



universität
wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Diagnostik des phonologischen Gedächtnisses
im Kindergartenalter mittels neuartigen im Vergleich zu
vertrautem Material: Konstruktion und psychometrische
Analyse eines Itempools“

verfasst von / submitted by

Sarah Macura, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Master of Science (MSc)

Wien, 2017 / Vienna 2017

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:

Ass.-Prof. Dr. Pia Deimann

Danksagung

In erster Linie danke ich herzlich Frau Ass.-Prof. Dr. Deimann und Frau Ass.-Prof. Dr. Kastner-Koller für die Möglichkeit, meine Masterarbeit in einem für mich sehr interessanten Forschungsbereich verwirklichen können, für ihre intensive Betreuung, Unterstützung und wertvollen, konstruktiven Anregungen.

Ein großer Dank gilt auch den pädagogischen Leiterinnen der *St. Nikolausstiftung* und der *Kinder in Wien* (KIWI) für die Genehmigung zu Datenerhebungen an den Kindergärten, ebenso wie den Kindergartenleiterinnen, Kindergartenpädagoginnen und dem gesamten Team der Kindergärten, welche durch freundliche und hilfsbereite Kooperation die Durchführung vorliegender Untersuchung unterstützten. Von Herzen danke ich allen Kindern, die durch ihre Mitwirkung das Fundament für die vorliegende Untersuchung bildeten sowie mir viel Freude bei den Erhebungen und praktische Erfahrungen ermöglichten. Ebenso danke ich den Eltern der Kinder für ihr Einverständnis zur Teilnahme an der Untersuchung.

Bei meinen Kolleginnen Minna Felkel und Ina Stacher möchte ich mich für die positive Zusammenarbeit während der Planungs- und Erhebungsphasen und bei der Generierung eines gemeinsamen Datenpools sowie für den freundlichen Austausch und die gegenseitige Motivierung bedanken. Meinen Freundinnen Ricarda Berger und Henriette Schulz danke ich für das sorgfältige Korrekturlesen meines englischen Abstracts.

Zutiefst dankbar bin ich meiner Familie, die mir stets in jeder Hinsicht eine wichtige Stütze war. Ich bedanke mich innigst bei meiner Mutter, Erika, für die fortwährende, liebevolle emotionale und große instrumentelle Unterstützung, für die inspirierenden Gespräche sowie das gewissenhafte Korrekturlesen. Ein lieber Dank gilt meinem Neffen Ben, der mir mit seiner Mitarbeit einen fröhlichen Start in die Voruntersuchungen schenkte. Ich bedanke mich auch herzlich bei meiner Schwester Julia, meinem Bruder Michele und meiner Schwägerin Beate für die Hilfe bei der Rekrutierung der Stichprobe für Vor- und Probeuntersuchungen. Ein liebes Dankeschön richtet sich auch an meinen Freund, Manuel, der mir in arbeitsintensiven Phasen im Rahmen meiner Masterarbeit mit Geduld und Verständnis zur Seite stand und mich bei der Verwirklichung meiner Ziele unterstützte.

Inhaltsverzeichnis

Abstract	7
Deutsch.....	7
Englisch.....	8
1. Einleitung und theoretischer Hintergrund	9
1.1 Phonologisches Gedächtnis.....	9
1.1.1 Struktur und Prozesse.....	9
1.1.2 Entwicklung des phonologischen Gedächtnisses.....	12
1.1.3 Biologische und umweltbedingte Einflussfaktoren.....	13
1.1.4 Entwicklungsrelevante Funktionen des phonologischen Gedächtnisses.	14
1.2 Psychologische Diagnostik des phonologischen Gedächtnisses	15
1.2.1 Kunstwörter versus Ziffernfolgen in der Diagnostik bei Kindergartenkindern.	17
1.2.2 Mechanismen und Fähigkeitskomponenten beim Nachsprechen von Kunstwörtern.	18
1.3 Entwicklung von Nachsprechaufgaben mit Kunstwörtern.....	20
1.3.1 Itemkonstruktion.	20
1.3.2 Psycholinguistischer Kontext in der Konstruktion von Kunstwörtern.....	21
1.3.3 Aufgabendesign.....	23
1.3.4 Messinstrumente mit Kunstwörtern.	24
1.4 Vorliegende Untersuchung.....	27
2. Methode.....	29
2.1 Versuchsplan / Untersuchungsdesign.....	29
2.2 Stichprobenbeschreibung	30
2.3 Messinstrumente.....	33
2.3.1 Phonologisches Gedächtnis bei neuartigem Material.	33
2.3.2 Phonologisches Gedächtnis bei vertrautem Material.	36
2.3.3 Konstruktferne Fähigkeitsdimensionen: Grob - und Visuomotorik.....	36
2.4 Vorgehen	37
2.5 Datenanalyse	38
3. Ergebnisse	42
3.1 Fehlende Daten.....	42
3.2 Gedächtnisleistungen in Teilstichproben	42
3.3 Vorläufige psychometrische Analysen der Gesamtdaten.....	44
3.3.1 Psychometrische Qualität nach probabilistischer Testtheorie.....	44

3.3.2 Psychometrische Qualität nach klassischer Testtheorie.....	46
3.4 Selektiertes Itemset: Psychometrische Qualität im Vergleich mit vertrautem Material	47
3.4.1 Leistungen beim selektierten Itemset.....	48
3.4.2 Überprüfung des selektierten Itemsets nach probabilistischer Testtheorie.....	50
3.4.3 Itemanalyse des selektierten Itemsets nach klassischer Testtheorie.	50
3.4.4 Objektivität.....	52
3.4.5 Reliabilität.	52
3.4.6 Validität.....	52
3.5 Zusatzergebnisse: Analyse des Arbeitsverhaltens bei KWN-Aufgaben	54
4. Diskussion	54
4.1 Eignung des KWN-Itemsets zur Erfassung des phonologischen Gedächtnisses im Rahmen eines Entwicklungstests	55
4.2 Praktische Implikationen.....	59
4.3 Limitationen	64
4.4 Ausblick	65
4.5 Conclusio.....	68
5. Literaturverzeichnis.....	69
6. Tabellenverzeichnis.....	96
7. Abbildungsverzeichnis	96
8. Anhang	97
8.1 Anhang A: Ergänzende Stichprobeninformationen	97
8.2 Anhang B: Informationen an Eltern und KindergartenleiterInnen / -pädagogInnen und Einverständniserklärung.....	98
8.3 Anhang C: Protokollbögen und bereitgestellte Erklärungen/Instruktionen	100
8.4 Anhang D: Voraussetzungstests.....	107
8.5 Anhang E: Ergänzende Deskriptive Statistiken und interferenzstatistische Analysen	110
8.6 Anhang F: Itemspezifische Rasch-Analysen (Wald-Tests) vor Itemselektion	115
8.7 Anhang G: Itemkennwerte	125
8.8 Anhang H: Beurteilerinnenübereinstimmung und Intracoder-Reliabilität.....	139
8.9 Anhang I: Vorschläge zur Modifikation der Items	143

Abstract

Deutsch

Das phonologische Gedächtnis bei Kindern von 3 bis 6 Jahren ist zentral für Lernprozesse, für die kognitive und sprachliche Entwicklung und eine wichtige Vorläuferfähigkeit für spätere schriftsprachliche und rechnerische Fertigkeiten. Eine frühzeitige Identifikation von Defiziten im phonologischen Gedächtnis ist daher grundlegend und erfordert Aufgaben mit hoher psychometrischer Qualität, die der Erfahrungswelt junger Kinder entsprechen. Ziel vorliegender Untersuchung war, einen Itempool mit neuartigem phonologischen Material in Form einer Nachsprechaufgabe mit Kunstwörtern zu konstruieren, welche hinsichtlich der Eignung für ein allgemeines Entwicklungstestverfahren einer Gedächtnisaufgabe mit vertrautem Material gegenübergestellt wurde. Die Analytestichprobe umfasste $N = 58$ Kinder im Alter von 3;1 bis 5;5 Jahren. Diese wurden randomisiert einem von zwei Itemsets ($n_1 = 31$, $n_2 = 27$) mit jeweils 20 konstruierten 2- bis 6-silbigen Kunstwörtern zugeteilt, bei gemeinsamer Anwendung von etablierten Ziffernfolgen sowie konstruktfernen Aufgaben aus dem Wiener Entwicklungstest. Psychometrische Analysen der Itemsets und diverser Bewertungsmodi mittels probabilistischer und klassischer Methoden wurden angewandt. Eine Itemselektion eines Itemsets (bei n_1) resultierte in 12 Items mit weitgehender Gültigkeit des Rasch-Modells, guter Auswertungsobjektivität, hoher Reliabilität ($\alpha = .81$, $r_{tt} = .85$) und Konstruktvalidität, indiziert durch hohe Korrelationen von Nachsprechleistungen mit Ziffernspannen bei fehlendem Zusammenhang mit motorischen Leistungen. Nachsprechleistungen beider Gedächtnisaufgaben ergaben einen linearen Alterstrend, mit kontinuierlicherem Verlauf bei Ziffernfolgen, sowie große signifikante Unterschiede zwischen 3- und 4-jährigen Kindern. Die Diskussion der Ergebnisse umfasst den Bezug zur Sprachentwicklung. Der kombinierte Einsatz von Nachsprechaufgaben mit Kunstwörtern und Ziffernfolgen erscheint hinsichtlich einer differenzierten Diagnostik des phonologischen Gedächtnisses vielversprechend.

Englisch

In children aged 3 to 6 years phonological memory is pivotal for learning processes, for cognitive and language development and is an important precursor skill for later literacy and numeracy. Therefore, it is crucial to identify deficits in phonological memory at an early point of time. This requires tasks of high psychometric quality that match young children's world of experience. The aim of the current study was the construction and psychometric analysis of an item pool consisting of novel phonological material in the form of a nonword repetition task which was compared with a recall task of familiar phonological material regarding its suitability as a task for a general developmental test. The analysis sample comprised $N = 58$ children aged 3;1 to 5;5 years. They were randomly assigned to one of two item sets ($n_1 = 31$, $n_2 = 27$), each consisting of 20 constructed 2- to 6-syllable nonwords. A well-established digit recall task and theoretically dissimilar measures of the *Wiener Entwicklungstest* were administered alongside. Psychometrical analysis on the item sets and diverse scoring modes using probabilistic and classical methods were applied. An item selection of one item set (based on n_1) resulted in 12 items which widely met expectations of the Rasch model, showed good objectivity of scoring, high reliability ($\alpha = .81$, $r_{tt} = .85$) and construct validity indicated by high correlation with digit span and no correlation with motor scores. Both repetition scores showed a linear age trend which was more continuous for digits. There were large significant differences between 3- and 4-year-old children. The discussion of results includes the reference to language development. The combined administration of nonword repetition tasks along with digit recall tasks seems to be promising in terms of a differentiated diagnostic assessment of phonological memory.

Einleitung und theoretischer Hintergrund

Kinder im Alter von 3 bis 6 Jahren (Kindergartenkinder) befinden sich in einer Phase rapider, sich kontinuierlich ausdifferenzierender Entwicklung, außergewöhnlicher Gehirnplastizität und Sensitivität für Umwelteinflüsse (Fitzpatrick & Pagani, 2012; Marsh, Gerber, & Peterson, 2008). In dieser Phase lernen sie täglich durch Zuhören sprachklanglicher (phonologischer) Informationen aus kurzen neuen Sprechereignissen, z.B. in Gesprächen, gemeinschaftlichen Spielsituationen und der Beschäftigung mit Informations- und Kommunikationstechnologien (Gupta & Tisdale, 2009; Havy, Nazzi, & Bertoncini, 2013; Hoffman, Teale, & Paciga, 2014; Kraus & White-Schwoch, 2015). Kindergartenkinder können Wörter unmittelbar nach einmaliger Darbietung mühelos nachsprechen und dadurch Wissen sowie Wortschatz erweitern (Brady & Goodman, 2014; Côte, Rouleau, Macoir, 2014; Kan & Kohnert, 2012; Krishnan, Alcock, Carey, Bergström, Karmiloff-Smith, & Dick, 2017; Storkel, 2015). Das phonologische Gedächtnis bildet die Grundlage für die Fähigkeit, phonologische Informationen wie ein neues Wort bzw. dessen willkürlich seriell angeordnete Silbensequenzen sowie Konsonanten und Vokale als kleinste Lauteinheiten (Phoneme) – aufgrund der Flüchtigkeit akustischen Inputs (Hartley & Houghton, 1996) – in Echtzeit verarbeiten, kurzfristig speichern und abrufen zu können; infolgedessen ist das phonologische Gedächtnis zentral für Entwicklungsschritte von Kindergartenkindern im kognitiven, sprachlichen und sozial-kommunikativen Bereich (Pugin, Metz, Stauffer, Wolf, Jenni, & Huber, 2014; Rose, Feldman, & Jankowski, 2009; von Goldammer, Mähler, Bockmann, & Hasselhorn, 2010; White-Schwoch et al., 2015).

Phonologisches Gedächtnis

Das phonologische Gedächtnis ist eine von drei phonologischen Verarbeitungsfähigkeiten, welche außerdem die Sensitivität für die Lautstruktur und bewusste Operationsfähigkeit mit Silben und Phonemen einer gesprochenen Sprache (phonologische Bewusstheit bzw. metaphonologische Fähigkeit) sowie die Schnelligkeit beim Zugriff auf die Wissensbasis (das mentale Lexikon) im Langzeitgedächtnis (Benennungsgeschwindigkeit) umfassen (Lonigan et al., 2009; Ptok et al., 2008; Preßler, Könen, Hasselhorn, & Krajewski, 2014; Wagner & Torgesen, 1987).

Struktur und Prozesse.

Hinsichtlich theoretischer Rahmenmodelle zum phonologischen Gedächtnis gilt das kapazitätslimitierte Mehrkomponentenmodell zum Arbeitsgedächtnis (Baddeley, 2000, 2007; Baddeley & Hitch, 1974) bei Kindergartenkindern als das am gründlichsten untersuchte sowie – nach Schumacker & Lomax (2004) – mit akzeptabler bis guter Anpassungsgüte nachgewiesene (Alloway, Gathercole, & Pickering, 2006; de Abreu, Conway, & Gathercole, 2010; Gathercole,

Pickering, Ambridge, & Wearing, 2004; Vugs, Knoors, Cuperus, Hendris, & Verhoeven, 2015). Nach dem Modell erfolgen im Arbeitsgedächtnis modalitätsspezifische intentionale oder inzidentelle Informationsaufnahme und -verarbeitung (Enkodierung), kurzfristige Speicherung sowie expliziter oder impliziter Abruf in zwei separat arbeitenden Subsystemen, einem für visuell-räumliche Informationen (*Räumlich-visueller Notizblock*) und einem für klangliches (z.B. Melodien), phonologisches und in einen artikulatorischen Code konvertiertes schriftliches Material (*Phonologische Schleife*) (Claessen, Leitão, Kane, & Williams, 2013). Die modalitätsspezifischen Subsysteme stehen in Interaktion mit einem modalitätsübergreifenden Steuerungs- und Kontrollsystem für Gedächtnis- und Aufmerksamkeitsprozesse (*zentrale Exekutive*), theoretisch und empirisch überschneidend mit dem Konstrukt der exekutiven Funktionen (Schmid, Zoelch, & Roebers, 2008), sowie mit einer Instanz zur Integration multimodaler Informationen aus den Subsystemen in Rückkoppelung mit dem Langzeitgedächtnis (*episodischer Puffer*), und sind vor dem 6. Lebensjahr funktionell nicht zuverlässig von diesen differenzierbar (Allo-way, Gathercole, Willis, & Adams, 2004; Gathercole & Pickering, 2000; Pickering & Gathercole, 2001). Die spätestens ab dem 4. Lebensjahr verfügbare phonologische Schleife umfasst als Input-Output-System eine – aufgrund seiner passiven, unstrategischen Funktionen als Kurzzeitgedächtnis konzeptualisierte – Speicherinstanz (*phonologischer / phonetischer Speicher*), welche Enkodierung, temporäre Speicherung in sprachlich basierter Form und Abruf in einem Zeitintervall von 1.5 bis 2 Sekunden ermöglicht, sowie einen, ab dem 7. Lebensjahr aktiv werdenden, *subvokalen artikulatorischen Kontrollprozess*, welcher mentale Wiederholungen (*subvokales rehearsal*) ermöglicht und dadurch erlaubt, Repräsentationen im *phonologischen Speicher* über dessen Behaltensintervall hinaus aufrechtzuerhalten und vor dem Zerfall zu bewahren sowie – ab dem Erwerb von Lesefähigkeiten – visuell aufgenommene schriftsprachliche Informationen in sprachliche zu konvertieren (Baddeley, Thomas, & Buchanan, 1975; Cowan, 2016; Hummel & French, 2016; Kiese-Himmel, 2011; Oberauer & Lewandowsky, 2010; Roebers & Zoelch, 2005).

Nach aktuellen, in Sprachverarbeitungsmodelle integrierten Konzeptionen zum phonologischen Gedächtnis mit Annahmen zu separaten Zwischenspeichern für Sprachwahrnehmung (Input) sowie -produktion (Output) oder unterschiedlichen Verarbeitungs- und Speicherungsrou-ten beim Nachsprechen von lexikalischen, bekannten sowie von neuen Wörtern (z.B. Jacquemot & Scott, 2006), ist für das Nachsprechen phonologischer Informationen ohne lexikalischer oder semantischer Verarbeitung ein spezifischer Mechanismus zur Konvertierung von phonologischem Input in Output sowie vice versa verantwortlich. Gestützt auf neuro-bildge-

bende Studien (z.B. Fiebach, Friederici, Smith, & Swinney, 2007; Majerus et al., 2010) indizieren Netzwerk- und interaktive Aktivierungsmodelle zur Sprachverarbeitung (z.B. Brown, Preece, & Hulme, 2000; Burgess & Hitch, 2005; Gupta, 2003; Grossberg & Kazerounian, 2011) zudem spezialisierte, hoch vernetzte Prozessmodule für separate Verarbeitung und Repräsentationsbildung der seriellen Anordnung von Phonemen und Phonemsequenzen über zeitlich-, kontextbasierte oder positionelle Codes (*Reihenfolgeinformationen*) sowie der phonologischen, die Lautgestalt (sublexikalische), Wortform (lexikalische) und Wortbedeutung (semantische) betreffende Charakteristika (*Iteminformationen*) über davon abgrenzbare Codes und temporäre Aktivierung von Sprachnetzwerken im Langzeitgedächtnis (Côte et al., 2014; Leclercq & Majerus, 2010; Perez, Majerus, & Poncelet, 2012; Majerus & D'Argembeau, 2011; Majerus & Boukebza, 2013). Mechanismen für Reihenfolge- und für Iteminformationen liegen diversen Gedächtnisphänomenen sowie mit einem Leistungsabfall einhergehenden Interferenzprozessen beim Abruf (Output- bzw. retroaktive Interferenz) zugrunde (Hurlstone, Hitch, & Baddeley, 2014); dazu zählen der, innerhalb einer umgekehrt u-förmigen seriellen Positionskurve, scharfe monotone Leistungsabfall ab Items in initialer Position (*primacy effect*) bei geringerem Leistungsaufschwung bei Items in finaler Position (*recency effect*), sowie die sich verringern- de Abrufgenauigkeit bei längeren Itemsequenzen (*Sequenz-Längeneffekt*; Baddeley et al., 1975), phonologisch ähnlichen Einsilbern, Silben oder Phonemen einer seriellen Sequenz (*phonologischer Item-Ähnlichkeitseffekt*; Baddeley, 1986), aufgabenirrelevanter Hintergrundsprache (*Irrelevanter Spracheffekt*; Colle, 1980; Colic & Welsh, 1976) und experimentell induzierten fortwährenden Wiederholungen irrelevanter Sprachreize während des Abrufs (*Artikulatorischer Unterdrückungseffekt*; Jones & Macken, 1995; Murray, 1967).

Im Rahmen typischer Aufgaben zum seriellen Abruf phonologischer Informationen ansteigenden Umfangs (*serial recall tasks*) umfassen etablierte Indikatoren für die Speicherkapazität des phonologischen Gedächtnisses die maximale Anzahl an chronologisch korrekt reproduzierten Folgen vertrauter einsilbiger Wörter, wie Ziffern von 1 bis 9 (ausgenommen 7) (*Gedächtnis- bzw. Ziffernspanne*), sowie die Leistung beim Nachsprechen von, in der Silbenanzahl variierenden, künstlichen Wörtern (*Kunstwörtern*) – inkonsistent je nach Wortähnlichkeit oder Artikulierbarkeit auch als Pseudo- oder Nichtwörter bezeichnet – welche als neuartige, nicht-lexikalische phonologische Sequenzen ohne semantischem Wert neben der Größe des phonetischen Speichers die Dimension der Verarbeitungsgenauigkeit messen können (Archibald & Griebeling, 2016; Perez et al., 2012; Schuchardt, Maehler, & Hasselhorn, 2011; Mathieu et al., 2016; von Goldammer et al., 2010).

Entwicklung des phonologischen Gedächtnisses.

Quer- und längsschnittliche Untersuchungen zeigen vom Kindergartenalter bis in die Adoleszenz lineare, kontinuierliche, insgesamt 2- bis 3-fache Kapazitätswachse im phonologischen Gedächtnis, mit graduell bis asymptotischem Verlauf ab einem Alter von 10 bis 11 bzw. 14 bis 15 Jahren (Alloway et al., 2006; Gathercole, Pickering et al., 2004; Nevo & Breznitz, 2013; Waber et al., 2007). Kindergartenkinder erbringen im Durchschnitt eine Merkleistung von ca. 3 Ziffern und können sich – bei signifikanten Leistungsunterschieden in Ganzjahres- sowie in Halbjahresschritten – mit 3 Jahren ca. ein bis drei Ziffern, mit 4 Jahren ca. drei bis vier Ziffern und mit 5 Jahren ca. vier Ziffern merken (Kastner-Koller & Deimann, 2012). Bei Kunstwörtern und Ziffernfolgen zeigen sich die substantiellsten Leistungssteigerungen – nach Cohen (1988) – mit großer Effektstärke im Alter von 3 auf 4 Jahre, bei Ziffernfolgen um ca. eine bis ca. zwei Standardabweichungen und bei Kunstwörtern um ca. eine bis ca. drei Standardabweichungen (z.B. Ziffernfolgen: Cohens d mit Hedges-Korrektur $[g] = 1.05-1.96$; Kunstwörter $g = 0.89-3.17$; Dollaghan & Campbell, 2009; Polišenská & Kapalková, 2014; Jones, Gobet, & Pine, 2007; Guiberson & Rodríguez, 2013; Ulset, Vitaro, Brendgen, Bekkhus, & Borge, 2017). Im Alter von 4 auf 5 Jahre zeigen sich mittlere bis große Leistungsanstiege, bei Ziffernfolgen sowie Kunstwörtern um ca. eine halbe bis eine Standardabweichung (z.B. Ziffernfolgen: $g = 0.57-0.86$; Kunstwörter: $g = 0.53-0.92$; Polišenská & Kapalková, 2014; Radeborg, Barthelom, Sjöberg, & Sahlen, 2006; Risse & Kiese-Himmel, 2009; Schmid et al., 2008; Sundström, Samuelsson, & Lyxell, 2014); weitere Befunde indizieren eine Stagnation der phonologischen Gedächtnisleistungen zwischen dem Alter von 4 bis 6 Jahren (Gray, 2006; Florit, Roch, Altoè & Levorato, 2009; Kapalková, Polišenská, & Vicenová, 2013; Majerus, Poncelet, Greffe, & van der Linden, 2006; Visu-Petra, Stanciu, Benga, Miclea & Cheie, 2014).

Zur Erklärung des Entwicklungswachstums im phonologischen Gedächtnis werden diverse Faktoren herangezogen. Neurologische Reifungsprozesse mit einhergehenden anatomischen, strukturellen und neurophysiologischen Veränderungen sowie zunehmenden neuronalen Verschaltungen, v.a. des Assoziationskortexes, können einen geringen Varianzanteil der kapazitären Leistungsanstiege im Kindergartenalter erklären (Brown & Jernigan, 2012; Cowan, 2010; Messer, Verhagen, Boom, Mayo, Leseman, 2015). Vor dem Hintergrund der sich bei Kindern im Alter von 2 bis 16 Jahren in vier Zyklen mit jeweils zwei Phasen vollziehenden Entwicklung fluider kognitiver Fähigkeiten gelten in der ersten Phase des Zyklus 2- bis 6-jähriger Kinder, bis zum Alter von 4 Jahren, eine ansteigende Informationsverarbeitungsgeschwindigkeit einhergehend mit einer Kapazitätsentlastung als Prädiktor für Leistungswachse im phonologischen Gedächtnis; in der zweiten Phase, ab 4 Jahren, ermöglichen Verknüpfungen von zuvor

konstruierten Repräsentationseinheiten sowie Arbeitsgedächtnisveränderungen mit Zuwächsen an exekutiven Funktionen, insbesondere der Inhibition irrelevanter Informationen sowie des Monitorings, ein Entwicklungswachstum im phonologischen Gedächtnis (Demetriou et al., 2014; Rose, Feldman, & Jankowski, 2009; Schmid et al., 2008). Das phonologische Gedächtnis entwickelt sich zudem analog zu Sprachgebrauch und -vertrautheit sowie als Ergebnis eines zunehmenden Umfangs an linguistischem, phonotaktischem und semantischem Wissen, welches besonders Kindern ab dem Alter von 4 bis 5 Jahren mit abgeschlossenen Phonemerwerb (Bowey, 2006) eine höhere Präzision bei der Repräsentation phonologischer Informationen und stabilere Gedächtnisspuren (*lexical restructuring* Hypothese; Metsala & Walley, 1998), Rekonstruktionsprozesse zur Wiederherstellung zerfallender phonologischer Gedächtnisspuren (*re-dintegration*) sowie kapazitätsentlastende automatische Gruppierungen und Rekodierungen von Sequenzen als bedeutungsvolle Einheiten (*chunking*) ermöglicht (Côte et al., 2014; Messer et al., 2015; Gobet et al., 2001; Grube, Lingen & Hasselhorn, 2008; Keren-Portnoy, Vihman, DePaolis, Whitaker, Williams, 2010; Storkel & Adlof, 2009). Der Einsatz von Strategien zur Steigerung der phonologischen Gedächtnisleistung unterliegt im Kindergartenalter einem Nutzungsdefizit (*utilization deficiency*) und ist auf den ab 4 Jahren erfolgenden gezielten, willentlichen Einsatz selektiver Aufmerksamkeit begrenzt (Bertrand & Camos, 2015; Camos & Barrouillet, 2011). Eine qualitative Entwicklung von Gedächtnisstrategien, deren kompetenter, regelmäßiger Einsatz und Effizienz in Kombination mit dem Bewusstsein eigener Gedächtnisfähigkeiten (metakognitivem Wissen) erfolgt ab dem Grundschulalter, ebenso wie der spontane Einsatz des, als Hauptquelle für bedeutsame Kapazitätsanstiege ab dem 7. Lebensjahr erachteten, subvokalen *rehearsals* mit einer effizienteren Nutzung der zeitlichen Begrenzung des phonologischen Speichers auf Basis von Steigerungen der Sprechgeschwindigkeit (Grube et al., 2008; Kiese-Himmel, Auberlen, & von Steinbüchel, 2012; Kron-Sperl, Schneider, Hasselhorn, 2008).

Biologische und umweltbedingte Einflussfaktoren.

Neben dem Lebensalter besteht die Annahme, dass genetische Anlagen phonologische Gedächtnisleistungen determinieren, gestützt auf empirisch festgestellten erblichen Komponenten bei Ziffernfolgen sowie insbesondere bei Nachsprechaufgaben mit Kunstwörtern (Kunstwörter-Nachsprechaufgaben, KWN-Aufgaben) (Bishop, Adams, & Norbury, 2006; Bishop, Holt, Line, McDonald, McDonald, & Watt, 2012; Barry, Yasin, & Bishop, 2007; Newbury et al., 2009; Peter et al., 2011; Truong et al., 2016). Hinsichtlich geschlechtlicher Einflüsse ist die Befundlage bei Kindergartenkindern heterogen und indiziert bei Ziffernfolgen sowie KWN-Aufgaben vergleichbare Leistungen von Mädchen und Buben (Kastner-Koller & Deimann, 2012; Kiese-

Himmel et al., 2012; Polišenská & Kapalková, 2014; Roy, Chiat, & Dodd, 2014) sowie kleine Geschlechtseffekte zugunsten von Mädchen (Kaushanskaya, Gross, & Buac, 2013; Lehmann, Quaiser-Pohl, & Jansen, 2014; Schöler & Schäfer 2004; Ulset et al., 2017; Lange, Euler, & Zaretsky, 2016) oder zugunsten von Buben, welche zu einem großen Anteil durch Geschlechtsunterschiede in der Informationsverarbeitungsgeschwindigkeit erklärbar sind (Schmid et al., 2008).

Trotz der als weitgehend unabhängig von demografischen sowie umweltbedingten Faktoren geltenden Leistungen im phonologischen Gedächtnis (Polišenská & Kapalková, 2014), zeigen diese Abhängigkeit vom Ausmaß der Spracheinwirkung, häuslichen Beschäftigungen wie Bilderbücher vorlesen (Duncan & Paradis, 2016; Silvén & Rubinov, 2010) sowie der Dauer des Kindergartenbesuchs (Niklas, Schmiedeler, Pröstler, & Schneider, 2011). Zudem können kurzzeitige Gedächtnis-Förderprogramme phonologische Gedächtnisleistungen mit kleinen bis mittelgroßen Interventionseffekten steigern (z.B. $g = 0.23-0.62$; Passolunghi & Costa, 2014; Rojas-Barahona, Förster, Moreno-Ríos, & McClelland, 2015). Empirisch stehen bei Ziffernfolgen sowie Kunstwörtern Befunde zu unabhängigen Leistungen vom sozioökonomischen Status (Balladares, Marshall, & Griffiths, 2016; Bonifacci, Barbieri, Tomassini, & Roch, 2017; Engel, Santos, & Gathercole, 2008; Kastner-Koller & Deimann, 2012; Law, McBean, & Rush, 2011) Befunde zu niedrigeren Leistungen bei Kindergartenkindern aus Familien mit niedrigerem sozioökonomischen Status gegenüber, die sich als Effekt des Wortschatzes erweisen (Chiat & Polišenská, 2016; Roy & Chiat, 2013, Seeff-Gabriel, Chiat, & Roy, 2008; Ulset et al., 2017). Ebenso als Effekt des Wortschatzes zu erachten sind auch die in Diskrepanz zu Befunden mit unabhängigen Leistungen hinsichtlich der Erstsprache (Bonifacci et al., 2017; Kiese-Himmel, et al., 2012) stehenden höheren Leistungen bei kumulativen Spracherfahrungen bzw. bi- oder multilingual aufwachsenden Kindergartenkindern (Summers, Bohman, Gillam, Pen, & Bedore, 2010; Thordardottir & Brandeker, 2013; Wild & Fleck, 2013) sowie Kindergartenkindern, deren Erst- statt Zweitsprache der Zielsprache des Testmaterials entspricht (Chiat & Polišenská, 2016; Messer, Leseman, Boom, & Mayo, 2010; Schöler & Schäfer, 2004).

Entwicklungsrelevante Funktionen des phonologischen Gedächtnisses.

Einhergehend mit der Annahme, dass die primäre Funktion des phonologischen Gedächtnisses in der Unterstützung des langfristigen Erwerbs phonologischer Strukturen einer Sprache liegt (Baddeley, Gathercole, & Papagno, 1988) und eine Grundvoraussetzung für den primären Sprach- sowie Fremdspracherwerb darstellt (Côte et al., 2014; Hu & Schuele 2005; Nora et al., 2012), erweist sich das phonologische Gedächtnis – v.a. durch den sequenziellen Reihenfolge-

mechanismus sowie auch durch den individuellen Beitrag des Mechanismus für Iteminformationen (Perez et al., 2012; Majerus et al., 2006) – insbesondere bei jungen Kindern als mittelstarker Prädiktor für das Erlernen neuer Wörter und weitere sprachliche Fähigkeitsaspekte, wie Sprachverständnis, Äußerungslänge und syntaktische Komplexität (Boyle, Lindell, & Kidd, 2013; Carpenter & Drabick, 2011; Chiat & Roy, 2007; Gathercole, Tiffany, Briscoe, Thorn, & The ALSPAC Team, 2005; Montgomery, Magimairaj, & Finney, 2010; Silvén & Rubinov, 2010; van Daal, Verhoeven, van Leeuwe, & van Balkom, 2008). Das phonologische Gedächtnis gilt zudem als reines Maß für das allgemeine Lernvermögen eines Kindes, entscheidend für das Entwicklungswachstum kognitiver Fähigkeiten (Alloway & Alloway, 2010; Preßler, Krajewski, & Hasselhorn, 2013) und als grundlegende Vorläuferfähigkeit sowie robuster, gegenüber dem sprachlichen Intelligenzquotienten überlegener, Prädiktor niedriger bis mittlerer Effektstärke für den Erwerb der Kulturtechniken des Lesens und Rechtschreibens im Schuleintritt (Alloway & Alloway, 2010; National Early Literacy Panel [NELP], 2008; Perez et al., 2012; von Goldammer et al., 2010) und rechnerischer Fähigkeiten (Alloway & Alloway, 2010; Passolunghi & Costa, 2014; Passolunghi, Lanfranchi, Altoè, & Sollazzo, 2015). Infolgedessen ist das phonologische Gedächtnis ein wesentlicher Prädiktor für das Schulleistungsniveau (Alloway, Gathercole, Kirkwood, & Elliott, 2009; Cowan, Ricker, Clark, Hinrichs, & Bret, 2015; Gathercole et al., 2004). Defizite im phonologischen Gedächtnis können aufgrund einhergehender Schwierigkeiten bei der Bewältigung von Aufgaben mit mehreren Schritten, beim Merken von Aufforderungen und Anweisungen sowie in der sozial-linguistischen Interaktion zudem mit einem Risiko für die Entwicklung von Verhaltensauffälligkeiten verbunden sein (Carpenter & Drabick, 2011; Gathercole, Durling, Evans, Jeffcock, & Stone, 2008).

Psychologische Diagnostik des phonologischen Gedächtnisses

Einhergehend mit den für die Kindesentwicklung zentralen Funktionen gilt eine defizitäre Kapazität des phonologischen Gedächtnisses in der psychologischen Diagnostik bei Kindergartenkindern als Hinweis für spezifische Sprachentwicklungsstörungen (SSES) bzw., nach der Internationalen statistischen Klassifikation der Krankheiten und verwandter Gesundheitsprobleme 10 (ICD-10), für *Umschriebene Entwicklungsstörungen des Sprechens und der Sprache* (F80.) sowie als Risiko für *Umschriebene Entwicklungsstörungen schulischer Fertigkeiten* (F81.) (Alloway & Alloway, 2010; Carroll, Mundy, & Cunningham, 2014; Jackson, Leita, & Claessen, 2016; Helland & Morken, 2015; Krishnan, Watkins, & Bishop, 2016; Montgomery et al., 2010; Reichenbach, Bastian, Rohrbach, Gross, & Sarrar, 2016; van Bergen, de Jong, Maassen, Krikhaar, Plakas, & van der Leij, 2014). Niedrige Leistungen im phonologischen Gedächtnis zeigen sich zudem bei weiteren neurologischen Entwicklungsstörungen, wie Down

Syndrom (Laws & Gunn, 2004; Næss, Lervåg, Lyster, & Hulme, 2015) und Autismus-Spektrum-Störungen (Harper-Hill, Copland & Arnott, 2013; Whitehouse, Barry, & Bishop, 2007; Wilson et al., 2013). Insbesondere für das Kindergartenalter, einer für die Ausbildung von schulischen Vorläuferfähigkeiten, Identifikation von Sprachentwicklungsrückständen und -störungen sowie Risikofaktoren für die weitere kognitive Entwicklung entscheidenden Periode (Renner, Mickley, & Friedrichshain, 2015; Roebers & Zoelch, 2005; White-Schwoch et al., 2015), ist daher eine frühzeitige Identifikation von Defiziten im phonologischen Gedächtnis mittels Messinstrumenten höchstmöglicher psychometrischer Qualität essentiell, um aufbauend auf diagnostischen Ergebnissen Fördermaßnahmen einleiten zu können (Blackwell, Cepeda, & Munakata, 2009; Drigas, Kokkalia, & Lytras, 2015; Pagani, Fitzpatrick, Archambault, et al., 2010; Renner et al., 2015; von Goldammer et al., 2010). In der Praxisanwendung stellt die Überprüfung des phonologischen Gedächtnisses bei Kindergartenkindern häufig einen Teil von Entwicklungstestbatterien zur Erfassung der Entwicklung in verschiedenen Funktionsbereichen dar (Jones & Macken, 2015; Renner et al., 2015; Roebers & Zoelch, 2005), wie Ziffernfolgen im Wiener Entwicklungstest (WET; Kastner-Koller & Deimann, 2012); ebenso werden sie häufig als verarbeitungsabhängiges Maß in Sprachentwicklungstests oder -screenings eingesetzt (Dispaldro, Leonard, & Deevy, 2013; Deevy, Weil, Leonard, & Goffman, 2010; Mathieu, Lindner, Lomako, Gagarina, Held, & Valentin-Klein, 2016; Moyle, Heilmann, & Finneran, 2014), wie Kunstwörter im Sprachentwicklungstest für drei- bis fünfjährige Kinder (SETK 3-5; Grimm, 2015; Subtest *Phonologisches Arbeitsgedächtnis für Nichtwörter*; PGN).

Herausforderungen für eine reliable und valide Diagnostik des phonologischen Gedächtnisses stellen bei Kindergartenkindern die Interferenz der in Testsituationen gezeigten Gedächtnisleistungen bzw. Zielfähigkeiten mit kindbezogenen, für die Erfassung und Bewältigung der Aufgaben erforderlichen, Fähigkeiten wie Aufgabenverständnis, Aufmerksamkeit, Konzentration, Ausdauer, Motivation und Interaktion mit dem/der TestleiterIn dar (Renner & Mickley, 2015; Schneider & Sodian, 1990). Eine weitere Herausforderung in der Erfassung des phonologischen Gedächtnisses sind die, insbesondere bei Kindern bis zum 4. Lebensjahr hohen, Verweigerungsraten (z.B. 1.8 bis 14.3%; Polišenská & Kapalková, 2014; Chiat & Roy, 2007; Roy & Chiat, 2004; Seeff-Gabriel et al., 2008), die u.a. auf Fähigkeitsdefizite (Hoff, Core, & Bridges, 2008) oder kurze Aufmerksamkeitsspannen, Schüchternheit, Ängstlichkeit und Müdigkeit zurückzuführen sind (Dohmen, Chiat, & Roy, 2013). Zur Schaffung optimaler Bedingungen für fähigkeitsentsprechende Nachsprechleistungen ist daher neben einer möglichst kurzen Testdauer (Polišenská & Kapalková, 2014) das Aufgabendesign sowie die Wahl des Aufgabenmaterials essenziell (de Abreu, Baldassi, Puglisi, & Befi-Lopes, 2013).

Kunstwörter versus Ziffernfolgen in der Diagnostik bei Kindergartenkindern.

Folgen aus bekannten, lexikalisch basierten einsilbigen Wörtern wie Ziffernfolgen, welche die älteste Methode zur Erfassung des phonologischen Gedächtnisses darstellen, sind mit Vorteilen einer Vorbeugung von Fremdheitseffekten und von einer möglichen Beanspruchung von Kapazitätsressourcen durch Bedeutungszuschreibungen verbunden, sowie mit Nachteilen einer lexikalisch-semanticen Verarbeitung und damit Konfundierung von Kurzzeit- mit Langzeitgedächtnis sowie einer unzureichenden Ausschöpfung phonologischer Verarbeitungsressourcen (Acheson & MacDonalds, 2009; Archibald & Griebeling, 2016; Jones & Macken, 2015; Mathieu et al., 2016). Dem gegenüber stellen Kunstwörter nicht-lexikalische Items dar, für deren Speicherung, Abruf und Nachsprechen aufgrund der Neuartigkeit des Materials keine lexikalischen Repräsentationen oder Wissen um bekannte Bewegungsmuster (motorische Programme) zur Verfügung stehen (Jones & Whiterstone, 2011; Kapalková et al., 2013; Vance, Stackhouse, & Wells, 2005). Infolgedessen gelten KWN-Aufgaben als sensitiveres Maß für den Speicherumfang des phonologischen Gedächtnisses unter Minimierung von Langzeitgedächtniseinflüssen sowie relativer Unabhängigkeit vom kulturellen und linguistischen Hintergrund (Adams & Gathercole, 1995; Gathercole, 2006; Law, Mahr, Schneeberg, & Edwards, 2016; Jackson et al., 2016). Kunstwörter beanspruchen zudem neurokognitive Kernprozesse für Input- und Outputverarbeitung sowie -speicherung in höherem Ausmaß als Ziffernfolgen und stellen ein förderdiagnostisches Prozess- statt Produktmaß dar (Casserly & Pisoni, 2013; Coady & Evans, 2008; Claessen et al., 2013; Jones, Gobet, Freudenthal, Watson, & Pine, 2014; Jones & Whiterstone, 2011; Moyle et al., 2014).

Durch Simulation der Basisaufgabe im Wortschatzerwerb junger Kinder mit der Tendenz, neue, durch einen Erwachsenen vorgesprochene Wörter nachzusprechen, gelten KWN-Aufgaben zudem als entwicklungsorientierte Anforderung für Kindergartenkinder, die Aufgabenverständnis, Motivation und Interesse sowie höhere Kooperationsbereitschaft im Vergleich zu Ziffernfolgen begünstigen kann (Archibald, 2008; Armon-Lotem & Meir, 2016; Chiat & Roy, 2007; Jones et al., 2014; Radeborg et al., 2006). Einhergehend mit der Kongruenz zwischen KWN-Aufgaben und der phonologischen Komponente beim Erlernen neuer Wörter als Voraussetzung für Einträge im lexikalischen Langzeitspeicher (Baddeley et al., 1998) gilt die Kunstwörter-Nachsprechleistung (KWN-Leistung) bei Kindergartenkindern als informativster Prädiktor mittlerer Effektstärke für Wortschatzerwerb und -erweiterung (Law et al., 2016; Gathercole, 2006; Leclercq & Majerus, 2010). Zudem erweisen sich KWN-Aufgaben – bei um ca. eineinhalb Standardabweichungen niedrigeren KWN-Leistungen von Kindergartenkindern mit SSES gegenüber gleichaltrigen normal entwickelten Kindern (z.B. $g = 1.42 - 1.65$; Bishop et al., 2012;

Dispaldro et al., 2013; Nash, Hulme, Gooch, & Snowling, 2013; Vugs, Hendriks, Cuperus, & Verhoeven 2014) – als reliable und größte diagnostische Schlüsselmarker zur Identifizierung von SSES (Bishop, North, & Donlan, 1996; Henry, Messer, & Nash, 2012; Montgomery et al., 2010; Petruccelli, Bavin, & Bretherton, 2012; Reichenbach et al., 2016). Die Sensitivität und Spezifität bei der Identifizierung von SSES mittels Kunstwörter ist – bei Interpretation von *Receiver-Operater-Characteristic*-Statistiken nach Swets (1988) – gut bis exzellent, Ziffernfolgen überlegen und steigt besonders bei Kunstwörtern ab drei Silben sowie mit Konsonantenclustern und niedrigerer Wortähnlichkeit (Armon-Lotem & Meir, 2016; Deevy et al., 2010; Dispaldro et al., 2013; Girbau & Schwartz, 2007; Kapalková et al., 2013; Thordardottir et al., 2011; Shahm Mahmood, Jalaie, Soleymani, Haresabadi, & Nemati, 2016). Überdies erweisen sich KWN-Leistungen im Vergleich zu Ziffernspannen in der Vorhersage von Rechtschreib- und Wortleseleistung sowie Lesegenauigkeit, -flüssigkeit und -verständnis als konsistentere und unabhängige Prädiktoren (Binamé & Poncelet, 2016; Carroll et al., 2014; Catts, Nielsen, Bridges, Liu, & Bontempo, 2014; Papadimitrioua & Vlachos, 2014; von Goldammer et al., 2010).

Mechanismen und Fähigkeitskomponenten beim Nachsprechen von Kunstwörtern.

Die zentrale Funktion des phonologischen Gedächtnisses bei KWN-Aufgaben, insbesondere des Mechanismus für serielle Reihenfolgeinformationen in Speicherung und Abruf der sequenziellen Anordnung von Phonemen und Silben (Coady & Evans, 2008; Gathercole, 2006; Majerus & Boukebza, 2013), stützt sich bei Kindergartenkindern auf Hauptkomponentenanalysen mit hohen Ladungen (z.B. .68 bis .89) auf einem, zusammen mit Ziffern- und/oder Wortfolgen extrahierten, Gedächtnisfaktor, gegenüber allgemeinen Sprachfaktoren, metaphonologischen Fähigkeiten oder zentral-exekutiven kognitiven Fähigkeiten (Alloway et al. 2004; Boets, Wouters, van Wieringen, & Ghesquière, 2006; Keilmann, Moein, & Schöler, 2012; Schmid et al., 2008). Die Konstruktvalidität stützt sich zudem auf Befunde zu mittleren bis hohen Zusammenhängen (mit und ohne Kontrolle des Lebensalters) zwischen KWN-Leistungen und Ziffernspannen (z.B. $r = .32$ bis $.61$; de Abreu et al., 2010; Papadimitrioua & Vlachos, 2014; Passolunghi et al., 2015; Schmid et al., 2008; Spencer, Kaschak, Jones, & Lonigan, 2015; von Goldammer et al., 2010; Vugs et al., 2015). Analog zu Annahmen einer Kapazitätsbegrenzung des phonologischen Gedächtnisses und eines zeitlichen Zerfalls von Gedächtnisspuren (Baddeley, 2003; Gathercole, Willis, Emslie, & Baddeley, 1992) zeigen sich beim Nachsprechen von Kunstwörtern Sequenz-Längeneffekte in Form von höheren Leistungen bei kurzen Kunstwörtern mit zwei bis drei Silben und sinkenden Leistungen mit steigender Silbenanzahl aufgrund des höheren Zeiterfordernisses bei Enkodierung und Artikulation während des Abrufs, sowie

phonologische Ähnlichkeitseffekte (Boerma et al., 2015; Bortolini et al., 2006; Dispaldro et al., 2013; Majerus & Boukebza, 2013; Polišenská & Kapalková, 2014; Sundström et al., 2014).

