

# TEILFÄHIGKEITEN DES LESENS BEI GEÜBTEN LESERN/INNEN: VALIDIERUNG EINES COMPUTERGESTÜTZTEN DIAGNOSTIKUMS<sup>1</sup>

*Johannes Naumann, Tobias Richter & Nadine van Holt<sup>2</sup>*

## *Zusammenfassung*

In zwei Studien wurde eine erste Erprobung eines auf Basis des Strategiemodell von Van Dijk & Kintsch (1983) entwickelten computergestützten Instruments zur Diagnose von Teilfähigkeiten des Lesens (*PL-Lesen*) bei Personen mit einem hohen Lesefähigkeitsniveau (z. B. Studierenden) vorgenommen. In Studie 1 ( $N = 116$ ) ergaben sich für die Skalen des Instruments zufriedenstellende interne Konsistenzen und eine erwartungskonforme Faktorenstruktur. In Studie 2 ( $N = 35$ ) konnte für 2 von 3 inhaltlichen Thematiken gezeigt werden, daß sich die Vorhersage der Bewältigung klassischer Leseverständnistests verbessern läßt, wenn nicht nur das bereichsspezifische Vorwissen, sondern auch die Testwerte von PL-Lesen als Prädiktoren herangezogen werden.

## **Lesefähigkeitsdiagnostik bei Erwachsenen**

Für den deutschsprachigen Raum liegen zwar etliche Instrumente zur Diagnose von Lesefähigkeiten bei Schulkindern vor (z. B. WLLP, Küspert & Schneider, 1998), es fehlt jedoch bislang ein Diagnostikum, das zwischen Erwachsenen mit einem relativ hohen Niveau an Lesefähigkeiten (z. B. Studierenden) differenzieren könnte. Ein solches (Forschungs-) Instrument kann vor allem in psycholinguistischen und pädagogisch-psychologischen Untersuchungen zum Textverstehen bzw. Lernen mit Texten sinnvoll eingesetzt werden (vgl. z. B. Christmann, Groeben, Flender, Naumann & Richter, 1999).

### *Warum ein neues Instrument zur Erfassung von Lesefähigkeiten?*

Im angelsächsischen Bereich gibt es eine jahrzehntelange Tradition standardisierter Leseverständnistests. Verfahren wie der Nelson Denny Reading Test oder Verbal Scholastic Aptitude Test werden an US-amerikanischen Colleges oft routinemäßig zur Vorhersage akademischer Leistungen eingesetzt und verfügen teilweise über eine bemerkenswerte prädiktive Validität (vgl. Flippo & Schumm, 2000). Kognitionspsychologisch betrachtet ist die Konstruktvalidität dieser standardisierten Tests allerdings völlig unklar, da an der Beantwortung der typischen Aufgaben (z. B. Multiple-Choice-Verständnisfragen im Anschluß an die Textlektüre) eine Reihe sehr heterogener Prozesse beteiligt sind, die in der Messung konfundiert sind. Die kognitionspsychologische Forschung zu Lesefähigkeiten bemüht sich daher seit den 70er Jahren darum, Unterschiede zwischen guten und schlechten Lesern/innen auf isolierte Prozeß- und Strukturcharakteristika zurückzuführen und damit

---

<sup>1</sup> Poster präsentiert auf dem 42. Kongreß der Deutschen Gesellschaft für Psychologie, 24.- 29. September 2000, Jena.

