte, als

die Gruppen mit einem mittleren und hohen Lerngewinn. Dies wird auch in der nächsten Abbildung (Abb.5) ersichtlich.

Tabelle 2

Mittelwerte und Standardabweichung der Prätestleistung in den Gruppen: niedriger/ mittlerer/ hoher Lerngewinn

Gruppe	Prätestleistung	
	M_P	SD_P
Niedriger Lerngewinn	1.06	1.62
Mittlerer Lerngewinn	0.12	1.27
Hoher Lerngewinn	0.31	0.95

Anmerkung. M_P = Mittelwert der Prätestleistung, SD_P = Standardabweichung der Prätestleistung

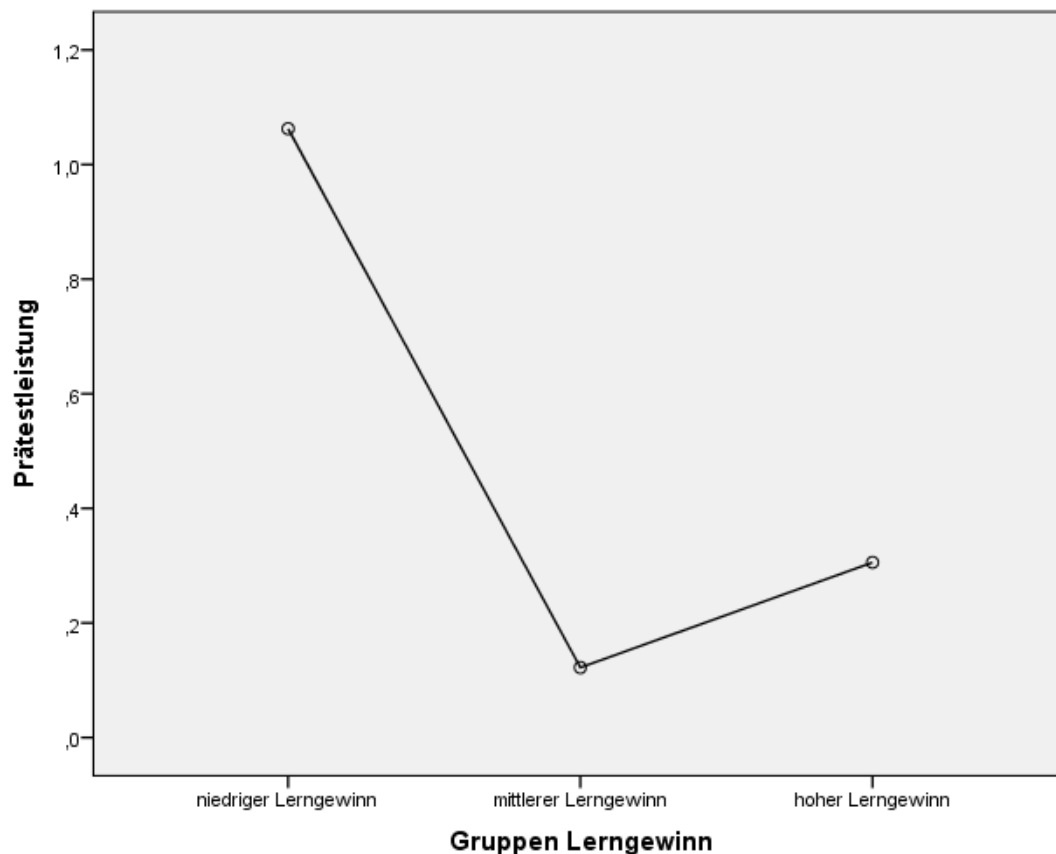


Abb. 5: Mittelwerte der Prätestleistung für die Gruppen: niedriger/mittlerer/hoher Lerngewinn

Hypothese 4. Mit der vierten Hypothese sollte untersucht werden, ob sich die Testpersonen mit niedrigem/ mittlerem/ hohem Lerngewinn bezüglich ihres Alters signifikant

voneinander unterscheiden. Die Hypothese, dass sich die Testpersonen mit einem niedrigen/mittleren/hohen Lerngewinn bezüglich ihres Alters signifikant voneinander unterscheiden, $H(2)=1.425$, $p=0.49$, konnte jedoch nicht bestätigt werden.