Im Vergleich zu einer vorwiegend segmentalen *bottom-up*-Verarbeitung neuartiger ebenso wie vertrauter phonologischer Informationen vor dem Alter von 4 Jahren (Ainsworth, Welbourne, & Hesketh, 2016; Vance et al., 2005), wirkt beim Nachsprechen von Kunstwörtern bei älteren Kindern modulatorisches *top-down* Feedback durch Wortschatz und sublexikalisches Wissen dem Zerfall oder der Degradierung temporärer phonologischer Repräsentationen entgegen (Grossberg & Kazerounian, 2016; Jones et al., 2014; Kapalková et al., 2013; Metsala & Chisholm, 2010; Moyle et al., 2014; Tanida, Ueno, Ralph, & Saito, 2015). Nach Sprachverarbeitungs- sowie Netzwerkmodellen (z.B. Jacquemot & Scott, 2006; Stackhouse & Wells, 1997) erfordert das Nachsprechen von Kunstwörtern, anders als das Nachsprechen von Ziffernfolgen, Prozesse der Inputverarbeitung, z.B. akustische Wahrnehmung, Analyse (Identifizierung, Diskrimination und Segmentierung) von Phonemen und Silbenstruktur sowie Phonemsynthese, welche der Speicherung und dem Abruf im phonologischen Gedächtnis vorausgehen, sowie diesen nachfolgende Prozesse der Outputgenerierung, z.B. motorische Programmierung, Erstellung eines motorischen Plans und Ausführung von Artikulationsbewegungen (Acheson & MacDonald, 2009; Coady & Evans, 2008; Jones & Whiterstone, 2011; Mathieu et al., 2016).

Zusätzlich zu individuumsspezifischen Fähigkeitsmerkmalen, wie intaktem Hörvermögen, phonologischer Bewusstheit, rezeptiven und expressiven Sprachfähigkeiten, oromotorischen Fähigkeiten und Artikulationsreife (Krishnan et al., 2017; Majerus et al., 2006; Tattersall, Nelson, & Tyler, 2014; Vance et al., 2005; Vugs et al., 2015), determinieren Itemeigenschaften KWN-Leistungen bei normal entwickelten Kindern, mit einer bei Kindergartenkindern alterskorreliert ansteigenden Tendenz und in höherem Ausmaß bei Kindern mit SSES (Coady & Evans, 2008; Coady, Mainela-Arnold, & Evans, 2013; Jones, Tamburelli, Watson, Gobet, & Pine, 2010; Goswami, Barnes, Mead, & Power, 2016). Höhere KWN-Leistungen zeigen sich bei Kunstwörtern mit subjektiv eingestufte Ähnlichkeit zu realen Wörtern in der nativen Sprache bzw. hoher Anzahl an lexikalischen Nachbarn (*Wortähnlichkeits-/Lexikalitätseffekt*; Jones et al., 2010; Jones et al., 2014; Metsala & Chisholm, 2010; Munson, Kurtz, & Windsor, 2005), in der nativen Sprache häufig auftretenden Phonemkombinationen (*Effekt der phonotaktischen Wahrscheinlichkeit*; MacRoy-Higgins & Dalton, 2015; Messer et al., 2010; Gray, Pittman, & Weinhold, 2014), einfachen Silben statt Silben mit Konsonantenclustern (*Effekt der phonologischen bzw. segmentalen Komplexität*; Smith, Goffman, Sasisekaran, & Weber-Fox, 2012), akzentuierten Silben und für die Zielsprache typischen Akzentuierungsmustern (*Effekt der prosodischen Struktur bzw. Akzent-Kongruenzeffekt*; Moyle et al., 2014; Sundström et al., 2014;

Tanida et al., 2015) sowie bei (im Akzentmuster) gebundener Darbietung statt silbenweiser Darbietung mit monotoner Intonation (Archibald & Gathercole, 2007). Demzufolge ist eine sorgfältige Konstruktion von Kunstwörtern grundlegend.

Entwicklung von Nachsprechaufgaben mit Kunstwörtern

Itemkonstruktion.

Zentrales Kriterium von Kunstwörtern ist der Ausschluss eines semantischen Wertes in der Zielsprache; Variationsmöglichkeit in der Itemkonstruktion besteht hinsichtlich der Silbenanzahl (Kunstwortlänge), Items pro Silbenanzahl, Sprachspezifität, Berücksichtigung des Erwerbsalters von Phonemen, Kontrolle von, mittels LaiInnenratings eingestufte, Wortähnlichkeit und Kontrolle von ordinalen Positionen von Phonemen, sowie hinsichtlich der Silbenstruktur (durch Konsonanten am Silbenende geschlossene Silben versus offene Silben mit Vokal am Silbenende; phonologisch einfache Silben aus Verbindungen von Konsonant [K] und Vokal [V], KVK und VK versus phonologisch komplexe Silben aus Verbindungen von KKV, KKVK und KVKK) (Archibald & Gathercole, 2007; Guiberson & Rodríguez, 2013; Jones, 2016; Polišenská & Kapalková, 2014). In der Festlegung von Kunstwortlängen ist zu berücksichtigen, dass eine Silbe im Durchschnitt ca. 2.8 Phoneme umfassen sollte (Gobet et al., 2007), die mittlere Sprechgeschwindigkeit bei Kindergartenkindern ca. zwei bis dreieinhalb Silben pro Sekunde beträgt (z.B. 2.43 bis 3.56 Silben pro Sekunde; Eaton & Rater, 2013; Sawyer, Chon, & Ambrose, 2008; Walker & Archibald, 2006) und bei 1-silbigen Kunstwörtern häufig Deckeneffekte sowie ab 5-silbigen Kunstwörtern bei jungen Kindern Bodeneffekte auftreten, jeweils einhergehend mit geringen differentiellen Beiträgen entsprechender Kunstwortlängen (Chiat, 2015; Graf Estes, Evans, & Else-Quest, 2007).

Unter Berücksichtigung von zielsprachlich zulässigen, artikulierbaren Phonemkombinationen (phonotaktischen Regeln) sowie weiteren, KWN-Leistungen determinierenden Besonderheiten einer Zielsprache können sprachspezifische Kunstwörter informativere Ergebnisse und höhere diagnostische Sensitivität in der Identifizierung von SSES erzielen (Graf-Estes, Evans, & Else-Quest, 2007; Jones et al., 2010). Konstruktionsmethoden sprachspezifischer Kunstwörter umfassen die Kombination von Konsonantencluster einschließenden Phonemsequenzen mit zielsprachlichen Affixen sowie den Austausch von Vokalen in existenten Wörtern, mit dem Nachteil einer höheren Wortähnlichkeit, oder den Austausch von Silbenanfängen (Onsets), Silbenkernen (Nuklei) oder Silbenenden (Kodas) diverser existenter Wörter (Gathercole, Willis, Baddeley, & Emslie, 1994; Keuleers & Brysbaert, 2010; Mathieu et al., 2016; Ulbrich, Alday, Knaus, Orzechowska, & Wiese, 2016). Mit dem Ziel einer kulturfreien Messung entstehen

sprachunspezifische Kunstwörter durch Kombination von Phonemsequenzen unter Ausschluss von zielsprachlichen Affixen und – einhergehend mit niedrigeren Itemschwierigkeiten – von Konsonantenclustern sowie häufig von spät erworbenen Phonemen (Dollaghan & Campbell, 1998), verbunden mit Vorteilen einer Minimierung von Wortähnlichkeit, Langzeitgedächtniseinflüssen und Anforderungen an die artikulationsmotorische Umsetzung beim Nachsprechen von Kunstwörtern und mit Nachteilen einer möglichen Verletzung der Phonotaktik und unzureichenden Abbildung phonologischer Verarbeitungs- und Speicheranforderungen an junge Kinder, deren Vokabular in vielen Sprachen Mehrfachkonsonanz, schwierige Phoneme und komplexe Silben mit Di- und Trigraphen umfassen soll (Baird, Dworzynski, Slonims, & Simonoff, 2010; Roy & Chiat, 2004; Ulrich, 2016).

Psycholinguistischer Kontext in der Konstruktion von Kunstwörtern.

Die phonetische Entwicklung betreffend können Kinder hinsichtlich Konsonanten, die Großteils nach Vokalen erworben werden, jene in der vorderen Mundhöhle vor jenen in der hinteren bilden (Kapalková et al., 2013; Weinrich & Zehner, 2011). Bei Betrachtung des altersspezifischen Erwerbs von Konsonanten und initialen Konsonantenclustern bei deutschsprachigen Kindern (Tabelle 1; vgl. Fox & Dodd, 2001) wird deutlich, dass 75% der Kinder vor dem Alter von 3 Jahren und 0 Monaten (3;0 Jahren) alle, bzw. 90% der Kinder fast alle, einfachen Konsonanten erworben haben, und Cluster aus zwei Konsonanten bis zum Ende des 4. Lebensjahres und Cluster aus drei oder mehr Konsonanten bis zum Alter von 4;5 Jahren erworben werden. Die phonologische Entwicklung erfolgt ab 18 bis 24 Lebensmonaten mit zunehmendem Wissenserwerb zur korrekten und regelkonformen Anwendung von Lauten und gilt im Alter von 4;0 bis 4;6 Jahren als weitgehend abgeschlossen; regelhafte Abweichungen (phonologische Prozesse) können bis zum Alter von 5;0 Jahren auftreten und sind zum Teil Effekte phonologischer Verarbeitungs- und Speicherprozesse (Vihman, DePaolis, & Keren-Portnoy, 2014; Weinrich & Zehner, 2011). Tabelle 2 gibt einen Überblick über altersspezifische, einzelne Phoneme, Phonemgruppen oder Silbenstrukturen betreffende, physiologische phonologische Prozesse von deutschsprachigen Kindern, unterteilt in Substitutions-, Silbenstruktur- und Assimilationsprozesse (Fox-Boyer, 2016). Zu den häufigsten phonologischen Prozessen zählen Ersetzungen, besonders hinsichtlich spät entwickelter und zugunsten häufiger Phoneme, und Silbenstrukturprozesse, v.a. Konsonantenclusterreduktionen; Assimilationen sowie Hinzufügungen gelten als seltener auftretende Fehler (Burke & Coady, 2015; James, Ferguson, & Butcher, 2015; Masso, McLeod, Baker, & McCormack, 2016).

Tabelle 1

Erwerb von Konsonanten und initialer Konsonantencluster bei deutschsprachigen Kindern

Alter	75% der Kinder erworben	90% der Kinder erworben
1;6-1;11	m, b, p, d, t, n	m, b, d,
2;0-2;5	v, h, z/s	b, n
2;6-2;11	f, l, j, η, x, ʁ, g, k, pf	v, f, l, t, η, x, h, k, z/s
3;0-3;5	ç, ts, bl, bʁ, fl, dʁ, tʁ, gl, kl	J, ʁ, g, pf, fʁ, kl
3;6-3;11	ʃ, gʁ, kʁ, kv, ʃm, ʃn, ʃʁ, ʃp, ʃv	ts, bl, bʁ, fl, gl, gʁ
4;0-4;5	kn, ʃl, ʃpʁ, ʃtʁ, ʃt,	ç, dʁ, tʁ, kʁ, kn, kv, l, ʃm, ʃn, ʃʁ, ʃp, ʃv, ʃt
4;6-4;11		ʃ, ʃpʁ, ʃtʁ

Anmerkung. Adaptiert nach nach Fox & Dodd (2001)

Tabelle 2

Physiologische phonologische Prozesse bei deutschsprachigen Kindern

Phonologischer Prozess	Alter für das Auftreten phonologischer Prozesse					
	2;0-2;5	2;6-2;11	3;0-3;5	3;6-3;11	4;0-4;5	4;6-4;11
Silbenstrukturprozesse						
Tilgung (silben-)finaler Konsonanten	■■■■■					
Tilgung unbetonter Silben	■■■■■	■■■■■				
Reduktion initialer Konsonantencluster	■■■■■	■■■■■	■■■■■	■■■■■		
Assimilationen						
	■■■■■	■■■■■				
Ersetzungsprozesse						
Plosifizierung	■■■■■					
Glottale Ersetzung von /ʁ/	■■■■■					
Deaffrizierung	■■■■■					
Vorverlagerung von /η/ zu [n]	■■■■■					
Rückverlagerung von /ʃ/ zu [ç]	■■■■■	■■■■■				
Vorverlagerung von /k/, /g/ zu [t], [d]	■■■■■	■■■■■	■■■■■			
Vorverlagerung von /ç/ zu [s]	■■■■■	■■■■■	■■■■■	■■■■■		
Vorverlagerung von /ʃ/ zu [s]	■■■■■	■■■■■	■■■■■	■■■■■	■■■■■	■■■■■

Anmerkung. Adaptiert nach Fox-Boyer (2016).

Hinsichtlich des Wortakzents in der deutschen Sprache gilt der Trochäus, mit Hauptakzent auf der vorfinalen Silbe, als häufigster bei zweisilbigen Wörtern (Janßen, 2003; Schmidt-Kassow, Rothermich, Schwartze, & Kotz, 2011). Nach der *Dreisilbenregel* (Vennemann, 1991; Wiese, 2000) erfolgt die Zuweisung des Hauptakzents vom rechten Wortrand ausgehend zum linken auf eine der drei finalen Silben in Abhängigkeit von quantitätssensitiven Informationen zu Silbenstruktur bzw. -schwere (je nach Definition geschlossene Silben, komplexe Silben oder lange Vokale) sowie von lexikalischem Wissen und Wissen um statistische Regelmäßigkeiten in der Akzentuierung (z.B. keine Betonung von Reduktionssilben und von den meisten deutschen Suffixen, Betonung fremdsprachlicher Suffixe) (Altmann & Kemmerling, 2005; Domahs, Plag, & Carroll, 2014; Domahs, Wiese, Bornkessel-Schlesewsky, & Schlesewsky, 2008; Domahs, Klein, Huber, & Domahs, 2013; Riedl, Wiese, Dellwo, & Wittig, 2014; Röttger et al., 2012; Vennemann, 2010). Die Zuweisung von Nebenakzenten erfolgt auf Basis der *Primary-Accent-First-Theory* (van der Hulst, 1997, zitiert nach van der Hulst, 2009) nach jener von Hauptakzenten, vom linken Wortrand aus zum rechten und am häufigsten auf starke initiale Silben, unter Berücksichtigung einer gleichmäßigen Rhythmusstruktur sowie u.a. der Regel, dass nicht mehr als zwei akzentuierte Silben aufeinander folgen dürfen (Knaus & Domahs, 2009; Knaus, Wiese, & Domahs, 2011).

Aufgabendesign.

Zur Sicherstellung eines hohen Standardisierungsgrades und einer konsistenten Darbietung von Kunstwörtern erfolgt diese, insbesondere bei Schulkindern, häufig über Audioaufzeichnungen (Archibald, 2008; Polišenská & Kapalková, 2014), in einigen Messinstrumenten auch mit verzerrten Darbietungen zur Erfassung der Verarbeitungsgenauigkeit (Schuchart et al., 2011). Bei Kindergartenkindern findet in der Regel eine unvermittelte Darbietung der Items statt, mit Nachteilen möglicher Inkonsistenzen und einer bei sichtbarem Sprechen erleichterten phonologischen Verarbeitung durch Hervorhebung von akustisch nicht salienten Differenzierungen (Havy, Foroud, Fais, & Werker, 2017), sowie bei Vorteilen einer höheren Flexibilität in der Anwendung und Förderung von Aufmerksamkeit sowie Kooperationsbereitschaft der Kinder als Ergebnis der Interaktion (Polišenská & Kapalková, 2014). Die Darbietungsform erfolgt in diesem Kontext häufig in Kombination mit interaktiven spielerischen Elementen, z.B. dem Rollen von Bällen oder Rufen von Puppen (Chiat & Roy, 2007; Grimm, 2015; Guiberson & Rodríguez, 2013; Stokes & Klee, 2009), die mit Interferenzen von Objekten mit der Gedächtnisleistung verbunden sein können (Brady & Goodman, 2014).

Die häufigsten Ansätze zur Bewertung von KWN-Leistungen umfassen die Ermittlung des Prozentsatzes an korrekten Phonemen (*percentage of phonemes correct*, PPC) und die Bewertung

des gesamten Kunstwortes mit Bildung eines Summenscores (*whole-item scoring*, Gesamt-Item-Bewertung), welche bei zeit- und aufwandsökonomischem Vorteil mit der PPC-Bewertung vergleichbare, ausreichend informative Ergebnisse liefert bzw. sich als effizienter, sensibler und klinisch praktikabler erweist (Boerma et al., 2015; Dispaldro et al., 2013; Graf Estes et al., 2007; Guiberson & Rodríguez, 2013; Polišenská & Kapalková, 2014; Roy & Chiat, 2004). Alternative Methoden zur Bewertung von KWN-Leistungen umfassen auf Phonemebene eine segmentale Bewertung der Anzahl korrekter Vokale, Konsonanten und Konsonantencluster, auf Silbenebene jene der Bewahrung der Silbenstruktur und auf Kunstwortebene jene der Anzahl korrekt reproduzierter Silben; kombinierte Methoden umfassen die Bewertung des gesamten Kunstwortes in Kombination mit der Ermittlung des Prozentsatzes an Silbenverlust und abweichenden Phonemen sowie zweiphasige Bewertungsprozesse mit einer segmentalen in Anschluss an eine Gesamt-Item-Bewertung (Boerma et al., 2015; Roy & Chiat, 2004; Tanida et al., 2015). Standardmäßig erfolgt eine konservative Bewertung der KWN-Leistungen, nach welcher Items als gelöst verrechnet werden, wenn Antworten alle Phoneme in korrekter Reihenfolge enthalten; einige Messinstrumente tolerieren abweichende Antworten, z.B. in Form von konsistenten Ersetzungen (Artikulationsfehlern), regionalen Dialekten, phonetischen Abweichungen, physiologischen phonologischen Prozessen (entwicklungstypischen Fehlern) und Hinzufügungen (Boerma et al., 2015; Chiat & Roy, 2007; Dollaghan & Campbell, 1998; Duncan & Paradis, 2016; Kohnert, Windsor, & Yim, 2006).

Messinstrumente mit Kunstwörtern.

Tabelle 3 gibt einen Überblick über Itemeigenschaften und Aufgabendesign ausgewählter, im deutschsprachigen und englischsprachigen Raum für den Altersbereich von Kindergartenkindern eingesetzter Messinstrumente zur Erfassung des phonologischen Gedächtnisses mit Kunstwörtern; der Anteil an Messinstrumenten mit sprachspezifischen Kunstwörtern, mit hoher Anzahl an Konsonantenclustern und geschlossenen Silben (Ulrich, 2016) sowie zum Teil mit zielsprachlichen Suffixen, ist mit dem Anteil an Messinstrumenten mit sprachunspezifischen Kunstwörtern (einfache Phonemkombinationen) – unabhängig von deren Aktualität – vergleichbar. Hinsichtlich der Itemeigenschaften zeigt Tabelle 3 eine weite Spannbreite der Itemanzahl von Messinstrumenten (13 bis 40 Kunstwörter), die Silbenanzahl liegt zwischen 1 und 8 Silben. Im SETK 3-5 (Grimm, 2015) sind für Kinder im Alter von 3;0 bis 3;11 Jahren eine geringere Item- und Silbenanzahl als für ältere Kindergartenkinder festgelegt. Einige Messinstrumente enthalten Probeitems, deren Wiederholung zur Gewährleistung von Aufgabenverständnis, anders als bei Testitems, zulässig ist.

Tabelle 3

Itemeigenschaften etablierter Messinstrumente im deutsch- und englischsprachigen Raum

Messinstrument	Items		Vorgabe		Bewertung	
(Alter)	Anzahl	Silben sprachsp.			Verrechnung	BM
Messinstrumente (aus Test- und -screeningverfahren ²) im deutschsprachigen Raum						
<i>PGN</i> aus	13	2-4				konservativ;
SETK 3-5 ¹ bzw.	(3;0-3;11)		ja	Akzent.,	Gesamt-I.,	Ungültigkeit bei
SSV ²	18	2-5		Probeitem	Summe	konsistenten
(3;0-5;11 Jahre)	(4;0-5;11)					Ersetzungsf.
<i>Gedächtnis Pho-</i>						Toleranz:
<i>nologisch</i> aus	15	2-8	nein	silbenweise,	Gesamt-I.,	konsistente
IDS-P ¹ (3;0-5;11)				monoton	Summe	Artikulationsf.
<i>Kunstwörter</i> aus				Akzent.,		Toleranz:
AGTB 5-12 ¹	24	2-4	ja	klare oder	Gesamt-I.,	Dehnungsf.,
(5;0-12;11)	(5;0-6;11)			verzerrte	Summe	dialektale
				Darbietung		Abweichungen
<i>Pseudowörter-</i>						
<i>Nachsprechen</i> aus						
BISC ² (letztes	11	4-6	ja	Akzent.	Gesamt-I.,	konservativ
Vorschuljahr)					Summe	
<i>Nachsprechen</i>						
<i>von Kunstwörtern</i>						Toleranz:
aus IDIS ¹ und	9	2-4	nein	Akzent.	Gesamt-I.,	diagnostizierte
HASE ²					Summe	Sprachf.
(4;6-6;11 Jahre)						
Mottier-Test						
(4-13 Jahre)	30	2-6	nein	silbenweise,	Gesamt-I.,	konservativ
				monoton	Summe	
Im englischsprachigen Raum/international eingesetzte Messinstrumente						
<i>PSRep Test</i>					Gesamt-I.,	Toleranz:
aus <i>Early repeti-</i>					Summe;	konsistente
<i>tion battery</i>	18	1-3	nein	Akzent.,	Anzahl an	Artikulationsf.,
(2;0-4;0)				Probeitems	Auslassungen	entwicklungs-
						typische F.,

						dialektale Abweichungen
<i>CNRep</i> (4-11 Jahre)	40	2-5	ja	Akzent., Probeitem	Gesamt-I., Summe	konservativ
<i>NRT</i> (4-11Jahre)	16	1-4	nein	Akzent.	Prozentsatz korrekter Sil- ben; PPC	Toleranz: Verzerrungen, Hinzufügungen
<i>COST Action IS0804</i> (kein festgelegter Al- tersbereich)						
Quasiuniverseller Test	16	2-5	nein	Akzent. (quasi-neut- rale oder sprachsp.)	Gesamt-I., Summe oder PPC	Toleranz: entwicklungs- typische F.
Sprachspezifi- sche Tests	24	2-4	ja	Akzent.		

Anmerkung. BM = Bewertungsmodus, Sprachsp. = sprachspezifisch, Akzent. = Akzentmuster (gebunden gesprochen), Gesamt-I. Summe = Geamtitem-Bewertung mit Bildung eines Summenscores, F. = Fehler, SETK 3-5 = Sprachentwicklungstest für drei- bis fünfjährige Kinder (Grimm, 2015), SSV = Sprachscreening für das Vorschulalter (Grimm, 2017), IDIS = Inventar diagnostischer Informationen bei Sprachentwicklungsauffälligkeiten (Schöler, 1999), HASE = Heidelberger Auditives Screening in der Einschulungsuntersuchung (Schöler & Brunner, 2008), BISC = Bielefelder Screening zur Früherkennung von Lese-Rechtschreibschwierigkeiten (Jansen, Mannhaupt, Marx, Skowronek, 2002), IDS-P = *Intelligence and Development Scales – Preschool* (Grob, Reimann, Gut, & Frischknecht, 2013), AGTB 5-12 = Arbeitsgedächtnistestbatterie für Kinder von 5–12 Jahren (Hasselhorn et al. 2012), Mottier-Test (Mottier, 1951), PSRep Test = The Preschool repetition test, *Early repetition battery* (Seeff-Gabriel, Chiat, & Roy, 2008), CNRep = *The children's test of nonword repetition* (Gathercole, Willis, Baddeley, & Emslie, 1994), NRT = *Nonword repetition Test* (Dollaghan & Campbell, 1998), *COST Action IS0804* (2013): Quasiuniverseller Test, sprachspezifische Tests = je nach Sprache unterschiedliche AutorInnen.

Die Darbietung erfolgt in den meisten Messinstrumenten mit Kunstwörtern in einem Akzentmuster gebunden, beim Mottier-Test (Mottier, 1951) und bei an diesem orientierten Messinstrumenten silbenweise im 0.5- oder 1.0-Sekunden-Takt. Die Bewertung erfolgt vorwiegend durch Verrechnung des gesamten Kunstwortes, in vielen Messinstrumenten unter Toleranz von Abweichungen.

In einem Großteil der wissenschaftlichen Untersuchungen im englischsprachigen sowie internationalen Raum findet eines von zwei etablierten, standardisierten Messinstrumenten Anwendung: *The children's test of nonword repetition* (CNRep; Gathercole et al., 1994) mit sprachspezifischen Kunstwörtern sowie Gesamt-Item-Bewertung und der *Nonword Repetition Test* (NRT; Dollaghan & Campbell, 1998) mit sprachunspezifischen Kunstwörtern und segmentaler Bewertung; international wurden innerhalb der *European Cooperation in Science and Technology* (COST) Action IS0804 (2013) quasi-universelle und analog dazu sprachspezifische KWN-Aufgaben, übergreifend für typologisch unterschiedliche Sprachen und für die Testung von bilingual aufwachsenden Kindern erstellt (Chiat, 2015). Im deutschsprachigen Raum ist der Mottier-Test (Mottier, 1951) das älteste Messinstrument. Da der Mottier-Test unzureichende Durchführungsobjektivität und verschiedene Normierungen aufweist, die in mindestens 10.0% der Fälle unterschiedliche Ergebnisse hinsichtlich auffälliger und unauffälliger Leistungen ergeben und für 16.7% der mit dem SETK 3-5 (Grimm, 2015) oder einem anderen standardisierten Messinstrument mit Kunstwörtern als auffällig eingestuften Kinder unauffällige Ergebnisse liefert (Ulrich, 2016), ist dieser anderen, standardisierten Messinstrumenten gegenüberzustellen.

Vorliegende Untersuchung

Die theoretische und empirische Befundlage hinsichtlich des phonologischen Gedächtnisses indiziert die Bedeutsamkeit reliabler, valider und entwicklungsorientierter Testaufgaben. KWN-Aufgaben entsprechen in besonderem Maß der Erfahrungswelt von Kindergartenkindern und ermöglichen im Vergleich zu Ziffernfolgen eine Erfassung der für Wortschatzerwerb und -erweiterung zentralen Sprachinputverarbeitungs- und Outputgenerierungsfähigkeiten sowie eine Ziffernfolgen übertreffende Sensitivität und Spezifität in der Identifizierung von SSES. In zahlreichen empirischen Studien erfolgte ein gemeinsamer Einsatz von Ziffernfolgen und Kunstwörtern, häufig bei Kindern ab 4 Jahren bzw. Vorschulkindern, es existieren jedoch wenige Untersuchungen mit Gegenüberstellungen dieser Maße bzw. deren Eignung für das Altersspektrum 3;0 bis 5;11-jähriger Kinder und hinsichtlich eines gemeinsamen Einsatzes in der Praxis. Zu nennen sind zwei Diplomarbeiten (Gruber, 1992; Varga, 1996), in welchen Leistungen bei Ziffernfolgen und neuartigem Material in Form von sinnfreien Silbenkombinationen

verglichen wurden und eine Replikation mit einer Anpassung des Testmaterials an aktuelle Messinstrumente mit Kunstwörtern erforderten. Da bei Kunstwörtern Leistungen von der phonetischen und phonologischen Entwicklung abhängig sind, ist in einigen Messinstrumenten mit Kunstwörtern für das gesamte Altersspektrum eine, im Vergleich zur konservativen Bewertung von Kunstwörtern, aufwändigere Bewertung von Antworten mit Toleranz von allgemeinen entwicklungstypischen Fehlern im Kindergartenalter vorgesehen, ohne diese auf das konkrete Erwerbsalter von Phonemen bzw. das Alter der Überwindung phonologischer Prozesse abzustimmen. Eine Forschungslücke ist diesbezüglich eine Untersuchung altersgruppenspezifischer Bewertungsregeln, ein Vergleich entsprechender Bewertung mit einer einfacher zu administrierenden und ökonomischen konservativen Bewertung sowie ein Vergleich zwischen Protokollierungen in Testsituationen und Transkriptionen mit dem Ziel, für die Anwendung in der Praxis eine möglichst effiziente und informative Bewertung festzulegen.

Ziel der vorliegenden Arbeit ist es, mit der Konstruktion von Kunstwörtern, die hinsichtlich Itemmerkmalen und Vorgabemodalität aktuellen Messinstrumenten entsprechen, sowie mit deren psychometrischen Analyse in einer Stichprobe aus Kindergartenkindern, neuartiges Testmaterial zu erstellen, das der Erfahrungswelt von Kindergartenkindern entspricht, Voraussetzungen für Motivation, Aufmerksamkeit und Kooperationsbereitschaft bietet und eine eindimensionale, psychometrischen Gütekriterien entsprechende sowie förderdiagnostische Erfassung des phonologischen Gedächtnisses in dieser Altersgruppe ermöglicht, als entwicklungsdiagnostisches Instrument entsprechende altersbedingte Kapazitätswachstums abbildet, und sich, mit Potenzial zur Unterscheidung zwischen Kindern mit normaler Entwicklung, Entwicklungsrückständen und -vorsprüngen, für den Wiener Entwicklungstest (WET; Kastner-Koller & Deimann, 2012) als Ergänzung zu einem Subtest mit Ziffernfolgen eignet. Der Vergleich der psychometrischen Qualität bei Anwendung des standardmäßig eingesetzten konservativen Bewertungsmodus mit jener bei Anwendung entwicklungstypischer oder anderer Abweichungen berücksichtigenden Bewertungsmodi soll der Festlegung auf einen informativen sowie ökonomischen und objektiven Bewertungsmodus dienen, ebenso zielt der Vergleich von Ergebnissen bei protokollierten und transkribierten Antworten auf Ökonomie. Zur Beantwortung der vorrangigen Fragestellung, „Inwieweit eignet sich neuartiges Material im Vergleich zu vertrautem zur Erfassung des phonologischen Gedächtnisses für ein im Kindergartenalter einzusetzendes allgemeines Entwicklungstestverfahren?“, erfolgt die Untersuchung folgender Fragestellungen: 1. „Welche Kunstwörter können nach psychometrischen Kriterien im Itempool belassen werden?“, 2. „Inwieweit entspricht ein selektiertes Itemset entwicklungspsychologischen Grundlagen und psychometrischen Gütekriterien?“, 3. „Ermöglicht die Anwendung des konservativen

Bewertungsmodus eine mit toleranteren Bewertungsmodi vergleichbare oder höhere psychometrische Qualität und in welchem Ausmaß stimmen Bewertungen von Transkriptionen mit jenen von Protokollierungen in Testsituationen überein?“ Als Nebenbedingung wurden qualitative Analysen von Beobachtungen des Verhaltens der Kinder während der Testsituationen und von, an Testsituationen anschließende, Verhaltensbeurteilungen vorgenommen.

Methode

Versuchsplan / Untersuchungsdesign

Der vorliegenden querschnittlich angelegten Hauptuntersuchung diene eine Voruntersuchung der Überprüfung des konstruierten Itempools auf Funktionalität, Akzeptanz und Eignung für die altersbezogene Zielpopulation mit einer darauf aufbauenden Adaptierung und – zur Begünstigung von konstanter Aufmerksamkeit und Motivation – Aufteilung in zwei aus dem Itempool zusammengestellten Itemsets unter Parallelisierung ausgewählter Itemmerkmale. Ausgenommen der Voruntersuchung erfolgten Untersuchungsplanung und Datenerhebung zur vorliegenden Arbeit in Kooperation mit Masterarbeitsuntersuchungen von Felkel (in Vorbereitung) und Stacher (in Vorbereitung) zur fein- bzw. grobmotorischen Entwicklung, unter Einsatz einheitlicher Messinstrumente und Erstellung eines gemeinsamen Datenpools. In der Hauptuntersuchung erfolgte die Erfassung des phonologischen Gedächtnisses mittels zweier Aufgabentypen – zur Vorbeugung von Ermüdungs- und Vermeidung bzw. Konstanthaltung von Lerneffekten – an zwei unterschiedlichen Testterminen mit einer für den ersten Termin randomisierten Zuteilung entsprechender Messinstrumente, sodass ein Teil der Kinder zum ersten Testtermin Kunstwörter (Protokollbogen A) und der andere Teil Ziffernfolgen (Protokollbogen B) erhielt und beim zweiten Testtermin das jeweils andere Messinstrument. Eine Randomisierung erfolgte auch in der Zuteilung eines von zwei Itemsets an Kunstwörtern (A1 und A2). Der für 3-jährige Kinder vorgesehene Verzicht auf die Vorgabe von Kunstwörtern der höchsten Silbenanzahl basierte auf der empirischen Befundlage sowie Voruntersuchungsergebnissen und ergab für entsprechende Altersgruppe aus dem Design resultierende fehlende Werte (*missings by design*). Die Beantwortung der Fragestellungen erfolgte im Rahmen eines korrelativen Designs sowie mit den abhängigen Variablen der KWN-Leistung und Ziffernspanne im Rahmen eines *between-subjects*-Designs mit den unabhängigen Variablen Altersgruppe in Ganzjahresschritten, Geschlecht, Messinstrument beim ersten Testtermin, Version des Itemsets und Testleiterin sowie eines *within-subjects*-Designs mit dem Innersubjektfaktor Mittelwert pro Silbenanzahl.

Stichprobenbeschreibung

Die Voruntersuchung erfolgte bei einer Gelegenheitsstichprobe von sechs, über den Verwandten- und Bekanntenkreis rekrutierten, Kindergartenkindern (1 Mädchen, 5 Buben) im Alter von 40 bis 69 Monaten ($M = 56 \pm 11$ Monate), basierend auf Aufklärungen über Untersuchungsziele und -ablauf, Freiwilligkeit der Teilnahme und mündliches Einverständnis der Eltern.

Die für die Hauptuntersuchung durchgeführte Stichprobenziehung, mit einer nach Alter in Ganz- sowie Halbjahresschritten und Geschlecht geschichteten Zufallsauswahl von 3;0 bis 5;11-jährigen Kindern mit ausreichenden Deutschkenntnissen, basierte auf einer TeilnehmerInnenrekrutierung über pädagogische Leiterinnen sowie Kindergartenleiterinnen und -pädagogInnen disponibler Kindergartenstandorte der privaten Trägerorganisationen St. Nikolausstiftung und Kinder in Wien (KIWI) unter Aushändigung von Informationsblättern zu Zielen und Ablauf der Untersuchung. Der daraus resultierende Datensatz wurde aufgrund eines unterproportional niedrigeren Anteils an Mädchen durch Daten einer Gelegenheitsstichprobe an weiblichen, über den Verwandten- und Bekanntenkreis rekrutierten Kindergartenkindern ergänzt, die im Rahmen von Probetestdurchführungen unter Einschluss aller für die Hauptuntersuchung erforderlichen Messinstrumente erhoben wurden. Jegliche Teilnahme basierte auf von Eltern bzw. Erziehungsberechtigten unterzeichneten Einverständniserklärungen in Kombination mit Informationsblättern, sowie auf Freiwilligkeit und unter Zusicherung von Anonymität in Dokumentation und Datenverarbeitung.

Die 66 Kinder aus Wiener Kindergärten umfassende Ausgangsstichprobe reduzierte sich infolge von Datenausfällen durch fehlende Kooperationsbereitschaft dreier Kinder bei den KWN-Aufgaben sowie infolge des Ausschlusses der fünf Kinder umfassenden Altersgruppe 5;6-5;11-Jähriger zur Vermeidung einer Ergebnisverzerrung aufgrund niedriger, für Kinder der höchsten Altersgruppe nicht repräsentativer Leistungen in Aufgaben zum phonologischen Gedächtnis sowie in konstruktfernen Aufgaben. Die durch entsprechenden Datenausschluss resultierende Analysestichprobe umfasste 58 Kinder (24 Mädchen, 34 Buben) im Alter von 37 bis 65 Monaten (3;1-5;5 Jahren) mit einem durchschnittlichen Alter von 52 Monaten (4;4 Jahren $\pm$ 8 Monate), zu 41.4% aus Kindergärten der St. Nikolausstiftung, zu 50.0% aus KIWI-Kindergärten und zu 1.7% aus dem Bekanntenkreis. Die Teilstichprobe, die Itemset 1 erhielt (Teilstichprobe 1), umfasste 31 Kinder (16 Mädchen, 15 Buben) im Alter von 38 bis 65 Monaten mit einem durchschnittlichen Alter von 54 Monaten (4;5 Jahren $\pm$ 7 Monate), die zu 38.7% aus Kindergärten der St. Nikolausstiftung, zu 54.8% aus KIWI-Kindergärten und zu 6.5% aus dem Bekanntenkreis stammten; die Teilstichprobe mit Itemset 2 (Teilstichprobe 2) umfasste 27 Kinder (9 Mädchen, 18 Buben) zwischen 37 und 65 Monaten mit einem durchschnittlichen Alter von

50 Monaten (4;2 Jahren $\pm$ 9 Monate), zu 44.4% aus Kindergärten der St. Nikolausstiftung, zu 44.4% aus KIWI-Kindergärten und zu 11.1% aus dem Bekanntenkreis. Die absoluten und relativen Häufigkeiten der Kinder in Bezug auf Altershalbjahresgruppen, Standorte der Kindergärten und Testleiterinnen sind Anhang A zu entnehmen.

Eindimensionale χ^2 -Tests ergaben für Teilstichprobe 1 Gleichverteilung hinsichtlich der Altersstufen in Halbjahresschritten, $\chi^2(4) = 6.90$, $p = .141$, des Geschlechts, $\chi^2(1) = 0.03$, $p = .857$, der Testleiterinnen, $\chi^2(2) = 1.23$, $p = .542$, sowie des Gedächtnisaufgabentyps beim ersten Testtermin, bei einer knapp signifikanten Ungleichverteilung des Alters in Ganzjahresschritten, $\chi^2(1) = 6.45$, $p = .040$; für Teilstichprobe 2 zeigte sich Gleichverteilung des Alters in Halbjahresschritten, $\chi^2(4) = 1.70$, $p = .790$, in Ganzjahresschritten, $\chi^2(2) = 2.00$, $p = .368$, des Geschlechts, $\chi^2(1) = 3.00$, $p = .083$, der Testleiterinnen, $\chi^2(2) = 0.22$, $p = .895$, sowie des Gedächtnisaufgabentyps beim ersten Testtermin, $\chi^2(1) = 0.93$, $p = .336$. Tabelle 4 zeigt die Verteilung des Geschlechts und des Aufgabentyps beim ersten Testtermin auf das Alter in Ganzjahresschritten. Zweidimensionale χ^2 -Tests zeigten stochastische Unabhängigkeit hinsichtlich der Altersgruppen in Ganzjahresschritten und des Geschlechts der Kinder für Itemset 1, $\chi^2(2) = 1.93$, $p = .381$, und Itemset 2, $\chi^2(2) = 0.00$, $p > .999$, sowie hinsichtlich der Altersgruppen in Ganzjahresschritten und des Gedächtnisaufgabentyps beim ersten Testtermin für Itemset 1, $\chi^2(2) = 1.39$, $p = .499$, und Itemset 2, $\chi^2(2) = 4.22$, $p = .121$.

Hinsichtlich der sozialen Schichtzugehörigkeit, ermittelt über die Berufstätigkeit der Väter und bei Alleinerzieherinnen der Mütter, zeigte sich, bei sonst vergleichbarer Verteilung mit der Grundgesamtheit, eine Unterrepräsentation der niedrigen Sozialschicht (Tabelle 5). Die Mehrheit der Kinder (mit Angaben zur Lebenssituation) lebte mit beiden Elternteilen im gemeinsamen Haushalt (92.7% in der Gesamtstichprobe, 93.1% in Teilstichprobe 1, 92.0% in Teilstichprobe 2), 7.3% mit alleinerziehenden Müttern (6.9% in Teilstichprobe 1, 8.0% in Teilstichprobe 2). Hinsichtlich Geschwister lebte 29.1% der Kinder zusammen mit mindestens einem älteren Geschwisterteil (31.0% in Teilstichprobe 1, 24.0% in Teilstichprobe 2), 34.5% der Kinder mit mindestens einem jüngeren Geschwisterteil (36.6% in Teilstichprobe 1, 32.0% in Teilstichprobe 2) und 38.2% der Kinder lebte ohne Geschwister im Haushalt (33.3% in Teilstichprobe 1, 44.0% in Teilstichprobe 2). Die durchschnittliche Dauer des Kindergartenbesuchs betrug in Teilstichprobe 1 ca. 29 Monate ($\pm$ 8 Monate) und in Teilstichprobe 2 ca. 27 Monate ($\pm$ 12 Monate). Ein Großteil der Kinder besuchte den Kindergärten ganztags (74.1% in der Gesamtstichprobe, 87.1% in Teilstichprobe 1, 59.3% der Kinder in Teilstichprobe 2) im Vergleich zu halbtags (19.0% in der Gesamtstichprobe, 6.5% in Teilstichprobe 1, 33.3% der Kinder in Teilstichprobe 2).

Tabelle 4

Kreuztabelle mit Alter und Geschlecht sowie mit Aufgaben beim ersten Testtermin

	Alter (Ganzjahre)								
	Itemset 1			Itemset 2			Itemset 1 und 2		
	3	4	5	3	4	5	3	4	5
Geschlecht									
Mädchen	3	7	5	4	3	2	7	10	7
Buben	4	10	2	8	6	4	12	16	6
1. Testtermin									
Kunstwörter	4	6	2	8	3	5	12	9	7
Ziffernfolgen	3	11	5	4	6	1	7	17	6

Tabelle 5

Soziale Schichtzugehörigkeit – Berufstätigkeit des Vaters/der Mutter in der Stichprobe und in der Grundgesamtheit

Beruf des Vaters / der Mutter	Vorliegende Untersuchung			Grundgesamtheit
	n (%)			Österreich ^e
	Gesamt	Teilstichprobe 1	Teilstichprobe 2	%
Selbstständigkeit, freier Beruf	9 (15.5)	6 (19.4)	3 (11.1)	13.7
Hohe Qualifikation ^a	15 (25.9)	7 (22.6)	8 (29.6)	25.4
Mittlere Qualifikation ^b	23 (39.7)	11 (35.5)	12 (44.4)	38.8
Niedrige Qualifikation ^c	5 (8.6)	3 (9.7)	2 (7.4)	20.3
Lehrling	1 (1.7)	1 (3.2)	0 (0.0)	2.7
Keine Angabe	5 (8.6)	3 (9.7)	2 (7.4)	

Anmerkung. Gesamt = gesamte Analysestichprobe, Teilstichprobe 1 = Vorgabe von Itemset 1, Teilstichprobe 2 = Vorgabe von Itemset 2.

^aAngestellte/r/Beamter, führende Tätigkeiten, Meister, Vorarbeiter, ^bAngestellte/r/Beamter, Facharbeiter, ^cAngestellte/r/Beamter Hilfstätigkeit, angelernte/r ArbeiterIn, ^eerrechnet aus der Beschäftigtenstatistik 2016 der Statistik Austria (Arbeitsmarktstatistiken 2016: Detailergebnisse C Erwerbspersonen und D Erwerbstätigkeit, Abgerufen unter https://www.statistik.at/web_de/statistiken/menschen_und_gesellschaft/arbeitsmarkt/erwerbsstatus/index.html).

Messinstrumente

Die festgelegte Durchführungsreihenfolge der Messinstrumente diene der Erfassung folgender Konstrukte: 1. Feinmotorik, 2. phonologisches Gedächtnis bei Kunstwörtern und 3. Grobmotorik in Protokollbogen A, sowie 1. Visuomotorik, 2. phonologisches Gedächtnis bei Ziffernfolgen und 3. Grobmotorik in Protokollbogen B, mit den Zielen der Herstellung einer abwechslungsreichen Situation durch Aufgaben höheren Aufforderungscharakters zu Beginn und am Ende der Testung sowie der Begünstigung von Aufmerksamkeit und Konzentration bei Aufgaben zum phonologischen Gedächtnis. Zur Erfassung des phonologischen Gedächtnisses bei vertrautem Material sowie konstruktferner Fähigkeiten erfolgte die Durchführung von vier aus insgesamt 14 Subtests des WET (Kastner-Koller & Deimann, 2012), welches als allgemeines Entwicklungstestverfahren für Kinder im Alter von 3 bis 6 Jahren der Erfassung des Entwicklungsstandes in allen relevanten Funktionsbereichen dient. Basierend auf einer repräsentativen Stichprobe aus deutschen und österreichischen Kindern ($N = 1245$) stellt der WET Normen für Altershalbjahresschritte in C-Werten für alle Subtests sowie in C-Werten und Standardwerten für einen Gesamtentwicklungsscore zur Verfügung und erfüllt psychometrische Anforderungen hinsichtlich Objektivität, Reliabilität sowie (Inhalts-, Konstrukt- und Kriteriums-)Validität. Hinsichtlich der Interpretation von Normwerten im WET (Kastner-Koller & Deimann, 2012) indizieren C-Werte von 4 bis 6 eine normale Entwicklung, C-Werte von 0 und 1 einen massiven Entwicklungsrückstand, C-Werte von 2 und 3 Förderbedarf, C-Werte von 7 und 8 eine gute Entwicklung und C-Werte von 9 und 10 einen Entwicklungsfortschritt. Die Erhebung soziodemografischer Daten erfolgte im Rahmen eines Elternfragebogens. Protokollierte Verhaltensbeobachtungen sowie anschließend an Testungen durchgeführte Verhaltensbeurteilungen im Rahmen eines für die Untersuchung konstruierten Bogens dienten der Erhebung qualitativer Aspekte zu den KWN-Aufgaben.

Phonologisches Gedächtnis bei neuartigem Material.