<sup>2</sup> Universität zu Köln. Die hier vorgestellten Untersuchungen wurde im Rahmen des DFG-Forschungsprojekts "Verarbeitungsstrategien von traditionellen (linearen) Buchtexten und zukünftigen (nichtlinearen) Hypertexten (PD Dr. U. Christmann, Heidelberg, Prof. Dr. N. Groeben, Köln) durchgeführt. Wir danken der DFG für ihre Unterstützung. Wir danken weiterhin Frau Julia Herfordt für ihre Unterstützung bei der Datenerhebung.

kognitionspsychologisch beschreibbar zu machen (Daneman, 1982; Perfetti, 1985). Die in diesem Zusammenhang entwickelten Maße, z. B. die Lesespanne als dynamisches Maß der Arbeitsgedächtniskapazität (Just & Carpenter, 1992), erfassen jedoch nur sehr eng umgrenzte Aspekte von Lesefähigkeit. Insbesondere können sie nicht dem Umstand Rechnung tragen, daß Defizite in bestimmten Teilprozessen des Lesens durch die erfolgreiche Bewältigung anderer Teilprozesse, in der Regel auf höheren Ebenen des Textverstehens, ausgeglichen werden können (für einen Überblick s. Kintsch, 1998, S. 282-289).

*Das Instrument "Print Literacy - Lesen" (PL-Lesen)*

*Aufbau.* Das von uns entwickelte computergestützte Instrument PL-Lesen orientiert sich vor diesem Hintergrund an einem hierarchischen Modell des Textverstehens, dem Strategiemodell von van Dijk und Kintsch (1983). Es enthält sieben Subtests, mit denen sowohl eine detaillierte als auch kognitionspsychologisch transparente Erfassung von Teilfähigkeiten des Lesens im Sinne der bei van Dijk und Kintsch (1983) unterschiedenen Strategietypen möglich sein soll (vgl. Abbildung 1). Vier der Subtests sind hierarchieniedrigen Prozessen (propositionale Strategien und Strategien der lokalen Kohärenzbildung) zugeordnet, die drei übrigen sollen die Bewältigung hierarchiehoher Prozesse (vor allem makropropositionale und rhetorische Strategien) erfassen (s. ausführlicher Richter & Naumann, 2000).

	Subtests	Aufgaben- und Itemtyp. Antwortkategorien	Zentrale Prozesse (van Dijk & Kintsch, 1983)
hierarchieniedrige Prozesse	Satzverifikation	Begriffliche Aussagen ( <i>wahr / falsch</i> ) "Erdbeeren sind eine rote süß schmeckende Gemüsesorte."	propositionale Strategien
	Sinnhaftigkeit (Sätze)	Aussagen ( <i>sinnvoll / sinnlos</i> ) "Wenn man verdorbene Speisen isst, verfärbt sich die Haut oft klingvoll."	propositionale Strategien lokale Kohärenzbildung
	Wortschatz	Bedeutungserklärungen ( <i>korrekt / falsch</i> ) "Der Ausdruck 'Determinante' bedeutet unter anderem 'bestimmender Faktor.'"	propositionale Strategien
	Sinnhaftigkeit (Satzfolgen)	Satzfolgen ( <i>sinnvoll / sinnlos</i> ) "Dietrich Dörner hat sich um die Problemlöseforschung verdient gemacht." ( <i>Erster Satz</i> ) "Dafür wurde sie mit dem angesehenen Leibnizpreis ausgezeichnet." ( <i>Zweiter Satz</i> )	lokale Kohärenzbildung
hierarchiehohe Prozesse	Textverständnis	Beurteilung von Aussagen nach Textlektüre ( <i>imText enthalten / nicht enthalten</i> )	makropropositionale Strategien
	Fakten/Meinungen	Unterscheiden von Fakten und Meinungen nach Textlektüre ( <i>Tatsache / Meinung, 5-stufige Skala</i> )	makropropositionale Strategien rhetorische Strategien
	Textimplikationen	Beurteilung möglicher Implikationen nach Textlektüre ( <i>Implikation / keine Implikation</i> )	makropropositionale Strategien rhetorische Strategien

Abbildung 1: Subtests von *PL-Lesen* mit Aufgabentypen und Zuordnungen zu den Prozesstypen des Strategiemodells

*Bearbeitung.* Alle Aufgaben mit der Ausnahme eines Subtests (Bearbeitung mit der Maus) sind mit zwei Antworttasten zu beantworten. Die Bearbeitung nimmt insgesamt etwa eine halbe Stunde in Anspruch.