Hypothese 5. Betreffend der Fragestellung, ob es zwischen den drei Gruppen niedriger/mittlerer/hocher Lerngewinn bezüglich des Geschlechts signifikante Unterschiede gibt, hat die Analyse gezeigt, dass es keine signifikanten Unterschiede gibt, $\chi^2(2, n = 52)=1.361$, $p=.506$.

Hypothese 6. Nach der Durchführung des Chi-Quadrat Tests hat sich gezeigt, dass sich die Testpersonen mit einem niedrigen Lerngewinn/ mit einem mittleren Lerngewinn/ mit einem hohen Lerngewinn bezüglich ihrer Nationalität (Österreich versus alle anderen Nationalitäten) nicht signifikant voneinander unterscheiden $\chi^2(2, n = 52)= 1.118$, $p=.572$.

7. Diskussion

Das Ziel der vorliegenden Studie war es, einen Lerntest für den Untertest „Formale Folgerichtigkeit“ des AID 3 zu entwickeln, der in der Lage ist die Lernfähigkeit von Kindern zu erfassen. Um festzustellen, ob mit dem Lerntest die Lernfähigkeit erfasst werden kann, wurde untersucht, ob damit ein signifikant höherer Lerngewinn zu erzielen ist, der deutlich über den reinen Testwiederholungsgewinn hinausgeht. Auch sollte untersucht werden, ob mit dem Lerntest eine zusätzliche diagnostische Information erhoben werden kann, die durch den standardmäßigen Untertest nicht erfasst wird.

Die Ergebnisse zeigen, dass nachdem alle Kinder denselben Test ohne Lerntestintervention bereits einmal bearbeitet hatten, die Kinder, die anschließend mit dem Lerntest getestet wurden, einen signifikant höheren Lerngewinn erzielten, als die Kinder, die mit dem standardmäßigen Untertest wiederholt getestet wurden. Es konnte auch bestätigt werden, dass sich die Posttestwerte des Lerntests nicht mit genügend hoher Wahrscheinlichkeit aus den Prätestwerten der Standardtestung voraussagen lassen. Die Ergebnisse der Studie deuten somit darauf hin, dass es mit dem Lerntest möglich ist, die Lernfähigkeit zu erfassen, als auch dem Prätest und damit dem Untertest „Formale Folgerichtigkeit“ eine neue diagnostische Information hinzuzufügen. Somit kann aus den Ergebnissen geschlossen werden, dass die Umsetzung des Lerntestkonzepts nach Guthke beim Untertest „Formale Folgerichtigkeit“ des AID 3 erfolgreich war.

Neben diesen Ergebnissen der Studie wurde zusätzlich untersucht, ob sich die Testpersonen der Experimentalgruppe mit einem niedrigen Lerngewinn/mit einem mittleren Lerngewinn/mit einem hohen Lerngewinn bezüglich ihrer Prätestleistungen, ihrem Alter, ihrem Geschlecht und ihrer Nationalität (Österreich versus alle andere Nationalitäten) signifikant voneinander unterscheiden. Es hat sich gezeigt, dass es keine signifikanten Unterschiede zwischen den Gruppen weder in Bezug auf die Prätestleistung, auf das Alter, auf das Geschlecht, noch in Bezug auf die Nationalität (Österreich versus alle anderen Nationalitäten) gab. Bei der Untersuchung der Fragestellung, ob es signifikante Unterschiede zwischen den Testpersonen mit einem niedrigen/mittleren/hohen Lerngewinn bezüglich der Prätestleistung gibt, war das Ergebnis knapp nicht signifikant. So konnte eine Tendenz festgestellt werden, dass diejenigen Testpersonen mit einem niedrigen Lerngewinn, zuvor eine höhere Prätestleistung erzielten, als diejenigen Testpersonen mit einem mittleren oder hohen Lerngewinn. Aus dieser Tendenz könnte man möglicherweise ableiten, dass diejenigen Personen, die schon eine relativ hohe Leistung im Prätest erzielen, weniger stark von dem Lerntest profitieren, als die Testpersonen, die im Prätest schlechtere Leistungen zeigen. Diese