Der im Rahmen vorliegender Masterarbeit entwickelte Itempool umfasste, nach mehrfachen Adaptierungen, 40 sprachspezifische Kunstwörter (z.B. Laponiteful) mit 2 bis 6 Silben deren Parallelisierung hinsichtlich initialer Konsonanten, Silbenstruktur (offene/geschlossene Silben) und segmentaler Komplexität (mit/ohne Konsonantencluster), Akzentuierung sowie Anzahl und Arten an möglichen entwicklungstypischen Fehlern in zwei Itemsets mit jeweils 20 (vier 2-silbigen, fünf 3-silbigen, vier 4-silbigen, vier 5-silbigen und drei 6-silbigen) Kunstwörtern resultierte. Als Indizien für die Akzeptanz und Tauglichkeit der Items für die Zielpopulation galten die von Kindern der Voruntersuchung gezeigte Kooperationsbereitschaft, Eifer, positive und spielerische Äußerungen zu den Kunstwörtern, Lachen und das spontane Erfinden neuer

Kunstwörter. Die auf Voruntersuchungsergebnissen aufbauenden Adaptierungen von Erhebungsmaterial und Aufgabendesign umfassten Vereinfachungen der Hälfte an 5- und 6-silbigen Kunstwörtern durch Reduktion von Konsonantenclustern, die auf 2- bis 5-silbige Items begrenzte Vorgabe für 3-jährige Kinder aufgrund niedriger Lösungshäufigkeiten, die Vorgabe eines Probeitems und die Verdeutlichung der Neuartigkeit der „Wörter“ in der Instruktion wegen Instruktionsverständnisproblemen sowie den bedarfsabhängigen statt ursprünglich vorgesehenen standardmäßigen Einsatz eines Teddybären aufgrund von Interferenzen mit der Gedächtnisleistung oder Ablehnung bei älteren Kindergartenkindern.

Konstruktion.

Die Entwicklung des Itempools für die vorliegende Untersuchung erfolgte in Anlehnung an Konstruktionsprinzipien des Subtests *PGN* aus dem SETK 3-5 (Grimm, 2015). Dieser ist ein im deutschsprachigen Raum hinsichtlich Testgütekriterien am besten überprüfte und in der Praxis am häufigsten eingesetzte sowie durch die Bund-Länder Initiative zur Sprachförderung, Sprachdiagnostik und Leseförderung empfohlene Sprachentwicklungstest (Bildung durch Sprache und Schrift [BiSS], 2016; Melzer, Rißling, Petermann, 2015; Sachse & Suchodoletz, 2013). Die Konstruktionsprinzipien des Itempools waren ebenso analog zum CNRep (Gathercole et al., 1994) und sprachspezifischen Tests der COST Action IS0804 (2013). Die Konstruktion der Kunstwörter erfolgte, bei Ausschluss von lexikalischen Morphemen, durch Kombination von Onsets, Nuklei und Koda unterschiedlicher Wörter aus dem deutschen, kindorientierten Sprachinput mit deutschsprachlich zugelassenen Phonemsequenzen. Bei 3- bis 6-silbigen Kunstwörtern erfolgten entsprechende Kombinationen zu einfachen und komplexen Silben mit Konsonantenclustern an initialen, medialen und/oder finalen Positionen und maximal einem deutschsprachlichen Suffix am Kunstwortrand; die für die jüngsten Altersgruppen relevantesten 2-silbigen Kunstwörter basierten auf Phonemkombinationen unter Ausschluss von spät erworbenen Phonemen und Suffixen. Der Rückgriff auf die zum Kernkorpus des Digitalen Wörterbuchs der deutschen Sprache (DWDS) zählenden Datenbank dlexDB (Heister et al., 2011) diente der Einhaltung phonotaktischer Regeln des Deutschen und der Kontrolle phonotaktischer Häufigkeit. Zur Vorbeugung von Übungseffekten kam innerhalb des gesamten Itempools jede Silbe mit mehr als zwei Phonemen nur einmal vor, ebenso wie Konsonanten innerhalb einer Sequenz. Die ordinalen Positionen von Vokalen in In- und Auslautpositionen wurden bei Kunstwörtern derselben Silbenanzahl systematisch variiert. Die Festlegung einer Akzentuierung für die Kunstwörter erfolgte unter Berücksichtigung des Wortakzents im Deutschen.

Durchführung.

Die Durchführung erforderte Protokollbogen und Stift zur lautgetreuen Protokollierung von Antworten der Kinder in der Testsituation und einen Tonträger, um Antworten der Kinder für eine nachträgliche lautgetreue Transkription aufzuzeichnen. Nach Erklärung des Tonträgers erfolgte mit der Aufforderung, „lustige Wörter“ nachzusprechen die Vorgabe eines Probeitems zur Sicherung des Aufgabenverständnisses und anschließend die Vorgabe der Testitems. Kunstwörter wurden mit freundlichem Blickkontakt und leicht gesenktem Kopf in der Reihenfolge am Protokollbogen (Item 1 bis Item 17 bei 3-Jährigen bzw. bis Item 20 bei 4- bis 5-Jährigen) und in der gekennzeichneten Akzentuierung gebunden vorgesprochen, Hauptakzente durch erhöhte Sprechlautstärke und lang ausgesprochenem Vokal, Nebenakzente durch minimal erhöhte Sprechlautstärke. Wiederholungen des Probeitems bei ausbleibender Reaktion und der Instruktion bei mangelndem Verständnis waren zulässig, bei Testitems Wiederholungen und Korrekturen unzulässig. Bei niedriger Kooperationsbereitschaft der Kinder bestand die Möglichkeit zur Einbeziehung eines Teddybären aus dem vorangegangenen Subtest. Protokollbögen mit den Itemsets inklusive wörtlicher Instruktionen, Angaben zu Rahmenbedingungen und Anweisungen zum Umgang mit möglichen herausfordernden Situationen bzw. Testablauf oder Leistung beeinträchtigenden Verhaltensweisen (z.B. aufgabenirrelevante Fragen und Äußerungen) sowie zur Förderung von Aufgabenorientierung und Arbeitshaltung sind in Anhang C.

Die Bewertung beschränkte sich auf die Testitems. Die dichotome Verrechnung erfolgte auf Ebene der Kunstwörter (Gesamt-Item-Bewertung), bei den protokollierten Antworten unter Anwendung des konservativen Bewertungsmodus (Standardbewertung, St) (0 = fehlerhafte/keine Antwort; 1 = vollständig korrekte Antwort) und bei den transkribierten Antworten zusätzlich zur Anwendung des Bewertungsmodus St unter Anwendung von einander ausschließenden Bewertungsmodi mit der Toleranz entwicklungstypischer phonologischer Fehler (E) bei den Altersgruppen 3;0-4;11, mit jeweils spezifischen, für die Altersgruppen 3;0-3;5, 3;0-3;11 oder 3;0-4;11 relevanten Abweichungen (E1 – E6), der Toleranz von Hinzufügungen (H) und von kombinierten Bewertungsmodi mit der Toleranz von entwicklungstypischen Fehlern und konsistenten Artikulationsfehlern (EK) sowie der Toleranz von entwicklungstypischen Fehlern, konsistenten Artikulationsfehlern und Hinzufügungen (EKH). Beschreibungen der Bewertungsmodi, Auswertungsregeln und -beispiele sind im Anhang C. Die Veränderung der Reduktionssilben /-er/ zu [-a] und /-el/ zu [-l] (z.B. Blesamda statt Blesamder, Listorkl statt Listorkel) wurde in allen Bewertungsmodi toleriert. Die Anzahl der korrekt nachgesprochenen Kunstwörter entsprach dem erreichten Summenscore (maximal 17 bei 3-Jährigen, maximal 20

bei 4- bis 5-Jährigen), die Silbenanzahl des letzten, korrekt reproduzierten Kunstwortes entsprach der Silbenspanne.

Phonologisches Gedächtnis bei vertrautem Material.

Der, dem Funktionsbereich Lernen und Gedächtnis zugeordnete, Subtest *Zahlen Merken* aus dem WET (Kastner-Koller & Deimann, 2012) diente der Erfassung des phonologischen Gedächtnisses bei vertrautem Material. Der Subtest umfasst 10 Testitems mit Ziffernfolgen der Länge von zwei bis sechs Ziffern (z.B. 3-8-6). Die Ziffernfolgen wurden, beginnend mit zwei Ziffern, adaptiv vorgegeben, in Ein-Sekunden-Abständen mit monotoner Intonation vorgesprochen, und sollten vom Kind unmittelbar in der Darbietungsreihenfolge nachgesprochen werden. Beim korrekten, seriellen Nachsprechen der Ziffernfolge einer Länge (zwei Versuche für jede Länge) wurde mit der um eine Ziffer längeren Ziffernfolge fortgeführt; als Abbruchkriterium galt das fehlerhafte Nachsprechen zweier Ziffernfolgen einer Länge. Bei dichotomer Verrechnung (0 = falsche/keine Antwort, 1 = korrekte Antwort) der Items erfolgte die Ermittlung des Umfangs der längsten korrekt reproduzierten Ziffernfolge (Ziffernspace) und der Anzahl der insgesamt vorgegebenen Ziffernfolgen sowie eine Ermittlung von Normwerten (C-Werte). Aufgrund der Vorgabemodalität des Subtests *Zahlen Merken* waren Reliabilitätsanalysen auf Schätzungen der Retest-Reliabilität (r_{tt}) begrenzt; diese lagen in der Normierungsstichprobe des WET bei .67. Die konvergente Validität des Subtests *Zahlen Merken* stützt sich auf hohe Korrelationen mit konstruktnahen Messinstrumenten, wie dem Subtest *Zahlennachsprechen* ($r = .75, p < .01$) und *Wortreihe* ($r = .64, p < .01$) aus der *Kaufman Assessment Battery for Children*, Deutsche Version (K-ABC; Kaufman & Kaufman, 1991) und mit dem Subtest *PGN* ($r = .57, p < .01$) und *Satzgedächtnis* ($r = .50, p < .01$) aus dem SETK 3-5 (Grimm, 2015). Die Kriteriumsvalidität des Subtests *Zahlen Merken* stützt sich auf hohe Korrelationen mit Leistungen im Subtest *Unmittelbares Reproduzieren-numerisch-vorwärts* aus dem Adaptiven Intelligenzdiagnostikum (AID 3; Kubinger & Holocher-Ertl, 2014) in der Schuleingangsphase ($r = .66, p < .01$) sowie auf Untersuchungen an klinischen Stichproben (Kinder mit Down-Syndrom, Autismus).

Konstruktferne Fähigkeitsdimensionen: Grob - und Visuomotorik.

Zur Ermittlung der diskriminanten Validität wurden die Subtests *Turnen* mit 10 Items zur Erfassung der Grobmotorik (vorgezeigte Turnübungen nachahmen; Möglichkeit mehrerer Versuche) und *Nachzeichnen* mit 10 Items zur Erfassung der Visuomotorik (geometrische Figuren von Vorlagen nachzeichnen; festgelegte Vorgabereihenfolge von Items) aus dem WET (Kastner-Koller & Deimann, 2012) eingesetzt und Items dichotom verrechnet. Die innere Konsistenz (Cronbachs α) betrug in der vorliegenden Gesamtstichprobe für den Subtest *Turnen* .72 (.76 in

Teilstichprobe 1 und .67 in Teilstichprobe 2) und für den Subtest *Nachzeichnen* .76 (.77 in Teilstichprobe 1 und .72. in Teilstichprobe 2)

Vorgehen

Der Konstruktion des Itempools, basierend auf Analysen des Testmaterials standardisierter Messinstrumente und – hinsichtlich der Akzentuierung – zugehöriger Audioaufzeichnungen, folgte dessen Überprüfung im Rahmen einer Voruntersuchung, welche nachmittags in privaten Räumlichkeiten der Familien getesteter Kinder aus dem Verwandten- und Bekanntenkreis stattfand, vorwiegend unter Anwesenheit eines Elternteils, mittels Einsatzes eines Teddybären und einer spielerischen Instruktion zum Nachsprechen der „Wörter in der Sprache des Teddys“. Mit dem Ziel, bei den KWN-Aufgaben hinsichtlich Aussprache und Akzentuierung eine hohe Durchführungsobjektivität zu erreichen, erhielten die, an der Datenerhebung beteiligten, Masterandinnen für die vorliegende Untersuchung erstellte Tonaufzeichnungen der vorgesprochenen Kunstwörter und es erfolgte eine Einschulung in die Bewertung der Kunstwörter anhand von fiktiven, auf den Protokollbögen eingetragenen, Beispielen möglicher Antworten (siehe Anhang C). Vor der Datenerhebung erfolgten Einarbeitungen der Masterandinnen (= Testleiterinnen) in die standardisierte Durchführung der eingesetzten Aufgaben und pro Masterandin Durchführungen von mindestens zwei Probetestungen bei Kindern aus dem Bekanntenkreis in deren privaten Räumlichkeiten, welche hinsichtlich der eingesetzten Messinstrumente und Durchführung mit den Hauptuntersuchungen ident waren. An pädagogische Leiterinnen der St. Nikolausstiftung und KIWI-Kindergärten wurden schriftliche Ansuchen um Genehmigung zur Datenerhebung inklusive Informationsblättern zu Zielen und Ablauf der Untersuchung übermittelt. Es folgten persönliche Kontaktaufnahmen mit Kindergartenleiterinnen disponibler Kindergartenstandorte in Wien unter Aushändigung von Informationsblättern für MitarbeiterInnen der Kindergärten und für Eltern bzw. Erziehungsberechtigte mit einem Abschnitt zur schriftlichen Teilnahme-Einverständniserklärung (siehe Anhang B).

Die Datenerhebung erstreckte sich von Anfang Februar bis Ende März 2017 und erfolgte an drei Kindergartenstandorten der St. Nikolausstiftung, aufgeteilt auf jeweils eine Masterandin, sowie an zwei KIWI-Kindergärten, aufgeteilt auf jeweils zwei Masterandinnen. Die Erhebung soziodemografischer Daten erfolgte über Elternfragebögen, die Kindergartenpädagoginnen an Eltern aushändigten und bei Retournierung an die Masterandinnen übergaben. Die Testleiterinnen erhoben pro Kind Daten an zwei Testterminen, die jeweils ca. 30 Minuten dauerten und innerhalb eines geringstmöglichen Tagesintervalls stattfanden. Mit Ausnahme eines Kindes aus der Stichprobe für die Probetestungen, bei dem die Aufgaben beider Protokollbögen an einem

Tag durchgeführt wurden, betrug das Intervall zwischen den zwei Testterminen in der Gesamtstichprobe 1 bis 15 Tage ($M = 4.91 \pm 3.70$ Tage), in Teilstichprobe 1 variierte es zwischen 1 bis 15 Tagen ($M = 4.90 \pm 3.87$ Tage) und in Teilstichprobe 2 zwischen 1 bis 14 Tagen ($M = 4.93 \pm 3.57$ Tage). Antworten der Kinder bei den KWN-Aufgaben wurden auf private Mobiltelefone aufgezeichnet. Die Testungen wurden ausschließlich bei Kindern mit retournierter schriftlicher Einverständniserklärung eines/r Erziehungsberechtigten in, von den Kindergärten zur Verfügung gestellten, Räumlichkeiten durchgeführt und erfolgten, zur Erreichung der höchsten Leistungsfähigkeit der Kinder, vornehmlich vormittags zwischen 9 und 12 Uhr, in Abstimmung auf Essens- und Schlafenszeiten der jeweiligen Kindergruppen und mit höchstem Stellenwert auf Bedürfnisse und Wohlbefinden der Kinder. Die Masterandinnen nahmen für die von ihnen getesteten Kinder Kodierungen von Testergebnissen, inklusive Transkriptionen und Bewertungen von Antworten bei den KWN-Aufgaben vor (ausgenommen der Bewertungsmodi EK und EKH, die von der für die vorliegende Untersuchung verantwortlichen Masterandin angewandt wurden). Dateneingaben in das Statistikprogramm SPSS und in Excel erfolgten ebenso durch die Masterandinnen. Für Schätzungen der Beurteilerinnenübereinstimmungen für die KWN-Leistungen wählte jede Masterandin aus der Stichprobe der Testkinder ihrer Datenerhebung nach dem Zufallsprinzip je ein Kind aus jeder Altershalbjahresgruppe unter Herstellung eines Geschlechterverhältnisses von 50:50 aus und übermittelte Tonaufzeichnungen der Kinder den zwei anderen Masterandinnen, welche diese unabhängig voneinander blind transkribierten und nach den Bewertungsmodi St, E und H bewerteten. Statistische Analysen für die vorliegende Untersuchung erfolgten in Bezug auf spezifizierte Fragestellungen, unabhängig von Fragestellungen der anderen Masterandinnen.

Datenanalyse

Die Datenanalysen erfolgten aufgrund nicht spezifizierter Hypothesen mangels experimentellen Designs mittels zweiseitiger Testungen bei einer Irrtumswahrscheinlichkeit von $\alpha = .05$. Die angestrebte Testmacht von $1 - \beta = .80$ konnte bei Korrelationsanalysen erreicht werden; bei Mittelwerts- und Varianzanalysen lag die Testmacht vorliegender Untersuchung wegen Unterschreitung der nach (Power-)Analysen mit G*Power Version 3.1.9.2 (Faul, 2014) ermittelten erforderlichen Teilstichproben- und Subgruppenstichprobengrößen unter der angestrebten Testmacht (erforderliche Stichprobengrößen bei Erwartung großer Effekte und bei gleich großem Verhältnis der Subgruppenstichprobengrößen $n = 51$ bis 67 Kinder). Zur Vermeidung einer weiteren Verringerung der Power erfolgten das Lebensalter betreffende Analysen in Ganzjahres- statt Halbjahresschritten. Dem Großteil der Datenanalysen diente die Statistiksoftware IBM SPSS Amos Version 24.0 (IBM Corporation, 2016), testtheoretischen Analysen anhand von

probabilistischen Methoden diente das Statistikprogramm *R* Version 3.3.3 (R Core Team, 2017) unter Anwendung des Softwarepakets *extended Rasch modeling* (eRm) Version 0.15-7 (Mair, Hatzinger, & Maier, 2016; Poinstingl, Mair & Hatzinger, 2007) und zweifachen, nicht-parametrischen Varianzanalysen das Programm mKVA Version 1.4.20 (Häusler & Kubinger, 2001). *Missings by design* bei 3-jährigen Kindern (nicht vorgegebene Items 18 bis 20) wurden für Reliabilitätsanalysen und nicht-parametrische Überprüfungen auf Gültigkeit des dichotomen logistischen Testmodells von Rasch (1 Parameter-Model, 1-PL Modell; Rasch, 1960) mittels logischer Imputation durch den Wert Null ersetzt. Ein Großteil der Analysen wurde beim vollständigen und selektierten Itemset durchgeführt, ein Teil der Analysen ausschließlich beim selektierten Itemset. Hinsichtlich Voraussetzungsanalysen erfolgten Überprüfungen auf Normalverteilung für Korrelationsanalysen grafisch mittels Histogrammen, P-P-Plots und z-Transformationen der Schiefe- und Kurtosiskoeffizienten, mit der Annahme einer Voraussetzungsverletzung bei z-Werten ± 1.95 ($\alpha = .05$), für Mittelwertsunterschieds- und Varianzanalysen mittels Shapiro-Wilk-Tests (Shapiro & Wilk, 1965; Shapiro, Wilk, & Chen, 1968) aufgrund, im Vergleich zu Kolmogorov-Smirnov-Tests, höherer Power bei kleinen Stichprobengrößen ($n < 50$) (Ghasemi & Zahediasl, 2012). Überprüfungen auf Varianzhomogenität erfolgten mittels Levene-Tests (Levene, 1960). Ergebnisse der Voraussetzungstests sind Anhang D zu entnehmen, jene der Überprüfungen auf Schiefe und Kurtosis – wenn nicht im Rahmen der Datenanalysen angeführt – sind in Anhang E. Fehler in den Antworten bei den KWN-Itemsets wurden durch Bildung der Differenz zu dem nach dem konservativen Bewertungsmodus St gebildeten Summenscore ermittelt bzw. für die kombinierten Bewertungsmodi EK und EKH durch Bildung der Differenzen zu Bewertungsmodus St sowie E und/oder und H.

Parametrische Überprüfungen der Eindimensionalität der vollständigen Itemsets bzw. der Gültigkeit des 1-PL Modells (Rasch, 1960) lieferten Schätzungen anhand des bedingten Maximum-Likelihood-Algorithmus auf Basis unvollständiger Datenmatrizen und erfolgten – getrennt nach Itemset und Bewertungsmodus – anhand des internen Trennungskriteriums Median des Rohscores (niedrig versus hoch) sowie der externen Trennungskriterien Geschlecht und Alter (dichotomisiert in jüngere [$\leq 4;5$ Jahre] und ältere [$\geq 4;6$ Jahre] Kinder) mittels Likelihood-Quotiententests (LQT) nach Andersen (Andersen, 1973) sowie z-Tests nach Wald (1943) und grafischer Modelltests. A priori Adjustierungen des Signifikanzniveaus von $\alpha = .05$ auf α/Anzahl an Trennungskriterien im Rahmen von LQTs nach Andersen (vgl. Koller, Alexandrowicz, und Hatzinger, 2012) dienten der Verringerung einer Alpha-Fehler-Kumulierung bei abhängigen Tests. Die, für kleine Stichprobengrößen geeigneten, nicht-parametrischen, auf Bootstrapping

basierenden Rasch-Analysen mittels (quasi-)exakter Tests nach Ponocny (2001) unter Anwendung des Markov-Chain Monte-Carlo Algorithmus (Verhelst, 2008) wurden a posteriori durchgeführt, mit den zur Verringerung oder Eliminierung von seriellen Abhängigkeiten und Standardfehlern geeigneten Einstellungen $burn_in = 1000$, $step = 64$, und n_eff , $seed = 0$ (vgl. Verhelst, 2008) zur Generierung von 128000 Matrizen ($step \times [burn_in + n_eff]$). Tests T_{10} galten der Überprüfung auf spezifische Objektivität und Subgruppeninvarianz anhand derselben Trennungskriterien wie in parametrischen Rasch-Analysen, T_{11} und T_{md} der Überprüfung auf lokal stochastische Unabhängigkeit und/oder Eindimensionalität. Itemanalysen nach klassisch-testtheoretischen Methoden basierten auf Ermittlung der, bei dichotomen Items durch die mittlere Lösungshäufigkeit definierten, Itemschwierigkeit (p_i) (Koeffizienten $< .20$ = hohe Itemschwierigkeiten; Koeffizienten $> .80$ = niedrige Itemschwierigkeiten), der mittleren Standardabweichung bzw. Itemvarianz $Var(x_i)$ und der Itemtrennschärfe (r_{it}) als korrigierte Korrelation zwischen einem Item und dem gesamten Itemset. Die Itemselektion erfolgte nach konventionellen Cutoffs (Bühner, 2006; Field, 2013; Fisseni, 2004) mit dem Ziel, bei Elimination von Items mit inakzeptablen bzw. niedrigen Itemtrennschärfen ($r_{it} \leq .29$) eine Zusammenstellung aus Items mit mittleren ($.30 \leq r_{it} \leq .50$) bis hohen ($r_{it} \geq .51$) Itemtrennschärfen und einer breiten Streuung von Itemschwierigkeitskoeffizienten zwischen $.20$ und $.80$ zu erhalten, unter Berücksichtigung des, hinsichtlich der inhaltlichen Validität erforderlichen, Erhalts von Items jeder Silbenanzahl.

Schätzungen der Beurteilerinnenübereinstimmung zwischen unabhängigen Bewertungen dreier Kodiererinnen erfolgten auf Itemebene mittels Fleiss'Kappa (κ) (1971) unter Anwendung des SPSS-Makros „STATS FLEISS KAPPA.spe“; der Interpretation ermittelter Koeffizienten dienten Cutoffs von Fleiss (1981). Schätzungen der inneren Konsistenz basierten auf den, aus Kovarianzmatrizen resultierenden, unstandardisierten Cronbachs α s, jene der Split-half-Reliabilität auf Spearman-Brown-Korrekturen. Nach Kriterien für die Individualdiagnostik wurde der Cutoff für die angestrebte Reliabilität auf Koeffizienten $> .8$ gesetzt (Döring & Bortz, 2016; Moosbrugger, 2012; Yousfi & Steyer, 2006), die Interpretation von Reliabilitätskoeffizienten erfolgte nach Nunnally und Bernstein (2006). Schätzungen der Intracoder-Reliabilität für Bewertungen der protokollierten und der von Tonaufzeichnungen transkribierten Antworten derselben Testleiterinnen (jeweils Bewertungsmodus St) erfolgten auf Itemebene mittels Cohens κ , die Interpretation von Koeffizienten nach Cutoffs von Fleiss (1981). Den Überprüfungen auf konvergente und diskriminante Validität der Itemsets dienten partielle Korrelationen bei Kontrolle des Lebensalters in Monaten bzw. bei Verletzung der Normalverteilungsvoraussetzungen partielle Rangkorrelationen nach Spearman durch Erstellung einer Rangkorrelationsmatrix und deren Einfügung in die partielle Korrelationsprozedur mittels „MATRIX IN“-Syntax Befehl.

Die Überprüfung auf Unterschiede zwischen Korrelationskoeffizienten erfolgte bei abhängigen Korrelationskoeffizienten mittels *z*-Tests bei Korrekturen nach Lenhard und Lenhard (2014) mit der Software Psychometrica.

Überprüfungen auf Effekte der Version des Itemsets und des Alters erfolgten für die vollständigen Itemsets mittels zweifaktorieller Varianzanalysen (ANOVAs) und paarweiser Vergleiche im Rahmen einer *simple effect analysis* mit LSD-Korrektur bzw. bei Verletzungen der Normalverteilungsvoraussetzung und/oder der Varianzhomogenität mittels zweifacher Rangvarianzanalysen sowie paarweiser exakter Mann-Whitney-Tests. Den Überprüfungen auf Alterseffekte sowie auf Testleiterinneneffekte im selektierten Itemset dienten einfaktorielle ANOVAs und – bei deren Signifikanz – wiederholte Kontrasttests, zur Vermeidung eines Verlusts an Power ausschließlich in Bezug auf interessierende Vergleiche, sowie polynomiale Kontrasttests bzw. bei Verletzungen der Normalverteilungs- und/oder der Varianzhomogenitätsvoraussetzung Kruskal-Wallis-Tests und paarweise exakte Mann-Whitney-Tests. Analysen auf Geschlechtseffekte und Effekte des Testtermins erfolgten mittels *t*-Tests für unabhängige Stichproben bzw. bei Verletzung der Normalverteilungsvoraussetzung mittels exakter Mann-Whitney-Tests. Die Überprüfung auf einen Kunstwortlängeneffekt (basierend auf gebildeten Mittelwerten pro Silbenanzahl) erfolgte aufgrund von Verletzungen der Normalverteilungsvoraussetzung nicht-parametrisch mittels Friedman-Test und paarweiser Vergleiche zwischen Mittelwerten aller Silbenanzahlen mittels exakter Wilcoxon-Rangsummen-Tests. Für jegliche Kontrasttests und paarweise Vergleiche wurden Bonferroni-Adjustierungen des Signifikanzniveaus vorgenommen ($\alpha = .05/\text{Anzahl durchgeführter Tests}$).

Als Maß für Effektstärken wurde für Globaltests einfaktorieller ANOVAs Omega-Quadrat (ω^2) als weniger verzerrtes Maß ermittelt und als Anteil erklärter Varianz nach Cohen (1988) interpretiert, für paarweise- bzw. Mittelwertsvergleiche Cohens *d* unter Anwendung der Hedges-Korrektur der Teilstichprobengrößen (Hedges *g*) (Cumming, 2012; Lakens, 2013) mit Interpretation der Koeffizienten nach Cohen (1988). Für zweifaktorielle ANOVAs und nicht-parametrische Tests wurden Effektstärken ausschließlich für Kontrasttests bzw. für paarweise Vergleiche signifikanter Modelltests durch approximative Ermittlung von *r* durch Bildung des Quotienten aus der standardisierten Teststatistik *z* und der Quadratwurzel der Fallzahl angegeben und nach Cutoffs von Cohen (1988) für Korrelationskoeffizienten interpretiert. Tabelle 6 gibt einen Überblick über die für vorliegende Untersuchung herangezogenen Cutoffs zur Interpretation diverser Koeffizienten und Effektstärken.

Tabelle 6

Cutoffs zur Interpretation von Koeffizienten bzw. Effektstärken

		mangelhaft ¹ / inakzeptabel ²	genügend ¹ / akzeptabel ²	gut ¹ / hoch ²	sehr gut ¹ / exzellente ²	klein/ niedrig	mittel	groß/ hoch
Itemtrennschärfe						<.30	.30-.50	>.50
Itemschwierigkeit						< .20	.20-.80	> .80
Inter-/Intracoder- Reliabilität ^a	Fleiss' / Cohens κ	< .40	.40	.60	.75			
Reliabilität ^b	α , r_{tt}	< .70	.70	.80	.90			
Korrelation ^c	r , r_s					.10	.30	.50
Varianzanalyse ^c	ω^2					.01	.06	.14
Mittelwertsdifferenz ^c	g					0.2	0.5	0.8

Anmerkung. ^a(Fleiss, 1981), ^b(Nunnally & Bernstein, 1994), ^c(Cohen, 1988), ¹für κ , ²für α , r_{tt}

Ergebnisse

Fehlende Daten

Der Anteil fehlender Daten mangels Kooperationsbereitschaft von Kindern (keine verbalen Reaktionen) betrug beim Subtest *Zahlen Merken* 1.72% und bei KWN-Aufgaben insgesamt 5.17% bzw. 0.00% bei Itemset 1 und 11.1% bei Itemset 2. Kinder, welche das Nachsprechen von Kunstwörtern verweigerten, waren heterogenen Alters und Geschlechts (zwei Mädchen: 4;5 und 5;0 Jahre, ein Bub: 3;1 Jahre); bei zwei der Kinder wurden die Kunstwörter beim ersten Testtermin, bei einem Kind beim zweiten Testtermin vorgegeben. Ein Kind hatte beim *Zahlen Merken* ein überdurchschnittliches ein weiteres hatte ein unterdurchschnittliches Ergebnis und ein Kind verweigerte die Kooperation beim *Zahlen Merken*. Neben der fehlenden Kooperationsbereitschaft bei den KWN-Aufgaben und beim *Zahlen Merken* verweigerte ein Kind einen Subtest zur Grobmotorik und ein weiteres Kind verweigerte einzelne Grobmotorik-Aufgaben.

Gedächtnisleistungen in Teilstichproben

Die Durchführungszeit variierte bei Itemset 1 zwischen 2 Minuten und 24 Sekunden (2:24 Minuten) und 6:53 Minuten ($M = 3:49$, $SD = 1.05$), bei Itemset 2 zwischen 2:10 und 5:42 Minuten ($M = 3.40$, $SD = 0:55$). Tabelle 7 enthält altersübergreifende sowie -spezifische deskriptive Statistiken zu KWN-Summenscores unter Anwendung des Bewertungsmodus St sowie zu Ziffernspannen und C-Werten im *Zahlen Merken*; für geschlechtsspezifische deskriptive Statistiken siehe Anhang E. Schiefe- und Kurtosisparameter zeigten in allen Alters- und Geschlechts-

gruppen beider Stichproben normalverteilte KWN-Summenscores, hinsichtlich der Ziffernspanne zeigten diese, bei sonst normalverteilten Leistungen, signifikant linksschiefe und schmalgipfelige Verteilungen der Leistungen bei 4-Jährigen und bei Buben der Stichprobe 1 sowie bei 3- und 5-Jährigen und beiden Geschlechtsgruppen der Stichprobe 2. Die mittleren Ziffernspannen in Stichproben vorliegender Untersuchung waren mit Mittelwerten der Normstichprobe des WET (Kastner-Koller & Deimann, 2012) in Bezug auf den Gesamtwert sowie Altersgruppen großteils vergleichbar; Abweichungen in Form von niedrigeren Ziffernspannen waren in Stichprobe 2 numerisch höher als in Stichprobe 1 bzw. am höchsten bei 4-Jährigen entsprechender Stichprobe. Mittlere gerundete C-Werte im *Zahlen Merken* indizierten, bei sonst für beide Teilstichproben im Normalbereich liegenden Leistungen, Förderbedarf bei 4-Jährigen der Stichprobe 2.

Tabelle 7

Stichprobenspezifische Gedächtnisleistungen bei Kunstwörtern (BM St) und Ziffernfolgen

	KWN St _{sum}				Zahlen Merken _{span}				Zahlen Merken _{C-Wert}			
	Alter (Ganzjahre)			Gesamt	Alter (Ganzjahre)			Gesamt	Alter (Ganzjahre)			Gesamt
	3	4	5		3	4	5		3	4	5	
Stichprobe 1												
<i>n</i>	7	17	7	31	6	17	7	30	6	17	7	30
<i>M</i>	3.71	8.71	11.14	8.13	2.50	3.24	3.86	3.23	3.83	4.06	5.00	4.23
<i>SD</i>	1.70	2.85	4.02	3.87	0.55	0.97	0.69	0.94	1.17	1.64	2.00	1.65
Sk.	-0.23	-0.01	-0.58	0.18	0.00	-2.39***	0.17	-1.32***	-0.67	-0.78	-0.18	-0.25
Ku.	-0.54	0.81	1.45	-0.41	-3.33	7.86***	0.34	3.86***	-0.45	0.81	-0.15	0.47
Stichprobe 2												
<i>n</i>	12	9	6	27	12	9	6	27	12	9	6	27
<i>M</i>	5.17	6.33	10.00	6.63	2.42	2.33	3.33	2.59	4.17	2.67	3.67	3.56
<i>SD</i>	2.62	2.65	1.41	3.01	0.90	1.32	0.52	1.05	1.12	1.58	1.86	1.55
Sk.	0.91	0.53	0.00	0.19	-1.95***	-1.62*	0.97	-1.68***	-0.39	-1.25	0.72	-0.51
Ku.	0.84	-0.64	-0.30	-1.16	4.37***	0.74	-1.88	2.44***	-0.06	0.17	-1.88	0.48
Normierungsstichprobe WET: Ziffernspannen im <i>Zahlen Merken</i>												
<i>n</i>					415	463	367	1245				
<i>M</i>					2.38-	3.39-	3.78-	3.37				
					2.87	3.52	4.05					

Anmerkung. Sk. = Schiefe, Ku. = Kurtosis, sum = Summenscore, span = Gedächtnisspanne, .

* $p < .05$, ** $p < .01$, *** $p < .001$.

Deskriptivstatistisch sowie bei Analyse von Profildiagrammen (Abbildung 1) zeigte sich in Stichprobe 1 im Vergleich zu Stichprobe 2 über Altersgruppen in Ganzjahresschritten ein deutlicher linearer Anstieg der mittleren KWN-Summenscores (Bewertungsmodus St) sowie Ziffernspannen im *Zahlen Merken*. Interferenzstatistische Ergebnisse hinsichtlich der Effekte Alter, Itemset bzw. Teilstichprobe, Geschlecht und Testtermin sowie deskriptive Statistiken zu fehlerhaften Antworten sind Anhang E zu entnehmen.

Vorläufige psychometrische Analysen der Gesamtdaten

Psychometrische Qualität nach probabilistischer Testtheorie.

Ergebnisse von LQTs nach Andersen zur parametrischen Modellkontrolle der Itemsets 1 und 2 basierten auf automatischem Ausschluss von nicht schätzbaren Items wegen ungünstiger Verteilungen von Antwortmustern in Subgruppen; die Anzahl der in die Parameterschätzungen aufgenommenen Items geht durch $df+1$ hervor (Tabelle 8). Erste Parameterschätzungen ($\alpha_{adj} = .017$) zeigten für Kunstwörter des Itemsets 1 bei Anwendung aller Bewertungsmodi bzw. für jene des Itemsets 2 bei Anwendung der Bewertungsmodi E, H und EKH hinsichtlich aller Trennungskriterien Raschhomogenität in Form von Subgruppeninvarianz; signifikante Modellabweichungen ergaben sich für Itemset 2 bei Anwendung der Bewertungsmodi St und EK hinsichtlich des internen Trennungskriteriums Median.

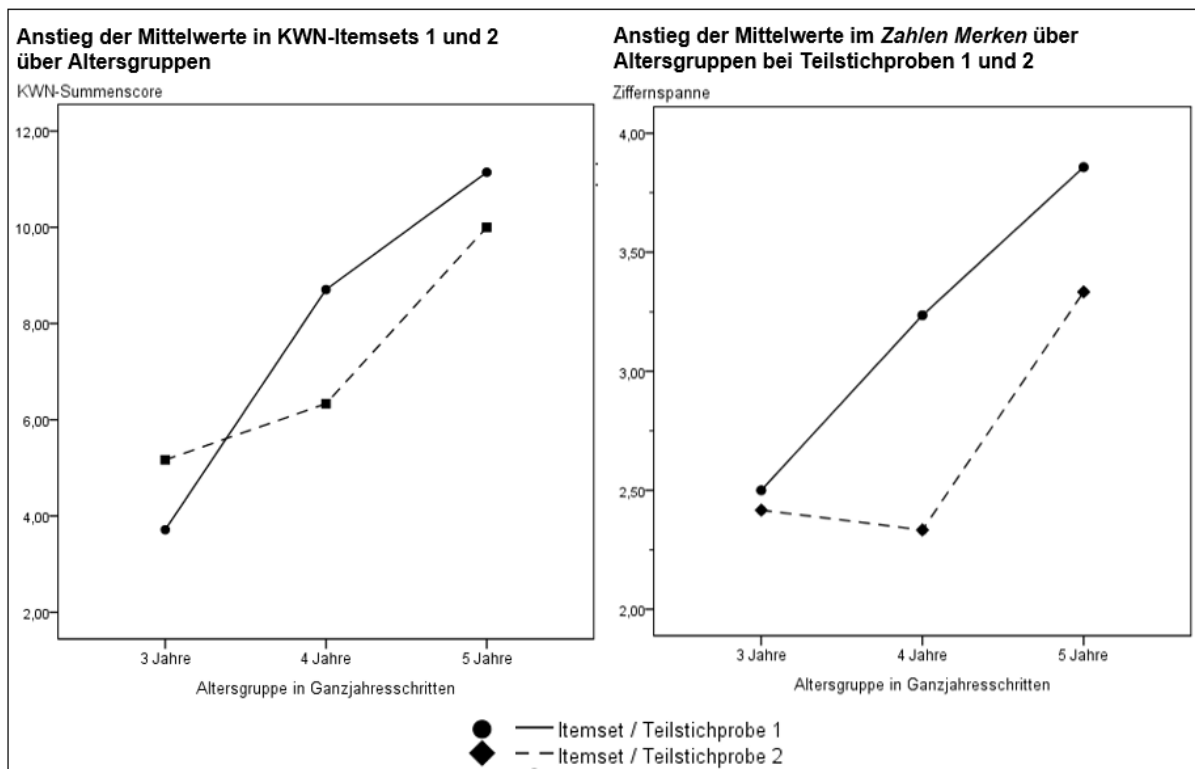


Abbildung 1. Alterstrends in KWN-Summenscores und Ziffernspannen in Itemsets 1 und 2

Ergebnisse itemspezifischer Analysen mittels z-Tests nach Wald sind in Anhang F. Bei nicht-parametrischen Rasch-Analysen des, sich als rasch-homogen erwiesenen, Itemsets 1 bei Bewertungsmodus St konnten für alle Items Schätzungen erfolgen. Der Test $T_{10}(\alpha_{\text{adj}} = .017)$ ergab mit nicht signifikanten Ergebnissen nach den Trennungskriterien Median ($n_{\text{Leistung} < \text{Median}} = 11$, $n_{\text{Leistung} > \text{Median}} = 20$, $p = .422$), Geschlecht ($n_{\text{Mädchen}} = 16$, $n_{\text{Buben}} = 15$, $p = .651$) und Alter ($n_{\text{Jüngere}} = 13$, $n_{\text{Ältere}} = 18$, $p = .054$) Raschhomogenität hinsichtlich der Annahmen spezifischer Objektivität und Subgruppeninvarianz. Nicht signifikante Ergebnisse der Tests T_{11} ($p = .101$) sowie T_{md} mit einer nicht signifikanten Korrelation der Personenscores zwischen 2- bis 3-silbigen sowie 4- bis 6-silbigen Items, $r = .72$, $p = .876$, zeigten Gültigkeit des 1-PL-Modells hinsichtlich lokal stochastischer Unabhängigkeit und/oder Eindimensionalität der Messungen.

Tabelle 8

Likelihood-Quotienten-Tests: Itemsets 1 und 2, getrennt nach Bewertungsmodus

Prüfgrößen aus den Likelihood-Quotiententests nach Trennungskriterien												
	Intern (Median)				Alter				Geschlecht			
	χ^2_{emp}	χ^2_{krit}	df	p	χ^2_{emp}	χ^2_{krit}	df	p	χ^2_{emp}	χ^2_{krit}	df	p
Itemset 1												
St	6.77	28.81	15	.964	15.59	23.12	11	.157	28.49	32.93	18	.424
E	7.14	31.57	17	.982	22.64	24.57	12	.031	21.45	32.93	18	.257
H	10.12	27.41	14	.753	12.49	19.68	11	.328	13.39	32.93	18	.768
EK	13.31	30.20	16	.861	27.50	27.41	14	.017	23.58	32.93	18	.169
EKH	15.39	30.20	16	.497	26.99	26.00	13	.017	16.91	32.93	18	.530
Itemset 2												
St	28.49	26.00	13	.008	13.74	28.81	15	.545	19.04	30.20	16	.267
E	25.16	26.00	13	.022	12.57	27.41	14	.561	13.50	30.20	16	.636
H	22.16	26.00	13	.053	13.15	28.81	15	.590	22.27	30.20	16	.135
EK	30.15	26.00	13	.004	13.65	27.41	14	.476	12.50	30.20	16	.709
EKH	21.52	28.81	15	.121	12.10	27.41	14	.598	14.52	30.20	16	.558

Anmerkung. St, E, H, EK, EKH = Bewertungsmodi, df + 1 = Anzahl geschätzter Items.

^a $\alpha_{\text{adj}} = .017$.

Psychometrische Qualität nach klassischer Testtheorie.

Tabelle 9 enthält Statistiken der pro Bewertungsmodus aggregierten Itemkennwerte (für item-spezifische Statistiken siehe Anhang G) sowie der Reliabilitätsschätzungen und partiellen Rangkorrelationen (Kontrolle des Lebensalters in Monaten) der KWN-Summscores und -Silbenspannen mit Ziffernspannen. Die Items erwiesen sich durchschnittlich als mittelschwer, die mittleren Itemtrennschärfen verringerten sich numerisch von Bewertungsmodus St und H bis zu E, EK und EKH und lagen bei Itemset 1 im akzeptablen Bereich und bei Itemset 2 im Verwerfungsbereich. Die höchsten Itemtrennschärfen zeigten sich bei Bewertungsmodus St und H des Itemsets 1. Das Itemset 1 wies mit Cronbachs $\alpha > .7$ akzeptable innere Konsistenzen mit dem numerisch höchsten Koeffizienten für den Bewertungsmodus St und bis zu Bewertungsmodus EKH linear sinkendem Muster auf, das Itemset 2 inakzeptable innere Konsistenzen; überlappende Konfidenzintervalle indizierten keine signifikanten Unterschiede zwischen inneren Konsistenzen der Bewertungsmodi pro Itemset sowie zwischen Itemsets 1 und 2.

Tabelle 9

Itemkennwerte und Reliabilitätsschätzungen der Itemsets vor Itemselektion

	Itemkennwerte $M(SD)$			Reliabilität		Validität: $r_{ZM_{span}}$	
	p_i	$Var(x_i)$	r_{it}	Cronbachs α [KI]	r_{tt}	KWN_{sum}	KWN_{span}
Itemset 1							
St	.41	.42	.36	.80 [.67, .89]	.81	.49**	.33
E	.43	.42	.32	.79 [.67, .89]	.80	.48**	.39*
H	.43	.41	.36	.76 [.62, .87]	.79	.52**	.31
EK	.44	.42	.30	.74 [.60, .86]	.81	.43*	.40*
EKH	.47	.41	.30	.73 [.58, .85]	.79	.51**	.42*
Itemset 2							
St	.37	.44	.25	.66 [.44, .82]	.80	.22	.40*
E	.40	.44	.22	.66 [.44, .82]	.76	.30	.26
H	.39	.44	.25	.61 [.36, .79]	.71	.28	.32
EK	.40	.44	.23	.63 [.39, .80]	.71	.28	.26
EKH	.42	.44	.23	.62 [.37, .80]	.63	.33	.17

Anmerkung. St, E, H, EK, EKH = Bewertungsmodi, p_i = Itemschwierigkeit, $Var(x_i)$ = Itemvarianz, r_{it} = Itemtrennschärfe, r_{tt} = Split-half-Reliabilität, r = partielle Rangkorrelation, ZM = Zahlen Merken, KWN = Kunstwörter nachsprechen, sum = Summscore, span = Gedächtnisspanne.

* $p < .05$, ** $p < .01$, *** $p < .001$.

Die konvergente Validität betreffend zeigten sich bei Itemset 1 mittlere bis hohe, signifikant positive partielle Korrelationen zwischen KWN-Summscores und Ziffernspannen im *Zahlen Merken*. Zwischen KWN-Silbenspannen und Ziffernspannen zeigten sich vorwiegend mittlere, signifikant positive partielle Korrelationen; bei einigen Bewertungsmodi waren entsprechende Zusammenhänge nicht signifikant. Bei Itemset 2 waren partielle Korrelationen von KWN-Leistungen mit Ziffernspannen nicht signifikant, ausgenommen der mittleren, signifikant positiven Korrelation zwischen KWN-Silbenspanne und Ziffernspanne bei Bewertungsmodus St. Innerhalb des Itemsets 1 zeigte sich kein signifikanter Unterschied zwischen Korrelationen der KWN-Summscores und Ziffernspannen bei Bewertungsmodus St und jenen bei Bewertungsmodus E, $z = 0.23$, $p = .410$, Bewertungsmodus H, $z = -0.69$, $p = .246$, Bewertungsmodus EK, $z = 0.73$, $p = .232$, und Bewertungsmodus EKH, $z = -0.16$, $p = .439$.