*Testwerte.* Registriert werden Akkuratheit, Entscheidungszeiten und Lesezeiten, die zu integrierten Testwerten kombiniert werden. Wir nehmen an, daß die verschiedenen Teilprozesse des Lesens als effizient angesehen werden können, wenn sie sowohl zu einem korrekten Ergebnis führen als auch schnell bewältigt werden. Daher wird in vier Subtests ein Item als gelöst bewertet, wenn die Aufgabe korrekt und innerhalb der unteren beiden Quartile der Item-spezifischen Entscheidungszeitverteilung bearbeitet wurde; die in diesen Studien verwendeten "Normen" wurden anhand einer separaten Stichprobe ( $N = 36$ ) ermittelt. Für die drei Subtests, die eine Lektüre von Texten erfordern, wird der Median des Produkts von Entscheidungszeiten  $\times$  Lesezeiten als Cut-off verwendet.

### *Validierung von PL-Lesen*

In den hier vorgestellten Studien wurde PL-Lesen erprobt und verschiedene Validitätsaspekte untersucht.

*Studie I.* In Studie I standen Aspekte der Konstruktvalidität im Vordergrund. Gemäß der Unterscheidung von hierarchieniedrigen und hierarchiehohen Prozessen des Strategiemodells sollten Subtests, die auf dieselbe Prozeßklasse bezogen sind, untereinander höher korreliert sein als mit Subtests der jeweils anderen Prozeßklasse. Zusätzlich wurden Zusammenhänge mit Selbsteinschätzungsdaten der eigenen Lesefähigkeiten ermittelt.

*Studie II.* In Studie II wurde im Sinne einer Kriteriumsvalidierung des Instruments geprüft, ob die Subtests von PL-Lesen Varianz in (Paper-Pencil-)Leseverständnistests aus drei verschiedenen Inhaltsbereichen aufklären können, auch wenn das bereichsspezifische Vorwissen auspartialisiert wird.

## **Methode**

### *Stichproben*

*Studie I.* An Studie I nahmen 116 Studierende (74 Frauen und 42 Männer) der Universität zu Köln mit einem mittleren Alter von 25 Jahren ( $SD = 4.6$ ) und einer mittleren Semesterzahl von 4.6 ( $SD = 2.2$ ) teil. In der Mehrzahl handelte es sich um Studierende der Psychologie (70%), die entweder für die Bescheinigung von Versuchspersonenstunden oder für ein geringes Entgelt an der Untersuchung teilnahmen.

*Studie II.* An Studie II nahmen 35 Studierende (20 Frauen und 15 Männer) der Universität zu Köln (70% Studierende der Psychologie) mit einem mittleren Alter von 27 Jahren ( $SD = 6.8$ ) und einer mittleren Semesterzahl von 5.1 ( $SD = 2.7$ ) teil. Für die Teilnahme wurden Versuchspersonenstunden bescheinigt; zusätzlich wurden unter den Teilnehmern/innen drei Geldpreise verlost.

## Durchführung

*Studie I.* PL-Lesen wurde in Einzelsitzungen dargeboten. Im Anschluß an die Darbietung des Tests am Computer wurde ca. der Hälfte der Stichprobe (66 Pbn) ein Fragebogen ausgehändigt, der neben Informationen zu tatsächlichen Leseaktivitäten über eine kurze Skala (6 Items in Aussageform, 5-stufige Ratingskalen) Selbsteinschätzungen der Bewältigung verschiedener Leseanforderungen erfragt. Die Pbn wurden ausführlich über den Untersuchungszweck informiert.