Fragestellung kann jedoch in der vorliegenden Untersuchung nicht eindeutig beantwortet werden, da das Ergebnis eben knapp nicht signifikant war. Es wäre sicherlich interessant, wenn zukünftige Studien versuchen diese Frage zu klären, da es allgemein wichtig ist, mehr darüber zu erfahren, für welche Personen der Lerntest besonders geeignet ist. Die bisherigen Erkenntnisse, dass vor allem Kinder kultureller Minderheiten stärker von Lerntests profitieren als die Kinder, die keiner kulturellen Minderheit in dem Land, wo die Studie durchgeführt wurde, angehören (Hessels & Hamers, 1993; Tzuriel & Kaufman, 1999), konnte in der vorliegenden Studie so nicht bestätigt werden. So haben sich keine signifikanten Unterschiede zwischen den Gruppen mit einem niedrigen Lerngewinn/mittleren Lerngewinn/hohen Lerngewinn bezüglich der Nationalität (Österreich versus alle anderen Nationalitäten) ergeben. Möglicherweise kann man aber die beiden referierten Studien mit der vorliegenden Untersuchung auch nicht gänzlich vergleichen, da bei dieser Studie in der Gruppe „andere Nationalitäten“ höchstwahrscheinlich nicht nur Immigrantenkinder oder Kinder kultureller Minderheiten miteinbezogen wurden, wie es bei den beiden anderen Studien der Fall war. In dieser Studie wurden die Kinder lediglich gefragt: „Aus welchem Land kommst du?“. Es wurde also nicht nach objektiven Kriterien entschieden, wann ein Kind in die Gruppe „Österreich“ oder in die Gruppe „andere Nationalität“ fällt, sondern es beruht lediglich auf der subjektiven Einschätzung und Angabe des Kindes selbst. So kann es sein, dass Kinder ein anderes Herkunftsland als Österreich angaben, aber trotzdem bereits in Österreich geboren, aufgewachsen und hier kulturell und sozial verwurzelt sind.

Bedeutend für die gesamte Studie ist, dass erstmals in einen bestehenden adaptiven Intelligenztest ein Lerntest integriert wurde, der in der Lage ist, die Lernfähigkeit zu erfassen. Trotzdem sollen an dieser Stelle auch ein paar Limitationen der Studie nicht unerwähnt bleiben. Bezüglich der Untersuchungsdurchführung ist anzumerken, dass die Pause zwischen der ersten und der zweiten Testung möglicherweise nicht lang genug war. So wurde während der Testungen festgestellt, dass sich manche Kinder noch sehr gut bei der zweiten Testung an die Items des Untertests „Formale Folgerichtigkeit“ erinnern konnten. Durch diesen Erinnerungseffekt kann nicht ausgeschlossen werden, dass dadurch Verzerrungen der Testleistungen einiger Versuchspersonen beim Posttest zustande gekommen sind. So haben sich einige Kinder gelangweilt gezeigt, dass sie die „selben“ Aufgaben wieder bearbeiten sollen. Möglicherweise haben sich dabei, durch die empfundene Langeweile, einige Kinder nicht mehr vollständig auf die Aufgaben konzentriert. So fiel auf, dass manche Kinder sich das Aufgabenmaterial nicht mehr genau angesehen haben und nur noch versucht haben aus ihrer Erinnerung die Antworten zu reproduzieren, die sie zuvor schon bei der ersten Testung

gegeben haben. Um dies in Zukunft zu verhindern, wird für weitere Studien empfohlen, dass bei Untersuchungen wie dieser, die zweite Testung, also der Posttest, an einem anderen Tag durchgeführt wird, damit die Erinnerung an die jeweiligen Aufgaben minimiert wird.

