

Institut und Poliklinik für Kinder- und Jugendpsychiatrie und Psychotherapie der Ludwig-  
Maximilians-Universität München

Kommissarischer Direktor: Prof. Dr. med. H.-J. Möller

**WIE VERGLEICHBAR SIND ERGEBNISSE VON RECHTSCHREIB-  
UND LESETESTS?**

Dissertation

zum Erwerb des Doktorgrades der Medizin

an der Medizinischen Fakultät der

Ludwig-Maximilians-Universität zu München

vorgelegt von  
Susanne Steinhauser  
aus München  
2006

Mit Genehmigung der Medizinischen Fakultät  
der Universität München

**Berichterstatter:** Prof. Dr. med. W. v. Suchodoletz

**Mitberichterstatter:** Prof. Dr. R. Engel  
Prof. Dr. R. v. Kries

**Dekan:** Prof. Dr. med. D. Reinhardt

**Tag der mündlichen Prüfung:** 06.07.2006

<b>1. EINLEITUNG.....</b>	<b>5</b>
1.1. Die Lese- und Rechtschreibstörung.....	5
1.1.1. Diagnostische Kriterien nach ICD-10 (F81).....	6
1.1.2. Klinische Relevanz einer Lese- Rechtschreibstörung.....	7
<b>2. THEORETISCHER HINTERGRUND.....</b>	<b>10</b>
2.1. Diagnostik einer LRS.....	10
2.1.1. Diagnostisches Vorgehen in der Praxis.....	10
2.1.2. Probleme bei der operationalisierten Diagnostik einer LRS.....	13
2.2. Standardisierte Testverfahren zur Überprüfung von Rechtschreib- und Lesefähigkeit und der Intelligenz.....	14
<b>3. FRAGESTELLUNG.....</b>	<b>21</b>
<b>4. METHODE.....</b>	<b>22</b>
4.1. Stichprobe.....	22
4.1.1. Rekrutierung.....	22
4.1.2. Einschlusskriterien und Ausschlusskriterien.....	22
4.1.3. Beschreibung der Stichprobe.....	23
4.2. Fragebögen.....	25
4.3. Testverfahren.....	26
4.3.1. Der Diagnostische Rechtschreibetest, DRT 1 bis 5.....	26
4.3.2. Die Weingartener Grundwortschatz Rechtschreib-Testreihe, WRT1+ bis WRT3+.....	28
4.3.3. Der Westermann Rechtschreibtest, WRT4/5 und WRT 6+.....	30
4.3.4. Der Salzburger Lese-Rechtschreibtest, SLRT.....	31
4.3.5. Die Hamburger Schreib-Probe, HSP 1+ bis 5-9.....	32
4.3.6. Der Zürcher Lesetest, Zürcher.....	33
4.3.7. Die Würzburger Leise Leseprobe, WLLP.....	34
4.3.8. Knuspels Leseaufgaben, KNUSPEL-L.....	34
4.3.9. Grundintelligenztest Skala 1, CFT1 und Skala 2, CFT20.....	35
4.4. Studienablauf.....	36
4.4.1. Anwendung der Testverfahren.....	36
4.4.2. Testungsabfolge und Parallelformen.....	38

4.4.3. Durchführung der Testverfahren.....	38
4.5. Statistik.....	39
<b>5. ERGEBNISSE.....</b>	<b>42</b>
5.1. Erhält man mit den Rechtschreibtest gleiche Ergebnisse?.....	42
5.1.1. Unterscheiden sich bei der HSP die Auswertung nach Worttreffer und die Auswertung nach Graphemtreffer in ihren Ergebnissen?.....	42
5.1.2. Wie unterscheiden sich die Rechtschreibtests verschiedener Klassenstufen?..	43
5.1.3. Findet sich ein Zusammenhang zwischen DRT 1, WRT 1+, HSP 1+ u. dem Erreichen des kritischen Wertes für nicht-lauttreue Fehler beim SLRT 1?.....	50
5.2. Erhält man mit den Lesetests gleiche Ergebnisse?.....	51
5.2.1. Erbringen die Subtests des SLRT, sowie die Subtests des Züricher gleiche Ergebnisse?.....	51
5.2.2. Wie unterscheiden sich die Lesetests verschiedener Klassenstufen?.....	54
5.3. Kommen Lehrer, Eltern und Testverfahren bei der Einschätzung von Schul- leistungen zu gleichen Ergebnissen?.....	59
5.3.1. Finden sich Unterschiede in der Leistungseinschätzung durch die Eltern und durch die Lehrer?.....	59
5.3.2. Finden sich Unterschiede in der Leistungseinschätzung durch die Eltern, bzw. durch die Lehrer und durch standardisierte Testverfahren?.....	60
5.4. Wie hoch ist die diagnostische Zuverlässigkeit von LRS-Tests zu bewerten?.....	62
5.4.1. Wie stark hängt die Diagnose „Lese-Rechtschreibstörung, F81“ vom eingesetzten Testverfahren ab?.....	62
5.4.2. Wie hoch ist die Trefferquote einzelner LRS-Tests einzuschätzen?.....	63
<b>6. DISKUSSION.....</b>	<b>70</b>
6.1. Zuverlässigkeit der gewählten Methode.....	70
6.2. Vergleichbarkeit der Ergebnisse von LRS-Tests.....	75
6.3. Die Übereinstimmung zwischen Testergebnissen, Lehrerurteil und Elternurteil.....	86
6.4. Bewertung der diagnostischen Zuverlässigkeit von LRS-Tests.....	92
6.5. Schlussfolgerung für die Praxis.....	98
<b>7. ZUSAMMENFASSUNG.....</b>	<b>103</b>
<b>8. ABKÜRZUNGSVERZEICHNIS.....</b>	<b>109</b>
<b>9. LITERATURVERZEICHNIS.....</b>	<b>110</b>
<b>10. ANHANG.....</b>	<b>121</b>

# 1. EINLEITUNG

## 1.1. Die Lese- und Rechtschreibstörung

Lernstörungen stellen in unserer modernen immer komplexer werdenden Gesellschaft ein relevantes Problem dar, was weit reichende sozioökonomische Folgen mit sich bringt.

Schon derzeit bestehen, trotz hoher Arbeitslosigkeit in Deutschland, Probleme Arbeitsplätze mit höheren Qualifikationsanforderungen zu besetzen. Entsprechend intensiv war die Reaktion der Öffentlichkeit auf die Ergebnisse der PISA-Studie, die für das deutsche Bildungswesen in den Grundfertigkeiten Lesen und Rechnen einen als unbefriedigend empfundenen unterdurchschnittlichen Platz ergab (Deutsches PISA-Konsortium, 2001).

Der Begriff "Legasthenie" wurde 1916 von Paul Ranschburg als Synonym für Leseschwäche eingeführt. Heute unterscheidet das "Internationale Klassifikationsschema für psychische Störungen (ICD-10)" der Weltgesundheitsorganisation (WHO) die Isolierte Rechtschreibstörung und die Lese-Rechtschreibstörung.

Bei der Lese- Rechtschreibstörung liegt eine relevante Beeinträchtigung des Lese- und Rechtschreibprozesses, bei ansonsten befriedigenden schulischen Leistungen und normaler bzw. überdurchschnittlicher Intelligenz vor, die sich typischerweise vor dem 6. Schuljahr manifestiert.

Die Lese-Rechtschreibstörung ist die häufigste Entwicklungsstörung des Kindesalters, die auf die schulische und soziale Entwicklung erheblichen Einfluss nimmt. Trotz ausreichend Unterricht und kognitiven Fähigkeiten können einfache Wörter nicht gelesen und (ohne Buchstabenverwecheln, Buchstabeninversion, Silbenweglassen und Regelfehler) nicht geschrieben werden (Esser & Schmidt 2002<sup>2</sup>, Huber & Gross 2005).

Häufig gehen der Lese- Rechtschreibstörung im Vorschulalter primäre Entwicklungsstörungen des Sprechens und der Sprache voraus. Danach stehen die Leseschwierigkeiten (Worte lesen zu lernen oder gelesene Worte wieder zu erkennen und sie vorzulesen) im Vordergrund. Meist ist auch das Erlernen der Rechtschreibung erschwert.

Je älter die Kinder werden, desto mehr dominiert die Rechtschreibstörung. Das Lesen kann häufig ausreichend erlernt werden, bleibt im Vergleich zur altersgleichen Schulpopulation aber verlangsamt. Die Schulnoten bewegen sich im Bereich zwischen mangelhaft und ungenügend.

Die Lese-Rechtschreibstörung ist ein Störungsbild, das wesentlich die persönliche, emotionale und soziale Entwicklung der Betroffenen prägt. Durch die hohe Persistenz der Störung (Klicpera & Gasteiger Klicpera 1993, Bruck 1998, Shaywitz et al. 1999) wird die persönliche und soziale Entwicklung bis ins Erwachsenenalter maßgeblich beeinflusst (Haffner et al., 1998).

Sekundäre Komorbidität wie Verhaltensstörungen sowie emotionale und psychosomatische Störungen sind während der Schulzeit häufig.

Als Ursache einer LRS muss eine psychische oder organische Erkrankung ausgeschlossen sein. Auch eine geistige Behinderung oder eine Verzögerung in der Aneignung der Kulturtechniken Lesen und Schreiben, häufig vor dem Hintergrund sozioökonomischer Benachteiligung oder Belastung, darf für die Diagnose LRS nicht vorliegen.

### **1.1.1. Diagnostische Kriterien nach ICD-10 (F81)**

In der Internationalen Klassifikation von Erkrankungen (ICD 10) wird die Lese und Rechtschreibstörung im Abschnitt V. (Psychische und Verhaltensstörungen) unter dem Diagnoseschlüssel F 81 wie folgt definiert:

„Das Hauptmerkmal ist eine umschriebene und eindeutige Beeinträchtigung in der Entwicklung der Lesefertigkeiten, die nicht allein durch das Entwicklungsalter, Visusprobleme oder unangemessene Beschulung erklärbar ist. Das Leseverständnis, die Fähigkeit, gelesene Worte wieder zu erkennen, vorzulesen und Leistungen bei Aufgaben, für welche Lesefähigkeit benötigt werden, können sämtlich betroffen sein. Mit Lesestörungen gehen häufig Rechtschreibstörungen einher. Diese persistieren oft bis in die Adoleszenz, auch wenn einige Fortschritte im Lesen gemacht wurden. Umschriebenen Entwicklungsstörungen des Lesens gehen Entwicklungsstörungen des Sprechens oder der Sprache voraus. Während der Schulzeit sind begleitende Störungen im emotionalen und Verhaltensbereich häufig (...). Die Leseleistungen

des Kindes müssen unter dem Niveau liegen, das auf Grund des Alters, der altersgemeinen Intelligenz und der Beschulung zu erwarten ist (...). In der späteren Kindheit und im Erwachsenenalter sind die Rechtschreibprobleme meist größer als Defizite in der Lesefähigkeit“.

### **1.1.2. Klinische Relevanz einer Lese- Rechtschreibstörung**

Kinder, die durch eine Rechtschreib- und Leseschwäche in ihrer schulischen Leistung beeinträchtigt sind, nehmen im Patientengut der Kinder- und Jugendpsychiatrie einen großen Anteil ein.

Es gibt verschiedene Definitionsmöglichkeiten einer LRS, die zu unterschiedlich hohen Häufigkeiten führen. Falldefinitionen, die sich an inhaltlichen Kriterien orientieren, zeigen eine Prävalenz der LRS zwischen 6 - 9 % (Strehlow and Haffner, 2002). WARNKE spricht in seinem aktuellen Leitfaden davon, dass von einer Prävalenz für die Schülerpopulation zwischen 4 - 8 % auszugehen sei. Aber auch er spricht von großen Schwankungen der Häufigkeit je nach Definitionskriterien (Warnke et al., 2004). Unter den „II. Bekanntmachungen der Bayerischen Staatsministerien für Unterricht und Kultus und Wissenschaft, Forschung und Kunst, 1999“ wird eine Prävalenz von 7 - 10 % aller Schüler im Einschulungsalter, die Schwierigkeiten beim Erlernen des Lesens haben, beschrieben. Es handelt sich somit um eine häufige Erkrankung im Kindes- und Jugendalter.

Die Fähigkeiten des Rechtschreibens und Lesens werden in unserer modernen, vor allem leistungsorientierten Gesellschaft von allen Seiten gefordert. Es konnte sogar gezeigt werden, dass die Rechtschreibleistung eine höhere Korrelation mit dem Schulerfolg zeigt als der nichtsprachliche IQ. Lediglich ein Drittel der Kinder mit einer Rechtschreibschwäche erreicht den Übergang in weiterführende Schulen, nur etwa 7 - 9 % absolvieren einen Realschulabschluss und ganze 2 % einen Hochschulabschluss (Haffner et al., 1998). Die Schullaufbahnen entsprechen in Ihrem Niveau etwa denen von minderbegabten Kindern in einem IQ-Bereich von 70 – 85 % (Schmid, 2004b).

Entsprechend lässt sich zusammenfassen, dass eine LRS im Grundschulalter langfristig mit einer ungünstigen Prognose für den schulischen und den beruflichen Erfolg, sowie auch für die seelische Entwicklung verbunden ist.

Nachteilig wirken sich immer eine verzögerte Diagnosestellung und schlechte sozioökonomische Lebensverhältnisse aus (Klicpera & Gasteiger Klicpera, 1995). Mitunter spielt hierbei auch die hohe Persistenz der LRS, die in Untersuchungen belegt wurde, eine entscheidende Rolle (Klicpera et al., 1993), Bruck 1998, Strehlow 1998, Shaywitz 1999).

Im Rahmen einer Therapie können durch gezielte Übungsbehandlungen die Lese- und Rechtschreibfähigkeiten verbessert werden, so dass die meisten Kinder und Jugendlichen ausreichend das Lesen und auch hinreichend verständlich zu schreiben lernen (Warnke 2002, v. Suchodoletz 2003, Warnke 2003). Bei einer Nachuntersuchung einer Stichprobe von lese-rechtschreibgestörten Kindern im Erwachsenenalter, deren psychische und soziale Entwicklung einen sehr guten Verlauf genommen hat, geht man davon aus, dass neben dem hohen IQ und dem hohen Berufsprestige der Väter, sich die lang andauernde Förderung positiv auf die Entwicklung ausgewirkt hat (Schulte-Körne et al., 2003).

Neben den Übungsbehandlungen ist ein sehr wichtiger Therapiebaustein die Psychoedukation (Leitlinien zur Diagnostik und Therapie von psychischen Störungen im Säuglings-, Kindes- und Jugendalter, 2003), schulische Interventionen und bei entsprechender Komorbidität eine frühe psychotherapeutische Behandlung.

Eine frühzeitige Diagnosestellung ist zur Vermeidung, bzw. für eine gezielte Therapie sekundärer Komorbiditäten, wie Aufmerksamkeits- und Konzentrationsstörungen, motorische Unruhe, Vermeidungsverhalten bei wachsendem Störungsbewusstsein (depressiv-ängstliche, oppositionell-aggressiv) und Schulangst, die sich bei Kindern mit schwerer Lese- und Rechtschreibstörung äußerst schnell entwickeln können, dringend erforderlich (Eggers et al., 2004). Grundschüler mit Lese-Rechtschreibschwierigkeiten sind im Vergleich zu ihren Mitschülern doppelt so häufig psychisch auffällig (Esser et al., 2002).

Der hohe Grad der komorbiden Auffälligkeiten erfordert eine umfassende und sorgfältige kinder- und jugendpsychiatrische Untersuchung, um die notwendigen Hinweise für entsprechende Förderung und Hilfestellung zu bekommen.

Die Diagnostik beginnt mit der psychiatrischen Exploration, dem klinischen Untersuchungsbefund, gegebenenfalls ergänzenden apparativen Untersuchungen, und die Einsicht in Schulhefte und Zeugnisse.



Für die Diagnosestellung unverzichtbar ist eine präzise testdiagnostische Untersuchung. Eine Diagnostik mit standardisierten Testverfahren ist eine geforderte Voraussetzung für eine exakte Diagnose und eine spezielle auf Schwächen orientierte Behandlung der Lese- Rechtschreibstörung (Deutsche Leitlinien für Kinder- und Jugendpsychiatrie und Psychotherapie 2003, Warnke et al. 2004).

Es gibt eine Vielzahl von Testverfahren, die im Praxisalltag angewandt werden, die sich aber in ihren Testkonzepten und Gütekriterien wesentlich unterscheiden.

Es soll nun im Rahmen einer Studie überprüft werden, ob unterschiedliche Tests zur Erfassung einer Lese-Rechtschreibstörung vergleichbare Ergebnisse erbringen und in welchem Ausmaß die Stellung der Diagnose „Lese-Rechtschreibstörung“ durch die Testauswahl beeinflusst werden kann.

## 2. THEORETISCHER HINTERGRUND

### 2.1. Diagnostik einer LRS

#### 2.1.1. Diagnostisches Vorgehen in der Praxis

Die Deutsche Gesellschaft für Kinder- und Jugendpsychiatrie und Psychotherapie empfiehlt die multiaxiale Diagnostik, wie sie ganz allgemein zur Diagnostik psychischer Störungen im Kindes- und Jugendalter angewandt wird (Remschmidt, 2006).

Diese Vorgehensweise bildet die Basis – in Bayern und in anderen Bundesländern – für das ärztliche Gutachten zur Eingliederungshilfe nach § 35a SGB VIII. Es kann auch als Gutachten zur Vorlage in der Schule verwendet werden, damit die Bestimmungen des Legastheneierlasses des Bayerischen Staatsministeriums für Unterricht und Kultus vom 16. November 1999 für den betroffenen Schüler in Kraft treten können (II. Bekanntmachungen der Bayerischen Staatsministerien für Unterricht und Kultus und Wissenschaft, Forschung und Kunst).

Untersuchungsbereiche(Achsen) der multiaxialen Diagnostik
<ul style="list-style-type: none"><li>• Psychische Gesundheit (klinisch-psychiatrische Befindlichkeit, Achse 1)</li><li>• Lesen und Rechtschreiben sowie weitere umschriebene Entwicklungsstörungen, nämlich die der Sprache, Motorik und des Rechnens (Achse 2)</li><li>• Intelligenz (Intelligenzniveau, Achse 3)</li><li>• Körperliche Gesundheit (die körperliche und neurologische Entwicklung, Achse 4)</li><li>• Psychosoziales Lebensmilieu (assoziierte aktuelle psychosoziale Umstände, Achse 5)</li><li>• Psychosoziale Anpassung und Eingliederung (die globale Beurteilung des psychosozialen Funktionsniveaus, Achse 6).</li></ul>

Die sich daraus ergebenden Leitlinien regulieren die Exploration der Eltern und Lehrer, die Exploration des Kindes/des Jugendlichen, die Verwendung standardisierter Fragebögen für Eltern, für das Kind/den Jugendlichen und für Lehrer sowie schließlich den Einsatz testpsychologischer Verfahren.

Beim Diktieren von Testwörtern (z.B. im Rahmen einer testpsychologischen Untersuchung) kann sich das Kind auf die einzelnen Wörter konzentrieren und das Ergebnis kann dadurch eventuell besser ausfallen als die Rechtschreibung in der Schule, wo ganze Texte geschrieben werden bzw. Aufsätze. Die Rechtschreibung ist in diesen Fällen für den Schüler häufig zweitrangig, da er sich primär auf den Inhalt konzentriert und die Rechtschreibung vernachlässigt

(Noterdaeme & Breuer-Schaumann 2003, Warnke et al. 2002). Aus diesem Grund sind die Ergebnisse der Exploration zum schulischen Werdegang unabdingbar. Hierunter fallen die aktuelle Symptomatik im Lesen und Rechtschreiben (Grad der Beeinträchtigung/Belastung des Patienten und seines Umfeldes) und die Schullaufbahn (u. a. Zeugnisbeurteilung im Lesen und Diktat diskrepant zu anderen Schulleistungen, auffälliges Allgemeinverhalten im Unterricht).

Ein weiterer Schritt ist die Abklärung psychischer Komorbiditäten und möglicher Differentialdiagnosen. Unter häufige Begleiterkrankungen, die primär oder sekundär auftreten können, fallen umschriebene Entwicklungsstörungen (z. B. die Rechenstörung, Sprachstörung), tiefgreifende Entwicklungsstörungen, die Hyperkinetische Störung – Aufmerksamkeitsstörung, Emotionale Störungen, Störungen des Sozialverhaltens, Angst und andere Störungen (z. B. Schlafstörungen). Eine übersichtliche Tabelle über Sekundärsymptomatiken findet sich in einem Übersichtsartikel von NIEBERGALL (1987).

Durch die Interaktion von Primärstörung, Schulversagen und sekundärer Überformung entwickelt sich häufig ein sehr komplexes Störungsbild, dessen exakte Einordnung erhebliche diagnostische Schwierigkeiten und Unsicherheiten aufwirft.

Ausreichend Möglichkeiten zum Wissenserwerb müssen gegeben sein. Wichtig ist die Abgrenzung von einer allgemeinen Retardierung infolge eines Mangelmilieus (sozioökologischer Benachteiligung bzw. Belastung).

Organische Erkrankungen oder eine geistige Behinderung müssen ausgeschlossen werden. Bei der Lese- und Rechtschreibstörung ist der Erwerb vom frühen Entwicklungsalter an beeinträchtigt im Gegensatz zu einem Verlust einer bereits erworbenen Fertigkeit. In Verbindung mit der körperlichen-neurologischen Untersuchung ist immer eine augenärztliche Untersuchung indiziert. Bei Auffälligkeiten in der Sprachentwicklung muss die Hörleistung überprüft werden.

Unverzichtbar für eine LRS-Diagnose sind schließlich die testpsychologischen Untersuchungen zum Niveau des Lesens und Rechtschreibens, sowie der Intelligenzentwicklung.

Nach den Grundsätzen der evidenzbasierten Medizin sollte im Weiteren die Diagnose der LRS mit operationalisierten Methoden gestellt werden. Diese basieren auf einer umfangreichen Testdiagnostik mittels Rechtschreib-, Lese- und Intelligenztests.

Im Folgenden werden Methoden vorgestellt, die alle auf dem doppelten Diskrepanzkriterium aufbauen:

1. Diskrepanz: Das Niveau im Lesen und in der Rechtschreibung ist mangelhaft oder ungenügend im Vergleich zu Kindern gleicher Klassenstufe.
2. Diskrepanz: Das Niveau im Lesen und in der Rechtschreibung ist wesentlich niedriger als das gemessene Intelligenzniveau.

Die Richtlinien für ein LRS-Gutachten nach § 35a SGB VIII (Empfehlungen einer Arbeitsgruppe des Bayerischen Sozialministeriums 1996) und den aktuellen Leitlinien der Deutschen Gesellschaft für Kinder und Jugendpsychiatrie und –psychotherapie stellen eine standardisierte und im Praxisalltag etablierte Methode dar.

Eine Diskrepanz zwischen der allgemeinen intellektuellen Begabung und dem Versagen im Lesen und Rechtschreiben ist aufzuzeigen. Eine Teilleistungsstörung ist anzunehmen, wenn das Testergebnis einem T-Wert von  $\leq 37$  entspricht und eine T-Wert Diskrepanz zum Gesamt IQ  $\geq 12$  besteht. Alternativ empfiehlt sich eine Diskrepanz von mindestens 1,5 Standardabweichungen zwischen relativ höherem IQ-Wert und relativ niedrigeren Lese- bzw. Rechtschreibtestwerten. Nach den Kriterien von ICD-10 ist für die Feststellung der Entwicklungsstörung ein Intelligenzquotient  $> 70$  vorauszusetzen.

Im unterdurchschnittlichen und überdurchschnittlichen Intelligenzbereich verliert dieses Modell an Sensitivität, da hier eine höhere, bzw. niedrigere Diskrepanz zum Erfüllen des Kriteriums notwendig ist. Diese Fehlerquelle bei Kindern mit einem hohen oder niedrigen Intelligenzniveau greift das so genannte Regressionsmodell auf. Es sieht Tabellen vor, die die Korrelation von Rechtschreibung bzw. Lesen und Intelligenz berücksichtigen. SCHULTE-KÖRNER beschreibt einen Vorteil gegenüber IQ-Diskrepanzmodellen in der Vermeidung von Verzerrungen in den Extrembereichen der Verteilung der Rechtschreib- und Leseleistung. Als Resultat ergibt sich, dass im Vergleich zum einfachen Diskrepanzmodell für überdurchschnittlich Intelligente eine höhere, und für unterdurchschnittlich Intelligente eine niedrigere

Diskrepanz zum Erfüllen des Kriteriums notwendig ist. Bei durchschnittlicher Intelligenz führen Diskrepanz- und Regressionsmodell zu identischen Ergebnissen (Schulte-Körne, 2001).

### **2.1.2. Probleme bei der operationalisierten Diagnostik einer LRS**

In der aktuellen Legasthenieforschung stellt die Diskussion über Modelle und Falldefinitionen nach wie vor ein zentrales und ungelöstes Problem dar (Deimel 2002, Meyer 2000). Abgesehen von der Forschung ist bei praktischen Fragen des Umgangs mit Lese-Rechtschreibschwierigkeiten wie Benotung, Förderungs- oder Therapiebedarf die Frage der Falldefinition von grundlegender Bedeutung.

Durchzusetzen scheinen sich momentan Modelle mit doppeltem Diskrepanzkriterium. Die Rechtschreibleistung soll von der Altersnorm und dem individuellen Intelligenzniveau abweichen. Unklarheiten herrschen noch über die notwendige Höhe der Abweichung. Die Angaben und Empfehlungen erstrecken sich hier von minimal einer bis maximal zwei Standardabweichungen (Strehlow and Haffner, 2002). Die Einbeziehung von Regressionsmodellen wird meist empfohlen, da sie den Bereich über- und unterdurchschnittlich intelligenter Kinder besser erfassen (s. 2.1.1.), (Schulte-Körne, 2001).

Nach den Leitlinien der Deutschen Gesellschaft für Kinder und Jugendpsychiatrie und –psychotherapie muss bei Verwendung des doppelten Diskrepanzkriteriums das Ergebnis der Rechtschreibleistung  $\leq 37$  T-Wert Punkte betragen und eine T-Wert Diskrepanz zum Gesamt-IQ von  $\geq 12$  T-Wert Punkten bestehen. Dies entspricht inhaltlich in etwa der schulischen Definition mangelhafter Rechtschreibung ( $\geq$  Note 5) bei ansonsten mindestens befriedigenden Schulleistungen ( $\leq$  Note 3).

Da Schulnoten von 5 oder schlechter wesentlich über die weitere Schullaufbahn entscheiden, erscheint dieses Modell inhaltlich bedeutsam und sinnvoll. Auch STREHLOW konstatierte, dass sich Definitionen und Wahl der cut-off Grenzen nicht mit Hilfe statistischer Methodik klären lassen, sondern inhaltlicher Überlegungen und eines Bezugs zu Außenkriterien bedürfen (Strehlow and Haffner, 2002).

Die ICD-Forschungskriterien fordern eine Diskrepanz von 2 Standardabweichungen, was die Fallzahlen auf unter 2 % sinken lässt. Da eine deutliche Unterdiagnostizierung der LRS im Praxisalltag beobachtet wurde, konnte sich dieses Modell nicht durchsetzen. (Strehlow and Haffner, 2002).

Ein wesentliches Problem aller operationalisierten Methoden sind die zeitlich bedingten gegenläufigen Leistungsveränderungen im Rechtschreiben und der Intelligenz. Während die Normen im deutschen Sprachraum für Rechtschreibverfahren sinken, sind die Normen bei Intelligenzverfahren über die Jahrzehnte angestiegen. Beide Effekte wirken additiv im Sinne einer extremen Überschätzung der Legastheniehäufigkeit. Zum Beispiel erreicht man bei Verwendung der sehr strengen ICD-10 Forschungskriterien und veralterten Normen (z. B. CFT20 Normierung 1977, Rechtschreibtest RT Normierung 1968) eine Fallzahl von 25 % im Vergleich zu einer Fallhäufigkeit von 1,7 % bei Verwendung aktuellerer Normen (Stand 1995) (Strehlow and Haffner, 2002).

Zusammenfassend kann man sagen, dass allen Definitionen von Diskrepanzmodellen gleich ist, dass eine gewisse Restunsicherheit in der Diagnosestellung bleibt und deshalb die Lese-Rechtschreibstörung immer auch eine klinische Diagnose bleiben muss. Das Hauptmerkmal sollte auf der inhaltlichen Betrachtung und somit der individuellen Bedeutung und der für den Alltag angezeigten Beeinträchtigung des Betroffenen liegen.

## **2.2. Standardisierte Testverfahren zur Überprüfung von Rechtschreib- und Lesefähigkeit und der Intelligenz**

Die Diagnose selbst und auch individuelle Therapieförderprogramme werden durch die Auswahl der operationalisierten Methode, weitgehend jedoch durch die Qualität der Testverfahren bestimmt.

Es gibt eine Vielzahl von Testverfahren, die zur Diagnostik eingesetzt werden.

Besonderem Interesse galt in dieser Studie, ob häufig in der testpsychologischen Praxis angewandte Verfahren zu denselben Ergebnissen in der Einschätzung der Schulleistungen gelangen und somit zu den richtigen Diagnosen beitragen.

Die Testverfahren unterscheiden sich meist hinsichtlich ihrer Grundkonzeption und vor allem hinsichtlich ihrer Gütekriterien. Während die Validität und Reliabilität bei Lese- und Rechtschreibtests normalerweise im üblichen akzeptablen Bereich liegen, gibt es bei den Normierungen erhebliche Unterschiede.

Bei der Normierung wird zwischen zwei Aspekten unterschieden. Dabei handelt es sich einerseits um das Alter und andererseits um die Repräsentativität der Normen. Letzteres betrifft die Zusammensetzung der Normierungsstichprobe (Stadt vs. Land, Gewichtung der Bundesländer etc.), sowie die Frage, inwieweit es überhaupt, z. B. für das Ende der zweiten Klasse allgemein gültige Rechtschreibnormen in Deutschland geben kann. Eine besondere Schwierigkeit besteht darin, dass es keinen Wortschatz gibt, der bundesweit alters-, klassen- und schultypspezifisch als allgemein verbindlich gilt (Hemminger, 2000). Somit kann man davon ausgehen, dass zwischen den Bundesländern Unterschiede in den Grundwortschatzlisten bestehen. Rechtschreibtests sollten Grundwortschatzlisten prüfen, die in den verschiedenen Bundesländern möglichst übereinstimmen.

Die Basis des Lernens bildet der Grundwortschatz. Bei seiner Erstellung sind mehrere Aspekte zu beachten. Ein Grundwortschatz muss die häufigsten Wörter enthalten und diese Wörter müssen gleichzeitig die wichtigsten Rechtschreibfälle repräsentieren. Zudem ist die Fehlerträchtigkeit einzelner Wörter zu berücksichtigen. Menzel (1985) analysierte insgesamt 20.041 Fehlerwörter aus Schüleraufsätzen. Er stellte fest, dass viele fehlerträchtige Wörter in den Grundwortschatzen nicht enthalten sind.

Und neben Unterschieden zwischen den Bundesländern gibt es natürlich auch Schwankungen zwischen einzelnen Schulen und auch zwischen Klassen innerhalb einer Schule. Es ist nicht klar, worauf diese Unterschiede im Detail zurückzuführen sind, es ist aber evident, dass diese übergeordneten Faktoren relevant für die Interpretation eines individuellen Testergebnisses sind. Ein unterdurchschnittlicher Prozentrang in einem Rechtschreibtest erhält eine andere Bedeutung, wenn man weiß, dass die ganze Schulklasse unterdurchschnittlich ist. Diese Information müsste der Diagnostiker haben, er hat sie aber normalerweise nicht.

Der zweite zentrale Aspekt der Normierung ist das Alter der Normen. Bei einigen Tests liegt die Normierung schon viele Jahre zurück, bei anderen wurde sie nur in Österreich oder der Schweiz durchgeführt.

Es kann also nicht unbedingt davon ausgegangen werden, dass mit unterschiedlichen Testverfahren gleich lautende Ergebnisse zu erhalten sind.

Sucht man in der Literatur nach aktuellen Validierungen von LRS-Tests, findet man meist nur Vergleiche mit Testverfahren oder Schulnoten, die im Rahmen der Testkonzeption durchgeführt wurden und oft schon viele Jahre zurück liegen. Meist wurden die Testvergleiche dann nur mit einem einzelnen Testverfahren und nicht klassenübergreifend durchgeführt.

Eine der aktuellsten Untersuchungen wurde im Jahre 2001 von TACKE durchgeführt. Da sich die Hinweise häuften, dass die Hamburger Schreibprobe häufig zu Fehldiagnosen führt, wurde ein Vergleich der HSP mit dem DRT und der WRT an einer Stichprobe von 248 Drittklässlern durchgeführt. Es konnte gezeigt werden, dass die HSP im Gegensatz zu den beiden anderen Verfahren zu erheblichen Fehldiagnosen führt (Tacke et al., 2001).

Die Verunsicherung im Bereich von Schultests ist sehr groß. Ergebnisse einer Umfrage zur Praxis der psychologischen Testdiagnostik in der deutschen Kinder- und Jugendpsychiatrie zeigen, dass sowohl Aktualisierungen und Neunormierungen gängiger Verfahren als auch adäquate Verfahren im Bereich Lese-, Rechtschreibdiagnostik vermisst werden (Boelte et al., 2000).

Aufgrund dieser allgemein herrschenden Unsicherheit bezüglich der Güte der verschiedenen Rechtschreib- und Lesetests, entstand der Entschluss, im Rahmen einer organisierten Studie einen Vergleich der gängigsten Testverfahren durchzuführen. Es sollte gezeigt werden, inwieweit in der Testpraxis häufig angewandte Testverfahren die Rechtschreib- bzw. Leseleistungen eines Kindes gleichermaßen einschätzen. Auf dieser Grundlage sollte dann eine kritische Betrachtung und Zusammenfassung von Vor- und Nachteilen vollzogen werden. Hierauf könnte dann wiederum eine Überarbeitung, bzw. Aktualisierung einzelner Untersuchungsmethoden aufbauen.

Nachfolgend eine Übersicht der für diese Untersuchung ausgewählten Testverfahren. Aus Gründen der Übersichtlichkeit wird bei der Nennung eines bestimmten Tests nicht jedes Mal die Literaturreferenz genannt. Alle im Text erwähnten Tests sind aber im Literaturverzeichnis aufgeführt.



## Rechtschreibtests

Bei den Rechtschreibtests wählten wir den Diagnostischen Rechtschreibtest für die 1. – 5. Klasse, die Weingartener Grundwortschatz Rechtschreib-Testreihe für die 1. – 3. Klasse, den Westermann Rechtschreibtest für die 4. – 5. Klasse, den Salzburger Lese-Rechtschreibtest für die 1. – 4. Klasse und die Hamburger Schreib-Probe für die 1. - 5. Klasse. Hinsichtlich ihrer Gütekriterien zeigen sie enorme Unterschiede.

Der WRT4/5 und WRT6+ bieten nur sehr alte Eichungen. Die Normstichprobe des SLRT ist ausschließlich aus Kindern zusammengesetzt, die aus den österreichischen Bundesländern Salzburg und Oberösterreich stammten.

Die Hamburger Schreib-Probe ist ein neueres Testverfahren, das die 1. – 9. Klasse abdeckt. Ihr Testkonzept weicht von den anderen Testverfahren ab. Illustrationen veranschaulichen die zu schreibenden Testwörter bzw. –sätze, die vom Lehrer vorgelesen werden. Basis der HSP bildet die Analyse von spezifischen Rechtschreibstrategien (Unterscheidung in orthographische, morphematische und alphabetische Strategie). Eine quantitative Auswertung (Fehleranzahl bzw. Anzahl korrekter Wörter 2. - 9. Klasse, bzw. korrekter Graphemtreffer 1. - 9. Klasse) ist ebenfalls möglich.

Im Gegensatz dazu werden beim WRT (Weingartener und Westermann), und beim DRT Wörter nach Diktat in Lückentexte eingefügt. Neben der quantitativen Auswertung (Fehleranzahl bzw. Anzahl korrekter Wörter) ist beim DRT und bei der Weingartener Rechtschreibtestreihe auch eine qualitative Fehleranalyse möglich, die bestimmte Fehlerschwerpunkte wie z.B. Wahrnehmungs-, Merk- und Regelfehler, identifiziert.

Der SLRT unterscheidet zwischen drei Fehlerarten: Fehler der lauttreuen Schreibung, orthographische Fehler und Verstöße gegen die Groß- und Kleinschreibung. Nur die orthographischen Fehler können in Prozentzahlen angegeben werden. Für die anderen Fehlerarten wird ein kritischer Wert angegeben.

Kinder mit einer Lese- und Rechtschreibstörung weisen im Vergleich zu Kindern ohne Teilleistungsschwäche vermehrt Phonemfehler (Verstöße gegen lautgetreue Schreibweise), Regelfehler (Verstöße gegen regelhafte Abweichungen von der lautgetreuen Schreibweise) und Speicherfehler (Verstöße gegen Abweichungen vom Regelhaften) auf.

Die in dieser Studie verwendeten Testverfahren bieten alle eine quantitative und eine mehr oder minder ausführliche qualitative Auswertung der Rechtschreibleistung an, so dass die für eine Lese- und Rechtschreibschwäche spezifischen Fehlerarten ermittelt werden können.

### Lesetests

Einleitend möchten wir darauf hinweisen, dass natürlich nur die Testverfahren berücksichtigt werden konnten, die zum Zeitpunkt der Studiendurchführung bereits veröffentlicht waren.

Zwei äußerst viel versprechende Lesetestverfahren, von denen eines 2003 und 2005 veröffentlicht wurde und das andere sich noch in Druck befindet, sollen hier zusätzlich kurz Erwähnung finden, auch wenn sie in unseren Untersuchungen leider noch nicht mitaufgenommen werden konnten.

2003 und 2005 erschien das Salzburger Lese-Screening für die Klassenstufen 1 – 4 und 5 – 8. Eine Liste inhaltlich einfacher Sätze, die auf das Wissen der Schülerinnen und Schüler abgestimmt wurde, muss möglichst schnell gelesen und der Wahrheitsgehalt jedes Satzes beurteilt werden. Es dient rein als Screeningverfahren, auf das bei auffälligem Ergebnis ausführlichere Lesetests wie z. B. der SLRT aufbauen sollen.

Ein anderes Verfahren, das sich noch in Druck befindet ist „Ein Leseverständnistest für Elementarschüler (ELFE)“ von SCHNEIDER und LENHARD. ELFE ist ein Leseverständnistest, der in den ersten sechs Schulklassen eingesetzt werden kann. Das Verfahren ist sowohl als Computerprogramm, als auch als Papier- und Bleistiftversion zur Gruppentestung verfügbar. Das Computerprogramm wird vom Kind selbstständig durchgeführt. Im Anschluss an den Test erfolgt eine automatische Auswertung und Diagnosestellung, sodass das Verfahren auch von testdiagnostisch ungeschultem Personal eingesetzt werden kann. Im Vordergrund des Tests steht die Erfassung des Leseverständnisses und nicht des orthographischen Wissens oder der Artikulationsfähigkeiten.

Insgesamt war das Angebot der Lesetests zum Zeitpunkt der Studiendurchführung, vor allem derjenigen, die über den Grundschulbereich hinausgehen, eher spärlich und teilweise von fragwürdiger Güte. Es gibt den Zürcher Lesetest für die 2. - 5. Klasse, den Salzburger Lese- und Rechtschreibtest 1. - 4. Klasse, die Würzburger Leise Leseprobe und Knuspels Leseaufgaben für die 1.- 4. Klasse.

Der Zürcher Lesetest weist für Deutschland keine und für die Schweiz seit 1981 keine Eichstichprobe mehr auf. Es finden sich keine Angaben zur Validität. Da der Zürcher der einzige Lesetest ist, der über den Grundschulbereich hinausgeht und eine breite Anwendung in der testpsychologischen Praxis genießt, wurde er in diese Studie mit aufgenommen.

Der Salzburger Lesetest weist gute Gütekriterien auf. Auch bei diesem Test liegt für Deutschland keine Eichung vor (Eichung 1994 - 1997 an österreichischen Kindern). Da auch dieser Test eine große Anwendung im deutschen Praxisalltag findet, wurde er ebenfalls in die Studie mit einbezogen. Beide Testverfahren messen die Lesegenauigkeit (Anzahl der Fehler) und das Lesetempo (Zeit in Sekunden).

Zwei neuere Testverfahren für den Grundschulbereich sind die Würzburger Leise Leseprobe und Knuspels Leseaufgaben. Die WLLP misst sehr spielerisch und mit einer hohen Motivationskomponente die Lese- (=Dekodier-)geschwindigkeit auf Wortebene. Dieses Verfahren ist in seiner Durchführung sehr ökonomisch und zeigt gute Gütekriterien. Diese Punkte sprachen eindeutig für eine Mitaufnahme in diese Studie.