Selektiertes Itemset: Psychometrische Qualität im Vergleich mit vertrautem Material

Eine Itemselektion nach klassischen Methoden erfolgte bei Itemset 1 in Bezug auf den Bewertungsmodus St aufgrund von entsprechenden, nach probabilistischer und klassischer Testtheorie, numerisch günstigeren psychometrischen Ergebnissen im Vergleich zu Stichprobe 2, in welcher bei 4-Jährigen auf Basis der Leistungen bei Ziffernfolgen unterdurchschnittliche Leistungen im phonologischen Gedächtnis anzunehmen waren, sowie im Vergleich zu anderen Bewertungsmodi. Die Itemselektion resultierte in einem um 40.0% der Items reduzierten Itemset, bestehend aus zwei 2-silbigen (Item 1, 4), drei 3-silbigen (Item 5, 6, 9), drei 4-silbigen (Item 10, 11, 12), drei 5-silbigen Kunstwörtern (Item 14, 16, 17) sowie einem 6-silbigen Kunstwort (Item 18). Summscores im selektierten Itemset wiesen eine hohe, signifikant positive partielle Korrelation (Kontrolle des Lebensalters in Monaten) mit Summscores im vollständigen Itemset unter Anwendung desselben Bewertungsmodus auf, $r = .94$, $p < .001$. Ein signifikanter Friedman-Test zeigte von der Silbenanzahl abhängige KWN-Leistungen, $\chi^2(4) = 63.22$, $p < .001$; paarweise Vergleiche mittels Wilcoxon-Rangsummen-Tests ($\alpha_{\text{adj}} = .005$) indizierten, dass 2-silbige Kunstwörter ($Mdn = 1.00$) signifikant häufiger korrekt reproduziert wurden als 3-silbige ($Mdn = 0.33$), $z = -4.29$, $p < .001$, $r = -.77$, 4-silbige ($Mdn = 0.33$), $z = -4.36$, $p < .001$, $r = -.78$, 5-silbige ($Mdn = 0.33$), $z = -4.74$, $p < .001$, $r = -.85$, und 6-silbige Kunstwörter ($Mdn = 0.00$), $z = -4.24$, $p < .001$, $r = -.82$, und 3-silbige Kunstwörter signifikant häufiger korrekt reproduziert wurden als 5-silbige, $z = -2.98$, $p = .003$, $r = .54$, und 6-silbige Kunstwörter, $z = -2.77$, $p = .003$, $z = -.53$. Keine signifikanten Unterschiede zwischen KWN-Leistungen ergaben sich zwischen 3- und 4-silbigen Kunstwörtern, $z = -1.39$, $p = .243$, $r = .24$, 4- und 5-silbigen, $z = -1.78$, $p = .108$, $r = .32$, 4- und 6-silbigen, $z = -2.35$, $p = .017$, $r = .45$, sowie 5- und 6-silbigen Kunstwörtern, $z = -1.80$, $p = .085$, $r = .32$.

Leistungen beim selektierten Itemset.

Die gesamten sowie alters- und geschlechtsspezifischen Leistungen der Teilstichprobe 1 im selektierten Itemset sind Tabelle 10 zu entnehmen. Schiefe- und Kurtosiskoeffizienten zeigten, bei Normalverteilung der KWN-Summenscores in allen Subgruppen, signifikant linksschiefe und schmalgipfelige Verteilungen der KWN-Silbenspannen bei 4- und 5-Jährigen und damit überproportional hohe Leistungen entsprechender Subgruppen (für Schiefe- und Kurtosiskoeffizienten zu Ziffernspannen siehe *Gedächtnisleistungen in Teilstichproben*). Einem *t*-Test für unabhängige Stichproben zufolge waren KWN-Summenscores des selektierten Itemsets zwischen Kindern, die Kunstwörter zum ersten Testtermin ($M = 4.75$, $SD = 2.25$) und Kindern, welche diese zum zweiten Testtermin erhielten ($M = 5.58$, $SD = 2.99$) vergleichbar, $t(29) = -0.73$, $p = .473$, $g = 0.27$. Eine einfaktorielle ANOVA ergab einen signifikanten großen Effekt des Alters auf KWN-Summenscores, $F(2, 28) = 14.53$, $p < .001$, $\omega^2 = .50$, mit linearem Trend großer Effektstärke bei polynomialem Kontrast, $F(1, 28) = 25.50$, $p < .001$, $\omega^2 = .44$.

Tabelle 10

Gedächtnisleistungen von Teilstichprobe 1 nach Itemselektion

	Alter (Ganzjahre)			Geschlecht)		Gesamt
	3	4	5	Mädchen	Buben	
<i>n</i>	7	17	7	15	16	31
KWN _{sum}						
<i>M</i>	1.43	5.94	7.43	5.47	5.06	5.26
<i>SD</i>	1.51	2.05	3.10	2.88	3.32	3.07
Range	0-4	2-10	2-11	0-11	0-10	0-11
Schiefe	1.00	0.24	-0.84	0.12	-0.08	-0.03
Kurtosis	-0.20	-0.10	0.23	0.01	1.25	-0.84
KWN _{span}						
<i>M</i>	1.86	4.71	5.00	4.33	3.94	4.13
<i>SD</i>	1.46	0.92	1.41	1.68	1.73	1.69
Range	0-4	2-6	2-6	0-6	0-6	0-6
Schiefe	-0.11	-1.52*	-1.98*	-1.35*	-1.04	-1.11
Kurtosis	-0.67	4.02**	4.40**	2.04	-0.09	0.46

Anmerkung. KWN = Kunstwörter nachsprechen, sum = Summencore, span = Gedächtnisspanne.

* $p < .05$, ** $p < .01$, *** $p < .001$.

Wiederholte Kontraste ($\alpha_{\text{adj}} = .025$) zeigten, dass die Anzahl an korrekt reproduzierten Kunstwörtern bei 4-Jährigen mit großem Effekt signifikant höher als bei 3-Jährigen war, $t(28) = 4.52$, $p < .001$, $g = 2.36$, und sich zwischen 4- und 5-Jährigen nicht unterschied, $t(28) = 1.49$, $p = .147$, $g = 0.63$. Ein Kruskal-Wallis-Test zeigte hinsichtlich der KWN-Silbenspannen ein vergleichbares Muster (siehe Anhang E). Hinsichtlich der Ziffernspannen zeigte ein Kruskal-Wallis-Test einen signifikanten Alterseffekt, $\chi^2(2) = 9.84$, $p = .007$, und Mann-Whitney-Tests ($\alpha_{\text{adj}} = .025$) ergaben bei 4-Jährigen signifikant höhere Ziffernspannen als bei 3-Jährigen ($Mdn_{3 \text{ Jahre}} = 2.50$, $Mdn_{4 \text{ Jahre}} = 3.00$), $U = 19.50$, $z = -2.42$, $p = .013$, $r = -.51$, bei keinem signifikanten Unterschied zwischen 4- und 5-Jährigen ($Mdn_{5 \text{ Jahre}} = 4.00$), $U = 37.00$, $z = -1.59$, $p = .171$, $r = -.32$. Grafisch (Abbildung 2) war bei Ziffernspannen ein kontinuierlicherer Verlauf im Alterstrend zu erkennen als bei KWN-Summenscores und -Silbenspannen, bei welchen 3-Jährige niedrigere Leistungen als bei Ziffernspannen zeigten und sich ein steilerer Anstieg von 3 auf 4 Jahre ergab. Mädchen und Buben zeigten auf Basis eines t -Tests für unabhängige Stichproben keinen Unterschied in KWN-Summenscores, $t(29) = -0.36$, $p = .720$, $g = 0.23$, ebenso wie – auf Basis von Mann-Whitney-Tests – in Silbenspannen ($Mdn_{\text{Mädchen}} = Mdn_{\text{Buben}} = 5.00$), $U = 105.50$, $z = -0.60$, $p = .575$, $r = -0.11$, und in Ziffernspannen ($Mdn_{\text{Mädchen}} = Mdn_{\text{Buben}} = 3.00$), $U = 87.50$, $z = -1.13$, $p = .302$, $r = -0.21$.

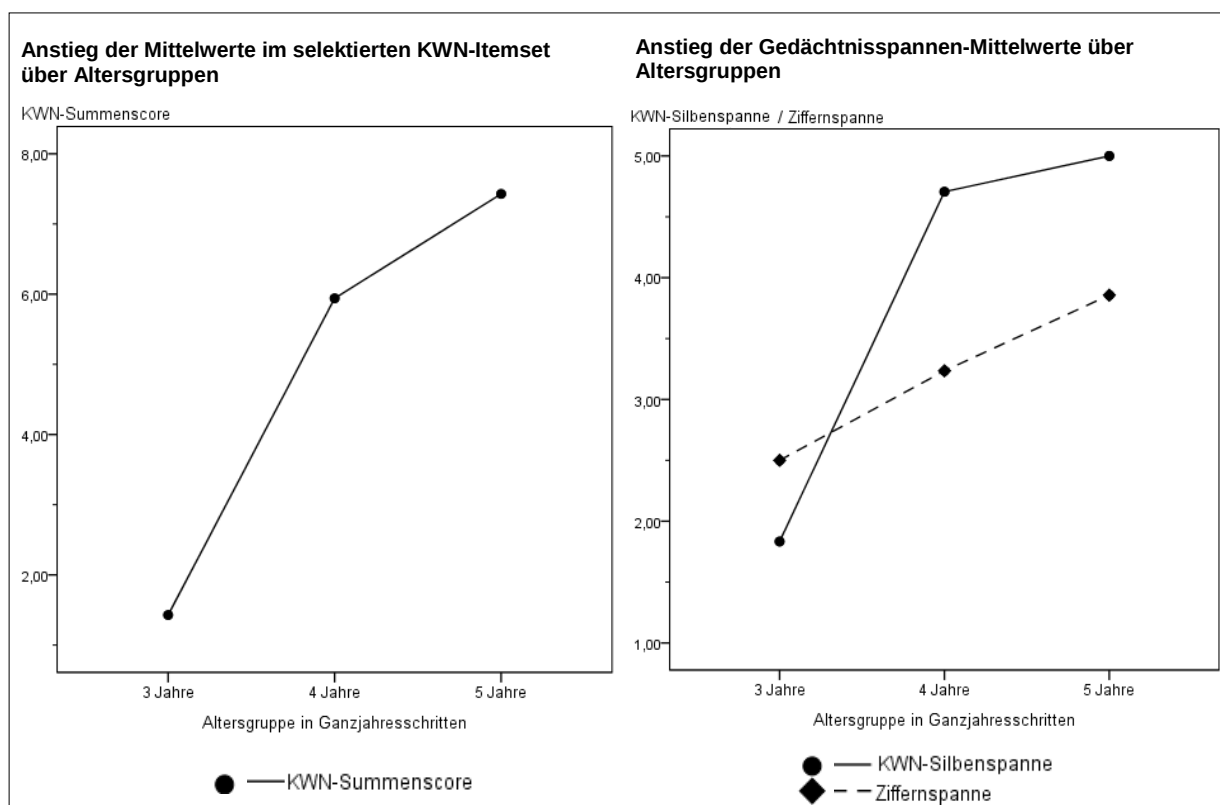


Abbildung 2. Alterstrends in KWN-Leistungen (selektiertes Itemset) und Ziffernspannen

Überprüfung des selektierten Itemsets nach probabilistischer Testtheorie.

Im Rahmen nicht-parametrischer Rasch-Analysen des selektierten Itemsets indizierten nicht signifikante Tests T_{10} ($\alpha_{\text{adj}} = .017$) hinsichtlich des internen Trennungskriteriums Median ($n_{\text{Leistung} < \text{Median}} = 14$, $n_{\text{Leistung} > \text{Median}} = 17$, $p = .664$) sowie des externen Trennungskriteriums Geschlecht ($p = .503$) Rasch-Homogenität; ein knapp signifikantes Ergebnis hinsichtlich des Trennungskriteriums Alter ($p = .016$) indizierte eine Verletzung der Subgruppeninvarianz bzw. eine differentielle Itemfunktion für jüngere und ältere Kinder. Nicht signifikante Tests T_{11} ($p = .408$) sowie T_{md} mit einer Korrelation von $r = .64$ ($p = .599$) zwischen Personenscores bei kurzen (2- bis 3-silbigen) und längeren (4- bis 6-silbigen) Kunstwörtern zeigten lokal stochastische Unabhängigkeit bzw. Eindimensionalität der Messungen.

Itemanalyse des selektierten Itemsets nach klassischer Testtheorie.

Tabelle 11 enthält für den gesamten Altersbereich sowie – basierend auf altersgruppenspezifischen Selektionen von Items mit Null-Varianzen – getrennt nach Altersgruppen geschätzte Itemkennwerte mit einem Überblick der Verteilungen von Items auf Trennschärfenbereiche; Tabelle 12 zeigt analog dazu silbenanzahlspezifische mittlere Itemschwierigkeiten und -trennschärfen. Das selektierte Itemset umfasste bei Anwendung des Bewertungsmodus St Itemschwierigkeiten von .16 bis .90, -varianzen von .30 bis .51 und -trennschärfen von .29 bis .53, deren Mittelwert – ausgenommen inakzeptabel hoher Itemschwierigkeit bzw. niedriger -trennschärfe beim 6-silbigen Item – insgesamt sowie silbenspezifisch innerhalb des Akzeptanzbereichs lag. Für 3-Jährige erwiesen sich Kunstwörter des selektierten Itemsets im Durchschnitt als eher schwierig und Kunstwörter ab 3 Silben als zu schwierig, für 4- und 5-Jährige waren Items im Durchschnitt mittelschwierig. In der Altersgruppe 4-Jähriger variierten Schwierigkeitskoeffizienten mit der Silbenanzahl und es ergab sich eine inakzeptabel hohe Itemschwierigkeit beim 6-silbigen Item; in der Altersgruppe 5-Jähriger waren Items jeder Silbenlänge mittelschwierig (ausgenommen 2-silbiger Kunstwörter mit 100% Lösungswahrscheinlichkeit) und Schwierigkeitskoeffizienten bei 4- und 5-silbigen Items ident. Bei altersübergreifend sowie für 3- und 5-Jährige im akzeptablen Bereich liegenden mittleren Itemtrennschärfen zeigten die Items für 4-Jährige eine hohe durchschnittliche Itemtrennschärfe. Differenzierungsfähigkeiten von Items variierten altersgruppenspezifisch; für 3-Jährige waren Item 6 und alle 5- und 6-silbigen Items unzureichend trennscharf, für 5-Jährige alle 2-silbigen Items sowie zwei 5-silbige Items (14, 17) und das 6-silbige, die sich bei 4-Jährigen – ebenso wie der Großteil der Items – als hoch trennscharf erwiesen.

Tabelle 11

Mittlere Itemkennwerte und Verteilungen von Items auf Trennschärfenbereiche

		Nummern der Items mit Trennschärfen im Bereich von							M_{rit}
	Item- anzahl	M_{pi}	< .30	.30-.39	.40-.49	.50-.59	.60-.69	> .70	
Gesamt	12	.44	18	4, 16, 17	1, 14	5, 6, 9, 10, 11, 12			.46
Alter									
3	5	.28	6		1	4	5, 12		.48
4	11	.45				18	10	4, 5, 6, 9, 12, 14, 16, 17	.76
5	10	.54	14, 17, 18		9	6, 11	10, 16	5, 12	.49

Anmerkung. p_i = Itemschwierigkeit, r_{it} = Itemtrennschärfe.

Tabelle 12

Silbenanzahlspezifische mittlere Itemtrennschärfen

	Gesamt		Alter (Ganzjahre)					
			3		4		5	
	p_i	r_{it}	p_i	r_{it}	p_i	r_{it}	p_i	r_{it}
Silben								
2	.86	.43	.50	.47	.88	.86	n/a	n/a
3	.45	.52	.14	.42	.61	.77	.62	.50
4	.38	.52	.14	.60	.39	.77	.52	.69
5	.29	.41	n/a	n/a	.32	.77	.52	.31
6	.16	.29	n/a	n/a	.12	.58	.43	.10

Anmerkung. p_i = Itemschwierigkeit, r_{it} = Itemtrennschärfe, n/a = nicht anwendbar/verfügbar.

Objektivität.

Einfaktorielle ANOVAs zeigten Testleiterinnenunabhängigkeit hinsichtlich KWN-Summenscores, $F(2, 28) = 0.86$, $p = .433$, $\omega^2 = .04$, und Kruskal-Wallis-Tests hinsichtlich KWN-Silbenspannen, $\chi^2(2) = 3.68$, $p = .328$, sowie hinsichtlich Ziffernspannen im *Zahlen Merken*, $\chi^2(2) = 3.28$, $p = .849$. Die Auswertungsobjektivität betreffend erreichten drei Kodiererinnen bei der Transkription und Bewertung von 19.4% der Daten an KWN-Leistungen (ein 3-jähriges Kind, drei 4-jährige und zwei 5-jährige Kinder) zu Itemset 1 mit einer Spannweite von Fleiss' κ von .44 bis 1.00 (exklusive Item 1 wegen Null-Varianzen) genügende bis sehr gute Konkordanzen bzw. mit einem mittleren Fleiss' κ von .69 ($SD = 0.18$) eine durchschnittlich gute Konkordanz bei 31% an fehlerhaften Bewertungen (für itemspezifische Koeffizienten siehe Anhang H).

Reliabilität.

Altersgruppenübergreifende sowie -spezifische Reliabilitätsschätzungen sind Tabelle 13 zu entnehmen. Das selektierte Itemset zeigte altersübergreifend eine gute innere Konsistenz und Split-half-Reliabilität; altersgruppenspezifisch ergaben sich, bei Ausschlüssen von Items mit Nullvarianzen, vergleichbare oder numerisch höhere Reliabilitätskoeffizienten und bei 5-Jährigen die höchste Reliabilität mit einer hohen inneren Konsistenz und Split-half-Reliabilität, bei 3-Jährigen eine akzeptable innere Konsistenz und Split-half-Reliabilität und bei 4-Jährigen eine inakzeptable innere Konsistenz und Split-half-Reliabilität einhergehend mit großen Standardmessfehlern und breiten Konfidenzintervallen. Die Überprüfung der Intracoder-Reliabilität zur Schätzung der Zuverlässigkeit von protokollierten Antworten bei KWN-Aufgaben in der Testsituation ergab mit einem mittleren Cohens κ von .83 ($SD = .11$, Streubreite = .66 – 1.00) eine gute durchschnittliche Konkordanz zwischen Bewertungen der derselben, in Testsituationen protokollierten und zu einem späteren Zeitpunkt von Tonträger-Aufnahmen transkribierten, Antworten.

Validität.

Tabelle 14 enthält Interkorrelationen bei Auspartialisierung des Alters in Monaten zwischen KWN-Summenscores, KWN-Silbenspannen, Ziffernspannen im *Zahlen Merken* sowie Summenscores in Subtests zur Grobmotorik (*Turnen*) und Visuomotorik (*Nachzeichnen*). Hinsichtlich der konvergenten Validität zeigten sich hohe, signifikant positive partielle Korrelationen zwischen KWN-Summenscores und Ziffernspannen ($p = .005$) sowie mittlere, signifikant positive partielle Korrelationen zwischen KWN-Silbenspannen und Ziffernspannen ($p = .039$). Die diskriminante Validität betreffend zeigte sich Unabhängigkeit der KWN-Summenscores und -Silbenspannen von konstruktfernen Grobmotorik-Summenscores (KWN-Summenscore:

$p = .272$, KWN-Silbenspanne: $p = .589$) bzw. der KWN-Summenscores von konstruktfernen Visuomotorik-Summenscores ($p = .084$), die am höchsten mit dem Lebensalter korrelierten und mittlere, signifikant positive Korrelationen mit KWN-Silbenspannen ($p = .015$) aufwiesen. ´

Tabelle 13

Reliabilitätsschätzungen des selektierten Itemset

	Gesamt ^a	Alter (Ganzjahre)		
		3	4	5
Cronbachs α [KI]				
Gesamt-Itemset	.81 [.69, .90]	.61 [-.01, .92]	.49 [.04, .78]	.79 [.45, .96]
Selektion Alter	—	.70 ^b [.11, .94]	.49 ^c [.04, .78]	.80 ^d [.48, .96]
Split half	.85	.75	.60	.95

Anmerkung. ^a $k = 12$ Items, ^b $k = 5$ Items, ^c $k = 11$ Items, ^d $k = 10$ Items.

Tabelle 14

Interkorrelationen zwischen Rohwerten in konstrukt-nahen und -fernen Aufgaben bei Kontrolle des Alters

	1. KWN _{sum} ^a	2. KWN _{span} ^b	3. Zahlen Merken _{span} ^b	4. Turnen ^a	5. Nachzeichnen
1. KWN _{sum} ^a	–	.74***	.51**	-.21	.32
2. KWN _{span} ^b		–	.38*	-.10	.44*
3. Zahlen Merken _{span}			–	.04	.12
4. Turnen ^a				–	.21
5. Nachzeichnen ^a					–
6. Alter (Monate) ^{c, d}	.57*** ^c	.63**** ^d	.60**** ^d	.63*** ^c	.70*** ^c

Anmerkung. Oberhalb der Diagonale: Partielle Korrelationen bei Auspartialisierung des Alters in Monaten, unterhalb der Diagonale: Korrelationen mit Lebensalter in Monaten, sum = Summenscore, span = Gedächtnisspanne.

^apartielle Korrelation nach Pearson, ^bpartielle Rangkorrelation nach Spearman, ^cProdukt-Moment-Korrelation nach Pearson, ^dRangkorrelation nach Spearman.

* $p < .05$, ** $p < .01$, *** $p < .001$.

Zusatzergebnisse: Analyse des Arbeitsverhaltens bei KWN-Aufgaben

Basierend auf Verhaltensbeobachtungen und -beurteilungen zeigten alle Kinder Instruktionsverständnis. Die meisten Kinder zeigten ohne Einsatz eines Teddybären Kooperationsbereitschaft bzw. Kinder, welche das Nachsprechen von Kunstwörtern verweigerten, waren nicht durch den Teddybären zum Antworten zu motivieren. Anfängliche und ca. bis zur Hälfte der Items scheinbare Motivation, Interesse und Aufmerksamkeit der Kinder (z.B. Lachen, aufgabenbezogene Äußerungen) sanken im Verlauf der Durchführung, bei einer Zunahme an Ermüdungserscheinungen und Ablenkbarkeit der Kinder sowie deren Erkundigung nach der Anzahl noch ausstehender Kunstwörter bis zum Testende. Einige Kinder stellten Analogien zu existenten deutschen Wörtern her (z.B. Bantuf [Probeitem] „klingt wie Pantoffel“).

Diskussion

Ziel vorliegender Untersuchung war, durch Konstruktion und Analyse eines Itempools mit Kunstwörtern inklusive multipler Bewertungsmodi, ein entwicklungspsychologisch fundiertes, psychometrischen Standards entsprechendes Itemset und Aufgabendesign zur Erfassung des phonologischen Gedächtnisses im Kindergartenalter mittels neuartigen Materials zu entwickeln und dieses, hinsichtlich der Eignung für ein allgemeines Entwicklungstestverfahren für die entsprechende Zielgruppe, vertrautem Material gegenüberzustellen. Zugehörige Zielsetzungen wurden dabei im Wesentlichen erreicht. Erstens konnte mit der an aktuellen und etablierten KWN-Aufgaben orientierten Konstruktion zweier paralleler Itemsets zu je 20 Kunstwörtern mit zwei bis sechs Silben auf Basis probabilistischer und klassischer psychometrischer Analysen an zwei Teilstichproben ($n_1 = 31$, $n_2 = 27$) 3;1- bis 5;5-jähriger Kinder aus Wiener Kindergärten ein Itemset ausgewählt und bei der zugrundeliegenden Teilstichprobe (n_1) mittels klassischer Methoden eine Itemselektion vorgenommen werden. Zweitens indizierten probabilistische und klassische testtheoretische Analysen, Varianz-, Mittelwerts-, Konkordanz- und Korrelationsanalysen für das 12 Kunstwörter umfassende selektierte Itemset eine altersübergreifende Streuung von Itemtrennschärfen über akzeptable, vorwiegend mittlere Trennschärfenbereiche bei einer Variation von Itemschwierigkeiten zwischen .16 bis .90, eindimensionale Messungen bei Diskriminationsfähigkeit zwischen Kindern mit hohen und niedrigen Fähigkeiten, voraussetzender Durchführungs- und guter Auswertungsobjektivität sowie hoher Reliabilität und Konstruktvalidität. Die Reduktion um 40.0% an Items war mit einem Anstieg der Reliabilität des vollständigen Itemsets von $\alpha = .80$ bzw. $r_{tt} = .81$ auf $\alpha = .81$ bzw. $r_{tt} = .85$ im selektierten Itemset verbunden und lässt für dieses – im Sinne der Ökonomie und Zumutbarkeit für Kindergarten-

kinder – bei kürzerer Durchführungsdauer eine vergleichbare oder minimal höhere Messgenauigkeit zu. KWN-Leistungen erwiesen sich als unabhängig vom Geschlecht der Kinder; hinsichtlich des linearen Alterstrends in Ganzjahresschritten zeigte sich ein großer Anstieg an KWN-Leistungen von 3- auf 4-jährige Kinder. Dasselbe Muster zeigte sich bei Ziffernspannen, mit einem kontinuierlicheren Verlauf im Alterstrend von 3- bis 5-jährigen Kindern. Drittens erwies sich der am einfachsten zu administrierende, ökonomische, konservative Bewertungsmodus hinsichtlich psychometrischer Kriterien als vergleichbar mit aufwändigeren Bewertungsmodi. Konkordanzen zwischen den durch dieselben Testleiterinnen in Testsituationen vorgenommenen Protokollierungen und späteren Transkriptionen von Antworten bei KWN-Aufgaben indizierten für simultan zur Durchführung protokollierte Antworten eine gute Zuverlässigkeit.

Eignung des KWN-Itemsets zur Erfassung des phonologischen Gedächtnisses im Rahmen eines Entwicklungstests

Die theoretisch und empirisch fundierte Operationalisierung und Itemkonstruktion bildet die Grundlage für die Annahme, dass der Itempool bzw. das selektierte Itemset die Fähigkeitsdimension des phonologischen Gedächtnisses im Kindergartenalter reflektiert. In Bezug auf das Multikomponentenmodell zum Arbeitsgedächtnis (Baddeley, 1986, 2000, 2007; Baddeley & Hitch, 1974) und auf entsprechende Studien im Kindergartenalter bildet es damit den phonologischen Kurzzeitspeicher exklusive des artikulatorischen Kontrollprozesses ab (Cowan, 2016; Kiese-Himmel, 2011; Oberauer & Lewandowsky, 2010; Roebers & Zoelch, 2005). Die Inhaltsvalidität des selektierten Itemsets stützt sich zudem auf die mit ansteigender Silbenanzahl ermittelten numerischen Verringerungen der KWN-Leistungen bei signifikantem Sequenz-Längeneffekt von 2- auf 3-silbige Items, welche theoretische Annahmen des Multikomponentenmodells (Baddeley, 1986, 2000, 2007; Baddeley & Hitch, 1974) zur Kapazitätsbegrenzung des phonologischen Speichers und zum zeitlichen Zerfall von Gedächtnisspuren innerhalb von 1.5 bis 2 Sekunden widerspiegeln (Baddeley et al., 1975; Baddeley, 2003; Gathercole et al., 1992) und sich mit empirischen Befunden zu ansteigenden Leistungen von kürzeren auf längere Kunstwörter decken (Boerma et al., 2015; Dispaldro et al., 2013; Polišenská & Kapalková, 2014; Sundström et al., 2014). Der Befund, dass Leistungsanstiege nur von 2- auf 3-silbige Kunstwörter substanziell waren, lässt – analog zu den pro Silbenanzahl unterschiedlichen Itemschwierigkeiten – vermuten, dass die KWN-Leistungen nicht nur von der Itemlänge, sondern auch von anderen Itemmerkmalen abhängig waren. Der empirischen Befundlage zufolge sind Effekte der Wortähnlichkeit (z.B. Suffixe deutscher Wörter), phonologischen bzw. segmentalen Komplexität, phonotaktischen Wahrscheinlichkeit und prosodischen Struktur in Betracht zu ziehen (Jones et al., 2010; Jones et al., 2014; MacRoy-Higgins & Dalton, 2015; Messer

et al. 2010; Gray et al., 2014; Metsala & Chisholm, 2010; Moyle et al., 2014; Smith et al., 2012; Tanida et al., 2015). Da sich im Itemset 2-silbige Kunstwörter von 3- und höher-silbigen insbesondere durch das Fehlen komplexer Silben unterscheiden und eine Variation weiterer Itemmerkmale auch Kunstwörter anderer Silbenanzahlen betraf, ist ein Effekt der phonologischen bzw. segmentalen Komplexität am wahrscheinlichsten.

Die auch bei Konstanthaltung des Lebensalters hohen bzw. mittleren Korrelationen von Ziffernspannen mit KWN-Summscores bzw. mit KWN-Silbenspannen decken sich mit empirischen Befunden (de Abreu et al., 2010; Papadimitrioua & Vlachos, 2014; Passolunghi et al., 2015; Schmid et al., 2008; Spencer et al., 2015; von Goldammer et al., 2010; Vugs et al., 2015) und indizieren, dass Zusammenhänge unabhängig vom Lebensalter auf das den Messinstrumenten zugrundeliegende Konstrukt des phonologischen Gedächtnisses zurückzuführen sind. Die diskriminante Validität des KWN-Itemsets stützt sich auf die ermittelte Unabhängigkeit der KWN-Summscores von Grobmotorik- und Visuomotorik-Leistungen. Die nicht perfekten Zusammenhänge zwischen Leistungen im KWN-Itemset und Ziffernspannen indizieren, dass diese zum Teil unterschiedliche Facetten des phonologischen Gedächtnisses erfassen und lassen annehmen, dass das KWN-Itemset im gemeinsamen Einsatz mit Ziffernfolgen inkrementelle Validität und Nützlichkeit aufweisen kann.

Eine Erklärung für die mögliche Erfassung unterschiedlicher Facetten durch Kunstwörter und Ziffernfolgen bietet die These, dass Kunstwörter als neuartiges Material die Größe des phonetischen Speichers und die Verarbeitungsgenauigkeit abgrenzbar von der mittels vertrauter Einsilber gemessenen Gedächtniskapazität erfassen (Archibald & Griebeling, 2016; Perez et al., 2012; Schuchardt et al., 2011; Mathieu et al., 2016; von Goldammer et al., 2010). Bei Kunstwörtern ist zudem eine höhere Beanspruchung des Mechanismus des phonologischen Gedächtnisses für Reihenfolgeinformationen (Côte et al., 2014; Leclercq & Majerus, 2010; Perez et al., 2012; Majerus & D'Argembeau, 2011) zu vermuten, da sequenzielle Anordnungen von neuartigen Silben und zusätzlich von Phonemen zu repräsentieren und abzurufen sind (Coady & Evans, 2008; Gathercole, 2006; Majerus & Boukebza, 2013). Auf Basis von Modellen zur Sprachverarbeitung und neuro-bildgebenden Studien (Brown et. al., 2000; Burgess & Hitch, 2005; Fiebach et al., 2007; Gupta, 2003; Grossberg and Kazerounian, 2011; Majerus et al., 2010) ist überdies anzunehmen, dass Kunstwörter wegen ihrer Neuartigkeit mit nicht-lexikalischen Verarbeitungsrouten einhergehende Inputverarbeitungsprozesse und damit das übergeordnete Konstrukt phonologischer Verarbeitungsfähigkeiten sowie Outputgenerierungsprozesse im höheren Ausmaß beanspruchen als Ziffernfolgen, deren Nachsprechleistung unter

Rückgriff auf bestehende lexikalisch-semantic Repräsentationen und motorische Programme erfolgt (Acheson & MacDonald, 2009; Coady & Evans, 2008; Mathieu et al., 2016; Kapalková et al., 2013; Jacquemot & Scott, 2006; Jones et al., 2014; Vance, et al., 2005).

Hinsichtlich der Eignung des selektierten Itemsets mit Kunstwörtern für ein allgemeines Entwicklungstestverfahren bildeten in vorliegender Untersuchung KWN-Leistungen einen linearen Alterstrend ab und das Alter in Ganzjahren konnte 50% der Varianz der KWN-Leistungen erklären. Die mit dem Lebensalter ansteigenden phonologischen Gedächtnisleistungen bei Kunstwörtern und Ziffernfolgen können auf Basis der empirischen Befundlage auf neurologische Reifungsprozesse sowie insbesondere auf Anstiege in Informationsverarbeitungsgeschwindigkeit, exekutiven Funktionen und Umfang an Sprachwissen zurückzuführen sein (Brown & Jernigan, 2012; Cowan, 2010; Jones et al., 2014; Messer et al. 2015; Keren-Portnoy et al., 2010; Schmid et al., 2008). Der große Anstieg der KWN-Summenscores um 2.36 Standardabweichungen von 3- auf 4-jährige Kinder bei nicht signifikanten entwicklungsbedingten Leistungsunterschieden zwischen 4- und 5-jährigen Kindern zeigte sich in ähnlichem Muster bei Ziffernspannen. Dies deckt sich mit empirischen Studienergebnissen, welche für das Kindergartenalter die substantiellsten Anstiege im phonologischen Gedächtnis von 3 auf 4 Jahre um ca. eine bis drei Standardabweichungen bei mäßigen Leistungsanstiegen um ca. eine halbe bis eine Standardabweichung zwischen 4- und 5-Jährigen, oder Stagnation der phonologischen Gedächtnisleistungen zwischen dem Alter von 4 bis 6 Jahren indizieren (Dollaghan & Campbell, 2009; Florit et al., 2009; Gray, 2006; Guiberson & Rodríguez, 2013; Kapalková et al., 2013; Polišenská & Kapalková, 2014; Jones et al., 2007; Majerus et al., 2006; Radeborg et al., 2006; Risse & Kiese-Himmel, 2009; Schmid et al., 2008; Sundström et al., 2014; Ulset et al., 2017; Visu-Petra et al., 2014).

Entsprechende Erkenntnisse sind insbesondere vor dem Hintergrund der Sprachentwicklung zu diskutieren, die ab dem Alter von 4 Jahren weitgehend abgeschlossen ist (Bowey, 2006; Vihman et al., 2014; Weinrich & Zehner, 2011). Im Vergleich zu 3-Jährigen verfügen 4-Jährige daher über höhere linguistische Spracherfahrungen und einen größeren Umfang an Sprachwissen, welche ihnen nach der *lexical restructuring*-Hypothese (Metsala & Walley, 1988) Rekonstruktionsprozesse, präzisere phonologische Repräsentationen und stabilere Gedächtnisspuren durch Verknüpfungen zwischen kurzfristigen phonologischen Repräsentationen mit lexikalischen Repräsentationen im Langzeitgedächtnis ermöglichen (Grube et al., 2008; Tanida et al., 2015; Ueno, 2012; Ueno et al., 2014). Gestützt wird diese Annahme durch den Befund, dass Kinder unter 4 Jahren phonologische Informationen mittels *bottom-up*-Prozessen verarbeiten und Kinder erst ab dem Alter von 4 Jahren bei solchen Informationen von *top-down*-Feedback

profitieren können (Ainsworth et al., 2016; Vance et al., 2005). In Einklang mit entsprechenden Annahmen steht auch die bei den Kunstwörtern ermittelte, das 1-PL-Modell (Rasch, 1960) verletzende, knapp signifikante differentielle Itemfunktion hinsichtlich des Alters, welche indiziert, dass bei jüngeren und älteren Kindern die Lösungswahrscheinlichkeit von Items nicht nur von der phonologischen Gedächtniskapazität, sondern auch von anderen Dimensionen abhängt. Nach der These, dass die Entwicklung fluider kognitiver Fähigkeiten innerhalb eines Zyklus von 2- bis 6-jährigen Kindern in zwei Phasen zu unterteilen ist, in eine mit 2- bis 3-Jährigen und in eine mit 4- bis 6-Jährigen (Demetriou et al., 2014; Rose et al. 2008; Schmid et al., 2008), kann zudem vermutet werden, dass 4- und 5-Jährige hinsichtlich des phonologischen Gedächtnisses eine homogenere Gruppe darstellen, die gegenüber 3-Jährigen von kapazitätsunterstützenden Steigerungen in der Inhibition sowie der Fähigkeit zum gezielten Einsatz selektiver Aufmerksamkeit profitieren können (Bertrand & Camos, 2015; Camos & Barrouillet, 2011).

Der bei Kunstwörtern im Vergleich zu Ziffernfolgen deutlicher ausgeprägte Leistungsanstieg von 3- auf 4-Jährige ist durch diverse Ansätze zu erklären. Die bei Kunstwörtern zusätzlich zum phonologischen Gedächtnis beanspruchten Fähigkeitsdimensionen lassen vermuten, dass KWN-Aufgaben an 3-Jährige höhere Anforderungen als Ziffernfolgen stellen. Es ist anzunehmen, dass KWN-Leistungen 3-Jähriger mit Unreife in metaphonologischen Fähigkeiten und – unter Berücksichtigung des Phonemerwerbs (vgl. Fox & Dodd, 2001) – mit Unreife in der Artikulationsmotorik (Kapalková et al., 2013; Weinrich & Zehner, 2011) sowie mit alterstypischen physiologischen phonologischen Prozessen (vgl. Fox-Boyer, 2016) konfundiert sind. Vor dem Hintergrund, dass der phonologische Speicher bei Kindergartenkindern empirisch nicht von Aufmerksamkeitssteuerungs- und -kontrollprozessen abgrenzbar ist und entsprechende Entwicklung mit einer Steigerung phonologischer Gedächtnisleistungen einhergeht (Alloway et al., 2004; Bertrand & Camos, 2015; Camos & Barrouillet, 2011, Gathercole & Pickering, 2000; Demetriou et al., 2014; Pickering & Gathercole, 2001; Rose et al. 2008; Schmid et al., 2008) kann zudem angenommen werden, dass Kunstwörter wegen ihrer höheren Komplexität im Vergleich zu Ziffernfolgen – analog zur Verarbeitungsgenauigkeit – ein höheres Ausmaß an selektiver Aufmerksamkeit und Inhibitionsprozessen erfordern und niedrigere Leistungen 3-Jähriger infolge entsprechender Defizite entstehen. In Bezug auf die Erfassung unterschiedlicher Facetten des phonologischen Gedächtnisses durch Ziffernfolgen und Kunstwörter besteht eine weitere Erklärung für die Alterstrends in der Annahme, dass sich die Gedächtniskapazität für Sequenzen an vertrauten Informationen kontinuierlicher entwickelt und ab einem Alter von 4 Jahren ein wesentliches Entwicklungswachstum im Umfang des phonetischen Speichers und in der Verarbeitungsgenauigkeit erfolgt.

Die bei Kunstwörtern im Vergleich zu Ziffernfolgen höheren Anforderungen dienen auch als Erklärung für die in vorliegender Untersuchung ermittelten, vom Alter unabhängigen, insgesamt höheren Verweigerungsraten bei den KWN-Aufgaben, welche sich numerisch mit der unteren Grenze an Verweigerungsraten in der Literatur decken (Polišenská & Kapalková, 2014; Chiat & Roy, 2007; Roy & Chiat, 2004; Seeff-Gabriel et al., 2008) und in Übereinstimmung mit Annahmen zu Fremdheitseffekten bei neuartigem Testmaterial und Thesen zu Verweigerungen infolge von Fähigkeitsdefiziten, niedrigen Aufmerksamkeitsspannen und Ängstlichkeit stehen (Dohmen, Chiat, & Roy, 2013; Hoff et al., 2008; Mathieu et al., 2016). Es ist auszuschließen, dass die auf das Itemset 2 beschränkte Verweigerung bei uneingeschränkter Kooperationsbereitschaft bei Itemset 1 ein Effekt des Materials war, da entsprechende Kinder bereits das Nachsprechen des für beide Itemsets einheitlichen Probeitems verweigert hatten.

Gegenüber dem Nachteil möglicher höherer Verweigerungsraten besteht der Vorteil der höheren Anforderungen bei Kunstwörtern durch Beanspruchung von neurokognitiven Kernprozessen der Sprachinputverarbeitung sowie Outputgenerierungsprozessen (Casslerly & Pisoni, 2013; Coady & Evans, 2008; Claessen et al., 2013; Jones et al., 2014; Jones & Whiterstone, 2011; Moyle et al., 2014) in der Schaffung einer Kongruenz zur phonologischen Komponente beim Erlernen neuer Wörter (Baddeley et al., 1998). Auf Basis dessen ist anzunehmen, dass die mit Kunstwörtern verbundenen höheren Anforderungen dafür verantwortlich sind, dass sich KWN-Leistungen als informativste Prädiktoren für Wortschatzerwerb und -erweiterung erweisen (Law et al., 2016; Gathercole, 2006; Leclercq & Majerus, 2010) und KWN-Aufgaben als größte diagnostische Schlüsselmarker bei sensibler und spezifischer Identifizierung von SSES gelten (Bishop et al., 1996; Deevy et al., 2010; Dispaldro et al., 2013; Girbau & Schwartz, 2007; Henry et al., 2012; Montgomery et al., 2010; Kapalková et al., 2013; Petrucci et al., 2012; Reichenbach et al., 2016; Shahm Mahmood et al., 2016; Thordardottir et al., 2011) ebenso wie als konsistente, unabhängige Prädiktoren für Lese- und Rechtschreibfähigkeiten in der Schuleingangsphase (Binamé & Poncelet, 2016; Carroll et al., 2014; Catts et al., 2014; Papadimitriou & Vlachos, 2014; von Goldammer et al., 2010).

Praktische Implikationen

Trotz der weitgehend symmetrisch verteilten KWN-Summenscores innerhalb von Altersgruppen und der Annahme, dass diese das Potenzial haben, Differenzierungen zwischen durchschnittlichen und unterdurchschnittlichen bzw. überdurchschnittlichen Leistungen zu ermöglichen, sind auf Basis vorliegender Untersuchungsergebnisse zum Alterstrend, Itemkennwerten und zur ermittelten differentiellen Itemfunktion – analog zur empirischen Befundlage (Chiat,

2015; Graf Estes et al., 2007) – Bodeneffekte bei 3-Jährigen und Deckeneffekte bei 4- bzw. 5-jährigen Kindern bei 2-silbigen Items zu erwarten. Einer Begrenzung der KWN-Aufgaben auf den Altersbereich 4- bis 5-Jähriger wegen zu hoher Itemschwierigkeiten für 3-Jährige stehen die zentrale Bedeutsamkeit einer differenzierten Diagnostik des phonologischen Gedächtnisses sowie einer frühzeitigen, insbesondere durch Kunstwörter mit hoher bis exzellenter Sensitivität ermöglichten, Identifikation von Sprachentwicklungsstörungen und von Risiken für schulische Leistungsschwächen in einer für Umwelteinflüsse sensitiven Phase gegenüber (Blackwell et al., 2009; Drigas et al., 2015; Renner et al., 2015; Roebers & Zoelch, 2005; White-Schwoch et al., 2015; von Goldammer et al., 2010). Zudem indizieren die in entsprechender Altersgruppe geschätzten mittleren Itemtrennschärfen 2-silbiger und einfacherer 3- und 4-silbiger Items sowie die akzeptable Reliabilität bei fünf Items prinzipiell Eignung ausgewählter Kunstwörter für 3-Jährige. Eine praktische Implikation der Ergebnisse vorliegender Untersuchung besteht im Erfordernis altersgruppenspezifischer und silbenanzahlabhängiger Start- und Enditems bei altersübergreifenden 3- und 4-silbigen *linking-items*. Für 3-Jährige ist ein Start bei 2-silbigen und Ende bei maximal 4-silbigen Items indiziert, für 4-Jährige ein Start bei 2-silbigen Items und Ende beim 6-silbigen Item und bei 5-Jährigen ein Start bei 3-silbigen Items und Ende beim 6-silbigen Item, mit einem Einsatz desselben 2-silbigen Probeitems für alle Altersgruppen.

In Hinblick auf den Einsatz im Rahmen eines Entwicklungstests und dem damit häufigen Anwendungsbereich bei Kindern mit Entwicklungsauffälligkeiten sind für 3-Jährige zudem zusätzlich die Inklusion weiterer 2- und 3-silbiger Items (z.B. aus Itemset 2) oder die Vereinfachung komplexer Silben von 3- und 4-silbigen Items anzudenken, um die niedrigen phonologischen Verarbeitungskapazitäten 3-Jähriger nicht zu überfordern, Frustration und Motivationsverluste vorzubeugen und Differenzierungen im unteren Leistungsbereich zu ermöglichen; eine Erweiterung des vorzugebenden Itemsets ist dabei auf ein Minimum zusätzlicher Items zu beschränken, um mit dem Ziel einer geringfügigen Erhöhung der Reliabilität Motivationsverluste und Ermüdungserscheinungen zu vermeiden. Zur Möglichkeit der Differenzierung im oberen Leistungsbereich bei 5-Jährigen ist, wegen niedrigerer Itemtrennschärfen bei identen Itemschwierigkeiten mit Kunstwörtern einer niedrigeren Silbenlänge, eine komplexere Gestaltung von 4- bis 6-silbigen Kunstwörtern durch Inklusion weiterer Konsonantencluster indiziert. Trotz der geschätzten inakzeptablen Reliabilität bei 4-Jährigen sind wegen hoher Itemtrennschärfen in dieser Altersgruppe keine weiteren Ausschlüsse von Items indiziert. Die inakzeptable Zuverlässigkeit bei 4-Jährigen, die hinsichtlich der KWN-Leistungen die heterogenste Gruppe darstellten, könnte einen Effekt der Stichprobe darstellen oder auf in dieser Alters-

gruppe stattfindende Ausdifferenzierungen des phonologischen Gedächtnisses und zu unterschiedlichen Zeitpunkten erfolgenden Vollendungen der Sprachentwicklung zurückzuführen sein. Daher ist insbesondere bei 4-Jährigen eine Ermittlung von Reliabilitäten in Halbjahresschritten indiziert. Hinsichtlich Modifikationen der Items sind zudem die Reduktionssilben zu verändern, da bei fast allen Kindern – analog zum Deutschen – eine Änderung der Laute (/er/ zu [-a]) oder eine Tilgung der Konsonanten am Ende entsprechender Silben (/el/ zu [-l]) erfolgte, einhergehend mit einer Verringerung der Länge bzw. Aussprechdauer der Kunstwörter. Trotz der Möglichkeit einer beliebigen Vorgabereihenfolge der Kunstwörter auf Basis von Ergebnissen probabilistischer Analysen, ist auf Basis der Itemschwierigkeiten innerhalb von Kunstwörtern einer Silbenanzahl eine geringfügige Modifikation der Vorgabereihenfolge durch Anordnung vom einfachsten bis zum schwierigsten Item innerhalb einer Silbenanzahl anzudenken. Eine Umsetzung vorgeschlagener Modifikationen der Items sind Anhang I zu entnehmen. Hinsichtlich der Verrechnung können KWN-Summenscores auf Basis von Rasch-Analysen als suffiziente Statistik erachtet werden und sind in der praktischen Anwendung auch wegen deren symmetrischen Verteilungen der Mittelwerte und größeren Varianzen in Altersgruppen gegenüber Silbenspannen vorzuziehen.