Des Weiteren ist unklar, warum bei den drei Testpersonen der Kontrollgruppe, die nachträglich ausgeschlossen wurden, so hohe Leistungssteigerungen auftraten. Da diese Testpersonen lediglich mit dem standardmäßigen Untertest „Formale Folgerichtigkeit“ wiederholt getestet wurden und keine Hilfestellungen bekamen, ist diese Fähigkeitssteigerung zu hinterfragen. Möglicherweise ist dies auf Eigenschaften der Testpersonen zurückzuführen, da keine Durchführungsfehler bei der Untersuchung auffielen, die dieses Testergebnis erklären könnten. Ein möglicher Grund für diese enormen Leistungssteigerungen könnte vielleicht sein, dass durch Sprachverständnisschwierigkeiten die Instruktion bei der ersten Testung nicht sofort verstanden und erst im Verlauf der Testung die Aufgabenstellung für die Testpersonen klarer wurde. Diese Annahme würden auch die Verhaltensbeobachtungen während der Testungen bekräftigen, da bei ein paar Kindern auffiel, dass diese Schwierigkeiten beim Verständnis der verbalen Instruktion hatten und erst bei der Bearbeitung der Aufgaben die Aufgabenstellung verstanden wurde. Es ist aber nicht klar, ob diese Beobachtung auch auf die Kinder, die ausgeschlossen wurden, zutraf. Mit Sicherheit lässt sich die Frage, weshalb solche Leistungssteigerungen auftraten, nicht beantworten.

Abgesehen davon fiel auf, dass der Lerntest, der wie auch der standardmäßige Untertest „Formale Folgerichtigkeit“ aus 3 Aufgabengruppen zu je 3 Items besteht, in seiner Testlänge sehr kurz war. Damit ist gemeint, dass bei einigen Kindern erst sehr spät bei der Testung die Hilfestellungen des Lerntests zur Anwendung kamen und so einige Kinder nicht mehr die Möglichkeit hatten, von den Hilfestellungen zu profitieren. Interessant wäre es bei zukünftigen Studien zu untersuchen, ob der Unterschied bezüglich des Lerngewinns zwischen der Versuchsgruppe und der Kontrollgruppe mit einer steigenden Anzahl an vorgegebenen Items noch größer wird und das Ergebnis an Signifikanz gewinnt.

Nichtsdestotrotz, hat sich gezeigt, dass der Lerntest die Lernfähigkeit erfassen kann und, dass das entwickelte Hilfesystem mit seinen Hilfestellungen für viele Kinder förderlich war. Obwohl der Lerntest nicht als Kurzzeitleerntest konzipiert ist, ist er im Vergleich zu anderen Langzeitleerntests um ein Vielfaches ökonomischer. Auch wenn die Lernfähigkeit über die Posttest-Prätest-Fähigkeitsparameterdifferenz erhoben wird und dabei zwei Testungen zur Messung der Lernfähigkeit erforderlich sind, so beanspruchen diese zwei Testungen nur eine geringe Durchführungsdauer. Ebenfalls fällt bei diesem Lerntest die Trainingsphase weg, da sie direkt in die zweite Testung integriert wurde. Für Praktiker, die

die Lernfähigkeit bei der Durchführung des AID 3 miterheben wollen, wäre zu empfehlen, dass diese den standardmäßigen Untertest „Formale Folgerichtigkeit“ gleich zu Beginn der Testung vorgeben und erst nach der Durchführung aller anderen ausgewählten Untertests, die zweite Testung mit dem Lerntest vornehmen. Bestenfalls geschieht die zweite Testung in einer anderen Testsitzung.

Da die Studie zeigt, dass mit einem relativ geringen Aufwand die Lernfähigkeit erfasst werden kann, wird ein wesentliches Argument der Kritiker am dynamischen Testen widerlegt, dass Lerntests zu aufwendig und zu zeitintensiv sind. Durchaus wird in Zukunft noch viel Forschung im Bereich des dynamischen Testens nötig sein, damit Lösungen für vorhandene messmethodische und psychometrische Probleme gefunden werden können. Es ist aber an der Zeit, dass dem Lerntestkonzept und dem dynamischen Testen allgemein in der Forschung, wie auch in der Praxis, wieder mehr Aufmerksamkeit geschenkt wird, da damit viele zuvor schon erwähnte Vorteile verbunden sind.

Zusammenfassend ist festzuhalten, dass es in dieser Arbeit gelungen ist, einen Lerntest zu entwickeln, der in der Lage ist die Lernfähigkeit von Testpersonen zu erfassen. Dabei wurde erstmals ein Lerntest erfolgreich in eine adaptive Intelligenztestbatterie integriert. Damit der Lerntest in der Praxis Einzug hält, könnte als nächster Schritt angedacht werden, Eich Tabellen für den Lerngewinn zu erstellen, um die Lernfähigkeit explizit messbar zu machen.