*Studie II.* PL-Lesen wurde in Einzel- oder Zweiersitzungen dargeboten. Zusätzlich bearbeiteten alle Pbn drei Leseverständnistests aus verschiedenen Inhaltsbereichen: Psychologie ("Signalentdeckungstheorie"), Literaturwissenschaft ("Weimarer Klassik") und Medizin/Pharmazie ("Pflanzliche Arzneimittel"). Bei den Leseverständnistests war unter der Zeitvorgabe von 20 Minuten jeweils ein Sachtext zu lesen und ein Paper-Pencil-Leseverständnistest mit 14-16 Items zu bearbeiten. In zwei Teilen wurden zunächst textnah formulierte Aussagen vorgegeben, für die zu entscheiden war, ob sie im Text enthalten waren oder nicht ("Rekognitionsfragen"), sowie Aussagen, für die zu entscheiden war, ob sie Implikationen des Texts darstellen oder nicht ("Inferenzfragen"); beide Fragetypen wurden aus psychometrischen Gründen zu je einer Skala zusammengefaßt. Für alle drei Texte wurde zudem verstehensrelevantes Vorwissen (Kenntnis zentraler Begriffe) mit je 6-8 Items mit offenem Antwortformat erhoben. Die Reihenfolge von PL-Lesen und Vorwissenstests/Leseverständnistests wurde ebenso wie die Reihenfolge der Leseverständnistests randomisiert. Die Pbn wurden ausführlich über den Untersuchungszweck informiert.

## Ergebnisse von Studie 1

### Reliabilitäten

Tabelle 1: Mittelwerte für Akkuratheit, Latenz und integrierte Testwerte sowie Reliabilitäten (Cronbach's  $\alpha$ ) der integrierten Testwerte

Subtest	$k_{\text{Items}}$	Akkuratheit	Latenz [ms]	Integrierte Testwerte	$\alpha$
		$M (SD)$	$M (SD)$	$M (SD)$	
Satzverifikation	15	14.04 (1.13)	2879 (911)	8.22 (4.46)	.87
Sinnhaftigkeit (Sätze)	15	14.47 (1.03)	3907 (1286)	8.02 (4.60)	.88
Wortschatz	15	11.97 (1.60)	4853 (1422)	6.94 (3.11)	.68
Sinnhaftigkeit (Satzfolgen)	14	13.42 (0.89)	3287 (1291)	8.31 (3.92)	.85
Textverständnis	14	11.45 (1.49)	6732 (2030)	5.02 (3.69)	.84
Fakten/Meinungen	18	12.75 (1.85)	10335 (5311)	8.49 (6.64)	.89
Textimplikationen	10	6.58 (2.05)	6252 (2377)	3.17 (2.89)	.83

Wie Tabelle 1 entnommen werden kann, werden insbesondere die hierarchieniedrigen Subtests (mit Ausnahme des Wortschatztests) in der Regel sehr akkurat bearbeitet, während die Entscheidungszeiten bei allen Subtests stark streuen. Die internen Konsistenzen der integrierten Testwerte sind zufriedenstellend bis auf den Wortschatztest, der eine höchstens akzeptable Reliabilität aufweist.

### Skaleninterkorrelationen

Die Skaleninterkorrelationen in Tabelle 2 zeigen ein deutliches Muster: Hierarchiehohe und -niedrige sind untereinander jeweils substantiell bis hoch korreliert; Korrelationen zwischen hierarchiehohen und -niedrigen Tests finden sich dagegen nur in 9 von 12 Fällen und sind durchweg geringer.

Tabelle 2: Skaleninterkorrelationen und Faktorladungen (Hauptkomponentenanalyse, Varimax-Rotation)

	SV	Si I	WS	Si II	TV	FM	FI	F II
Satzverifikation (SV)							<b>.86</b>	.13
Sinnhaftigkeit Sätze (Si I)	.69**						<b>.86</b>	.13
Wortschatz (WS)	.62**	.70**					<b>.82</b>	.25
Sinnhaftigkeit Satzfolgen (Si II)	.61**	.61**	.57**				<b>.79</b>	.08
Textverständnis (TV)	.25**	.30**	.37**	.22*			.26	<b>.70</b>
Fakten/Meinungen (FM)	.24*	.19*	.17	.20*	.49**		.11	<b>.87</b>
Textimplikationen (TI)	.14	.22*	.30**	.12	.46**	.70**	.06	<b>.89</b>

Anmerkung. FI: Erster Faktor, F II: Zweiter Faktor.