Aufgrund der Abhängigkeit der KWN-Leistungen von der sprachlichen Entwicklung erfordern Ergebnisinterpretationen von KWN-Leistungen und darauf aufbauende Indikationsstellungen in der Praxis die Berücksichtigung sprachlicher Fähigkeiten der Kinder und Informationen zur Sprachentwicklung und zu psychologischen, audiologischen und logopädischen Untersuchungen. Zur Interpretation niedriger Leistungen bei KWN-Aufgaben sind TestanwenderInnen u.a. – analog zur bereitgestellten Tabelle in vorliegender Untersuchung – Angaben zu entwicklungs-spezifischen Fehlern zur Verfügung zu stellen. Auf Basis der empirischen Befundlage ist bei Kindern ab 4 Jahren nicht nur bei Ziffernfolgen, sondern auch bei KWN-Aufgaben eine mögliche Konfundierung des phonologischen Kurzzeitgedächtnisses mit dem Langzeitgedächtnis zu berücksichtigen (Jones et al., 2014; Metsala & Chisholm, 2010; Moyle et al., 2014; Tanida et al., 2015).

Weitere Praxisimplikationen umfassen Aspekte der Objektivität, Reliabilität und Ökonomie. Die varianzanalytisch ermittelten von Testleiterinnen unabhängigen KWN-Leistungen sind auf die Bereitstellung von wörtlichen Instruktionen, Angaben zu Rahmenbedingungen, erforderlichen Materialien, zum Umgang mit etwaig auftretenden Herausforderungen in den Testsituationen sowie von Tonaufzeichnungen der vorgesprochenen Kunstwörter zur Standardisierung der Akzentuierung zurückzuführen. Für den zukünftigen Einsatz der Kunstwörter sind zur Er-

reichung einer entsprechenden Durchführungsobjektivität analoge Standardisierungen erforderlich und für TestleiterInnen das Einüben der Akzentuierung mittels Tonaufzeichnungen für eine standardisierte Vorgabe der Kunstwörter essentiell. Hinsichtlich der Auswertungsobjektivität bei KWN-Aufgaben ergab sich bei der lautgetreuen Transkription und Bewertung der Kunstwörter, trotz der als durchschnittlich gut einzustufenden Übereinstimmung zwischen drei Kodiererinnen, fast ein Drittel an Fehleinschätzungen, welche bei dem eindeutig anzuwendenden konservativen Bewertungsmodus (1 = idente Wiedergabe, 0 = jegliche Abweichungen) auf einen Ermessensspielraum bei der lautgetreuen Transkription zurückzuführen ist. Diesem Ermessensspielraum konnte in vorliegender Untersuchung trotz bereitgestellter schriftlicher und persönlich instruierter Auswertungsrichtlinien sowie konkreter Auswertungsbeispiele nicht entgegengewirkt werden. Es ist anzunehmen, dass – analog zum Erfordernis einer guten phonologischen Verarbeitung von Kunstwörtern bei Testkindern (Bowey, 2006) – die Verarbeitung der Antworten genaue Phonemdiskriminationen der TestleiterInnen erfordern und z.T. vom subjektivem Höreindruck abhängen.

Während bei Ziffernfolgen Auswertungsobjektivität aufgrund der, mit der Bekanntheit des Materials verbundenen, phonologischen Klarheit der Antworten anzunehmen ist, besteht zur Optimierung der Auswertungsobjektivität von KWN-Aufgaben für die Praxis das Erfordernis sorgfältiger Trainings von TestleiterInnen zur Schulung der Phonemdifferenzierung und -identifizierung und lautgetreuen Protokollierung von Antworten. Zur Ermöglichung eigenverantwortlicher Trainings von TestleiterInnen sind Tonträgeraufzeichnungen mit Beispielen von u.a. undeutlichen und schwierig einzuschätzenden Antworten sowie entsprechende Lösungen zusätzlich zu schriftlichen Auswertungsbeispielen bereitzustellen. Auf Basis der durchschnittlich als ausgezeichnet einzustufenden Zuverlässigkeit von in Testsituationen protokollierten Antworten erscheint in der Anwendung ein Verzicht auf Transkriptionen angemessen. Die 17% an Fehleinschätzungen zwischen protokollierten und transkribierten Kodierungen sind auf Transkriptionsfehler, basierend auf unterschiedlicher Qualität der Tonträgeraufzeichnungen, oder auf unzureichende Ressourcen zur sorgfältigen Protokollierung in Testsituationen zurückzuführen. Für die Praxis sind daher Tonaufzeichnungen mit guter Aufzeichnungsqualität zu empfehlen, um – bei Überforderung in einer Testsituation oder Unsicherheiten in der akustischen Wahrnehmung – in der Testsituation vorgenommene Protokollierungen und Bewertungen von Items durch wiederholtes Anhören der Antworten oder durch das Einholen einer zweiten Meinung absichern zu können.

Zur Minimierung von Verweigerungsraten infolge hoher Anforderungen und etwaiger Fremdheitseffekte bei Kunstwörtern und zur Motivierung zur Mitarbeit bei Maximierung von Aufmerksamkeit, Konzentration und Ausdauer, welche für eine reliable und valide Diagnostik des phonologischen Gedächtnisses zentral sind (Renner & Mickley, 2015; Schneider & Sodian, 1990; Radeborg et al., 2006), ist für die praktische Anwendung der KWN-Aufgaben eine spielerische Gestaltung in Form einer Rahmengeschichte – zum Ausschluss von Interferenzeffekten (Brady & Goodman, 2014) – ohne Testmaterialien anzudenken.

Die vorliegende Untersuchung lieferte ein neu entwickeltes Messinstrument mit deutschsprachlich orientierten Kunstwörtern, das eine objektive, reliable und valide Erfassung des phonologischen Gedächtnisses bei Kindergartenkindern unter Annäherung an Ökonomie und Zumutbarkeit für die Zielpopulation ermöglichen kann. Vor dem Hintergrund der Ergebnisse vorliegender Untersuchung bei systematischem Vergleich zwischen Leistungen bei Kunstwörtern und Ziffernfolgen sowie der theoretischen und empirischen Befundlage ist davon auszugehen, dass sich das KWN-Itemset für ein allgemeines Entwicklungstestverfahren eignet, im kombinierten Einsatz mit Ziffernfolgen wertvolle Informationen zu Fähigkeitsaspekten des phonologischen Gedächtnisses liefert, Nützlichkeit aufweist und eine differenzierte Diagnostik des phonologischen Gedächtnisses ermöglicht. Bei kombiniertem Einsatz des KWN-Itemsets mit Ziffernfolgen können entsprechende Testergebnisse in Zusammenhang gesetzt und gemeinsam interpretiert werden; bei diskrepanten Ergebnissen kann das KWN-Itemset als prozessdiagnostisches Maß (Casserly & Pisoni, 2013; Claessen et al., 2013) Hinweise auf akustische und sprachliche Fähigkeiten, wie Hörvermögen, metaphonologische Fähigkeiten, Sprachwissen und artikulationsmotorische Fähigkeiten liefern sowie bei qualitativer Auswertung Hinweise auf Sprachfehler und förderdiagnostische Aussagen ermöglichen. In Abhängigkeit von spezifischen psychologisch-diagnostischen Fragestellungen kann die Vorgabe zudem auf eine der beiden Aufgabentypen beschränkt werden und bei fehlender Kooperationsbereitschaft von Kindern bei einer der Aufgaben, auf Basis der jeweils anderen, Aussagen zum phonologischen Gedächtnis getroffen werden. Ein weiterer Beitrag vorliegender Untersuchung für die praktische Anwendung besteht im Vergleich von Bewertungsmodi, u.a. mit der Entwicklung altersgruppenspezifischer Bewertungsregeln in Anlehnung an entwicklungstypische phonologische Prozesse, und in der Ermöglichung einer effizienten Messung. Bei Verzicht auf Transkriptionen der Antworten bei KWN-Aufgaben und bei Anwendung des ökonomischen konservativen Bewertungsmodus sind reliable und valide Ergebnisse zu erwarten.

Limitationen

Die kleinen und hinsichtlich des Alters ungleich großen Teilstichproben, die Beschränkung der Altersgruppe 5-Jähriger auf 5;0-5;5-jährige Kinder sowie die eingeschränkte Stichprobenrepräsentativität aufgrund der, in Bezug auf die österreichische Grundgesamtheit, unterrepräsentierten niedrigsten sozialen Schichtzugehörigkeit der Eltern sind mit eingeschränkter Aussagekraft der Ergebnisse und Schlussfolgerungen verbunden. Die mit den Teilstichprobengrößen einhergehende niedrige Power für die Untersuchung bei erwarteten vorwiegend großen Effekten kann den nicht signifikanten Anstiegen in phonologischen Gedächtnisleistungen bei 4- und 5-jährigen Kindern zugrunde liegen, ebenso wie der Ausschluss der kleinen, nicht repräsentativen Substichprobe 5;6- bis 5;11-jährigen Kinder. Die Ergebnisse und ungünstigeren Kennwerte für Itemset 2 können durch die in Aufgaben zum phonologischen Gedächtnis sich als leistungsschwach erwiesenen 4-Jährigen in Teilstichprobe 2 verzerrt sein und als Effekt der Stichprobe statt als Effekt des Testmaterials interpretiert werden. Da in vorliegender Untersuchung etwaige diagnostizierte Sprachfehler und Sprachentwicklungsstörungen nicht erhoben wurden, ist in Betracht zu ziehen, dass die vorliegende Stichprobe Kinder mit entsprechenden Beeinträchtigungen umfasste und Ergebnisse verzerrt bzw. für die Normalpopulation nicht repräsentativ sind. Einhergehend mit der guten statt exzellenten Auswertungsobjektivität kann überdies nicht mit Sicherheit angenommen werden, dass die kodierten Ergebnisse beim KWN-Itemset wahre Werte widerspiegeln. Analog dazu stellt die fehlende Standardisierung der Tonträger bzw. der Transkription entsprechender Aufzeichnungen eine Limitation dar. Die Auswertungsobjektivität ist zudem an der, auf erfolgten Transkriptionen und Bewertungen basierenden, geringen Stichprobengröße von $n = 6$ Kinder zu relativieren.

Die ausschließliche Untersuchung von Kindern mit Deutschkenntnissen und die an der deutschen Sprache orientierte Konstruktion der Kunstwörter stellt eine Einschränkung der Generalisierbarkeit vorliegender Untersuchungsergebnisse und Schlussfolgerungen für einen Einsatz des Messinstruments außerhalb des deutschsprachigen Raumes und für einen Einsatz bei, im deutschsprachigen Raum lebenden, Kindern mit Migrationshintergrund und Deutsch als Fremdsprache dar, insbesondere für Kinder mit nicht-indoeuropäischer Erstsprache (z.B. Türkisch), für die höhere Anforderungen gegenüber Kindern mit Deutsch als Erstsprache und bilingual deutsch aufwachsenden Kindern zu erwarten sind (Chiat & Polišenká, 2016; Messer et al., 2010; Schölar & Schäfer, 2004; Summers et al., 2010; Thordardottir & Brandeker, 2013; Wild & Fleck, 2013). Die Spezifika der, in der Anzahl an Konsonantenclustern variierenden, Kunstwörter limitiert zudem die Generalisierbarkeit der Reliabilität und diagnostischen Aussagekraft auf spezifische klinische Gruppen, wie Kinder mit Defiziten in der Sprachentwicklung

und Artikulationsfähigkeit, deren niedrige Leistungen bei Kunstwörtern nicht bzw. nicht ausschließlich niedrige Gedächtnisfähigkeiten widerspiegeln. Für die praktische Anwendung ist daher, zusätzlich zur fraglichen Eindimensionalität und Fairness über Altersbereiche hinweg aufgrund der höheren artikulatorischen Anforderungen für jüngere Kinder, eine mögliche Beeinträchtigung der Eindimensionalität und Fairness des KWN-Itemsets hinsichtlich der Erstsprache und Sprachentwicklung von Kindern zu beachten.

Eine mögliche Einschränkung der ermittelten psychometrischen Qualität hinsichtlich der Generalisierbarkeit auf andere Stichproben bzw. die Zielpopulation ist auch unter Berücksichtigung der Stichprobenabhängigkeit der ermittelten Itemkennwerte, Reliabilitätsschätzungen und Validitätskoeffizienten in Betracht zu ziehen. Die bei Ziffern- und Silbenspannen mit fehlenden Normalverteilungen verbundenen nicht-parametrischen statistischen Analysen stellen eine weitere methodische Limitation der Aussagekraft der Ergebnisse dar.

Auf Basis qualitativ erfasster sinkender Motivation und Aufmerksamkeit der Kinder im Verlauf der Vorgabe der Kunstwörter sind bei Kunstwörtern höherer Silbenanzahl niedrigere Leistungen u.a. auch auf die Position in der Vorgabereihenfolge und entsprechende Motivations- und Aufmerksamkeitsverluste zurückzuführen. Eine Limitation stellt auch die Streubreite des Zeitintervalls von einem bis 15 Tagen zwischen ersten und zweiten Testterminen dar, die keine eindeutigen Aussagen zulässt, ob bei Ziffernfolgen und Kunstwörtern Übungs- und Lerneffekte auftreten können bzw. die für einige Kinder mit einer Bevorzugung oder Benachteiligung verbunden gewesen sein könnte. Trotz der ermittelten Unabhängigkeit der Leistungen vom Testtermin bleibt zudem offen, ob Übungs- oder Lerneffekte bei einer Vorgabe von KWN-Aufgaben und von Ziffernfolgen am selben Tag im Rahmen eines Entwicklungstests stattfinden können. Obwohl KWN-Aufgaben eine Situation im Erwerb neuer Wörter simulieren (Chiat & Roy, 2007; Radeborg et al., 2006), ist in Betracht zu ziehen, dass diese die Fähigkeit eines Kindes in solchen Situationen unterschätzen, da ein Kontext, der für ein dargebotenes Wort in natürlichen Situationen gegeben ist (z.B. Referenz zu einem Objekt oder Ereignis), bei KWN-Aufgaben fehlt (Vihman et al., 2014).

Ausblick

Nachfolgende Forschungsarbeiten erfordern größere Stichproben unter einem ausgeglichenen Verhältnis von Altersgruppen in Ganz- sowie in Halbjahresschritten, eine Inklusion 5;6- bis 5;11-jähriger Kinder sowie Repräsentativität hinsichtlich der sozialen Schichtzugehörigkeit der Eltern. Das selektierte Itemset ist mit altersspezifischen Start- und Enditems sowie *linking items*

im Rahmen parametrischer Rasch-Analysen, bei Dichotomisierung sowie Dreiteilung hinsichtlich des Alters und in Bezug auf die soziale Schichtzugehörigkeit zu analysieren. Zur Vorbeugung von Frustration sind in zukünftigen Untersuchungen geeignete Cutoffs für Abbruchkriterien für Kinder mit niedrigen Fähigkeiten anzudenken. Schätzungen der Itemkennwerte, Reliabilitäts- und Validitätsschätzungen sind auf Basis größerer Stichproben nach Altersgruppen getrennt vorzunehmen, um Überschätzungen zu vermeiden. In zukünftigen Forschungsarbeiten sind zudem psychometrische Analysen bei Kindern mit Entwicklungsrückständen und -störungen durchzuführen, um zu bestimmen, ob auch in der vornehmlichen Anwendungsgruppe von Entwicklungstests objektive, reliable und valide Ergebnisse erzielt werden können.

Hinsichtlich des Testmaterials ist die Aussprechdauer der Kunstwörter auf Basis von Aufzeichnungen mehrerer TestleiterInnen mittels eines computerisierten Sprachanalyseprogramms (*computerized speech lab*) zu messen und für Items pro Silbenanzahl eventuell durch Exklusion oder Inklusion von Konsonanten(-clustern) zu standardisieren, um mit der Silbenanzahl bedeutsam ansteigende Anforderungen an das phonologische Gedächtnis erfassen zu können. Die Wortähnlichkeit der Kunstwörter ist durch LaiInnenratings, vorzugsweise durch die Anwendungsgruppe (Kindergartenkinder), schätzen zu lassen und im Bedarf durch Modifikationen an Items auszubalancieren. Bei Aufzeichnungen der Antworten bei den KWN-Aufgaben sind Tonträger, deren Anbringung in der Testsituation (z.B. Abstand und Höhe zwischen Kind und TestleiterIn) sowie Auswertungen der Tonaufzeichnungen (z.B. über spezifizierte Geräte und Kopfhörer) zu standardisieren.

Zur Sicherung der Interpretationsobjektivität ist die Erhebung von Normierungsdaten erforderlich, um Normwerte und kritische Differenzen für bedeutsame Unterschiede zwischen Leistungen bei Ziffernfolgen und beim KWN-Itemset sowie weitere Interpretationsschritte hinsichtlich Förderungsmöglichkeiten oder dem Erfordernis nachfolgender Untersuchungen, z.B. zur Identifikation einer SSSES, zu ermitteln und für zukünftige Anwendungen zur Verfügung zu stellen. Aufgrund der Abhängigkeit von KWN-Leistungen von akustischen und sprachlichen Zugangsfähigkeiten sind in nachfolgenden Studien mittels Eltern- oder KindergartenpädagogInnenfragebögen Vorinformationen hinsichtlich medizinisch abgeklärten Hörvermögens, Innenohrschwerhörigkeiten, auditiven Sinnesstörungen, der Sprachentwicklung (z.B. erstes Wort, Satzumfang), einer diagnostizierten SSSES, Sprechfehlern und logopädischen Behandlungen zu erheben, um von entsprechenden Beeinträchtigungen betroffene Kinder a priori für eine Normstichprobe auszuschließen. Hörscreenings mit realen deutschen Wörtern, die spät erworbene Konsonanten und Konsonantencluster umfassen, sowie die Aufzeichnung spontaner Sprach-

samples können demselben Zweck dienen. Entsprechendes Vorgehen erlaubt zudem eine Kontrolle um allgemeine Sprachfähigkeiten der Kinder und dadurch eine reine Erfassung des phonologischen Gedächtnisses. Um eine faire Erfassung des phonologischen Gedächtnisses mittels des KWN-Itemsets zu ermöglichen, sind für Kinder mit diagnostizierten Sprachentwicklungsstörungen bzw. Sprechfehlern spezifische Auswertungsregeln und separate Normwerte erforderlich.

Für die Anwendung im Rahmen des WET (Kastner-Koller & Deimann, 2012) ist in Folgeuntersuchungen bei einer Durchführung aller Subtests die faktorielle Validität zu bestimmen und in Hauptkomponentenanalysen zu ermitteln, ob das KWN-Itemset höher auf einem Faktor zusammen mit dem Subtest *Zahlen Merken* oder Subtests zur Sprache lädt. Zur Bestimmung der differenziellen Validität des KWN-Itemsets sind Mittelwerts- oder Varianz- sowie ROC-Kurven-Analysen zur Ermittlung von Sensitivität und Spezifität bei normal entwickelten Kindern sowie Kindern mit diversen SSES und/oder Sprachentwicklungsverzögerungen indiziert; entsprechende Analysen sind auch beim Subtest *Zahlen Merken* durchzuführen und die Sensitivität und Spezifität zur Identifizierung von SSES beider Aufgabentypen einander gegenüberzustellen. Zur Analyse auf Fairness hinsichtlich der Erstsprache sind Mittelwerts- oder Varianzanalysen bei deutsch- und fremdsprachigen und bilingual aufwachsenden Kindern vorzunehmen. Zudem sind in zukünftigen Forschungen Untersuchungen an klinischen Gruppen mit Entwicklungsstörungen, wie Down Syndrom und Autismus-Spektrum-Störungen, durchzuführen. Zur Überprüfung auf Übereinstimmungsvalidität sind Zusammenhänge der KWN-Leistungen mit Außenkriterien über Elternfragebögen zu ermitteln, z.B. mit Fragen zu den von Eltern eingeschätzten Fähigkeiten eines Kindes, erstmalig dargebotene Wörter nachzusprechen, sich Aufträge, Inhalte von Hörkassetten oder vom gemeinsamen Bilderbuchlesen zu merken, sowie bei akustisch basierten Spielen wie „Ich packe meinen Rucksack“. Das KWN-Itemset erfordert zudem eine Validierung an etablierten Messinstrumenten mit Kunstwörtern.

Zur Bestimmung der prognostischen Validität des KWN-Itemsets bedarf es Längsschnittstudien. Es sind Zusammenhänge von KWN-Leistungen zu unterschiedlichen Messzeitpunkten sowie mit dem Wortschatz, der Sprachentwicklung in der Erst- und in Fremdsprachen zu einem späteren Erhebungszeitpunkt bzw. mit Lese- und Rechtschreibleistungen in der Schuleingangsphase zu ermitteln. Mittels multipler Regressionen ist unter Kontrolle von Wortschatz und metaphonologischen Fähigkeiten der einzigartige Beitrag des KWN-Itemsets zur Vorhersage entsprechender Fähigkeiten im Vergleich zum *Zahlen Merken* zu bestimmen. In einem randomisierten, kontrollierten Studiendesign sind zudem Interventionseffekte von Maßnahmen zur Förderung des phonologischen Gedächtnisses auf KWN-Leistungen zu ermitteln.

Conclusio

Das entwickelte KWN-Itemset kann als rasch-homogenes, objektives, reliables, valides, effizientes und ökonomisches Messinstrument zur Erfassung des phonologischen Gedächtnisses erachtet werden, das interindividuelle und entwicklungsbedingte Unterschiede im phonologischen Gedächtnis bzw. dessen Wechselwirkung mit Sprachfähigkeiten sensitiv erfassen kann. Die Abbildung eines kontinuierlicheren altersabhängigen Anstiegs der KWN-Leistungen ist durch Einführung altersgruppen- und silbenanzahlspezifischer Start- und Enditems bei phonologischer Vereinfachung von Items für 3-Jährige und komplexerer Gestaltung für 5-Jährige zu erwarten. Für den kombinierten Einsatz mit Ziffernfolgen im Rahmen eines allgemeinen Entwicklungstestverfahrens sind wegen der Neuartigkeit des Materials inkrementelle Validität des KWN-Itemsets zu erwarten. Höhere Anforderungen bei Kunstwörtern im Vergleich zu Ziffernfolgen betreffen Kinder ebenso wie TestleiterInnen (bei der Vorgabe und Auswertung) und sind zugunsten des Informationsgewinns zu tolerieren. Es ist anzunehmen, dass Ziffernfolgen Hinweise auf die Fähigkeit liefern, sich eine Reihe an vertrauten Informationseinheiten kurzfristig zu merken, z.B. aus mehreren Teilen bestehende Aufträge sowie Aufgaben mit mehreren Schritten, hingegen die Anzahl an korrekt reproduzierten Kunstwörtern Hinweise auf die Basiskompetenz liefert, sich neuartige phonologische Sequenzen wie neue Wörter in einer Sprache nach einmaliger Darbietung kurzfristig zu merken um diese in der Folge langfristig zu erwerben (Baddeley et al., 1998; Côte et al., 2014; Hu & Schuele 2005; Nora et al., 2012). Es ist zu erwarten, dass mit dem KWN-Itemset Entwicklungsrückstände im phonologischen Gedächtnis und Defizite in spezifischen, für die Sprachentwicklung relevanten Fähigkeiten identifiziert werden können. Damit soll die Voraussetzung für eine Einleitung von Fördermaßnahmen geschaffen werden, um gedächtniskapazitätsabhängige schulische und sprachliche Leistungsschwächen vorzubeugen, Kindern optimale Voraussetzungen für die Erweiterung des Wortschatzes, zur Bewältigung von sprachlichen und kognitiven Anforderungen in Schule und Alltag, den Erwerb von Kulturtechniken, kommunikative Kompetenzen und Teilhabe am gesellschaftlichen Leben zu ermöglichen (Blackwell, Cepeda, & Munakata, 2009; Drigas, Kokkalia, & Lytras, 2015; von Goldammer et al., 2010).

Literaturverzeichnis

- Acheson, D. J. & MacDonald, M. C. (2009). Verbal working memory and language production: Common approaches to the serial ordering of verbal information. *Psychological bulletin*, 135(1), 50-68. doi: 10.1037/a0014411
- Adams, A. M. & Gathercole, S. E. (1995). Phonological working memory and speech production in preschool children. *Journal of Speech, Language, and Hearing Research*, 38(2), 403-414. doi:10.1044/jshr.3802.403
- Ainsworth, S., Welbourne, S., & Hesketh, A. (2016). Lexical restructuring in preliterate children: Evidence from novel measures of phonological representation. *Applied Psycholinguistics*, 37(4), 997-1023. doi: 10.1017/S0142716415000338
- Alloway, T. P. & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of experimental child psychology*, 106(1), 20-29. doi: 10.1016/j.jecp.2009.11.003
- Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and visuospatial short-term and working memory in children: Are they separable?. *Child development*, 77(6), 1698-1716. doi: 10.1111/j.1467-8624.2006.00968.x
- Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child development*, 80(2), 606-621. doi: 10.1111/j.1467-8624.2009.01282.x
- Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A. M. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of experimental child psychology*, 87(2), 85-106. doi: 10.1016/j.jecp.2003.10.002
- Altmann, H. & Kemmerling, S. (2005). *Wortbildung fürs Examen* (2. überarb. Aufl.) Göttingen: Vandenhoeck & Ruprecht.
- Andersen, E. B. (1973). A goodness of fit test for the Rasch model. *Psychometrika*, 38, 123-140. doi: 10.1007/BF02291180
- Archibald, L. M. D. (2008). The promise of nonword repetition as a clinical tool. *Canadian Journal of Speech-Language Pathology and Audiology*, 32(1), 21-28. Abgerufen von http://cjslpa.ca/files/2008_CJSLPA_Vol_32/CJSLPA_2008_Vol_32_No_01_Spring.pdf#page=22

- Archibald, L. M. D. & Gathercole, S. E. (2007). Nonword repetition in specific language impairment: More than a phonological short-term memory deficit. *Psychonomic bulletin & review*, 14(5), 919-924. doi: 10.3758/BF03194122
- Archibald, L. M. D. & Griebeling, K. H. (2016). Rethinking the connection between working memory and language impairment. *International journal of language & communication disorders*, 51(3), 252-264. doi: 10.1111/1460-6984.12202
- Armon-Lotem, S. & Meir, N. (2016). Diagnostic accuracy of repetition tasks for the identification of specific language impairment (SLI) in bilingual children: evidence from Russian and Hebrew. *International journal of language & communication disorders*, 51(6), 715-731. doi: 10.1111/1460-6984.12242
- Baddeley, A. (1998). Recent developments in working memory. *Current opinion in neurobiology*, 8(2), 234-238. doi: 10.1016/S0959-4388(98)80145-1
- Baddeley, A. (2000). The episodic buffer: a new component of working memory?. *Trends in cognitive sciences*, 4(11), 417-423. doi: 10.1016/S1364-6613(00)01538-2
- Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4, 829-839. doi: 10.1038/nrn1201
- Baddeley, A. (2007). *Working Memory, Thought and Action*. Oxford, UK: Oxford Univ. Press.
- Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.
- Baddeley, A. D. & Hitch, G. (1974). Working memory. *Psychology of learning and motivation*, 8, 47-89.
- Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of verbal learning and verbal behavior*, 14(6), 575-589. doi: 10.1016/S0079-7421(08)60452-1
- Baddeley, A., Gathercole, S., & Papagno, C. (1998). The Phonological Loop as a Language Learning Device. *Psychological Review*, 105(1), 158-173. doi: 10.1037/0033-295X.105.1.158
- Baird, G., Dworzynski, K., Slonims, V., & Simonoff, E. (2010). Memory impairment in children with language impairment. *Developmental Medicine & Child Neurology*, 52(6), 535-540. doi: 10.1111/j.1469-8749.2009.03494.x

- Balladares, J., Marshall, C., & Griffiths, Y. (2016). Socio-economic status affects sentence repetition, but not non-word repetition, in Chilean preschoolers. *First Language*, 36(3), 338-351. doi: 10.1177/0142723715626067
- Barry, J. G., Yasin, I., & Bishop, D. V. (2007). Heritable risk factors associated with language impairments. *Genes, Brain and Behavior*, 6(1), 66-76. doi: 10.1111/j.1601-183X.2006.00232.x
- Bertrand, R. & Camos, V. (2015). The role of attention in preschoolers' working memory. *Cognitive Development*, 33, 14-27. doi: 10.1016/j.cogdev.2014.10.002
- Binamé, F. & Poncelet, M. (2016). Order short-term memory capacity predicts nonword reading and spelling in first and second grade. *Reading and Writing*, 29(1), 1-20. doi: 10.1007/s11145-015-9577-9
- Bishop, D. V., Adams, C. V., & Norbury, C. F. (2006). Distinct genetic influences on grammar and phonological short-term memory deficits: evidence from 6-year-old twins. *Genes, Brain and Behavior*, 5(2), 158-169. doi: 10.1111/j.1601-183X.2005.00148.x
- Bishop, D. V., Holt, G., Line, E., McDonald, D., McDonald, S., & Watt, H. (2012). Parental phonological memory contributes to prediction of outcome of late talkers from 20 months to 4 years: a longitudinal study of precursors of specific language impairment. *Journal of neurodevelopmental disorders*, 4(3), 1-12. doi: 10.1186/1866-1955-4-3
- Bishop, D. V., North, T., & Donlan, C. H. R. I. S. (1996). Nonword repetition as a behavioural marker for inherited language impairment: Evidence from a twin study. *Journal of child Psychology and Psychiatry*, 37(4), 391-403. doi: 10.1111/j.1469-7610.1996.tb01420.x
- Boerma, T., Chiat, S., Leseman, P., Timmermeister, M., Wijnen, F., & Blom, E. (2015). A Quasi-Universal Nonword Repetition Task as a Diagnostic Tool for Bilingual Children Learning Dutch as a Second Language. *Journal of Speech, Language, and Hearing Research*, 58(6), 1747-1760. doi: 10.1044/2015_JSLHR-L-15-0058
- Boets, B., Wouters, J., van Wieringen, A., & Ghesquière, P. (2006). Auditory temporal information processing in preschool children at family risk for dyslexia: Relations with phonological abilities and developing literacy skills. *Brain and Language*, 97(1), 64-79. doi: 10.1016/j.bandl.2005.07.026

- Bonifacci, P., Barbieri, M., Tomassini, M., & Roch, M. (2017). In few words: linguistic gap but adequate narrative structure in preschool bilingual children. *Journal of Child Language*, 1-28. doi: 10.1017/S0305000917000149
- Bortolini, U., Arfé, B., Caselli, C. M., Degasperi, L., Deevy, P., & Leonard, L. B. (2006). Clinical markers for specific language impairment in Italian: the contribution of clitics and non-word repetition. *International Journal of Language & Communication Disorders*, 41(6), 695-712. doi: 10.1080/13682820600570831
- Bowey, J. A. (2006). Clarifying the phonological processing account of nonword repetition. *Applied Psycholinguistics*, 27(4), 548-552. doi: 10.1017/S0142716406220393
- Boyle, W., Lindell, A. K., & Kidd, E. (2013). Investigating the role of verbal working memory in young children's sentence comprehension. *Language Learning*, 63(2), 211-242. doi: 10.1111/lang.12003
- Brady, K. W. & Goodman, J. C. (2014). The type, but not the amount, of information available influences toddlers' fast mapping and retention of new words. *American Journal of Speech-Language Pathology*, 23(2), 120-133. doi:10.1044/2013_AJSLP-13-0013
- Brown, G., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological review*, 107(1), 127-181. doi: 10.1037/0033-295X.107.1.127
- Brown, T. T. & Jernigan, T. L. (2012). Brain development during the preschool years. *Neuropsychology review*, 22(4), 313-333. doi: 10.1007/s11065-012-9214-1
- Bühner, M. (2006). *Einführung in die Test- und Fragebogenkonstruktion* (2. Aufl.). München: Pearson Studium.
- Burgess, N. & Hitch, G. (2005). Computational models of working memory: putting long-term memory into context. *Trends in cognitive sciences*, 9(11), 535-541. doi: 10.1016/j.tics.2005.09.011
- Burke, H. L. & Coady, J. A. (2015). Nonword repetition errors of children with and without specific language impairments (SLI). *International Journal of Language & Communication Disorders*, 50(3), 337-346. doi: 10.1111/1460-6984.12136
- Camos, V. & Barrouillet, P. N. (2011). Developmental Change in Working Memory Strategies: From Passive Maintenance to Active Refreshing. *Developmental Psychology*, 47(3), 898-904. doi: 10.1037/a0023193

- Carpenter, J. L. & Drabick, D. A. (2011). Co-occurrence of linguistic and behavioural difficulties in early childhood: A developmental psychopathology perspective. *Early child development and care*, 181(8), 1021-1045. doi: 10.1080/03004430.2010.509795
- Carroll, J. M., Mundy, I. R., & Cunningham, A. J. (2014). The roles of family history of dyslexia, language, speech production and phonological processing in predicting literacy progress. *Developmental Science*, 17(5), 727-742. doi: 10.1111/desc.12153
- Casserly, E. D. & Pisoni, D. B. (2013). Nonword Repetition as a Predictor of Long-Term Speech and Language Skills in Children With Cochlear Implants. *Otology & Neurotology*, 34(3), 460-470. doi: 10.1097/MAO.0b013e3182868340
- Catts, H. W., Nielsen, D. C., Bridges, M. S., Liu, Y. S., & Bontempo, D. E. (2014). Early identification of reading disabilities within an RTI framework. *Journal of learning disabilities*, 48(3), 281-297. doi: 10.1177/0022219413498115
- Chiat, S. (2015). Nonword Repetition. In: S. Armon-Lotem, J. de Jong, & N. Meir (Hrsg.), *Methods for assessing multilingual children: disentangling bilingualism from Language Impairment* (S. 125–150). Bristol: Multilingual Matters.
- Chiat, S. & Polišenská, K. (2016). A Framework for Crosslinguistic Nonword Repetition Tests: Effects of Bilingualism and Socioeconomic Status on Children's Performance. *Journal of Speech, Language, and Hearing Research*, 59(5), 1179-1189. doi: 10.1044/2016_JSLHR-L-15-0293
- Chiat, S. & Roy, P. (2007). The Preschool Repetition Test: An evaluation of performance in typically developing and clinically referred children. *Journal of Speech, Language, and Hearing Research*, 50(2), 429-443. doi: 10.1044/1092-4388(2007/030)
- Claessen, M., Leitão, S., Kane, R., & Williams, C. (2013). Phonological processing skills in specific language impairment. *International journal of speech-language pathology*, 15(5), 471-483. doi: 10.3109/17549507.2012.753110
- Coady, J. A. & Evans, J. L. (2008). Uses and interpretations of non-word repetition tasks in children with and without specific language impairments (SLI). *International Journal of Language & Communication Disorders*, 43(1), 1-40. doi: 10.1080/13682820601116485
- Coady, J. A., Mainela-Arnold, E., & Evans, J. L. (2013). Phonological and lexical effects in verbal recall by children with specific language impairments. *International journal of language & communication disorders*, 48(2), 144-159. doi: 10.1111/1460-6984.12005

- Cohen J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale: Lawrence Erlbaum.
- Colic, H. A. & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of Verbal Learning and Verbal Behavior*, 15(1), 17-31. doi: 10.1016/S0022-5371(76)90003-7
- Colle, H. A. (1980). Auditory encoding in visual short-term recall: Effects of noise intensity and spatial location. *Journal of Verbal Learning and Verbal Behavior*, 19(6), 722-735. doi: 10.1016/S0022-5371(80)90403-X
- COST Action IS0804. (2013). *Language Impairment in a Multilingual Society: Linguistic Patterns and the Road to Assessment*. Abgerufen von <http://www.bi-sli.org/>
- Côté, I., Rouleau, N. & Macoir, J. (2014). New Word Acquisition in Children: Examining the Contribution of Verbal Short-term Memory to Lexical and Semantic Levels of Learning. *Applied Cognitive Psychology*, 28(1), 104-114. doi: 10.1002/acp.2961
- Cowan, N. (2010). Multiple concurrent thoughts: The meaning and developmental neuropsychology of working memory. *Developmental neuropsychology*, 35(5), 447-474. doi: 10.1080/87565641.2010.494985
- Cowan, N. (2016). Working memory maturation: can we get at the essence of cognitive growth?. *Perspectives on Psychological Science*, 11(2), 239-264. doi: 10.1177/1745691615621279
- Cowan, N., Ricker, T. J., Clark, K. M., Hinrichs, G. A., & Glass, B. A. (2015). Knowledge cannot explain the developmental growth of working memory capacity. *Developmental science*, 18(1), 132-145. doi: 10.1111/desc.12197
- Cumming G. (2012). *Understanding the New Statistics: Effect sizes, Confidence Intervals, and Meta-Analysis*. New York, NY: Routledge.
- de Abreu, P. M. E., Baldassi, M., Puglisi, M. L., & Befi-Lopes, D. M. (2013). Cross-linguistic and cross-cultural effects on verbal working memory and vocabulary: Testing language-minority children with an immigrant background. *Journal of Speech, Language, and Hearing Research*, 56(2), 630-642. doi: 10.1044/1092-4388(2012/12-0079)
- de Abreu, P. M. E., Conway, A. R., & Gathercole, S. E. (2010). Working memory and fluid intelligence in young children. *Intelligence*, 38(6), 552-561. doi: 10.1016/j.intell.2010.07.003

- Deevy, P., Weil, L. W., Leonard, L. B., & Goffman, L. (2010). Extending use of the NRT to preschool-age children with and without specific language impairment. *Language, Speech, and Hearing Services in Schools*, 41(3), 277-288. doi: 10.1044/0161-1461(2009/08-0096)
- Demetriou, A., Spanoudis, G., Shayer, M., van der Ven, S., Brydges, C. R., Kroesbergen, E., ... & Swanson, H. L. (2014). Relations between speed, working memory, and intelligence from preschool to adulthood: Structural equation modeling of 14 studies. *Intelligence*, 46, 107-121. doi: 10.1016/j.intell.2014.05.013
- Dispaldro, M., Leonard, L. B., & Deevy, P. (2013). Real-word and nonword repetition in Italian-speaking children with specific language impairment: A study of diagnostic accuracy. *Journal of Speech, Language, and Hearing Research*, 56(1), 323-336. doi: 10.1044/1092-4388(2012/11-0304)
- Dohmen, A., Chiat, S., & Roy, P. (2013). Nonverbal imitation skills in children with specific language delay. *Research in developmental disabilities*, 34(10), 3288-3300. doi: 10.1016/j.ridd.2013.06.004
- Dollaghan, C. A. & Campbell, T. F. (2009). How well do poor language scores at ages 3 and 4 predict poor language scores at age 6?. *International Journal of Speech-Language Pathology*, 11(5), 358-365. doi: 10.1080/17549500903030824
- Dollaghan, C. & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research*, 41(5), 1136-1146. doi: 10.1044/jslhr.4105.1136
- Domahs, U., Klein, E., Huber, W., & Domahs, F. (2013). Good, bad and ugly word stress—fMRI evidence for foot structure driven processing of prosodic violations. *Brain and language*, 125(3), 272-282. doi: 10.1016/j.bandl.2013.02.012
- Domahs, U., Plag, I., & Carroll, R. (2014). Word stress assignment in German, English and Dutch: Quantity-sensitivity and extrametricality revisited. *The Journal of Comparative Germanic Linguistics*, 17(1), 59-96. doi: 10.1007/s10828-014-9063-9
- Domahs, U., Wiese, R., Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). The processing of German word stress: evidence for the prosodic hierarchy. *Phonology*, 25(01), 1-36. doi: 10.1017/S0952675708001383

- Döring, N. & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5. vollständig überarb., akt. und erw. Aufl.). Berlin, Heidelberg: Springer-Verlag GmbH.
- Duncan, T. S. & Paradis, J. (2016). English language learners' nonword repetition performance: The influence of age, L2 vocabulary size, length of L2 exposure, and L1 phonology. *Journal of Speech, Language, and Hearing Research*, 59(1), 39-48. doi:10.1044/2015_JSLHR-L-14-0020
- Eaton, C. T. & Ratner, N. B. (2013). Rate and phonological variation in preschool children: Effects of modeling and directed influence. *Journal of Speech, Language, and Hearing Research*, 56(6), 1751-1763. doi: 10.1044/1092-4388(2013/12-0171)
- Eid, M., Gollwitzer, M., & Schmitt, M. (2011). *Statistik und Forschungsmethoden Lehrbuch*. Weinheim: Beltz.
- Engel, P. M. J., Santos, F. H., & Gathercole, S. E. (2008). Are Working Memory Measures Free of Socioeconomic Influence?. *Journal of Speech, Language, and Hearing Research*, 51(6), 1580-1587. doi: 10.1044/1092-4388(2008/07-0210)
- Faul, (2014). *G*Power Version 3.1.9.2*. Deutschland: Universität Kiel. Abgerufen von <http://www.gpower.hhu.de/>
- Felkel, M. (in Vorbereitung). *Überarbeitung und testtheoretische Überprüfung der Skala Lernbär des Wiener Entwicklungstests* (Nicht veröffentlichte Masterarbeit), Universität Wien, Österreich.
- Fiebach, C. J., Friederici, A. D., Smith, E. E., & Swinney, D. (2007). Lateral Inferotemporal Cortex Maintains Conceptual—Semantic Representations in Verbal Working Memory. *Journal of Cognitive Neuroscience*, 19(12), 2035-2049. doi: 10.1162/jocn.2007.19.12.2035
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4. Aufl.). London, England: Sage.
- Fisseni, H.-J. (2004). *Lehrbuch der psychologischen Diagnostik* (3. erw. und überarb. Aufl.). Göttingen: Hogrefe.
- Fitzpatrick, C. & Pagani, L. S. (2012). Toddler working memory skills predict kindergarten school readiness. *Intelligence*, 40(2), 205-212. doi: 10.1016/j.intell.2011.11.007

- Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological bulletin*, 76(5), 378-382. doi: 10.1037/h0031619
- Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2. Aufl.). New York: Wiley.
- Florit, E., Roch, M., Altoè, G., & Levorato, M. C. (2009). Listening comprehension in pre-schoolers: The role of memory. *British Journal of Developmental Psychology*, 27(4), 935-951. doi: 10.1348/026151008X397189
- Fox, A. V. & Dodd, B. (2001). Phonologically disordered German-speaking children. *American Journal of Speech-Language Pathology*, 10(3), 291-307. doi: doi:10.1044/1058-0360(2001/026)
- Fox-Boyer, A. (2016). *Kindliche Aussprachestörungen. Phonologischer Erwerb – Differenzialdiagnostik - Therapie* (7. Aufl.). Idstein: Schulz-Kirchner Verlag GmbH.
- Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, 27(4), 513-543. doi: 10.1017/S0142716406060383
- Gathercole, S. E. & Pickering, S. J. (2000). Assessment of Working Memory in Six-and Seven-Year-Old Children. *Journal of Educational Psychology*, 92(2), 377-390. doi: 10.1037/0022-0663.92.2.377
- Gathercole, S. E., Durling, E., Evans, M., Jeffcock, S., & Stone, S. (2008). Working memory abilities and children's performance in laboratory analogues of classroom activities. *Applied Cognitive Psychology*, 22(8), 1019-1037. doi: 10.1002/acp.1407
- Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, 40(2), 177-190. doi: 10.1037/0012-1649.40.2.177
- Gathercole, S. E., Tiffany, C., Briscoe, J., & Thorn, A. (2005). Developmental consequences of poor phonological short-term memory function in childhood: a longitudinal study. *Journal of child Psychology and Psychiatry*, 46(6), 598-611. doi: 10.1111/j.1469-7610.2004.00379.x
- Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory*, 2(2), 103-127. doi: 10.1080/09658219408258940

- Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental psychology*, 28(5), 887-898. doi: 10.1037/0012-1649.28.5.887
- Ghasemi, A. & Zahediasl, S. (2012). Normality Tests for Statistical Analysis: A Guide for Non-Statisticians. *International Journal of Endocrinology and Metabolism*, 10(2), 486-489. doi: 10.5812/ijem.3505
- Girbau, D. & Schwartz, R. G. (2008). Phonological working memory in Spanish–English bilingual children with and without specific language impairment. *Journal of Communication Disorders*, 41(2), 124-145. doi: 10.1016/j.jcomdis.2007.07.001
- Gobet, F., Lane, P. C., Croker, S., Cheng, P. C., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in cognitive sciences*, 5(6), 236-243. doi: 10.1016/S1364-6613(00)01662-4
- Goswami, U., Barnes, L., Mead, N., Power, A. J., & Leong, V. (2016). Prosodic Similarity Effects in Short-Term Memory in Developmental Dyslexia. *Dyslexia*, 22(4), 287-304. doi: 10.1002/dys.1535
- Graf Estes, K. G., Evans, J. L., & Else-Quest, N. M. (2007). Differences in the nonword repetition performance of children with and without specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, 50(1), 177-195. doi: 10.1044/1092-4388(2007/015)
- Gray, S. (2006). The relationship between phonological memory, receptive vocabulary, and fast mapping in young children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 49(5), 955-969. doi: 10.1044/1092-4388(2006/069)
- Gray, S., Pittman, A., & Weinhold, J. (2014). Effect of phonotactic probability and neighborhood density on word-learning configuration by preschoolers with typical development and specific language impairment. *Journal of Speech, Language, and Hearing Research*, 57(3), 1011-1025. doi:10.1044/2014_JSLHR-L-12-0282
- Grimm, H. (2015). *Sprachentwicklungstest für drei- bis fünfjährige Kinder (SETK 3-5)* (2. überarb. und neu normierte Aufl.). Göttingen: Hogrefe.
- Grimm, H. (2017). *Sprachscreening für das Vorschulalter: Kurzform des SETK 3-5* (2. überarb. und neu normierte Aufl.). Göttingen: Hogrefe.