8. Literaturverzeichnis

- Beckmann, J. F. (2003). Statusdiagnostik. In K. D. Kubinger & R.S. Jäger (Hrsg.), *Schlüsselbegriffe der Psychologischen Diagnostik* (1. Aufl., S. 393–394). Weinheim: Beltz.
- Beckmann, J. F. (2004). Nachruf auf Jürgen Guthke. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 25(2), 117–119. doi:10.1024/0170-1789.25.2.117
- Caffrey, E., Fuchs, D., & Fuchs, L. S. (2008). The predictive validity of dynamic assessment: A review. *The Journal of Special Education*, 41(4), 254–270. doi:10.1177/0022466907310366
- Cole, M., John-Steiner, V., Scribner, S., & Souberman, E. (Hrsg.). (1978). *Mind in society: The development of higher psychological processes*. L. S. Vygotsky. Oxford, England: Harvard University Press.
- De Beer, M. (2005). Development of the learning potential computerised adaptive test (LPCAT). *South African Journal of Psychology*, 35(4), 717–747. doi:10.1177/008124630503500407
- Elliott, J. (2003). Dynamic assessment in educational settings: Realising potential. *Educational Review*, 55(1), 15–32. doi:10.1080/00131910303253
- Feuerstein, R., Rand, Y. A., & Hoffman, M. B. (1979). *The dynamic assessment of retarded performers: The learning potential assessment device, theory, instruments, and techniques*. Baltimore: University Park Press.
- Grigorenko, E. L. (2009). Dynamic assessment and response to intervention: Two sides of one coin. *Journal of Learning Disabilities*, 42(2), 111–132. doi:10.1177/0022219408326207
- Grigorenko, E. L., & Sternberg, R. J. (1998). Dynamic testing. *Psychological Bulletin*, 124(1), 75–111. doi:10.1037/0033-2909.124.1.75
- Guthke J. (1972). *Zur Diagnostik der intellektuellen Lernfähigkeit*. Berlin: Verlag der Wissenschaften.

- Guthke J. (1978). *Ist Intelligenz meßbar?: Einführung in Probleme der psychologischen Intelligenzforschung & Intelligenzdiagnostik*. Berlin: Deutscher Verlag der Wissenschaften.
- Guthke, J. (1992). Learning tests- the concept, main research findings, problems and trends. *Learning and individual differences*, 4(2), 137–151. doi:10.1016/1041-6080(92)90010-C
- Guthke, J. (Hrsg.). (1995). *Adaptive computergestützte Intelligenzlerntestbatterie (ACIL)*. Wien: Dr. Schuhfried.
- Guthke, J., Beckmann, J. F., & Wiedl, K. H. (2003). Dynamik im dynamischen Testen. *Psychologische Rundschau*, 54(4), 225–232. doi:10.1026/0033-3042.54.225
- Guthke, J., Jäger, C., & Schmidt, L. (1983). *Lerntestbatterie „Schlußfolgerndes Denken“*. Berlin: Psychodiagnostisches Zentrum.
- Guthke, J., & Stein, H. (1996). Are learning tests the better version of intelligence tests? *European Journal of Psychological Assessment*, 12(1), 1–13. doi:10.1027/1015-5759.12.1.1
- Guthke, J., & Wiedl, K.H. (1996). *Dynamisches Testen. Zur Psychodiagnostik der intraindividuellen Variabilität; Grundlagen, Verfahren und Anwendungsfelder*. Göttingen: Hogrefe.
- Hagenmüller, B. (2011). *Entwicklung des Untertests „Formale Folgerichtigkeit“ zur Erfassung von Reasoning in der Intelligenz-Testbatterie AID 3*. Unveröffentlichte Diplomarbeit, Universität Wien.
- Hamers, J. H., Hessels, M. G. P., & Pennings, A. H. (1996). Learning potential in ethnic minority children. *European journal of psychological assessment*, 12(3), 183–192. doi:10.1027/1015-5759.12.3.183
- Haywood, H. C., Brown, A. L., & Wingenfeld, S. (1990). Dynamic approaches to psychoeducational assessment. *School Psychology Review*, 19(4), 411–422. Verfügbar unter <https://www.researchgate.net/>