Knuspels Leseaufgaben ist ein Testverfahren, das die Basisfertigkeiten der Lesefähigkeit und das Leseverständnis mit Hilfe einer sehr spielerischen und motivierenden Aufmachung und dem „Knuspel-Wesen“ als Leitfigur überprüft. Es weist gute und aktuelle Gütekriterien auf. Dies überzeugte, trotz sehr langer Bearbeitungsdauer von durchschnittlich 50 Minuten, dieses Testverfahren ebenfalls mit in die Studie aufzunehmen.

Bei Kindern, die unter einer Lese- und Rechtschreibstörung leiden, dominieren beim Vorlesen eine verzögerte Lesegeschwindigkeit, ein verzögerter Start oder Zeilenverlust beim Vorlesen, falsche Wiedergabe von Worten bzw. Wortteilen (auslassen, ersetzen, verdrehen, hinzufügen), Wort- und Buchstabenvertauschen, sowie die Unfähigkeit, Gelesenes wiederzugeben, aus Gelesenem Schlüsse zu ziehen und/oder Zusammenhänge zu sehen.

Bei Untersuchungen leserechtschreibschwacher Kinder fanden sich breite und ausgeprägte Defizite beim Speichern phonologischer Wortformen (Nachsprechen von Pseudowörtern, Lernen neuer Namen, Wiedererkennen existierender Wörter). Auch zeigten sie ein deutlich verlangsamtes Tempo beim Abruf hoch vertrauter Wörter (Mayringer & Wimmer, 1999).

Beim Zürcher Lesetest, wie auch beim Salzburger Lesetest gehen die Lesezeit und die Lesefehler in die Bewertung ein und bei der WLLP wird die Wortlesegeschwindigkeit überprüft. Knuspels Leseaufgaben prüft grundlegende Lesefertigkeiten wie Rekodieren und Dekodieren, das zum Lesenlernen unverzichtbare Hörverstehen und das Leseverstehen.

Somit sollten die bei einer LRS vorhandenen Lesedefizite durch diese Testverfahren ausreichend erkannt werden.

### Intelligenztests

Zur Ermittlung der Intelligenz wurde der Grundintelligenztest Skala 1 (CFT1) und Skala 2 (CFT 20) gewählt. Sie sind zur Messung der allgemeinen intellektuellen Leistungsfähigkeit sehr reliabel und valide, sowie in der Durchführung sehr ökonomisch (als Gruppentest durchführbar, Testdurchführungsdauer im mittleren Bereich, schnelle Auswertung). Der CFT wird in den Leitlinien der Deutschen Gesellschaft für Kinder- und Jugendpsychiatrie und –psychotherapie der AWMF zur Intelligenzdiagnostik für Lese- und Rechtschreibstörungen empfohlen (Warnke and Hemminger, 2004).

### 3. FRAGESTELLUNG

In dieser Studie sollen folgende Fragestellungen überprüft werden:

- (1) Sind Ergebnisse von LRS-Tests untereinander vergleichbar?
  - a) Wie hoch korrelieren die Ergebnisse miteinander?
  - b) Beurteilen unterschiedliche LRS-Tests die Lese- bzw. Rechtschreibfähigkeit von Kindern gleich streng?
  
- (2) Wie groß ist die Übereinstimmung zwischen Testergebnissen, Lehrer- und Elternurteil?
  
- (3) Wie hoch ist die diagnostische Zuverlässigkeit von LRS-Tests zu bewerten?
  - a) Wie stark hängt die Diagnose „Lese-Rechtschreibstörung, F81“ vom eingesetzten Testverfahren ab?
  - b) Wie hoch ist die Trefferquote einzelner LRS-Tests einzuschätzen?

## **4. METHODE**

### **4.1. Stichprobe**

#### **4.1.1. Rekrutierung**

Die Studienpopulation besteht aus Kindern, die sich für sechs Wochen (29.03 - 09.05.2001) in der Fachklinik für Kinder- und Jugendmedizin in Hochried in Behandlung befanden. Die Klinik Hochried ist eine Kurklinik mit den Schwerpunktindikationen Adipositas (E66), Entwicklungsstörungen mit dem Schwerpunkt Lese-Rechtschreibstörung (F80 - F89), Hyperkinetische Störung (F90), Asthma bronchiale (J45), Atopisches Ekzem (L20) und Skoliose (M41).

Die Kinder und Jugendlichen werden interdisziplinär von den Fachbereichen Medizin, Psychologie und Sozialpädagogik betreut. Eine integrierte Schule bietet die Möglichkeit spezifischer Fördermaßnahmen bei Teilleistungsstörungen und Hyperkinetischem Syndrom.

Vier Wochen vor Beginn der Kurbehandlung wurden alle Eltern mit Kindern zwischen 7 - 12 Jahren, die sich in der Klinik Hochried für den Zeitraum 29.03.-09.05.2001 angemeldet hatten (insgesamt 64), mit einem Elternbrief, einem Anamnesebogen, einem Schreiben an die Lehrer zur Leistungseinschätzung der Kinder und einer Einverständniserklärung (s. Anlage) angeschrieben.

58 Eltern gaben die Einwilligung zur Teilnahme ihres Kindes. Davon mussten 10 Kinder aufgrund nicht passender Klassenstufen, unregelmäßiger Medikamenteneinnahmen und körperlicher Behinderung von der Studie ausgeschlossen werden. 6 Eltern lehnten eine Teilnahme ab.

#### **4.1.2. Einschlusskriterien und Ausschlusskriterien**

Da Testverfahren der 1. - 5. Klasse untersucht werden sollten, war Einschlusskriterium entweder der Besuch einer Grund- oder Sprachheilschule oder einer Schule für Sprachbehinderte in den Klassenstufen 1 - 4, der Besuch einer Förderschule in den Klassenstufen 3 – 6 (Testung mit Tests zwei Klassenstufen unter ihrer eigentlichen Klassenstufe) oder der Besuch einer

Haupt- oder Realschule oder Gymnasium in der 5. Jahrgangsstufe. Die Schulen mussten deutschsprachig sein.

Als Ausschlusskriterien wurden Sinnesbeeinträchtigungen sowie organische und zentralnervöse Erkrankungen angesehen, die Schreiben und/oder Lesen erheblich beeinträchtigen. Beginn oder Änderungen einer medikamentösen Therapie während der Diagnostikphase führten zum Ausschluss aus der Studie.

17 % der initial in die Studie aufgenommenen Kinder, mussten aufgrund nicht passender Klassenstufen, nicht konstanter Medikamenteneinnahmen und/oder körperlicher Behinderung von der Studie ausgeschlossen werden.

#### 4.1.3. Beschreibung der Stichprobe

##### **Alter, Geschlecht, Muttersprache, Klassenstufe, Schultyp:**

In der folgenden Tabelle sind das Alter, die Geschlechterverteilung und die Muttersprache der Studienteilnehmer angegeben. Zusätzlich sind die Größe der verschiedenen Klassenstufen und der Anteil der einzelnen Schularten an der Gesamtstichprobe aufgeführt.

Alter in Jahren	n	Geschlecht	n	Muttersprache	n	Schultyp	n	Schulklasse	n
7 Jahre	7	Männlich	35	Deutsch	44	Grundschule	33	1.Klasse	9
8 Jahre	8	Weiblich	13	Mehrsprachig (Deutsch als Schulsprache)	4	Sprachheilschule	1	2.Klasse	5
9 Jahre	7					Schule für Lernbehinderte	1	3.Klasse	9
10 Jahre	10					Hauptschule	6	4.Klasse	14
11 Jahre	12					Realschule	2	5.Klasse	10
12 Jahre	4					Montessori	1	6.Klasse	1
						Förderschule	4		

Die Geschlechterverteilung ist w : m = 1 : 3. 92 % der Gesamtstichprobe gaben deutsch als alleinige Muttersprache an, 8 % wurden mehrsprachig erzogen. Das jüngste Kind war 7, das älteste 12 Jahre alt. Die Studienteilnehmer besuchten Grundschulen, eine Sprachheilschule, eine Schule für Lernbehinderte, Hauptschulen, Realschulen, eine Montessorischule und Förderschulen.

### **Klinikdiagnosen:**

Die Studienteilnehmer waren aufgrund verschiedener Krankheitsbilder in der Klinik Hochried in Behandlung. Dort erhielten alle Kinder eine Hauptdiagnose (1. Diagnose), sowie in 77 % der Fälle eine Nebendiagnose (2. Diagnose) zugewiesen. Im Folgenden wird die Häufigkeitsverteilung der 1. und 2. Diagnosen der Studienteilnehmer dargestellt.

<b>1.Diagnose</b>	Prozent %	<b>2.Diagnose</b>	Prozent %
Lese-Rechtschreibschwäche	33	Andere nicht- neuropsychiatrische Krankheitsbilder	62
Andere nicht-neuropsychiatrische Krankheitsbilder	31	Lese-Rechtschreibschwäche	14
Hyperkinetisches Syndrom	25	Allgemeine Entwicklungsretardierung	14
Entwicklungsstörung des Sprechens und der Sprache	8	Hyperkinetisches Syndrom	8
Kopfschmerzen	2	Emotionale Störung	3

Bei den ersten Diagnosen ist die Lese-Rechtschreibschwäche mit 33 % führend, gefolgt von nicht- neuropsychiatrischen Krankheitsbildern mit 31 % und dem Hyperkinetischem Syndrom mit 25 %. Bei den nicht-neuropsychiatrischen Krankheitsbildern dominieren Erkrankungen des atopischen Formenkreises und des muskuloskeletalen Systems. Entwicklungsstörungen des Sprechens und der Sprache sind zu 8 % vertreten und ein Kind litt unter Kopfschmerzen.

Die zweiten Diagnosen der Kinder sind zu mehr als der Hälfte (62 %) nicht-neuropsychiatrische Krankheitsbilder. Die anderen 38 % setzten sich aus der Lese-Rechtschreibschwäche, Allgemeinen Entwicklungsretardierungen, dem Hyperkinetischen Syndrom und einer Emotionalen Störung zusammen. Bei 23 % der Kinder wurde keine 2. Diagnose festgestellt.

### **Medikamente:**

Über die Hälfte (56 %) der an der Studie teilnehmenden Kinder erhielt eine Dauermedikation. 21 % aller Teilnehmer erhielten primär ZNS-wirksame Medikamente wie Methylphenidat, Tricyclische Antidepressiva und Nootropika. Unter sonstige nicht primär ZNS wirksame Medikamente fallen inhalativ verabreichte Corticoide und  $\beta_2$ -Sympathomimetika, Cromoglycinsäure, ACC, Oxymetazolin-Nasentropfen, Dimeticon-Tropfen, Dimetinden-Tropfen, Penicillin, Erythromycin, Cotrimoxazol, Nystatincreme, Mesalazin, Aciclovir, Ciclosporin und Azathioprin.



Medikamente	[%]
keine medikamentöse Therapie	44
sonstige nicht primär ZNS –wirksame Medikamente	35
Methylphenidat	13
Methylphenidat + sonstige nicht primär ZNS –wirksame Medikamente	4
Methylphenidat + Nootropikum	2
tricyclische Antidepressiva	2

## 4.2. Fragebögen

### Anamnesebogen für die Eltern (s. Anhang):

Die Eltern wurden in ihrem Anamnesebogen zur Schullaufbahn (Schultyp, Schulwechsel, Klassenwiederholungen, Schulausfall wegen Krankheit) ihres Kindes befragt. Des Weiteren wurden Fragen zu emotionalen oder körperlichen Beschwerden bezüglich des Schulbesuches gestellt.

Die Schulleistungen sollten in einer Tabelle von gut, über durchschnittlich bis schlecht von den Eltern beurteilt werden.

Es wurden Fragen zur Hausaufgabensituation und zu inner- oder außerschulischen Förderungen gestellt. Das Verhalten der Kinder wurde mittels der Connor-Skala erfragt. Bezüglich der Entwicklung des Kindes enthält der Bogen Fragen über den Schwangerschaftsverlauf, die Geburt, den gegenwärtigen Entwicklungsstand und Sprachauffälligkeiten. Anschließend wurden Fragen zu früheren Erkrankungen gestellt und nach eventuellen augen- und ohrenärztliche Untersuchungen und deren Ergebnissen gefragt. Im letzten Teil wurde auf die aktuelle Familiensituation, auf den Beruf der Eltern und auf die zu Hause gesprochene Sprache eingegangen. Die abschließende Frage bezog sich auf eine bekannte Lese-Rechtschreibschwäche innerhalb der Familie.

Der Fragebogen diente in dieser Studie zur Erhebung der persönlichen Daten, zur Überprüfung der Ein- und Ausschlusskriterien und zur Beurteilung der Schulleistungen durch die Eltern.

### **Bogen zur Beurteilung der Schulleistungen durch die Lehrer (s. Anhang):**

Die Eltern erhielten den Bogen für die Lehrer zusammen mit dem primären Anschreiben per Post. Die Lehrer trugen in einer Notenskala von 1 (sehr gut) bis 6 (ungenügend) die aktuellen Schulleistungen der Studienteilnehmer ihrer Klasse ein. 94 % der Lehrerurteile wurden ausgefüllt zurückgesandt, bzw. von den Eltern bei Studienbeginn mitgebracht. Drei Schulen verweigerten die Mitarbeit.

Das Lehrerurteil diente zur Überprüfung der Übereinstimmung zwischen Lehrerurteil, Elternurteil und Testergebnissen.

## **4.3. Testverfahren**

### **4.3.1. Der Diagnostische Rechtschreibetest, DRT 1 bis 5**

Der Diagnostische Rechtschreibetest bietet Rechtschreibtests für die 1.- 5. Klasse an. Alle Tests sind als Lückentexte aufgebaut. Es liegen zwei Parallelformen gleichen Schwierigkeitsgrades vor. Der Test kann als Gruppen-, wie auch als Einzeltest durchgeführt werden. Die Auswertung kann sowohl quantitativ als auch qualitativ erfolgen. Die quantitative Auswertung bietet eine Beurteilung der Rechtschreibleistung, wobei die qualitative Auswertung zusätzlich eine Bestimmung der Fehlerschwerpunkte als Grundlage für darauf aufbauende gezielte Fördermaßnahmen ermöglicht.

Der Anwendungszeitraum erstreckt sich auf das Ende des ersten, Beginn und Ende des zweiten und dritten Schuljahres und Beginn und Mitte des vierten und fünften Schuljahres. Für die Anwendungszeiträume des DRT liegen Normentabellen vor, die die Gesamtleistung in Prozenträngen und beim DRT 4 und 5 zusätzlich in T-Werten darstellen. Die DRT 2, 3 und 4 bieten getrennte Normentabellen für deutsche und ausländische Kinder an. Darüber hinaus sind für den DRT Sonderschulnormen für Lernbehinderte der 3./4., 4./5. und 5./6. Klasse vorhanden.

Der **DRT 1** umfasst 30 Wörter, die in zwei kleine Geschichten eingebaut sind und nach Diktat von den Kindern aufgeschrieben werden. Der Anwendungszeitraum ist der letzte Monat des ersten und die ersten drei Monate des zweiten Grundschuljahres. In Sonderschulen für Lern-

behinderte kann der Test in der 3./4. Klasse angewandt werden. Die Durchführungszeit beträgt 30 - 45 Minuten.

	Objektivität*	Paralleltest-zuverlässigkeit Form A/B*	Halbierungs-Zuverlässigkeit*	Äquivalenz d. Parallelformen (nach Wilks)*	Validität (Korrelation mit dem Lehrerurteil)	Eichung 1989**
n	20 Tests+ 5 Auswerter	267	347(FormA) 330(FormB)	267	54 Klassen	1488
Ergebnis (Korrelation=r)	r=0,99	r=0,89	r=0,94(FormA) r=0,96(FormB)	Nicht signifikant verschieden	r=0,27-0,97(jede Klasse gesondert berechnet)	

\*Die folgenden Angaben beziehen sich nur auf die Auswertung nach Wortfehler.

\*\*Die Stichprobe setzt sich aus 11 Bundesländern zusammen.

Der **DRT 2** enthält 32 Wörter, die nach Diktat von den Kindern in Satzlücken geschrieben werden. Der Anwendungszeitraum umfasst die letzten 2 Monate der 2. Klasse und die ersten 2 Monate der 3. Klasse Grundschule. Bei lernbehinderten Sonderschülern kann der Test in der 4./5. Klasse angewandt werden. Die Durchführungszeit beläuft sich auf 25 - 30 Minuten.

	Paralleltest-zuverlässigkeit Form A/B*	Halbierungs-Zuverlässigkeit*	Äquivalenz d. Parallelformen (nach Wilks)*	Validität (Korrelation mit dem Lehrerurteil)	Eichung 1995***
n	234(FormA) 118(FormB)	2313	236	86 Klassen	2313
Ergebnis (Korrelation=r)	r=0,90(FormA) r=0,91(FormB)	r=0,90	höchst signifikant verschieden**	r=0,20-0,95 (jede Klasse gesondert berechnet)	

\* Die folgenden Angaben beziehen sich nur auf die Auswertung nach Wortfehler.

\*\*danach wurde ein Wort ausgetauscht, das die Ursache für diese Ungleichheit gewesen sein sollte.

\*\*\* Die Stichprobe setzt sich aus 14 Bundesländern zusammen.

Der **DRT 3** umfasst 44 Wörter, die nach Diktat von den Kindern in Satzlücken eingefügt werden. Der Anwendungszeitraum liegt in den letzten 4 Monaten des 3. und in den ersten 3 Monaten des 4. Jahres Grundschule. Sonderschulen für Lernbehinderte können diesen Test für die 5./6. Klasse verwenden. Die Durchführungszeit liegt zwischen 25 - 35 Minuten.

	Paralleltest-zuverlässigkeit Form A/B*	Halbierungs-Zuverlässigkeit*	Äquivalenz d. Parallelformen (nach Wilks)*	Validität (Korrelation mit dem Lehrerurteil)	Eichung**
n	179(FormA) 102(FormB)	160	204	156 Klassen	1. 1982, 2. 1995
Ergebnis (Korrelation=r)	r=0,920(FormA) r=0,923(FormB)	r=0,949	nicht signifikant verschieden	r=0,48-0,97(jede Klasse gesondert berechnet)	1. 3690, 2. 2234

\* Die Angaben beziehen sich nur auf die Auswertung nach Wortfehler.

\*\* Die Stichprobe setzt sich aus 10 Bundesländern zusammen.

Der **DRT 4** umfasst 42 Wörter, die nach Diktat in Satzlücken eingefügt werden. Der Anwendungszeitraum beläuft sich auf die Monate Oktober bis Januar der 4. Klasse Grundschule. Mit

Sonderschülern kann der Test je nach Leistungsfähigkeit von der 5. bis 8. Klasse durchgeführt werden. Die Durchführungszeit beträgt 45 Minuten.

	Halbierungs- zuverlässigkeit	Äquivalenz d. Parallelformen	Validität (Korrelation mit dem Schreiben v. 500, 550, 900 und 1000 Wörtern d. Grundwortsch.)	Eichung 1992*
n	1101(FormA) 1047(FormB)	--	15 und 25 Schülern pro Gruppe	2148
Ergebnis (Korrelation=r)	r=0,92 (FormA+B)	n.s.	r=0,68-0,94(Auswertung pro Gruppe)	

\* Die Stichprobe setzt sich aus 13 Bundesländern zusammen.

Der **DRT 5** beinhaltet einen Lückentext mit 51 Wörtern, die nach Diktat von den Kindern geschrieben werden. Der Anwendungszeitraum sind die Monate Oktober bis Januar der 5. Klasse. Der Test bietet Normentabellen für die Hauptschule, die Realschule, das Gymnasium und zusammengefasst für die Mittel-, Regel- und Sekundarschule an (Individualnormen und Klassennormen). Die Durchführungsdauer beträgt 45 Minuten.

	Halbierungs- zuverlässigkeit	Äquivalenz d. Parallelformen	Validität (Korrelation mit dem Schreiben v. 550, 750, 1000 und 1200 Wörtern d. Grundwortsch.)	Eichung 1993*
n	1503(FormA) 1628(FormB)	--	16 und 29 Schülern pro Gruppe	3131
Ergebnis (Korrelation=r)	r=0,93 (FormA+B)	n.s.	r=0,71-0,95(Auswertung pro Gruppe)	

\*Die Stichprobe setzt sich aus 11 Bundesländern zusammen.

#### 4.3.2. Die Weingartener Grundwortschatz Rechtschreib-Testreihe, WRT1+ bis WRT3+

Die Weingartener Grundwortschatz Rechtschreib-Testreihe beinhaltet Testverfahren zur Messung der Rechtschreibleistung der 1. – 4. Klasse. Alle Tests sind als Lückentexte aufgebaut. Es liegen Parallelformen vor und es können sowohl Gruppen- als auch Einzeltestungen vorgenommen werden. Der Test gibt darüber Auskunft, welche individuellen, gruppen- oder klassenspezifischen Rechtschreibschwierigkeiten vorliegen, welche Lernfortschritte einzelne Schüler verzeichnen und für welche Schüler die Teilnahme an besonderen Fördermaßnahmen in Frage kommt. Die Normen können am Ende des ersten Schuljahres, jeweils zu Beginn, in der Mitte und am Ende des zweiten und dritten Schuljahres sowie zu Beginn des vierten Schuljahres angewandt werden. Für die Anwendungszeiträume des WRT liegen Normentabellen vor, die die Gesamtleistung in Prozenträngen und in T-Werten darstellen. Der WRT bietet dem Untersucher Normentabellen für deutsche und ausländische Kinder an. Die Testgütekriterien wurden nur für den in dieser Studie relevanten Testzeitpunkt angegeben.

Der **WRT 1+** testet die Rechtschreibleistung mit einem, in eine Geschichte eingearbeiteten Lückentext mit 25 fehlenden Wörtern. Anwendungszeitraum ist die 30. - 40. Unterrichtswoche am Ende des 1. Schuljahres und die 2. - 14. und 15. - 27. Unterrichtswoche in der Mitte des 2. Schuljahres. Die Testdurchführung beträgt 45 Minuten.

	Testhalbierungsreliabilität FormA/B	Konsistenzanalyse	Reliabilität/ Parallelformen Ende 1.Kl.	Validität, Ende 1.Kl Lehrerurteil als Außenkriterium FormA/B	Eichung Ende 1.Kl.* 1992-1994
n	500	500	1306	1245(FormA) 884(FormB)	2085(AForm) 1669(BForm)
Ergebnis Korrelation(r)	0,92(FormA) 0,91(FormB)	0,96	0,89	0,64(FormA) 0,63(FormB)	

\* Die Stichprobe setzt sich aus 12 Bundesländern zusammen.

Der **WRT 2+** hat den Lückentext in mehrere Geschichten eingearbeitet. Es werden 43 Wörter diktirt. Anwendungszeitraum ist die 30. - 40. Unterrichtswoche des zweiten Schuljahres und die 2. - 6. und 7. - 20. Unterrichtswoche des dritten Schuljahres. Die Dauer des Testes beläuft sich auf 45 Minuten.

	Testhalbierungsreliabilität FormA/B	Konsistenzanalyse FormA/B	Reliabilität/ Parallelformen Klasse 2	Retest-Reliabilität FormA/B (Intervall 8-9Wochen)	Validität Deutschnote +Diktat als Außenkriterium 2.Kl	Validität DRT2 als Aussenkriterium	Eichung Ende 2.Kl.* 1991/1992
N	450	450	2929	325/319	1587-1851	ca.100	6690
Ergebnis Korrelation(r)	0,94 (FormA) 0,93 (FormB)	0,94 (FormA) 0,94 (FormB)	0,93	0,88 (FormA) 0,88 (FormB)	0,65-0,69	0,84-0,86	

\* Die Stichprobe setzt sich aus 17 Bundesländern zusammen.

Der **WRT 3+** bietet eine Lang (55 Wörter)- und eine Kurzform (16 Wörter) an. Beides sind Lückentexte. Anwendungszeitraum ist die 30. - 40. Unterrichtswoche des dritten Schuljahres und die 2. - 12. Unterrichtswoche des vierten Schuljahres. Die Testdauer liegt bei der Kurzform bei 15 - 20 Minuten und bei der Langform bei 45 Minuten.

	Testhalbierungsreliabilität LangformA/B	Konsistenzanalyse LangformA/B	Reliabilität/ Parallelformen Klasse 3. und 4.	Validität Deutschnote+Diktat als Aussenkriterium LangformA/B	Validität (Testvorformen A/B/C) DRT3 als Aussenkriterium	Eichung* Ende 3.Kl. 1989-1991
n	450	450	2477	1423-2009 (LangformA) 1528-2168 (LangformB)	113	2973
Ergebnisse Korrelation(r)	0,96 (LangformA) 0,93 (LangformB)	0,95 (LangformA) 0,95 (LangformB)	0,91	0,65-0,74 (LangformA) 0,61-0,72 (LangformB)	0,81-0,86	

\*Die Stichprobe setzt sich aus 12 Bundesländern zusammen.

### 4.3.3. Der Westermann Rechtschreibtest, WRT4/5 und WRT 6+

Der **WRT 4/5** kann als Einzel- oder als Gruppentest durchgeführt werden und bietet zwei Parallelformen gleichen Schwierigkeitsgrades. Die quantitative Auswertung spiegelt die Gesamtleistung wieder. Für den Anwendungszeitraum liegen Normentabellen vor, die die Gesamtleistung in Prozenträngen und T-Werten darstellen. Der WRT 4/5 testet die Rechtschreibung mittels 34 Wörter, die nach Diktat in einzelne Sätze eingesetzt werden. Die Testdauer beträgt 30 Minuten. Anwendungszeitraum ist die 16. - 24. und die 32. - 40. Unterrichtswoche in der 4. Klasse und die 2. - 12. Unterrichtswoche in der 5. Klasse. Die Testgütekriterien werden nur für die Mitte der 4. Klasse angegeben.

	Testhalbierungsreliabilität Mitte 4.Klasse FormA	Reliabilität/ Parallelformen Mitte 4.Klasse	Validität Mitte 4.Klasse, DRT3 als Aussenkriterium	Eichung* Mitte 4.Klasse vor 1980 (keine genaue Angabe)
n	1217	315	284	1217
Ergebnisse Korrelation(r)	0,93	0,95	0,95	

\*in 2.Auflage seit 1980 lieferbar. Die Stichprobe setzt sich aus 5 Bundesländern zusammen.

Der **WRT 6+** kann als Gruppen- oder Individualtest durchgeführt werden. Die Gesamtleistung wird durch eine quantitative Auswertung ermittelt. Gesamtnormen und schulformbezogene Normen liegen für das Ende der 5. bis Anfang der 7. Klasse vor. Ende 7. Anfang 8. Klasse bietet der Test Normen für die Hauptschule. Anwendungszeitraum sind jeweils die letzten bzw. die ersten drei Monate eines Schuljahres vom Ende der 5. Klasse bis zum Ende der 7. Klasse. Der WRT 6+ überprüft 40 Wörter, die nach Diktat von den Kindern in einen Lückentext eingetragen werden. Der Test dauert 30 Minuten.

	Reliabilität/Konsistenz-Analyse 6.Klasse	Eichung* 5./6. Klasse vor 1980 (keine genaue Angabe)
N	638	2219
Ergebnisse Korrelation(r)	0,84-0,92	

\*In Anwendung seit 1980. Die Stichprobe setzt sich aus 7 Bundesländern zusammen.

#### 4.3.4. Der Salzburger Lese-Rechtschreibtest, SLRT

Der Lesetest (SLRT-LT) ist ein Einzeltest und dauert klassenabhängig zwischen 10 und 20 Minuten. Grundlage ist die Beherrschung von Buchstaben-Lautbeziehungen. Der Test bietet zwei Parallelversionen vom gleichen Schwierigkeitsgrad an. Während die Kinder die Testwörter aus einer Lesemappe laut vorlesen, kann der Testleiter die Leseleistung auf einem Protokollbogen festhalten. Eine Testmappe gliedert sich in mehrere Subtests, die für die verschiedenen Klassenstufen Testmaterial enthalten. Eine Übersicht gibt die folgende Tabelle:

Übungsblatt Wortlesen	Häufige Wörter	Zusammengesetzte Wörter	Text lang	Text kurz	Übungsblatt Pseudowortlesen	Wortunähnliche Pseudowörter	Wortähnliche Pseudowörter
f. alle Klassen	f. alle Klassen	3.+4. Klasse	3.+4.Klasse	1.+2. Klasse	f. alle Klassen	f. alle Klassen	2.-4.Klasse

Die Subtests erfassen sowohl Defizite des synthetischen, lautierenden Lesens (Lesen von Pseudowörtern), als auch Defizite in der automatischen, direkten Worterkennung (Erfassung von Lesefehlern und Lesezeit beim Wortlesen). Mit der Lesezeit, die für jeden Subtest benötigt wird, kann die Gesamtleistung ermittelt werden, die als entscheidendes Kriterium des Lesetests bewertet wird. Dafür stehen Normentabellen für die 1. Klasse, 2. Klasse 1. + 2. Halbjahr, 3. und 4. Klasse zur Verfügung. Zur Bewertung der Fehleranzahl wird ein kritischer Wert angegeben, der einem Prozentrang von 10 entspricht.

	Reliabilität/Parallel-Testmethode, 2.-4. Klasse	Validität: Deutschnote als Außenkriterium, 4.Klasse	Eichung 1.Auflage 1997
n	39-50	63	1600*
Ergebnisse Korrelation(r)	Fehler:0,14-0,66 Zeit:0,84-0,99	0,37-0,54 (Zeit) 0,46 (Fehler)	

\* Die Stichprobe setzt sich aus Salzburg und Oberösterreich zusammen.

Der Salzburger **Rechtschreibtest (SLRT-RT)** kann als Gruppen- oder als Einzeltest durchgeführt werden. Es liegen Parallelversionen gleichen Schwierigkeitsgrades vor. Der Test hat insgesamt 49 Testwörter, wobei die 1. und 2. Klasse nur die ersten 25 bearbeitet. Die Kinder bekommen einen Bogen mit Lückensätzen, in die sie die Testwörter nach Diktat eintragen. Die Auswertung erfolgt nach 3 Fehlerkategorien: Den nicht lauttreuen Fehlern (N-Typ), den

orthographischen Fehlern (O-Typ) und den Verstößen gegen die Groß- und Kleinschreibung (G-Typ). N- und O-Fehler schließen einander aus. Der G-Fehler ist unabhängig von beiden Fehlerarten. Für die O-Fehler kann in den Normentabellen der entsprechende Prozentrang abgelesen werden. Für N- und G-Fehler wird nur ein kritischer Wert (PR=10) angegeben.

	Reliabilität/Parallel-Testmethode, 2.-4. Klasse O-Fehler,N-Fehler,G-Fehler	Validierung Lehrerurteil als Außenkriterium	Eichung 1.Auflage 1997
n	39-50	1165 2.Klasse 2.Halbjahr	2731*
Ergebnisse Korrelation(r)	O-Fehler:0,74-0,90 N-Fehler:0,19-0,93 G-Fehler:0,59-0,88	204 von den Lehrern als schwach bezeichnete Kinder, schnitten bei den O-Fehlern signifikant schlechter ab als unauffällige Kinder (p<0,001)	

\* Die Stichprobe setzt sich aus Salzburg und Oberösterreich zusammen.

#### 4.3.5. Die Hamburger Schreib-Probe, HSP 1+ bis 5 - 9

Die Hamburger Schreib-Probe bietet Rechtschreibtests ab Mitte der 1. bis zum Ende der 9. Klasse an. Zu Beginn des Testes liest der Testleiter alle Wörter vor, damit sie danach jeder Schüler in seinem eigenen Tempo schreiben kann. Eine Bildvorgabe soll sicherstellen, dass die Bedeutung der zu schreibenden Wörter richtig erfasst wird. Nachfragen ist immer erlaubt. Man kann den Test als Gruppen- oder Einzeltest benützen. Die Bearbeitung des Testheftes beansprucht weniger als eine Unterrichtsstunde.

In den Klassen 2 - 5 kann die Auswertung sowohl nach Wort-, wie auch nach Graphemtreffern erfolgen. In der 1. Klasse ist ausschließlich eine Auswertung nach Graphemtreffern möglich. Die Beherrschung der Rechtschreibstrategien (alphabetische-, orthographische-, morphematische- und wortübergreifende Strategie) kann mit einer differenzierteren Auswertung aufgeschlüsselt werden. Durch eine weit reichende Differenzierung im unteren Leistungsbe- reich ist die HSP auch zur Leistungsüberprüfung und Planung von Fördermaßnahmen im Be- reich der Sonderpädagogik einsetzbar. Für den Sonderschulbereich soll die Version verwen- det werden, die zwei Klassen unter der für die Regelschule vorgesehenen Fassung liegt. Für alle Anwendungsbereiche der HSP liegen Normentabellen vor, die eine Darstellung der Wort- und Graphemtreffer in Prozenträngen und T-Werten ermöglichen.



	<b>HSP1+</b>	<b>HSP2</b>	<b>HSP3</b>	<b>HSP4/5</b>	<b>HSP5-9</b>
	HSP-M1:Jan.-Febr. d. 1.Kl HSP-E1:Mai-Juni d. 1.Kl HSP-M2:Dez.-Jan. d. 2.Kl	Letzten 3.Monate d. 2.Kl	Letzten 3.Monate d. 3.Kl	15.-23.Unterrichtswoche d. 4Kl. 33.-40.Unterrichtswoche d. 4Kl 1.-12. Unterrichtswoche d. 5Kl*	Letzten 3 Mo- nate der Klas- sen 5,6,7,9
Eichung** 1987-1993	n=577 (Mitte 1.Kl)	n=1470	n=1188	n=946 (Mitte 4.Kl)	n=1212 (5.Kl)

\*Es liegen Gesamtnormen und schulformbezogene Normen vor.

\*\*Die Stichproben stammen aus 11 verschiedenen Bundesländern

	Reliabilität / Interne Konsistenz	Stabilität	Validität, Rechtsschreibleistung in Deutschaufsätzen als Außenkriterium
n		400	292, Ende 4.Kl
Ergebnisse Korrelation(r)	Wörter:0,81-0,96 Grapheme:0,92-0,99	0,52-0,93, je nach Abstand d. Erhebungszeitpunkte	0,78-0,81

#### 4.3.6. Der Zürcher Lesetest, Zürcher

Der Zürcher Lesetest kann zur Einstufung der Lesefähigkeit ab Beginn der 2. Klasse bis zur 6. Klasse verwendet werden. Der Test wurde als Einzeltest konzipiert. Die Kinder erhalten Lesekarten mit einzelnen Wörtern und Leseabschnitten. Es gibt drei Wort-Lesekarten und sechs Textlesekarten mit unterschiedlichen Schwierigkeitsgraden für die verschiedenen Jahrgangsstufen. Der Versuchsleiter notiert während des Vorlesens Lesefehler und die Lesezeit auf einem Protokollbogen. Werden noch nicht alle Buchstaben-Lautbeziehungen beherrscht, kann die Buchstabentafel eingesetzt werden. Die Zeit und die Fehler für das Wort- und das Textlesen können mit Normentabellen in Prozentrangstufen (>75 / 51-75 / 26-50 / 16-25 / 11-15 / 6-10 / 1-5 / <1) dargestellt werden. Für die 4. Klasse liegen nur wenig differenzierte Tabellen mit den Prozentrangstufen über 75 / 25-75 / < 25 vor. Für die 5. und 6. Klasse bietet der Test Prozentrangstufen für die Gesamtleistung der Lesegeschwindigkeit und der Lesegenauigkeit (seit der 5. Auflage).

	<b>2.Klasse</b>	<b>3.Klasse</b>	<b>4.Klasse</b>	<b>5.Klasse</b>
	1.Quartal(Ende Juni) 2.Quartal(Ende Sept.) 4.Quartal(Anfang März)	1.Quartal(Ende Juni) 2.Quartal(Ende Sept.) 4.Quartal(Anfang März)	1.Quartal d. 4.Kl	5. und 6.Kl
Eichungsjahr*: Anzahl	1973/1974 n=158-189	1973/1974 n=105-192	1967 n=121	1981 n=129

\* Die Eichung erfolgte ausschließlich an Züricher Schülern.

	Reliabilität/Retestverfahren (Intervall 1 Woche)	Validität
N	29 (2. Klasse) / 26 (3. Klasse)	keine Angabe
Ergebnisse Korrelation(r)	0,77-0,92 (2. Klasse) / 0,57-0,86 (3. Klasse)	

#### 4.3.7. Die Würzburger Leise Leseprobe, WLLP

Die WLLP erfasst die Leseleistung, indem sie die Dekodier(=Lese-)geschwindigkeit ermittelt. Sie kann in den letzten zwei Monaten (Juni, Juli) der Grundschulklassen 1 - 4 verwendet werden. Eine Durchführung als Gruppen- wie auch als Einzeltest ist möglich. Es liegen Pseudo-Parallelförmungen A/B vor. Der Test ist als Multiple-Choice-Test in Speed-Variante aufgebaut. Einem geschriebenen Wort sind jeweils vier Bildalternativen gegenübergestellt und es sollen in fünf Minuten möglichst viele korrespondierende Bilder markiert werden. Insgesamt können 140 Aufgaben gelöst werden. Die gesamte Testdurchführung dauert 15 Minuten. Zur Leistungsbewertung liegen Normentabellen für den Anwendungszeitraum getrennt nach Geschlecht vor. Bei der Eichstichprobe, die 1997 durchgeführt wurde, sind 2.820 Kinder aus sieben deutschen und einem österreichischen Bundesland getestet worden.

	Reliabilität Paralleltest	Reliabilität Retestmethode(Intervall 14 Wochen)	Validität Lehrerurteil als Außenkriterium
n	35-59	109-137	67-74
Ergebnisse Korrelation(r)	<b>1.KI:</b> 0,87 <b>2.KI:</b> 0,92 <b>3.KI:</b> 0,93 <b>4.KI:</b> 0,82	<b>1.KI:</b> 0,75 <b>2.KI:</b> 0,81 <b>3.KI:</b> 0,88 <b>4.KI:</b> 0,78	<b>1.KI:</b> -0,75* <b>2.KI:</b> -0,58* <b>3.KI:</b> -0,39* <b>4.KI:</b> -0,56*

\*Negative Korrelationen, da Lehrerbeurteilung 1(sehr gut) - 6(ungenügend)

#### 4.3.8. Knuspels Leseaufgaben, KNUSPEL-L

Knuspels Leseaufgaben basieren auf einem theoretischen Modell der Leselernentwicklung und gliedern sich in 4 Subtests:

- Ø Subtest 1 / Hörverstehen: Verstehen mündlich gestellter Fragen und Aufforderungen
- Ø Subtest 2 / Rekodieren: Erkennen von lautgleichen Wörtern
- Ø Subtest 3 / Dekodieren: Erkennen von Wortbedeutungen
- Ø Subtest 4 / Leseverstehen: Verstehen schriftlich gestellter Fragen und Aufforderungen

Somit sollen grundlegende Lesefertigkeiten, das Leseverstehen und das zur Beurteilung des Leseverstehens notwendige Hörverstehen gemessen werden. Es kann anhand der Einzelfertigkeiten ein Score „Vorläuferfertigkeiten“ (Subtest 1+2+3) und ein Score „Lesefähigkeit“ (Subtest 1+2+3+4) berechnet werden. Der Test beinhaltet Normen zu allen Einzelfertigkeiten und Scores für den Anwendungszeitraum Ende 1. bis Ende 4. Schuljahr (in Halbjahresschritt-

ten). Man kann auswählen zwischen mono- oder multilingualen Klassenstufennormen und klassenstufenübergreifenden Normen. Die Leseleistung wird in Prozenträngen und T-Werten dargestellt und die Auswertung kann quantitativ, wie auch qualitativ erfolgen. Der Test kann als Gruppen- oder Einzeltest Anwendung finden und enthält zwei Pseudoparalleltestformen (Form A/B). Die Testdauer beträgt für beinahe alle Jahrgangsstufen etwa eine Schulstunde. Nur in der ersten Klasse ist sie etwas länger, kann aber ohne Probleme auf zwei Schulstunden aufgeteilt werden. Mit dem Knuspel-Wesen als Leitfigur bemüht sich der Test um eine sehr kindgerechte Testatmosphäre.