- Grossberg, S. & Kazerounian, S. (2011). Laminar cortical dynamics of conscious speech perception: Neural model of phonemic restoration using subsequent context in noise. *The Journal of the Acoustical Society of America*, 130(1), 440-460. doi: 10.1121/1.3589258
- Grube, D., Lingen, M., & Hasselhorn, M. (2008). Entwicklung des phonologischen Arbeitsgedächtnisses: Zur Rolle von Rehearsal und Lexikalität für den Ähnlichkeitseffekt. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 40(4), 200-207. doi: 10.1026/0049-8637.40.4.200
- Gruber, K. (1992). *Konstruktion eines Itempools für die Dimensionen "Motorik", "Wahrnehmung" und "Gedächtnis" für 3- bis 6-jährige Kinder* (Nicht veröffentlichte Diplomarbeit). Universität Wien, Österreich.
- Guiberson, M. & Rodriguez, B. L. (2013). Classification accuracy of nonword repetition when used with preschool-age spanish-speaking children. *Language, Speech, and Hearing Services in Schools*, 44(2), 121-132. doi: 10.1044/0161-1461(2012/12-0009)
- Gupta, P. & Tisdale, J. (2009). Word learning, phonological short-term memory, phonotactic probability and long-term memory: towards an integrated framework. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1536), 3755-3771. doi: 10.1098/rstb.2009.0132
- Harper-Hill, K., Copland, D., & Arnott, W. (2013). Do spoken nonword and sentence repetition tasks discriminate language impairment in children with an ASD?. *Research in Autism Spectrum Disorders*, 7(2), 265-275. doi: 10.1016/j.rasd.2012.08.015
- Hartley, T. & Houghton, G. (1996). A linguistically constrained model of short-term memory for nonwords. *Journal of memory and language*, 35(1), 1-31. doi: 10.1006/jmla.1996.0001
- Hasselhorn, M., Schumann-Hengsteler, R., Grube, D., König, J., Mähler, C., & Schmid, I. (2012). *Arbeitsgedächtnistestbatterie für Kinder von 5 bis 12 Jahren (AGTB 5-12)*. Göttingen: Hogrefe.
- Häusler, J. & Kubinger, K. D. (2003). *Zur Trennschärfe der two-way-layout Rangvarianzanalyse*. Abgerufen unter: http://typo3.univie.ac.at/fileadmin/user_upload/inst_entw_dia/diag/soft_down/statistik/zweifache_rangvarianzanalyse.pdf

- Havy, M., Foroud, A., Fais, L., & Werker, J. F. (2017). The Role of Auditory and Visual Speech in Word Learning at 18 Months and in Adulthood. *Child Development*. Online Vorveröffentlichung. doi: 10.1111/cdev.12715
- Havy, M., Nazzi, T., & Bertoncini, J. (2013). Phonetic processing during the acquisition of new words in 3-to-6-year-old French-speaking deaf children with cochlear implants. *Journal of Communication Disorders*, 46(2), 181-192. doi: doi.org/10.1016/j.jcomdis.2012.12.002
- Helland, T. & Morken, F. (2015). Neurocognitive development and predictors of L1 and L2 literacy skills in dyslexia: A longitudinal study of children 5–11 years old. *Dyslexia*, 22(1), 3-26. doi: 10.1002/dys.1515
- Henry, L. A., Messer, D. J., & Nash, G. (2012). Phonological and visuospatial short-term memory in children with specific language impairment. *Journal of Cognitive Education and Psychology*, 11(1), 45. doi: 10.1891/1945-8959.11.1.45
- Hoff, E., Core, C., & Bridges, K. (2008). Non-word repetition assesses phonological memory and is related to vocabulary development in 20-to 24-month-olds. *Journal of Child Language*, 35(4), 903-916. doi: 10.1017/S0305000908008751
- Hoffman, J. L., Teale, W. H., & Paciga, K. A. (2014). Assessing vocabulary learning in early childhood. *Journal of Early Childhood Literacy*, 14(4), 459-481. doi: 10.1177/146879841350118
- Hu, C. F. & Schuele, C. M. (2005). Learning nonnative names: The effect of poor native phonological awareness. *Applied Psycholinguistics*, 26(3), 343-362. doi: 10.1017/S0142716405050204
- Hummel, K. M. & French, L. M. (2016). Phonological memory and aptitude components: Contributions to second language proficiency. *Learning and Individual Differences*, 51, 249-255. doi: 10.1016/j.lindif.2016.08.016
- Hurlstone, M. J., Hitch, G. J., & Baddeley, A. D. (2014). Memory for serial order across domains: An overview of the literature and directions for future research. *Psychological Bulletin*, 140(2), 339-373. doi: 10.1037/a0034221
- IBM Corporation. (2016). *IBM SPSS statistics for windows, version 24.0*. Armonk, NY: IBM Corporation.

- Jackson, E., Leita, S., & Claessen, M. (2016). The relationship between phonological short-term memory, receptive vocabulary, and fast mapping in children with specific language impairment. *International Journal of Language & Communication Disorders*, 51(1), 61-73. doi: 10.1111/1460-6984.12185
- Jacquemot, C. & Scott, S. K. (2006). What is the relationship between phonological short-term memory and speech processing?. *Trends in cognitive sciences*, 10(11), 480-486. doi: 10.1016/j.tics.2006.09.002
- James, D. G., Ferguson, W. A., & Butcher, A. (2015). Assessing children's speech using picture-naming: The influence of differing phonological variables on some speech outcomes. *International journal of speech-language pathology*, 1-14. doi: 10.3109/17549507.2015.1101159
- Jansen, H., Mannhaupt, G., Marx, H., Skowronek, H. (2002). *Bielefelder Screening zur Früherkennung von Lese-Rechtschreibschwierigkeiten (BISC)* (2. überarb. Aufl.). Göttingen: Hogrefe.
- Janßen, U. (2003). Untersuchungen zum Wortakzent im Deutschen und Niederländischen (Dissertation). Abgerufen von urn:nbn:de:hbz:061-20040802-000911-2
- Jones, D. M. & Macken, W. J. (1995). Phonological Similarity in the Irrelevant Speech Effect: Within-or Between-Stream Similarity?. *Learning, Memory*, 21(1), 103-115. doi: 10.1037/0278-7393.21.1.103
- Jones, G. (2016). The influence of children's exposure to language from two to six years: The case of nonword repetition. *Cognition*, 153, 79-88. doi: 10.1016/j.cognition.2016.04.017
- Jones, G. & Macken, B. (2015). Questioning short-term memory and its measurement: Why digit span measures long-term associative learning. *Cognition*, 144, 1-13. doi: 10.1016/j.cognition.2015.07.009
- Jones, G. & Witherstone, H. L. (2011). Lexical and sublexical knowledge influences the encoding, storage, and articulation of nonwords. *Memory & cognition*, 39(4), 588-599. doi: 10.3758/s13421-010-0045-0
- Jones, G., Gobet, F., & Pine, J. M. (2007). Linking working memory and long-term memory: a computational model of the learning of new words. *Developmental Science*, 10(6), 853-873. doi: 10.1111/j.1467-7687.2007.00638.x

- Jones, G., Gobet, F., Freudenthal, D., Watson, S. E., & Pine, J. M. (2014). Why computational models are better than verbal theories: the case of nonword repetition. *Developmental science*, 17(2), 298-310. doi: 10.1111/desc.12111
- Jones, G., Tamburelli, M., Watson, S. E., Gobet, F., & Pine, J. M. (2010). Lexicality and frequency in specific language impairment: Accuracy and error data from two nonword repetition tests. *Journal of Speech, Language, and Hearing Research*, 53(6), 1642-1655. doi: 10.1044/1092-4388(2010/09-0222)
- Kan, P. F. & Kohnert, K. (2012). A growth curve analysis of novel word learning by sequential bilingual preschool children. *Bilingualism: Language and Cognition*, 15(3), 452-469. doi: 10.1017/S1366728911000356
- Kapalková, S., Polišenská, K., & Vicenová, Z. (2013). Non-word repetition performance in Slovak-speaking children with and without SLI: novel scoring methods. *International journal of language & communication disorders*, 48(1), 78-89. doi: 10.1111/j.1460-6984.2012.00189.x
- Kastner-Koller, U. & Deimann, P. (2012). *Wiener Entwicklungstest (WET)* (3. überarb. und erw. Aufl.). Göttingen: Hogrefe.
- Kaufman, A. S. & Kaufman, N. L. (1993). *Kaufman Assessment Battery für Children*. Frankfurt: Swets & Zeitlinger.
- Kaushanskaya, M., Gross, M., & Buac, M. (2013). Gender differences in child word learning. *Learning and individual differences*, 27, 82-89. doi: 10.1016/j.lindif.2013.07.002
- Keilmann, A., Moein, G., & Schöler, H. (2012). Werden mit dem SETK 3–5 klinisch diagnostizierte Sprachentwicklungsstörungen erfasst?. *HNO*, 60(1), 63-71. doi: 10.1007/s00106-011-2408-7
- Keren-Portnoy, T., Vihman, M. M., DePaolis, R. A., Whitaker, C. J., & Williams, N. M. (2010). The role of vocal practice in constructing phonological working memory. *Journal of Speech, Language, and Hearing Research*, 53, 1280-1293. doi: 10.1044/1092-4388(2009/09-0003)
- Keuleers, E. & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior research methods*, 42(3), 627-633. doi:10.3758/BRM.42.3.627

- Kiese-Himmel, C. (2011). Welche phonologische Arbeitsgedächtnisleistung trennt Kinder mit AVWS von unauffälligen Kindern?. *HNO*, 59(3), 292-300. doi: 10.1007/s00106-010-2198-3
- Kiese-Himmel, C., Auberlen, S., & von Steinbüchel, N. (2012). Ausgewählte Sprachentwicklungsstandsfacetten von Kindergartenkindern mit Migrationshintergrund in Deutschland. *Zeitschrift für medizinische Psychologie*, 21(2), 80-88. doi: 10.3233/ZMP-2012-210004
- Kohnert, K., Windsor, J., & Yim, D. (2006). Do language-based processing tasks separate children with language impairment from typical bilinguals?. *Learning Disabilities Research & Practice*, 21(1), 19-29. doi: 10.1111/j.1540-5826.2006.00204.x
- Koller, I., Alexandrowicz, R., & Hatzinger, R. (2012). *Das Rasch Modell in der Praxis: Eine Einführung mit eRm*. Wien: facultas.wuv, UTB.
- Kraus, N. & White-Schwoch, T. (2015). Unraveling the biology of auditory learning: a cognitive–sensorimotor–reward framework. *Trends in cognitive sciences*, 19(11), 642-654. doi: 10.1016/j.tics.2015.08.017
- Krishnan, S., Alcock, K. J., Carey, D., Bergström, L., Karmiloff-Smith, A., & Dick, F. (2017). Fractionating nonword repetition: The contributions of short-term memory and oromotor praxis are different. *PLoS One*, 12(7), 1-18. doi: 10.1371/journal.pone.0178356
- Krishnan, S., Watkins, K. E., & Bishop, D. V. (2016). Neurobiological basis of language learning difficulties. *Trends in cognitive sciences*, 20(9), 701-714. doi: 10.1016/j.tics.2016.06.012
- Kron-Sperl, V., Schneider, W., & Hasselhorn, M. (2008). The development and effectiveness of memory strategies in kindergarten and elementary school: Findings from the Würzburg and Göttingen longitudinal memory studies. *Cognitive Development*, 23(1), 79-104. doi: 10.1016/j.cogdev.2007.08.011
- Kubinger, K.D. & Holocher-Ertl, S. (2014). *Adaptives Intelligenz Diagnostikum - Version 3.1 (AID 3)*. Göttingen: Beltz.
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in psychology*, 4(863), 1-12. doi: 10.3389/fpsyg.2013.00863

- Lange, B. P., Euler, H. A., & Zaretsky, E. (2016). Sex differences in language competence of 3-to 6-year-old children. *Applied Psycholinguistics*, 1-22. doi: 10.1017/S0142716415000624
- Law, F., Mahr, T., Schneeberg, A., & Edwards, J. (2016). Vocabulary size and auditory word recognition in preschool children. *Applied Psycholinguistics*, 38(1), 89-125. doi: 10.1017/S0142716416000126
- Law, J., McBean, K., & Rush, R. (2011). Communication skills in a population of primary school-aged children raised in an area of pronounced social disadvantage. *International Journal of Language & Communication Disorders*, 46(6), 657-664. doi: 10.1111/j.1460-6984.2011.00036.x
- Laws, G. & Gunn, D. (2004). Phonological memory as a predictor of language comprehension in Down syndrome: a five-year follow-up study. *Journal of Child Psychology and Psychiatry*, 45(2), 326-337. doi: 10.1111/j.1469-7610.2004.00224.x
- Leclercq, A. L. & Majerus, S. (2010). Serial-Order Short-Term Memory Predicts Vocabulary Development: Evidence from a Longitudinal Study. *Developmental Psychology*, 46(2), 417-427. doi: 10.1037/a0018540
- Lehmann, J., Quaiser-Pohl, C., & Jansen, P. (2014) Correlation of motor skill, mental rotation, and working memory in 3- to 6-year-old children. *European Journal of Developmental Psychology*, 11(5), 560-573, doi: 10.1080/17405629.2014.888995
- Levene, H. (1960). Robust tests for equality of variances. In I. Olkin, S. G. Churye, W. Hoeffding, W. G. Madow, & H. B. Mann (Hrsg.), *Contributions to probability and statistics* (S. 278-292). Stanford, CA: Stanford University Press.
- Heister, J., Würzner, K. M., Bubenzner, J., Pohl, E., Hanneforth, T., Geyken, A., & Kliegl, R. (2011). dlexDB—eine lexikalische Datenbank für die psychologische und linguistische Forschung. *Psychologische Rundschau*, 62(1), 10-20. doi: 10.1026/0033-3042/a000029
- Lenhard, W. & Lenhard, A. (2014). *Hypothesis Tests for Comparing Correlations*. Abgerufen am 16.10.2017 von <https://www.psychometrica.de/correlation.html>. Bibergau (Deutschland). doi: 10.13140/RG.2.1.2954.1367
- Lonigan, C. J., Anthony, J. L., Phillips, B. M., Purpura, D. J., Wilson, S. B., & McQueen, J. D. (2009). The nature of preschool phonological processing abilities and their relations

- to vocabulary, general cognitive abilities, and print knowledge. *Journal of educational psychology*, 101(2), 345. doi: 10.1037/a0013837
- Lunkenheimer, E. & Wang, J. (2017). It's OK to Fail: Individual and Dyadic Regulatory Antecedents of Mastery Motivation in Preschool. *Journal of Child and Family Studies*, 26(5), 1481-1490. doi: 10.1007/s10826-016-0633-0
- MacRoy-Higgins, M. & Dalton, K. P. (2015). The Influence of Phonotactic Probability on Nonword Repetition and Fast Mapping in 3-Year-Olds With a History of Expressive Language Delay. *Journal of Speech, Language, and Hearing Research*, 58(6), 1773-1779. doi:10.1044/2015_JSLHR-L-15-0079
- Mair, P., Hatzinger, R., & Maier M. J. (2016). *eRm: Extended Rasch modeling*. Abgerufen von <http://CRAN.R-project.org/package=eRm>.
- Majerus, S. & Boukebza, C. (2013). Short-term memory for serial order supports vocabulary development: New evidence from a novel word learning paradigm. *Journal of Experimental Child Psychology*, 116(4), 811-828. doi: 10.1016/j.jecp.2013.07.014
- Majerus, S. & D'Argembeau, A. (2011). Verbal short-term memory reflects the organization of long-term memory: Further evidence from short-term memory for emotional words. *Journal of Memory and Language*, 64(2), 181-197. doi: 10.1016/j.jml.2010.10.003
- Majerus, S., D'Argembeau, A., Perez, T. M., Belayachi, S., van der Linden, M., Collette, F., ... & Maquet, P. (2010). The commonality of neural networks for verbal and visual short-term memory. *Journal of cognitive neuroscience*, 22(11), 2570-2593. doi: 10.1162/jocn.2009.21378
- Majerus, S., Poncelet, M., Greffe, C., & van der Linden, M. (2006). Relations between vocabulary development and verbal short-term memory: The relative importance of short-term memory for serial order and item information. *Journal of experimental child psychology*, 93(2), 95-119. doi: 10.1016/j.jecp.2005.07.005
- Marsh, R., Gerber, A. J., & Peterson, B. S. (2008). Neuroimaging studies of normal brain development and their relevance for understanding childhood neuropsychiatric disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, 47(11), 1233-1251. doi: 10.1097/CHI.0b013e318185e703
- Masso, S., McLeod, S., Baker, E., & McCormack, J. (2016). Polysyllable productions in preschool children with speech sound disorders: Error categories and the Framework of

- Polysyllable Maturity. *International journal of speech-language pathology*, 18(3), 272-287. doi: 10.3109/17549507.2016.1168483
- Mathieu, J., Lindner, K., Lomako, J., & Gagarina, N. (2016). „Wo bist du, kleiner Monster?“ Sprachspezifische nonword repetition Tests zur Differenzierung von bilingualen typisch entwickelten Kindern und entsprechenden Risikokindern für USES. *Forschung Sprache*, 4(1), 5-24. Abgerufen von <https://www.forschung-sprache.eu/heft-archiv/jahrgang-4-2016/heft-1/>
- Melchers, P. & Preuß, U. (1991). *Kaufman Assessment Battery for Children*. Frankfurt/Main: Swets & Zeitlinger.
- Melzer, J., Rißling, J. K., & Petermann, F. (2015). Sprachdiagnostik im Vorschulalter. *Monatsschrift Kinderheilkunde*, 163(1), 58-66. doi: 10.1007/s00112-014-3269-4
- Messer, M. H., Leseman, P. P., Boom, J., & Mayo, A. Y. (2010). Phonotactic probability effect in nonword recall and its relationship with vocabulary in monolingual and bilingual preschoolers. *Journal of Experimental Child Psychology*, 105(4), 306-323. doi: 10.1016/j.jecp.2009.12.006
- Messer, M. H., Verhagen, J., Boom, J., Mayo, A. Y., & Leseman, P. P. (2015). Growth of verbal short-term memory of nonwords varying in phonotactic probability: A longitudinal study with monolingual and bilingual children. *Journal of Memory and Language*, 84, 24-36. doi: 10.1016/j.jml.2015.05.001
- Metsala, J. L. & Chisholm, G. M. (2010). The influence of lexical status and neighborhood density on children's nonword repetition. *Applied Psycholinguistics*, 31(3), 489-506. doi: 10.1017/S0142716410000081
- Renner, G. & Mickley, M. (2015). Berücksichtigen deutschsprachige Intelligenztests die besonderen Anforderungen von Kindern mit Behinderungen?. *Praxis der Kinderpsychologie und Kinderpsychiatrie*, 64(2), 88-103. doi: 10.13109/prkk.2015.64.2.88
- Montgomery, J. W., Magimairaj, B. M., & Finney, M. C. (2010). Working memory and specific language impairment: An update on the relation and perspectives on assessment and treatment. *American journal of speech-language pathology*, 19(1), 78-94. doi: 10.1044/1058-0360(2009/09-0028)
- Mottier, G. (1951). Mottier-Test. Über Untersuchungen der Sprache lesegestörter Kinder. *Folia Phoniatica*, 3(3), 170-177. doi: 10.1159/000262507

- Moosbrugger, H. (2012). Klassische Testtheorie (KTT). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 103-117) (2. akt. und überarb. Aufl.). Berlin: Springer Verlag.
- Moyle, M. J., Heilmann, J. J., & Finneran, D. A. (2014). The role of dialect density in non-word repetition performance: An examination with at-risk African American pre-school children. *Clinical linguistics & phonetics*, 28(9), 682-696.
doi:10.3109/02699206.2014.882990
- Munson, B., Kurtz, B. A., & Windsor, J. (2005). The influence of vocabulary size, phonotactic probability, and wordlikeness on nonword repetitions of children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research*, 48(5), 1033-1047. doi: 10.1044/1092-4388(2005/072)
- Næss, K. A. B., Lervåg, A., Lyster, S. A. H., & Hulme, C. (2015). Longitudinal relationships between language and verbal short-term memory skills in children with Down syndrome. *Journal of experimental child psychology*, 135, 43-55. doi: 10.1016/j.jecp.2015.02.004
- Nash, H. M., Hulme, C., Gooch, D., & Snowling, M. J. (2013). Preschool language profiles of children at family risk of dyslexia: continuities with specific language impairment. *Journal of Child Psychology and Psychiatry*, 54(9), 958-968. doi: 10.1111/jcpp.12091
- Lonigan, C. J., Schatschneider, C., & Westberg, L. (2008). Identification of children's skills and abilities linked to later outcomes in reading, writing, and spelling. In National Center for Family Literacy, *Developing early literacy: Report of the National Early Literacy Panel* (S. 55–106). Washington, DC: National Institute for Literacy. Abgerufen von <https://www.nichd.nih.gov/publications/pubs/documents/NELPReport09.pdf>
- Nevo, E. & Breznitz, Z. (2013). The development of working memory from kindergarten to first grade in children with different decoding skills. *Journal of experimental child psychology*, 114(2), 217-228. doi: 10.1016/j.jecp.2012.09.004
- Newbury, D. F., Winchester, L., Addis, L., Paracchini, S., Buckingham, L. L., Clark, A., ... & Goodyer, I. M. (2009). CMIP and ATP2C2 modulate phonological short-term memory in language impairment. *The American Journal of Human Genetics*, 85(2), 264-272. doi: 10.1016/j.ajhg.2009.07.004

- Niklas, F., Schmiedeler, S., Pröstler, N., & Schneider, W. (2011). Die Bedeutung des Migrationshintergrunds, des Kindergartenbesuchs sowie der Zusammensetzung der Kindergartengruppe für sprachliche Leistungen von Vorschulkindern. *Zeitschrift für Pädagogische Psychologie*, 25(2), 115–130. doi: 10.1024/1010-0652/a000032
- Nora, A., Hultén, A., Karvonen, L., Kim, J. Y., Lehtonen, M., Yli-Kaitala, H., ... & Salmelin, R. (2012). Long-term phonological learning begins at the level of word form. *NeuroImage*, 63(2), 789-799. doi: 10.1016/j.neuroimage.2012.07.026
- Nunnally, J. C. & Bernstein, I. H. (2006) *Psychometric theory*. (3. Aufl.) New York: Lightning Source.
- Papadimitriou, A. M. & Vlachos, F. M. (2014). Which specific skills developing during preschool years predict the reading performance in the first and second grade of primary school?. *Early Child Development and Care*, 184(11), 1706-1722. doi: 10.1080/03004430.2013.875542
- Passolunghi, M. C. & Costa, H. M. (2014). Working memory and early numeracy training in preschool children. *Child Neuropsychology*, 22(1), 81-98. doi: 10.1080/09297049.2014.971726
- Passolunghi, M. C., Lanfranchi, S., Altoè, G., & Sollazzo, N. (2015). Early numerical abilities and cognitive skills in kindergarten children. *Journal of experimental child psychology*, 135, 25-42. doi: 10.1016/j.jecp.2015.02.001
- Perez, T. M., Majerus, S., & Poncelet, M. (2012). The contribution of short-term memory for serial order to early reading acquisition: Evidence from a longitudinal study. *Journal of Experimental Child Psychology*, 111(4), 708-723. doi: 10.1016/j.jecp.2011.11.007
- Peter, B., Raskind, W. H., Matsushita, M., Lisowski, M., Vu, T., Berninger, V. W., ... & Brkanac, Z. (2011). Replication of CNTNAP2 association with nonword repetition and support for FOXP2 association with timed reading and motor activities in a dyslexia family sample. *Journal of neurodevelopmental disorders*, 3(1), 39-49. doi: 10.1007/s11689-010-9065-0
- Petrucelli, N., Bavin, E. L., & Bretherton, L. (2012). Children with specific language impairment and resolved late talkers: Working memory profiles at 5 years. *Journal of Speech, Language, and Hearing Research*, 55(6), 1690-1703. doi: 10.1044/1092-4388(2012/11-0288)

- Poinstingl, H., Mair, P., & Hatzinger, R. (2007). *Manual zum Softwarepackage eRm (extended Rasch modeling). Anwendung des Rasch-Modells (1-PL Modell)*. Lengerich, Deutschland: Pabst Science.
- Polišenská, K. & Kapalková, S. (2014). Improving child compliance on a computer-administered nonword repetition task. *Journal of Speech, Language, and Hearing Research*, 57(3), 1060-1068. doi: 10.1044/1092-4388(2013/13-0014)
- Ponocny, I. (2001). Nonparametric goodness-of-fit tests for the Rasch model. *Psychometrika*, 66(3), 437-459. doi: 10.1007/BF02294444
- Preßler, A. L., Könen, T., Hasselhorn, M., & Krajewski, K. (2014). Cognitive preconditions of early reading and spelling: a latent-variable approach with longitudinal data. *Reading and Writing*, 27(2), 383-406. doi: 10.1007/s11145-013-9449-0
- Preßler, A. L., Krajewski, K., & Hasselhorn, M. (2013). Working memory capacity in preschool children contributes to the acquisition of school relevant precursor skills. *Learning and Individual Differences*, 23, 138-144. doi: 10.1016/j.lindif.2012.10.005
- Ptok, M., Berendes, K., Gottal, S., Grabherr, B., Schneeberg, J., & Wittler, M. (2008). Phonologische Verarbeitung. *Monatsschrift Kinderheilkunde*, 156(9), 860-866. doi: 10.1007/s00112-008-1770-3
- Pugin, F., Metz, A. J., Stauffer, M., Wolf, M., Jenni, O. G., & Huber, R. (2014). Working memory training shows immediate and long-term effects on cognitive performance in children. *F1000Research*, 3, 82-82. doi: 10.12688/f1000research.3665.2
- R Core Team (2017). *R: A Language and Environment for Statistical Computing*. Wien, Österreich: R Foundation for Statistical Computing. Abgerufen von <https://www.R-project.org>
- Radeborg, K., Barthelom, E., Sjöberg, M., & Sahlén, B. (2006). A Swedish non-word repetition test for preschool children. *Scandinavian journal of psychology*, 47(3), 187-192. doi: 10.1111/j.1467-9450.2006.00506.x
- Rasch, G. (1960). *Probabilistic Models for some Intelligence and Attainment Tests*. Kopenhagen: Danish Institute for Educational Research.
- Reichenbach, K., Bastian, L., Rohrbach, S., Gross, M., & Sarrar, L. (2016). Cognitive functions in preschool children with specific language impairment. *International journal of pediatric otorhinolaryngology*, 86, 22-26. doi: 10.1016/j.ijporl.2016.04.011

- Renner, G., Mickley, M., & Friedrichshain, S. Z. (2015). Intelligenzdiagnostik im Vorschulalter. CHC-theoretisch fundierte Untersuchungsplanung und Cross-battery-assessment. *Frühförderung interdisziplinär*, 34(2), 67-83. doi: 10.2378/fi2015.art07d
- Riedl, L., Wiese, R., Dellwo, V., & Wittig, A. (2014). Die Leistung im Memorieren und Nachsprechen von Pseudowörtern: Eine Untersuchung zum Wortakzent im Deutschen. In: M. Steinbach, G. Grewendorf, & A. v. Stechow, *Linguistische Berichte* (Bd. 240) (S. 447-470). Hamburg: Helmut Böske Verlag.
- Risse, T. & Kiese-Himmel, C. (2009). Der Mottier-Test. Teststatistische Überprüfung an 4- bis 6-jährigen Kindern. *HNO*, 57(5), 523-528. doi: 10.1007/s00106-008-1708-z
- Roebers, C. M. & Zoelch, C. (2005). Erfassung und Struktur des phonologischen und visuell-räumlichen Arbeitsgedächtnisses bei 4-jährigen Kindern. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 37(3), 113-121. doi: 10.1026/0049-8637.37.3.113
- Rojas-Barahona, C. A., Förster, C. E., Moreno-Ríos, S., & McClelland, M. M. (2015). Improvement of working memory in preschoolers and its impact on early literacy skills: a study in deprived communities of rural and urban areas. *Early Education and Development*, 26(5-6), 871-892. doi: 10.1080/10409289.2015.1036346
- Rose, S. A, Feldman, J. R., & Jankowski, J. J. (2009). A cognitive approach to the development of early language. *Child Development*, 80(1), 134–150. doi: 10.1111/j.1467-8624.2008.01250.x
- Röttger, T. B., Domahs, U., Grande, M., & Domahs, F. (2012). Structural factors affecting the assignment of word stress in German. *Journal of Germanic Linguistics*, 24(1), 53-94. doi: 10.1017/S1470542711000262
- Roy, P. & Chiat, S. (2004). A prosodically controlled word and nonword repetition task for 2- to 4-year-olds: evidence from typically developing children. *Journal of Speech, Language, and Hearing Research*, 47(1), 223-234. doi: 10.1044/1092-4388(2004/019)
- Roy, P., Chiat, S., & Dodd, B. (2014). Language and socioeconomic disadvantage: From research to practice. London, UK: City University London. Abgerufen von <http://open-access.city.ac.uk/4989/>
- Sawyer, J., Chon, H., & Ambrose, N. G. (2008). Influences of rate, length, and complexity on speech disfluency in a single-speech sample in preschool children who stutter. *Journal of fluency disorders*, 33(3), 220-240. doi: 10.1016/j.jfludis.2008.06.003

- Schneider, W. & Sodian, B. (1990). Gedächtnisentwicklung im Vorschulalter: „Theoriewandel“ im kindlichen Verständnis des Lernens und Erinnerns?. In M. Knopf & W. Schneider (Hrsg.), *Entwicklung. Festschrift zum 60. Geburtstag von Franz Emanuel Weinert* (S. 45-64). Göttingen: Hogrefe.
- Schmid, C., Zoelch, C., & Roebers, C. M. (2008). Das Arbeitsgedächtnis von 4-bis 5-jährigen Kindern. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 40(1), 2-12. doi: 10.1026/0049-8637.40.1.2
- Schmidt-Kassow, M., Rothermich, K., Schwartz, M., & Kotz, S. A. (2011). Did you get the beat? Late proficient French-German learners extract strong-weak patterns in tonal but not in linguistic sequences. *Neuroimage*, 54(1), 568-576. doi: 10.1016/j.neuroimage.2010.07.062
- Schöler, H. (1999). *IDIS - Inventar diagnostischer Informationen bei Sprachentwicklungsauffälligkeiten*. Heidelberg: Edition S im Universitätsverlag C. Winter.
- Schöler, H. & Brunner, M. (2008). *HASE. Heidelberger Auditives Screening in der Einschulungsuntersuchung* (2. erw. Auflage). Wertingen: Westra.
- Schuchardt, K., Maehler, C., & Hasselhorn, M. (2011). Functional deficits in phonological working memory in children with intellectual disabilities. *Research in developmental disabilities*, 32(5), 1934-1940. doi: 10.1016/j.ridd.2011.03.022
- Seeff-Gabriel, B. K., Chiat, S., & Roy, P. (2008). *Early repetition battery*. London: Pearson Assessment.
- Shahmahmood, T. M., Jalaie, S., Soleymani, Z., Haresabadi, F., & Nemati, P. (2016). A systematic review on diagnostic procedures for specific language impairment: The sensitivity and specificity issues. *Journal of research in medical sciences: the official journal of Isfahan University of Medical Sciences*, 21(67), 1-16. doi: 10.4103/1735-1995.189648
- Shapiro, S. S. & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4), 591-611. doi: 10.2307/2333709
- Shapiro, S. S., Wilk, M. B., & Chen, H. J. (1968). A comparative study of various tests for normality. *Journal of the American Statistical Association*, 63(324), 1343-1372. doi: 10.1080/01621459.1968.10480932

- Silvén, M. & Rubinov, E. (2010). Language and preliteracy skills in bilinguals and monolinguals at preschool age: Effects of exposure to richly inflected speech from birth. *Reading and Writing*, 23(3-4), 385-414. doi: 10.1007/s11145-009-9206-6
- Smith, A., Goffman, L., Sasisekaran, J., & Weber-Fox, C. (2012). Language and motor abilities of preschool children who stutter: evidence from behavioral and kinematic indices of nonword repetition performance. *Journal of fluency disorders*, 37(4), 344-358. doi: 10.1016/j.jfludis.2012.06.001
- Spencer, M., Kaschak, M. P., Jones, J. L., & Lonigan, C. J. (2015). Statistical learning is related to early literacy-related skills. *Reading and writing*, 28(4), 467-490. doi: 10.1007/s11145-014-9533-0
- Stacher, I. (in Vorbereitung). *Validierung des Subtests Turnen aus dem Wiener Entwicklungstest* (Nicht veröffentlichte Masterarbeit), Universität Wien, Österreich.
- Stackhouse, J. & Wells, B. (1997). *Children's speech and literacy difficulties: A psycholinguistic framework*. London: Whurr.
- Storkel, H. L. (2015). Learning from input and memory evolution: Points of vulnerability on a pathway to mastery in word learning. *International journal of speech-language pathology*, 17(1), 1-12. doi: 10.3109/17549507.2014.987818
- Storkel, H. L. & Adlof, S. M. (2009). The effect of semantic set size on word learning by preschool children. *Journal of Speech, Language, and Hearing Research*, 52(2), 306–320. doi:10.1044/1092-4388(2009/07-0175)
- Summers, C., Bohman, T. M., Gillam, R. B., Peña, E. D., & Bedore, L. M. (2010). Bilingual performance on nonword repetition in Spanish and English. *International Journal of Language & Communication Disorders*, 45(4), 480-493. doi: 10.3109/13682820903198058
- Sundström, S., Samuelsson, C., & Lyxell, B. (2014). Repetition of words and non-words in typically developing children: The role of prosody. *First Language*, 34(5), 428-449. doi: 10.1177/0142723714550213
- Tanida, Y., Ueno, T., Ralph, M. A. L., & Saito, S. (2015). The roles of long-term phonotactic and lexical prosodic knowledge in phonological short-term memory. *Memory & cognition*, 43(3), 500-519. doi: 10.3758/s13421-014-0482-2

- Tattersall, P. J., Nelson, N. W., & Tyler, A. A. (2014). Associations among nonword repetition and phonemic and vocabulary awareness: Implications for intervention. *Child Language Teaching and Therapy*,. doi: 10.1177/0265659014554719
- Thordardottir, E. & Brandeker, M. (2013). The effect of bilingual exposure versus language impairment on nonword repetition and sentence imitation scores. *Journal of Communication Disorders*, 46(1), 1-16. doi: 10.1016/j.jcomdis.2012.08.002
- Thordardottir, E., Kehayia, E., Mazer, B., Lessard, N., Majnemer, A., Sutton, A., ... & Chilingaryan, G. (2011). Sensitivity and specificity of French language and processing measures for the identification of primary language impairment at age 5. *Journal of Speech, Language, and Hearing Research*, 54(2), 580-597. doi: doi:10.1044/1092-4388(2010/09-0196)
- Truong, D. T., Shriberg, L. D., Smith, S. D., Chapman, K. L., Scheer-Cohen, A. R., DeMille, M. M. C., ... & Gruen, J. R. (2016). Multipoint genome-wide linkage scan for nonword repetition in a multigenerational family further supports chromosome 13q as a locus for verbal trait disorders. *Human genetics*, 135(12), 1329-1341. doi: 10.1007/s00439-016-1717-z
- Ulbrich, C., Alday, P. M., Knaus, J., Orzechowska, P., & Wiese, R. (2016). The role of phonotactic principles in language processing. *Language, Cognition and Neuroscience*, 31(5), 662-682. doi: 10.1080/23273798.2015.1136427
- Ulrich, T. (2016). Sprachtherapeutische Diagnostik mit dem Mottier-Test. Viele Normierungen und viele Fragezeichen. *Forum Logopädie*, 30, 22-29. doi: 10.2443/skv-s-2016-53020160203
- Ulset, V., Vitaro, F., Brendgen, M., Bekkhus, M., & Borge, A. I. (2017). Time spent outdoors during preschool: Links with children's cognitive and behavioral development. *Journal of Environmental Psychology*, 52, 69-80. doi: 10.1016/j.jenvp.2017.05.007
- van Bergen, E., de Jong, P. F., Maassen, B., Krikhaar, E., Plakas, A., & van der Leij, A. (2014). IQ of four-year-olds who go on to develop dyslexia. *Journal of learning disabilities*, 47(5), 475-484. doi: 10.1177/0022219413479673
- van Daal, J., Verhoeven, L., van Leeuwe, J., & van Balkom, H. (2008). Working memory limitations in children with severe language impairment. *Journal of communication disorders*, 41(2), 85-107. doi: 10.1016/j.jcomdis.2007.03.010

- van der Hulst, H. (2009). Brackets and grid marks, or theories of primary accent and rhythm, In E. Raimy & C. E. Cairns (Hrsg.), *Contemporary Views on Architecture and Representations in Phonology* (S. 225-246). Cambridge: MIT Press.
- Vance, M., Stackhouse, J., & Wells, B. (2005). Speech-Production Skills in Children Aged 3-7 Years. *International Journal of Language and Communication Disorders*, 40(1), 29-48. doi: 10.1080/13682820410001716172
- Varga, B. (1996). *Gedächtnisentwicklung im Vorschulalter: Konstruktion von Testitems zur Überprüfung der Gedächtnisentwicklung im Vorschulalter (für 3- bis 6-jährige Kinder) (Nicht veröffentlichte Diplomarbeit)*, Universität Wien, Österreich.
- Vennemann, T. (1991). Skizze der deutschen Wortprosodie. *Zeitschrift für Sprachwissenschaft*, 10(1), 86-111. doi: 10.1515/zfsw.1991.10.1.86
- Vennemann, T. (2010): Die Silbe in Akzent und Rhythmus. In: U. Bredel, T. von der Becke, D. Cramm, M. Krüßmann, S. Zepnik, L. Dummer-Smoch, R. Klicpera, K. P. Kuhn, C. Noack et al. (Hrsg.), *Die Silbe im Anfangsunterricht Deutsch: Festschrift zum zehnjährigen Jubiläum des Lehrgangs ABC der Tiere – Silbenmethode mit Silbentrenner* (S. 85-106). Offenburg: Mildenerger.
- Verhelst, N. D. (2008). An efficient MCMC algorithm to sample binary matrices with fixed marginals. *Psychometrika*, 73(4), 705-728. doi: 10.1007/S11336-008-9062-3
- Vihman, M. M., DePaolis, R. A., & Keren-Portnoy, T. (2014). The role of production in infant word learning. *Language Learning*, 64(2), 121-140. doi: 10.1111/lang.12058
- Visu-Petra, L., Stanciu, O., Benga, O., Miclea, M., & Cheie, L. (2014). Longitudinal and concurrent links between memory span, anxiety symptoms, and subsequent executive functioning in young children. *Frontiers in psychology*, 5(433), 1-21. doi: 10.3389/fpsyg.2014.00443
- von Goldammer, A., Mähler, C., Bockmann, A. K., & Hasselhorn, M. (2010). Vorhersage früher Schriftsprachleistungen aus vorschulischen Kompetenzen der Sprache und der phonologischen Informationsverarbeitung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*.doi: 10.1026/0049-8637/a000005
- Vugs, B., Hendriks, M., Cuperus, J., & Verhoeven, L. (2014). Working memory performance and executive function behaviors in young children with SLI. *Research in Developmental Disabilities*, 35(1), 62-74. doi: 10.1016/j.ridd.2013.10.022

- Waber, D. P., De Moor, C., Forbes, P. W., Almli, C. R., Botteron, K. N., Leonard, G., ... & Rumsey, J. (2007). The NIH MRI study of normal brain development: performance of a population based sample of healthy children aged 6 to 18 years on a neuropsychological battery. *Journal of the International Neuropsychological Society*, 13(5), 729-746. doi: 10.1017/S1355617707070841
- Wagner, R. K. & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological bulletin*, 101(2), 192. doi: 10.1037/0033-2909.101.2.192
- Wald, A. (1943). Tests of statistical hypothesis concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, 54(3), 426-482. doi: 10.2307/1990256
- Walker, J. F. & Archibald, L. M. D. (2006). Articulation rate in preschool children: a 3-year longitudinal study. *International Journal of Language & Communication Disorders*, 41(5), 541-565. doi: 10.1080/10428190500343043
- Weinrich, M. & Zehner, H. (2011). *Phonetische und phonologische Störungen bei Kindern* (S. 1-36). Berlin: Springer.
- Whitehouse, A. J., Barry, J. G., & Bishop, D. V. (2007). The broader language phenotype of autism: a comparison with specific language impairment. *Journal of Child Psychology and Psychiatry*, 48(8), 822-830. doi: 10.1111/j.1469-7610.2007.01765.x
- White-Schwoch, T., Carr, K. W., Thompson, E. C., Anderson, S., Nicol, T., Bradlow, A. R., ... & Kraus, N. (2015). Auditory processing in noise: A preschool biomarker for literacy. *PLoS biology*, 13(7), 1-17. doi: 10.1371/journal.pbio.1002196
- Wiese, R. (1996). Phonological versus morphological rules: On German umlaut and ablaut. *Journal of Linguistics*, 32(1), 113-135. doi: 10.1017/S0022226700000785
- Wild, N. & Fleck, C. (2013). Neunormierung des Mottier-Tests für 5-bis 17-jährige Kinder mit Deutsch als Erst-oder als Zweitsprache. *Praxis Sprache*, 3, 152-158. Abgerufen von <https://www.schulpsychologie-sg.ch/pic-pdf-temp/Neunormierung%20Mottier-Test.pdf>

Tabellenverzeichnis

Tabelle 1. <i>Erwerb von Konsonanten und initialer Konsonantencluster bei deutschsprachigen Kindern</i>	22
Tabelle 2. <i>Physiologische phonologische Prozesse bei deutschsprachigen Kindern</i>	22
Tabelle 3. <i>Itemeigenschaften etablierter Messinstrumente im deutsch- und englischsprachigen Raum</i>	25
Tabelle 4. <i>Kreuztabelle mit Alter und Geschlecht sowie mit Aufgaben beim ersten Testtermin</i>	32
Tabelle 5. <i>Soziale Schichtzugehörigkeit –Berufstätigkeit des Vaters/der Mutter in der Stichprobe und in der Grundgesamtheit</i>	32
Tabelle 6. <i>Cutoffs zur Interpretation von Koeffizienten bzw. Effektstärken</i>	42
Tabelle 7. <i>Stichprobenspezifische Gedächtnisleistungen bei Kunstwörtern (BM St) und Ziffernfolgen</i>	43
Tabelle 8. <i>Likelihood-Quotienten-Tests: Itemsets 1 und 2, getrennt nach Bewertungsmodus</i>	45
Tabelle 9. <i>Itemkennwerte und Reliabilitätsschätzungen der Itemsets vor Itemselektion</i>	46
Tabelle 10. <i>Gedächtnisleistungen von Teilstichprobe 1 nach Itemselektion</i>	48
Tabelle 11. <i>Mittlere Itemkennwerte und Verteilungen von Items auf Trennschärfenbereiche</i>	51
Tabelle 12. <i>Silbenanzahlspezifische mittlere Itemtrennschärfen</i>	51
Tabelle 13. <i>Reliabilitätsschätzungen des selektierten Itemset</i>	53
Tabelle 14. <i>Interkorrelationen zwischen Rohwerten in konstrukt-nahen und -fernen Aufgaben bei Kontrolle des Alters</i>	53

Abbildungsverzeichnis

<i>Abbildung 1. Alterstrends in KWN-Summenscores und Ziffernspannen in Itemsets 1 und 2</i>	44
<i>Abbildung 2. Alterstrends in KWN-Leistungen (selektiertes Itemset) und in Ziffernspannen</i>	49

Anhang

Anhang A: Ergänzende Stichprobeninformationen

Tabelle A1

Absolute und relative Häufigkeiten hinsichtlich Altersgruppen in Halbjahresschritten, Kindergartenstandort und Testleiterin

	n (%)		
	Gesamt	Teilstichprobe 1	Teilstichprobe 2
Alter in Halbjahren: n (%)			
3;0 – 3;5	9 (15.5)	2 (6.5)	7 (25.9)
3;6 – 3;11	10 (17.2)	5 (16.1)	5 (18.5)
4;0 – 4;5	12 (20.7)	6 (19.4)	6 (22.2)
4;6 – 4;11	14 (24.1)	11 (35.5)	3 (11.1)
5;0 – 5;5	13 (22.4)	7 (22.6)	6 (22.2)
Kindergarten			
St. Nikolausstiftung Breitensee	9 (15.5)	5 (16.1)	4 (14.8)
St. Nikolausstiftung Altotterkring	8 (13.8)	4 (12.9)	4 (14.8)
St. Nikolausstiftung St. Hubertus	7 (12.1)	3 (9.7)	4 (14.8)
KIWI Breitensee	17 (29.3)	11 (34.5)	6 (22.2)
KIWI Brünnerstraße	12 (20.7)	6 (19.4)	6 (22.2)
Rekrutierung Bekanntenkreis	5 (8.6)	2 (6.5)	3 (11.1)
Testleiterin			
Minna Felkel	19 (32.8)	10 (32.3)	9 (33.3)
Sarah Macura	21 (36.2)	13 (41.9)	8 (29.6)
Ina Stacher	18 (31.0)	8 (25.8)	10 (37.0)

Anhang B: Informationen an Eltern und KindergartenleiterInnen / -pädagogInnen und Einverständniserklärung

	universität wien	Fakultät für Psychologie			
		Institut für Angewandte Psychologie: Gesundheit, Entwicklung und Förderung			
Sehr geehrte KindergartenleiterInnen und PädagogInnen! Vielen herzlichen Dank für Ihre Bereitschaft zur Kooperation bei den Erhebungen im Rahmen von Masterarbeiten der Universität Wien! Wir möchten Sie hiermit über die Untersuchungen informieren. Ziele: ▪ Es sollen Informationen zur Entwicklung der Motorik und des phonologischen Gedächtnisses bei Kindern im Alter von 3;0 bis 5;11 Jahren erhoben und Aufgaben zur Diagnostik dieser Bereiche überprüft und optimiert werden. Dabei geht es nicht darum, den individuellen Entwicklungsstand der Kinder einzuschätzen. Zielgruppe: ▪ Die Stichprobe für die Untersuchung sollte 30 Kinder im oben genannten Altersbereich umfassen, die zu gleichen Teilen auf die beiden Geschlechter und auf folgende sechs, in Halbjahresschritten differenzierte, Altersgruppen aufgeteilt sind: 3;0-3;5, 3;6-3;11, 4;0-4;5, 4;6-4;11, 5;0-5;5 und 5;6-5;11 Jahre. Ablauf: ▪ Jedem Kind werden von den Masterandinnen an zwei Terminen á 30 Min. pro Kind (Zeitintervall: ca. eine Woche) spielerische Aufgaben zur Grobmotorik (z.B. Ball fangen), Feinmotorik (z.B. Anziehen eines Teddybären), Grafomotorik (Nachzeichnen von geometrischen Figuren) und zum phonologischen Gedächtnis (Nachsprechen von Zahlen und Silben) vorgegeben. Die Antworten der Kinder zu den Gedächtnisaufgaben sollen auf Tonträger aufgezeichnet werden. Bei den Testungen stehen die Bedürfnisse der Kinder selbstverständlich im Vordergrund. Ergänzend sollen Eltern einen Fragebogen zu Verhaltensweisen ihrer Kinder sowie zu soziodemografischen Daten ausfüllen. Ort: ▪ Wir möchten den Kindern die Aufgaben gerne im Kindergarten, in einem ruhigen Raum mit Platz zum Turnen vorgeben und bitten Sie um Bereitstellung eines solchen Raumes. Zeit: ▪ Die Erhebungen werden ab Februar 2017 beginnen. Die Testungen sollen vormittags zwischen 8 und 12 Uhr stattfinden. Selbstverständlich werden wir auf die Abläufe in Ihrem Kindergarten Rücksicht nehmen. Die Teilnahme an der Studie erfolgt freiwillig und nur mit Einverständnis der Eltern. Alle Ergebnisse werden ausschließlich für wissenschaftliche Zwecke verwendet und vollkommen anonym behandelt. Nach Beendigung des Projekts werden die Aufzeichnungen auf den Tonträgern wieder gelöscht. Wir bitten Sie, die Elternbriefe, Fragebögen und Einverständniserklärungen an die Eltern entsprechender Kinder weiterzuleiten. Für alle Fragen stehen wir Ihnen jederzeit sehr gerne telefonisch oder per E-Mail zur Verfügung! Mit freundlichen Grüßen <table border="0"><tr><td>Minna Felkel, BSc a1204128@unet.univie.ac.at 0660/7733139</td><td>Sarah Macura, BSc a1202091@unet.univie.ac.at 0699/19932910 (Masterandinnen)</td><td>Ina Stacher, BSc a1007413@unet.univie.ac.at 0650/6015280</td></tr></table> Ass. Prof. Dr. Ursula Kastner-Koller (Betreuerinnen der Masterarbeiten) Ass. Prof. Dr. Pia Deimann Ass. Prof. Dr. Pia Deimann – T +43-1-4277-47277 E-Mail pia.deimann@univie.ac.at Ass. Prof. Dr. Ursula Kastner-Koller T +43-1-4277-47261 E-Mail ursula.kastner-koller@univie.ac.at			Minna Felkel, BSc a1204128@unet.univie.ac.at 0660/7733139	Sarah Macura, BSc a1202091@unet.univie.ac.at 0699/19932910 (Masterandinnen)	Ina Stacher, BSc a1007413@unet.univie.ac.at 0650/6015280
Minna Felkel, BSc a1204128@unet.univie.ac.at 0660/7733139	Sarah Macura, BSc a1202091@unet.univie.ac.at 0699/19932910 (Masterandinnen)	Ina Stacher, BSc a1007413@unet.univie.ac.at 0650/6015280			

Abbildung B1. Informationsblatt für KindergartenleiterInnen und -pädagogInnen



Liebe Eltern!