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

Eine Hauptkomponentenanalyse (Abbruchkriterium: Kaiser-Kriterium) resultiert in zwei Faktoren (Varianzaufklärung 71%), die nach Varimax-Rotation klar interpretierbar sind: Der erste Faktor (Varianzaufklärung 40%) faßt die hierarchieniedrigen, der zweite Faktor (Varianzaufklärung 31%) die hierarchiehohen Subtests zusammen (Tabelle 2).

### Korrelationen mit Selbsteinschätzungsdaten

Zwei der vier hierarchieniedrigen Subtests, Wortschatztest und Sinnhaftigkeit (Satzfolgen) sind moderat ( $r = .21$ ) mit der Selbsteinschätzungsskala zu Lesefähigkeiten korreliert (signifikant bei  $\alpha = .05$ , einseitig;  $n = 66$ ). Für zwei der drei hierarchiehohen Subtests, Textverständnis und Fakten/Meinungen, ergeben sich etwas höhere Korrelationen ( $r = .31$  bzw.  $r = .33$ ) mit den Selbsteinschätzungsdaten (signifikant bei  $\alpha = .01$ , einseitig). Faktor 2, der die hierarchiehohen Tests bündelt, ist in Höhe von .36 mit den Selbsteinschätzungsdaten korreliert (signifikant bei  $\alpha = .01$ , einseitig). Die Subtests Satzverifikation, Sinnhaftigkeit (Sätze) sowie Textimplikationen sowie Faktor 2, der die hierarchieniedrigen Subtests zusammenfaßt, weisen keine Zusammenhänge mit den Selbsteinschätzungsdaten auf.

## Ergebnisse von Studie II

Den offenen Antworten der drei Vorwissenstests wurden von zwei Kodierern/innen nach den Kriterien Korrektheit und Verstehensrelevanz Punktwerte von 0-2 zugewiesen (Intercoder-Reliabilität  $\kappa = .70$ ; Spearman's  $\rho = .80$ , bestimmt anhand von  $3 \times 3$  Vorwissenstests). Bei abweichenden Kodierungen wurden die Urteile einer dritten Kodiererin zur Entscheidung herangezogen. Die internen Konsistenzen (Cronbach's  $\alpha$ ) liegen für die Vorwissenstests im Bereich von .71 bis .81, für die Leseverständnistests im Bereich von .74 bis .80.

Den Interkorrelationen von Vorwissenstests und Lesetests lassen sich Hinweise auf die Validität der vorgenommenen Operationalisierungen entnehmen: So haben z. B. Vorwissens- und Leseverständnistests, die derselben Thematik zugehörig sind, jeweils die höchsten paarweisen Zusammenhänge (von .34 medizinisch-pharmazeutischen Thematik bis .47 bei der psychologischen Thematik, alle  $p < .05$ ). Bei den Vorwissenstests sind lediglich der Vorwissenstests zur psychologischen Thematik und der Vorwissenstests zur medizinisch-pharmazeutischen Thematik miteinander korreliert ( $r = .46$ ,  $p < .01$ ), bei den Lesetests korreliert der Lesetest zur psychologischen Thematik zu .50 mit dem Lesetest zur literaturwissenschaftlichen Thematik ( $p < .01$ ) und zu .29 mit dem Lesetest zur medizinisch-pharmazeutischen Thematik ( $p < .05$ ).

### *Multiple Zusammenhänge der Leseverständnistests mit den Subtests von PL-Lesen (bei Ausparialisierung des Vorwissens)*

Wie sich Tabelle 3 entnehmen läßt, tragen für die psychologische Thematik (Testtext zur Signalentdeckungstheorie) weder die hierarchieniedrigen noch die hierarchiehohen Subtests zur Varianzaufklärung der Leseverständnis-Testwerte bei; lediglich das Vorwissen erweist sich als guter Prädiktor.