	Retest* 1)Intervall:1-3W 2)Intervall:6Mo 3)Intervall:12Mo	Paralleltest* 1)Intervall:1-3W 2)Intervall:6Mo 3)Intervall:12Mo	Validität* Lehrerurteil als Außenkriterium**	Eichung 1994-1997 in drei verschiedenen Bundesländern
n	1)78-83 2)131 3)59	1)172-223 2)185 3)70	<b>1.KI:</b> 196, 509 <b>2.KI:</b> 244-541 <b>3.KI:</b> 288-677 <b>4.KI:</b> 321-568	4746
Ergebnisse Korrelation(r)	1)0,75-0,83 2)0,81-0,82 3)0,81-0,82	1)0,79-85 2)0,75-80 3)0,84-0,89	<b>1.KI:</b> 0,45-0,47 <b>2.KI:</b> 0,56-0,67 <b>3.KI:</b> 0,60-0,70 <b>4.KI:</b> 0,62-0,70	

\*Die Koeffizienten der Knuspel-Scores I+II

\*\*Die Korrelation mit anderen ähnlichen Testverfahren (ASVT, AST2+3, WRT2+/3+, PSB, KT 3-4 u.a.) wurde zusätzlich überprüft.

#### 4.3.9. Grundintelligenztest Skala 1, CFT1 und Skala 2, CFT20

Der **CFT 1** stellt eine partielle Adaption des „Culture Fair Intelligence Tests-Scale 1“ von CATTELL dar. Der CFT 1 bestimmt die Grundintelligenz, d. h. die Fähigkeit des Kindes, in neuartigen Situationen und anhand von sprachfreiem, figuralem Material, Denkprobleme zu erfassen, Beziehungen herzustellen, Regeln zu erkennen, Merkmale zu identifizieren und rasch wahrzunehmen. Die Erfassungen und Lösungen von nonverbalen Problemstellungen werden dabei ermittelt. Der CFT 1 kann als Gruppen- wie auch als Einzeltest durchgeführt werden. Es liegen zwei Pseudoparallelformen A/B vor. Der Test gliedert sich in 5 Untertests.

1. Subtest: Substitution
2. Subtest: Labyrinth
3. Subtest: Klassifikationen
4. Subtest: Ähnlichkeiten
5. Subtest: Matrizen

Die durchschnittliche Testdauer beträgt zwischen 35 - 50 Minuten. Normen in Form von Prozenträngen, T-Werten und Intelligenzquotienten liegen für die Altersstufen 5;6 - 9;5 vor.

	Reliabilität: a) Aller Untertests b) Summe der 3 letzten Untertests:	Validität HAWIK als Außenkriterium	Eichung 1995, in zwei Bundesländern
Ergebnisse Korrelation(r)	a) 0,65-0,86 b) 0,90-0,96	Mit Handlungsteil: 0,66 Mit Verbalteil: 0,48	n=1500

Der **CFT20** stellt die überarbeitete deutsche Adaption des „Culture Fair Intelligence Test – Scale 2“ von CATTELL dar. Ziel dieses Testes ist die Erfassung der allgemeinen intellektuellen Leistungsfähigkeit im Sinne der Catell’schen „General Fluid Ability“, die als Fähigkeit, komplexe Beziehungen in neuartigen Situationen wahrnehmen und innerhalb einer bestimmten Zeit erfassen zu können, interpretiert werden kann. Der CFT 20 soll die Entfaltungsmöglichkeiten eines Individuums möglichst milieunabhängig erfassen und ist aus diesem Grunde ein sprachunabhängiger Test mit einer Auswahl anschaulich-figuralen Formen. Der Test ist in zwei Testteile mit je 4 Untertests aufgebaut („Reihenfortsetzen“, „Klassifikationen“, „Matriizen“ und „topologische Schlussfolgerungen“). Die Testdauer beträgt etwa eine Stunde. Es bestehen zwei Pseudoparallelförmigen A/B. Gruppen-, wie auch Einzeltestungen sind möglich. Die Antworten werden auf einem gesonderten Antwortbogen in Multiple-Choice-Form notiert. Altersnormen stehen von 8;7 bis 18 Jahren, sowie für Erwachsene von 18 bis 70 Jahren (Teil 1) zur Verfügung. Außerdem beinhaltet der Test Klassenstandardwerte für die 5. - 10. Klasse und Schulstandardwerte für die Grundschule und Berufsschulen. Die Rohwerte können in IQ-, T-Werte und Prozentrangnormen umgewandelt werden.

	Testwiederholung Intervall: 2W	Splithalf- Methode	Äquivalenzkoeffizient Zwischen 1.+2. Teil	Validität	Eichung 1977
Ergebnisse Korrelation (r)	0,77	0,95	0,78	Landungsanteil des Subtestes Klassifikationen mit dem „General Fluid Ability Faktor“ r=0,58-0,78 Korr. mit and. Intelligenztests (PSB,HAWIK,CPM,WIP u.a.) r=0,64	n=4400*

\*Die Stabilität der CFT 20 Normen wurde anhand großer Stichproben in den Jahren 1978 (n=12818), 1981 (n=10604), 1988/89 (n=2300), 1995 (n=2300) 1996 (n=12330) in den 4., 5., 6. Klassen überprüft. Es gab zu keinen Messzeitpunkten Anzeichen für eine bedeutsame Veränderung.

## 4.4. Studienablauf

### 4.4.1. Anwendung der Testverfahren

Jedes Kind der 1. - 4. Grundschulklasse oder der 3. - 6. Förderklasse nahm an 4 Rechtschreibtests (DRT, SLRT u. HSP 1. - 4. Klasse, WRT1+ - 3+, WRT4/5), an 3 - 4 Lesetests (SLRT, WLLP u. Knuspel 1. - 4. Klasse, Zürcher 2. - 4. Klasse,) und an einem altersentsprechenden

nonverbalen Intelligenztest (CFT1 oder CFT20) teil. Die Kinder der 5. Klasse nahmen an 3 Rechtschreib- (DRT 5, WRT 6+, HSP 5), einem Lese- (Züricher 5. Klasse) und einem nonverbalen Intelligenztest (CFT20) teil.

Beim WRT 3+ kann man zwischen einer Lang- und einer Kurzform wählen. Für diese Studie wurde die Langform mit 55 Testwörtern verwendet. Beim Salzburger Rechtschreibtest wurde nur die Auswertung für orthographische Fehler in den Berechnungen erfasst, da nur hierfür Prozenträge vorliegen. Bei der HSP wurde bis auf die 1. Klasse die Auswertung nach Worttreffer verwendet. Für die 1. Klasse bietet dieses Testverfahren ausschließlich eine Auswertung nach Graphemtreffer an. Knuspels Leseaufgaben bietet bei seiner Auswertung zwei Lesescores an, die eine Zusammenfassung der Subtests darstellen. Bei unserem Testvergleich wurden nur die Lesescores berücksichtigt.

Welche Testverfahren in welcher Klasse verwendet wurden, zeigt die folgende Tabelle. Wurde ein Testverfahren während eines Schuljahres zu mehreren Zeitpunkten normiert, so wurde in der Auswertung die Normierung verwendet, deren Erhebungszeitpunkt am nächsten zu dem Untersuchungszeitpunkt lag.

Bei einigen Testverfahren kann man unter verschiedenen Normentabellen (deutsch, ausländisch, schulspezifisch, u. a.) wählen. Wir verwendeten einheitlich für alle Schultests die klassenstufenspezifischen Individualnormen für deutsche Kinder. Bei den Intelligenztests wurden die Altersnormen verwendet. Die Kinder von Förderschulen wurden mit Tests, die zwei Klassenstufen unter ihrer eigentlichen Klassenstufe liegen, getestet. Die Auswertung erfolgte, soweit wählbar quantitativ.

## Zeitpunkt der Studiendurchführung: 29.03.2001 – 09.05.2001 (26. - 30. Unterrichtsw.)

	<u>1. Klasse</u>	<u>Verwendete Normierung</u>	<u>2. Klasse</u>	<u>Verwendete Normierung</u>	<u>3. Klasse</u>	<u>Verwendete Normierung</u>
<b>DRT</b>	DRT 1	Letzter Monat d. 1.Schuljahres(36.-39.*)	DRT 2	Letzten 2 Mo. d. 2.Schulj., (34.-39.*)	DRT 3	Letzten 4 Mo. d. 3.Schulj., (27.-39.*)
<b>WRT</b>	WRT 1+	30.-40.Unterrichtsw. d. 1. Schuljahres	WRT 2+	30.-40.Unterrichtsw. d. 2. Schuljahres	WRT 3+	30-40.Unterrichtsw. d. 3. Schuljahres
<b>SLRT-RT</b>	SLRT 1&2	Ende d. 1. Klasse (genaue Angabe fehlt*)	SLRT 1&2	2.Halbjahr d. 2. Kl., (22.-39.*)	SLRT 3&4	3.Klasse
<b>HSP</b>	HSP1+	M1:Januar-Februar d. 1.Kl, (15.-21.*)	HSP2	Letzten 3.Monate d. 2.Kl, (29.-39.*)	HSP3	Letzten 3.Monate d. 3.Kl, (29.-39.*)
<b>SLRT-LT</b>	SLRT 1. Klasse	Ende d. 1. Klasse (genaue Angabe fehlt*)	SLRT 2.Klasse	2.Halbjahr d. 2. Kl., (22.-39.*)	SLRT 3.Klasse	3.Klasse
<b>Zürcher</b>			Zürcher 2.Klasse	Ende 4.Quartal d. 2.Klasse, (29.-39.*)	Zürcher 3.Klasse	Ende 4.Quartal d. 3.Klasse, (29.-39.*)
<b>WLLP</b>	WLLP	Letzten 2 Monate der 1.Klasse, (34.-39.*)	WLLP	Letzten 2 Monate der 2.Klasse, (34.-39.*)	WLLP	Letzten 2 Monate d. 3.Klasse, (34.-39.*)
<b>Knuspel</b>	Knuspel Ende1.Kl.	Ende 1.Klasse, (22.-39.*)	Knuspel Ende 2.Kl.	Ende 2.Klasse, (22.-39.*)	Knuspel Ende3.Kl.	Ende 3.Klasse, (22.-39.*)

	<u>4. Klasse</u>	<u>Verwendete Normierung</u>	<u>5. Klasse</u>	<u>Verwendete Normierung</u>
<b>DRT</b>	DRT4	Oktober – Januar, (4.-18.*)	DRT5	Oktober - Januar, (4.-18.*)
<b>WRT</b>	WRT4/5	16.-24.Unterrichtswoche	WRT6+	Letzten 3 Monate der 5.Klasse, (29.-39.*)
<b>SLRT-RT</b>	SLRT 3&4	4.Klasse	--	--
<b>HSP</b>	HSP4/5	15.-23.Unterrichtswoche d. 4Kl.	HSP5-9	Letzten 3 Monate d. 5.Kl, (29.-39.*)
<b>SLRT-LT</b>	SLRT 4.Klasse	4.Klasse	--	--
<b>Zürcher</b>	Zürcher 4.Klasse	1.Quartal d. 4.Kl, (1.-11.*)	Zürcher 5/6.Kl.	5./6. Klasse
<b>WLLP</b>	WLLP	Letzten 2 Monate der 4.Klasse, (34.-39.*)	<b>Tabelle:</b> Darstellung der verwendeten Tests u. Normierungszeitpunkte in den einzelnen Jahrgangsstufen; * Unterrichtswoche.	
<b>Knuspel</b>	Knuspel, Ende4.Kl.	Ende 4.Klasse, (22.-39.*)		

### 4.4.2. Testungsabfolge und Parallelformen

Es fand eine Randomisierung der Testungsreihenfolge statt, um einen zeitlich bedingten Folgefehler auszuschließen.

Stellt ein Testverfahren zwei Gruppen A/B für eine Jahrgangsstufe zur Verfügung, so wurden beide Gruppen gleichermaßen randomisiert.

### 4.4.3. Durchführung der Testverfahren

Die Daten wurden im Zeitraum 29.03. - 09.05.2001 erhoben. Die Testverfahren wurden unter der Betreuung von Prof. Dr. W. v. Suchodoletz, von einer Psychologie- und einer Sozialpädagogikpraktikantin und der Autorin dieser Arbeit durchgeführt. Der Testzeitpunkt der Recht-

schreib- und Lesetests lag am Vormittag zwischen 8.45 und 12.00 Uhr. Die Intelligenztests wurden sowohl am Vormittag zwischen 8.45 und 12.00 Uhr als auch am Nachmittag zwischen 13.00 und 14.00 Uhr durchgeführt. Zur Durchführung der Rechtschreib- und Lesetests bei einem Studienteilnehmer wurden im Schnitt 12 Tage (incl. Wochenende) benötigt. Als Testräume dienten ein kleiner und ein mittelgroßer Raum im Kommunikationszentrum und zwei Klassenzimmer in der Schule der Klinik Hochried. Die Teilnehmerstärke eines Tests variierte zwischen einem und sieben Kindern.

#### **4.5. Statistik**

Die Gesamtleistungen der Testverfahren wurden in T-Werten aufgeführt. Bietet ein Testverfahren zur Darstellung der Gesamtleistung nur Prozenträge an, so wurden die Prozenträge mittels einer Transformationstabelle in T-Werte umgerechnet (Schmid, 2004). Bei den Lesetests (SLRT/Zürcher) wurde der Mittelwert der einzelnen Untertests zur Darstellung der Gesamtleistung als Punktwert gebildet.

Bei einer Gruppenstärke  $n \geq 30$  wurde die zu untersuchende Variable mittels des Kolmogorov-Smirnow-Test auf Normalverteilung überprüft. Bei einer Gruppenstärke  $n < 30$  wurde davon ausgegangen, dass die zu untersuchende Variable keine Normalverteilung aufweist.

Als deskriptive Statistik wurde bei einer Stichprobengröße  $\geq 30$  der Mittelwert, der Standardfehler des Mittelwertes, die Standardabweichung, das 95 % Konfidenzintervall und die Größe der Stichprobe angegeben. Bei einer Stichprobengröße  $< 30$  sind der Median mit 1. und 3. Quartile, das Minimum, das Maximum und die Größe der Stichprobe dargestellt.

Zur Berechnung der Korrelation wurde bei Intervallskalierung und Normalverteilung der Testdaten die Produkt-Moment-Korrelation nach Pearson verwendet. War eine der beiden Variablen ordinalskaliert oder nicht normalverteilt, so wurde die Rangkorrelation nach Spearman berechnet. Die Höhe einer Korrelation richtet sich nach dem Korrelationskoeffizienten  $r$ . Eine sehr geringe Korrelation liegt bei  $r \leq 0,2$ , eine geringe Korrelation bei  $r > 0,2 - 0,5$ , eine mittlere Korrelation bei  $r > 0,5 - 0,7$ , eine hohe Korrelation bei  $r > 0,7 - 0,9$  und eine sehr hohe Korrelation bei  $r > 0,9 - 1$  vor.

Der Vergleich von zwei Testergebnissen hinsichtlich ihrer Mittelwerte, erfolgte bei normalverteilten Variablen mittels des T-Tests bei gepaarten Stichproben. War die Stichprobengröße  $n < 30$  und konnte somit die Annahme der Normalverteilung nicht aufrechterhalten werden, so wurde der Wilcoxon-Test zum nicht-parametrischen Vergleich zweier abhängiger Stichproben verwendet. Wurde ein Vergleich von mehr als zwei unabhängigen Variablen bei nicht normalverteilten Daten notwendig, wurde die Rangvarianzanalyse nach KRUSKAL und WALLIS verwendet.

Zur Beurteilung der Güte eines diagnostischen Verfahrens wurden die Sensitivität, die Spezifität, die Gesamttrefferquote und der RATZ-Index (**R**elativer **A**nstieg der **T**refferquote gegenüber der **Z**ufallstrefferquote) berechnet. Die Sensitivität eines diagnostischen Verfahrens für eine bestimmte Krankheit ist der Anteil der zu Recht erkannten Kranken. Die Spezifität eines diagnostischen Verfahrens für eine bestimmte Krankheit ist der Anteil der zu Recht als gesund bezeichneten Gesunden (Gaus et al., 1999). Die Gesamttrefferquote ist ebenfalls ein Maß für die Qualität der Vorhersage. Sie berechnet sich als Quotient aus der Anzahl der valid klassifizierten Kinder (= Summe der valid positiven und valid negativen Kinder) geteilt durch die Anzahl aller klassifizierten Kinder. Anschließend wird der Quotient in Prozentwerte umgerechnet und gibt an, bei wie viel Prozent der Kinder die Vorhersage eingetroffen ist. Eine Gesamttrefferquote sollte immer mit der Zufallstrefferquote verglichen werden, um die Signifikanz der Vorhersage zu prüfen. Der RATZ-Index (**R**elativer **A**nstieg der **T**refferquote gegenüber der **Z**ufallstrefferquote) nach Marx bietet sich hierfür an. Dieser erlaubt, den „Gewinn“ zu bestimmen, den die Trefferquote mit Hilfe der Prädiktorvariablen gegenüber einer Zufallsauswahl hat, und erlaubt so, verschiedene Vorhersagegüten miteinander zu vergleichen. „Der RATZ-Index berechnet sich aus der Gesamttrefferquote GQ minus der Zufallstrefferquote ZQ dividiert durch die Maximaltrefferquote MQ minus der Zufallstrefferquote ZQ. Mit 100 multipliziert ergibt sich der Prozentwert  $RATZ=100 \times (GQ-ZQ)/(MQ-ZQ)$ “ (Marx, 2000).

Vorraussetzung zur Berechnung der Güteparameter ist ein Goldstandard, den in dieser Studie das Lehrerurteil darstellt. Die Lehrer erhielten im Rahmen der Studie einen Beurteilungsbogen mit einer Notenskala von 1 (sehr gut) bis 6 (ungenügend), auf dem sie die aktuellen Schulleistungen ihrer Schüler und Studienteilnehmer eintrugen.

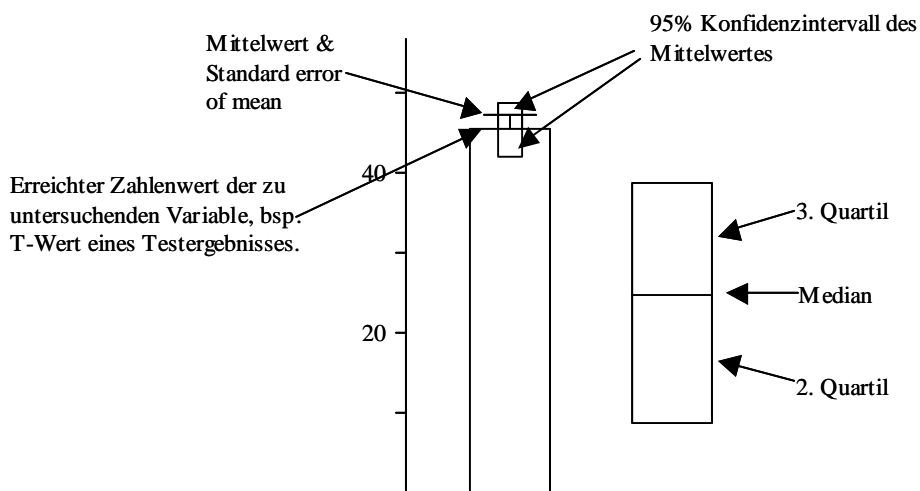


Definition des Goldstandards:

⇒ Mangelhafte Schulleistung („Einschätzung“ der Lehrer  $\geq 4,5$  und  $\geq 5$  )

⇒ mangelhaftes Testergebnis (T-Wert  $\leq 37$ )

Die untenstehende Abbildung erläutert die Systematik der graphischen Darstellung der Ergebnisse: Links im Beispieldiagramm sind die Elemente der Darstellung bei Verwendung eines Mittelwertes erläutert, rechts wird die sog. *box-plot*-Darstellung definiert.



## 5. ERGEBNISSE

### 5.1. Erhält man mit den Rechtschreibtests gleiche Ergebnisse?

#### 5.1.1. Unterscheiden sich bei der HSP die Auswertung nach Worttreffer und die Auswertung nach Graphemtreffer in ihren Ergebnissen?

Die Hamburger Schreibprobe bietet, bis auf die 1. Klasse, sowohl eine Auswertung nach Worttreffer, wie auch nach Graphemtreffer an. In der 1. Klasse kann ausschließlich nach Graphemtreffer ausgewertet werden. Wir bevorzugten die Auswertung nach Worttreffer, da alle anderen Rechtschreibtests ebenso nach Worttreffer ausgewertet werden. Voraussetzung hierfür sollte sein, dass sich die beiden Auswertungsformen nicht in ihren Ergebnissen unterscheiden. Folglich sollte nun untersucht werden, ob sich diese beiden Auswertungsmöglichkeiten in ihren Ergebnissen unterscheiden. Da wie bereits erwähnt nur für die 2. – 5. Klasse beide Auswertungsformen angeboten werden, beziehen sich unsere Berechnungen auch nur auf diese Klassenstufen.

Variable	Korrelationskoeffizient	Signifikanz (2-seitig)
HSP wo. / HSP gr.	0,96	<0.001

**Tabelle 1: Höhe der Korrelation zwischen HSP Graphem- u. Worttreffer d. 2.-5. Klasse:** n=39, (Korrelationskoeffizient nach Pearson).

		HSP wo.	HSP gr.
N		39	39
Mittelwert + SEM		42,1±1,5	42,1±1,6
Standardabweichung		9,2	10,1
95 % Konfidenzintervall des Mittelwerts	Untergrenze	39,1	38,8
	Obergrenze	45,1	45,3

**Tabelle 2: Deskriptive Statistik:** HSP Graphemtreffer u. HSP Worttreffer 2.-5. Klasse in T-Werten.

Variablen	Signifikanz (2-seitig)
HSP gr. / HSP wo.	0,96

**Tabelle 3: Vergleich der Auswertungen nach Graphem- und Worttreffer d. 2.-5. Klasse:** n=39,n.s.; (T-Test bei gepaarten Stichproben).

Zwischen der Auswertung nach Graphemtreffer und der Auswertung nach Worttreffer der Hamburger Schreibprobe besteht ein höchst signifikanter, sehr hoher positiver Zusammenhang von 0,96 (s. Tab. 1). Beim Vergleich der beiden Auswertungsformen zeigt sich kein signifikanter Unterschied (s. Tab. 2 u. 3). Somit wurde im Folgenden für die Klassen 2 - 4 die

Auswertung nach Worttreffer und nur für die 1. Klasse die Auswertung nach Graphemtreffer verwendet.

### 5.1.2. Wie unterscheiden sich die Rechtschreibtests verschiedener Klassenstufen?

#### Klassen 1 – 4 zusammen

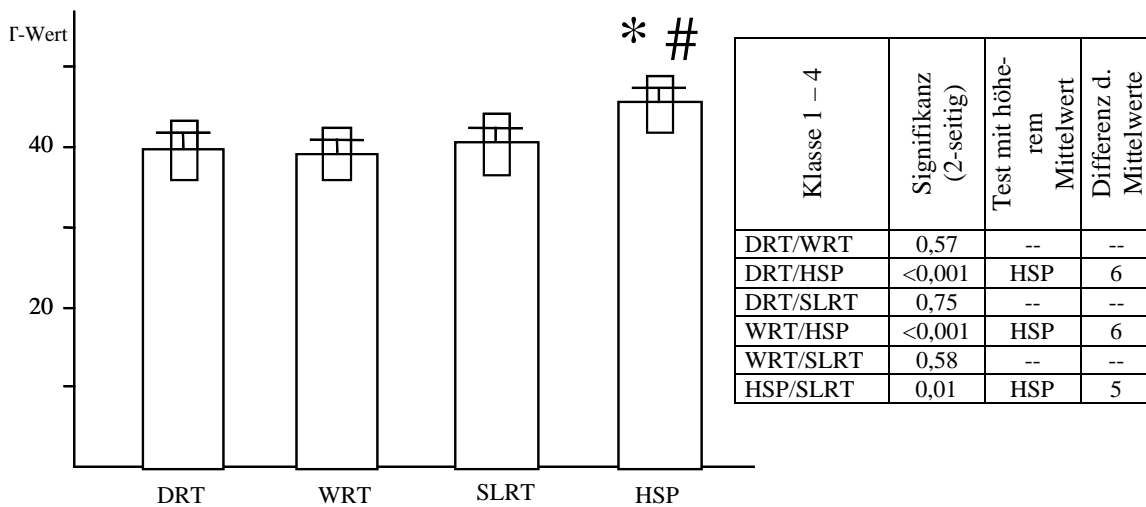
Um einen Überblick über die Tendenzen der einzelnen Testreihen während der Grundschuljahre zu erhalten, wurden zu Beginn die Testergebnisse der 1. – 4. Klasse zu Mittelwerten zusammengefasst. Die Weingartener Grundwortschatz Rechtschreib-Testreihe deckt nur die ersten 3 Jahrgangsstufen ab. Für die 4. Klasse verwendeten wir den Westermann Rechtschreibtest. Durch die Zusammenfassung aller vier Jahrgangsstufen, wurden auch diese beiden Testverfahren zusammengefasst. Die Testverfahren sind von zwei verschiedenen Autoren verfasst. In ihrem Aufbau sind sie sich jedoch sehr ähnlich. Alle anderen Testreihen bieten Verfahren für alle vier Jahrgangsstufen an. Es sollte untersucht werden, ob ein Zusammenhang zwischen den Ergebnissen der Rechtschreibtestreihen besteht und wie streng die verschiedenen Testverfahren die Rechtsschreibleistungen bewerten.

Klasse 1 – 4	Korrelationskoeffizient	Signifikanz (2-seitig)
DRT/WRT	0,86	<0,001
DRT/HSP	0,70	<0,001
DRT/SLRT	0,50	0,001
WRT/HSP	0,66	<0,001
WRT/SLRT	0,36	<0,05
HSP/SLRT	0,40	0,01

**Tabelle 4:**  
Höhe der Korrelation zwischen DRT, WRT(1+-3+ u. 4/5), SLRT, HSP der 1.-4. Klasse: n=39, (Korrelationskoeffizient nach Pearson).

Klasse 1 – 4		DRT	WRT	SLRT	HSP
N		39	39	39	39
Mittelwert+SEM		39,9±1,7	39,4±1,6	40,5±1,8	45,6±1,7
Standardabweichung		10,8	9,8	10,9	10,6
95 % Konfidenzintervall des Mittelwerts	Untergrenze	36,4	36,2	36,9	42,1
	Obergrenze	43,4	42,6	44	49

**Tabelle 5: Deskriptive Statistik:** DRT, WRT(1+-3+ u. 4/5), SLRT, HSP der 1.-4. Klasse in T-Werten.



**Abbildung 1: Paarvergleich von DRT, WRT(1+3+ u. 4/5), SLRT, HSP der 1.-4. Klasse:**  
 n=39, MW ± SEM, 95 % CI,  
 \* p<0,001: vs. WRT1+-3+ u. 4/5, DRT 1-4,  
 # p=0,01: vs. SLRT1-4,  
 (T-Test für gepaarte Stichproben).

Der DRT korreliert mittel bis hoch (0,5 - 0,9) mit den anderen Rechtschreibtests. Vergleichbar hiermit ist der WRT (WRT 1+ - 3+, WRT 4/5), der eine Korrelation zwischen 0,4 - 0,9 mit den anderen Test aufweist. Die Hamburger Schreibprobe zeigt eine geringe bis mittlere Übereinstimmung (0,4 - 0,7) und am schlechtesten schneidet der Salzburger Rechtschreibtest mit einer Korrelation zwischen 0,4 - 0,5 ab (s. Tab. 4).

Im Vergleich zeigen sich zwischen der HSP und dem DRT, WRT (WRT 1+ - 3+, WRT 4/5), sowie dem SLRT sehr bis höchst signifikante Unterschiede. Die Kinder der 1. - 4. Klasse erreichen in der Hamburger Schreibprobe deutlich höhere T-Werte als in den Rechtschreibtest WRT (WRT 1+ - 3+, WRT 4/5), DRT und SLRT (s. Tab. 5, Abb. 1).

### **Klassen 1 – 3 zusammen**

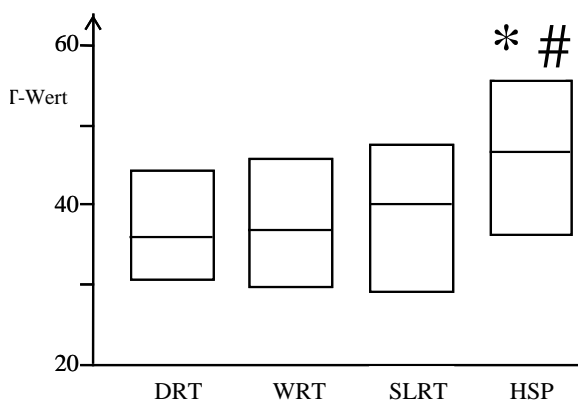
Da der WRT 1+ - 3+ und der WRT 4/5 verschiedenen Testreihen angehören, die zwar nur geringe Unterschiede in ihrem formellen Aufbau aufweisen, jedoch nicht zu vernachlässigende Unterschiede hinsichtlich ihrer Gütekriterien, wurden zusätzlich die Klassen 1 – 3 untersucht. Auch hier sollte gezeigt werden, ob ein Zusammenhang zwischen den Ergebnissen der Rechtschreibtests besteht und ob die Rechtschreibleistungen durch die Testverfahren gleich streng bewertet werden.

Klasse 1 – 3	Korrelationskoeffizient	Signifikanz (2-seitig)
DRT/WRT	0,85	<0,001
DRT/HSP	0,75	<0,001
DRT/SLRT	0,34	0,09
WRT/HSP	0,70	<0,001
WRT/SLRT	0,21	0,30
HSP/SLRT	0,36	0,08

**Tabelle 6:**  
**Höhe der Korrelation zwischen DRT, WRT, SLRT und HSP der 1. – 3. Klasse:** n=26, (Rangkorrelation nach Spearman).

		DRT	WRT	SLRT	HSP
N		26	26	26	26
Median		36	37	40	46,5
Perzentile	25	31	30	29,5	36,5
	75	44,3	45,8	47,5	55,5
Minimum		24	24	24	23
Maximum		70	65	59	68

**Tabelle 7:**  
**Deskriptive Statistik:** DRT, WRT, SLRT, HSP der 1.-3. Klasse in T-Werten.



Rechtschreibtests	Signifikanz (2-seitig)	Test mit höherem Mittelwert	Differenz d. Mittelwerte
DRT/WRT	0,46	--	--
DRT/HSP	<0,01	HSP	11
DRT/SLRT	0,65	--	--
WRT/HSP	<0,01	HSP	10
WRT/SLRT	0,98	--	--
HSP/SLRT	<0,05	HSP	7

**Abbildung 2: Paarvergleich von DRT, WRT, SLRT, HSP der 1.-3. Klasse:**

n=26, Median, 1. u. 3. Quartil,  
\*p<0,01: vs. DRT, WRT,  
# p<0,05: vs. SLRT  
(Wilcoxon -Test).

Die Tendenz, die bei Zusammenfassung der 1. - 4. Klasse gezeigt werden konnte, verstärkt sich nun bei alleiniger Betrachtung der ersten drei Grundschuljahre (s. Tab. 6). DRT, WRT und HSP zeigen untereinander signifikant hohe Zusammenhänge (0,7 - 0,9). Lediglich der Salzburger Rechtschreibtest korreliert sehr gering bis gering (0,2 - 0,3) mit den anderen Testverfahren.

Beim Vergleich finden sich erneut signifikante bis sehr signifikante Unterschiede zwischen der HSP und dem DRT, WRT, sowie SLRT. Die Kinder der 1. - 3. Klasse erreichen in der Hamburger Schreibprobe deutlich höhere T-Werte als in den Rechtschreibtest WRT, DRT und SLRT (s. Tab. 7, Abb. 2).

### Klassen 4 – 5 zusammen

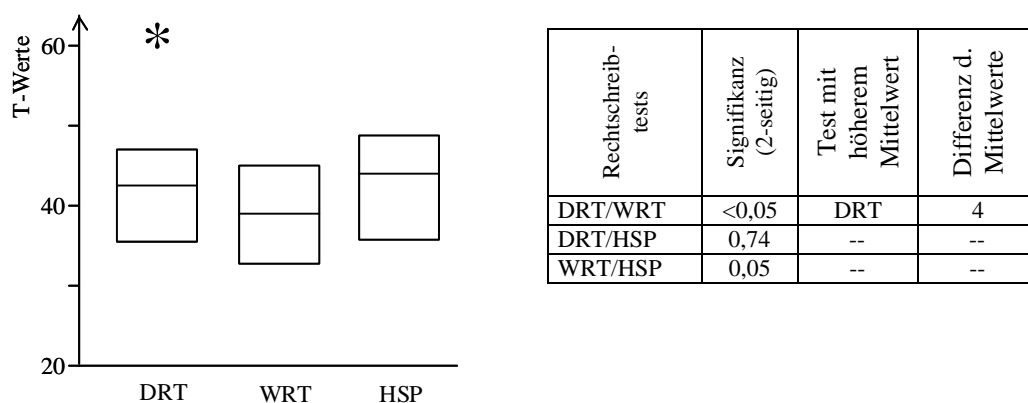
Da der WRT 1+ - 3+ und der WRT 4/5 verschiedenen Testreihen zugehörig sind, die zwar nur geringe Unterschiede in ihrem formellen Aufbau aufweisen, jedoch nicht zu vernachlässigende Unterschiede hinsichtlich ihrer Gütekriterien, sollten nun nach Untersuchung der Klassen 1 – 3 zusätzlich die Klassen 4 – 5 untersucht werden. Ziel war die Klärung einerseits des Zusammenhangs zwischen den Ergebnissen der Rechtschreibtests und andererseits der Strenge der Beurteilung der Testverfahren.

Klasse 4 – 5	Korrelationskoeffizient	Signifikanz (2-seitig)
DRT/WRT	0,79	<0,001
DRT/HSP	0,64	0,001
WRT/HSP	0,60	<0,05

**Tabelle 8: Höhe der Korrelation zwischen DRT, WRT und HSP der 4.–5. Klasse:**  
n=22 (Rangkorrelation nach Spearman).

		DRT	WRT	HSP
n	Gültig	22	22	22
Median		42,5	39	44
Perzentile	25	35,5	32,8	35,8
	75	47	45	48,8
Minimum		20	27	25
Maximum		62	51	55

**Tabelle 9: Deskriptive Statistik:** DRT, WRT, HSP der 4.-5. Klasse in T-Werten.



**Abbildung 3: Paarvergleich von DRT, WRT, HSP der 4.-5. Klasse (in T-Werten):**

n=22, Median, 1. u. 3. Quartil,

\* p<0,05: vs. WRT,  
(Wilcoxon -Test).

Die verwendeten Rechtschreibtests (DRT, WRT, HSP) korrelieren mittel bis hoch untereinander (0,6 - 0,8). (s. Tab. 8).

Es besteht ein signifikanter Unterschied zwischen dem DRT und dem WRT. In dem DRT Rechtschreibtest erreichen die Kinder der 4. – 5. Klasse höhere T-Werte als in dem WRT Rechtschreibtest (s. Tab. 9, Abb. 3).

### Klassen 1 – 5 getrennt

Nun war natürlich interessant, ob die Unterschiede, die sich in der Zusammenfassung der Grundschuljahre aufzeigen ließen, auch für die einzelnen Klassen zutreffen würden. Eine getrennte Darstellung der einzelnen Klassen sollte dies zeigen. Zur Beantwortung dieser Frage musste also der Zusammenhang zwischen den Ergebnissen der Rechtschreibtests und die Ergebnisse hinsichtlich ihrer Strenge innerhalb der einzelnen Klassenstufen überprüft werden.

Rechtschreibtests	1. Klasse		2. Klasse	
	Korrelationskoeffizient	Signifikanz (2-seitig)	Korrelationskoeffizient	Signifikanz (2-seitig)
DRT/WRT	0,96	<0,001	0,98	<0,001
DRT/HSP	0,78	0,01	0,93	<0,01
DRT/SLRT	-0,80	<0,01	0,80	<0,05
WRT/HSP	0,68	<0,05	0,95	0,001
WRT/SLRT	-0,88	<0,05	0,87	0,01
HSP/SLRT	-0,54	0,13	0,84	<0,05

**Tabelle 10: Höhe der Korrelation zwischen DRT, WRT, SLRT und HSP der 1. u. 2. Klasse:**  
n(1.Kl.)=9, n(2.Kl.)=7 (Rangkorrelation nach Spearman).

Rechtschreibtests	3. Klasse		4. Klasse	
	Korrelationskoeffizient	Signifikanz (2-seitig)	Korrelationskoeffizient	Signifikanz (2-seitig)
DRT/WRT	0,70	<0,05	0,76	<0,01
DRT/HSP	0,85	<0,01	0,91	<0,001
DRT/SLRT	0,98	<0,001	0,80	0,001
WRT/HSP	0,82	<0,01	0,65	<0,05
WRT/SLRT	0,72	<0,05	0,71	<0,01
HSP/SLRT	0,87	0,001	0,77	<0,01

**Tabelle 11: Höhe der Korrelation zwischen DRT, WRT, SLRT und HSP der 3. u. 4. Klasse:**  
n(3.Kl.)=10, n(4.Kl.)=13 (Rangkorrelation nach Spearman).

Rechtschreibtests	5. Klasse	
	Korrelationskoeffizient	Signifikanz (2-seitig)
DRT/WRT	0,77	<0,05
DRT/HSP	0,35	0,36
WRT/HSP	0,09	0,83

**Tabelle 12: Höhe der Korrelation zwischen DRT, WRT, SLRT und HSP der 5. Klasse:**  
n=9 (Rangkorrelation nach Spearman).

		1. Klasse				2. Klasse				3. Klasse			
		DRT	WRT	SLRT	HSP	DRT	WRT	SLRT	HSP	DRT	WRT	SLRT	HSP
N	Gültig	9	9	9	9	7	7	7	7	10	10	10	10
Median		39	39	47,5	57	35	38	36	30	36	32	38	46
Perzentile	25	33,5	31,5	34,8	41	29,5	30	29,5	27	28,1	27,5	29,5	41
	75	42,5	47,5	56,5	62	49	51	40	51	44,3	43,5	44,5	49
Minimum		24	29	24	28	26,5	30	24	23	24	24	24	32
Maximum		53	62	59	68	55	53	47,5	58	70	65	59	55

		4. Klasse				5. Klasse		
		DRT	WRT	SLRT	HSP	DRT	WRT	HSP
N	Gültig	13	13	13	13	9	9	9
Median		43	40	40	47	41	38	36
Perzentile	25	38	33	29,5	42,5	34,5	32,5	32,5
	75	49	47,5	47,5	51	47	40,5	42,5
Minimum		20	27	24	27	29	28	25
Maximum		62	51	59	55	48	44	46

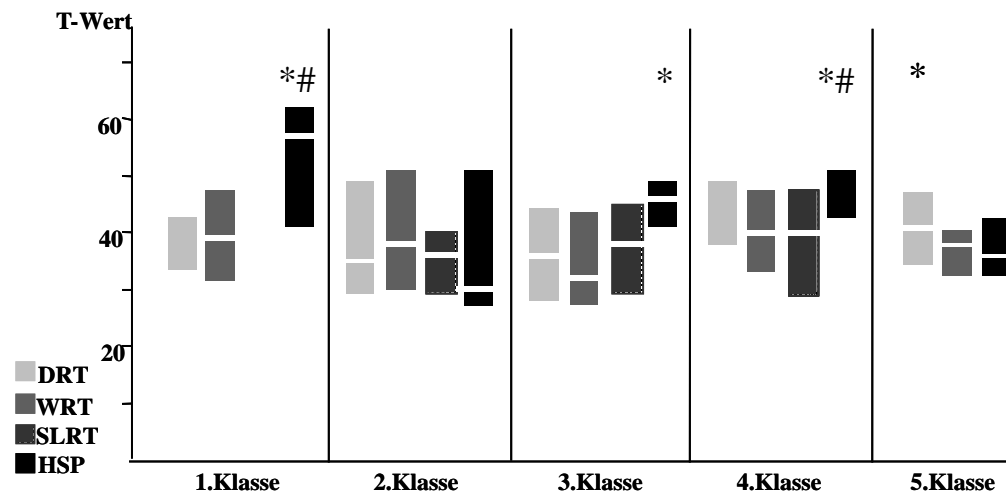
**Tabelle 13:**  
**Deskriptive Statistik:** DRT, WRT, SLRT, HSP in T-Werten, Klasse 1 – 5 getrennt dargestellt.