- Haywood, H. C., & Tzuriel, D. (2002). Applications and challenges in dynamic assessment. *Peabody Journal of Education*, 77(2), 40–63. Verfügbar unter <https://www.researchgate.net/>
- Hessels, M. G. P. (1997). Low IQ but high learning potential: Why Zeyneb and Moussa do not belong in. *Educational and Child Psychology*, 14(4), 121–136. Verfügbar unter <https://www.researchgate.net>
- Hessels, M. G. P., & Hamers, J. H. M. (1993). A learning potential test for ethnic minorities. In J.H.M.Hamers, K. Sijtsma & A.J.J.M. Ruijsenaars (Hrsg.), *Learning Potential Assessment: theoretical, methodological and practical issues* (S.285–311) Amsterdam: Swets & Zeitlinger.
- Kirschenbaum, R. J. (1998). Dynamic assessment and its use with underserved gifted and talented populations. *Gifted Child Quarterly*, 42(3), 140–147. doi:10.1177/001698629804200302
- Kubinger, K. D. (2009). *Psychologische Diagnostik. Theorie und Praxis psychologischen Diagnostizierens* (2. Aufl.). Göttingen: Hogrefe Verlag.
- Kubinger, K. D., & Holocher-Ertl, S. (2014). *Adaptives Intelligenz Diagnostikum 3 (AID 3)*. Weinheim: Beltz.
- Lauchlan, F., & Elliott, J. (2001). The psychological assessment of learning potential. *British Journal of Educational Psychology*, 71(4), 647–665. doi:10.1348/000709901158712
- Lidz, C. S. (1995). Dynamic assessment and the legacy of L.S. Vygotsky. *School Psychology International*, 16(2), 143–153. doi:10.1177/0143034395162005
- Lidz, C. S. (1996). Dynamic Assessment Approaches. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Hrsg.), *Contemporary intellectual assessment: Theories, tests, and issues* (S.281–296). New York: Guilford.
- Meumann, E. (1922). *Vorlesungen zur Einführung in die Experimentelle Pädagogik und ihren psychologischen Grundlagen: 1*. Leipzig: Engelmann.
- Murphy, R., & Maree, D. J. F. (2006). A review of South African research in the field of dynamic assessment. *South African Journal of Psychology*, 36(1), 168–191. doi:10.1177/008124630603600110

- Rasch, D., Kubinger, K. D., & Moder, K. (2011). The two sample *t*-test: pre-testing its assumptions does not pay off. *Statistical Papers*, 52, 219–231. doi:10.1007/s00362-009-0224-x
- Renz, T. (2018). *Unterschiede in der subjektiven Einschätzung der eigenen Leistung beim adaptiven und konventionellen Testen am Beispiel des AID 3*. Unveröffentlichte Masterarbeit, Universität Wien.
- Ruijsenaars, A.J.J.M., Castelijns J. H. M., & Hamers, J. H. M. (1993). The validity of learning potential tests. In J.H.M.Hamers, K. Sijtsma & A.J.J.M. Ruijsenaars (Hrsg.), *Learning potential assessment: theoretical, methodological and practical issues* (S.69–82). Amsterdam: Swets & Zeitlinger.
- Schack, T., & Guthke, J. (2003). Dynamic testing. *International Journal of Sport and Exercise Psychology*, 1(1), 40–60. doi:10.1080/1612197X.2003.9671703
- Schreiber, M. (2002). *Zur Optimierung der Testing-the-Limits-Strategie in der Demenzfrüherkennung: Der Adaptive Figurenfolgen-Lerntest (ADAFI)*. Unveröffentlichte Dissertation, Universität Wuppertal.
- Thorndike, E. L. (1922). *An introduction to the theory of mental & social measurement*. New York: Teachers College, Columbia University
- Tzuriel, D. (1989). Inferential cognitive modifiability in young socially disadvantaged and advantaged children. *International Journal of Dynamic Assessment and Instruction*, 1, 65–80.
- Tzuriel, D. (2000). Dynamic assessment of young children: Educational and intervention perspectives. *Educational Psychology Review*, 12(4), 385–435. doi:10.1023/A:1009032414088
- Tzuriel, D. (2001). Dynamic assessment of learning potential. In J. J. W. Andrews, D. H. Saklofski & H. L. Janzen (Hrsg.), *Handbook of psychoeducational assessment: ability, achievement, and behavior in children* (S. 451–497). San Diego: Academic Press.
- Tzuriel, D., & Kaufman, R. (1999). Mediated learning and cognitive modifiability: Dynamic assessment of young Ethiopian immigrant children to Israel. *Journal of Cross-Cultural Psychology*, 30(3), 359–380. doi:10.1177/0022022199030003005