Um Kindern mit Problemen helfen zu können, ist es wichtig das Verhalten und die Fähigkeiten von altersentsprechend entwickelten Kindern zu kennen und zu wissen, in welchem Alter diese Kinder fähig sind, bestimmte Aufgaben zu lösen.

Aus diesem Grund werden im Rahmen von Masterarbeiten an der Fakultät für Psychologie Erhebungen in Wiener Kindergärten durchgeführt.

Es ist geplant, dass Kinder im Altersbereich von 3-6 Jahren an spielerischen Aufgaben in den Bereichen Grobmotorik (Turnen), Feinmotorik (Anziehen eines Teddybären), Grafomotorik (Nachzeichnen) und phonologisches Gedächtnis (Nachsprechen von Silben und Zahlen) teilnehmen. Die Antworten der Kinder zu den Gedächtnisaufgaben sollen auf Tonträger aufgezeichnet werden. Nach Beendigung des Projekts werden die Aufzeichnungen auf den Tonträgern wieder gelöscht.

Die Erhebungen finden an zwei Tagen im Kindergarten statt und die erhobenen Daten werden anonym behandelt. Alle Aufgaben sind selbstverständlich auf das Alter der Kinder abgestimmt und bereiten den Kindern erfahrungsgemäß viel Spaß.

Ergänzend bitten wir Sie, einen kurzen Fragebogen auszufüllen, bei dem Sie einschätzen, womit sich Ihr Kind gerne beschäftigt.

Wir würden uns sehr freuen, wenn Ihr Kind an unserer Untersuchung teilnehmen darf!

Mit freundlichen Grüßen,

Minna Felkel, BSc

Sarah Macura, BSc
(Masterandinnen)

Ina Stacher, BSc

Ass. Prof. Dr. Ursula Kastner-Koller

Ass. Prof. Dr. Pia Deimann

(Betreuerinnen der Masterarbeiten)

EINVERSTÄNDNISERKLÄRUNG

Ich bin einverstanden, dass mein Kind.....
geboren am....., an der Untersuchung im Kindergarten teilnimmt
und seine Antworten zu den Gedächtnisaufgaben auf Tonträger aufgezeichnet werden.

.....
Unterschrift des Erziehungsberechtigten

Ass. Prof. Dr. Pia Deimann – T +43-1-4277-47277 E-Mail pia.deimann@univie.ac.at

Ass. Prof. Dr. Ursula Kastner-Koller T +43-1-4277-47261 E-Mail ursula.kastner-koller@univie.ac.at

Abbildung B2. Informationsblatt für Eltern mit Einverständniserklärung

Anhang C: Protokollbögen und bereitgestellte Erklärungen/Instruktionen

Kunstwörter Merken

Material: Protokollbogen, Stift, Handy/Tonträger; **Kurzbeschreibung:** Kunstwörter der Reihe nach gebunden (nicht silbenweise), mit der gekennzeichneten Akzentuierung*, freundlichem Blickkontakt und leicht gesenktem Kopf vorsagen und das Kind soll diese unmittelbar nachsprechen → Start: Übungsitem; Testende: 3;0-3;11 Jahre → Item 17; 4;0-5;11 Jahre → Item 20; Reproduktionen der Kinder während der Durchführung auf Tonträger aufzeichnen und lautgetreu protokollieren; nach Abschluss der Testung: Transkription und Bewertung
*Hauptakzente (erhöhte Lautstärke, lang ausgesprochener Vokal): hochgestelltes Apostroph, fett und unterstrichen; Nebenakzente (leicht erhöhte Lautstärke): tiefgesetztes Apostroph, unterstrichen

Instruktion und Durchführung:

Handy/ Tonträger erklären: **„Jetzt kommt ein Spiel, da möchte ich mir alles, was du sagst, behalten und spiele es deshalb hier darauf!“** → Handy/Tonträger mit dem Display nach unten etwas abseits platzieren und Aufmerksamkeit auf Testleiterin zurücklenken: **„Das Spiel geht so: ich sage dir jetzt ein paar ganz lustige Wörter vor, die du sicher noch nie gehört hast. Ich bin neugierig, ob du sie dir merken kannst. Du musst mir ganz genau zuhören, und wenn ich mit einem Wort fertig bin, sollst du es genau nachsagen. Jetzt kommt das erste lustige Wort:...“** → Übungsitem: bei ausbleibender Reaktion → 1 x wiederholen; Nach Übungsitem: **„Du gibst dir wirklich viel Mühe, das ist super! Jetzt kommen noch mehr lustige Wörter, hör mir ganz genau zu...“**. → Testitems (keine Korrekturen, keine Wiederholung)

Niedrige Kooperationsbereitschaft → Einbeziehen des Lernbären:

- **„Schau mal, der Teddy wünscht sich so sehr, dass du ein paar der Wörter zu ihm sagst, weil die Wörter in seiner Sprache sind. Also, ich sage dir immer ein Wort in Teddys Sprache vor und du hörst mir ganz genau zu. Dann sagst du dieses Wort genauso zum Teddy.“**

Ausbleibende Reaktion → Ermutigung:

- **„Jetzt bist du an der Reihe!“; „Ich möchte das Wort so gerne von dir hören!“; „Ich bin ganz neugierig, wie du das nachsagen kannst!“**

Verzerrtes Aussprechen oder Angleichung an existierende Wörter → Hinweis / Aufforderung zu Genauigkeit:

- **„Das sind alles Wörter, die du nicht kennst. Sprich sie bitte genauso nach, wie ich sie dir vorsage!“; „Das Wort soll genauso klingen, wie ich es sage!“**

Häufige Bestärkung, ca. nach der Hälfte und am Ende jedes Silbenlängenblocks:

- **„Toll, wie du mitmachst!“; „Du bemühst dich wirklich gut, weiter so!“; „Es ist toll, wie du dabei bist!“; „Super, wie tüchtig/eifrig du mitmachst!“; „Ich finde toll, wie du das machst!“ / „Der Teddy findet toll...“; „Es gefällt mir, wie du dich bemühst!“ / „Der Teddy freut sich sehr, wie...“**

Äußerungen des Kindes hinsichtlich der Schwierigkeit der Items → Bestärkung und Ermutigung:

- **„Stimmt, das ist ein ganz schön schwieriges Wort! Das nächste fällt dir vielleicht leichter!“; „Nur wenige schaffen das, es nachzusagen. Versuchen wir ein anderes!“; „Es ist in Ordnung, nicht jedes Wort nachsagen zu können, das ist auch mehr ein Wort für ältere Kinder!“; „Die Wörter nachzusagen ist echt schwierig. Ich finde toll, wie du dich bemühst!“**

Niedrige Aufmerksamkeit → Lenkung der Aufmerksamkeit **vor** dem Vorsprechen des nächsten Testitems:

- **„Gut zuhören, jetzt kommt noch ein lustiges Wort...“; „Pass gut auf, das nächste Wort ist...“; „Schau mich an, jetzt sag ich dir das nächste lustige Wort...!“**

Aufgabenirrelevante Äußerungen/Fragen des Kindes → auf diese Eingehen und späteren Zeitpunkt anbieten:

- **„Ich freu mich schon, wenn du mir das nach diesem Spiel erzählst/zeigst!“; „Das sage ich dir nachher, zuerst spielen wir das zu Ende!“**

Niedrige Motivation → Ziele setzen und mit Anreizen verbinden:

- **„Wenn wir diese Aufgabe beendet haben, machen wir gemeinsam Bewegung! Aber jetzt sagen wir noch ein paar der lustigen Wörter/spielen wir das zu Ende!“; „Wenn wir mit den Wörtern fertig sind, kannst du mir ein lustiges Wort sagen, das ich nachsprechen soll / bringst du dem Teddy ein lustiges Wort bei!“; „Wenn wir mit allem fertig sind, habe ich eine kleine Belohnung für dich!“**

Frage des Kindes nach Bedeutung der Wörter:

- **„Ich habe dir in der lustigen Sprache gerade gesagt: Du bist ein liebes/r, tolles/r Mädchen/Bub!“**

Abbildung C1. Protokollbogen A: Vorangestellte Kurzinstruktionen

Kunstwörter Merken (3;0-5;11)											
Item / Protokollierte Antwort		Transkribierte Antwort	0/1								
			St	E1 a	E2 a, b	E3 a, b	E4 a, b	E5 a-c	E6 a-c	E a-c	H
	Übung: Ban -tuf										
2 Silben	1. Mul -dop										
	2. Baf -mil										
	3. Lab -din										
	4. Tem -fol										
3 Silben	5. Pro -fa-ling										
	6. Dra- nig -sche										
	7. Ble- sam -der										
	8. Nes- böl -ung										
	9. Kno-sep- tist										
4 Silben	10. Fa- dri -lu-sen										
	11. Mi-,ref-ko- nal										
	12. ,Fru-da- lis -bek										
	13. ,Glo-me- rupf -tig										
5 Silben	14. ,Tin-gu-fo- ra -sel										
	15. ,Dul-ka- mo -si-ner										
	16. ,La-po- ni -te-ful										
	17. ,Bir-lo- sta -ke-tu										
Ende 3;0-3;11											
6 Silben	18. ,Pe-ki-sa-,fu-lo- richt										
	19. ,Su-ta-le- zi -bo-kan										
	20. ,Lon-ba-re- no -pi-gum										
Summe:											

Abbildung C2. Protokollbogen A: Itemset 1

Hochgestelltes Apostroph, fett und unterstrichen = Hauptakzent,

tiefgesetztes Apostroph und unterstrichen = Nebenakzent.

^aAltersgruppe 3;0-3;5, ^bAltersgruppe 3;6-3;11, ^cAltersgruppe 4;0-4;11.

Kunstwörter Merken (3;0-5;11)											
Item / Protokollierte Antwort		Transkribierte Antwort	0/1								
			St	E1 a	E2 a, b	E3 a, b	E4 a, b	E5 a-c	E6 a-c	E	H
	Übung: Ban -tuf										
2 Silben	1. ' <u>Neb</u> -luf										
P	2. ' <u>Fen</u> -dal										
	3. ' <u>Lip</u> -nat										
	4. ' <u>Wol</u> -bem										
3 Silben	5. Ku-' <u>bon</u> -heit										
	6. Mer-' <u>gul</u> -bisch										
	7. Gli-' <u>wan</u> -tas										
	8. Li-' <u>stor</u> -kel										
	9. Fra-del-' <u>ponk</u>										
4 Silben	10. Bru-' <u>fil</u> -sa-keit										
	11. , <u>Ne</u> -pa-gil-' <u>mat</u>										
	12. , <u>Sa</u> -ken-' <u>bol</u> -tich										
	13. , <u>Rol</u> -ka-' <u>lens</u> -bar										
5 Silben	14. , <u>Wap</u> -si-' <u>tu</u> -fe-lan										
	15. , <u>Mu</u> -fa-' <u>go</u> -be-ri										
	16. , <u>Il</u> -bo-' <u>kleu</u> -ta-fon										
	17. , <u>Ku</u> -zel-' <u>bra</u> -to-nis										
Ende 3;0-3;11											
6 Silben	18. , <u>Ro</u> -ple-ha-, <u>ti</u> -so-' <u>lut</u>										
	19. , <u>Nu</u> -sa-le-' <u>fa</u> -be-rit										
	20. , <u>Pil</u> -do-ke-' <u>lu</u> -wa-sor										
Summe:											

Abbildung C3. Protokollbogen A: Itemset 2

Hochgestelltes Apostroph, fett und unterstrichen = Hauptakzent,
 tiefgesetztes Apostroph und unterstrichen = Nebenakzent.

^aAltersgruppe 3;0-3;5, ^bAltersgruppe 3;6-3;11, ^cAltersgruppe 4;0-4;11

Bewertung (dichotom; Summenscore)

Höchstpunktzahl für 3;0-3;11 Jahre: 17 Rohwertpunkte; für 4;0-5;11 Jahre: 20 Rohwertpunkte.

Die **protokollierten** Antworten sind nur nach der Standardbewertung (St) auszuwerten.

Für die **transkribierten** Antworten erfolgt eine mehrfache Bewertung und die Bildung von drei Summenscores (St, E und H). Für die Altersgruppen 3;0-4;11 (a-c) werden Antworten entwicklungsspezifisch ausgewertet (E1-E6, E), während für die Altersgruppe 5;0-5;11 (d) nur die Kategorien St und H aktiv bewertet werden und die Werte in E immer jenen von St entsprechen (Übertragung am Protokollbogen nicht erforderlich, in SPSS wird die Spalte St in E kopiert). Die Kategorien E1 bis E6, E und H sind „Toleranzkategorien“, d.h. in diesen Kategorien werden das Auftreten betreffender Fehler in den Antworten toleriert und in diesem Fall die jeweiligen Items als gelöst bewertet.

Die Bewertung E1 bis E6 (abhängig von der genauen Altersgruppe; Kennzeichnung ^{a-c}) dient der Nachvollziehbarkeit bzw. persönlichen Unterstützung für die Bewertung E. In **E1 bis E6** wird der **Wert 1** eingetragen, wenn der betreffende entwicklungstypische Fehler in der Antwort auftritt (auch bei gemeinsamem Auftreten eines anderen entwicklungstypischen Fehlers) und kein nicht-entwicklungstypischer Fehler auftritt; in **E** wird der **Wert 1** eingetragen, wenn in **E1 bis E6 mind. einmal der Wert 1** steht. Infolgedessen bedeutet ein Wert 1 in E, dass eine fehlerhafte Antwort nur durch (einen oder mehrere) entwicklungstypische „Fehler“ zustande gekommen ist.

Vorgehen:

- wenn Wert 1 in Spalte St → Wert 1 wird in alle Felder übertragen
- in nicht-schattierten Feldern → Neubewertung, wenn in Spalte St der Wert 0 steht
- in schattierten Feldern → Übertragung des Wertes aus der Spalte St; keine Neubewertung
- Allgemein (auch in der Kategorie St) werden Veränderungen der Silbenendungen „-er“ zu „-a“ (z.B. Blesamder → „Blesamda“) oder „-el“ zu „-l“ (z.B. Listorkel → „Listorkl“) nicht als Fehler bewertet.

❖ **St... Standardbewertung:**

- 1 = vollständig korrekte Reproduktion (alle Laute und Silben in der richtigen Reihenfolge)
- 0 = fehlerhafte Antwort (z.B. Ersetzungen, Hinzufügungen, Auslassungen)

❖ **H... Bewertung mit Hinzufügungen**

- 1 = korrekte Reproduktion und Antwort mit Hinzufügungen (Laute, Silben; z.B. Muldopt, Balftmil)
- 0 = andere Fehler (z.B. Ersetzungen, Auslassungen, entwicklungstypische Fehler)

❖ **E1 – E6... Entwicklungs- und fehlerspezifische Bewertung (nur für 3;0-4;11 J. (a-c)):**

- 1 = korrekte Reproduktion und Antwort mit dem betreffenden entwicklungstypischen Fehler der jeweiligen Kategorie; auch wenn zusätzlich weitere entwicklungstypische Fehler für die betreffende Altersgruppe auftreten wird der Wert 1 in der Kategorie eingetragen (Tabelle EB);
- 0 = andere, nicht-entwicklungstypische Fehler (z.B. H)
 - ➔ In nicht-schattierten Felder (= von Fehlerkategorie betroffene Items): Neubewertung der Antworten für relevante Altersgruppe**; in schattierten Feldern: Übertragung der Werte aus Spalte St

❖ **E... Entwicklungsspezifische Bewertung gesamt:**

- 1 = korrekte Reproduktion und Antwort mit ausschließlich Fehlern E1 bis E6 oder SE;
- 0 = nicht-entwicklungstypische Fehler in der Antwort (z.B. H)
 - ➔ wenn in Spalten E1 bis E6 und/oder SE der Wert 1 eingetragen ist → Übertragung des Wertes 1; sonst Übertragung der Werte aus Spalte St

****Altersgruppen für die eine entwicklungsspezifische Bewertung oder Bewertung SE erfolgt:**

- a... Bewertung für Altersgruppe **3;0-3;5**
- b... Bewertung für Altersgruppe **3;6-3;11**
- c... Bewertung für Altersgruppe **4;0-4;11**

Abbildung C4. Angaben zur Bewertung der Kunstwörter Teil 1

Bewertungskategorien E1-E6, betroffene Items und relevante Altersgruppen

Entwicklungstypische Fehler in der Reproduktion → Bewertung = 1	Beispiel	Betroffene Items	Alter		
			^a 3;0-3;5	^b 3;6-3;11	^c 4;0-4;11
E1 Vorverlagerung von /k/ → /t/; /g/ → /d/ (Achtung: in umgekehrter Reihenfolge (z.B. /t/ → /k/) ist dies ein Fehler)	<u>K</u> ubonheit → <u>T</u> ubonheit <u>G</u> lomerupftig → <u>D</u> omerupftig	Protokollbogen A_I: 5, 6, 8, 9, 11-15, 17, 18-20 Protokollbogen A_II: 5-13, 15-17, 20	X		
E2 Reduktion der Konsonantenverbindungen am Silben-Anfangsrand (Onset) (Achtung: Reduktion von Konsonantenverbindungen am Silbenende sind Fehler)	<u>F</u> radelpong → <u>F</u> adelpong	Protokollbogen A_I: 5-7, 9, 10, 12, 13, 17 Protokollbogen A_II: 7-10, 16-18	X	X	
E3 Kontaktassimilation „dr“ → „gr“	<u>D</u> ranigsche → <u>G</u> ranigsche	Protokollbogen A_I: 6	X	X	
E4 Vorverlagerung von „ch“ → „s“	Sakenbolt <u>i</u> ch → Sakenbolt <u>i</u> s	Protokollbogen A_I: 12 Protokollbogen A_II: 18	X	X	
E5 Vorverlagerung von „sch“ → „s“	Dranig <u>s</u> che → Dranig <u>s</u> e	Protokollbogen A_I: 6 Protokollbogen A_II: 6	X	X	X
E6 Sonorierungen (Ersetzungen von stimmlosen Lauten durch deren stimmhafte Gegenspieler /p/ → /b/; /t/ → /d/; /k/ → /g/; /z/ → /s/ Entstimmungen (Ersetzungen von stimmhaften Lauten durch deren stimmlose Gegenspieler /b/ → /p/; /d/ → /t/; /g/ → /k/; /s/ → /z/	<u>T</u> inguforasel → <u>D</u> inguforasel <u>K</u> noseptist → <u>G</u> nose <u>b</u> tist Sutale <u>z</u> ibokan → Sutale <u>s</u> ibokan <u>B</u> lesamder → <u>P</u> lezamder <u>D</u> ulkamosiner → <u>T</u> ulkamosiner <u>G</u> lomerupftig → <u>K</u> lomerupftig	bei allen Items möglich	X	X	X

Beispiele für entwicklungsspezifische Bewertung:

- Item: Dranigsche → Antwort: „Dranigse“ (Vorverlagerung von „sch“ → „s“) → in der Standardbewertung und bei nicht-entwicklungsspezifischen Bewertungsmodi wird der Wert 0 eingetragen; wenn das Kind in Altersgruppe a-c ist, wird in E5 und in E der Wert 1 eingetragen.
- Item: Glomerupftig → Antwort: „Gomerupftig“ (Reduktion von Konsonantenverbindungen) → in der Standardbewertung und bei nicht-entwicklungsspezifischen Bewertungsmodi wird der Wert 0 eingetragen; wenn das Kind in Altersgruppe a oder b ist, wird in E2 und in E der Wert 1 eingetragen. Wenn das Kind in Altersgruppe c ist, wird in allen Bewertungskategorien der Wert 0 eingetragen.
- Item: Profaling → Antwort: „Polfagling“ → nach allen Bewertungsmodi und in allen entwicklungsspezifischen Bewertungskategorien (E1-E6 und E) wird der Wert 0 eingetragen, da die fehlerhafte Antwort nicht nur durch entwicklungstypische Fehler (E2), sondern auch durch einen nicht-entwicklungstypischen Fehler (H) zustande gekommen ist.

Abbildung C5. Angaben zur Bewertung der Kunstwörter Teil 2

Kind A: Alter = 3;3 Jahre (Altersgruppe a)

Kunstwörter Marken (3;0-5;11)															
Item / Protokolierte Antwort				Transkribierte Antwort		0/1									
						St	E1 a	E2 a,b	E3 a,b	E4 a,b	E5 a-c	E6 a-c	E a-c	H	
	Übung: 'Ban-tuf ✓				✓										
2 Silben	1. 'Mul-dop ✓	1		✓	1								1	1	
	2. 'Baf-mil ✓	1		✓	1								1	1	
	3. 'Lab-din ×	0		Nabdin	0								0	0	
	4. 'Tem-fo/ ×	0		Tenfo	0								0	0	
3 Silben	5. 'Pro-fa-ling ×	0		Pofaling	0		1						1	0	
	6. D/a-'nig-sche ×	0		Danigsche	0		1						1	0	
	7. Ble-'sam-der ✓	1		Blesamda	1								1	1	
	8. Nes-'böl-ung ✓	1		✓	1								1	1	
	9. Klo-sep-'tist Polifist? ×	0		Polizipst	0								0	0	
4 Silben	10. Fa-'fri-ly-sef ×	0		Fafrilose	0								0	0	
	11. Mi-'ref-ko-'nal ×	0		Mireftonal	0	1							1	0	
	12. 'Fru-da-'lis-bek ✓	1		✓	1								1	1	
	13. Glo-me-'rupf-tig ×	0		Klomerupftig	0							1	1	0	
5 Silben	14. 'Tin-gu-fo-'fa-sef ×	0		Tingufolase	0								0	0	
	15. 'Dul-ka-'mo-si-ner- ta ×	0		Dulkamosita	0								0	0	
	16. 'La-po-'ni-te-fyl ×	0		Lapolitefol	0								0	0	
	17. 'Bir-lo-'sta-ke-tu ×	0		Birlostaketor	0								0	0	
Ende 3;0-3;11															
6 Silben	18. 'Pe-ki-sa-'fu-lo-'richt														
	19. 'Su-ta-le-'zi-bo-kan														
	20. 'Lon-ba-re-'no-pi-gum														
5					Summe:	5							8	5	

Abbildung C6. Bewertungsbeispiel Protokollbogen 1

Kind B: Alter = 4;9 Jahre

Kunstwörter Merken (3;0-5;11)												
Item / Protokollierte Antwort			Transkribierte Antwort		0/1							
					St	E1 a	E2 a, b	E3 a, b	E4 a, b	E5 a-c	E6 a-c	E
	Übung: 'Ban-tuf ✓			✓								
2 Silben	1. 'Neb-luf ✓	1		✓	1							1 1
	2. 'Fen-dal ✓	1		✓	1							1 1
	3. 'Lip-nat ✓	1		✓	1							1 1
	4. 'Wol-beh _n	0		Wolben	0							0 0
3 Silben	5. Ku-'bon-heit _{bon}	0		Kubon bon heit	0							0 1
	6. Mer-'gul-bis _{ch} x	0		Mergulbis	0					1		1 0
	7. Gli-'wan-tas ✓	1		✓	1							1 1
	8. Li-'stor-kel ✓	1		Listorkel	1							1 1
	9. Fra-del-'ponk ✓	1		✓	1							1 1
4 Silben	10. Bu-'fil-sa-keit x	0		Bufilsakeit	0							0 0
	11. 'Ne-pa-gil-'mat ✓	1		✓	1							1 1
	12. 'Sa-ken-'bol-tich _s x	0		Sakenboltis	0							0 0
	13. 'Rol-ka-'lens-bar ✓	1		✓	1							1 1
5 Silben	14. 'Wap-si-'tu-fe-lan _t x	0		Tapsitufelan	0							0 0
	15. 'Mu-fa-'go-be-ri ✓	1		✓	1							1 1
	16. 'Il-bo-'kleu-ta-fon _{läufa} x	0		Ilboläufafon	0							0 0
	17. 'Ku-zel-'bra-to-nis _{zi ballolis} x	0		Kuzibrallolis	0							0 0
Ende 3;0-3;11												
6 Silben	18. 'Ro-plē-ha-'ji-so-'lut _{i h n}	0		Roplihanisonut	0							0 0
	19. 'Nu-sa-le-'fa-be-rit ✓	1		✓	1							1 1
	20. 'Pil-do-ke-'tu-wa-so _{f s}	0		Pildofesuwaso	0							0 0
10					Summe:	10						11 11

Abbildung C7. Bewertungsbeispiel Protokollbogen 2

Anhang D: Voraussetzungstests

Tabelle D1

Shapiro-Wilk-Tests zur Überprüfung auf Normalverteilung

Unabhängige Variable (UV)	Vollständiges Itemset / Ziffernfolgen						Selektiertes Itemset		
Abhängige Variable	Itemset 1			Itemset 2					
Subgruppen (UV)	W	df	p	W	df	p	W	df	p
Alter									
KWN-Summenscore St									
3 Jahre	0.92	7	.468	0.93	12	.391	0.84	7	.107
4 Jahre	0.95	17	.468	0.96	9	.747	0.97	17	.812
5 Jahre	0.95	7	.734	0.98	6	.960	0.95	7	.731
KWN-Silbenspanne St									
3 Jahre	0.87	7	.183	0.78	12	.006	0.90	7	.330
4 Jahre	0.78	17	.001	0.92	9	.364	0.78	17	.001
5 Jahre	0.72	7	.006	0.50	6	< .001	0.72	7	.006
Ziffernspanne									
3 Jahre	0.68	6	.004	0.68	12	.001			
4 Jahre	0.64	17	< .001	0.54	9	< .001			
5 Jahre	0.84	7	.009	0.64	6	.001			
Geschlecht									
KWN-Summenscore St									
Mädchen	0.98	15	.920	0.96	9	.780	0.98	15	.985
Buben	0.92	16	.179	0.94	18	.333	0.93	16	.223
KWN-Silbenspanne St									
Mädchen							0.86	15	.001
Buben							0.77	16	.021
Ziffernspanne									
Mädchen	0.86	15	.025	0.68	18	< .001			
Buben	0.79	15	.002	0.72	9	.002			
Aufgabe 1. Testtermin									
KWN-Summenscore St									
Kunstwörter	0.96	12	.799	0.91	16	.136	0.96	12	.707
Ziffernfolgen	0.93	19	.165	0.88	11	.093	0.98	19	.931

Ziffernspanne

Kunstwörter	0.82	12	.015	0.68	16	< .001
Ziffernfolgen	0.80	18	.002	0.76	11	.003

Testleiterinnen

KWN-Summenscore St

Minna Felkel				0.97	10	.874
Sarah Macura				0.95	12	.589
Ina Stacher				0.91	8	.331

Ziffernspannen

Minna Felkel	0.80	10	.015
Sarah Macura	0.85	12	.039
Ina Stacher	0.80	8	.027

Sequenz-Längen-Effekt

2 Silben				0.55	31	< .001
3 Silben				0.88	31	.002
4 Silben				0.85	31	< .001
5 Silben				0.78	31	< .001
6 Silben				0.45	31	< .001

Tabelle D2

Levene-Tests zur Überprüfung auf Varianzhomogenität

Unabhängige Variable Abhängige Variable	Vollständiges Itemset / Zahlen Merken				Selektiertes Itemset			
	<i>F</i>	df ₁ /df	df ₂	<i>p</i>	<i>F</i>	df ₁ /df	df ₂	<i>p</i>
Alter x Version Itemset								
KWN-Summenscore St	0.86	5	52	.552				
KWN-Silbenspanne St	0.86	5	52	.515				
Ziffernspanne	1.14	5	51	.352				
Alter								
KWN-Summenscore St					1.56	2	28	.227
KWN-Silbenspanne St					0.68	2	28	.517
Ziffernspanne								
Geschlecht								
						α		
KWN-Summenscore St								
Itemset 1	0.00	29		.964	0.54		29	.469
Itemset 2	0.34	25		.565				
KWN-Silbenspanne St								
Itemset 1	0.02	28		.878	0.39	29	29	.537
Itemset 2	0.00	25		.967				
Ziffernspanne								
Itemset 1	0.31	29		.584				
Itemset 2	0.07	25		.801				
Aufgabe 1. Testtermin								
KWN-Summenscore St								
Itemset 1	0.92	29		.347	0.14	29		.713
Itemset 2	3.35	25		.079				
Ziffernspanne								
Itemset 1	1.83	29		.556				
Itemset 2	0.51	25		.484				
Testleiterinnen								
KWN-Summenscore St					1.56	2	28	.227
KWN-Silbenspanne St					0.68	2	28	.517

Anhang E: Ergänzende Deskriptive Statistiken und interferenzstatistische Analysen

Tabelle E1

Deskriptive Statistiken der Teilstichproben getrennt nach Bewertungsmodus

	KWN-Summenscore					KWN-Silbenspanne				
	Itemset 1									
	St	E	H	EK	EKH	St	E	H	EK	EKH
<i>n</i>	31	31	31	31	31	31	31	31	31	31
<i>M</i>	8.13	8.58	8.65	8.84	9.45	4.58	4.71	4.65	4.74	4.84
<i>SD</i>	3.87	3.58	3.81	3.50	3.41	1.21	1.16	1.11	1.13	1.04
Range	1-17	1-17	1-17	1-17	1-17	2-6	2-6	2-6	2-6	2-6
Schiefe	0.18	0.14	0.07	-0.04	-0.24	-0.69	-0.89	-0.63	-0.96	-0.81
SE Schiefe	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42	0.42
<i>z</i> Schiefe	0.43	0.34	0.17	0.08	0.57	1.64	2.12	1.48	2.27	1.93
Kurtosis	-0.41	-0.02	-0.44	0.22	0.40	-0.44	0.20	-0.39	0.58	0.48
SE Kurtosis	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82	0.82
<i>z</i> Kurtosis	0.50	0.03	0.54	0.27	0.49	0.54	0.25	0.47	0.71	0.59
	Itemset 2									
	St	E	H	EK	EKH	St	E	H	EK	EKH
<i>n</i>	27	27	27	27	27	27	27	27	27	27
<i>M</i>	6.63	7.15	6.96	7.19	7.63	4.52	4.56	4.56	4.56	4.63
<i>SD</i>	3.01	2.85	3.06	2.92	2.91	0.94	0.93	0.93	0.93	0.88
Range	2-12	3-12	2-12	3-13	3-14	3-6	3-6	3-6	3-6	3-6
Schiefe	0.19	0.06	0.04	0.16	0.11	-0.36	-0.48	-0.48	-0.48	-0.61
SE Schiefe	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45	0.45
<i>z</i> Schiefe	0.41	0.14	0.08	0.36	0.23	0.81	1.07	1.07	1.07	1.35
Kurtosis	-1.16	-1.11	-1.03	-0.92	-0.59	-0.71	-0.62	-0.62	-0.21	-0.21
SE Kurtosis	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87
<i>z</i> Kurtosis	1.33	1.27	1.18	1.05	0.68	0.81	0.71	0.24	0.24	0.24

Anmerkung. St, E, H, EK, EKH = Bewertungsmodi.

Tabelle E2

Deskriptive Statistiken zu Gedächtnisleistungen getrennt nach Geschlecht und Leistungen in konstruktfernen Aufgaben, Itemset 1 nach Itemselektion

	KWN				Zahlen Merken				Turnen	Nach- zeichnen
	Mädchen		Buben		Mädchen		Buben			
	sum	span	sum	span	span	C	span	C	sum	sum
<i>n</i>	15	15	16	16	15	15	15	15	31	31
<i>M</i>	5.47	4.33	5.06	3.94	3.47	4.47	3.00	4.00	6.65	5.58
<i>SD</i>	2.88	1.68	3.32	1.73	0.74	1.55	1.07	1.77	2.35	2.31
Range	0-11	0-6	0-10	0-6	2-5	2-8	0-4	0-6	2-10	1-9
Schiefe	0.12	-1.35	-0.08	-1.04	0.13	0.53	-1.62	-0.71	-0.20	-0.25
SE Schiefe	0.58	0.58	0.564	0.56	0.58	0.58	0.58	0.58	0.42	0.42
z Schiefe	0.21	2.32	0.13	1.83	0.22	0.92	2.79	1.22	0.48	0.59
Kurtosis	0.01	2.04	1.25	-0.09	0.18	0.45	3.64	0.19	-0.99	-0.84
SE Kurtosis	1.12	1.12	1.09	1.09	1.12	1.12	1.12	1.12	0.82	0.82
z Kurtosis	0.01	1.82	1.14	0.08	0.16	0.40	3.24	0.17	-0.70	-1.20

Anmerkung. sum = Summenscore, span = Gedächtnisspanne.

Tabelle E3

Spezifische fehlerhafte Antworten bei den KWN-Itemsets, getrennt nach Altersgruppen

Bewertungs- modus	% Kinder mit ≥ 1 Fehler der Kategorien (<i>M</i> , <i>SD</i> der Anzahl an Fehlern)					
	Itemset 1			Itemset 2		
	3 Jahre	4 Jahre	5 Jahre	3 Jahre	4 Jahre	5 Jahre
E	71.4 (2.20, 1.10)	17.6 (1.00, 0.00)	n/a	58.3 (1.43, 0.53)	33.3 (1.33, 0.58)	n/a
H	28.6 (2.00, 0.00)	47.1 (1.13, 0.35)	42.9 (1.67, 1.15)	41.7 (1.20, 0.45)	22.2 (1.00, 0.00)	16.7 (1.00, 0.00)
EK/K	42.9 (2.67, 0.58)	0.0 (0.00, 0.00)	n/a	16.7 (1.00, 0.00)	0.0	n/a
E, K und H	14.3 (1.00, 0.00)	11.8 (1.00, 0.00)	n/a	16.7 (1.00, 0.00)	11.1 (1.00, 0.00)	n/a

Anmerkung. n/a = nicht anwendbar/verfügbar.

Tabelle E4

Zweifaktorielle ANOVA und paarweise Vergleiche zu KWN-Summenscores zwischen Itemsets und Altersgruppen

	<i>F</i>	df ₁	df ₂	<i>p</i>	Paarweise Vergleiche ($\alpha_{\text{adj}} = .025$)							
KWN-Summen-												
score												
Version Itemset	0.81	1	52	.371								
/Stichprobe												
Alter	19.1	2	52	<.001	3 vs. 4 Jahre				4 vs. 5 Jahre			
	<i>t</i>	df		<i>p</i>	<i>g</i>	<i>t</i>	df		<i>p</i>	<i>g</i>		
	Itemset 1											
	4.08	28		<.001	1.93	1.99	28		.052	0.76		
	Itemset 2											
	0.97	24		.336	0.44	2.56	24		.001	1.63		
Alter x Itemset	2.52	2	52	.090								

Anmerkung. Paarweise Vergleiche im Rahmen einer *simple effect analysis*.

Tabelle E5

Zweifache Rangvarianzanalyse und Mann-Whitney-Tests (paarweise Vergleiche) zu Ziffernspannen zwischen Itemsets und Altersgruppen

	χ^2	df	p	Paarweise Vergleiche ($\alpha_{\text{adj}} = .025$)							
Ziffernspanne											
Version Itemset	5.25	1	.020								
/Stichprobe											
Alter	14.28	2	.001	3 vs. 4 Jahre				4 vs. 5 Jahre			
	<i>U</i>	<i>z</i>	<i>p</i>	<i>r</i>	<i>U</i>	<i>z</i>	<i>p</i>	<i>r</i>			
	Itemset 1										
	19.50	-2.42	.013	-.51	37.00	-1.59	.171	-.32			
	Itemset 2										
	47.50	-0.54	.636	-.12	14.00	-1.97	.132	-.51			
Alter x Itemset	3.03	2	.217								

Tabelle E6

Mittelwertsvergleiche (t-Tests für unabhängige Stichproben) der KWN-Summscores in Itemsets zwischen Mädchen und Buben

	Geschlecht						t-Test			
	Mädchen			Buben						
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	df	<i>p</i>	<i>g</i>
KWN-Sum-										
menscore										
Itemset 1		8.40	4.03		7.88	3.83	-0.37	29	.731	0.13
Itemset 2		7.11	3.44		6.39	2.85	-0.58	25	.568	0.24

Tabelle E7

Mann-Whitney-Tests zu Unterschieden in Ziffernspannen in Itemsets zwischen Mädchen und Buben

	Geschlecht						Mann-Whitney-Test			
	Mädchen			Buben						
	<i>n</i>	<i>Mdn</i>	<i>IQR</i>	<i>n</i>	<i>Mdn</i>	<i>IQR</i>	<i>U</i>	<i>z</i>	<i>p</i>	<i>r</i>
Ziffernspanne										
Itemset 1		3.00	4.00-3.00		3.00	4.00-3.00	87.50	-1.13	.302	-0.21
Itemset 2		3.00	3.00-2.50		3.00	3.00-2.00	74.00	-0.43	.731	-0.08

Anmerkung. IQR = Interquartile range ($Q_{.75}-Q_{.25}$)

Tabelle E8

Mittelwertsvergleiche (t-Tests für unabhängige Stichproben) der KWN-Summscores zwischen der Vorgabe des KWN-Itemsets beim 1. versus 2. Testtermin

	Vorgabe Kunstwörter beim:						t-Test			
	1. Testtermin			2. Testtermin						
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	df	<i>p</i>	<i>g</i>
KWN-Sum-										
menscore										
Itemset 1	12	7.75	4.27	19	8.37	3.70	-0.43	29	.672	0.16
Itemset 2	16	7.31	3.38	11	5.64	2.16	1.45	25	.160	0.57

Anmerkung. KWN-Summscore = Summscore beim Kunstwörter Nachsprechen, 1. versus 2. Testtermin = Vorgabe der Kunstwörter beim 1. Testtermin versus beim 2. Testtermin.

Tabelle E9

Mann-Whitney-Tests zu Unterschieden in Ziffernspannen in Itemsets zwischen der Vorgabe von Ziffernfolgen beim 1. versus 2. Testtermin

	Ziffernfolgen beim						Mann-Whitney-Test			
	1. Testtermin			2. Testtermin						
	<i>n</i>	<i>Mdn</i>	IQR	<i>n</i>	<i>Mdn</i>	IQR	<i>U</i>	<i>z</i>	<i>p</i>	<i>r</i>
Ziffernspanne										
Itemset 1	12	3.00	4.00-3.00	19	3.00	4.00-3.00	93.00	-0.69	.518	-.12
Itemset 2	16	3.00	3.00-3.00	11	3.00	3.00-0.00	71.00	-1.00	.302	-.19

Anmerkung. 1. versus 2. Testtermin = Vorgabe der Ziffernfolgen beim 1. Testtermin versus beim 2. Testtermin, IQR = *Interquartile range* ($Q_{.75}-Q_{.25}$)

Tabelle E10

Wilcoxon-Rangsummen-Test und Mann-Whitney-Tests (paarweise Vergleiche) zu Silbenspannen zwischen Altersgruppen im selektierten Itemset 1

	Wilcoxon-Rang-			Paarweise Vergleiche ($\alpha_{adj} = .025$)							
	summen-Test			3 vs. 4 Jahre				4 vs. 5 Jahre			
	<i>H</i>	<i>df</i>	<i>p</i>	<i>U</i>	<i>z</i>	<i>p</i>	<i>r</i>	<i>U</i>	<i>z</i>	<i>p</i>	<i>r</i>
KWN-Summen-											
score											
Alter	14.61	2	.001	5.50	-3.59	< .001	-.73	40.50	-1.33	.215	-.27

Anhang F: Itemspezifische Rasch-Analysen (Wald-Tests) vor Itemselektion

Tabelle F1

Z-Test nach Wald bei Itemset 1, Bewertungsmodus St

Item	TK Median		TK Alter		TK Geschlecht	
	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>
1. Muldop ^a	0.59	.553	n/a	n/a	-0.80	.422
2. Bafmil ^a	-0.42	.674	-0.17	.864	-1.73	.084
3. Labdin ^a	1.19	.236	1.63	.103	1.36	.173
4. Temfol ^a	n/a	n/a	-0.17	.864	-1.73	.084
5. Profaling ^a	-0.48	.565	0.82	.414	0.24	.812
6. Dranigsche ^a	0.14	.891	0.15	.879	0.18	.857
7. Blesamder ^a	0.37	.710	-1.10	.317	0.37	.712
8. Nesbölung ^a	-0.38	.703	1.51	.131	0.57	.569
9. Knoseptist ^a	-0.61	.540	-1.37	.170	1.30	.194
10. Fadrilusen ^a	-0.05	.964	-0.37	.711	0.77	.442
11. Mirefkonal ^a	-0.50	.621	n/a	n/a	1.22	.223
12. Frudalisbek ^a	-0.11	.912	1.98	.048	-0.96	.335
13. Glomerupftig ^a	-0.49	.626	n/a	n/a	-0.09	.932
14. Tinguforasel ^a	0.34	.736	-2.06	.039	1.10	.270
15. Dulkamosiner ^a	1.22	.221	0.25	.803	0.44	.660
16. Laponiteful ^a	n/a	n/a	n/a	n/a	-0.05	.958
17. Birlostaketu ^a	n/a	n/a	n/a	n/a	1.41	.159
18. Pekisafuloricht ^b	n/a	n/a	n/a	n/a	-1.15	.251
19. Sutalezibokan ^b	n/a	n/a	n/a	n/a	0.44	.657
20. Lonbarenopigum ^b	n/a	n/a	n/a	n/a	n/a	n/a

Anmerkung. TK = Teilungskriterium, n/a = nicht anwendbar/verfügbar (ungeeignete Antwortmuster).

^a*n* = 31, ^b *n* = 24.

Tabelle F2

Z-Test nach Wald bei Itemset 1, Bewertungsmodus E

Item	TK Median		TK Alter		TK Geschlecht	
	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>
1. Muldop ^a	0.29	.776	-0.45	.656	-1.39	.163
2. Bafmil ^a	0.00	.999	-0.95	.344	-1.63	.102
3. Labdin ^a	1.85	.064	1.35	.178	1.60	.109
4. Temfol ^a	-0.26	.795	-0.48	.632	-1.85	.064
5. Profaling ^a	-0.09	.929	1.65	.099	1.25	.211
6. Dranigsche ^a	-0.74	.460	1.99	.047	1.06	.289
7. Blesamder ^a	1.22	.223	-0.48	.629	1.19	.234
8. Nesbölung ^a	1.06	.291	0.54	.587	0.81	.421
9. Knoseptist ^a	-1.16	.246	-1.39	.163	0.88	.380
10. Fadrilusen ^a	0.77	.441	-0.21	.837	0.83	.409
11. Mirefkonat ^a	-0.89	.376	n/a	n/a	0.83	.409
12. Frudalibek ^a	0.53	.597	2.35	.019	-0.48	.635
13. Glomerupftig ^a	-0.08	.936	n/a	n/a	-0.04	.968
14. Tinguforasel ^a	0.51	.613	-1.80	.072	1.13	.260
15. Dulkamosiner ^a	0.71	.480	0.37	.714	0.49	.628
16. Laponiteful ^a	-0.92	.360	n/a	n/a	-0.4	.679
17. Birlostaketu ^a	-0.92	.360	-1.18	.240	1.331	.183
18. Pekisafuloricht ^b	-0.07	.922	n/a	n/a	-1.04	.297
19. Sutalezibokan ^b	n/a	n/a	n/a	n/a	-0.16	.877
20. Lonbarenopigum ^b	n/a	n/a	n/a	n/a	n/a	n/a

Anmerkung. TK = Teilungskriterium, n/a = nicht anwendbar/verfügbar (ungeeignete Antwortmuster).

^a*n* = 31, ^b *n* = 24.

Tabelle F3

Z-Test nach Wald bei Itemset 1, Bewertungsmodus H

Item	TK Median		TK Alter		TK Geschlecht	
	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>
1. Muldop ^a	0.56	.572	n/a	n/a	-0.77	.440
2. Bafmil ^a	n/a	n/a	-0.59	.553	-1.42	.155
3. Labdin ^a	0.03	.975	0.64	.522	0.93	.354
4. Temfol ^a	n/a	n/a	-0.20	.840	-1.13	.257
5. Profaling ^a	-1.18	.238	0.77	.441	-0.07	.942
6. Dranigsche ^a	-0.59	.553	0.09	.930	-0.15	.881
7. Blesamder ^a	0.60	.549	-0.39	.699	1.27	.204
8. Nesbölung ^a	0.02	.984	1.06	.290	-0.26	.796
9. Knoseptist ^a	-1.42	.156	-1.39	.166	0.94	.349
10. Fadrilusen ^a	0.10	.919	0.30	.764	0.02	.982
11. Mirefkonal ^a	-0.61	.544	n/a	n/a	0.89	.376
12. Frudalisbek ^a	-0.94	.350	2.23	.026	-0.83	.404
13. Glomerupftig ^a	0.42	.674	n/a	n/a	0.02	.988
14. Tinguforasel ^a	0.43	.665	-1.18	.073	1.22	.221
15. Dulkamosiner ^a	1.82	.069	0.47	.637	0.54	.587
16. Laponiteful ^a	-0.56	.573	n/a	n/a	0.06	.955
17. Birlostaketu ^a	n/a	n/a	n/a	n/a	1.50	.133
18. Pekisafuloricht ^b	0.43	.666	n/a	n/a	-1.04	.298
19. Sutalezibokan ^b	n/a	n/a	n/a	n/a	0.51	.613
20. Lonbarenopigum ^b	n/a	n/a	n/a	n/a	n/a	n/a

Anmerkung. TK = Teilungskriterium, n/a = nicht anwendbar/verfügbar (ungeeignete Antwortmuster).