Rechtschreib-Tests	1. Klasse			2. Klasse			3. Klasse		
	Signifikanz (2-seitig)	Test mit höherem Median	Differenz d. Mediane	Signifikanz (2-seitig)	Test mit höherem Median	Differenz d. Mediane	Signifikanz (2-seitig)	Test mit höherem Median	Differenz d. Mediane
DRT/WRT	0,24	--	--	0,55	--	--	0,73	--	--
DRT/HSP	<0,01	HSP	18	0,27	--	--	0,07	--	--
DRT/(SLRT)	(0,17)	--	--	0,29	--	--	0,48	--	--
WRT/HSP	<0,05	HSP	18	0,45	--	--	<0,05	HSP	14
WRT/(SLRT)	(0,44)	--	--	0,11	--	--	0,77	--	--
HSP/(SLRT)	(0,37)	--	--	0,61	--	--	0,01	HSP	8

Rechtschreib-Tests	4. Klasse			5. Klasse		
	Signifikanz (2-seitig)	Test mit höherem Median	Differenz d. Mediane	Signifikanz (2-seitig)	Test mit höherem Median	Differenz d. Mediane
DRT/WRT	0,29	--	--	0,05	DRT	3
DRT/HSP	<0,05	HSP	4	0,09	--	--
DRT/SLRT	0,35	--	--			
WRT/HSP	<0,01	HSP	7	0,68	--	--
WRT/SLRT	1,00	--	--			
HSP/SLRT	<0,05	HSP	7			

Beschreibung s. Abbildung 4





**Abbildung 4: Paarvergleich der Rechtschreibtests DRT, WRT(1+-3+ u. 4/5-6+), SLRT, HSP der 1.-5. Klasse, jede Klassenstufen gesondert aufgeführt:**  
 Median, 1. u. 3. Quartil,  
 1. Klasse: n = 9, \*p<0,01: HSP1+ vs. DRT1, #p<0,05: HSP1+ vs. WRT1+;  
 2. Klasse: n = 7;  
 3. Klasse: n = 10, \*p<0,05: HSP3 vs. WRT3+ , SLRT3;  
 4. Klasse: n = 13, \*p<0,01: HSP4/5 vs. WRT4/5, #p<0,05: HSP4/5 vs. DRT4, SLRT4;  
 5. Klasse: n = 9, \*p=0,05: DRT5 vs. WRT6+;  
 (Wilcoxon-Test ).

Der DRT, die WRT und die HSP für die 1. Klasse zeigen signifikante hohe bis sehr hohe Zusammenhänge untereinander (0,7 - >0,9). Sehr interessant ist, dass sich beim Salzburger Rechtschreibtest durchweg negative Korrelationen (-0,8 - -0,5) darstellen lassen. Ganz anders sieht es dagegen in der 2. Klasse aus. Alle Testverfahren korrelieren signifikant hoch bis sehr hoch (0,8 - >0,9) miteinander. Ähnliches ergeben die Ergebnisse für die 3. und die 4. Klasse (r = 0,7 - >0,9). Bei Untersuchung der 5. Klasse lässt die Korrelation zwischen WRT und DRT einen hohen Zusammenhang (0,8) erkennen, die Hamburger Schreibprobe korreliert hingegen gering bis sehr gering sowohl mit DRT als auch mit dem WRT (0,1 - 0,4) (s. Tab. 10 - 12).

Bei der Untersuchung der 1.Klasse fanden sich ein sehr signifikanter Unterschied zwischen HSP und DRT und ein signifikanter Unterschied zwischen HSP und WRT. Die Kinder der 1. Klasse erreichen mit der Hamburger Schreibprobe deutlich höhere T-Werte als mit den Rechtschreibtests DRT und WRT. Da der Salzburger Rechtschreibtest der 1. Klasse gemäß den Ergebnissen der Korrelation völlig Unterschiedliches zu messen scheint, ist ein Vergleich mit den anderen Testverfahren der 1. Klasse hinsichtlich ihrer Strenge wenig aussagekräftig.

Bei der Untersuchung der 2. Klasse ergab sich kein Unterschied zwischen den Rechtschreibtests. Der DRT, die WRT, der SLRT und die HSP der 2. Klasse messen also gleich streng. Die Diagnostik der 3. Klasse ergab signifikante Unterschiede zwischen HSP und SLRT und HSP und WRT. Die HSP zeigt deutlich höhere T-Werte als die Rechtschreibtests WRT und SLRT. Bei den Testverfahren der 4. Klasse zeigte sich ein sehr signifikanter Unterschied zwischen HSP und WRT und signifikante Unterschiede zwischen HSP und DRT und HSP und SLRT. Die Testergebnisse der HSP zeigen deutlich höhere T-Werte als die Testergebnisse des WRT, des DRT und des SLRT. Bei den Testverfahren für die 5. Klasse stellte sich ein signifikanter Unterschied zwischen DRT und WRT heraus. Beim DRT erhielten die Kinder höhere T-Werte als mit dem WRT (s. Tab. 13 u. Abb. 4).

Zusammenfassend konnte gezeigt werden, dass die Hamburger Schreibprobe in der 1., 3. und 4. Klasse im Vergleich zu den anderen Testverfahren deutlich mildere Bewertungen der Rechtschreibleistungen erbringt.

Der DRT für die 5. Klassen zeigt eine etwas mildere Bewertung als der WRT für ebenfalls die 5. Jahrgangsstufe. Der Unterschied ist allerdings sehr gering (Differenz der Mittelwerte in T-Werten = 3). Zusätzlich korreliert die HSP in der 5. Klasse nur sehr gering mit dem WRT und dem DRT.

Des Weiteren fällt auf, dass der Salzburger Rechtschreibtest in den Klassen 2 - 4 sehr gut mit den anderen Rechtschreibtests übereinstimmt, nur in der 1. Klasse etwas völlig Unterschiedliches zu messen scheint.

### **5.1.3. Findet sich ein Zusammenhang zwischen DRT 1, WRT 1+, HSP 1+ und dem Erreichen des kritischen Wertes für nicht-lauttreue Fehler beim SLRT 1?**

Zwischen dem SLRT und dem DRT, WRT, HSP in den Klassen 2 – 4 zeigen sich hohe Korrelationen. Demgegenüber findet sich in der 1. Klasse eine hohe negative Korrelation. Die nicht-lauttreuen Fehler werden in den von dem Testverfahren angebotenen Prozentrangdarstellungen, in denen nur die orthographischen Fehler zählen, nicht berücksichtigt. Für die nicht-lauttreuen Fehler sieht das Testverfahren nur einen kritischen Wert vor. Dieser liegt bei T-Wert = 37. Es folgte eine Überprüfung inwiefern ein Zusammenhang zwischen der Höhe

der nicht-lauttreuen Fehler, beziehungsweise des Erreichens des kritischen Wertes des SLRT 1 und den Ergebnissen der anderen Rechtschreibtests der 1. Klasse vorliegt.

Hierfür wurden die Ergebnisse des SLRT in 3 Gruppen aufgeteilt: Die erste Gruppe hat den kritischen Wert für nicht-lauttreue Fehler nicht erreicht (T-Wert >37), die zweite Gruppe hat ihn genau erreicht (T-Wert = 37) und die dritte Gruppe liegt mit ihrer Fehleranzahl darüber (T-Wert <37).

Rechtschreibtests	Korrelationskoeffizient	Signifikanz (2-seitig)
SLRT/DRT	0,60	0,09
SLRT /WRT	0,55	0,12
SLRT/HSP	0,69	<0,05

**Tabelle 14: Höhe der Korrelation zwischen DRT1, WRT1+, HSP1+ und SLRT1-„nicht-lauttreue Fehler“:** n=9 (Rangkorrelation nach Spearman).

Die Darstellung der Höhe des Zusammenhangs zwischen dem kritischen Wert für „nicht-lauttreue Fehler“ des SLRT und den Testverfahren DRT, WRT, HSP der 1. Klasse zeigt positiv mittlere Korrelationen (0,6 - 0,7). Eine Signifikanz zeigt nur der Zusammenhang zwischen der HSP und dem SLRT, was sicherlich auch in der sehr niedrigen Fallzahl begründet ist (s. Tab. 14).

## 5.2. Erhält man mit den Lesetests gleiche Ergebnisse?

### 5.2.1. Erbringen die Subtests des SLRT, sowie die Subtests des Zürcher gleiche Ergebnisse?

Sowohl der Salzburger- als auch der Zürcher-Lesetest sind aus mehreren Subtests aufgebaut, die unterschiedliche Fähigkeiten messen. Um beide Testverfahren besser miteinander vergleichen zu können, ist die Bildung eines Mittelwerts der einzelnen Subtests notwendig. Da bei einem relativ hohen Zusammenhang zwischen den Untergruppen und keinem signifikanten Unterschied innerhalb der Gruppen ein Mittelwert ohne zu erwartende Verzerrungen gebildet werden könnte, wurden diese Berechnungen durchgeführt. Da nicht jeder Subtest von allen Klassenstufen bearbeitet wurde, entstanden bei den Berechnungen unterschiedliche Fallzahlen.

**a.) Der Salzburger Lesetest und seine Untertests:**

Der Salzburger Lesetest setzt sich aus 6 Subtests zusammen, die in Abhängigkeit von der jeweiligen Klassenstufe bearbeitet werden.

Subtest	Häufige Wörter	Zusammengesetzte Wörter	Text lang	Text kurz	Wortunähnliche Pseudowörter	Wortähnliche Pseudowörter
Zuordnung der Subtests zu den Klassenstufen	f. alle Klassen	3.+4. Klasse	3.+4. Klasse	1.+2. Klasse	f. alle Klassen	2.-4. Klasse

Subtests des SLRT	n	Korrelatinskoeffizient	Signifikanz (2-seitig)
häufige Wörter / zusammengesetzte Wörter	23	0,87	<0,001
häufige Wörter / Text lang	23	0,82	<0,001
häufige Wörter / Text kurz	16	0,91	<0,001
häufige Wörter / wortunähnliche Pseudowörter	39	0,89	<0,001
häufige Wörter / wortähnliche Pseudowörter	30	0,81	<0,001
zusammengesetzte Wörter / Text lang	23	0,80	<0,001
zusammengesetzte Wörter / wortunähnliche Pseudowörter	23	0,86	<0,001
zusammengesetzte Wörter / wortähnliche Pseudowörter	23	0,81	<0,001
Text lang / wortunähnliche Pseudowörter	23	0,73	<0,001
Text lang / wortähnliche Pseudowörter	23	0,71	<0,001
Text kurz / wortunähnliche Pseudowörter	16	0,90	<0,001
Text kurz / wortähnliche Pseudowörter	7	0,97	<0,001
Wortunähnliche- u./ wortähnliche Pseudowörter	30	0,94	<0,001

**Tabelle 15: Korrelation zwischen den Subtests 1. – 4. Klasse des SLRT Lesetest (Rangkorrelation nach Spearman).**

		Häufige Wörter	Zusammengesetzte Wörter	Text lang	Text kurz	Wortunähnliche Pseudowörter	Wortähnliche Pseudowörter
N		39	39	39	16	39	39
Median		43,5	43,5	43,5	39	41	43,5
Perzentile	25	30	37,3	39,3	29,5	35	37
	75	57	49	53,4	46,3	54	57
Minimum		24	24	24	24	24	24
Maximum		64	61	64	61	64	64

**Tabelle 16: Deskriptive Statistik:** Subtests des SLRT Lesetest in T-Werten, Klasse 1 – 4

	n	Signifikanz (2-seitig)
Subtests des SLRT	16 u. 39	0,75

**Tabelle 17: Vergleich der Subtests der 1.-4. Klasse des SLRT Lesetest (Rangvarianzanalyse nach KRUSKALL und WALLIS).**

Zwischen den Subtests des Salzburger Lesetests herrscht eine signifikante hohe bis sehr hohe positive Korrelation (0,7 - >0,9). Ein signifikanter Unterschied besteht zwischen den Subtests nicht (s. Tab. 15 - 17).

**b) Der Zürcher Lesetest und seine Untertests:**

Dieser Test bietet für die 2. - 4. Klasse Wort-Lesekarten und Textlesearten mit unterschiedlichen Schwierigkeitsgraden für die verschiedenen Jahrgangsstufen. Die Bewertung erfolgt nach der Lesezeit und den Lesefehlern. Für die 5. Klasse bietet das Testverfahren Gesamt-

Prozentränge für die Lesezeit und die Lesefehler an, so dass diese gesondert auf ihren Zusammenhang überprüft wurden.

Zürcher 2.-4. Klasse	n	Korrelationskoeffizient	Signifikanz (2-seitig)
Wort-Zeit / Wort-Fehler	39	0,78	<0,001
Wort-Zeit / Text-Zeit	39	0,90	<0,001
Wort-Zeit / Text-Fehler	39	0,63	0,010
Wort-Fehler / Text-Zeit	39	0,86	<0,001
Wort-Fehler / Text-Fehler	39	0,82	<0,001
Text-Zeit / Text-Fehler	39	0,75	<0,001

**Tabelle 18: Korrelation zwischen den Subtests d. 2. - 4. Klasse des Zürcher Lesetests** (Rangkorrelation nach Spearman).

Zürcher 5. Klasse	n	Korrelationskoeffizient	Signifikanz (2-seitig)
Lesezeit/Lesefehler	9	0,52	0,14

**Tabelle 19: Korrelation zwischen den Subtests d. 5. Klasse des Zürcher Lesetests** (Rangkorrelation nach Spearman).

		Wort-Zeit	Wort-Fehler	Text-Zeit	Text-Fehler
N	Gültig	39	39	39	16
Median		39	36	36	34
Perzentile	25	30,6	29,5	29,5	29,5
	75	53,5	50	51,9	47
Minimum		24	22	24	24
Maximum		58	58	58	58

**Tabelle 20: Deskriptive Statistik:** Subtests des Zürcher Lesetest d. 2. - 4. Klasse in T-Werten.

	n	Signifikanz (2-seitig)
Subtests Zürcher, 2.-4. Klasse	39	0,73

**Tabelle 21: Vergleich der Subtests des Zürcher Lesetest d. 2.-4. Klasse (Wort-Zeit/Wort-Fehler/Text-Zeit/Text-Fehler),** (Rangvarianzanalyse nach KRUSKALL und WALLIS).

		Lesezeit	Lesefehler
N	Gültig	9	9
Median		39	36
Perzentile	25	34,9	29,9
	75	41,5	42,9
Minimum		25	25
Maximum		47	47

**Tabelle 22: Deskriptive Statistik:** Subtests des Zürcher Lesetest d. 5. Klasse in T-Werten.

	n	Signifikanz (2-seitig)
Subtests Züricher, 5. Klasse	9	0,13

**Tabelle 23: Vergleich der Subtests des Zürcher Lesetest d. 5. Klasse** (Wilcoxon -Test).

Überwiegend hohe positive Korrelationen (0,6 - 0,9) mit höchst signifikanten Koeffizienten zeigen sich auch zwischen den Subtests des Zürcher-Lesetest. Die mangelnde Signifikanz in der 5. Klasse ist wahrscheinlich auf die sehr niedrige Fallzahl zurückzuführen.

Es besteht kein signifikanter Unterschied zwischen den Subtests der 2. – 5. Klasse (s. Tab. 18-23).

### 5.2.2. Wie unterscheiden sich die Lesetests verschiedener Klassenstufen?

#### **Klasse 1 – 4 zusammen**

Es wurden die Ergebnisse der Lesetests SLRT, Zürcher, WLLP und Knuspels Leseaufgaben der 1. - 4. Klasse jeweils zu einem Mittelwert zusammengeführt, um die Tendenz eines Testverfahrens innerhalb der Grundschuljahre zu erfassen. Aus den Zahlenwerten (T-Werte) der Subtests von SLRT und Zürcher wurden, bei einer nachgewiesenen hohen Übereinstimmung zwischen den Subtests, Mittelwerte gebildet. Da der Zürcher Lesetest erst ab der 2. Grundschulklasse anwendbar ist, konnte die 1. Klasse nicht berücksichtigt werden.

Knuspels Leseaufgaben bietet Normierungstabellen sowohl für die einzelnen Subtests als auch für zwei Scores an, die eine Zusammenfassung dieser Subtests darstellen. Zwei Scores sind notwendig, da in der 1. Klasse Subtest 4 nicht bearbeitet wird und somit der Score Vorläuferfertigkeit (VF) nur die Subtests 1 - 3 zusammenfasst, wohingegen der Score Lesefähigkeit (LF) alle Subtests beinhaltet. Als Folge davon enthalten beide Gruppen, die durch die Berechnung der Scores entstehen, unterschiedliche Fallzahlen.

- Ø Subtest 1 / Hörverstehen: Verstehen mündlich gestellter Fragen und Aufforderungen
- Ø Subtest 2 / Rekodieren: Erkennen von lautgleichen Wörtern
- Ø Subtest 3 / Dekodieren: Erkennen von Wortbedeutungen
- Ø Subtest 4 / Leseverstehen: Verstehen schriftlich gestellter Fragen und Aufforderungen
- Ø *Score „Vorläuferfertigkeit“: Subtest 1+2+3*
- Ø *Score „Lesefähigkeit“: Subtest 1+2+3+4*

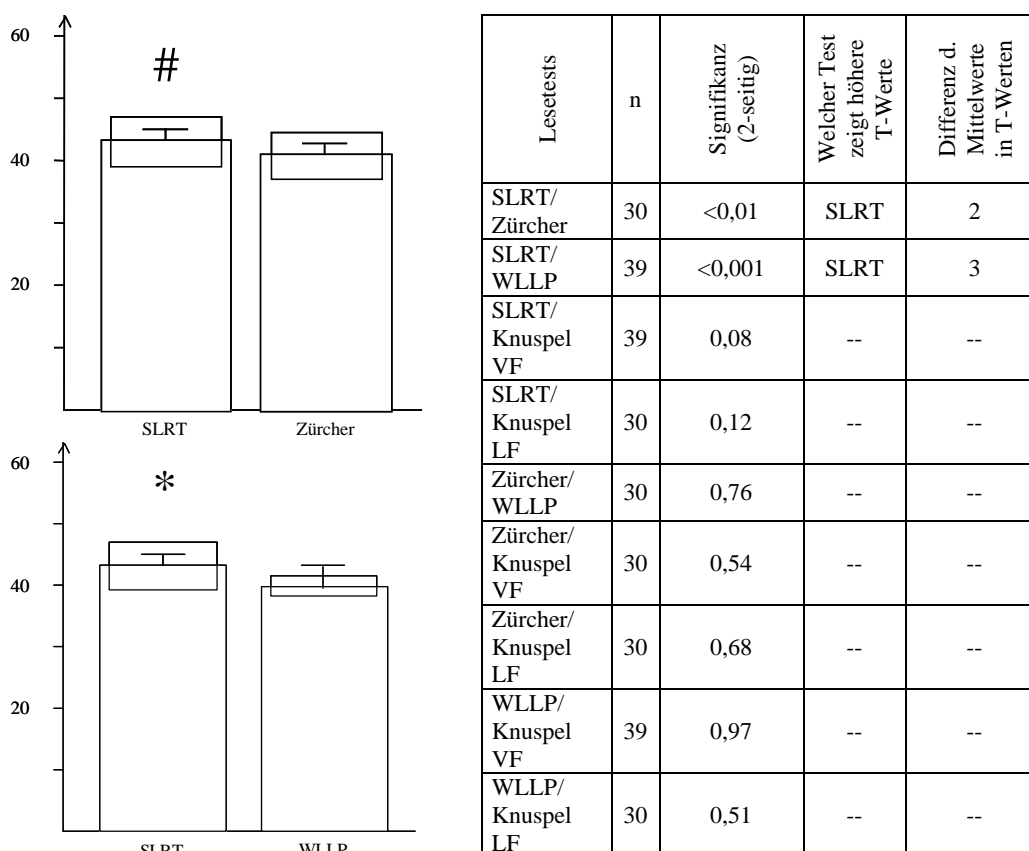
Von Interesse war nun der Zusammenhang zwischen den Lesetests und wie streng die Leseleistung von den Tests beurteilt wird.

Variable	n	Korrelationskoeffizient	Signifikanz (2-seitig)
SLRT / Zürcher	30	0,84	<0,001
SLRT / WLLP	39	0,87	<0,001
SLRT/Knuspel VF	39	0,50	<0,01
SLRT/Knuspel LF	30	0,66	<0,001
Zürcher / WLLP	30	0,72	<0,001
Zürcher/Knuspel VF	30	0,49	<0,01
Zürcher/Knuspel LF	30	0,68	<0,001
WLLP/Knuspel VF	39	0,52	0,001
WLLP/Knuspel LF	30	0,65	<0,001

**Tabelle 24: Höhe der Korrelation zwischen den Lesetests: SLRT 1.-4. Klasse, Zürcher 2.-4. Klasse, WLLP 1.-4. Klasse, Knuspel's Scores (Vorläuferfertigkeit-VF 1.-4.Klasse, Lesefähigkeit-LF 2.-4.Klasse), (Korrelationskoeffizient nach Pearson).**

		SLRT	Zürcher	WLLP	Knuspel's VF	Knuspel's LF
n		39	30	39	39	30
Mittelwert+SEM		43,2±1,9	41±1,8	39,9±1,7	40±1,6	41,6±2
Standardabweichung		11,6	9,8	10,4	10,2	11
95 % Konfidenzintervall des Mittelwerts	Untergrenze	39,4	37,4	36,6	35,8	37,5
	Obergrenze	47	44,6	43,3	43,8	45,8

**Tabelle 25: Deskriptive Statistik: SLRT 1.-4. Klasse, Zürcher 2.-4. Klasse, WLLP 1.-4. Klasse, Knuspel's Scores (Vorläuferfertigkeit-VF 1.-4. Klasse, Lesefähigkeit-LF 2.-4. Klasse) in T-Werten.**



**Abbildung 5: Paarvergleich der Lesetests: SLRT 1.-4. Klasse, Zürcher 2.-4. Klasse, WLLP 1.-4. Klasse, Knuspel's Scores (Vorläuferfertigkeit-VF 1.-4.Klasse, Lesefähigkeit-LF 2.-4.Klasse) :**

MW ± SEM, 95 % CI,

#p<0,01: vs. Zürcher und \*p<0,001: vs. WLLP.

(T-Test bei gepaarten Stichproben).

Zwischen den Lesetests SLRT, Zürcher und der Würzburger Leisen Leseprobe liegen relativ übereinstimmend hohe signifikante Korrelationen vor (0,7 - 0,9). Knuspels Leseaufgaben weichen in Bezug auf den Score Vorläuferfertigkeit mit Korrelationen bei maximal 0,5 davon ab. Der Score Lesefähigkeit erreicht wieder etwas höhere Werte ( $r = 0,7$ ) (s. Tab. 24).

Zusätzlich bestehen ein sehr signifikanter Unterschied zwischen dem Salzburger- und dem Zürcher-Lesetest, sowie ein höchst signifikanter Unterschied zwischen dem Salzburger Lesetest und der WLLP. Die Kinder erhielten stets durch den SLRT höhere T-Werte (s. Tab. 25 u. Abb. 5).

### **Klassen 1 – 4 getrennt**

Als Nächstes sollte überprüft werden, ob sich die Ergebnisse, die durch die Zusammenfassung der Grundschuljahre entstanden sind, in den einzelnen Klassen bestätigen lassen. Aus den Zahlenwerten (T-Werte) der Subtests von SLRT und Zürcher wurden wieder Mittelwerte gebildet. Von Interesse ist erneut die Höhe des Zusammenhangs zwischen den Lesetests und mit welcher Strenge sie in den Klassen 1 – 4 messen.



	Korrelatinskoeffizient	Signifikanz (2-seitig)
<b>SLRT/Zürcher</b>		
2. Klasse	0,64	0,119
3. Klasse	0,90	<0,001
4. Klasse	0,85	<0,001
<b>SLRT/WLLP</b>		
1. Klasse	0,80	<0,01
2. Klasse	0,89	<0,001
3. Klasse	0,94	<0,001
4. Klasse	0,79	<0,001
<b>SLRT/Knuspel VF</b>		
1. Klasse	0,66	0,05
2. Klasse	0,71	0,05
3. Klasse	0,55	0,08
4. Klasse	0,33	0,25
<b>SLRT/Knuspel LF</b>		
2. Klasse	0,70	0,05
3. Klasse	0,69	<0,05
4. Klasse	0,63	<0,05
<b>Zürcher/WLLP</b>		
2. Klasse	0,68	0,07
3. Klasse	0,83	<0,001
4. Klasse	0,78	<0,001
<b>Zürcher/Knuspel VF</b>		
2. Klasse	0,71	0,05
3. Klasse	0,69	<0,05
4. Klasse	0,37	0,19
<b>Zürcher/Knuspel LF</b>		
2. Klasse	0,68	0,07
3. Klasse	0,77	<0,01
4. Klasse	0,49	0,08
<b>WLLP/Knuspel VF</b>		
1. Klasse	0,50	0,17
2. Klasse	0,64	0,12
3. Klasse	0,59	0,08
4. Klasse	0,55	0,05
<b>WLLP/ Knuspel LF</b>		
2. Klasse	0,67	0,10
3. Klasse	0,82	<0,01
4. Klasse	0,66	<0,05

**Tabelle 26: Höhe der Korrelation zwischen SLRT 1.-4. Kl., Zürcher 2.-4. Kl., WLLP 1.-4. Kl., Knuspel's Scores (Vorläuferfertigkeit-VF 1.-4. Kl., Lesefähigkeit-LF 2.-4. Kl.), Klassenstufen getrennt aufgeführt: n(1.Kl.)=9, n(2.Kl.)=7, n(3.Kl.)=10, n(4.Kl.)=13, (Rangkorrelation nach Spearman).**

		1. Klasse			2. Klasse				
		SLRT	WLLP	Knuspel VF	SLRT	Zürcher	WLLP	Knuspel VF	Knuspel LF
N	Gültig	9	9	9	7	7	7	7	7
Median		36	35	40	38,1	43,5	40	44	41
Perzentile	25	27	26,5	36,5	31,1	34	35	34	37
	75	56,1	51	47,5	47	48,9	48	53	54
Minimum		26,5	22	24	25,4	29,5	26,5	31	36
Maximum		62	61	53	55,8	58,3	53	60	56

		3. Klasse					4. Klasse				
		SLRT	Zürcher	WLLP	Knuspel VF	Knuspel LF	SLRT	Zürcher	WLLP	Knuspel VF	Knuspel LF
N		10	10	10	10	10	13	13	13	13	13
Median		44,9	34,5	42,5	36,5	40	45,2	47,8	37	39	44
Perzentile	25	36,5	27,8	33,6	31	28	35,2	38	34,5	32,5	40
	75	58,8	49,3	49,8	46	44,3	52,8	49,9	45	40	46
Minimum		25,1	25,4	23	27	22	26,9	31,5	29,5	17	20
Maximum		62,8	58	56	70	71	60,6	55,8	62	58	58

**Tabelle 27: Deskriptive Statistik:** SLRT 1.-4. Kl., Zürcher 2.-4. Kl., WLLP 1.-4. Kl., Knuspel's Scores (Vorläuferfertigkeit-VF 1.-4. Kl., Lesefähigkeit-LF 2.-4. Kl.), Klassenstufen getrennt dargestellt in T-Werten.

	SLRT/ Zürcher	SLRT/ WLLP	SLRT/ Knuspel VF	SLRT/ Knuspel LF	Zürcher/ WLLP	Zürcher/ Knuspel VF	Zürcher/ Knuspel LF	WLLP/ Knuspel VF	WLLP/ Knuspel LF
Klassen 1.-4.: Signif. (2-seitig)		1. Kl.: 0,38	1. Kl.: 0,91					1. Kl.: 0,52	
		2. Kl.: 0,13	2. Kl.: 0,30	2. Kl.: 0,69	2. Kl.: 0,80	2. Kl.: 0,81	2. Kl.: 0,08	2. Kl.: 0,08	2. Kl.: 0,17
		<b>3. Kl.: &lt;0,01*</b>	<b>3. Kl.: &lt;0,05#</b>	3. Kl.: 0,11	<b>3. Kl.: &lt;0,05#</b>	3. Kl.: 0,43	3. Kl.: 0,77	3. Kl.: 0,77	3. Kl.: 0,59
		4. Kl.: 0,81	4. Kl.: 0,06	4. Kl.: 0,08	4. Kl.: 0,31	<b>4. Kl.: &lt;0,05#</b>	<b>4. Kl.: &lt;0,05#</b>	4. Kl.: 0,27	4. Kl.: 0,43
Welcher Test zeigt höhere T-Werte	3. Kl.: SLRT	3. Kl.: SLRT		3. Kl.: SLRT	4. Kl.: Zürcher	4. Kl.: Zürcher			
Differenz d. Mittel- werte in T-Werten	3. Kl.: 10	3. Kl.: 2		3. Kl.: 5	4. Kl.: 11	4. Kl.: 9			

**Tabelle 28: Paarvergleich der Lesetests SLRT 1.-4. Kl., Zürcher 2.-4. Kl., WLLP 1.-4. Kl. und Knuspel (Score Vorläuferfertigkeit-VF 1.-4. Kl., Lesefähigkeit-LF 2.-4. Kl.), jede Klassenstufe gesondert aufgeführt; bei vorhandener Signifikanz wird in Zeile 3 der Test ausgewiesen, der höhere T-Werte erreicht. In der darauffolgenden Zeile wird die Differenz der Mittelwerte in T-Werten für diesen Test angegeben:**

1. Klasse: n=9, n.s.
2. Klasse: n=7, n.s.
3. Klasse: n=10, \*p<0,01: SLRT vs. Zürcher, #p<0,05: SLRT vs. WLLP und SLRT vs. Knuspel-LF
4. Klasse: n=13, #p<0,05: Zürcher vs. WLLP und Zürcher vs. Knuspel-VF (Wilcoxon-Test).

Bei der Testung der 1. Klasse ergeben sich zwischen den Lesetests mittlere bis hohe (0,5 - 0,8) Korrelationen. Zwischen Lesetests der 2. Klassen findet man mittlere bis meist hohe (0,7 - 0,9) Zusammenhänge. Am besten ist die Übereinstimmung bei den 3. Klassen. Dort zeigen sich mittlere bis überwiegend hohe und sehr hohe Korrelationen (0,6 - 0,9). Etwas geringer fallen die Ergebnisse bei den 4. Klassen aus. Dort herrscht eine Korrelation zwischen 0,3 - 0,8.

Einzelne geringe Korrelationen finden sich erneut in Verbindung mit dem Score Vorläuferfertigkeit von Knuspels Leseaufgaben. Er korreliert mit den anderen Lesetests zwischen 0,3 - 0,7. Der Score Lesefähigkeit korreliert wiederum etwas höher mit den anderen Lesetests (r: 0,5 - 0,7) (s. Tabelle 26).

Beim Vergleich der Lesetests schneiden Kinder der 3. Klasse im Salzburger Lesetest signifikant besser ab. Am ausgeprägtesten zeigt sich der Unterschied ( $<0,01$ ) zwischen dem Salzburger- und dem Zürcher-Lesetest. Beim Vergleich des Zürcher Lesetest mit der WLLP, sowie mit dem Knuspel Score Vorläuferfertigkeit fällt ein signifikanter Unterschied bei den Viertklässlern auf. Der Zürcher bewertet die Leseleistung milder (s. Tab. 28).

### **5.3. Kommen Lehrer, Eltern und Testverfahren bei der Einschätzung von Schulleistungen zu gleichen Ergebnissen?**

#### **5.3.1. Finden sich Unterschiede in der Leistungseinschätzung durch die Eltern und durch die Lehrer?**

##### **1. – 5. Klasse**

Die Lehrer wurden um eine Leistungseinschätzung auf einer Skala von 1 - 6 (sehr gut - ungenügend) gebeten. Die Eltern erhielten einen Fragebogen, in dem sie die Schulleistungen mittels 3 Abstufungen von gut-durchschnittlich-schlecht einschätzen konnten. Nun sollte geklärt werden, ob ein Zusammenhang zwischen diesen beiden Bewertungen vorliegt. Dafür wurden die Grundschulklassen zusammengefasst. Die unterschiedlichen Fallzahlen entstehen durch teilweise unvollständige Bewertungen von Seiten der Lehrer wie auch der Eltern.

Lehrer- u. Elternbeurteilung	n	Korrelationskoeffizient	Signifikanz (2-seitig)
Diktat	41	0,57	<0,001
Lesen	43	0,55	<0,001
Abschreiben	41	0,50	<0,001
Aufsatz	31	0,34	0,06
Schriftbild	45	0,46	<0,01
Mathe	44	0,69	<0,001

**Tabelle 29: Höhe der Korrelationen zwischen den Beurteilungen d. Schulleistungen durch die Lehrer und den Beurteilungen d. Schulleistungen durch die Eltern: 1. – 5. Klasse, (Rangkorrelation nach Spearman).**

Die Zusammenhänge zwischen Lehrer- und Elternurteil bewegen sich im geringen bis mittlerem Bereich ( $r = 0,3 - 0,7$ ). Am besten stimmen beide im Bereich Mathematik überein, gefolgt von der Diktatleistung und der Leseleistung. (s. Tab. 29)

### **5.3.2. Finden sich Unterschiede in der Leistungseinschätzung durch die Eltern, bzw. durch die Lehrer und durch standardisierte Testverfahren?**

Es sollte untersucht werden, inwieweit die Lehrer-, bzw. Elterneinschätzung mit den Ergebnissen der Testverfahren übereinstimmt. Die "Einschätzung" der Diktatleistung durch die Lehrer/Eltern wurde mit den Rechtschreibtests, die "Einschätzung" der Leseleistung durch die Lehrer/Eltern mit den Lesetests auf einen Zusammenhang hin überprüft.

Da außerdem in dieser Studie gezeigt werden konnte, dass der SLRT in der 1. Klasse bei alleiniger Verwendung der Prozenträge zu enormen Fehldiagnosen führen kann, wurde die 1. Klasse bei den nun folgenden Untersuchungen ausgeklammert. Da für die nicht-lauttreuen Fehler nur sehr grobe Prozenträge (T-Wert  $<37$ ,  $= 37$  u.  $>37$ ) angegeben werden, konnte auch keine Zusammenfassung zu einem Mittelwert mit den Prozenträgen für orthographische Fehler erfolgen. Bei Knuspels Leseaufgaben wurden wie gehabt, die Scores Vorläuferfertigkeit (1. - 4. Klasse) und die Lesefähigkeit (2. - 4. Klasse) verwendet (s. 5.2.).

Test	Lehrerurteil-Diktat			Elternurteil-Diktat		
	n	Korrelationskoeffizient	Signifikanz (2-seitig)	n	Korrelationskoeffizient	Signifikanz (2-seitig)
DRT 1 - 5	42	-0,51	<0,001	45	-0,47	0,001
WRT 1+ - 3+	22	-0,70	<0,001	23	-0,79	<0,001
WRT 4/5 - 6+	22	-0,44	0,041	22	-0,20	0,37
HSP 1 - 5	42	-0,59	<0,001	45	-0,63	<0,001
SLRT 2 - 4	30	-0,51	<0,01	30	-0,61	<0,001

**Tabelle 30:**

**Zusammenhang zwischen Lehrer- u. Elternbeurteilung des Diktats und den Rechtschreibtests:**

Beurteilung d. Schulleistungen durch die Lehrer und Beurteilung d Schulleistungen durch die Eltern, Rechtschreibtests DRT 1-5, WRT 1+-3+, WRT 4/5-6+, SLRT 2-4, HSP 1-5, (Rangkorrelation nach Spearman).

Die gegenläufigen Bewertungsskalen (T-Werte: 20 (schlecht) – 80 (sehr gut) / „Einschätzung“ der Lehrer/Eltern 1 bzw. 3 (sehr gut) – 6 bzw. 3 (schlecht)) führten zu negativen Korrelationen. Aufgrund der unterschiedlichen Klassenstufen, die uns durch die einzelnen Testverfahren zur Verfügung standen, ergaben sich unterschiedliche Fallzahlen.

Die Korrelationen zwischen Lehrer und Testverfahren liegen zwischen  $r = 0,4 - 0,7$ , zwischen Eltern und Testverfahren zwischen  $r = 0,2 - 0,8$ . Beide weisen mit der Weingartener Grundwortschatz Rechtschreibtestreihe der 1. - 3. Klasse die höchste, und mit der Westermann Rechtschreibtestreihe für die 4. - 5. Klasse die geringste Übereinstimmung auf (s. Tab. 30).

Tests	Lehrerurteil-Lesen			Elternurteil-Lesen		
	n	Korrelationskoeffizient	Signifikanz (2-seitig)	n	Korrelationskoeffizient	Signifikanz (2-seitig)
SLRT 1.-4.	37	-0,78	<0,001	39	-0,58	<0,001
Zürcher 2.-5	39	-0,57	<0,001	39	-0,38	<0,05
WLLP 1.-4..	37	-0,80	<0,001	39	-0,44	<0,01
Knuspel 1.-4. Vorläuferfertigkeit	37	-0,48	<0,01	39	-0,34	<0,05
Knuspel 2.-4. Lesefähigkeit	30	-0,71	<0,001	30	-0,39	<0,05

**Tabelle 31:**

**Zusammenhang zwischen Lehrer u. Elternbeurteilung des Lesens und den Lesetests:**

Beurteilung d. Schulleistungen durch die Lehrer und Beurteilung d Schulleistungen durch die Eltern, Lesetests: Zürcher 2.-5., SLRT 1.-4., WLLP 1.-4., Knuspel Vorläuferfertigkeit 1.-4. Klasse u. Lesefähigkeit der 2.-4. Klasse, (Rangkorrelation nach Spearman).

Mit den Lesetests stimmen die Lehrer mehr als die Eltern überein. Bei den Lehrern liegen die Korrelationen mit  $r = 0,5 - 0,8$  im mittleren bis hohen Bereich, bei den Eltern hingegen liegen sie eher im niedrigen bis mittleren Bereich ( $r = 0,4 - 0,6$ ). Am meisten stimmt die Lehrerschaft mit der Würzburger Leise Leseprobe, die Eltern hingegen eher mit den Ergebnissen des Salzburger Lesetests überein. Einig sind sie sich allerdings dahingehend, dass sie beide die



Rechtschreib- tests		5. Klasse (9 Kinder)								Σ	
Rechtschreib- tests	DRT	x				x	x		x	x	5
	WRT4/5-6+	x				x	x		x	x	5
	HSP	x	x			x	x		x	x	6
Lesetests	Zürcher	x	x			x	x	x	x	x	7

**Tabelle 32: Verteilung und Anzahl der LRS- Diagnosen:** Testergebnis T-Wert  $\leq 37$  und T-Wert Diskrepanz zum Gesamt-IQ  $\geq 12$  ( $\geq 1,5$  Standardabweichungen).

\*Der Intelligenztest konnte bei diesem Kind leider nicht durchgeführt werden.

Tabelle 32 gibt eine Übersicht über die Häufigkeit einer LRS - Diagnose, ausschließlich ausgehend von der Lese- und Rechtschreib- Testleistung, innerhalb der Studienpopulation.

#### 5.4.2. Wie hoch ist die Trefferquote einzelner LRS-Tests einzuschätzen?

Zur Einschätzung der Trefferquote eines Testverfahrens bieten sich die Sensitivität und die Spezifität, die Gesamttrefferquote, sowie der RATZ-Index (**R**elativer **A**nstieg der **T**refferquote gegenüber der **Z**ufallstrefferquote) an. Sie sollten nun für die in dieser Studie verwendeten Testverfahren berechnet werden. Zur Berechnung der Parameter benötigt man einen so genannten „Goldstandard“. Am ehesten konnte dieser durch das Lehrerurteil (aktuelle „Einschätzung“ der Schulleistungen durch die Lehrer) repräsentiert werden. Die Lehrer erhielten im Rahmen der Studie einen Beurteilungsbogen mit einer Notenskala von 1 (sehr gut) bis 6 (ungenügend), auf dem sie die aktuellen Schulleistungen ihrer Schüler und Studienteilnehmer eintrugen.

Definition des „Goldstandard“:  $\Rightarrow$  Mangelhafte Schulleistung: „Einschätzung“ der Lehrer  $\geq 4,5$

$\Rightarrow$  Mangelhaftes Testergebnis: T-Wert  $\leq 37$

Alle getesteten Klassenstufen einer Testreihe wurden zusammengefasst. Bei der 1. Klasse des Salzburger Rechtschreibtest wurde der kritische Wert für nicht-lauttreue Fehler verwendet, da die alleinige Verwendung der Prozenträge zu Fehldiagnosen führte (s. 5.1.3.).

## Rechtschreibtests:

	1. Klasse (9 Kinder)	2. Klasse (7 Kinder)	3. Klasse (10 Kinder)	3. Klasse (13 Kinder)	4. Klasse (9 Kinder)
DRT	* * x x * * x	x x x x x	x x x x x x	x x x x x x x	* x * X
WRT1+ -3+	* * x x * * x	x x x x	x x x x x x		
WRT4/5 -6+				x x x x x	* * X
SLRT	x * * x X x * * x	x x x x x	x x x x x x	x x x x x x	
HSP	* * * * * * x	x x x x x		x x x x x	x * x * X x
Diktat	* * x x * * x	x x x x	x x x x x x	x x x x x x	x * x x * x x

**Tabelle 33:**

Überblick über die mangelhaften Rechtschreibleistungen („Einschätzung“ der Lehrer  $\geq 4,5$  und Testverfahren T-Werten  $\leq 37$ ) der Kinder der 1.-5. Klasse.