Zubin, J. (1950). Symposium on statistics for the clinician. *Journal of Clinical Psychology*, 6(1), 1–6. doi:10.1002/1097-4679(195001)6:1<1::AID-JCLP2270060102>3.0.CO;2-O

9. Abbildungsverzeichnis

1 Beispielaufgabe aus dem Untertest „Formale Folgerichtigkeit“ des AID 3	S. 19
2 Prä- und Posttestmittelwerte der Experimentalgruppe und der Kontrollgruppe	S. 30
3 Verteilung des Lerngewinns für die Experimentalgruppe und für die Kontrollgruppe	S. 31
4 Verteilung des Lerngewinns für die Experimentalgruppe	S. 32
5 Mittelwerte der Prätestleistung für die Gruppen: niedriger/mittlerer/hoher Lerngewinn	S. 33

10. Tabellenverzeichnis

1 Deskriptivstatistik der beiden Versuchsbedingungen	S. 28
2 Mittelwerte und Standardabweichung der Prätestleistung in den Gruppen: niedriger/ mittlerer/hocher Lerngewinn	S. 33

11. Anhang

11.1 Anhang A- Abstract (Deutsch)

Das Ziel der Studie war es einen Lerntest nach dem Lerntestkonzept von Guthke für den Untertest „Formale Folgerichtigkeit“ des Adaptiven Intelligenz Diagnostikums 3 (Kubinger & Holocher-Ertl, 2014) zu entwickeln. Überprüft werden sollte dabei, ob der Lerntest in der Lage ist, die Lernfähigkeit von Kindern zu erfassen und ob dieser dem standardmäßigen Untertest eine neue diagnostische Information hinzufügt. Des Weiteren sollte untersucht werden, ob sich die Testpersonen der Experimentalgruppe mit einem niedrigen/mittleren/hohen Lerngewinn hinsichtlich ihrer Prätestleistung, ihrem Alter, ihrem Geschlecht und ihrer Nationalität (Österreich versus alle anderen Nationalitäten) unterscheiden.

In einer experimentellen Studie wurden die Leistungen von insgesamt 101 Kindern, die für die Studie rekrutiert werden konnten, zu zwei Messzeitpunkten erhoben. Nachdem in der ersten Testung alle Kinder mit dem standardmäßigen Untertest „Formale Folgerichtigkeit“ getestet wurden, fand vor der zweiten Testung eine zufällig Zuteilung zur Experimentalgruppe oder zur Kontrollgruppe statt. Die zur Versuchsgruppe zugeteilten Kinder wurden in der zweiten Testung mit dem Lerntest getestet und die Kinder der Kontrollgruppe wiederholt mit dem standardmäßigen Untertest.

Die Ergebnisse zeigen, dass die Umsetzung des Lerntestkonzepts nach Guthke beim Untertest „Formale Folgerichtigkeit“ erfolgreich war. Es konnte gezeigt werden, dass mit dem Lerntest sowohl die Lernfähigkeit der Kinder erfasst wird als auch, dass der Lerntest eine neue diagnostische Information bereitstellt, die durch den standardmäßigen Untertest nicht erhoben wird. Es zeigten sich keine signifikanten Unterschiede zwischen den Gruppen: niedriger/mittlerer/hoher Lerngewinn bezüglich der Prätestleistung, des Alters, dem Geschlecht und der Nationalität. Bedeutend für diese Studie ist, dass erstmals in eine bereits bestehende adaptive Intelligenztestbatterie ein Lerntest integriert wurde.

11.2 Anhang B- Abstract (English)

The goal of the present study was to develop a learning test following the learning test concept by Guthke for the subtest “Formale Folgerichtigkeit” of the AID 3 (Kubinger & Holocher-Ertl, 2014). The study sought to review whether the learning test is able to assess the learning ability of children and whether it adds new diagnostic information to the standard subtest “Formale Folgerichtigkeit”. Furthermore, it was examined whether the test persons of the experimental group showed any differences with regard to a low/average/high learning gain on the basis of their pretest results, their age, their sex or their nationality.