^a*n* = 31, ^b *n* = 24.

Tabelle F4

Z-Test nach Wald bei Itemset 1, Bewertungsmodus EK

Item	TK Median		TK Alter		TK Geschlecht	
	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>
1. Muldop ^a	0.27	.784	-0.51	.607	-1.51	.132
2. Bafmil ^a	1.14	.255	-1.03	.302	-1.75	.080
3. Labdin ^a	1.77	.077	1.26	.207	1.39	.166
4. Temfol ^a	0.80	.422	-0.57	.569	-1.97	.049
5. Profaling ^a	0.15	.883	2.04	.041	1.39	.166
6. Dranigsche ^a	-0.78	.436	1.94	.053	0.86	.387
7. Blesamder ^a	1.07	.285	-0.48	.633	1.02	.309
8. Nesbölung ^a	0.15	.883	0.44	.657	0.59	.558
9. Knoseptist ^a	-1.26	.207	-1.33	.183	0.74	.457
10. Fadrilusen ^a	-0.06	.954	0.84	.401	1.35	.176
11. Mirefkonat ^a	-1.60	.111	-1.53	.126	1.35	.176
12. Frudalibek ^a	0.78	.438	2.70	.007	-0.33	.739
13. Glomerupftig ^a	-0.47	.636	-0.61	.540	0.39	.696
14. Tinguforasel ^a	0.01	.993	-1.31	.189	1.27	.205
15. Dulkamosiner ^a	0.62	.539	0.44	.660	0.39	.696
16. Laponiteful ^a	-1.02	.309	n/a	n/a	-0.52	.601
17. Birlostaketu ^a	n/a	n/a	-1.11	.267	1.21	.227
18. Pekisafuloricht ^b	-0.19	.852	n/a	n/a	-1.02	.307
19. Sutalezibokan ^b	n/a	n/a	n/a	n/a	-0.14	.891
20. Lonbarenopigum ^b	n/a	n/a	n/a	n/a	n/a	n/a

Anmerkung. TK = Teilungskriterium, n/a = nicht anwendbar/verfügbar (ungeeignete Antwortmuster).

^a*n* = 31, ^b *n* = 24.

Tabelle F5

Z-Test nach Wald bei Itemset 1, Bewertungsmodus EKH

Item	TK Median		TK Alter		TK Geschlecht	
	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>
1. Muldop ^a	0.45	.654	-0.65	.517	-1.27	.205
2. Bafmil ^a	0.09	.930	-1.10	.273	-1.54	.122
3. Labdin ^a	0.83	.406	0.73	.466	0.85	.395
4. Temfol ^a	2.10	.036	n/a	n/a	-0.91	.361
5. Profaling ^a	-1.43	.153	2.17	.030	0.95	.340
6. Dranigsche ^a	-0.74	.457	2.09	.037	0.35	.730
7. Blesamder ^a	0.87	.383	-0.09	.931	1.76	.079
8. Nesbölung ^a	0.51	.613	0.20	.839	-0.36	.720
9. Knoseptist ^a	-1.64	.102	-1.46	.145	0.42	.678
10. Fadrilusen ^a	0.51	.613	1.47	.142	1.25	.212
11. Mirefkonat ^a	-1.36	.173	-1.63	.103	1.02	.307
12. Frudalisbek ^a	0.03	.974	2.22	.027	-0.05	.959
13. Glomerupftig ^a	-0.60	.547	-0.56	.573	0.50	.617
14. Tinguforasel ^a	-0.59	.553	-1.40	.161	0.93	.355
15. Dulkamosiner ^a	1.45	.147	0.49	.626	0.50	.617
16. Laponiteful ^a	-1.15	.248	n/a	n/a	-0.40	.689
17. Birlostaketu ^a	n/a	n/a	-1.04	.298	1.31	.189
18. Pekisafuloricht ^b	0.85	.396	n/a	n/a	-0.90	.368
19. Sutalezibokan ^b	n/a	n/a	n/a	n/a	-0.51	.611
20. Lonbarenopigum ^b	n/a	n/a	n/a	n/a	n/a	n/a

Anmerkung. TK = Teilungskriterium, n/a = nicht anwendbar/verfügbar (ungeeignete Antwortmuster).

^a*n* = 31, ^b *n* = 24.

Tabelle F5

Z-Test nach Wald bei Itemset 2, Bewertungsmodus St

Item	TK Median		TK Alter		TK Geschlecht	
	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>
1. Nebluf ^a	2.95	.003	1.26	.207	-0.34	.732
2. Fendal ^a	n/a	n/a	-0.22	.830	n/a	n/a
3. Lipnat ^a	0.62	.533	-0.49	.625	-0.34	.732
4. Wolbem ^a	1.50	.134	1.48	.138	0.04	.970
5. Kubonheit ^a	n/a	n/a	-0.46	.647	1.73	.084
6. Mergulbisch ^a	-0.11	.909	-1.44	.149	1.64	.101
7. Gliwantas ^a	0.27	.785	0.38	.707	-1.07	.286
8. Listorkel ^a	0.26	.792	-0.16	.874	1.40	.162
9. Fradelponk ^a	-1.33	.185	-0.53	.595	0.61	.542
10. Brufilsakeit ^a	-0.95	.344	-1.44	.149	0.78	.437
11. Nepagilmat ^a	n/a	n/a	-0.33	.739	-1.32	.188
12. Sakenboltich ^a	-2.42	.016	-0.63	.531	-0.12	.903
13. Rolkalensbar ^a	0.88	.378	-0.09	.928	-1.67	.095
14. Wapsitufelan ^a	-0.77	.441	n/a	n/a	1.02	.309
15. Mufagoberi ^a	-0.43	.665	0.36	.718	-0.36	.716
16. Ilbokleutafon ^a	0.13	.894	0.77	.443	-1.00	.320
17. Kuzelbratonis ^a	n/a	n/a	n/a	n/a	0.39	.694
18. Roplehatisolut ^a	n/a	n/a	n/a	n/a	n/a	n/a
19. Nusalefaberit ^a	0.93	.353	1.68	.094	-1.01	.314
20. Pildokeluwasor ^a	n/a	n/a	n/a	n/a	n/a	n/a

Anmerkung. TK = Teilungskriterium, n/a = nicht anwendbar/verfügbar (ungeeignete Antwortmuster).

^a*n* = 27, ^b *n* = 15.

Tabelle F6

Z-Test nach Wald bei Itemset 2, Bewertungsmodus E

Item	TK Median		TK Alter		TK Geschlecht	
	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>
1. Nebluf ^a	2.59	.010	1.02	.309	-0.23	.822
2. Fendal ^a	n/a	n/a	n/a	n/a	n/a	n/a
3. Lipnat ^a	0.48	.634	-0.43	.667	0.04	.968
4. Wolbem ^a	1.91	.056	1.22	.223	0.13	.898
5. Kubonheit ^a	n/a	n/a	-0.62	.537	1.86	.064
6. Mergulbisch ^a	-0.52	.606	-0.42	.677	0.13	.898
7. Gliwantas ^a	0.39	.694	0.86	.389	0.28	.780
8. Listorkel ^a	1.23	.218	-0.07	.945	0.82	.414
9. Fradelponk ^a	-1.47	.142	-0.79	.427	0.64	.522
10. Brufilsakeit ^a	-1.60	.110	-1.34	.180	1.08	.282
11. Nepagilmat ^a	n/a	n/a	-0.54	.591	-1.27	.203
12. Sakenboltich ^a	-1.99	.047	-0.88	.379	-0.06	.951
13. Rolkalensbar ^a	0.69	.493	-0.36	.723	-1.61	.108
14. Wapsitufelan ^a	-0.91	.365	n/a	n/a	1.04	.299
15. Mufagoberi ^a	-0.09	.932	0.49	.622	-0.93	.351
16. Ilbokleutafon ^a	-0.09	.932	0.49	.622	-0.93	.351
17. Kuzelbratonis ^a	n/a	n/a	n/a	n/a	0.43	.670
18. Roplehatisolut ^a	n/a	n/a	n/a	n/a	n/a	n/a
19. Nusalefaberit ^a	0.83	.406	1.57	.116	-1.00	.318
20. Pildokeluwasor ^a	n/a	n/a	n/a	n/a	n/a	n/a

Anmerkung. TK = Teilungskriterium, n/a = nicht anwendbar/verfügbar (ungeeignete Antwortmuster).

^a*n* = 27, ^b *n* = 15.

Tabelle F7

Z-Test nach Wald bei Itemset 2, Bewertungsmodus H

Item	TK Median		TK Alter		TK Geschlecht	
	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>
1. Nebluf ^a	2.18	.029	1.20	.231	-0.25	.805
2. Fendal ^a	n/a	n/a	-0.23	.816	n/a	n/a
3. Lipnat ^a	0.49	.623	-0.55	.580	-0.25	.805
4. Wolbem ^a	0.71	.475	1.40	.161	0.15	.881
5. Kubonheit ^a	n/a	n/a	-0.23	.816	1.20	.232
6. Mergulbisch ^a	-0.68	.495	-1.55	.121	1.75	.080
7. Gliwantas ^a	1.00	.316	0.27	.786	-0.97	.334
8. Listorkel ^a	0.15	.883	-0.74	.461	1.75	.080
9. Fradelponk ^a	-1.51	.130	0.19	.850	0.36	.722
10. Brufilsakeit ^a	-0.68	.495	-1.55	.121	0.90	.370
11. Nepagilmat ^a	n/a	n/a	-0.42	.678	-1.26	.207
12. Sakenboltich ^a	-1.92	.055	-0.39	.696	-0.63	.528
13. Rolkalensbar ^a	1.89	.059	0.25	.803	-2.12	.034
14. Wapsitufelan ^a	-1.00	.316	n/a	n/a	1.11	.268
15. Mufagoberi ^a	-0.25	.804	0.56	.575	0.63	.528
16. Ilbokleutafon ^a	-0.68	.498	0.19	.850	-1.48	.140
17. Kuzelbratonis ^a	n/a	n/a	n/a	n/a	0.47	.640
18. Roplehatisolut ^a	n/a	n/a	n/a	n/a	n/a	n/a
19. Nusalefaberit ^a	0.92	.359	1.69	.091	-0.97	.332
20. Pildokeluwasor ^a	n/a	n/a	n/a	n/a	n/a	n/a

Anmerkung. TK = Teilungskriterium, n/a = nicht anwendbar/verfügbar (ungeeignete Antwortmuster).

^a*n* = 27, ^b *n* = 15.

Tabelle F8

Z-Test nach Wald bei Itemset 2, Bewertungsmodus EK

Item	TK Median		TK Alter		TK Geschlecht	
	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>
1. Nebluf ^a	2.63	.009	1.02	.308	-0.22	.827
2. Fendal ^a	n/a	n/a	n/a	n/a	n/a	n/a
3. Lipnat ^a	0.52	.606	-0.43	.669	0.04	.965
4. Wolbem ^a	2.30	.021	1.53	.127	0.41	.685
5. Kubonheit ^a	n/a	n/a	-0.62	.538	1.85	.064
6. Mergulbisch ^a	-0.46	.646	-0.42	.678	0.14	.886
7. Gliwantas ^a	0.47	.640	0.85	.397	0.32	.747
8. Listorkel ^a	1.31	.192	-0.08	.937	0.86	.393
9. Fradelponk ^a	-1.42	.156	-0.84	.401	0.70	.486
10. Brufilsakeit ^a	-2.29	.022	-1.35	.177	0.24	.814
11. Nepagilmat ^a	n/a	n/a	-0.63	.529	-1.22	.224
12. Sakenboltich ^a	-1.93	.053	-0.89	.371	-0.03	.974
13. Rolkalensbar ^a	0.76	.445	-0.41	.684	-1.57	.117
14. Wapsitufelan ^a	-0.85	.397	n/a	n/a	1.10	.270
15. Mufagoberi ^a	-0.01	.990	0.47	.639	-0.90	.369
16. Ilbokleutafon ^a	-0.01	.990	0.47	.639	-0.90	.369
17. Kuzelbratonis ^a	n/a	n/a	n/a	n/a	0.49	.622
18. Roplehatisolut ^a	n/a	n/a	n/a	n/a	n/a	n/a
19. Nusalefaberit ^a	0.88	.380	1.63	.104	-1.00	.319
20. Pildokeluwasor ^a	n/a	n/a	n/a	n/a	n/a	n/a

Anmerkung. TK = Teilungskriterium, n/a = nicht anwendbar/verfügbar (ungeeignete Antwortmuster).

^a*n* = 27, ^b *n* = 15.

Tabelle F9

Z-Test nach Wald bei Itemset 2, Bewertungsmodus EKH

Item	TK Median		TK Alter		TK Geschlecht	
	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>	<i>z</i>	<i>p</i>
1. Nebluf ^a	2.39	.017	1.21	.228	0.11	.916
2. Fendal ^a	n/a	n/a	n/a	n/a	n/a	n/a
3. Lipnat ^a	-0.21	.836	-0.55	.586	0.11	.916
4. Wolbem ^a	1.76	.078	1.40	.162	0.47	.637
5. Kubonheit ^a	n/a	n/a	-0.45	.652	1.31	.189
6. Mergulbisch ^a	0.63	.531	-0.56	.579	0.21	.835
7. Gliwantas ^a	0.11	.911	0.70	.484	0.40	.690
8. Listorkel ^a	1.25	.213	-0.70	.484	1.18	.238
9. Fradelponk ^a	-0.74	.458	-0.15	.882	0.40	.690
10. Brufilsakeit ^a	-0.86	.390	-1.17	.240	0.57	.571
11. Nepagilmat ^a	0.14	.886	-0.80	.426	-1.17	.244
12. Sakenboltich ^a	-1.98	.047	-0.70	.484	-0.58	.564
13. Rolkalensbar ^a	0.99	.323	-0.11	.916	-2.06	.039
14. Wapsitufelan ^a	-1.02	.306	n/a	n/a	1.18	.238
15. Mufagoberi ^a	-1.57	.117	0.59	.558	0.04	.969
16. Ilbokleutafon ^a	0.11	.911	-0.15	.882	-1.42	.155
17. Kuzelbratonis ^a	n/a	n/a	n/a	n/a	0.56	.573
18. Roplehatisolut ^a	n/a	n/a	n/a	n/a	n/a	n/a
19. Nusalefaberit ^a	0.45	.650	1.64	.101	-0.97	.334
20. Pildokeluwasor ^a	n/a	n/a	n/a	n/a	n/a	n/a

Anmerkung. TK = Teilungskriterium, n/a = nicht anwendbar/verfügbar (ungeeignete Antwortmuster).

^a*n* = 27, ^b *n* = 15.

Anhang G: Itemkennwerte

Itemkennwerte vor Itemselektion

Gesamtstichprobe.

Tabelle G1

Itemstatistiken - Itemset 1, Bewertungsmodus St

Item	p_i	$\text{Var}(x_i)$	r_{it}	Cronbachs Alpha wenn Item gelöscht
1. Muldop	.90	.30	.38	.79
2. Bafmil	.81	.40	.22	.79
3. Labdin	.68	.48	.27	.79
4. Temfol	.81	.40	.38	.79
5. Profaling	.52	.51	.52	.78
6. Dranigsche	.58	.50	.50	.78
7. Blesamder	.32	.48	.05	.81
8. Nesbölung	.61	.50	.15	.80
9. Knoseptist	.26	.45	.50	.78
10. Fadrilusen	.35	.49	.55	.77
11. Mirefkonal	.32	.48	.50	.78
12. Frudalisbek	.48	.51	.46	.78
13. Glomerupftig	.16	.37	.34	.79
14. Tinguforasel	.45	.51	.56	.77
15. Dulkamosiner	.19	.40	.14	.80
16. Laponiteful	.23	.43	.40	.78
17. Birlostaketu	.19	.40	.37	.79
18. Pekisafuloricht	.16	.37	.36	.79
19. Sutalezibokan	.06	.25	.28	.79
20. Lonbarenopigum	.03	.18	.29	.79

Anmerkung. $n = 31$, p_i = Itemschwierigkeit, $\text{Var}(x_i)$ = Itemvarianz, r_{it} = Itemtrennschwärfe.

Tabelle G2

Itemstatistiken - Itemset 1, Bewertungsmodus E

Item	p_i	$\text{Var}(x_i)$	r_{it}	Cronbachs Alpha wenn Item gelöscht
1. Muldop	.90	.30	.32	.75
2. Bafmil	.81	.40	.21	.75
3. Labdin	.68	.48	.31	.75
4. Temfol	.84	.37	.33	.75
5. Profaling	.61	.50	.33	.75
6. Dranigsche	.77	.43	.31	.75
7. Blesamder	.35	.49	-.01	.77
8. Nesbölung	.61	.50	.23	.75
9. Knoseptist	.26	.45	.48	.73
10. Fadrilusen	.35	.49	.51	.73
11. Mirefkonal	.32	.48	.49	.73
12. Frudalisbek	.52	.51	.35	.74
13. Glomerupftig	.16	.37	.33	.75
14. Tinguforasel	.45	.51	.52	.73
15. Dulkamosiner	.19	.40	.18	.76
16. Laponiteful	.23	.43	.38	.74
17. Birlostaketu	.23	.43	.24	.75
18. Pekisafuloricht	.16	.37	.36	.74
19. Sutalezibokan	.10	.30	.27	.75
20. Lonbarenopigum	.03	.18	.29	.75

Anmerkung. $n = 31$, p_i = Itemschwierigkeit, $\text{Var}(x_i)$ = Itemvarianz, r_{it} = Itemtrennschwärfe.

Tabelle G3

Itemstatistiken - Itemset 1, Bewertungsmodus H

Item	p_i	$\text{Var}(x_i)$	r_{it}	Cronbachs Alpha wenn Item gelöscht
1. Muldop	.90	.30	.40	.78
2. Bafmil	.84	.37	.29	.79
3. Labdin	.77	.43	.42	.78
4. Temfol	.87	.34	.38	.78
5. Profaling	.55	.51	.51	.77
6. Dranigsche	.61	.50	.54	.77
7. Blesamder	.39	.50	-.06	.81
8. Nesbölung	.68	.48	.09	.80
9. Knoseptist	.29	.46	.55	.77
10. Fadrilusen	.42	.50	.49	.77
11. Mirefkonal	.35	.49	.48	.78
12. Frudalisbek	.48	.51	.42	.78
13. Glomerupftig	.16	.37	.33	.79
14. Tinguforasel	.45	.51	.53	.77
15. Dulkamosiner	.19	.40	.14	.80
16. Laponiteful	.23	.43	.44	.78
17. Birlostaketu	.19	.40	.37	.78
18. Pekisafuloricht	.16	.37	.36	.78
19. Sutalezibokan	.06	.25	.24	.79
20. Lonbarenopigum	.03	.18	.27	.79

Anmerkung. $n = 31$, p_i = Itemschwierigkeit, $\text{Var}(x_i)$ = Itemvarianz, r_{it} = Itemtrennschwärfe.

Tabelle G4

Itemstatistiken - Itemset 1, Bewertungsmodus EK

Item	p_i	$\text{Var}(x_i)$	r_{it}	Cronbachs Alpha wenn Item gelöscht
1. Muldop	.90	.30	.29	.73
2. Bafmil	.81	.40	.18	.74
3. Labdin	.68	.48	.31	.73
4. Temfol	.84	.37	.21	.73
5. Profaling	.65	.49	.33	.73
6. Dranigsche	.77	.43	.36	.72
7. Blesamder	.35	.49	.03	.75
8. Nesbölung	.61	.50	.30	.73
9. Knoseptist	.26	.45	.44	.72
10. Fadrilusen	.42	.50	.44	.72
11. Mirefkonal	.39	.50	.45	.71
12. Frudalisbek	.55	.51	.34	.72
13. Glomerupftig	.19	.40	.30	.73
14. Tinguforasel	.48	.51	.44	.72
15. Dulkamosiner	.19	.40	.15	.74
16. Laponiteful	.23	.43	.35	.72
17. Birlostaketu	.23	.43	.27	.73
18. Pekisafuloricht	.16	.37	.34	.73
19. Sutalezibokan	.10	.30	.25	.73
20. Lonbarenopigum	.03	.18	.28	.73

Anmerkung. $n = 31$, p_i = Itemschwierigkeit, $\text{Var}(x_i)$ = Itemvarianz, r_{it} = Itemtrennschwärfe.

Tabelle G5

Itemstatistiken - Itemset 1, Bewertungsmodus EKH

Item	p_i	$\text{Var}(x_i)$	r_{it}	Cronbachs Alpha wenn Item gelöscht
1. Muldop	.90	.30	.29	.72
2. Bafmil	.84	.37	.22	.73
3. Labdin	.77	.43	.43	.71
4. Temfol	.90	.30	.19	.73
5. Profaling	.68	.48	.34	.72
6. Dranigsche	.81	.40	.41	.71
7. Blesamder	.42	.50	-.01	.74
8. Nesbölung	.68	.48	.21	.73
9. Knoseptist	.29	.46	.49	.70
10. Fadrilusen	.52	.51	.37	.71
11. Mirefkonal	.42	.50	.43	.71
12. Frudalisbek	.58	.50	.29	.72
13. Glomerupftig	.19	.40	.26	.72
14. Tinguforasel	.48	.51	.42	.71
15. Dulkamosiner	.19	.40	.14	.73
16. Laponiteful	.23	.43	.37	.71
17. Birlostaketu	.23	.43	.27	.72
18. Pekisafuloricht	.16	.37	.31	.72
19. Sutalezibokan	.13	.34	.23	.73
20. Lonbarenopigum	.03	.18	.25	.73

Anmerkung. $n = 31$, p_i = Itemschwierigkeit, $\text{Var}(x_i)$ = Itemvarianz, r_{it} = Itemtrennschwärfe.

Tabelle G6

Itemstatistiken - Itemset 2, Bewertungsmodus St

Item	p_i	$\text{Var}(x_i)$	r_{it}	Cronbachs Alpha wenn Item gelöscht
1. Nebluf	.63	.49	-.08	.68
2. Fendal	.78	.42	.46	.61
3. Lipnat	.59	.50	.23	.64
4. Wolbem	.56	.51	-.08	.68
5. Kubonheit	.74	.45	.09	.65
6. Mergulbisch	.37	.49	.41	.61
7. Gliwantas	.15	.36	.11	.66
8. Listorkel	.37	.49	.23	.64
9. Fradelponk	.30	.47	.27	.63
10. Brufilsakeit	.41	.50	.32	.62
11. Nepagilmat	.11	.32	.40	.62
12. Sakenboltich	.41	.50	.46	.60
13. Rolkalensbar	.22	.42	.14	.65
14. Wapsitufelan	.19	.40	.51	.60
15. Mufagoberi	.30	.47	.46	.60
16. Ilbokleutafon	.30	.47	.21	.64
17. Kuzelbratonis	.11	.32	.31	.63
18. Roplehatisolut	n/a	n/a	n/a	n/a
19. Nusalefaberit	.11	.32	.10	.65
20. Pildokeluwasor	n/a	n/a	n/a	n/a

Anmerkung. $n = 31$, n/a = nicht anwendbar/verfügbar, p_i = Itemschwierigkeit, $\text{Var}(x_i)$ = Itemvarianz, r_{it} = Itemtrennschwärfe.

Tabelle G7

Itemstatistiken - Itemset 2, Bewertungsmodus E

Item	p_i	$\text{Var}(x_i)$	r_{it}	Cronbachs Alpha wenn Item gelöscht
1. Nebluf	.63	.49	-.08	.68
2. Fendal	.78	.42	.46	.61
3. Lipnat	.59	.50	.23	.64
4. Wolbem	.56	.51	-.08	.68
5. Kubonheit	.74	.45	.09	.65
6. Mergulbisch	.37	.49	.41	.61
7. Gliwantas	.15	.36	.11	.65
8. Listorkel	.37	.49	.23	.64
9. Fradelponk	.30	.47	.27	.63
10. Brufilsakeit	.41	.50	.32	.62
11. Nepagilmat	.11	.32	.40	.62
12. Sakenboltich	.41	.50	.46	.60
13. Rolkalensbar	.22	.42	.14	.65
14. Wapsitufelan	.19	.40	.51	.60
15. Mufagoberi	.30	.47	.46	.60
16. Ilbokleutafon	.30	.47	.21	.64
17. Kuzelbratonis	.11	.32	.31	.63
18. Roplehatisolut	n/a	n/a	n/a	n/a
19. Nusalefaberit	.11	.32	.10	.65
20. Pildokeluwasor	n/a	n/a	n/a	n/a

Anmerkung. $n = 29$, n/a = nicht anwendbar/verfügbar, p_i = Itemschwierigkeit, $\text{Var}(x_i)$ = Itemvarianz, r_{it} = Itemtrennschwärfe.

Tabelle G8

Itemstatistiken - Itemset 2, Bewertungsmodus H

Item	p_i	$\text{Var}(x_i)$	r_{it}	Cronbachs Alpha wenn Item gelöscht
1. Nebluf	.63	.49	-.09	.68
2. Fendal	.78	.42	.49	.61
3. Lipnat	.59	.50	.30	.63
4. Wolbem	.56	.51	-.10	.68
5. Kubonheit	.78	.42	.06	.66
6. Mergulbisch	.37	.49	.33	.62
7. Gliwantas	.15	.36	.21	.64
8. Listorkel	.41	.50	.21	.64
9. Fradelponk	.37	.49	.33	.62
10. Brufilsakeit	.41	.50	.29	.63
11. Nepagilmat	.11	.32	.39	.62
12. Sakenboltich	.44	.51	.52	.60
13. Rolkalensbar	.26	.45	.09	.66
14. Wapsitufelan	.19	.40	.45	.61
15. Mufagoberi	.37	.49	.36	.62
16. Ilbokleutafon	.33	.48	.35	.62
17. Kuzelbratonis	.11	.32	.30	.63
18. Roplehatisolut	n/a	n/a	n/a	n/a
19. Nusalefaberit	.11	.32	.10	.65
20. Pildokeluwasor	n/a	n/a	n/a	n/a

Anmerkung. $n = 29$, n/a = nicht anwendbar/verfügbar, p_i = Itemschwierigkeit, $\text{Var}(x_i)$ = Itemvarianz, r_{it} = Itemtrennschwärfe.

Tabelle G9

Itemstatistiken - Itemset 2, Bewertungsmodus EK

Item	p_i	$\text{Var}(x_i)$	r_{it}	Cronbachs Alpha wenn Item gelöscht
1. Nebluf	.63	.49	-.12	.65
2. Fendal	.85	.36	.44	.57
3. Lipnat	.63	.49	.18	.60
4. Wolbem	.59	.50	-.17	.66
5. Kubonheit	.74	.45	.13	.61
6. Mergulbisch	.52	.51	.41	.57
7. Gliwantas	.30	.47	.06	.62
8. Listorkel	.41	.50	.01	.63
9. Fradelponk	.30	.47	.30	.59
10. Brufilsakeit	.44	.51	.62	.53
11. Nepagilmat	.11	.32	.42	.58
12. Sakenboltich	.41	.50	.48	.56
13. Rolkalensbar	.22	.42	.14	.61
14. Wapsitufelan	.19	.40	.43	.57
15. Mufagoberi	.33	.48	.43	.56
16. Ilbokleutafon	.30	.47	.14	.61
17. Kuzelbratonis	.11	.32	.25	.60
18. Roplehatisolut	n/a	n/a	n/a	n/a
19. Nusalefaberit	.11	.32	.07	.61
20. Pildokeluwasor	n/a	n/a	n/a	n/a

Anmerkung. $n = 29$, n/a = nicht anwendbar/verfügbar, p_i = Itemschwierigkeit, $\text{Var}(x_i)$ = Itemvarianz, r_{it} = Itemtrennschwärfe.

Tabelle G10

Itemstatistiken - Itemset 2, Bewertungsmodus EKH

Item	p_i	$\text{Var}(x_i)$	r_{it}	Cronbachs Alpha wenn Item gelöscht
1. Nebluf	.67	.48	-.12	.64
2. Fendal	.89	.32	.40	.57
3. Lipnat	.63	.49	.29	.58
4. Wolbem	.59	.50	-.20	.65
5. Kubonheit	.78	.42	.07	.61
6. Mergulbisch	.52	.51	.35	.57
7. Gliwantas	.30	.47	.16	.60
8. Listorkel	.44	.51	-.01	.63
9. Fradelponk	.37	.49	.39	.56
10. Brufilsakeit	.48	.51	.52	.54
11. Nepagilmat	.11	.32	.41	.57
12. Sakenboltich	.44	.51	.52	.54
13. Rolkalensbar	.26	.45	.07	.61
14. Wapsitufelan	.19	.40	.35	.57
15. Mufagoberi	.41	.50	.36	.57
16. Ilbokleutafon	.33	.48	.24	.59
17. Kuzelbratonis	.11	.32	.23	.59
18. Roplehatisolut	n/a	n/a	n/a	n/a
19. Nusalefaberit	.11	.32	.06	.61
20. Pildokeluwasor	n/a	n/a	n/a	n/a

Anmerkung. $n = 29$, n/a = nicht anwendbar/verfügbar, p_i = Itemschwierigkeit, $\text{Var}(x_i)$ = Itemvarianz, r_{it} = Itemtrennschwärfe.

Altersgruppenspezifische Itemkennwerte zu Bewertungsmodus St vor Itemselektion.

Tabelle G11

Altersgruppenspezifische Itemstatistiken - Itemset 1, Bewertungsmodus St

Item	3 Jahre ^a			4 Jahre ^b			5 Jahre ^c		
	p_i	$\text{Var}(x_i)$	r_{it}	p_i	$\text{Var}(x_i)$	r_{it}	p_i	$\text{Var}(x_i)$	r_{it}
1. Muldop	.57	.53	.08	n/a	n/a	n/a	n/a	n/a	n/a
2. Bafmil	.71	.49	.00	.76	.44	.20	n/a	n/a	n/a
3. Labdin	.57	.53	.08	.71	.47	.20	.71	.49	.54
4. Temfol	.43	.53	.78	.88	.33	-.09	n/a	n/a	n/a
5. Profaling	.14	.38	.42	.59	.51	.17	.71	.49	.74
6. Dranigsche	.14	.38	.12	.71	.47	.25	.71	.49	.54
7. Blesamder	.14	.38	.12	.35	.49	-.09	.43	.53	-.24
8. Nesbölung	.71	.49	-.20	.47	.51	.14	.86	.38	.74
9. Knoseptist	n/a	n/a	n/a	.29	.47	.35	.43	.53	.49
10. Fadrilusen	n/a	n/a	n/a	.47	.51	.38	.43	.53	.67
11. Mirefkonal	n/a	n/a	n/a	.41	.51	.37	.43	.53	.49
12. Frudalisbek	.14	.38	.42	.53	.51	.06	.71	.49	.74
13. Glomerupftig	n/a	n/a	n/a	.12	.33	.06	.43	.53	.32
14. Tinguforasel	n/a	n/a	n/a	.53	.51	.39	.71	.49	.34
15. Dulkamosiner	.14	.38	-.38	.24	.44	.11	.14	.38	.58
16. Laponiteful	n/a	n/a	n/a	.24	.44	.01	.43	.53	.67
17. Birlostaketu	n/a	n/a	n/a	.18	.39	.38	.43	.53	-.01
18. Pekisafuloricht	n/a	n/a	n/a	.12	.33	.12	.43	.53	.32
19. Sutalezibokan	n/a	n/a	n/a	.06	.24	.51	.14	.38	-.11
20. Lonbarenopigum	n/a	n/a	n/a	.06	.24	.51	n/a	n/a	n/a

Anmerkung. n/a = nicht anwendbar/verfügbar, p_i = Itemschwierigkeit, $\text{Var}(x_i)$ = Itemvarianz, r_{it} = Itemtrennschwärfe.

^a $n = 7$, ^b $n = 17$, ^c $n = 7$.

Altersgruppenspezifische Itemkennwerte zu Bewertungsmodus St vor Itemselektion.

Tabelle G12

Altersgruppenspezifische Itemstatistiken - Itemset 2, Bewertungsmodus St

Item	3 Jahre ^a			4 Jahre ^b			5 Jahre ^c		
	p_i	$\text{Var}(x_i)$	r_{it}	p_i	$\text{Var}(x_i)$	r_{it}	p_i	$\text{Var}(x_i)$	r_{it}
1. Nebluf	.50	.52	-.13	.78	.44	-.10	.67	.52	-.55
2. Fendal	.75	.45	.35	.67	.50	.53	n/a	n/a	n/a
3. Lipnat	.42	.51	.39	.67	.50	-.36	.83	.41	.06
4. Wolbem	.50	.52	.14	.56	.53	-.71	.67	.52	.21
5. Kubonheit	.83	.39	.06	.56	.53	.10	.83	.41	.49
6. Mergulbisch	.17	.39	.28	.44	.53	.68	.67	.52	-.74
7. Gliwantas	.17	.39	.38	.11	.33	-.03	.17	.41	-.28
8. Listorkel	.33	.49	.36	.33	.50	.31	.50	.55	-.56
9. Fradelponk	.25	.45	.02	.22	.44	.71	.50	.55	-.36
10. Brufilsakeit	.17	.39	.09	.56	.53	.20	.67	.52	.21
11. Nepagilmat	.08	.29	.64	.11	.33	.26	.17	.41	.49
12. Sakenboltich	.42	.51	.31	.11	.33	.58	.83	.41	.49
13. Rolkalensbar	n/a	n/a	n/a	.22	.44	-.13	.50	.55	-.36
14. Wapsitufelan	.08	.29	-.01	.11	.33	.58	.50	.55	-.36
15. Mufagoberi	.25	.45	.63	.22	.44	.572	.67	.52	-.09
16. Ilbokleutafon	.25	.45	.35	.22	.44	-.232	.50	.55	.15
17. Kuzelbratonis	n/a	n/a	n/a	.22	.44	.707	.17	.41	-.28
18. Roplehatisolut	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
19. Nusalefaberit	n/a	n/a	n/a	.22	.44	.086	.17	.41	-.28
20. Pildokeluwasor	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Anmerkung. n/a = nicht anwendbar/verfügbar, p_i = Itemschwierigkeit, $\text{Var}(x_i)$ = Itemvarianz, r_{it} = Itemtrennschwärfe.

^a $n = 12$, ^b $n = 9$, ^c $n = 6$.

Itemkennwerte nach Itemselektion.

Tabelle G13

Itemstatistiken – Itemset 1 nach Itemselektion, Bewertungsmodus St

Item	p_i	$\text{Var}(x_i)$	r_{it}	Cronbachs Alpha wenn Item gelöscht
1. Muldop	.90	.30	.46	.80
4. Temfol	.81	.40	.39	.80
5. Profaling	.52	.51	.52	.79
6. Dranigsche	.58	.50	.52	.79
9. Knoseptist	.26	.45	.51	.79
10. Fadrilusen	.35	.49	.52	.79
11. Mirefkonal	.32	.48	.52	.79
12. Frudalisbek	.48	.51	.53	.79
14. Tinguforasel	.45	.51	.49	.79
16. Laponiteful	.23	.43	.38	.80
17. Birlostaketu	.19	.40	.36	.80
18. Pekisafuloricht	.16	.37	.29	.81

Anmerkung. $n = 31$, p_i = Itemschwierigkeit, $\text{Var}(x_i)$ = Itemvarianz, r_{it} = Itemtrennschwärfe.

Altersgruppenspezifische Itemkennwerte zu Bewertungsmodus St nach Itemselektion.

Tabelle G14

Altersgruppenspezifische Itemstatistiken - Itemset 1, Bewertungsmodus St

Item	3 Jahre ^a			4 Jahre ^b			5 Jahre ^c		
	p_i	$\text{Var}(x_i)$	r_{it}	p_i	$\text{Var}(x_i)$	r_{it}	p_i	$\text{Var}(x_i)$	r_{it}
1 Muldop	.57	.54	.40	n/a	n/a	n/a	n/a	n/a	n/a
4 Temfol	.43	.54	.54	.88	.33	.86	n/a	n/a	n/a
5 Profaling	.14	.38	.60	.59	.51	.71	.71	.49	.82
6 Dranigsche	.14	.38	.23	.71	.47	.78	.71	.49	.54
9 Knoseptist	n/a	n/a	n/a	.53	.51	.81	.43	.54	.44
10 Fadrilusen	n/a	n/a	n/a	.29	.47	.70	.43	.54	.69
11 Mirefkonal	n/a	n/a	n/a	.47	.51	.74	.43	.54	.56
12 Frudalisbek	.14	.38	.60	.41	.51	.86	.71	.49	.82
14 Tinguforasel	n/a	n/a	n/a	.53	.51	.75	.71	.49	.16
16 Laponiteful	n/a	n/a	n/a	.24	.44	.71	.43	.54	.69
17 Birlostaketu	n/a	n/a	n/a	.18	.39	.86	.43	.54	.10
18 Pekisafuloricht	n/a	n/a	n/a	.12	.33	.58	.43	.54	.10

Anmerkung. n/a = nicht anwendbar/verfügbar, p_i = Itemschwierigkeit, $\text{Var}(x_i)$ = Itemvarianz, r_{it} = Itemtrennschwärfe.

^a $n = 7$, ^b $n = 17$, ^c $n = 7$.

Anhang H: Beurteilerinnenübereinstimmung und Intracoder-Reliabilität

Tabelle H1

Beurteilerinnenübereinstimmung für Itemset 1

Item	Bewertungsmodus											
	St				E				H			
	Fleiss κ	z	SE	p	Fleiss κ	z	SE	p	Fleiss κ	z	SE	p
1.	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
2.	0.17	0.72	0.24	.473	0.17	0.72	0.24	.473	0.20	0.85	0.24	.396
3.	0.68	2.88	0.24	.004	0.68	2.88	0.24	.004	0.68	2.88	0.24	.004
4.	0.77	3.25	0.24	.001	0.77	3.25	0.24	.001	0.68	2.88	0.24	.004
5.	0.77	3.25	0.24	.001	0.68	2.88	0.24	.004	0.77	3.25	0.24	.001
6.	0.56	2.36	0.24	.018	0.30	1.27	0.24	.205	0.56	2.36	0.24	.018
7.	1.00	4.24	0.24	<.001	0.78	3.29	0.24	.001	1.00	4.24	0.24	<.001
8.	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
9.	0.68	2.88	0.24	.004	0.68	2.88	0.24	.004	0.68	2.88	0.24	.004
10.	0.50	2.12	0.24	.034	0.68	2.88	0.24	.004	0.50	2.12	0.24	.034
11.	0.68	2.88	0.24	.004	0.68	2.88	0.24	.004	0.68	2.88	0.24	.004
12.	1.00	4.24	0.24	<.001	1.00	4.24	0.24	<.001	0.78	3.29	0.24	.001
13.	0.44	1.86	0.24	.063	0.44	1.86	0.24	.063	0.44	1.86	0.24	.063
14.	0.55	2.33	0.24	.020	0.55	2.33	0.24	.020	0.55	2.33	0.24	.020
15.	0.44	1.86	0.24	.063	0.44	1.86	0.24	.063	0.44	1.86	0.24	.063
16.	1.00	4.24	0.24	<.001	1.00	4.24	0.24	<.001	1.00	4.24	0.24	<.001
17.	0.44	1.86	0.24	.063	0.20	0.85	0.85	.396	0.44	1.86	0.24	.063
18.	0.66	2.55	0.26	.011	0.66	2.55	0.26	.011	0.66	2.55	0.26	.011
19.	1.00	3.87	0.26	<.001	1.00	3.87	0.26	<.001	1.00	3.87	0.26	<.001
20.	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Anmerkung. n Kinder = 6, n Kodiererinnen = 3, n/a = nicht anwendbar/verfügbar.

Tabelle H2

Beurteilerinnenübereinstimmung für Itemset 2

Item	Bewertungsmodus											
	St				E				H			
	Fleiss	z	SE	p	Fleiss	z	SE	p	Fleiss	z	SE	p
	κ				κ				κ			
1.	0.40	2.08	0.19	.038	0.40	2.08	0.19	.038	0.54	2.81	0.19	.005
2.	0.71	3.67	0.19	<.001	0.25	1.30	0.19	.194	0.71	3.67	0.19	<.001
3.	1.00	5.20	0.19	<.001	1.00	5.20	0.19	<.001	1.00	5.20	0.19	<.001
4.	0.85	4.40	0.19	<.001	0.85	4.40	0.19	<.001	0.85	4.40	0.19	<.001
5.	0.84	4.37	0.19	<.001	0.84	4.37	0.19	<.001	0.84	4.37	0.19	<.001
6.	0.69	3.60	0.19	<.001	0.70	3.64	0.19	<.001	0.85	4.43	0.19	<.001
7.	0.46	2.39	0.19	.017	0.61	3.19	0.19	.001	0.46	2.39	0.19	.017
8.	0.36	1.86	0.19	.063	0.36	1.86	0.19	.063	0.36	1.86	0.19	.063
9.	0.68	3.55	0.19	<.001	0.68	3.55	0.19	<.001	0.68	3.55	0.19	<.001
10.	0.47	2.43	0.19	.015	0.54	2.81	0.19	.005	0.47	2.43	0.19	.015
11.	1.00	5.20	0.19	<.001	1.00	5.20	0.19	<.001	1.00	5.20	0.19	<.001
12.	0.84	4.37	0.19	<.001	0.84	4.37	0.19	<.001	0.69	3.60	0.19	<.001
13.	1.00	5.20	0.19	<.001	1.00	5.20	0.19	<.001	1.00	5.20	0.19	<.001
14.	1.00	5.20	0.19	<.001	1.00	5.20	0.19	<.001	1.00	5.20	0.19	<.001
15.	0.42	2.19	0.19	.029	0.33	1.73	0.19	.083	0.29	1.50	0.19	.133
16.	0.26	1.37	0.19	.171	0.26	1.37	0.19	.171	0.26	1.37	0.19	.171
17.	1.00	5.20	0.19	<.001	1.00	5.20	0.19	<.001	1.00	5.20	0.19	<.001
18.	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
19.	0.66	2.28	0.29	.023	0.66	2.28	0.29	.023	0.66	2.28	0.29	.023
20.	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Anmerkung. n Kinder = 9, n Kodiererinnen = 3, n/a = nicht anwendbar/verfügbar.

Tabelle H3

Intracoder-Reliabilität: Übereinstimmung zwischen protokollierter und transkribierter Antwort (Bewertungsmodus St), Itemset 1

Item	Cohens κ	näherungs- weises T	SE	p
1. Muldop	1.00	5.75	0.00	<.001
2. Bafmil	0.89	5.15	0.11	<.001
3. Labdin	0.93	5.36	0.07	<.001
4. Temfol	0.74	4.28	0.14	<.001
5. Profaling	0.63	3.64	0.14	<.001
6. Dranigsche	0.82	4.79	0.10	<.001
7. Blesamder	0.80	4.61	0.11	<.001
8. Nesbölung	0.79	4.65	0.11	<.001
9. Knoseptist	1.00	5.75	0.00	<.001
10. Fadrilusen	0.75	4.45	0.11	<.001
11. Mirefkonal	0.93	5.37	0.07	<.001
12. Frudalisbek	0.82	4.79	0.10	<.001
13. Glomerupftig	0.89	5.15	0.11	<.001
14. Tinguforasel	0.82	4.69	0.10	<.001
15. Dulkamosiner	0.71	4.12	0.16	<.001
16. Laponiteful	0.76	4.40	0.13	<.001
17. Birlostaketu	0.90	5.22	0.09	<.001
18. Pekisafuloricht	0.79	4.19	0.14	<.001
19. Sutalezibokan	0.78	4.16	0.21	<.001
20. Lonbarenopigum	0.65	3.60	0.32	<.001

Anmerkung. n Kinder = 9, n Kodiererinnen = 3, n/a = nicht anwendbar/verfügbar.

Tabelle H4

Intracoder-Reliabilität: Übereinstimmung zwischen protokollierter und transkribierter Antwort (Bewertungsmodus St), Itemset 2

Item	Cohens κ	näherungs-wei- ses T	SE	p
1. Nebluf	0.91	0.09	4.67	<.001
2. Fendal	0.26	0.26	1.41	.158
3. Lipnat	0.72	0.14	3.84	<.001
4. Wolbem	0.53	0.17	2.69	<.001
5. Kubonheit	0.89	0.11	4.54	<.001
6. Mergulbisch	0.77	0.13	3.93	<.001
7. Gliwantas	0.81	0.13	4.19	<.001
8. Listorkel	0.77	0.13	3.92	<.001
9. Fradelponk	0.83	0.12	4.20	<.001
10. Brufilsakeit	0.77	0.12	4.04	<.001
11. Nepagilmat	0.83	0.16	4.23	<.001
12. Sakenboltich	0.76	0.13	3.92	<.001
13. Rolkalensbar	0.53	0.19	2.64	<.001
14. Wapsitufelan	0.90	0.10	4.51	<.001
15. Mufagoberi	0.74	0.14	3.78	<.001
16. Ilbokleutafon	0.61	0.15	3.18	.001
17. Kuzelbratonis	0.51	0.17	2.91	.004
18. Roplehatisolut	n/a	n/a	n/a	n/a
19. Nusalefaberit	0.82	0.17	3.33	.001
20. Pildokeluwasor	1.00	0.00	4.00	<.001

*Anmerkung.*n/a = nicht anwendbar/verfügbar.

Anhang I: Vorschläge zur Modifikation der Items

Tabelle I1

Vorschlag für Modifikationen von Items, Vorgabereihenfolge und Vorgabe für Altersgruppen

Silben- anzahl	Item	Vorgabe für Altersgruppen		
		3 Jahre	4 Jahre	5 Jahre
2	1 Muldop	x		
	4 Temfol	x	x	
3	5 Dalnigsche	x	x	x
	6 Profaling	x	x	x
	9 Knoseptist		x	x
4	10 Fadrilusan	x	x	x
	11 Mirefkonal		x	x
	12 Frudalisbek		x	x
5	14 Tinguforapsil		x	x
	16 Lasponiteful		x	x
	17 Birlostaketum		x	x
6	18 Pekisafuloricht		x	x