\*Lehrerbeurteilungen bezüglich der Rechtschreibleistung fehlen (Verweigerung der Schule, Montessorischule).

SENSITIVITÄT %	„Einschätzung“ der Lehrer $\geq 4,5$	davon Test auffällig	davon Test unauffällig	
61	23	14	9	DRT1-5
82	11	9	2	WRT1+-3+
42	12	5	7	WRT4/5-6+
39	23	9	14	HSP1-5
67	18	12	6	SLRT1-4

SPEZIFITÄT %	„Einschätzung“ der Lehrer $< 4,5$	davon Test unauffällig	davon Test auffällig	
79	19	15	4	DRT1-5
73	11	8	3	WRT1+-3+
88	8	7	1	WRT4/5-6+
90	19	17	2	HSP1-5
59	17	10	7	SLRT1-4

**Tabelle 34:**

### Sensitivität und Spezifität der Rechtschreibtests der 1.-5. Klasse:

Sensitivität und Spezifität der Rechtschreibtests:

-Darstellung wie viele der Kinder mit einer mangelhaften Diktatleistung im Test ebenfalls eine mangelhafte Leistung (Test auffällig) oder keine mangelhafte Leistung (Test unauffällig) zeigten.

-Darstellung wie viele der Kinder mit einer nicht-mangelhaften Diktatleistung im Test ebenfalls eine nicht-mangelhafte Leistung (Test unauffällig) oder eine mangelhafte Leistung (Test auffällig) zeigten.

Die Sensitivitäten der Rechtschreibtests liegen zwischen 39 - 82 %. Führend ist die Weingarterner Rechtschreib-Testreihe, 1. - 3. Klasse mit 82 %. Am schlechtesten (39 %) erkennt die Hamburger Schreibprobe die von den Lehrern als auffällig eingestufteten Kinder.

Etwas gehobener fallen die Spezifitäten der Testverfahren aus. Sie befinden sich im Bereich zwischen 59 - 90 %, angeführt von der Hamburger Schreibprobe mit 90 %. Das Schlusslicht bildet der Salzburger Rechtschreibtest mit 59 % (s. Tab. 34).



Testverfahren	Summe der valid positiven und valid negativen Kinder	Summer aller Kinder	Gesamttrefferquote %
DRT1-5	14+15	42	69
WRT1+-3+	9+8	22	77
WRT4/5-6+	5+7	20	60
HSP1-5	9+17	42	62
SLRT1-4	12+10	35	63

**Tabelle 35:**

**Darstellung der Gesamttrefferquote eines Testverfahrens.**

Valid positiv = von den Lehrern und im Test als auffällig eingestufte Kinder („Einschätzung“ der Lehrer  $\geq 4,5$  und Testverfahren T-Werten  $\leq 37$ )

Valid negativ = von den Lehrern und im Test als unauffällig eingestufte Kinder („Einschätzung“ der Lehrer  $< 4,5$  und Testverfahren T-Werten  $> 37$ )

Die Gesamttrefferquoten der Testverfahren variieren zwischen 60 und 77 %. Die WRT 1+ - 3+ führt mit 77 %. Das Schlusslicht bildet der WRT 4/5 - 6+ mit 60 % (s. Tab.35).

Testverfahren	RATZ-Index (Relativer Anstieg der Trefferquote gegenüber der Zufallstrefferquote) in %
DRT1-5	51
WRT1+-3+	60
WRT4/5-6+	17
HSP1-5	18
SLRT1-4	27

**Tabelle 36:**

**Darstellung des Relativen Anstiegs der Trefferquote gegenüber der Zufallstrefferquote eines Testverfahrens.**

Der RATZ-Index liegt bei den Rechtschreibtests zwischen 17 (WRT 4/5 - 6+) und 60 % (WRT 1+ - 3+) (s. Tab. 36).

**Lesetests:**

	1. Klasse (9 Kinder)				2. Klasse (7 Kinder)				3. Klasse (10 Kinder)				4. Klasse (13 Kinder)				5. Klasse (9 Kinder)												
SLRT		x	x	x	*	*	x		x			x			x	x													
Zürcher								x		x	x		x	x	x	x	x	x	x				*	*	x	x			
WLLP	x		x	x	x	*	*	x		x			x	x	x	x	x	x	x	x		x	x	x	x		*	*	
Knuspel VF		x		x		*	*	x	x			x	x	x	x	x	x		x			x	x	x	x		*	*	
Knuspel LF								x	x			x										x			x		*	*	
Lese- "Einschätzung" der Lehrer			x	x		*	*	x						x												x		*	*

**Tabelle 37: Überblick über die mangelhaften Leseleistungen („Einschätzung“ der Lehrer  $\geq 4,5$  und Testverfahren T-Wert  $\leq 37$ ) der Kinder der 1.-5. Klasse.**

\*Lehrerbeurteilungen bezüglich der Rechtschreibleistung fehlen (Verweigerung der Schule, Montessorischule).

<b>SENSITIVITÄT %</b>	<i>„Einschätzung“ der Lehrer <math>\geq 4,5</math></i>	<i>davon Test auffällig</i>	davon Test unauffällig	
80	5	4	1	Zürcher2.-5.
86	7	6	1	SLRT1.-4.
100	7	7	0	WLLP1.-4.
57	7	4	3	Knuspel VF1.-4.
100	4	4	0	Knuspel LF2.-4.

<b>SPEZIFITÄT %</b>	<i>„Einschätzung“ der Lehrer <math>&lt; 4,5</math></i>	<i>davon Test unauffällig</i>	davon Test auffällig	
63	32	20	12	Zürcher2.-5.
80	30	24	6	SLRT1.-4.
63	30	19	11	WLLP1.-4.
60	30	18	12	Knuspel VF1.-4.
81	26	21	5	Knuspel LF2.-4.

**Tabelle 38:**

**Sensitivitäten und Spezifitäten der Lesetests der Kinder der 1.-5. Klasse:**

Sensitivität und Spezifität der Lesetests:

-Darstellung wie viele der Kinder mit einer mangelhaften Leseleistung im Test ebenfalls eine mangelhafte Leistung (Test auffällig) oder keine mangelhafte Leistung (Test unauffällig) zeigten.

-Darstellung wie viele der Kinder mit einer nicht-mangelhaften Leseleistung im Test ebenfalls eine nicht-mangelhafte Leistung (Test unauffällig) oder eine mangelhafte Leistung (Test auffällig) zeigten.

An Sensitivität erreichen die Lesetests zwischen 57 - 100 %. Die Spitze bilden die Würzburger Leise Leseprobe und der Score Lesefähigkeit von Knuspels Leseaufgaben mit 100 %, wohingegen wieder einmal am Schluss der Score Vorläuferfertigkeit von Knuspels Leseaufgaben zu finden ist.

Die Spezifitäten variieren zwischen 60 - 81 %. Spitzenreiter ist wieder der Score Lesefähigkeit von Knuspels Leseaufgaben mit 81 %. Leider erreicht der Knuspel Score Vorläuferfertigkeit erneut die niedrigste Punktzahl (s. Tabelle 38).

<b>Testverfahren</b>	Summe der valid positiven und valid negativen Kinder	Summe aller Kinder	<b>Gesamttrefferquote %</b>
Zürcher2.-5.	4+20	37	65
SLRT1.-4.	6+24	37	81
WLLP1.-4.	7+19	37	70
Knuspel VF1.-4.	4+18	37	60
Knuspel LF2.-4.	4+21	30	83

**Tabelle 39:**

**Darstellung der Gesamttrefferquote eines Testverfahrens.**

Valid positiv = von den Lehrern und im Test als auffällig eingestufte Kinder („Einschätzung“ der Lehrer  $\geq 4,5$  und Testverfahren T-Werten  $\leq 37$ )

Valid negativ = von den Lehrern und im Test als unauffällig eingestufte Kinder („Einschätzung“ der Lehrer  $< 4,5$  und Testverfahren T-Werten  $> 37$ )

Die Gesamttrefferquoten der Lesetests liegen im Bereich zwischen 65 und 83 %. Knuspels Lesefähigkeit zeigt die höchste Gesamttrefferquote (83 %), gefolgt von dem SLRT (81 %), der WLLP (70 %), dem Zürcher (65 %) und Knuspels Vorläuferfertigkeit (60 %) (s. Tab. 39).

Testverfahren	RATZ-Index (Relativer Anstieg der Trefferquote gegenüber der Zufallstrefferquote) in %
Zürcher2.-5.	65
SLRT1.-4.	79
WLLP1.-4.	100
Knuspe VF1.-4.	25
Knuspe LF2.-4.	100

**Tabelle 40:**  
Darstellung des Relativen Anstiegs der Trefferquote gegenüber der Zufallstrefferquote eines Testverfahrens.

Nach Berechnung des RATZ-Index (Relativer Anstieg der Trefferquote gegenüber der Zufallstrefferquote) in Prozent führt Knuspels Lesefähigkeit und die WLLP mit 100 %, gefolgt vom SLRT (79 %) und dem Zürcher mit 65 %. Das Schlusslicht bildet Knuspels Vorläuferfertigkeit mit 25 % (s. Tab. 40).

Den Ergebnissen zufolge sind die Lesetests zuverlässiger als die Rechtschreibtests. Es scheint, dass der RATZ-Index stark von der Sensitivität abhängt, die bei den Rechtschreibtests in unseren Untersuchungen deutlich niedriger liegt als bei den Lesetests.

Dies liegt möglicherweise daran, dass die Diktatleistung von den Lehrern strenger als die Leseleistung beurteilt wird. Zur Überprüfung dieser Annahme soll nun der „Cut-off“ für die Lehrer-Rechtschreibbeurteilung höher gesetzt werden, auf  $\geq 5$ , damit eine Gleichbewertung von Lese- und Rechtschreibleistung durch die Lehrer gegeben ist und die Werte für RATZ, Sensitivität, Spezifität und Gesamttrefferquote vergleichbar werden.

Definition des „Goldstandard“:  $\Rightarrow$  Mangelhafte Schulleistung: „Einschätzung“ der Lehrer  $\geq 5$   
 $\Rightarrow$  Mangelhaftes Testergebnis: T-Wert  $\leq 37$

	1. Klasse (9 Kinder)	2. Klasse (7 Kinder)	3. Klasse (10 Kinder)	4. Klasse (13 Kinder)	6. Klasse (9 Kinder)
DRT	* * x x * * x x	x x x x x x	x x x x x x	x x x x x x x	* x * x x
WRT1+ -3+	* * x x * * x	x x x x	x x x x x x		
WRT4/5 -6+				x x x x x x x	* * x
SLRT	x * * x x x * * x x	x x x x x	x x x x x x	x x x x x x	
HSP	* * * * * x x	x x x x	x x x x x x	x x x x x	x * x * x x
Diktat	* * x x * * x x		x x x x x x	x x x x x x	x * x x * x x

**Tabelle 41:**  
**Überblick über die mangelhaften Rechtschreibleistungen („Einschätzung“ der Lehrer  $\geq 5$  und Testverfahren T-Wert  $\leq 37$ ) der Kinder der 1.-5. Klasse.**  
 \*Lehrerbeurteilungen bezüglich der Rechtschreibleistung fehlen (Verweigerung der Schule, Montessorischule).

SENSITIVITÄT %	„Einschätzung“ der Lehrer $\geq 5$	Test auffällig	Test unauffällig	
65	20	13	7	DRT1-5
89	9	8	1	WRT1+-3+
46	11	5	6	WRT4/5-6+
40	20	8	12	HSP1-5
73	15	11	4	SLRT1-4

SPEZIFITÄT %	„Einschätzung“ der Lehrer $< 5$	Test unauffällig	Test auffällig	
77	22	17	5	DRT1-5
69	13	9	4	WRT1+-3+
89	9	8	1	WRT4/5-6+
86	22	19	3	HSP1-5
65	20	13	7	SLRT1-4

**Tabelle 42:**  
**Sensitivitäten und Spezifitäten der Rechtschreibtests der Kinder der 1.-5. Klasse:**

Sensitivität und Spezifität der Rechtschreibtests:

-Darstellung wie viele der Kinder mit einer mangelhaften Diktatleistung im Test ebenfalls eine mangelhafte Leistung (Test auffällig) oder keine mangelhafte Leistung (Test unauffällig) zeigten.

- Darstellung wie viele der Kinder mit einer nicht-mangelhaften Diktatleistung im Test ebenfalls eine nicht-mangelhafte Leistung (Test unauffällig) oder eine mangelhafte Leistung (Test auffällig) zeigten.

Die Sensitivitäten der Rechtschreibtests liegen zwischen 40 - 89 %. Führend ist die Weingartener Rechtschreib-Testreihe, 1. - 3. Klasse mit 89 %. Am schlechtesten (40 %) erkennt die Hamburger Schreibprobe die von den Lehrern als auffällig eingestuften Kinder wieder.

Etwas gehobener fallen die Spezifitäten der Testverfahren aus. Sie befinden sich im Bereich zwischen 65 - 89 %, angeführt vom Westermann Rechtschreibtest mit 89 %. Das Schlusslicht bildet der Salzburger Rechtschreibtest mit 65 % (s. Tab. 42).

Testverfahren	Summe der valid positiven und valid negativen Kinder	Summer aller Kinder	Gesamttrefferquote %
DRT1-5	13+17	42	71
WRT1+-3+	8+9	22	77
WRT4/5-6+	5+8	20	65
HSP1-5	8+19	42	64
SLRT1-4	11+13	35	69

**Tabelle 43:**

**Darstellung der Gesamttrefferquote eines Testverfahrens.**

Valid positiv = von den Lehrern und im Test als auffällig eingestufte Kinder („Einschätzung“ der Lehrer  $\geq 5$  und Testverfahren T-Werten  $\leq 37$ )

Valid negativ = von den Lehrern und im Test als unauffällig eingestufte Kinder („Einschätzung“ der Lehrer  $< 5$  und Testverfahren T-Werten  $> 37$ )

Die Gesamttrefferquoten der Testverfahren variieren zwischen 64 und 77 %. Die WRT 1+ - 3+ führt mit 77 %. Das Schlusslicht bildet die HSP 1 - 5 mit 64 % (s. Tab. 43).

Testverfahren	RATZ-Index (Relativer Anstieg der Trefferquote gegenüber der Zufallstrefferquote) in %
DRT1-5	39
WRT1+-3+	76
WRT4/5-6+	22
HSP1-5	19
SLRT1-4	45

**Tabelle 44:**

**Darstellung des Relativen Anstiegs der Trefferquote gegenüber der Zufallstrefferquote eines Testverfahrens.**

Ist der „Goldstandard“ als mangelhafte Schulleistung „Einschätzung“ der Lehrer  $\geq 5$  definiert, ergibt der RATZ-Index bei den Rechtschreibtests Werte zwischen 19 % (HSP 1 - 5) und 76 % (WRT 1+ - 3+) (s. Tab. 44).

## **6. DISKUSSION**

### **6.1. Zuverlässigkeit der gewählten Methode**

In der vorliegenden Studie wurden Kinder getestet, die sich aufgrund unterschiedlichster Diagnosen in der Klinik Hochried in medizinischer, psychologischer und sozialpädagogischer Behandlung befanden. Man muss davon ausgehen, dass sie durch ihre jeweiligen Krankengeschichten (u. a. chronische Erkrankungen) besonders belastet sind und somit nicht den „Durchschnittschüler“ repräsentieren können. Für den Vergleich der Testverfahren, war dies zu vernachlässigen, da sich die individuelle Belastung idealerweise in allen Testverfahren gleichermaßen widerspiegeln sollte. Wichtiger war, dass sich innerhalb der untersuchten Population verschiedene Leistungsniveaus fanden. Die Präsenz der verschiedenen Schulformen ließ dies annehmen.

Über die Hälfte (56 %) der Schüler erhielt während der Testungen einer Dauermedikation. 21 % davon waren primär ZNS- gängige Medikamente. Man muss annehmen, dass Medikamente und vor allem primär ZNS- gängige Medikamente ein Kind in Testsituationen beeinflussen. Der mögliche Einfluss eines Medikamentes auf eine Testsituation wurde relativ konstant gehalten, indem die Medikation während der Testungsphase nicht verändert werden durfte.

Die Stärke der vertretenen Schulklassen variierte zwischen 7 Schülern in der 2. Klasse und 13 Schülern in der 4. Klasse. Die unterschiedliche Zusammensetzung war uns durch die Zusammensetzung der in der Klinik Hochried stationär behandelten Kinder vorgegeben.

Zweisprachig aufgewachsene Kinder wurden nicht ausgeschlossen oder extra gruppiert, da es sich auch hier wiederum um einen Einfluss handelt, der auf alle Testverfahren gleichermaßen einwirkt. Die Schulen mussten deutschsprachig sein.

Der durch das Testverfahren vorgegebene Anwendungszeitraum sollte sich mit dem Zeitraum der Studiendurchführung decken, beziehungsweise nur wenig davon abweichen. Die Güte eines Testverfahrens kann natürlich nur innerhalb des vorgesehenen Anwendungszeitraums gewährleistet werden. In der Beurteilung der Ergebnisse wurde der Testungszeitpunkt, wenn er denn



Die Lese-Rechtschreibtests enthalten für die einzelnen Klassenstufen unterschiedliche Aufgaben. Um zu überprüfen, ob sich die Zuverlässigkeit der Tests zwischen den Varianten für die einzelnen Klassenstufen unterscheidet, wurden die Ergebnisse für die einzelnen Klassenstufen getrennt dargestellt. Die Gruppen der einzelnen Klassenstufen wurden dadurch sehr klein und eine Aussage erschwert.

Der Zürcher und der Salzburger Lesetest sind aus verschiedenen Subtests zusammengesetzt, die je einzelne Punktwerte mit entsprechenden Prozentwerten ergeben. Zum Vergleich mit den anderen Testverfahren, war ein Gesamtwert als T-Wert notwendig. So wurde aus den Prozentwerten ein Mittelwert gebildet und mittels Transformationstabelle in einen T-Wert umgewandelt. Wir konnten zeigen (s. 5.2.1.), dass die Subtests innerhalb eines Testverfahrens überwiegend hoch miteinander korrelieren (SLRT  $r = 0,7 - >0,9$ , Zürcher  $0,5 - 0,9$ ) und keine signifikanten Unterschiede zwischen den Subtests bestehen. Es ist somit nicht zu erwarten, dass durch die Bildung eines Mittelwertes aus den Prozenträngen, die durch die Subtests erreicht wurden, Verzerrungen entstehen. Entscheidend für dieses Vorgehen war, dass bei der operationalisierten Diagnostik für beispielsweise ein LRS-Gutachten auch ein Gesamtprozentrang notwendig ist und in der Praxis verwendet wird (Warnke et al. 2004, Leitlinien der Deutschen Gesellschaft für Kinder- und Jugendpsychiatrie und –psychotherapie 2003).

Um die Frage zu beantworten, ob verschiedene Lese- und Rechtschreibtests die Leistungen eines Kindes gleich bewerten, musste ein Kind innerhalb eines sehr kurzen Zeitraums mit allen Testinstrumenten in einer randomisierten Reihenfolge getestet werden. Ein kurzer Zeitraum war wichtig, da durch den zum Teil parallel laufenden Förderunterricht eine mögliche Verbesserung der Schulleistungen die Testergebnisse hätte beeinflussen können. Da sich die Kinder stationär in der Klinik Hochried befanden, war ein schneller Testdurchlauf möglich. Es ließ sich nicht ausschließen, dass ein durchgeführtes Testverfahren zu einer Leistungsveränderung führt und somit möglicherweise das Ergebnis des darauf folgenden Tests beeinflusst. Aus diesem Grund wurde die Testungsreihenfolge randomisiert.

Zur Beantwortung der Frage, wie zuverlässig ein LRS-Test eine Lese-Rechtschreibstörung erfasst, benötigten wir einen so genannten „Goldstandard“. Der „Goldstandard“ sollte ein zuverlässig-



siges Außenkriterium bezüglich der Einschätzung der Rechtschreib- und Leseleistung sein. Problematisch ist nun, dass es ein solch zuverlässiges Außenkriterium nicht gibt.

Zur Verfügung stehen die Beurteilung der Schulleistungen durch die Eltern und die Beurteilung der Schulleistungen durch die Lehrer. Beide unterliegen Störeinflüssen.

Was sollte ein Kind in welcher Klassenstufe bereits können? Den Eltern fehlt oft die Erfahrung, was sie von einem Kind in einer bestimmten Phase des Lesen und Schreibenlernens erwarten können. Isolierte Schulschwierigkeiten, wie z. B. eine Lese-Rechtschreibstörung, werden von Eltern signifikant seltener erkannt als durchgängige Schulprobleme (Schrader and Helmke, 2001). Eine erste Orientierung geben meist die Lehrer. Interessant ist, dass es Hinweise gibt, dass das Elternurteil durch das Lehrerurteil mehr beeinflusst wird als durch die Mitteilung von Ergebnissen standardisierter Lese- und Rechtschreibtests. Einer Befragung zufolge glaubten nur 50 % der Eltern von LRS-Kindern nach Mitteilung der Testergebnisse, dass ihre Kinder im Lesen und Schreiben weniger gut seien als Gleichaltrige (Schrader and Helmke, 2001).

Lehrer dagegen sind von dem Niveau der gerade unterrichteten Klasse abhängig. Auf Erfahrung aufbauende diagnostische Verfahren (Beobachtungen, Funktionsproben, informelle Tests) sind stets subjektiv. Oft fehlt ihnen die Erfahrung, was andere Schulklassen leisten oder wie sie sich verhalten (Sommer-Stumpenhorst, 1993).

Wie viel wissen die Lehrer über die individuellen Defizite ihrer Kinder? Eine Zusatzuntersuchung von Pisa ergab ein erschreckendes Ergebnis. Bei einer kleinen deutschen Stichprobe wurden Klassenlehrern der neunten Klasse Listen von Schülern vorgelegt und gebeten, besonders schwache Leser zu kennzeichnen. Zehn Prozent der schwachen Leser wurden von den Lehrern als schwach eingestuft (Otto, 2002).

Mehrere Untersuchungen konnten bereits nachweisen, dass Kriterien, nach denen Noten vergeben werden, sehr variieren. Im Jahre 2002 wurde eine Studie veröffentlicht, im Rahmen derer dieselben Aufsätze von verschiedenen Lehrern bewertet wurden. Die Streuung der Aufsatznoten für denselben Aufsatz reichte über mehrere Notenstufen (von „sehr gut“ bis „mangelhaft“), was für recht unterschiedliche Kriterien bei der Notenfindung spricht (Birkel and Birkel, 2002).

Interessant sind auch Ergebnisse aus dem Marburger Hochbegabtenprojekt, die sich auf die Qualität der Hochbegabtendiagnostik von Lehrern beziehen. Vergleichsgrundlage waren begabungsrelevante Testdaten. Die Lehrer waren nicht in der Lage, spezifische Fähigkeitsbereiche unabhängig voneinander einzuschätzen. Die empirisch gefundene Bandbreite der Übereinstimmungen zwischen Lehrerurteil und Testergebnis reicht von beinahe vollständigen Übereinstimmungen bis zu Nullkorrelationen (Wild, 1993).

All diese Untersuchungen zeigen, dass sowohl die Eltern-, wie auch die Lehrerberurteilung sehr kritisch einzuschätzende Parameter sind, die vielen Störeinflüssen unterliegen.

Auf der anderen Seite darf man aber auch den situativen und gesellschaftlichen Kontext nicht vergessen, der Schulnoten umgibt. Welcher Schulabschluss erfordert welche Leistungen? Letztendlich entscheidet die Schulnote über die Schulkarriere. Bei lese-rechtschreibschwachen Kindern ist der Schulerfolg sowohl im Grundschulalter, wie auch in den anschließenden Schulen erheblich beeinträchtigt. ESSER konnte zeigen, dass beispielsweise nur 3 % der Kinder mit einer Lese-Rechtschreibstörung das Gymnasium besuchen können, ein Viertel besucht die Realschule, über die Hälfte verbleibt auf der Hauptschule und ein Sechstel der Kinder muss eine Förderschule besuchen. Resultat des geringen Schulerfolges ist, dass im Alter von 18 Jahren mit einer um den Faktor 3 (12 % vs. 4 %) erhöhten Rate von Jugendarbeitslosigkeit zu rechnen ist (Esser et al., 2002).

Lehrkräfte haben die gesamte Persönlichkeit des Lernenden im Blick und beurteilen die Ergebnisse auf dem Hintergrund länger andauernder Lernprozesse und unter Einbeziehung der institutionellen und personellen Rahmenbedingungen (May, 2002<sup>2</sup>). Individuelle Schwierigkeiten wie beispielsweise Prüfungsangst, oder individuelle Tagesschwankungen können nur durch den Lehrer erkannt und in eine Beurteilung miteinbezogen werden (Warnke et al., 2002).

Es lässt sich also zusammenfassen, dass das Lehrerurteil, obwohl es die tatsächlichen Lese-Rechtschreibleistungen nur ungenau abbildet, einen wichtigen und unverzichtbaren Faktor bei der LRS-Diagnostik darstellt. Trotz seiner Ungenauigkeit wird es häufig als Außenkriterium zur Validierung von Lese-Rechtschreibtests herangezogen, da keine zuverlässigeren Kriterien zur Verfügung stehen.

Und auch für unsere Untersuchungen wählten wir aus diesen Gründen das Lehrerurteil als Goldstandard zur Überprüfung der Zuverlässigkeit von LRS-Tests aus. Die Lehrerbeurteilung wurde mittels einer Skala von 1 - 6 (sehr gut bis ungenügend) erfragt. Es handelt sich hierbei wie auch bei dem Anamnesebogen für die Eltern nicht um einen standardisierten Fragebogen. Für die Lehrer ist die Beurteilung von Schulleistungen mittels Notensystem eine vertraute und gängige Methode. Auch sind es die Noten, die über eine Schullaufbahn entscheiden. Die letzten Zeugnisse wurden nicht verwendet werden, da eine aktuelle Einschätzung vorliegen sollte.

Für die Eltern wurde eine vereinfachte Skalierung (gut-befriedigend-schlecht) verwendet, um eine Überforderung zu vermeiden.

## **6.2. Vergleichbarkeit der Ergebnisse von LRS-Tests**

Im Folgenden soll nun diskutiert werden, ob die ausgewählten Testinstrumente – Rechtschreibtest (DRT 1 - 5, WRT 1+ - 3+, WRT 4/5 - 6+, SLRT 1 - 4, HSP 1+ - 5) und Lesetests (Zürcher Lesetest 2 - 5, SLRT 1 - 4, WLLP 1 - 4 u. KNUSPEL-L 1 - 4) – die Rechtschreib- bzw. Leseleistung bei einem Kind gleichermaßen erfassen. Bei den Testverfahren wurden die in der Praxis gängigsten Verfahren ausgewählt (Hemminger, 2000), Leitlinien zur Diagnostik und Therapie von psychischen Störungen im Säuglings-, Kindes- und Jugendalter 2003, (Warnke and Hemminger, 2004), (Schmid, 2004a)).

Als Antwort auf die Frage wie vergleichbar die Ergebnisse von LRS-Tests sind, konnten wir feststellen, dass alle Rechtschreibtests über die Klassenstufen 1 - 5 weitgehend hoch ( $r = 0,6 - 0,9$ ) miteinander korrelieren. Auch die Lesetests für die 1. - 4. Klasse zeigten sich weitgehend übereinstimmend in ihren Ergebnissen ( $r = 0,7 - 0,9$ ). Die meisten Mittelwerte der Lese-Rechtschreibtestverfahren präsentierten sich ebenfalls vergleichbar.

Zusammenfassend kamen wir also zu dem Ergebnis, dass man mit verschiedenen Rechtschreib- und Lesetests weitgehend vergleichbare Beurteilungen erhält. Nun gibt es aber doch einzelne Testverfahren, die sich in ihrer Übereinstimmung von den Anderen unterscheiden.

Ein Testverfahren bildet in verschiedenen Klassenstufen hierbei eine Ausnahme: Mit der Hamburger Schreibprobe erreichten die Kinder signifikant bessere Ergebnisse als mit dem DRT, WRT (Weingartener und Westermann Rechtschreib-Testreihe) oder SLRT-RT.

Dieser Unterschied ließ sich in der 1. (HSP vs. DRT und WRT1+), 3. (HSP vs. SLRT und WRT3+), und 4. Klasse (HSP vs. DRT, WRT4/5 und SLRT), nachweisen. Die Unterschiede sind in der 1. und 3. Klasse am ausgeprägtesten.

Etwas geschwächt werden die Ergebnisse dadurch, dass die für die 1. und 4. Klasse verwendeten Normierungen der Hamburger Schreibprobe (1. Klasse: 15. - 21. und 4. Klasse: 15. - 23. Unterrichtswoche) früher liegen als der Testzeitpunkt (26. - 30. Unterrichtswoche). Eine beinahe identische Abweichung von den Normierungszeitpunkten findet sich beim WRT 4/5 (16. - 24. Unterrichtswoche). Der Diagnostische Rechtschreibtest weicht mit seinem Normierungszeitpunkt der 1. Klasse (34. - 39. Unterrichtswoche) um 6 Wochen nach hinten ab. Der DRT 4 hat wiederum eine sehr frühe Normierung in der 4. - 18. Unterrichtswoche.

Geht man nun davon aus, dass durch die unterschiedlich verteilten Normierungszeitpunkte bei der 1. Klasse die Ergebnisse der HSP 1 milder und die des DRT 1 strenger ausfallen, so bleibt immer noch die WRT 1+, deren Normierungszeitpunkt genau mit dem Untersuchungszeitpunkt übereinstimmt. Die Abweichung sowohl zwischen WRT und HSP, als auch zwischen DRT und HSP bezieht sich immerhin auf eine Differenz der Mediane von 18 T-Werten.

In der 3. Klasse weichen die Normierungen der Rechtschreibtests nicht von den Normierungszeitpunkten ab, so dass also davon ausgegangen werden kann, dass die HSP milder bewertet als der SLRT-RT oder die WRT 3+. TACKE et al. bestätigt dies für die 3. Klassen. Er konnte ebenfalls nachweisen, dass die HSP im Vergleich zu DRT und WRT die Rechtschreibleistung zu milde bewertet (Tacke et al., 2001). In der 4. Klasse finden sich nur früher als der Untersuchungszeitpunkt liegende Normierungszeitpunkte, so dass man auch hier davon ausgehen muss, dass die HSP zu einer milderen Bewertung der Rechtschreibleistungen gelangt.

Bei vielen Anwendern (auch in der Fachklinik Hochried) entstand der Eindruck, dass die HSP zu leicht ist, dass die Prozentwerte bzw. die T-Werte zu hoch ausfallen (Deimel 2002, Schulte-

Körne Forschungsgruppe Lese-Rechtschreibstörungen). In der Literatur wird diese Aussage teilweise kontrovers diskutiert (Tacke et al. 2001, May et al. 2001).

Die HSP ist 2002 mit neuer Normierung erschienen. Die Normwerte sind gegenüber der Vorversion aber nur unwesentlich strenger geworden, so dass die Kritik an diesem Verfahren noch nicht ausgeräumt scheint (Schulte-Körne, Forschungsgruppe Lese-Rechtschreibstörungen).

Insgesamt finden sich in der aktuelleren Literatur wenige Empfehlungen zur Verwendung der Hamburger Schreibprobe. HEMMINGER, der 2000 eine kritische Übersicht über testdiagnostische Verfahren veröffentlichte und der wiederum in Lehrbüchern oft zitiert wird (z. B. Eggers et al., 2004), erwähnt die HSP nicht. Auch in den aktuellen Leitlinien der deutschen Gesellschaft für Kinder- und Jugendpsychiatrie und –psychotherapie wird sie nicht erwähnt. Im aktuellen Brickenkamp Handbuch psychologischer und pädagogischer Tests (2002) wird die Hamburger Schreibprobe zur qualifizierten Erfassung des Rechtschreibniveaus ebenfalls nicht angeführt. WARNKE schreibt in seinem aktuellen diagnostischen Leitfaden, dass praktische Erfahrungen mit der HSP den Eindruck hinterlassen, dass die wahre Rechtschreibleistung tendenziell überschätzt wird (2004). Er bezieht sich hierbei allerdings auch noch auf die Ausgabe von 2000.

In einer großen Kinder- und Jugendpsychiatrischen Ambulanz wie z. B. in der Heckscher Klinik, Klinik für Kinder- und Jugendpsychiatrie und Psychotherapie in München, findet die Hamburger Schreib-Probe, obwohl sie zu den vorrätigen Testverfahren zählt, keine Verwendung. Als Grund hierfür wird vor allem die zu milde Bewertung der Rechtschreibleistung genannt. Rechtschreibschwache Kinder werden mit der Hamburger Schreibprobe nicht ausreichend erkannt und erhalten somit keine Förderung. Weiter wurde kritisiert, dass die HSP dem Diagnostiker im Vergleich zu anderen Rechtschreibtests weniger Aufschluss über das individuelle Fehlerprofil biete (z.B. hilft die HSP den Kindern bei der Groß- und Kleinschreibung durch die Vorgabe des bestimmten Artikel für das zu schreibende Testwort). Der Testaufbau, der von den anderen Testverfahren abweicht, wird eher als Doppelbelastung gesehen und nicht als Hilfe. Illustrationen veranschaulichen die zu schreibenden Testwörter bzw. –sätze, die vom Lehrer vorgelesen werden. Das Kurzzeitgedächtnis wird also zusätzlich beansprucht. Die Kinder dürfen jederzeit nachfragen, wenn sie ein Wort nicht mehr wissen, doch welches Kind meldet sich freiwillig in einer Gruppe und

gibt damit zu, dass es sich ein Wort nicht merken konnte. (Bauer 2005, Breuer 2005, Schmieder 2005).

Zusammenfassend kann die Verwendung der Hamburger Schreibprobe zur Überprüfung der Rechtschreibleistung nur eingeschränkt empfohlen werden. Ihre Ergebnisse sollten immer auf ihre Gültigkeit vor allem bezüglich einer zu milden Einschätzung der Rechtschreibleistung überprüft werden.

Nun zu weiteren Testverfahren, die von den anderen Rechtschreibtests in unserer Auswertung abzuweichen scheinen.

Der Diagnostische Rechtschreibtest scheint in der 5. Klasse im Vergleich zum WRT milder zu messen. Wahrscheinlich muss man hierbei von einem Artefakt ausgehen. Es ist anzunehmen, dass die Ursache in dem deutlich früheren Normierungszeitpunkt des DRT (4 -19 Unterrichtswoche) liegt. Passend hierzu ist auch, dass die Differenz der T-Wert Mediane nur 3 Punkte beträgt. Allerdings findet man in der Literatur Hinweise darauf, dass der WRT 6+ für den vorgesehenen Anwendungszeitraum zu hohe Anforderungen stellt und besser erst ab Mitte der 6. Klasse eingesetzt werden soll (Höger et al. 1999, Reuter-Liehr 2003).

Der SLRT- Rechtschreibtest erscheint in seiner Beurteilung der Rechtschreibleistung mit den anderen Rechtschreibtests vergleichbar, solange in der 1. Klasse nicht die orthographischen Fehler, sondern die nicht-lauttreuen Fehler als Hauptbewertungskriterium herangezogen werden. Die Prozentrangtabellen, die vom Testverfahren angeboten werden, beziehen sich nur auf die orthographischen Fehler. Bei alleiniger Verwendung der Prozentränge stellen sich jedoch zwischen Salzburger Rechtschreibtest und den anderen Rechtschreibtests in der 1. Klasse durchweg nur negative Korrelationen ( $r = -0,8 - -0,5$ ) dar. Für die nicht-lauttreuen Fehler und Fehler der Groß-Kleinschreibung bietet der Salzburger Rechtschreibtest nur einen kritischen Wert (T-Wert =37). Gerade bei Kindern der 1. Klasse, die die Rechtschreibung gerade erst zu lernen beginnen, liegt ein großer Anteil der Rechtschreibfehler in dem Bereich der nicht-lauttreuen Fehler. Verwendet man den kritischen Wert für die nicht-lauttreuen Fehler zur Beurteilung der Rechtschreibleistung, zeigt sich der SLRT auch in der 1. Klasse mit den anderen Testverfahren vergleichbar (Höhe der Korrelation zwischen DRT, WRT, HSP und SLRT-nicht-lauttreue Fehler der 1. Klasse  $r = 0,6 -$

0,7). Es muss also bei Verwendung dieses Testverfahrens unbedingt bedacht werden, dass Rechtschreibanfänger oder auch Kinder mit verzögertem Rechtschreibprozess wie bei einer Lese-Rechtschreibschwäche bei alleiniger Verwendung der Prozentränge für orthographische Fehler in ihrer Rechtschreibleistung überschätzt werden.

Es wird im Testheft empfohlen, die orthographischen Fehler als das Hauptbewertungskriterium heranzuziehen. Diese Empfehlung bezieht sich auf die Ergebnisse der Validierung, die im Rahmen der Testkonzeption durchgeführt wurde. Bei der Validierung, die an Zweitklässlern durchgeführt wurde, schnitten die von Lehrern als auffällig im Schreiben eingestuften Schüler im SLRT signifikant schlechter ab und der Unterschied war am deutlichsten bei den orthographischen Fehlern zu erkennen (Landerl, 2000).

DEIMEL schreibt 2002, dass der SLRT für den Rechtschreibteil nur sehr grobe Normen (Prozentrangbänder) zur Verfügung stellt und daher nicht empfohlen wird. In der Literatur wird der SLRT-RT vorrangig als Testinstrument bei bereits auffällig gewordenen Kindern empfohlen. Als Anwendungsbereich wird eine detaillierte Diagnose von Lese-Rechtschreibschwierigkeiten vor allem im unteren Leistungsbereich genannt, um spezifische Fördermaßnahmen zu erstellen, einzuleiten und durchzuführen (Brickenkamp 2002, Hemminger et al. 2000). Auch HASSELHORN und WARNKE nennen einen Vorteil des SLRT- Rechtschreibtest, dass er eine spezifische Diagnose von Schwierigkeiten beim lautorientierten Schreiben und dem orthographisch korrekten Schreiben ermöglicht (2000, 2004).

In der kinder- und jugendpsychiatrischen Praxis wird der SLRT Rechtschreibtest wenig verwendet. Als Grund hierfür wird vor allem die fehlende Gesamtnorm für alle Fehlerarten genannt. Ein Punktwert, wie er bei der operationalisierten Diagnostik zur Erstellung eines LRS-Gutachten notwendig ist, kann somit nicht erreicht werden kann. Die spezifische Fehleranalyse wird nicht als Vorteil gesehen (Schmieder, 2005).

Zusammenfassend kann festgehalten werden, dass sich in der Beurteilung der Rechtschreibleistung der Salzburger Rechtschreibtest von den anderen Testverfahren nicht zu unterscheiden scheint, solange alle drei Fehlerarten (orthographische, nicht-lauttreue und Fehler der Groß-Kleinschreibung) in die Bewertung des Testergebnisses mit einbezogen werden.

Nun zu den Lesetests, von denen sich auch einzelne Testverfahren in ihrer Beurteilung der Leseleistung von den Anderen unterscheiden.

Eine geringe Abweichung zu den anderen Testverfahren zeigt sich bei Knuspels Leseaufgaben und seinem Score Vorläuferfertigkeit ( $r = 0,5$  vs. SLRT, Zürcher, WLLP). In den Grundschuljahren liegt der Knuspel Score Vorläuferfertigkeit mit  $r = 0,5$  leicht unter der durchschnittlichen Übereinstimmung der anderen Lesetests ( $r = 0,7 - 0,9$ ). KNUSPEL-L misst neben dem Leseverständnis auch die Vorläuferfertigkeiten (Basisfertigkeiten der Lesefähigkeit: Hörverstehen, Rekodieren, Dekodieren), die mit den anderen Lesetests schlechter übereinstimmen.

Knuspels Leseaufgaben unterscheiden sich inhaltlich deutlich von den anderen Lesetests. Sie repräsentieren einen anspruchsvollen Begriff von Lesen im Grundschulalter und sie operationalisieren zugleich ein Prozessmodell des Lesens (Marx, 2000). Der Prozess zum Leseverstehen (Subtest 4) führt demzufolge über Rekodieren (Subtest 2) und Dekodieren (Subtest 3, Sinnzuweisen zur Lautung) einerseits und andererseits über das Hörverstehen (Subtest 1, auf der Grundlage von Sprachentwicklung und Kapazität des Arbeitsgedächtnisses). Und diese vier Instanzen bilden dann vier Subtests ab (Probst, 2003). Also überprüft das Testverfahren neben Basisfertigkeiten der Lesefähigkeit auch das Leseverständnis (Hemminger et al., 2000).