Tabelle 3: Multiple Regressionen des Vorwissens und der Lesefähigkeits-Faktoren auf das Leseverständnis für die psychologische Thematik (Signalentdeckungstheorie)

Prädiktor	Unvollständige Modelle				Vollständiges Modell	
	$\beta$	$t$	$\beta$	$t$	$\beta$	$t$
SDT-Vorwissen	.34	2.8**	.46	2.9**	.44	2.8**
Hierarchieniedrige Tests	.03	0.9		--	.18	-0.3
Hierarchiehohe Tests		--	.06	0.4	-.05	0.4
Modellgüte	$R = .50$		$R = .48$		$R = .50$	
	$R^2 = .25$		$R^2 = .23$		$R^2 = .25$	
	$R^2_{\text{adjust}} = .20$		$R^2_{\text{adjust}} = .18$		$R^2_{\text{adjust}} = .18$	
	$F(2, 32) = 5.2**$		$F(2, 32) = 4.7**$		$F(3, 31) = 3.4*$	

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

Bei den in Tabelle 4 wiedergegebenen Regressionsanalysen für die literaturwissenschaftliche Thematik (Testtext zur Weimarer Klassik) läßt sich durch Einbeziehung der hierarchie-

niedrigen Subtests eine Verbesserung der Vorhersage der Leseverständnis-Testwerte erzielen ( $\beta = .32$ ,  $t = 2.0$ ,  $p < .10$ ; unvollständiges Modell), während die hierarchiehohe Subtests keine Verbesserung erbringen.

Tabelle 4: Multiple Regressionen des Vorwissens und der Lesefähigkeits-Faktoren auf das Leseverständnis für die literaturwissenschaftliche Thematik (Weimarer Klassik)

Prädiktor	Unvollständige Modelle				Vollständiges Modell	
	$\beta$	$t$	$\beta$	$t$	$\beta$	$t$
Klassik-Vorwissen	.30	1.8 <sup>+</sup>	.37	2.2*	.31	1.9 <sup>+</sup>
Hierarchieniedrige Tests	.32	2.0 <sup>+</sup>		--	.39	1.9 <sup>+</sup>
Hierarchiehohe Tests		--	.12	0.7	-.11	-0.6
Modellgüte	$R = .51$		$R = .42$		$R = .52$	
	$R^2 = .26$		$R^2 = .18$		$R^2 = .27$	
	$R^2_{\text{adjust}} = .21$		$R^2_{\text{adjust}} = .13$		$R^2_{\text{adjust}} = .19$	
	$F(2, 32) = 5.6^{**}$		$F(2, 32) = 3.5^{**}$		$F(3, 31) = 3.7^{**}$	

<sup>+</sup>  $p < .10$ , \*  $p < .05$ , \*\*  $p < .01$ .

Lediglich bei den in Tabelle 5 wiedergegebenen Regressionsanalysen für die medizinisch-pharmazeutische Thematik (Testtext zu pflanzlichen Arzneimitteln) erweisen sich sowohl hierarchieniedrige als auch hierarchiehohe Subtests als gute Prädiktoren: Im vollständigen Modell tragen neben dem bereichsspezifischen Vorwissen ( $\beta = .37$ ,  $t = 2.8$ ,  $p < .01$ ) auch die hierarchieniedrigen Subtests ( $\beta = .38$ ,  $t = 2.3$ ,  $p < .05$ ) und – mit annähernd doppeltem Beta-Gewicht – die hierarchiehohe Subtests ( $\beta = -.76$ ,  $t = -4.5$ ,  $p < .001$ ) zur Varianzaufklärung ( $R^2_{\text{adjust}} = .42$ ) bei. Überraschend ist dabei allerdings der negative Einfluß der hierarchiehohe Subtests.