In an experimental study, the test performances of 101 children were assessed at two points of measurement. After testing all children with the standard subtest at the first point of measurement, they were randomly assigned to either the experimental group or the control group. At the second point of measurement, the children from the experimental group were tested with the learning test and the children from the control group repeated the standard subtest.

The results show that the learning test is able to assess the learning ability of children as well as it gathers new diagnostic information which is not assessed in the standard subtest. No significant differences were identified between the groups concerning a low/average/high learning gain with regard to pretest results, age, sex or nationality. The study is considerable in that it is the first time that a learning test was successfully integrated in an established adaptive intelligence test.

11.3 Anhang C- Einverständniserklärung der Eltern

Liebe Eltern und Erziehungsberechtigte,

Viele Kinder werden während ihrer Schullaufbahn aus den unterschiedlichsten Gründen mit psychologischen Verfahren untersucht. Damit die Schullaufbahnberatung und Schulpsychologie immer auf dem neuesten Stand der Wissenschaft ist und Verbesserungen der psychologischen Tests vorgenommen werden können, ist die Forschung in diesem Bereich sehr wichtig.

Im Rahmen unserer Masterarbeiten an der Fakultät für Psychologie an der Universität Wien führen wir beide unter der Leitung von Univ. Prof. Dr. Mag. Klaus D. Kubinger zwei Studien durch, wo wir in den nächsten Wochen mit Einverständnis der Direktion, in der Klasse ihres Kindes eine Testung mithilfe des AID 3 (Adaptives Intelligenz Diagnostikum) durchführen. Speziell für diese Testung werden den Kindern: Textrechenaufgaben vorgegeben; Figurenfolgen vorgegeben, die man um eine weitere Figur ergänzen soll; Würfel zur Verfügung gestellt, bei denen man geometrische Muster anhand einer Vorlage nachbauen soll; Fragen zum Alltagswissen gestellt.

Bei der ersten Studie geht es darum, wie Kinder je nach Art der Testung ihre eigene Leistung beim Testen einschätzen. Dabei werden zwei Arten der Testung miteinander verglichen: das adaptive Testen, wo die Aufgabenauswahl während der Testung an die Fähigkeit des Kindes angepasst wird und das konventionelle Testen, wo alle Kinder die gleichen Aufgaben bekommen. Interessant ist der Unterschied in der Selbsteinschätzung deshalb, weil sie u.a. Einfluss auf die Motivation der Kinder während der Testung und damit auch auf die gemessene Leistung selbst haben könnte.

Für die zweite Studie wurde für einer der vorher aufgeführten Aufgabengruppen ein spezieller Lerntest entwickelt. Ein Lerntest soll die Lernfähigkeit und das zukünftige Potential des Kindes ermitteln. Das Ziel dieser Studie ist zu prüfen, ob mit dem entwickelten Lerntest auch wirklich die Lernfähigkeit erfasst werden kann.

Wir wenden uns daher mit der Bitte an Sie, Ihr Kind an dieser Untersuchung teilnehmen zu lassen, vorausgesetzt natürlich, dass Sie und Ihr Kind damit einverstanden sind. Die Untersuchung findet während der Schulzeit statt und dauert zwischen einer halben Stunde und einer Stunde. Aus Erfahrung macht den Kindern die Mitarbeit an den Aufgaben viel Spaß. Die gewonnenen Daten werden im Sinne des Datenschutzrechtes ausschließlich für wissenschaftliche Zwecke genutzt. Die Anonymisierung sämtlicher Ergebnisse der

SchülerInnen findet bereits während der Testung statt. Wir bitten Sie, bis _____, mit Ihrer Unterschrift auf dem Abschnitt unten Ihr Einverständnis zur Teilnahme Ihres Kindes an den beschriebenen Untersuchungen zu erteilen.

Mit freundlichen Grüßen und herzlichem Dank im Voraus!

xxxxx & xxxxx

-----bitte abtrennen-----

Ich erkläre mich an der Teilnahme meiner Tochter/ meines Sohnes _____, die oder der die Klasse ____ besucht, geboren am _____, an der Testung der Masterarbeiten von xxxx & xxxx einverstanden.

Datum, Unterschrift _____