Betrachtet man das Testkonzept von Knuspels Leseaufgaben, handelt es sich also nur zu einem Teil um einen Lesetest. Das Verfahren beinhaltet vier Untertests, von denen sich drei mit Lesen im weiteren Sinn befassen. Der zweite und dritte Untertest erfasst auch sehr stark phonologische Fertigkeiten bzw. den passiven Wortschatz, und der erste Untertest hat mit Lesen nichts mehr zu tun (Ausführen mündlicher Anweisungen). Laut DEIMEL handelt es sich hierbei nicht um einen Lesetest im klassischen Sinn (2002).

Zu beachten ist, dass der Test mit seinen verschieden ausgerichteten Subtests sehr viele sprachliche Fähigkeiten mit beansprucht. Bei dem Untertest Rekodieren und Dekodieren ist eine Lösung nur durch eine vorangehende innerliche Versprachlichung möglich. Auch bei dem ersten Subtest „Hörverstehen“ wird das Sprachverständnis gefordert. Hierbei möchten wir auf die hohe Koinzidenz von LRS und Sprachstörungen hinweisen. 60 % der sprachgestörten Kinder weisen auch



eine Lese-Rechtschreib-Schwäche auf (Esser, 2002<sup>2</sup>). In der Praxis zeigt sich, dass vor allem sprachgestörte Kinder mit dieser Art Aufgaben völlig überfordert sind (Bauer, 2005).

In der testpsychologischen Praxis wird dieses Testverfahren kaum verwendet. Als Grund werden die Länge, die Komplexität und der Schwierigkeitsgrad (Basisfertigkeiten werden auf einem hohen Anspruchsniveau abgefragt), sowie die ausführliche Auswertung genannt. Die zusätzliche Forderung sprachlicher Fertigkeiten wird als Doppelbelastung gesehen. (Schmieder 2005, Bauer 2005, Breuer 2005).

Bei den Testungen zur Studiendurchführung zeigten die Kinder, verglichen mit den anderen Lese-Rechtschreibtestverfahren, wenig Freude an der Durchführung und wenig Motivation. Es entstand der Eindruck, dass sie vor der Länge und Komplexität der Aufgaben zurückschraken. Auch schien ihnen die Art der Aufgaben wenig vertraut. Zusammenfassend war unser Eindruck, dass dieses Testverfahren für Kinder, die Probleme im Lesen und Schreiben, vielleicht noch der Aufmerksamkeit und der Sprache haben, sehr schwierig ist und somit schnell zu einem Motivationsverlust führen kann. In der Studienpopulation bestand bei 33 % eine Lese-Rechtschreibschwäche, bei 25 % eine Aktivitäts- und Aufmerksamkeitsstörung und bei 8 % eine Sprachstörung.

Sehr vorteilhaft für das Testverfahren wirken sich allerdings seine guten Gütekriterien aus. Durchgeführt wurden umfassende Reliabilitäts- und Validitätsuntersuchungen, die die Stabilität des Verfahrens und seine ökologische und kriterienbezogene Validität auf allen Klassenstufen belegen (Marx, 2000).

Ein weiteres Testverfahren, das in seiner Bewertung auch von den anderen Lesetests abzuweichen scheint, ist der Salzburger Lesetest. Schaut man die einzelnen Klassen getrennt an, besteht dieser Unterschied nur in der 3. Klasse. Zwischen SLRT-LT und WLLP findet man die niedrigste Differenz der Mittelwerte (2 Punkte) und man muss relativierend hinzufügen, dass die Normierung der WLLP vier Wochen nach Ende der Untersuchungen beginnt. Alle anderen Lesetests weisen Normierungen für den Untersuchungszeitpunkt auf. Somit muss hierbei von einem Artefakt ausgegangen werden. Der Unterschied zwischen dem Salzburger Lesetest und dem Knuspel Score Lesefähigkeit ist gering. Bei Zusammenfassung der Grundschuljahre lässt sich dieser Unterschied nicht mehr nachweisen. Eine sichere Aussage kann daraus nicht abgeleitet werden.

Wichtig erscheint uns aber auch hier, die inhaltlichen Unterschiede der beiden Testverfahren zu beachten. Wie bereits erwähnt, überprüfen Knuspels Leseaufgaben neben Basisfertigkeiten der Lesefähigkeit auch das Leseverständnis (Hemminger et al., 2000). Der Knuspel Score Lesefähigkeit unterscheidet sich von dem Score Vorläuferfertigkeit in der zusätzlichen Messung des Leseverständnisses. Hierin liegt ein inhaltlicher Unterschied zum SLRT Lesetest, der die Laut-Graphemerkennung und die Lesesyntese und nicht das Leseverständnis überprüft. Ein weiterer Unterschied liegt natürlich in der Überprüfung der Basisfertigkeiten (Hörverstehen, Rekodieren, Dekodieren) des Leselernens, die bereits ausführlich diskutiert wurden.

Die größte Abweichung in der dritten Klasse besteht zwischen SLRT-LT und Zürcher (10 T-Wert Punkte). Die Ergebnisse des Zürcher im Vergleich zum SLRT Lesetest fallen etwas niedriger aus, die des SLRT im Vergleich etwas höher. Nach unseren Ergebnissen muss man am ehesten davon ausgehen, dass der SLRT in der 3. Klasse zu milde misst.

In der Literatur finden sich keinerlei Hinweise für eine zu milde Bewertung des SLRT-LT, wohingegen sich Hinweise finden, dass der Zürcher eine strengere Einschätzung der Rechtschreibleistung erbringen könnte. Die Normierungen des Zürcher Lesetest wurden zwischen 1973 und 1981 erhoben. Die Rechtschreibleistungen haben sich in den letzten 35 Jahren wesentlich verschlechtert (Strehlow and Haffner, 2002). Auch DEIMEL hebt diesen Effekt der Verschiebungen der Leistungsparameter in der Bevölkerung stark hervor. Er verwendet sogar eine Kategorisierung, nach der er ausschließlich anhand des Normierungsalters Testempfehlungen ausspricht. Testverfahren mit einer Normierung, die maximal 10 Jahre alt ist, können uneingeschränkt empfohlen werden. Ergebnisse von Testverfahren mit Normierungen älter als 10 Jahre, aber nicht älter als 15 Jahre sollten nur mit Vorsicht interpretiert werden. Bei Normierungen älter als 15 Jahre rät er generell von der Benutzung des Testverfahrens ab (Demel, 2002). Im Gegensatz zum Zürcher verfügt der SLRT über gute und aktuelle Gütekriterien. Die Normierung erfolgte 1997, ist also noch keine 10 Jahre alt.

Zugunsten des Zürcher spricht wiederum, dass er trotz veralteter Normen eine große Beliebtheit in der testpsychologischen Praxis erfährt und auch in aktuellen Leitlinien zur Lesediagnostik empfohlen wird (Warnke et al., 2004). In der testpsychologischen Praxis wird das Testverfahren weiterhin verwandt, da es schlechte Leser relativ sicher einzuschätzen scheint und es wenig gute

Alternativen gibt (Reuter-Liehr 2003, Schmieder 2005, Breuer 2005). Die Laut-Graphemerkennung und die Lesesyntese können sicher abgefragt werden. Die Pseudowörter werden dabei als hilfreich erlebt. Sie zeigen Defizite in der phonologischen Dekodierfähigkeit rasch auf, die als eine wesentliche Voraussetzung für den Lese- und Schreibprozess gesehen wird und bei der Lese-Rechtschreibschwäche oft schwach ausgebildet ist (Schulte-Körne et al. 1997, v. Suchodoletz 1999, Mayringer & Wimmer, 2000). Ob das Kind sinnentnehmend lesen kann, ist nach dem Lesen der Leseabschnitte des Zürcher Lesetests gut zu erfragen (Reuter-Liehr, 2003). Auch die einzelnen Subtests mit dazugehörigen Prozenträngen werden als Vorteil gesehen, da bei sehr schlechten Lesern nicht zwingend das gesamte Testverfahren durchzuführen ist, um einen Prozentrang zu erhalten (Schmieder 2005, Breuer 2005).

Zum Abschluss muss noch erwähnt werden, dass der Zürcher Lesetest in der 4. Klasse von der WLLP und Knuspels Vorläuferfertigkeit abweicht. Der Zürcher bewertet gegenüber der WLLP und dem Knuspel Score die Lesefähigkeit in der 4. Klasse milder. Dieser Unterschied lässt sich bei der Zusammenfassung der Grundschuljahre nicht mehr darstellen. Unserer Meinung nach muss auch hierbei aufgrund der unterschiedlichen Normierungszeitpunkte von einem Artefakt ausgegangen werden. Die Normierung des Zürcher Lesetest beginnt etwa 4 Wochen vor dem Zeitpunkt der Studiendurchführung. Die WLLP, die in ihrer Strenge mehr als Knuspels Leseaufgaben abweicht, hat den Beginn ihres Normierungszeitpunktes 1 Monat nach der Studiendurchführung. Die Normierungen von Knuspels Leseaufgaben decken sich mit der Studiendurchführung. Was ebenfalls für ein Artefakt spricht ist, dass die Normierungen des Zürcher für die vorangehenden Klassenstufen jeweils zum Studienzeitpunkt liegen und sich dabei keine solch milde Bewertung der Leseleistung im Vergleich zu den anderen Testverfahren zeigt.

Somit kann zusammengefasst werden, dass bis auf den SLRT-LT (und hier nur für die 3. Klasse) alle Lesetests relativ vergleichbare Ergebnisse erzielen und somit zur Erfassung der Leseleistung verwendet werden können.

#### FAZIT für die Praxis:

Vor Besprechung unserer Testverfahren erscheint es uns wichtig, den Praktiker auf einen Missstand hinzuweisen, der die meisten Testverfahren betrifft. Es sind die veralteten Normen. DEIMEL setzt in seinem Übersichtsartikel „Testverfahren zur Diagnostik der Lese-

Rechtschreibstörung – eine Übersicht“ (in Legasthenie: Zum aktuellen Stand der Ursachenforschung, der diagnostischen Methoden und der Förderkonzepte) sogar Zeitgrenzen, nach denen ein Testverfahren nicht mehr verwandt werden soll. Laut DEIMEL soll ein Testverfahren mit Normierungen älter als 10 Jahre, nur mit Vorsicht interpretiert werden, von Testverfahren älter als 15 Jahre ist abzuraten (Deimel 2002, s. auch die Veröffentlichung der Arbeitsgruppe von Schulte-Körne). Längerfristige Verschiebungen von Leistungsparametern in der Bevölkerung wurden sowohl im englischsprachigen wie auch im deutschsprachigen Raum belegt (Rodgers 1998, Strehlow & Haffner 2002, Zerahn-Hartung et al. 2002).

Nimmt man diese Normverschiebungen ernst, bleibt dem Diagnostiker nur noch eine geringe Auswahl an Lese-Rechtschreibtests übrig. Von den Testverfahren, die in unserer Studie untersucht wurden, besitzen Normierungen, die nicht älter als 10 Jahre sind, der DRT 2 u. 3, der SLRT, die WLLP und Knuspels Leseaufgaben. Unsere Empfehlung ist nicht, die Testverfahren mit den älteren Normierungen aus der psychologischen Diagnostik zu verbannen (was aufgrund der mangelnden Restauswahl auch nicht realistisch wäre), sondern dieses Wissen bei der Interpretation der Ergebnisse mit einzubeziehen.

Die Rechtschreibtest DRT 1 - 5, WRT 1+ - 3+ und WRT 4/5 – 6+ erbringen unseren Ergebnissen zufolge eine vergleichbare Beurteilung der Rechtschreibleistung. Vom formellen Aufbau (Lückentextdiktate) und der Testdurchführung weisen diese Testverfahren keine wesentlichen Unterschiede auf. REUTER-LIEHR schreibt, dass der WRT 2+ u. 3+ im Vergleich zum DRT 2 u. 3 eindeutig die besseren Testwörter beinhaltet. Da bei LRS-Kinder meist basale Probleme im Rechtschreiben, d.h. im lautgetreuen Schreiben und Lesen vorliegen, teilte REUTER-LIEHR diesen basalen Bereich in sogenannte 6 Phonemstufen, welche einen steigenden Schwierigkeitsgrad beinhalten, ein (2001, 2002). Sie stellt nun eine Gegenüberstellung der Phonemstufen von WRT 2+ u. 3+ und DRT 2 u. 3 vor. Durch diese Gegenüberstellung lässt sich sehr deutlich und anschaulich darstellen, dass der WRT 2+ u. 3+ eindeutig die besseren Testwörter beinhaltet. REUTER-LIEHR bezieht sich hierbei auf die Verteilung der Phonemstufen in der deutschen Orthographie (Reuter-Liehr, 2001) und Vorteile für die weitere Therapieplanung (Reuther-Lier, 2003).

Die Hamburger Schreibprobe stellt bei unseren Untersuchungen einen Sonderfall dar. Im Vergleich erhielten die Kinder mit der HSP in der 1., 3., und 4. Klasse, eine deutlich mildere Bewertung der Rechtschreibleistung. Entsprechendes findet sich auch in der aktuellen Literatur (Tacke et al. 2001, Deimel 2002, Warnke et al. 2004). Die HSP ist 2002 mit neuer Normierung erschienen. Die Normwerte sind gegenüber der Vorversion aber nur unwesentlich strenger geworden. (Schulte-Körne, Forschungsgruppe Lese-Rechtschreibstörungen).

Eine Einstufung von Kindern als lese-rechtschreibschwach mit dem SLRT-RT in der 1. Klasse ist problematisch. Bezieht man sich auf die orthographischen Fehler, die im Testhandbuch als das zentrale Auswertungskriterium genannt werden, so werden sehr viele Erstklässler, die vor allem nicht-lauttreue Fehler machen, unzureichend als rechtschreibschwach erkannt. Die orthographischen Fehler dürfen also nur zusammen mit den zwei weiteren Fehlerarten interpretiert werden. Ein Punktwert, wie er bei der operationalisierten Diagnostik zur Erstellung eines LRS-Gutachten notwendig ist, kann somit nicht erreicht werden. Unter Beachtung dieses Gesichtspunktes ist dieser Rechtschreibtest zur Diagnostik anwendbar. HASSELHORN und WARNKE nennen die drei Fehlerarten einen Vorteil des SLRT- Rechtschreibtest. Sie ermöglichen eine spezifische Diagnose von Schwierigkeiten beim lautorientierten Schreiben und dem orthographisch korrekten Schreiben (2000, 2004).

Auch die Lesetests für die 1. - 4. Klasse präsentierten sich weitgehend übereinstimmend in ihren Ergebnissen. Nur der Score Vorläuferfertigkeit von Knuspels Leseaufgaben korreliert etwas schlechter mit den anderen Lesetests. KNUSPEL-L misst neben dem Leseverständnis auch die Vorläuferfertigkeiten (Basisfertigkeiten der Lesefähigkeit: Hörverstehen, Rekodieren, Dekodieren), die mit den anderen Lesetests schlechter übereinstimmen. Es ist zu beachten, dass dieses Testverfahren inhaltlich und formell von den anderen Testverfahren abweicht. Während sich dieses Testverfahren an dem Modell der Leselernentwicklung orientiert, messen die klassischen Lesetests (SLRT, Zürcher) die Laut-Graphemerkennung und die Lesesyntese. Bei Knuspels Leseaufgaben werden sehr stark phonologische Fertigkeiten geprüft, bzw. passiver Wortschatz, das Hörverstehen und das Leseverständnis auf Wort- und Satzebene (Deimel, 2002). Sprachliche Fertigkeiten und auch die Aufmerksamkeit werden bei diesem Testverfahren beansprucht. Das Testverfahren ist leider durch seine Länge wenig ökologisch in seiner Durchführung und Auswertung.

Der Zürcher und der Salzburger Lesetest erfreuen sich weiterhin einer großen Beliebtheit im testpsychologischen Alltag. Als Vorteil wird gesehen, dass es sich um reine Funktionstests handelt, die die Laut-Graphemerkennung und die Lesesyntese sicher erfassen. Die Pseudowörter werden als sehr hilfreich erlebt, da sie zeigen, ob ein Kind wirklich lesen kann oder vorgelesene Wörter beispielsweise erraten hat (Reuter-Liehr, 2003). Schlechte Leser werden mit diesen Testverfahren sicher erfasst (Breuer 2005, Schmieder 2005, Bauer 2005). Auch in den aktuellen Leitlinien von WARNKE werden diese Testverfahren zur Überprüfung der Lesefähigkeit empfohlen (2004).

Zu beachten ist allerdings, dass sich bei den Drittklässlern Hinweise ergaben, dass der SLRT-LT im Vergleich zum Zürcher die Leseleistung milder bewertet.

Die WLLP ist ein sehr ökonomisches Testverfahren, das den Kindern, die an der Studie teilnahmen, sehr viel Spaß machte. Das anschauliche Bildmaterial und die Zeitkomponente wirken sehr motivierend. Das Bildmaterial erscheint dem Material, das in der Schule verwendet wird sehr ähnlich und den Kindern vertraut. Ein Nachteil dieses Testverfahrens ist, dass es keinerlei Aufschluss über den Leseprozess gibt. Es wird ausschließlich das Leseverständnis auf Wortebene überprüft. Man erhält aber keinerlei Aussage darüber, ob das Kind auch auf Textebene lesen kann. Beobachtungsmöglichkeiten während des Lesens (Buchstaben-Lautzuordnung, Lesegenauigkeit, Lesetempo) gehen verloren. Aus diesen Gründen wird der Test in der Praxis meist als weiterführendes Verfahren verwendet, beispielsweise um das Leseverständnis zusätzlich zu testen. Ein klarer Vorteil dieses Testverfahrens ist die ökonomische Durchführung.

### **6.3. Übereinstimmung zwischen Testergebnissen, Lehrerurteil und Elternurteil**

Nun sollte geklärt werden, ob die gängigen Testverfahren die Schulleistungen gleichermaßen einschätzen wie die Lehrer, die Noten vergeben und über das Weiterrücken in die nächste Jahrgangsstufe entscheiden. Auch sollte die Übereinstimmung der Testergebnisse mit dem Urteil der Eltern überprüft werden, wie auch der Zusammenhang zwischen Lehrer- und Elternurteil.

Unsere Auswertungen zeigen, dass Eltern und Lehrer in den Beurteilungen der Schulleistungen ihrer Kinder, bzw. Schüler mittelmäßig übereinstimmen. Am größten ist die Einigkeit bei der Mathematikleistung ( $r = 0,7$ ), gefolgt von Diktat- ( $r = 0,6$ ) und der Leseleistung ( $r = 0,6$ ). In der Regel unterscheidet sich das Elternurteil nicht sehr von der Lehrermeinung. Diese Aussage beruht auf Untersuchungen, die im Rahmen der „Iglu Studie“ erhoben wurden. (DIE ZEIT, 2003).

Mit den Testergebnissen allerdings weisen sowohl die Lehrer, wie auch die Eltern bezüglich der Beurteilung der Diktat- und Leseleistung unterschiedlich hohe Übereinstimmungen auf.

Die Übereinstimmung von Lehrern, wie auch Eltern mit den verschiedenen Rechtschreib-Testverfahren variierte beträchtlich (Lehrerurteil/Rechtschreibtests:  $r = 0,4 - 0,7$ , Elternurteil/Rechtschreibtests:  $r = 0,2 - 0,8$ ). Mit der WRT 1+ - 3+ (Lehrer  $r = 0,7$  u. Eltern  $r = 0,8$ ) zeigten sie die größte Übereinstimmung, gefolgt von der HSP (Lehrer  $r = 0,6$  u. Eltern  $r = 0,6$ ), dann dem SLRT (Lehrer  $r = 0,5$  u. Eltern  $r = 0,6$ ) und dem DRT (Lehrer  $r = 0,5$  u. Eltern  $r = 0,5$ ). Am schlechtesten stimmten beide mit dem WRT 4/5 - 6+ (Lehrer  $r = 0,4$  u. Eltern  $r = 0,2$ ) überein.

Bei den Lesetests stimmten die Lehrer besser mit den Testergebnissen überein als die Eltern (Lehrerurteil/Lesetests:  $r = 0,5 - 0,8$ , Elternurteil/Lesetests:  $r = 0,3 - 0,6$ ). Die Übereinstimmung zwischen Lehrerurteil und Tests ist für die meisten Lesetests besser ( $r = 0,7 - 0,8$ ) als für die Rechtschreibtests. Dies trifft nicht für den Zürcher und den Score Vorläuferfertigkeiten von Knuspels Leseaufgaben ( $r = 0,5 - 0,6$ ) zu.

GASTEIGER-KLICPERA et al. benennt in ihren Untersuchungen, bei denen sie 1.800 Kinder der 2. - 4. Klassenstufe über ein Jahr untersuchte, eine nur mäßige Übereinstimmung der Eltern- und Lehrereinschätzung mit den Testergebnissen. Außerdem zeigt sie in ihren Untersuchungen, dass das Urteil der Eltern vorrangig von den Lehrern und nicht vom testmäßig erfassten Leistungsstand geprägt wird. (Gasteiger-Klicpera et al., 2001).

Im Rahmen von Validierungen bei der Testkonzeption und vereinzelt bei weiterführenden Studien wurden ebenfalls die Übereinstimmungen zwischen Lehrerurteil und Testverfahren bestimmt.

Beim DRT 1 - 5 wurden für jede Klassenstufe ausführliche Validierungen bezogen auf die Schulleistungen vorgenommen. Die Übereinstimmungen von Lehrerurteil bzw. Schulnoten und Testergebnissen variieren beim DRT 1 - 3 über alle drei Klassenstufen hinweg deutlich (1. Klasse  $n = 54$   $r = 0,27 - 0,97$ , 2. Klasse  $n = 86$   $r = 0,20 - 0,95$ , 3. Klasse  $n = 156$   $r = 0,48 - 0,97$ ), wobei jede Klassenstufe gesondert berechnet wurde. Die Ergebnisse zeigen, dass die Fähigkeit, die Rechtschreibleistung eines Kindes einzuschätzen, bei den Lehrern unterschiedlich stark ausgebildet ist. Sie erstreckt sich von hoher Treffsicherheit ( $r = 0,97$ ) bis zur ungenügenden Einschätzung der Rechtschreibleistung ihrer Kinder ( $r = 0,20$ ). Auch bei unseren Ergebnissen sind die vorliegenden Unterschiede durch verschiedene Lehrkräfte möglicherweise mitbestimmt.

Eine aktuelle Untersuchung von 1999 an 248 Drittklässlern konnte eine Korrelation von  $r = 0,8$  zwischen DRT 3 und der Rechtschreibnote nachweisen (Tacke et al., 2001). Beim DRT 4 und 5 wurde die Validierung erweitert. Die Testergebnisse wurden zusätzlich mit der Höhe der Fehlerzahlen beim Schreiben eines Grundwortschatzes (500 - 1200 Wörter) korreliert ( $r = 0,7 - 0,9$ ). Die Übereinstimmung mit dem Lehrerurteil liegt für die 4. und 5. Klassen zwischen  $r = 0,7$  u.  $0,9$ . Im Rahmen einer Erweiterung des Konzeptes der Hamburger Schreibprobe wurden Schüler der 1. bis 4. Klasse mit dem DRT getestet und die Ergebnisse mit der Rechtschreibnote verglichen. Der Zusammenhang variierte zwischen  $r = 0,7$  u.  $0,9$  (May et al., 2001).

Ebenfalls im Rahmen der Validierung wurden beim WRT 1+ Testergebnisse mit dem Lehrerurteil, beim WRT 2+ und 3+ mit den Schulnoten und Diktatleistungen korreliert. Bei der gesamten Testreihe (WRT 1+ - 3+) liegen die Übereinstimmungen relativ konstant verteilt zwischen  $r = 0,6 - 0,7$ . Eine neuere Untersuchung bei 248 Schülern der 3. Klasse konnte zwischen den Testergebnissen des WRT 3+ und der Rechtschreibnote einen Zusammenhang von  $r = 0,8$  erbringen (Tacke et al., 2001).

Für die Westermann Rechtschreib-Testreihe sind uns keine Untersuchungen bekannt, die Testergebnisse auf einen Zusammenhang mit einem Lehrerurteil hin überprüft haben.

Die Testergebnisse der HSP von 292 Viertklässlern weisen mit der Rechtschreibleistung in Deutschaufsätzen einen Zusammenhang von  $r = 0,8$  auf (im Rahmen der Validierung). MAY zeigte außerdem bei seinen Anschlussuntersuchungen einen Zusammenhang zwischen der HSP 1



- 4 und der Rechtschreibnote von  $r = 0,7 - 0,9$  (2001). Dies entspricht der Untersuchung von TACKER, der eine Korrelation von  $r = 0,7$  zwischen HSP3 und der Rechtschreibnote aufführt (Tacke et al., 2001).

Die Validierung des SLRT Rechtschreibtest wurde an zweiten Klassen zu Beginn des 2. Halbjahres durchgeführt. Lehrer und Lehrerinnen wurden gebeten unter 1.165 Kindern diejenigen zu benennen, die spezifisch Schwierigkeiten im Lesen und Schreiben haben. Die 18 %, die von den Lehrkräften angegeben wurden, schnitten beim Rechtschreibtest signifikant schlechter ab als die unauffälligen Kinder. Besonders deutlich war der Unterschied bei den orthographischen Fehlern, wobei die nicht-lauttreuen Fehler und die Fehler der Groß-Kleinschreibung geringere Unterschiede aufwiesen.

Gutzelnig fand einen mittel-hohen Zusammenhang zwischen der Geschwindigkeit beim leisen Lesen eines Textes und den Leistungen beim Salzburger Lesetest (mit den Lesezeiten aller Subtests  $r = 0,6 - 0,8$ ). Des Weiteren setzte er bei einer Gruppe von 63 Schülern der 4. Schulstufe die Leseleistung beim Salzburger Lesetest in Bezug zur letzten Zeugnisnote in Deutsch. Zwischen der Lesezeit der einzelnen Subtests und der Deutschnote zeigten sich signifikante Korrelationen zwischen  $r = 0,4 - 0,5$  (Gutzelnig, 1995).

Bei der WLLP sind die Validierungen sehr gründlich für jede Klasse gesondert durchgeführt worden. Auffallend ist auch hier, dass die Korrelationen in Bezug auf das Lehrerurteil in den einzelnen Klassenstufen (pro Klassenstufe 67 - 74 Kinder) deutlichen Schwankungen unterliegen (1. Klasse  $r = 0,8$ ; 2. Klasse  $r = 0,6$ ; 3. Klasse  $r = 0,4$ ; 4. Klasse  $r = 0,6$ ). Wie bei den Untersuchungen zum DRT deutet dies erneut daraufhin, dass die Rechtschreibleistungen von verschiedenen Lehrkräften unterschiedlich eingeschätzt werden.

Auch beim Knuspel wurde eine sehr ausführliche Validierung für die Klassen 2 - 4 bei der Testkonzeption durchgeführt. Im Rahmen der Validierung wurden die Übereinstimmungen zwischen einem sogenannten Lehrerrating (Lehrerurteil bezogen auf Leseverstehen, Lesegenauigkeit und Lesetempo) und den Knuspel-Scores (Vorläuferfertigkeit und Lesefähigkeit) überprüft. Die Scores korrelieren etwa mittelmäßig ( $r = 0,6 - 0,7$ ) mit dem Lehrerurteil. Ab dem 3. Schuljahr wurden die Testergebnisse zusätzlich mit den Schulnoten auf ihre Übereinstimmung hin überprüft. In

Bezug auf das Lesen zeigen sich Korrelationen zwischen  $r = 0,6 - 0,7$  (bezüglich der Knuspel Scores).

Für den Zürcher Lesetest liegen keine Angaben zur Validierung vor. Außerdem finden sich keinerlei Angaben in der Literatur zu einer Überprüfung der Testergebnisse mit einem Lehrerurteil.

Es lässt sich zusammenfassen, dass für die meisten LRS-Tests Untersuchungen vorliegen, die ausreichend hohe Korrelationen mit dem Lehrerurteil aufweisen. Auffallend ist jedoch, dass bei einigen Tests sehr große Schwankungen in den Korrelationen von Lehrerurteil und Testergebnis zwischen den teilnehmenden Klassen und auch innerhalb einer Klassenstufe bestehen. Auch in unserer Studie zeigen sich bei den Lehrern, wie auch den Eltern unterschiedlich hohe Übereinstimmungen mit den verwendeten LRS-Tests.

Am ehesten deuten diese Ergebnisse daraufhin, dass die Fähigkeit, die Rechtschreibleistung eines Kindes einzuschätzen, bei den Lehrern unterschiedlich stark ausgebildet ist. Man muss auch davon ausgehen, dass sicherlich nicht alle Klassen nach der gleichen Methode unterrichtet werden. (Deimel, 2002).

Vergleicht man Gütekriterien für Testverfahren und Noten miteinander, überrascht einen dieses Ergebnis nicht. Testverfahren besitzen Gütekriterien mit klaren Definitionen und Bedeutungen (Objektivität, Reliabilität (Zuverlässigkeit, Messgenauigkeit) u. Validität). Schulnoten hingegen können dies nicht aufweisen. Schüler schreiben bundesweit Klassenarbeiten, die inhaltlich und formell unterschiedlich aufgebaut sind. Es gibt keine klaren Regelungen, dass Schüler vergleichbare Instruktionen zur Testbearbeitung erhalten und den Test unter vergleichbaren Bedingungen bearbeiten. Bei Klassenarbeiten liegen häufig keine exakten Auswertungsrichtlinien zur Beurteilung vor. Es gibt auch keine klaren Regeln, wie z.B. Prozentsätze korrekter Lösungen oder Fehler in Noten überführt werden sollen (Kornadt, 1978).

Lehrkräfte wurden in Schulen mit Hauptschulbildungsgang gefragt, welche ihrer Schülerinnen und Schüler besonders schwache Leser sind. Weniger als 15 % der Jugendlichen, die auf Grund ihrer Leistungsergebnisse in PISA als potenzielle Risikoschüler einzustufen sind, wurden von ihren Lehrern/-innen als solche erkannt. (Baumert, 2003). Auch wird die mangelnde diagnosti-

sche Kompetenz der Lehrer/-innen kritisiert. Mängel im Leseverständnis bleiben auch deshalb verborgen, weil die diagnostische Kompetenz der Lehrkräfte wenig ausgeprägt zu sein scheint. Nur bei jedem zehnten Schüler aus der Gruppe der 15-Jährigen, die in der Lesekompetenz nicht die erforderliche Minimalstufe II erreichen, vermuten die Lehrkräfte Lernschwächen dieses Ausmaßes. Effektive Förderung setzt voraus, dass die Lehrkräfte die Lernfortschritte beobachten und einschätzen können, so dass wirksame Hilfen eingeleitet werden können. Eine Einführung in die Diagnostik findet man bislang nur im Ausbildungsprogramm der sonderpädagogischen Lehrämter.

„Standardisierte Schulleistungstests werden in der Sekundarstufe I in weniger als einem Fünftel der Schulen zur Qualitätssicherung eingesetzt“ (PISA 2000, Faust 2003).

#### FAZIT für die Praxis:

Lehrer wie auch Eltern zeigen unterschiedlich hohe Übereinstimmungen mit den verwendeten LRS-Tests. In der Literatur finden sich für die meisten LRS-Tests Untersuchungen, die ausreichend hohe Korrelationen zwischen Testergebnis und Lehrerurteil nachweisen. Auffallend ist hierbei jedoch, dass bei einigen Tests die Höhe der Korrelation zwischen Lehrerurteil und Testergebnis zwischen den teilnehmenden Klassen und auch innerhalb einer Klassenstufe sehr schwankt. Am ehesten deuten diese Ergebnisse daraufhin, dass die Fähigkeit, die Rechtschreibleistung eines Kindes einzuschätzen, bei den Lehrern unterschiedlich stark ausgebildet ist. Mehrere Untersuchungen konnten nachweisen, dass sowohl die Eltern-, wie auch die Lehrerbeurteilung sehr kritisch einzuschätzende Parameter sind, die vielen Störeinflüssen unterliegen.

Es lassen sich aus den Zeugnissen und der Anamnese also nicht unbedingt die schwachen Schüler herausfiltern. Wenn man über die tatsächlichen Fähigkeiten eines Kindes eine Aussage treffen möchte, muss man die Fähigkeiten unbedingt auch mittels normierter Testverfahren einschätzen.

Im Idealfall sollten sich standardisierte Testverfahren und die Beurteilung durch den Lehrer ergänzen. Zum einen das Testverfahren, das die Objektivität bei der Beurteilung gut gewährleistet und zum anderen die Beurteilung durch den Lehrer, der zusätzlich den Schulverlauf und die individuelle Persönlichkeit mit einbezieht.

Ergebnisse von Testverfahren und das Lehrerurteil sollten von Ärzten, Psychologen, Lehrern und anderen Diagnostikern als gegenseitige Kontrollinstanzen betrachtet und ergänzt werden.

#### **6.4. Bewertung der diagnostischen Zuverlässigkeit von LRS-Tests**

Eine umschriebene Lese-Rechtschreibstörung (ICD-10: F81) ist durch eine Diskrepanz zwischen den allgemeinen kognitiven Fähigkeiten eines Kindes und dessen Lese-Rechtschreibleistungen gekennzeichnet.

Die Zuverlässigkeit einer psychometrisch gestellten Diagnose hängt zum einen von der Zuverlässigkeit der diagnostischen Kriterien und zum anderen von der Zuverlässigkeit der verwendeten psychometrischen Testverfahren ab.

Wir errechneten in unserer Studie die LRS- Diagnosen nach den Richtlinien für ein LRS-Gutachten nach § 35a SGB VIII (Empfehlung einer Arbeitsgruppe des Bayer. Sozialministeriums) u. den Leitlinien der Deutschen Gesellschaft für Kinder und Jugendpsychiatrie und – psychotherapie: „Eine Teilleistungsstörung ist anzunehmen, wenn das Testergebnis einem T-Wert von  $\leq 37$  entspricht und eine T-Wert Diskrepanz zum Gesamt-IQ  $\geq 12$  besteht“.

Wir haben uns für dieses Modell entschieden, da es ein in der testpsychologischen Praxis meist verwendetes und für ein LRS-Gutachten notwendiges Modell ist, das sich vorrangig an den Leistungsanforderungen im schulischen Alltag orientiert.

In unserer Studie nicht berücksichtigt wurde ein Regressionsmodell, wie es von SCHULTE-KÖRNE erarbeitet wurde. Der Vorteil des Regressionsmodells gegenüber dem einfachen IQ Diskrepanzmodell ist, dass Verzerrungen in den Extrembereichen der Verteilung der Rechtschreib- oder Leseleistung vermieden werden (Schulte-Körne 2001, Schulte-Körne Forschungsgruppe Lese-Rechtschreibstörungen).

In der Literatur wird die geforderte T-Wert-Differenz von 12 zwischen Intelligenztest und Lese-Rechtschreibtest kritisch gesehen, zumal diese entscheidend von dem durchgeführten Verfahren

zur Einschätzung der Intelligenzleistung und der Auswahl des Lese-Rechtschreibtests abhängt. Es wird ihr jedoch eine richtungweisende Bedeutung zugestanden (Reuter-Liehr, 2003). DEIMEL benennt die sich ständig verändernde Rechtschreibleistung der Kinder als Problem, zu deren Einschätzung der Bezug auf die Klassennorm das einzig objektive Kriterium darstellt. Das Kernsymptom einer Rechtschreibstörung ist eine im Vergleich zur Norm zu geringe Rechtschreibleistung. Dieses Maß ist kontinuierlich, d. h. es gibt erst einmal keine zwangsläufige Definition, ab wann genau eine Minderleistung als auffällig zu bezeichnen ist (Klicpera et al., 2003). Die Prävalenz einer LRS wird natürlich genau von einer solchen Definition bestimmt (Strehlow & Haffner, 2002). Auch werden fehlende Untersuchungen zur Reliabilität von Diskrepanzmodellen kritisiert (Deimel 2002, Schulte-Körne Forschungsgruppe Lese-Rechtschreibstörungen).

Ebenfalls sehr anhaltend wird international darüber diskutiert, ob für die Diagnostik die Intelligenz berücksichtigt werden sollte (Aaron 1997, Meyer 2000, Deimel 2002, Klicpera & Gasteiger Klicpera 2001). Viel wichtiger, und bei vielen Diskussionen um den IQ oft wenig berücksichtigt, erscheint uns die praktische Bedeutung, die eine Unterteilung in spezifische und allgemeine Rechtschreibschwäche hat. Für eine gezielte Therapieplanung (bzw. auch Schulplanung bei einer leistungsmäßigen Überforderung) ist die Einschätzung der intellektuellen Fähigkeiten eines Kindes ein notwendiger Bestandteil der kinder- und jugendpsychiatrischen Diagnostik. (Reuter-Liehr, 2003).

Es lässt sich zusammenfassen, dass die Zuverlässigkeit der diagnostischen Kriterien einer Lese-Rechtschreibstörung in der Literatur sehr kritisch diskutiert wird. Neben der Zuverlässigkeit der diagnostischen Kriterien hängt die Zuverlässigkeit einer psychometrisch gestellten Diagnose auch von der Zuverlässigkeit der verwendeten psychometrischen Testverfahren ab.

Als Antwort auf die Frage nach der diagnostischen Zuverlässigkeit von LRS-Tests kann nach unseren Untersuchungen zusammengefasst werden, dass bei Verwendung unterschiedlicher Testverfahren die Ergebnisse nicht immer übereinstimmend sind. In der Literatur wird die Abhängigkeit einer LRS-Diagnose vom eingesetzten Testverfahren bestätigt (Reuter-Liehr, 2003). Die Rechtschreib- wie auch Leseleistung ist als Maß ein kontinuierlicher Parameter. Setzt man nun eine „Cut-off“ Grenze, erhält man nur auffällig vs. unauffällig, aber keine differenzierte Einschätzung der Schulleistung. Dies trifft insbesondere für die Kinder zu, die mit ihrer Leistungsfä-

higkeit im Bereich der Grenze liegen. Kinder, die knapp oberhalb bzw. unterhalb der Grenze liegen, werden schon bei geringen Abweichungen der Ergebnisse in den verschiedenen Tests, unterschiedlichen Kategorien zugeordnet. DEIMEL rät auch in diesem Zusammenhang dazu, statistische Normwerte vor allem bei grenzwertigen Ergebnissen nicht sklavisch zu befolgen (2002).

Zwei Rechtschreibtests weichen deutlich von den anderen Testverfahren ab. Es zeigt sich, dass die Ergebnisse des Salzburger Rechtschreibtests in der 1. Klasse deutlich häufiger zu LRS-Diagnosen führen als die Bewertungen der anderen Rechtschreibtests. Mit dem SLRT-RT konnte bei 88 %, mit dem DRT bei 38 %, mit der WRT 1+ bei 25 % und mit der HSP bei 0 % der Studienteilnehmer der 1. Klasse eine LRS-Diagnose gestellt werden. Nur beim SLRT Rechtschreibtest für die 1. Klasse wurden die sehr groben Prozentrangbänder (<37, =37, >37) für die nicht-lauttreuen Fehler als Auswertungskriterium verwandt. Eine Einstufung von Kindern als rechtschreibschwach mit dem SLRT-RT in der 1. Klasse ist problematisch. Bezieht man sich auf die orthographischen Fehler, die im Testhandbuch als das zentrale Auswertungskriterium genannt werden, so werden sehr viele Erstklässler, die vor allem nicht-lauttreue Fehler machen, unzureichend als rechtschreibschwach erkannt (s. 6.2.). Wenn anstelle der orthographischen Fehler die nicht-lauttreuen Fehler als Beurteilungskriterium verwandt werden, werden zu viele Kinder als auffällig eingestuft. Die „cut-off“ Grenze für die nicht-lauttreuen Fehler liegt dementsprechend zu hoch. Es zeigt sich hiermit sehr deutlich, dass mit dem Salzburger Rechtschreibtest durch seine sehr groben Normen eine LRS-Diagnose nicht zuverlässig gestellt werden kann. Auch DEIMEL spricht sich aus diesen Gründen gegen eine Empfehlung dieses Testverfahrens aus (2002).

Wieder einmal stellt auch die Hamburger Schreibprobe einen Sonderfall dar. Sowohl in der 1. wie auch in der 3. Klasse werden durch die Ergebnisse der HSP deutlich weniger LRS-Diagnosen gestellt. In der 1. Klasse erhält durch die HSP keines der Kinder eine LRS-Diagnose (vgl. LRS Diagnose durch DRT: 38 % der Kinder, WRT 1+: 25 % der Kinder) und in der 3. Klasse sind es 10 % aller Kinder, die durch die HSP eine LRS Diagnose erhalten (vgl. LRS Diagnose durch DRT: 40 % der Kinder, WRT 1+: 50 % der Kinder, SLRT-RT: 40 % der Kinder). In der 4. Klasse zeigte sich ebenfalls ein Unterschied, jedoch weniger ausgeprägt (LRS Diagnose durch: HSP: 5 % der Kinder, DRT: 15 % der Kinder, WRT 4/5: 15 % der Kinder, SLRT-RT: 23 % der Kinder).