Tabelle 5: Multiple Regressionen des Vorwissens und der Lesefähigkeits-Faktoren auf das Leseverständnis für die medizinisch-pharmazeutische Thematik (pflanzliche Arzneimittel)

Prädiktor	Unvollständige Modelle				Vollständiges Modell	
	$\beta$	$t$	$\beta$	$t$	$\beta$	$t$
Arznei-Vorwissen	.36	2.2*	.41	2.9*	.37	2.8**
Hierarchieniedrige Tests	-.10	-0.6		--	.38	2.3*
Hierarchiehohe Tests		--	-.52	-3.7**	-.76	-4.5***
Modellgüte	$R = .36$		$R = .62$		$R = .69$	
	$R^2 = .13$		$R^2 = .39$		$R^2 = .47$	
	$R^2_{\text{adjust}} = .07$		$R^2_{\text{adjust}} = .35$		$R^2_{\text{adjust}} = .42$	
	$F(2, 32) = 2.3$		$F(2, 32) = 10.0^{***}$		$F(3, 31) = 9.2^{***}$	

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

## Diskussion

Die Ergebnisse von Studie I liefern ermutigende Resultate in bezug auf die Reliabilitäten (interne Konsistenzen) sowie die faktorielle Validität des entwickelten Instruments. Die Ergebnisse von Studie II sind differenzierter zu betrachten. Zunächst scheinen es die hierarchieniedrigen Lesefähigkeits-Komponenten zu sein, die zusätzlich zum thematischen Vorwissen zur Erklärung der Kriteriumsvarianz beitragen.

Die hierarchiehohen Lesefähigkeits-Komponenten kommen dagegen nur für die medizinisch-pharmazeutische Thematik zum Tragen, und hier mit einem hypothesendiskonformen Vorzeichen: Je besser die Probanden die hierarchiehohen Subtests bewältigen, umso schlechter bewältigen sie den Leseverständnistest für das pharmazeutische Thema (und umgekehrt). Eine mögliche post-hoc-Erklärung für dieses Resultat könnte sein, daß gerade für den pharmazeutischen Subtest in hohem Maße eine oberflächliche Strategie ("Einlesen" von Textinformationen) funktional war, für die die von den hierarchiehohen Tests abgedeckten Teilfähigkeiten eher störend wirkten. Schon wegen der geringen Stichprobengröße in Studie II sind weitere Validierungen vorgesehen.

## Literatur

- Christmann, U., Groeben, N., Flender, J., Naumann, J. & Richter, T. (1999). Verarbeitungsstrategien von traditionellen (linearen) Buchtexten und zukünftigen (nicht-linearen) Hypertexten. In N. Groeben (Hrsg.), *Lesesozialisation in der Mediengesellschaft* (S. 175-189). Tübingen: Niemeyer.
- Daneman, M. (1982). The measurement of reading comprehension: How not to trade construct validity for predictive power. *Intelligence*, 6, 331-345.
- Flippo, R. F. & Schumm, J. F. (2000). Reading tests. In R. F. Flippo & D. C. Caverley (Eds.), *Handbook of college reading and study strategy research* (pp. 403-472). Mahwah, NJ: Erlbaum.
- Just, M. A. & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99, 122-149.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.
- Küspert, P. & Schneider, W. (1998). WLLP: *Würzburger Leise Leseprobe*. Göttingen: Testzentrale.
- Perfetti, C. A. (1985). *Reading ability*. New York, NY: Oxford University Press.



Richter, T. & Naumann, J. (2000). Computer-based assessment of reading skills. *Proceedings of the 2<sup>nd</sup> Computers in Psychology Conference (CiP 2000), March 29-31, 2000, York* [WWW document]. Available URL: [http://cti-psy.york.ac.uk/cip2000-online/papers/fra\\_paper4.htm](http://cti-psy.york.ac.uk/cip2000-online/papers/fra_paper4.htm)

van Dijk, T. A. & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York, NY: Academic Press.

Johannes Naumann  
Universität zu Köln  
Psychologisches Institut  
Lehrstuhl Allgemeine Psychologie  
Herbert-Lewin-Str. 2  
50931 Köln

Tel.: 0221/470-3848

Fax: 0221/470-5002

E-mail: [johannes.naumann@uni-koeln.de](mailto:johannes.naumann@uni-koeln.de)