Dies entspricht unseren Ergebnissen, dass die HSP vor allem in der 1. und 3. Klasse die Rechtschreibleistung zu milde einschätzt, was durch andere Untersuchungen (Tacke et al., 2001) wie auch die klinische Erfahrung bestätigt wird. Eine Lese-Rechtschreibstörung kann somit mit der Hamburger Schreibprobe nicht zuverlässig diagnostiziert werden. 2002 erschien eine Neuauflage mit neuer Normierung. Die Normwerte sind gegenüber der Vorversion aber nur unwesentlich strenger geworden, so dass keine allzu großen Veränderungen zu erwarten sind (Schulte-Körne, Forschungsgruppe Lese-Rechtschreibstörungen).

Abschließend interessierten uns noch die Lese-Rechtschreibtests und ihre Trefferquoten bei der Ermittlung einer Lese-Rechtschreibschwäche. Mit der Berechnung der Sensitivität und der Spezifität sollte sich bei den einzelnen Testverfahren zeigen, wie zuverlässig in der Schule auffällig gewordene Kinder (mangelhafte Schulleistung: „Einschätzung“ der Lehrer  $\geq 4,5$ ) schlechte Testergebnisse (mangelhaftes Testergebnis: T-Wert  $\leq 37$ ), bzw. in der Schule unauffällige Kinder gute Testergebnisse erhalten.

Bei den Rechtschreibtests liegt die Sensitivität zwischen 39 - 82 %, was bedeutet, dass die Testverfahren Schüler mit Rechtschreibproblemen in der Schule unterschiedlich gut erkennen. Orientiert man sich an allgemeinen Konventionen in der psychologischen Diagnostik, die besagen, dass sowohl die Sensitivität als auch die Spezifität einen Wert von 80% nicht unterschreiten sollte, ansonsten ist davon auszugehen, dass der Test nicht hinreichend genau das misst, was er messen möchte (Petermann & Macha, 2005), so erreicht diesen Kennwert nur die Weingartener Rechtschreibtestreihe. Die Spezifitäten fallen etwas höher aus, aber auch sie zeigen eine deutliche Varianz bei den einzelnen Rechtschreibtests (59 - 90 %). Nur die HSP und der Westermann-Rechtschreibtests erkennen über 80 % der Kinder wieder, die von den Lehrern als unauffällig eingestuft wurden.

Zur genaueren Differenzierung der Aussagekraft der verschiedenen Testverfahren bezüglich tatsächlicher Schulprobleme berechneten wir die Gesamttrefferquote (als Maß für die Qualität der Vorhersage) und den RATZ-Index (**R**elativer **A**nstieg der **T**refferquote gegenüber der **Z**ufallstrefferquote). Natürlich gäbe es auch auf Grund des Zufalls eine gewisse Trefferquote. Deshalb wird eine Zufallstrefferquote bestimmt, die ausdrückt, wie hoch die Treffer-Wahrscheinlichkeit bei einer konkreten Untersuchungssituation wäre; sie kann als Untergrenze für die Güte der Ge-

samttrefferquote gelten. Die Gesamttrefferquote, die zwischen den Werten der Zufallstrefferquote und der Maximaltrefferquote liegt, gibt die Anzahl der als valide positiv und negativ klassifizierten Fälle an der Gesamtstichprobe an (Marx, 2000).

Die Gesamttrefferquote bei den Rechtschreibtests liegt zwischen 62 - 77 %. Die Berechnung des RATZ-Index erbrachte erneut bei den einzelnen Testverfahren große Unterschiede. Während der Westermann Rechtschreibtest, die Hamburger Schreibprobe und der Salzburger Rechtschreibtest Werte zwischen 17 - 27 % zeigen, erreichen der DRT und die Weingartener Rechtschreib Testreihe 51 % und 60 %.

Es gibt bisher noch kein interferenzstatistisches Verfahren, die Unterschiedlichkeit von RAZT-Indizes abzusichern. Marx (in Druck) schlägt deshalb in Analogie zur Interpretation der Effektstärke vor, den Schwankungsbereich des Indexes zu dritteln und wie folgt zu interpretieren: > 66 % sehr gute Vorhersage, 66 - 34 % gute Vorhersage und < 34 % nicht zufriedenstellende Vorhersage (Marx, 2000).

Somit erbringen nur der DRT und die WRT 1+ - 3+ gute Aussagen für schulische Rechtschreibschwierigkeiten, was insgesamt eher ein enttäuschendes Ergebnis ist.

Die Lesetests zeigen etwas bessere Ergebnisse. Die Sensitivität bewegt sich meist über 80 %, die Spezifität zwischen 60 - 81 %. Auch die Gesamttrefferquote, und vor allem der RATZ-Index lagen höher als bei den Rechtschreibtests. Die Gesamttrefferquote lag zwischen 65 - 83 %, der RATZ-Index meist über 65 % und sogar zweimal bei 100 % (WLLP und Knuspel Score Lesefähigkeit). Als einziger Lesetest lag KNUSPEL-L mit seinem Score Vorläuferfertigkeit bei 25 %, zeigte also eine nicht zufriedenstellende Aussage. Interessant ist, dass genau dieser Score von Knuspels Leseaufgaben auch auffällig schlechter mit den anderen Lesetests korreliert (vgl. dazu 6.2.).

Also erreichen die meisten Lesetests sehr gute bis gute Aussagen in Bezug auf Schulschwierigkeiten im Lesen. Diese Ergebnisse bedeuten, dass die Lesetests deutlich zuverlässiger sind als die Rechtschreibtests. Hierfür gibt es keine schlüssige Erklärung. Es scheint, dass der RATZ-Index stark von der Sensitivität abhängt, die bei den Rechtschreibtests in unseren Untersuchungen deut-



lich niedriger liegt als bei den Lesetests. Dies liegt möglicherweise daran, dass die Diktatleistung von den Lehrern strenger als die Leseleistung beurteilt wird. Von dieser Annahme geleitet, setzen wir eine neue Definition für den Goldstandard (mangelhafte Schulleistung (Diktat): „Einschätzung“ der Lehrer  $\geq 5$ ). Nun lagen die Sensitivitäten etwas höher, zeigten aber auch wie erwartet, weiterhin große Unterschiede zwischen den einzelnen Tests (40 - 89 %). Über 80 % erreichte auch hier nur die Weingartener Rechtschreib-Testreihe. Die Gesamttrefferquote veränderte sich kaum (64 - 71 %). Das Ergebnis des RAZ-Index zeigte geringe Unterschiede. Immerhin erreichen nun die meisten Rechtschreibtests (mit Ausnahme der HSP und dem WRT 4/5 - 6+) eine gute und die WRT 1+ - 3+ eine sehr gute Aussage in Bezug auf die Rechtschreibleistungen in der Schule.

Zusammenfassend zeigt sich jedoch, dass die Lesetests bessere Aussagen zu schulischen Lese-problemen erreichen als die Rechtschreibtests zu schulischen Problemen im Rechtschreiben.

#### FAZIT für die Praxis:

Bei der diagnostischen Zuordnung der einzelnen Kinder sind bei Verwendung unterschiedlicher Testverfahren die Ergebnisse nicht immer übereinstimmend. Ein Grund hierfür könnte die Wahl einer „cut-off“ Grenze sein, wie sie bei der operationalisierten Diagnostik notwendig ist. Dies trifft insbesondere für die Kinder zu, die mit ihrer Leistungsfähigkeit im Grenzbereich liegen. Kinder, die knapp oberhalb bzw. unterhalb der Grenze liegen, werden schon bei geringen Abweichungen der Ergebnisse in den verschiedenen Tests, zu unterschiedlichen Kategorien zugeordnet.

Eine Einstufung von Kindern als lese-rechtschreibschwach mit dem SLRT-RT in der 1. Klasse ist problematisch. Aufgrund eines fehlenden Gesamtwertes für alle drei Fehlerarten (orthographische Fehler, nicht-lauttreue Fehler und Fehler der Groß-Kleinschreibung) und aufgrund von Fehldiagnosen bei alleiniger Verwendung einzelner Fehlerarten, ist dieses Testverfahren zur Diagnostizierung einer LRS bei Erstklässlern nicht geeignet.

Auch bildet die HSP wieder einen Sonderfall. Sowohl in der 1. wie auch in der 3. Klasse werden durch die HSP deutlich weniger LRS-Diagnosen gestellt. Dies entspricht unseren Ergebnissen aus dem Testvergleich. Die HSP schätzt vor allem in der 1. und 3. Klasse die Rechtschreibleistung zu milde ein, was durch andere Untersuchungen (Tacke et al., 2001) wie auch die klinische

Erfahrung bestätigt wird. Eine Lese-Rechtschreibstörung kann somit mit der Hamburger Schreibprobe nicht zuverlässig diagnostiziert werden. 2002 erschien eine Neuauflage mit neuer Normierung. Die Normwerte sind gegenüber der Vorversion aber nur unwesentlich strenger geworden, so dass die Kritik an diesem Verfahren weiterhin fortbesteht (Schulte-Körne, Forschungsgruppe Lese-Rechtschreibstörungen).

Die Aussagekraft der LRS-Tests ist sehr unterschiedlich. Es zeigte sich, dass die Lesetests bessere Aussagen zu schulischen Leseproblemen erlangen als die Rechtschreibtests zu schulischen Problemen im Rechtschreiben. Lesetests sollten daher verstärkt eingesetzt werden.

### **6.5. Schlussfolgerung für die Praxis**

Die üblichen Testkriterien (Validität und Reliabilität) sind bei Lese- und Rechtschreibtests normalerweise unproblematisch. Die Validität ist in der Regel gegeben (Augenscheinvalidität), und Reliabilitätskoeffizienten liegen normalerweise im üblichen akzeptablen Bereich. Ein Aspekt jedoch, worin sich Testverfahren unterscheiden, ist die Normierung. Besonders von erfahrenen Diagnostikern wird meist das Alter der Normen als sehr wichtiges Qualitätsmerkmal hervorgehoben.

Man findet längerfristige Verschiebungen von Leistungsparametern in der Bevölkerung. Rechtschreib- und Lesetests werden mit zunehmendem Alter der Normen zu schwierig, während Intelligenztests im Laufe der Zeit zu leicht werden. Für die Intelligenz (bekannt als „Flynn-Effekt“) und, in umgekehrter Richtung, für Schulleistungstests ist dies im englischsprachigen Raum gut belegt (Rodgers, 1998). Für eine deutschsprachige Stichprobe wurde dies ebenfalls nachgewiesen (Strehlow & Haffner 2002, Zerahn-Hartung et al. 2002). Eine Hauptforderung an Schulleistungstests ist demnach, dass deren Normierung aktuell ist (Schulte-Körne, Forschungsgruppe Lese-Rechtschreibstörungen).

Normierungen sollen maximal 10 Jahre alt sein, damit die Testverfahren uneingeschränkt empfohlen werden können (Deimel, 2002). Von den Testverfahren, die in unserer Studie untersucht wurden, besitzen Normierungen, die nicht älter als 10 Jahre sind, nur der DRT 2 u. 3, der SLRT,

die WLLP und Knuspels Leseaufgaben. Da gängige Testverfahren mit alten Normierungen, wie beispielsweise der Zürcher Lesetest kaum vom Markt zu verbannen sind, raten wir dringend bei Verwendung solcher Testverfahren diese Normenverschiebung zu beachten und das Ergebnis stets zu hinterfragen. Hierfür ist viel Praxiserfahrung notwendig!

Die Rechtschreibtest DRT 1 - 5, WRT 1+ - 3+ und WRT 4/5 - 6+ schätzen die Schulleistungen vergleichbar ein, so dass sie gleichermaßen zur Diagnostik einer Rechtschreibstörung verwendet werden können. Vom formellen Aufbau (Lückentextdiktate) und der Testdurchführung weisen diese Testverfahren keine wesentlichen Unterschiede auf. Nach REUTER-LIEHR beinhaltet der WRT 2+ u. 3+ im Vergleich zum DRT 2 u. 3 die besseren Testwörter. Sie nimmt dabei auf die Verteilung der Phonemstufen in der deutschen Orthographie Bezug (Reuter-Liehr 2001, Reuther-Lier 2003).

Die Hamburger Schreibprobe schätzt die Rechtschreibleistung zu milde ein, so dass rechtschreibschwache Kinder nicht ausreichend erkannt werden. Dies gilt für alle Klassen mit Ausnahme der zweiten und der fünften Schulstufe. Dies entspricht sowohl dem Eindruck aus dem Praxisalltag, als auch der aktuellen Literatur (Tacke et al. 2001, Deimel 2002, Warnke et al. 2004). Die HSP ist 2002 mit neuer Normierung erschienen. Die Normwerte sind gegenüber der Vorversion aber nur unwesentlich strenger geworden. (Schulte-Körne, Forschungsgruppe Lese-Rechtschreibstörungen).

Der Salzburger Rechtschreibtest unterscheidet drei Fehlerarten (orthographische, nicht-lauttreue und Fehler der Groß-Kleinschreibung). Die orthographischen Fehler, die im Testhandbuch als das zentrale Auswertungskriterium genannt werden, dürfen nur zusammen mit den zwei weiteren Fehlerarten interpretiert werden, da es sonst vor allem bei rechtschreibschwachen Kindern, die noch viele nicht-lauttreue Fehler machen, zu Fehleinschätzungen kommt. Ein Punktwert, wie er bei der operationalisierten Diagnostik zur Erstellung eines LRS-Gutachten notwendig ist, kann somit nicht erreicht werden.

Die Ergebnisse von Lesetests sind größtenteils untereinander vergleichbar. Bei Knuspels Leseaufgaben, die sich am Modell der Leselernentwicklung orientieren, werden sehr stark phonologische Fertigkeiten geprüft, bzw. passiver Wortschatz, das Hörverstehen und das Leseverständnis

auf Wort- und Satzebene (Deimel, 2002). Somit ist es kein klassischer Lesetest wie der SLRT-LT oder der Zürcher. Vor allem der Score Vorläuferfertigkeit, der die Überprüfung der Vorläuferfertigkeiten beinhaltet, stimmt mit den anderen Lesetests schlechter überein. Die Testaufgaben sind sehr anspruchsvoll und können dadurch auf sehr leseschwache Kinder abschreckend wirken. Ein Nachteil stellt die nicht geringe Bearbeitungszeit von bis zu einer Stunde dar.

Beim Zürcher und dem Salzburger Lesetest handelt es sich um reine Funktionstests, die die Laut-Graphemerkennung und die Lesesyntese sicher erfassen. Als hilfreich gelten die Pseudowörter, da durch sie schwache Leser sicher erkannt werden können. Das Leseverständnis kann durch entsprechende Zusatzverfahren oder durch entsprechende Fragen nach Durchlesen der Textabschnitte sicher erfragt werden (Reuter-Liehr, 2003). Abweichende Ergebnisse zeigen die Testverfahren bei den dritten Klassen. Die Ergebnisse des Zürcher im Vergleich zum SLRT Lesetest fallen etwas niedriger aus, die des SLRT im Vergleich etwas höher. Am ehesten muss man davon ausgehen, dass der SLRT in der 3. Klasse zu milde misst.

Die WLLP ist ein sehr ökonomisches Testverfahren, das den Kindern, die an der Studie teilnahmen, sehr viel Spaß machte. Das Bildmaterial erscheint dem Material, das in der Schule verwendet wird sehr ähnlich und den Kindern vertraut. Ein Nachteil dieses Testverfahrens ist, dass es keinerlei Aufschluss über den Leseprozess gibt. Es wird ausschließlich das Leseverständnis auf Wortebene überprüft. Man erhält aber keinerlei Aussage darüber, ob das Kind auch auf Textebene lesen kann. Beobachtungsmöglichkeiten während des Lesens gehen verloren. Laut KÜSPERT (2000) vermag die WLLP sehr gut zwischen Lesern unterschiedlicher Kompetenzgrade zu differenzieren. Auch WARNKE empfiehlt die WLLP in seinem aktuellen Leitfaden zur Diagnostik einer LRS (2004).

Um zu überprüfen, welche der Tests die tatsächlichen Lese- und Rechtschreibleistungen der Kinder am besten widerspiegeln, wurde eine Einschätzung der Schulleistungen durch die Lehrer und durch die Eltern erfragt. Lehrer wie auch Eltern zeigen unterschiedlich hohe Übereinstimmungen mit den verwendeten LRS-Tests. In der Literatur finden sich für die meisten LRS-Tests Untersuchungen, die ausreichend hohe Korrelationen zwischen Testergebnis und Lehrerurteil nachweisen. Auffallend ist hierbei jedoch, dass bei einigen Tests die Höhe der Korrelation zwischen Lehrerurteil und Testergebnis zwischen den teilnehmenden Klassen und auch innerhalb einer Klas-

senstufe sehr schwankt. Am ehesten deuten diese Ergebnisse daraufhin, dass die Fähigkeit, die Rechtschreib- bzw. Leseleistung eines Kindes einzuschätzen, bei den Lehrern unterschiedlich stark ausgebildet ist. Mehrere Untersuchungen konnten nachweisen, dass sowohl die Eltern-, wie auch die Lehrerberurteilung sehr kritisch einzuschätzende Parameter sind, die vielen Störeinflüssen unterliegen. Auf der anderen Seite haben nur die Lehrkräfte die gesamte Persönlichkeit des Lernenden im Blick und beurteilen die Ergebnisse auf dem Hintergrund länger andauernder Lernprozesse und unter Einbeziehung der institutionellen und personellen Rahmenbedingungen.

Wenn man also über die tatsächlichen Fähigkeiten eines Kindes eine Aussage treffen möchte, sollten die Fähigkeiten auch mittels normierter Testverfahren geprüft werden. Im Idealfall sollten sich standardisierte Testverfahren und die Beurteilung durch den Lehrer ergänzen. Zum einen das Testverfahren, das die Objektivität bei der Beurteilung gut gewährleistet und zum anderen die Beurteilung durch den Lehrer, der zusätzlich den Schulverlauf und die individuelle Persönlichkeit mit einbezieht (May, 2002<sup>2</sup>).

Zuletzt setzten wir uns noch mit der diagnostischen Zuverlässigkeit von LRS-Tests auseinander. Wir errechneten in unserer Studie die LRS- Diagnosen nach den Richtlinien für ein LRS-Gutachten nach § 35a SGB VIII (Empfehlung einer Arbeitsgruppe des Bayer. Sozialministeriums) u. den Leitlinien der Deutschen Gesellschaft für Kinder und Jugendpsychiatrie und – psychotherapie (Testwert T-Wert  $\leq 37$  u. Diskrepanz zum Gesamt IQ  $\geq 12$  T-Wertpunkte). Dieses Modell wird in der testpsychologischen Praxis häufig verwendet und ist für ein LRS-Gutachten notwendig. Die Zuverlässigkeit der diagnostischen Kriterien einer Lese-Rechtschreibstörung wird in der Literatur sehr kritisch diskutiert. Es gibt verschiedene Modelle, wobei sich bisher keines klar durchsetzen konnte (Aaron 1997, Meyer 2000, Schulte-Körne 2001, Deimel 2002, Klicpera & Gasteiger Klicpera 2001).

Als Antwort auf die Frage nach der diagnostischen Zuverlässigkeit von LRS-Tests kann zusammengefasst werden, dass bei Verwendung unterschiedlicher Testverfahren die Ergebnisse nicht immer übereinstimmend sind. Die Rechtschreib- wie auch Leseleistung ist als Maß ein kontinuierlicher Parameter. Setzt man nun eine „Cut-off“ Grenze, erhält man nur auffällig vs. unauffällig, aber keine differenzierte Einschätzung der Rechtschreibleistung. Dies trifft insbesondere für die Kinder zu, die mit ihrer Leistungsfähigkeit im Bereich der Grenze liegen. Kinder, die knapp

oberhalb bzw. unterhalb der Grenze liegen, werden schon bei geringen Abweichungen der Ergebnisse in den verschiedenen Tests, zu unterschiedlichen Kategorien zugeordnet.

Zwei Testverfahren zeigten sich eindeutig unzuverlässig bei der Erhebung der LRS-Diagnosen. Wie bereits erwähnt, bietet der Salzburger Rechtschreibtest keinen Gesamtwert für alle Fehlerarten an, so dass nur eine getrennte Auswertung der einzelnen Fehlerarten erfolgen kann. Bezieht man sich auf die orthographischen Fehler, die im Testhandbuch als das zentrale Auswertungskriterium genannt werden, so werden sehr viele Erstklässler, die vor allem nicht-lauttreue Fehler machen, unzureichend als rechtschreibschwach erkannt. Wenn anstelle der orthographischen Fehler die nicht-lauttreuen Fehler als Beurteilungskriterium verwendet werden, werden zu viele Kinder als auffällig eingestuft. Die „cut-off“ Grenze für die nicht-lauttreuen Fehler liegt dementsprechend zu hoch.

Wie zu erwarten, erhalten durch die Hamburger Schreibprobe auffallend weniger Kinder eine Lese-Rechtschreibschwäche. Dies entspricht unseren Ergebnissen aus dem Testvergleich, wo die HSP im Vergleich zu den anderen Testverfahren die Rechtschreibleistung milder einschätzt. Eine Lese-Rechtschreibstörung kann somit mit der Hamburger Schreibprobe nicht zuverlässig diagnostiziert werden. Wie bereits erwähnt, erschien 2002 eine Neuauflage mit neuer Normierung. Die Normwerte sind gegenüber der Vorversion aber nur unwesentlich strenger geworden (Schulte-Körne, Forschungsgruppe Lese-Rechtschreibstörungen).

Zum Abschluss interessierte uns noch, wie gut die verschiedenen LRS-Test in der Schule im Lesen bzw. Rechtschreiben auffällig gewordene Kinder erkennen. Durch die Lesetests wurden die von den Lehrern als schwach eingestuften Kinder insgesamt etwas besser erkannt als von den Rechtschreibtests die rechtschreibschwachen Schüler. Lesetests sollten daher in der testdiagnostischen Praxis häufig eingesetzt werden.

## 7. ZUSAMMENFASSUNG

### Theoretischer Hintergrund

Kinder, die durch eine Rechtschreib- und Lesestörung in ihrer schulischen Leistung beeinträchtigt sind, nehmen im Patientengut der Kinder- und Jugendpsychiatrie einen großen Anteil ein. Eine frühzeitige Diagnosestellung stellt die Basis für gezielte Förderprogramme dar. Zur Diagnostizierung einer Lese- und Rechtschreibstörung werden standardisierte Testverfahren verwendet.

Die Testverfahren unterscheiden sich meist hinsichtlich ihrer Grundkonzeption und vor allem hinsichtlich ihrer Normierung. Bei der Normierung sollten zwei Aspekte beachtet werden. Es handelt sich hierbei um die Repräsentativität und das Alter der Normierung. Die Repräsentativität betrifft die Zusammensetzung der Normierungsstichprobe (Stadt vs. Land, Gewichtung der Bundesländer etc.), sowie die Frage, inwieweit überhaupt, z. B. für das Ende der zweiten Klasse allgemein gültige Rechtschreibnormen in Deutschland existieren. Eine besondere Schwierigkeit besteht darin, dass es keinen Wortschatz gibt, der bundesweit alters-, klassen- und schultypspezifisch als allgemein verbindlich gilt. Somit kann man davon ausgehen, dass zwischen den Bundesländern Unterschiede in den Grundwortschatzlisten bestehen. Neben Unterschieden zwischen den Bundesländern herrschen auch Schwankungen zwischen einzelnen Schulen und auch zwischen Klassen innerhalb einer Schule. Als zweiter entscheidender Aspekt soll nun auf das Alter der Normierung eingegangen werden. Während einige Testverfahren relativ aktuelle Normierungen aufweisen, bieten andere Testverfahren Normierungen, die teilweise über 30 Jahre alt sind. Es konnte durch Studien belegt werden, dass Rechtschreib- und Lesetests mit zunehmendem Alter der Normen zu schwierig werden.

Es herrscht eine große Verunsicherung bezüglich der Güte der verschiedenen Rechtschreib- und Lesetests, da nicht davon ausgegangen werden kann, dass mit unterschiedlichen Testverfahren gleichlautende Ergebnisse zu erhalten sind. Aus diesem Grund entschieden wir uns im Rahmen einer Studie zu einem Vergleich der gängigsten Testverfahren.

Es sollte untersucht werden, inwieweit die in der Testpraxis häufig angewandten Testverfahren die Rechtschreib- bzw. Leseleistungen eines Kindes übereinstimmend erfassen.

Auf dieser Grundlage sollte dann eine kritische Betrachtung und Zusammenfassung von Vor- und Nachteilen, auch der klassischen (Validität, Reliabilität, Objektivität) sowie der ökonomischen Gütekriterien (zeitliche Dauer, aufwändige Durchführung) vollzogen werden.

#### Fragestellung:

In dieser Studie sollen folgende Fragestellungen überprüft werden:

- 1) Sind Ergebnisse von LRS- Tests untereinander vergleichbar?
  - Wie hoch korrelieren die Ergebnisse miteinander?
  - Beurteilen unterschiedliche LRS- Tests die Lese- bzw. Rechtschreibfähigkeit von Kindern gleich streng?
- 2) Wie groß ist die Übereinstimmung zwischen Testergebnissen, Lehrer- und Elternurteil?
- 3) Wie hoch ist die diagnostische Zuverlässigkeit von LRS- Tests zu bewerten?
  - Wie stark hängt die Diagnose „Lese-Rechtschreibstörung“ vom eingesetzten Testverfahren ab?
  - Wie hoch ist die Trefferquote einzelner LRS-Tests einzuschätzen?

#### Methode:

Alle Kinder, die sich zum Studienzeitpunkt in der Klinik Hochried in Behandlung befanden, wurden in die Untersuchung miteinbezogen. Die Klinik Hochried ist eine Kurklinik mit den Schwerpunktindikationen Lese-Rechtschreibstörung (F81), Hyperkinetische Störung (F90) und andere Erkrankungen (vor allem Adipositas, Asthma bronchiale, Atopisches Ekzem, Skoliose). Untersucht wurden 48 Schüler/innen der 1. - 6. Klasse, die unterschiedliche Schultypen besuchten (Grund- und Sprachheilschulen, Haupt- und Realschulen, einer Montessorischule und Förderschulen).

Als Ausschlusskriterien wurden Sinnesbeeinträchtigungen sowie organische und zentralnervöse Erkrankungen angesehen, die Schreiben und/oder Lesen erheblich beeinträchtigen. Beginn oder Änderungen einer medikamentösen Therapie während der Diagnostikphase führten zum Ausschluss aus der Studie.

Als Rechtschreibtest wählten wir den Diagnostischen Rechtschreibtest für die 1. - 5. Klasse (DRT 1 - 5), die Weingartener Grundwortschatz Rechtschreib-Testreihe für die 1. - 3. Klasse (WRT 1+ - 3+), den Westermann Rechtschreibtest für die 4. - 5. Klasse (WRT 4/5 - 6+), den



Salzburger Rechtschreibtest für die 1. - 4. Klasse (SLRT-RT) und die Hamburger Schreibprobe für die 1. - 5. Klasse (HSP 1+ - 5).

Bei den Lesetests wurden der Zürcher Lesetest für die 2. - 5. Klasse (Zürcher Lesetest), der Salzburger Lesetest für die 1. - 4. Klasse (SLRT-LT), die Würzburger Leise Leseprobe für die 1. - 4. Klasse (WLLP) und Knuspels Leseaufgaben für die 1. - 4. Klasse (KNUSPEL-L) untersucht.

Jedes Kind der 1 - 4. Grundschulklasse oder der 3. - 6. Förderklasse nahm an 4 Rechtschreibtests (DRT, WRT 1+ - 3+, WRT 4/5, SLRT, HSP), an 3 - 4 Lesetests (SLRT, Zürcher, WLLP, KNUSPEL-L) und an einem altersentsprechenden nonverbalen Intelligenztest (CFT1 oder CFT20) teil. Bei den Kindern der 5. Klasse wurden 3 Rechtschreibtests (DRT, WRT 6+, HSP), ein Lesetest (Zürcher) und ein nonverbalen Intelligenztest (CFT20) durchgeführt. Kinder von Förderschulen wurden mit Testverfahren, die zwei Klassenstufen unter ihrer eigentlichen Klassenstufe liegen, getestet.

Zusätzlich erfolgte eine Beurteilung der Schulleistungen durch die Lehrer anhand einer sechsstufigen Notenskala, sowie durch die Eltern anhand einer dreistufigen Skala (gut-durchschnittlich-schlecht) im Rahmen eines Fragebogens.

### Ergebnisse:

Die meisten Rechtschreibtests erbringen unseren Ergebnissen zufolge eine vergleichbare Beurteilung der Rechtschreibleistung. Alle Rechtschreibtests über die Klassenstufen 1 - 5 korrelieren weitgehend hoch ( $r = 0,7 - 0,9$ ) miteinander. Auch die meisten Mittelwerte der T-Werte stellen sich vergleichbar dar. Eine Ausnahme zeigt sich jedoch in der Hamburger Schreibprobe. Im Vergleich zu den anderen Testverfahren bewertet die Hamburger Schreibprobe die Rechtschreibleistungen in der 1., 3. und 4. Klasse deutlich milder als die anderen Rechtschreibtests.

Auch die Lesetests für die 1. - 4. Klasse präsentierten sich weitgehend übereinstimmend in ihren Ergebnissen ( $r = 0,7 - 0,9$ ). Nur der Score Vorläuferfertigkeit von Knuspels Leseaufgaben korreliert etwas schlechter mit den anderen Lesetests ( $r = 0,3 - 0,7$ ). KNUSPEL-L misst neben dem Leseverständnis auch die Vorläuferfertigkeiten (Basisfertigkeiten der Lesefähig-

keit: Hörverstehen, Rekodieren, Dekodieren), die mit den anderen Lesetests schlechter übereinstimmen.

Es besteht bis auf eine Ausnahme kein nennenswerter Mittelwertunterschied zwischen den T-Werten der Lesetests, so dass davon auszugehen ist, dass die Normierungen der Testverfahren weitgehend übereinstimmen. Diese Abweichung zeigt sich bei den dritten Klassen. Die Ergebnisse des Zürcher im Vergleich zum SLRT Lesetest fallen etwas niedriger aus, die des SLRT im Vergleich etwas höher. Am ehesten muss man davon ausgehen, dass der SLRT in der 3. Klasse zu milde misst.

Um zu überprüfen, welche der Tests die tatsächlichen Lese- und Rechtschreibleistungen der Kinder am besten widerspiegeln, wurde eine Einschätzung der Schulleistungen durch die Lehrer und durch die Eltern erfragt. Mehrere Untersuchungen konnten nachweisen, dass sowohl die Eltern-, wie auch die Lehrerbeurteilung sehr kritisch einzuschätzende Parameter sind, die vielen Störeinflüssen unterliegen. Auf der anderen Seite haben nur die Lehrkräfte die gesamte Persönlichkeit des Lernenden im Blick und beurteilen die Ergebnisse auf dem Hintergrund länger andauernder Lernprozesse und unter Einbeziehung der institutionellen und personellen Rahmenbedingungen. Wie die Erfahrung aber insgesamt gezeigt hat, bildet das Lehrerurteil die tatsächlichen Lese-Rechtschreibleistungen nur ungenau ab. Trotzdem wird es häufig als Außenkriterium zur Validierung von Lese-Rechtschreibtests herangezogen, da keine zuverlässigeren Kriterien zur Verfügung stehen. Auch in dieser Studie wird das Lehrerurteil zur Abschätzung der Treffsicherheit der Testverfahren bei der Erfassung eines LRS-Kindes benutzt.

Lehrer wie auch Eltern zeigen unterschiedlich hohe Übereinstimmungen mit den verwendeten LRS-Tests. Bei den Rechtschreibtests stimmen Lehrer, wie auch Eltern am besten ( $r = 0,7 - 0,8$ ) mit dem Weingartener Rechtschreibtest und am schlechtesten ( $r = 0,2 - 0,4$ ) mit dem Westermann Rechtschreibtest überein. Alle weiteren Rechtschreibtests zeigen etwa mittelgradige Übereinstimmungen mit dem Lehrer- bzw. Elternurteil ( $r = 0,5 - 0,6$ ). Bei den Lesetests stimmen die Lehrer besser mit den Testergebnissen überein als die Eltern. Die Übereinstimmung zwischen Lehrerurteil und Tests ist für die meisten Lesetests besser ( $r = 0,7 - 0,8$ ) als für die Rechtschreibtests. Dies trifft nicht für den Zürcher und den Score Vorläuferfertigkeiten von Knuspels Leseaufgaben ( $r = 0,5 - 0,6$ ) zu.

Bei der diagnostischen Zuordnung der einzelnen Kinder sind bei Verwendung unterschiedlicher Testverfahren die Ergebnisse nicht immer übereinstimmend. Dies trifft insbesondere für die Kinder zu, die mit ihrer Leistungsfähigkeit im Grenzbereich liegen. Ein Grund hierfür könnte die Wahl einer „cut-off“ Grenze sein, wie sie bei der operationalisierten Diagnostik notwendig ist (eine Teilleistungsstörung ist anzunehmen, wenn das Testergebnis einem T-Wert von  $\leq 37$  entspricht und eine T-Wert Diskrepanz zum Gesamt-IQ  $\geq 12$  besteht). Kinder, die knapp oberhalb bzw. unterhalb der Grenze liegen, werden schon bei geringen Abweichungen der Ergebnisse in den verschiedenen Tests, unterschiedlichen Kategorien zugeordnet.

Eine Einstufung von Kindern als rechtschreibschwach mit dem SLRT-RT in der 1. Klasse ist problematisch. Bezieht man sich auf die orthographischen Fehler, die im Testhandbuch als das zentrale Auswertungskriterium genannt werden, so werden sehr viele Erstklässler, die vor allem nicht-lauttreue Fehler machen, unzureichend als rechtschreibschwach erkannt. Wenn anstelle der orthographischen Fehler die nicht-lauttreuen Fehler als Beurteilungskriterium verwandt werden, werden zu viele Kinder als auffällig eingestuft. Die „cut-off“ Grenze für die nicht-lauttreuen Fehler liegt dementsprechend zu hoch.

Wieder bildete die Hamburger Schreibprobe einen Sonderfall, da mit ihr besonders wenige Schüler eine LRS-Diagnose erhielten. Ihre Normierung führt zu einer zu milden Bewertung der Rechtschreibleistung.

Abschließend interessierte uns noch die Trefferquote einzelner LRS-Tests in Bezug auf die Erfassung einer Lese-Rechtschreibschwäche. Die Trefferquote der Lesetests für die Schüler, die von den Lehrern als schlechte Leser eingestuft wurden, ist als relativ hoch einzustufen (RATZ-Index 65 - 100). Eine Ausnahme bildet nur der Score Vorläuferfertigkeit von Knuspels Leseaufgaben (RATZ-Index 25). Demgegenüber ist die Trefferquote der Rechtschreibtests für Schüler, die von den Lehrern als rechtschreibschwach eingestuft wurden, eher gering. Immerhin erreichten die meisten Rechtschreibtests (mit Ausnahme der HSP und dem WRT 4/5 - 6+) eine gute (RATZ Index 30 - 45) und die WRT 1+ - 3+ eine sehr gute Aussage (RATZ-Index 76).

### Schlussfolgerung:

Die meisten Lese- und Rechtschreibtest zeigten in unseren Untersuchungen relativ gute Übereinstimmungen hinsichtlich ihrer Ergebnisse. Bei einigen Tests zeigten sich jedoch Abweichungen, die beachtet werden sollten.

Als Ausnahme ergibt die Hamburger Schreibprobe eine bessere Einschätzung der Rechtschreibleistung, was auf unterschiedliche Normwerte zurückzuführen ist. Dies führt auch dazu, dass ein Kind mit der HSP seltener als rechtschreibschwach eingestuft wird als mit den anderen Testverfahren. 2002 erschien eine Neuauflage mit neuer Normierung. Die Normwerte sind gegenüber der Vorversion aber nur unwesentlich strenger geworden.

Zur Einstufung rechtschreibgestörter Kinder in der 1. Klasse ist der SLRT-Rechtschreibtest wenig geeignet. Als zentrales Auswertungskriterium nennt das Testverfahren die Anzahl der orthographischen Fehler, was gerade bei Erstklässlern, die noch sehr viele nicht-lauttreue Fehler machen, zu einer besseren Einschätzung der Rechtschreibleistung führt und somit wenig geeignet erscheint.

Wenn es darum geht, Kinder mit schlechten Lese- und Rechtschreibleistungen gemäß dem Eindruck der Lehrer/innen zu erfassen, zeigt sich, dass die Lesetests eine höhere Trefferquote bezüglich der Leseleistung als die Rechtschreibtests bezüglich der Rechtschreibleistung erzielen. Lesetests sollten daher in der Praxis obligat eingesetzt werden.

Wie die Ergebnisse aber auch zeigen, ist die diagnostische Zuordnung der Testverfahren mit gewissen Fehlern behaftet. Bei Kindern, bei denen Testergebnis und Lehrerurteil deutlich voneinander abweichen, sind deshalb außer der formalen Testergebnisse weitere klinische Beurteilungskriterien bei der diagnostischen Zuordnung heranzuziehen. Ergebnisse von Testverfahren und das Lehrerurteil sollten von Ärzten, Psychologen, Lehrern und anderen Diagnostikern als gegenseitige Kontrollinstanzen betrachtet und ergänzt werden.

## 8. ABKÜRZUNGSVERZEICHNIS

CFT 1 / CFT 20	Grundintelligenztest Skala 1 und Skala 2
CI	Konfidenzintervall
d. h.	das heißt
DRT 1 - 5	Diagnostische Rechtschreibtest für die 1. – 5. Klasse
et al.	und Andere
HSP 1+ - 5-9	Hamburger Schreib-Probe für die 1. – 9. Klasse
ICD 10	International classification of disease, Nr 10
IQ	Intelligenzquotient
Kl.	Klasse
KNUSPEL-L	Knuspels Leseaufgaben
LF	Lesefähigkeit (Score des KnuspeL-L)
LRS	Lese- und Rechtsschreibstörung
MW	Mittelwert
n.s.	nicht signifikant
SEM	Standard error of mean
SLRT	Salzburger Lese-Rechtschreibtest
SLRT-RT	Salzburger Rechtschreibtest
SLRT-LT	Salzburger Lesetest
VF	Vorläuferfertigkeit (Score des KnuspeL-L)
vs.	versus
WLLP	Würzburger Leise Leseprobe
WRT 1+ - 3+	Weingartener Grundwortschatz Rechtschreib-Testreihe für die 1. – 4. Klasse
WRT4/5 - 6+	Westermann Rechtschreibtest für die 4. – 7. Klasse
z. B.	zum Beispiel
ZLT	Zürcher Lesetest
ZNS	Zentrales Nervensystem

## 9. LITERATURVERZEICHNIS

- Aaron PG.: The impending demise of the discrepancy formula. *Review of Educational Research*. 67(4), 461 - 502, 1997.
- Auer, M.; Gruber, G.; Mayringer, H.; Wimmer, H.: *Salzburger Lese-Screening für die Klassenstufen 5-8*. Hogrefe Verlag, 1.Auflage 2005.
- Bauer, S.: *Persönliche Befragung über Vor- und Nachteile von Lese- Rechtschreibtests und deren Verwendung in der Ambulanz für Entwicklungsstörungen der Heckscher Klinik, Klinik für Kinder- und Jugendpsychiatrie und Psychotherapie München, 2005*.
- Baumert, J.: *PISA 2000 Die Studie im Überblick. Grundlagen, Methoden und Ergebnisse*. Politische Studien, Zweimonatszeitschrift für Politik und Zeitgeschehen, Sonderheft 3/2003 - *Bildung: Standards Tests Reformen*, 54. Jahrgang. Hanns Seidel Stiftung - Atwerb-Verlag KG, 2003.
- Baving, L.; Schmidt, M.H.: *Testpsychologie zwischen Anspruch und Wirklichkeit am Beispiel der Intelligenzdiagnostik*. *Zeitschrift für Kinder- und Jugendpsychiatrie- und Psychotherapie* 28, 163 - 176, 2000.
- II. *Bekanntmachungen der Bayerischen Staatsministerien für Unterricht und Kultus und Wissenschaft, Forschung und Kunst*. KWMBI I Nr. 23, 379 - 383, 1999.
- Birkel, P.: *WRT 1+ Weingartener Grundwortschatz Rechtschreib-Test für erste und zweite Klassen*. Hogrefe Verlag, 1995.
- Birkel, P.: *WRT 2+ Weingartener Grundwortschatz Rechtschreib-Test für zweite und dritte Klassen*. Hogrefe Verlag, 1994.
- Birkel, P.: *WRT 3+ Weingartener Grundwortschatz Rechtschreib-Test für dritte und vierte Klassen*. Hogrefe Verlag, 1994.

- Birkel, P.; Birkel, C.: Wie einig sind sich Lehrer bei der Aufsatzbeurteilung? Eine Replikationsstudie zur Untersuchung von Rudolf Weiss. *Psychologie in Erziehung und Unterricht* 49, 219 - 224, 2002.
- Boelte, S.; Adam-Schwebe, S.; Englert, E.; Schmeck, K.; Poustka, F.: Zur Praxis der psychologischen Testdiagnostik in der deutschen Kinder- und Jugendpsychiatrie: Ergebnisse einer Umfrage. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie* 28, 151 - 161, 2000.
- Breuer A.E.: Persönliche Befragung über Vor- und Nachteile von Lese- Rechtschreibtests und deren Verwendung in der Ambulanz für Entwicklungsstörungen der Heckscher Klinik, Klinik für Kinder- und Jugendpsychiatrie und Psychotherapie München, 2005.
- Brickenkamp Handbuch psychologischer und pädagogischer Tests Teil 1, (Hrsg.) Brähler, E.; Holling, H.; Leutner, D.; Petermann, F.. Hogrefe Verlag, 2002.
- Bruck, M.: Outcomes of adults with childhood histories of dyslexia. In: C. Hulme; R. M. Joshi (Hrsg.). *Reading and Spelling: Development and Disorders*. Lawrence Erlbaum Associates, Mahwah, NJ, London, 179 – 200, 1998.
- Cohen, J.: Learning disabilities and the college student: identification and diagnosis. *Adolesc.Psychiatry* 11, 177 - 198, 1983.
- Deimel, W.: Testverfahren zur Diagnostik der Lese-Rechtschreibstörung – eine Übersicht. In *Legasthenie: Zum aktuellen Stand der Ursachenforschung der diagnostischen Methoden und der Förderkonzepte*. Herausgegeben von Gerd Schulte-Körne in Zusammenarbeit mit dem Bundesverband Legasthenie. Winkler Verlag Bochum, 2002.
- Deimel, W.: Diagnostik der Lese-Rechtschreibstörung. In *Legasthenie: Zum aktuellen Stand der Ursachenforschung der diagnostischen Methoden und der Förderkonzepte*. Herausgegeben von Gerd Schulte-Körne in Zusammenarbeit mit dem Bundesverband Legasthenie. Winkler Verlag Bochum, 2002.

- Leitlinien zur Diagnostik und Therapie von psychischen Störungen im Säuglings-, Kindes- und Jugendalter. Deutsche Gesellschaft für Kinder- und Jugendpsychiatrie und Psychotherapie, Berufsverband der Ärzte für Kinder- und Jugendpsychiatrie und Psychotherapie in Deutschland, Bundesarbeitsgemeinschaft der leitenden Klinikärzte für Kinder- und Jugendpsychiatrie und Psychotherapie. Deutscher Ärzte-Verlag Köln, 2003.
- Deutsches PISA-Konsortium. Baumert, J.; Klieme, E.; Neubrand, M.; Schiefle, U.; Scheider, W.; Stanat, P.; Tillmann, K-J.; Weiß, M. (Hrsg.), PISA 2000.
- Dilling, H. et al.: Internationale Klassifikation psychischer Störungen: Klinisch diagnostische Leitlinien. Huber Bern, 2004.
- Dilling, H. et al.: Internationale Klassifikation psychischer Störungen: Diagnostische Kriterien für Forschung und Praxis. Huber Bern, 2004.
- Eggers, C.; Fegert, J.M.; Resch, R.: Psychiatrie und Psychotherapie des Kindes- und Jugendalters. Springer Verlag, 2004.
- Esser, G.; Wyschkon, A.; Schmidt, M.H.: Was wird aus Achtjährigen mit einer Lese- und Rechtschreibstörung. Ergebnisse im Alter von 25 Jahren. Zeitschrift für Klinische Psychologie- und Psychotherapie 31, 235 - 242, 2002.
- Esser, G.; Schmidt, M.H.: Umschriebene Entwicklungsstörungen. In G. Esser (Hrsg.) Lehrbuch der klinischen Psychologie und Psychotherapie des Kinder- und Jugendalters. S. 134 - 151, Thieme Verlag, 2002.
- Faust, G.: PISA und die Grundschule – Interpretation der Befunde und mögliche Konsequenzen. Politische Studien, Zweimonatszeitschrift für Politik und Zeitgeschehen, Sonderheft 3/2003 - Bildung: Standards Tests Reformen, 54. Jahrgang. Hanns Seidel Stiftung - Atwerb-Verlag KG, 2003.
- Gasteiger Klicpera, B.; Klicpera, C.; Schabmann, A.: Wahrnehmung der Schwierigkeiten lese- und rechtschreibschwacher Kinder durch die Eltern: Pygmalion im Wohnzimmer. Praxis der Kinderpsychologie und Kinderpsychiatrie 50, 622 - 639, 2001.



- Gaus, W.; Hingst, V.; Mattern, R.; Reinhardt, G.; Seidel, H. J.; Sonntag, H.-G.: Ökologisches Stoffgebiet 3, 54 - 54, 1999.
- Giovannini, S.; Haffner, J.; Parzer, P.; Steen, R.; Klett, M.; Resch, F.: Verhaltensauffälligkeiten bei Erstklässlern aus Sicht der Eltern und der Lehrerinnen. *Praxis der Kinderpsychologie und Kinderpsychiatrie* 54(2), 104 - 125, 2005.
- Grisseemann, H.; Baumberger, W.: ZLVT 4 - 6, Zürcher Leseverständnistest für das 4. bis 6. Schuljahr. Hogrefe Verlag, 2000.
- Grund, M.; Haug, G.; Naumann, C.L.: DRT 4 Diagnostischer Rechtschreibtest für 4. Klassen. Hogrefe Verlag, 1994.
- Grund, M.; Haug, G.; Naumann, C.L.: DRT 4 Diagnostischer Rechtschreibtest für 4. Klassen. Hogrefe Verlag, 2003.
- Grund, M.; Haug, G.; Naumann, C.L.: DRT 5 Diagnostischer Rechtschreibtest für 5. Klassen. Hogrefe Verlag, 1995.
- Grund, M.; Haug, G.; Naumann, C.L.: DRT 5 Diagnostischer Rechtschreibtest für 5. Klassen. Hogrefe Verlag, 2003.
- Gutzelnig, B.: Das Leseverständnis: Eine Untersuchung unter Berücksichtigung des Salzburger Lesetests. Unveröffentlichte Diplomarbeit, Universität Salzburg, 1995.
- Haffner, J.; Zerahn-Hartung, C.; Pfuller, U.; Parzer, P.; Strehlow, U.; Resch, F.: Effects and significance of specific spelling problems in young adults - empirical studies in an epidemiologic patient sample. *Zeitschrift f. Kinder- und Jugendpsychiatrie u. Psychotherapie* 26(2), 124 - 135, 1998.
- Hemminger, U.; Roth, E.; Schneck, S.; Jans, T.; Warnke, A.: Testdiagnostische Verfahren zur Überprüfung der Fertigkeiten im Lesen, Rechtschreiben und Rechnen. Eine kritische Übersicht. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie* 28, 188 - 201, 2000.

- Höger, C. et al.: Anforderungen an Diagnostik und Behandlung umschriebener Entwicklungsstörungen schulischer Fertigkeiten. Zur Diskussion gestellt. Niedersächsischer Arbeitskreis Kinder- und Jugendpsychiatrie. 79 - 84, 1999.
- Huber, G.; Gross, G.: Psychiatrie. Schattauer Verlag, 2005.
- Klicpera, C.; Schabmann, A.; Gasteiger-Klicpera, B.: Lesen- und Schreibenlernen während der Pflichtschulzeit: Eine Längsschnittuntersuchung über die Häufigkeit und Stabilität von Lese- und Rechtschreibschwierigkeiten in einem Wiener Schulbezirk. Zeitschrift für Kinder- und Jugendpsychiatrie 21, 214 - 225, 1993.
- Klicpera, Ch.; Gasteiger Klicpera, G.: Lesen und Schreiben: Entwicklung und Schwierigkeiten (unter Mitarbeit von Alfred Schabmann). Verlag Hans Huber, 1993.
- Klicpera, Ch.; Gasteiger Klicpera, G.: Psychologie der Lese-Rechtschreibschwierigkeiten. Entwicklung, Ursachen, Förderung. Beltz Weinheim, 1995.
- Klicpera, Ch.; Gasteiger Klicpera, G.: Macht Intelligenz einen Unterschied? Rechtschreiben und phonologische Fertigkeiten bei diskrepanten und nichtdiskrepanten Lese/Rechtschreibschwierigkeiten. Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie 29(1), 37 - 49, 2001.
- Klicpera, C.; Schabmann, A.; Gasteiger-Klicpera, B.: Legasthenie. Verlag Ernst Reinhardt, 2003.
- Kornadt, H.-J.: Abschlussbericht über die wissenschaftliche Begleituntersuchung zum Schulversuch Oberstufe Saar 1970 - 1977. S.283, 1978.
- Küspert, P.; Schneider, W.: WLLP Würzburger Leise Leseprobe. Hogrefe Verlag, 1998.
- Küspert, P.; Schneider, W.: WLLP Würzburger Leise Leseprobe. In Hasselhorn, M.; Schneider, W.; Marx, H.: Diagnostik von Lese-Rechtschreibschwierigkeiten. Tests und Trends. N. F. Band 1. Jahrbuch der pädagogisch-psychologischen Diagnostik. Hogrefe Verlag, 2000.

Landerl, K.; Wimmer, H.; Moser, E.: SLRT Salzburger Lese- und Rechtschreibtest. Hogrefe Verlag, 1997.

Landerl, K.: Der Salzburger Lese- und Rechtschreibtest (SLRT). In Hasselhorn, M.; Schneider, W.; Marx, H.: Diagnostik von Lese-Rechtschreibschwierigkeiten. Tests und Trends. N. F. Band 1. Jahrbuch der pädagogisch-psychologischen Diagnostik. Hogrefe Verlag, 2000.

Legasthenie Erlass von Bayern (unter [www.legasthenie.net](http://www.legasthenie.net)).

Leonhardt & Franz Wember: Grundfragen der Sonderpädagogik. S. 244 - 268; Beltz, 2003.

Lindner, M.; Grisseemann, H.: Zürcher Lesetest (ZLT), 6. Auflage. Hogrefe Verlag, 2000.

Marx, H.: KNUSPEL-L Knuspels Leseaufgaben. Hogrefe Verlag, 1998.

Marx, H.: Knuspels Leseaufgaben: Theorie, Umsetzung und Überprüfung. In Hasselhorn, M.; Schneider, W.; Marx, H.: Diagnostik von Lese-Rechtschreibschwierigkeiten. Tests und Trends. N. F. Band 1. Jahrbuch der pädagogisch-psychologischen Diagnostik. Hogrefe Verlag, 2000.

Marx, H.; Jansen, H.; Skowronek, H.: Prognostische, differentielle und konkurrente Validität des Bielefelder Screenings zur Früherkennung von Lese-Rechtschreibschwierigkeiten (BISC). In Hasselhorn, M.; Schneider, W.; Marx, H.: Diagnostik von Lese-Rechtschreibschwierigkeiten. Tests und Trends. N. F. Band 1. Jahrbuch der pädagogisch-psychologischen Diagnostik. Hogrefe Verlag, 2000.

May, P.: Lernförderlichkeit im schriftsprachlichen Unterricht. Effekte des Klassen und Förderunterrichts in der Grundschule auf den Lernerfolg. Ergebnisse der Evaluation des Projekts "Lesen und Schreiben für alle" (PLUS). Behörde für Schule, Jugend und Berufsbildung Hamburg, 2000.

May, P.: HSP 1-9 Hamburger Schreib-Probe für 1. bis 9. Klassen. Hogrefe Verlag, 2000.

May, P.: HSP 1-9 Hamburger Schreib-Probe für 1. bis 9. Klassen. Hogrefe Verlag, 2002.

- May, P.; Malitzky, V.; Vieluf, U.: Rechtschreibtests im Vergleich: Wie stellt man deren Güte fest und wie besser nicht? Anmerkungen zur Kritik von Tacke, Voelker und Lohmueller an der HSP. *Psychologie in Erziehung und Unterricht* 48, 146 - 152, 2001.
- May, P.: HSP 1 - 9 Hamburger Schreib-Probe für 1. bis 9. Klassen, Umrechnung von alten Strategie-Normen auf neue Werte. Unter [www.peter-may.de](http://www.peter-may.de), 2002.
- May, P.: Lernstandsdiagnose im Rechtschreiben – mit und ohne Test. *Grundschule* 34. Jg., Heft 05, S. 44 - 46, 2002<sup>2</sup>.
- Mayringer, H.; Wimmer, H.: Kognitive Defizite lese-rechtschreibschwacher Kinder. *Kindheit und Entwicklung* 8(3), 141 – 146, 1999.
- Mayringer, H.; Wimmer, H.: Pseudoword Learning by German-Speaking Children with Dyslexia: Evidence for a Phonological Learning Deficit. *Journal of Experimental Child Psychology* 75, 116 - 133, 2000.
- Mayringer, H.; Wimmer, H.: Salzburger Lese-Screening für die Klassenstufen 1 – 4. 1. Auflage, Hogrefe Verlag, 2003.
- Menzel, Wolfgang: *Rechtschreibunterricht*. Seelze, 1985.
- Meyer MS.: The ability-achievement discrepancy: Does it contribute to an understanding of learning disabilities? *Educational Psychology Review* 12(3), 315 - 337, 2000.
- Müller, R.: DRT 1 Diagnostischer Rechtschreibtest für 1. Klassen. Hogrefe Verlag, 1990.
- Müller, R.: DRT 1 Diagnostischer Rechtschreibtest für 1. Klassen. Hogrefe Verlag, 2003.
- Müller, R.: DRT 2 Diagnostischer Rechtschreibtest 2. Hogrefe Verlag, 1997.
- Müller, R.: DRT 2 Diagnostischer Rechtschreibtest 2. Hogrefe Verlag, 2003.
- Müller, R.: DRT 3 Diagnostischer Rechtschreibtest 3. Hogrefe Verlag, 1997.
- Müller, R.: DRT 3 Diagnostischer Rechtschreibtest 3. Hogrefe Verlag, 2003.

- Niebergall, G.: Diagnostische Aspekte der Legasthenie. Monatsschrift Kinderheilkunde 135, 297 - 301, 1987.
- Noterdaeme, M.; Breuer-Schaumann A.E.: Lesen und Schreiben Bausteine des Lebens, ein Übungsprogramm zum Schriftspracherwerb. Verlag modernes lernen Dortmund, 2003.
- OECD (Organisation für wirtschaftliche Zusammenarbeit und Entwicklung): Lernen für das Leben. Erste Ergebnisse der Internationalen Schulleistungsstudie Pisa 2000. 2001.
- Otto, J.: Nutzloses Aussortieren! Ein Gespräch mit dem Bildungsforscher Klaus-Jürgen Tillmann über alte Vorurteile und neue Schulmodelle. Zeit, 2002.
- Petermann, F. & Macha, T.: Psychologische Tests für Kinderärzte. Hogrefe Göttingen, 2005.
- Probst, H.: Unterrichtsbegleitende Diagnostik im schriftsprachlichen Lernbereich. In: Annette Leonhardt & Franz Wember: Grundfragen der Sonderpädagogik. S. 244 - 268, Beltz, 2003.
- Rathenow, P.: WRT 4/5 Westermann Rechtschreibtest 4/5. Hogrefe Verlag, 1980.
- Rathenow, P.; Laupenmuehlen, D.; Voegelé, J.: WRT 6+ Westermann Rechtschreibtest 6+. Hogrefe Verlag, 1980.
- Remschmidt, H.: Multiaxiales Klassifikationsschema für psychische Störungen des Kindes- und Jugendalters nach ICD-10 der WHO. Huber Bern, 2006.
- Reuter-Liehr, C.: Lautgetreue Lese-Rechtschreibförderung, Band 1. Eine Einführung in das strategiegeleitete Lernen zum Training von Phonemstufen auf der Basis des rhythmischen Syllabierens. Verlag Dr. Dieter Winkler Bochum, 2001.
- Reuter-Liehr, C.: Konsequenter strategiegeleitetes Lernen beim lese-rechtschreibschwachen Kind. In: Schulte-Körne, G. (Hrsg.) Legasthenie: Zum aktuellen Stand der Ursachenfor-

- schung, der diagnostischen Methoden und der Förderkonzepte. Verlag Dr. Dieter Winkler Bochum, 359 - 384, 2002.
- Reuter-Liehr, C.: Legasthenie Diagnose und Therapie in der kinder- und jugendpsychiatrischen Praxis. Forum der Kinder- und Jugendpsychiatrie und Psychotherapie 2, 2003.
- Richtlinien für LRS-Gutachten nach § 35a SGB VIII. Empfehlungen einer Arbeitsgruppe des Bayer. Sozialministeriums, 1996.
- Rodgers, J.L.: A critique of the Flynn Effect: Massive IQ gains, methodological artifacts, or both? Intelligence 26(4), 337 - 356, 1998.
- Schmid, R.G.: Lese- und Rechtschreibstörung. Monatsschrift Kinderheilkunde 152, 1111 - 1121, 2004.
- Schmieder, S.: Persönliche Befragung über Vor- und Nachteile von Lese- Rechtschreibtests und deren Verwendung in der Ambulanz für Entwicklungsstörungen der Heckscher Klinik, Klinik für Kinder- und Jugendpsychiatrie und Psychotherapie München, 2005.
- Schrader, F.-W.; Helmke, A.: Wahrnehmung der Schwierigkeiten lese-rechtschreibschwacher Kinder durch die Eltern: Pygmalion im Wohnzimmer. Praxis der Kinderpsychologie- und Kinderpsychiatrie 50(8), 622 - 639, 2001.
- Schulte-Körne, G.; Deimel, W.; Remschmidt, H.: Die Bedeutung der phonologischen Dekodierfähigkeit und orthographischem Wissen für die Rechtschreibfähigkeit Erwachsener. Sonderdruck aus: Zeitschrift für Klinische Psychologie 26(3), 210 - 217, 1997.
- Schulte-Körne, G.: Zur Diagnostik der Lese-Rechtschreibstörung. Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie 29, 113 - 116, 2001.
- Schulte-Körne, G.; Deimel, W.; Jungermann, M.; Remschmidt, H.: Nachuntersuchung einer Stichprobe von lese-rechtschreibgestörten Kindern im Erwachsenenalter. Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie 31(4), 267 - 276, 2003.

- Schulte-Körne, G.; Bartling, J.; Deimel, W.: Forschungsgruppe Lese-Rechtschreibstörungen. Unter [www.kjp.uni-marburg.de](http://www.kjp.uni-marburg.de), 2006.
- Shaywitz, S.E.; Fletcher, J.M.; Holohan, J.M.; Schneider, A.E.; Marchione, K.E.; Stuebing, K.K.; Francis, D.J.; Pugh, K.R. & Shaywitz, B.A.: Persistence of dyslexia: the connecticut longitudinal study at adolescence. *Pediatrics* 104, 1351 – 1359, 1999.
- Sommer-Stumpfenhorst, N.: Lese- und Rechtschreibschwierigkeiten vorbeugen und überwinden. Cornelsen Scriptor Verlag Frankfurt am Main, 1993.
- Steck, P.: Psychologische Testverfahren in der Praxis Ergebnisse einer Umfrage unter Testanwendern. *Diagnostica* 43, 267 - 284, 1997.
- Strehlow, U.: Der Verlauf der umschriebenen Lese-Rechtschreibschwäche. *Sprache-Stimme-Gehör* 22, 31 - 33, 1998.
- Strehlow, U.; Haffner, J.: Definitionsmöglichkeiten und sich daraus ergebende Häufigkeit der umschriebenen Lese- bzw. Rechtschreibstörung - theoretische Überlegungen und empirische Befunde an einer repräsentativen Stichprobe junger Erwachsener. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie* 30, 113 - 126, 2002.
- Suchodoletz, W. v.: Therapie der Lese-Rechtschreibstörung. Traditionelle und alternative Behandlungsverfahren im Überblick. Kohlhammer Verlag Stuttgart, 1999.
- Suchodoletz, W. v.: 100 Jahre LRS-Forschung – was wissen wir heute? *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie* 27(3), 199 – 206, 1999.
- Tacke, G.; Voelker, R.; Lohmueller, R.: Die Hamburger Schreibprobe: Probleme mit einem neuen Rechtschreibtest. *Psychologie in Erziehung und Unterricht* 48, 135 - 145, 2001.

- Wagner, K.R.; Steintraeter,C.: Test der Sprachtests. Wie valide ist der Wortschatz in Sprach- und Rechtschreibtests für die Grundschule. *Psychologie in Erziehung und Unterricht* 34, 145 - 150, 1987.
- Warnke, A.: Legasthenie. Diagnostik, Therapie und neue neuropsychologische Befunde zur Ätiologie. *Pädiatrische Praxis* 42, 11 - 22, 1991.
- Warnke, A.; Hemminger, U.; Roth, E.; Schneck, S.: *Legasthenie, Leitfaden für die Praxis*. Hogrefe Verlag, 2002.
- Warnke, A.: Umschriebene Entwicklungsstörungen des Lesens und der Rechtschreibung. In *Entwicklungspsychiatrie – Biopsychologische Grundlagen und die Entwicklung psychischer Störungen*, (Hrsg.) B. Herpertz-Dahlmann, F. Resch, M. Schulte-Markwort & A. Warnke. Schattauer Verlag Stuttgart, S. 404 - 436, 2003.
- Warnke, A.; Hemminger, U.; Plume, E.: *Lese-Rechtschreibstörungen. Leitfaden Kinder- und Jugendpsychotherapie, Band 6*. Hogrefe Verlag, 2004.
- Weiss, R.; Osterland, J.: *CFT 1 Grundintelligenztest Skala 1*. Hogrefe Verlag, 1997.
- Weiss, R. H.: *CFT 20 Grundintelligenztest Skala 2*. Hogrefe Verlag, 1998.
- Wild, K.P.: *Hochbegabendiagnostik durch Lehrer*. Dissertation, 1993.
- Zerahn-Hartung, C.; Strehlow, U.; Haffner, J.; Pfüller, U.; Parzer, P.; Resch, F.: Normverschiebungen bei Rechtschreibleistung und sprachfreier Intelligenz. *Praxis der Kinderpsychologie und Kinderpsychiatrie* 51, 281 - 297, 2002.
- DIE ZEIT: Note zwei minus. Deutsche Grundschüler lesen gut – und landen oft trotzdem auf der falschen Schule. Fragen an den Iglu-Experten Wilfried Bos. *DIE ZEIT* Nr.16, April 2003.



## **10. ANHANG**

- Elternbrief
- Anamnesebogen für die Eltern
- Einverständniserklärung
- Bogen zur Beurteilung der Schulleistungen durch die Lehrer

**Klinik Hochried**  
Fachklinik für  
Kinder- und Jugendmedizin  
Chefarzt Dr. Hermann Mayer

Anschrift: Hochried 1-12  
82418 Murnau  
Tel.: 08841/ 474-0  
Fax.: 08841/ 474-111  
Email: mayer@klinikhochried.de

**Ludwig-Maximilians-Universität München**  
Institut und Poliklinik für  
Kinder- und Jugendpsychiatrie  
komm. Direktor Prof. Dr. H.–J. Möller

Anschrift: Waltherstr. 23  
80337 München  
Tel. Nr.: 089/ 5160/ 3427  
Fax Nr.: 089/ 5160/ 4756  
Email: kjp@lrz.uni-muenchen.de

*Liebe Eltern,*

viele Kinder haben in der Schule Schwierigkeiten. Häufige Ursache ist eine Lese-Rechtschreibstörung. Durch eine frühzeitige Förderung kann diesen Kindern geholfen werden.

In Zusammenarbeit mit der Klinik Hochried bemüht sich die Kinder- und Jugendpsychiatrie der Universität München darum, dass Lese-Rechtschreibstörungen früher und zuverlässiger erkannt werden können. Mit der jetzigen Studie wollen wir untersuchen, mit welchen Methoden die Diagnose „Lese- und Rechtschreibstörung“ am besten und sichersten gestellt werden kann. Damit soll erreicht werden, dass die Kinder in Zukunft frühzeitiger Unterstützung erhalten.

Um dies zu erreichen, müssen wir viele Kindern untersuchen. Daher möchten wir auch Sie und Ihr Kind bitten, sich an der Studie zu beteiligen.

**Was erwartet Ihr Kind:** Es werden mehrere Rechtschreibe-, Lese- und allgemeine Entwicklungstests durchgeführt. Die Rechtschreibetests bestehen aus altersentsprechenden Wortdiktaten und die Lesefähigkeit wird durch Vorlesen ermittelt. Die allgemeine geistige Entwicklung wird spielerisch mit Bilderrätseln, Sortieraufgaben u. a. untersucht. Alle Ergebnisse werden für die Auswertung verschlüsselt und unterliegen der Schweigepflicht. Gerne erhalten Sie im Anschluss an die Untersuchung Auskunft über die Ergebnisse Ihres Kindes.

Die Untersuchungen können während des Aufenthaltes in Hochried durchgeführt werden.

**Wenn Ihr Kind bei unserem Projekt mitmachen darf, füllen Sie bitte die Einverständniserklärung aus und schicken uns diese und den ausgefüllten Fragebogen zu. Eine andere Möglichkeit ist, die ausgefüllten Unterlagen zu Ihrem ersten Besuch in der Klinik Hochried mitzubringen.**

**Wir wissen, dass wir bereits vor der Aufnahme Ihres Kindes viel von Ihnen verlangen, da Sie auch noch Fragebögen der Klinik auszufüllen haben. Selbstverständlich bekommen Sie dann aber auch spezielle und umfangreichere Untersuchungsergebnisse über Ihr Kind .**

In das Kuvert bitte:

- ü Einverständniserklärung
- ü Fragebogen zur Entwicklung und früheren Erkrankungen

Eine Nicht-Teilnahme an der Studie beeinflusst die Behandlung Ihres Kindes in Hochried nicht.

Falls Sie noch irgendwelche Fragen zu Inhalt oder Ablauf unserer Untersuchungen haben, können Sie sich gerne telefonisch an uns wenden. Sie erreichen uns von Montag bis Freitag von 8.00 bis 17.00 unter der Münchner Tel.-Nr.: **089/5160-3427**. Am Aufnahmetag, den 29. März 2001, steht Ihnen außerdem in der Klinik Hochried ein kompetenter Ansprechpartner für weitere Fragen zur Verfügung.

**Herzlichen Dank für Ihre Unterstützung.**

**Durch Ihre Mitarbeit helfen Sie, die Situation betroffener Kinder zu verbessern.**

Mit freundlichen Grüßen

Dr. med. Hermann Mayer

Prof. Dr. med. Waldemar v. Suchodoletz

Susanne Steinhauser

Dr. Hermann Mayer, Klinik Hochried, Fachklinik für Kinder- und Jugendmedizin, Murnau

Prof. Dr. med. W. v. Suchodoletz, Institut für Kinder- und Jugendpsychiatrie, LMU München

Cand. med. Susanne Steinhauser, Institut für Kinder- und Jugendpsychiatrie, LMU München

**UNTERSUCHUNG ZUR DIAGNOSTIK**  
**VON LESE-RECHTSCHREIBSTÖRUNGEN**

**Elternfragebogen zur Entwicklung und zu früheren Erkrankung**  
 bitte Zutreffendes ankreuzen

Datum: \_\_\_\_\_ . \_\_\_\_\_ . \_\_\_\_\_

Name des Kindes: \_\_\_\_\_

Vorname: \_\_\_\_\_

Geburtsdatum: \_\_\_\_\_ . \_\_\_\_\_ . \_\_\_\_\_

**Fragen zur Schule**

1) Besucht Ihr Kind derzeit eine Schule?

a) ja • nein , b) Klasse: \_\_\_\_\_  
 Schultyp (z.B. Regel-, Montessori-, Sprachbehindertenschule): \_\_\_\_\_

2) Hat es seit der 1.Klasse Schulwechsel gegeben?

ja • nein , wenn ja, wie viele: \_\_\_ in welcher Klasse: \_\_\_\_\_  
 bei Wechsel des Schultyps - welcher Wechsel: \_\_\_\_\_

3) Hat Ihr Kind eine oder mehrere Klassen wiederholt?

ja • nein , wenn ja, welche: \_\_\_\_\_

4) Wie oft konnte Ihr Kind in diesem Schuljahr (seit Sept. 2000) die Schule nicht besuchen (z.B. wegen Krankheit)?

länger als 1 Monat • etwa 1-4 Wochen , weniger als 1 Woche *f*

5) Wie oft konnte Ihr Kind die Schule wegen Unterrichtsausfall in diesem Schuljahr(seit Sept. 2000) nicht besuchen (z.B. Erkrankung des Lehrers) ?

länger als 1 Monat • etwa 1-4 Wochen , weniger als 1 Woche *f*

6) Geht Ihr Kind gerne zur Schule?

ja • nein ,

7) Hat Ihr Kind manchmal Angst vor dem Schulbesuch?

nie • selten , oft *f*

8) Hat Ihr Kind in Zusammenhang mit der Schule körperliche Beschwerden (Kopfschmerzen, Bauchschmerzen)?

nie • selten , oft *f*

9) Wird Ihr Kind wegen Schulproblemen gehänselt?

nie • selten , oft *f*

10) Wie schätzen Sie die Leistungen Ihres Kindes in der Schule ein?

	gut	durchschnittlich	schlecht	ich weiß es nicht
a) Lesen	•	,	<i>f</i>	% <sub>00</sub>
b) Diktat	•	,	<i>f</i>	% <sub>00</sub>
c) Abschreiben	•	,	<i>f</i>	% <sub>00</sub>
d) Schriftbild	•	,	<i>f</i>	% <sub>00</sub>
e) Aufsätze	•	,	<i>f</i>	% <sub>00</sub>
f) Mathematik(Rechnen)	•	,	<i>f</i>	% <sub>00</sub>
g) Mathematik(Textaufgaben)	•	,	<i>f</i>	% <sub>00</sub>
h) Zeichnen	•	,	<i>f</i>	% <sub>00</sub>
i) Sport	•	,	<i>f</i>	% <sub>00</sub>
j) Musik	•	,	<i>f</i>	% <sub>00</sub>

11) Wie lange ist Ihr Kind täglich mit Hausaufgaben beschäftigt (durchschnittlich)?

15-30 Minuten • etwa 1 Stunde , etwa 2 Stunden oder mehr *f*

12) Gibt es bei der Hausaufgabendurchführung Probleme, Streitigkeiten (macht ungern Hausaufgaben, weigert sich ...)?

nie • selten , oft *f*

13) Erhielt Ihr Kind innerhalb der Schule eine bestimmte Art von Förderung?

ja • nein ,

wenn ja, welche: \_\_\_\_\_ wie lange: \_\_\_\_\_

14) Erhielt Ihr Kind außerhalb der Schule eine bestimmte Art von Förderung?

ja • nein ,

wenn ja, welche: \_\_\_\_\_ wie lange: \_\_\_\_\_

### Fragen zum Verhalten

15) Ausmaß der Aktivität: Das zutreffende Kästchen mit einem *y* zu kennzeichnen

	überhaupt nicht (0)	ein wenig (1)	ziemlich (2)	sehr stark (3)
a) Ist unruhig oder übermäßig aktiv				
b) Ist erregbar, impulsiv				
c) Stört andere Kinder				
d) Bringt angefangene Dinge nicht zu einem Ende - kurze Aufmerksamkeitsspanne				
e) Ist ständig zappelig				
f) Ist unaufmerksam, leicht abgelenkt				
g) Erwartungen müssen umgehend erfüllt werden, ist leicht frustriert				
h) Weint leicht und häufig				
i) Zeigt schnellen und ausgeprägten Stimmungswechsel				
j) Hat Wutausbrüche, explosives und unvorhersagbares Verhalten				

**Fragen zur Entwicklung**

16) Gab es während der Schwangerschaft oder Geburt Auffälligkeiten oder Komplikationen?

ja •                      nein ,

wenn ja, welche? \_\_\_\_\_

17) Gab es in der Entwicklung Ihres Kindes Auffälligkeiten (z.B. in der Geschicklichkeit, im Verhalten, ...)?

ja •                      nein ,

wenn ja, welche? \_\_\_\_\_

18) Wie schätzen Sie den gegenwärtigen Entwicklungsstand Ihres Kindes in folgenden Bereichen im Vergleich zu Gleichaltrigen ein?

	besser	durchschnittlich	schlechter	ich weiß nicht
a) geistige Entwicklung	•	,	<i>f</i>	‰
b) motorische Geschicklichkeit	•	,	<i>f</i>	‰
c) Selbständigkeit	•	,	<i>f</i>	‰
d) Ausdauer im Spiel	•	,	<i>f</i>	‰

19) Hat oder hatte Ihr Kind Auffälligkeiten in der Sprache?

ja •                      nein ,                      Weiß ich nicht ‰

wenn ja, welche (mehrere Kreuze sind möglich)?

Stottern                      *f*                      undeutliches Sprechen                      †

Sprachhemmung                      „                      Probleme mit Satzbildung                      ‡

überhastetes Sprechen                      ...                      Lispeln                      Š

andere Sprachauffälligkeiten: ^ \_\_\_\_\_

wenn ja, in welchem Alter? \_\_\_\_\_

**Fragen zu früheren Erkrankungen**

20) Hatte Ihr Kind ernsthafte oder chronische Erkrankungen (z.B. Erkrankungen mit Bewusstlosigkeit, Kopfunfälle, epileptische Anfälle, Hirnentzündungen ...)?

ja •                      nein ,                      Weiß ich nicht ‰

wenn ja, welche? \_\_\_\_\_

21) Wurde Ihr Kind jemals augenärztlich untersucht?

ja •                      nein ,

wenn ja, Untersuchungsergebnis: keine *f* leichte „ ausgeprägte ... Sehstörung

22) Trägt Ihr Kind eine Brille?                      Ja •                      nein ,

23) Wurde bei Ihrem Kind ein Hörtest durchgeführt?

ja •                      nein ,

wenn ja, Untersuchungsergebnis: keine *f* leichte „ ausgeprägte ... Hörstörung

**Fragen zur Familie****24)** Wer hat den Bogen ausgefüllt?Mutter • Vater , beide Eltern *f*  
andere ^ \_\_\_\_\_**25)** Wie ist Ihr Familienstand?Fester Partner •  
Alleinerziehend ,**26)** Das Kind lebt:in der Familie •  
im Heim ,**27)** Welcher Art ist die derzeitige (Berufs-)Tätigkeit der Mutter / des Vaters?

	a) Mutter	b) Vater
Vollzeit	•	•
Teilzeit	,	,
Hausfrau / Hausmann	<i>f</i>	<i>f</i>

**28)** Welchen Berufsabschluß hat die Mutter / der Vater?

	a) Mutter	b) Vater
keinen Berufsabschluß	•	•
Lehre u. ä.	,	,
Hochschulabschluß u.ä.	<i>f</i>	<i>f</i>
sonstiges: _____	^	^

**29)** Anzahl der Geschwister: \_\_\_\_\_**30)** Welche Sprache wird bei Ihnen zu Hause gesprochen?deutsch • eine andere ^ , welche? \_\_\_\_\_  
mehrsprachig , welche? 1. \_\_\_\_\_ 2. \_\_\_\_\_ 3. \_\_\_\_\_**31)** Haben oder hatten Eltern oder Geschwister eine Lese-Rechtschreibschwäche?ja • nein , Weiß ich nicht %  
wenn ja, wer? \_\_\_\_\_

Wenn wir in diesem Fragebogen etwas nicht gefragt haben, das Sie bezüglich der Entwicklung Ihres Kindes für wichtig halten, dann haben Sie hier Platz für Ergänzungen.

---



---



---



---

Vielen Dank für Ihre Mühe!

**Antwort**

Ludwig-Maximilians-Universität München  
Institut für Kinder- und Jugendpsychiatrie  
Waltherstr. 23

80337 München

**Studie zur Diagnostik von Lese- und Rechtschreibstörungen**

Ich erlaube meinem Kind

Name: \_\_\_\_\_

Vorname: \_\_\_\_\_

geb. am: \_\_\_\_\_

Straße: \_\_\_\_\_

PLZ/Ort: \_\_\_\_\_

Telefon: \_\_\_\_\_

Schule/Einrichtung: \_\_\_\_\_

\_\_\_\_\_

Klasse: \_\_\_\_\_

an der Studie des Instituts für Kinder- und Jugendpsychiatrie der Universität München zur Diagnostik von Lese- und Rechtschreibstörungen teilzunehmen.

Mein Kind nimmt nicht an der Studie teil.

.....  
*Datum*

.....  
*Unterschrift des/der Sorgeberechtigten*



**Eine Bitte** an die Eltern der Kinder, die an der Studie teilnehmen dürfen.

***Liebe Eltern,***

wir möchten herausfinden, wie gut Lese-, Rechtschreib- und Entwicklungstests die tatsächlichen Leistungen der Kinder in der Schule widerspiegeln. Dazu ist es erforderlich, unsere Untersuchungsergebnisse mit den Einschätzungen der Lehrer zu vergleichen.

**Deshalb bitten wir Sie, dieses Schreiben dem/der Lehrer(in) Ihres Kindes zu geben.**

Vielen Dank für Ihre große Unterstützung

**Bitte ausfüllen**

Name des Kindes: \_\_\_\_\_

Vorname: \_\_\_\_\_

**Antwort**

Ludwig-Maximilians-Universität München  
Institut für Kinder- und Jugendpsychiatrie  
Herrn Prof. Dr. W. v. Suchodoletz  
Waltherstr. 23

80337 München

***Liebe(r) Lehrer(innen),***

in Zusammenarbeit mit der Klinik Hochried bemüht sich die Kinder- und Jugendpsychiatrie der Universität München darum, dass Lese-Rechtschreibstörungen früher und zuverlässiger erkannt werden können. Mit einer Studie wollen wir untersuchen, mit welchen Methoden die Diagnose „Lese-Rechtschreibstörung“ am besten und sichersten gestellt werden kann. Dazu überprüfen wir mehrere Lese-, Rechtschreib- und Entwicklungstests.

**Da das Lehrer(innen)urteil bei der Leistungseinschätzung der Kinder ein sehr wichtiger Bestandteil ist, möchten wir Sie bitten, die unten aufgeführte Skala für das betreffende Kind Ihrer Klasse (Name s. Rückseite) auszufüllen und uns Ihre Einschätzung seiner gegenwärtigen Fähigkeiten zuzusenden.**

Das zutreffende Kästchen (bzw. bei einer Einschätzung zwischen zwei Werten auf der Linie) mit einem  $\checkmark$  kennzeichnen

	1	2	3	4	5	6
Lesen	-----○-----	-----○-----	-----○-----	-----○-----	-----○-----	-----○-----
Diktat	-----○-----	-----○-----	-----○-----	-----○-----	-----○-----	-----○-----
Abschreiben	-----○-----	-----○-----	-----○-----	-----○-----	-----○-----	-----○-----
Schriftbild	-----○-----	-----○-----	-----○-----	-----○-----	-----○-----	-----○-----
Aufsatz	-----○-----	-----○-----	-----○-----	-----○-----	-----○-----	-----○-----
Mathematik	-----○-----	-----○-----	-----○-----	-----○-----	-----○-----	-----○-----

Vielen Dank!

Prof. Dr. W. v. Suchodoletz

Cand. med. Susanne Steinhauser

## **DANKSAGUNG**

Ich danke Herrn Professor Dr. med. Dr. med. habil. W. v. Suchodoletz für die Aufnahme am Institut für Kinder- und Jugendpsychiatrie und Psychotherapie, für sein Vertrauen in meine Arbeit und seine fortwährende Betreuung. Auch möchte ich mich bei den Mitarbeitern des Instituts für Kinder- und Jugendpsychiatrie und Psychotherapie für ihre stets offene und unterstützende Mitarbeit bedanken.

Ich möchte Herrn Dr. med. Hermann Mayer, Chefarzt der Klinik Hochried, Fachklinik für Kinder- und Jugendmedizin in Murnau und seinen Mitarbeitern für ihre sehr bereitwillige und großzügige Kooperation danken. Nur durch ihre offene und kollegiale Unterstützung konnte der praktische Teil dieser Promotion durchgeführt werden.

Ein besonderer Dank gilt Frau Privatdozentin Dr. med. Dr. med. habil. M. Noterdaeme, die mich vor allem gegen Ende der Fertigstellung dieser Promotion durch ihre fachliche Kompetenz sehr verständnisvoll und hilfreich unterstützte.

Nicht zuletzt danke ich ganz besonders meinen Eltern, die mich tatkräftig zu fast jeder Tageszeit sehr hilfreich unterstützten, sowie Andreas Bayer für den Rückhalt und die Unterstützung während der Durchführung dieser Promotion.



