

Prof. Dr. Eduard W. Kleber und Gerda Kleber

Differentieller Leistungs-Test – KE (DL-KE)

Mappe mit Handanweisung und Testmaterial 46,50 DM

Der differentielle Leistungstest – KE erfaßt Variablen des Leistungsverhaltens bei konzentrierter Tätigkeit. Die Variablen für den quantitativen und den qualitativen Leistungsaspekt sowie für Gleichmäßigkeit der Leistung stellen bedeutsame Prediktoren für Schulleistung und wichtige Persönlichkeitsvariablen dar. Sie können für eine valide Interpretation dann erfaßt werden, wenn die Probanden auf einem angemessenen Konzentrationsniveau belastet werden.

Der differentielle Leistungstest – KE stellt für die Eingangsstufe der Grundschule Bearbeitungsmaterial auf einem angemessenen Konzentrationsniveau zur Verfügung. Er ermöglicht Aussagen über das Leistungsverhalten von Schulanfängern bei konzentrierter Tätigkeit, läßt problematische Leistungsstörungen frühzeitig erkennen und gibt Hinweise für differentielle Förderung für Schulanfänger mit verschiedenen allgemeinen Leistungsstörungen.

Prof. Dr. Eduard W. Kleber, Gerda Kleber und Olaf Hans

Differentieller Leistungs-Test – KG (DL-KG)

Mappe mit Handanweisung und Testmaterial 58,- DM

Der Differentielle Leistungstest-KG schließt an die Tradition der Durchstreichtests von Bourdon und anderen an. Er bietet an Bearbeitungsmaterial, das für über 95% der Zielgruppen ein angemessenes (nicht zu schwieriges) Bearbeitungsniveau darstellt. Er ist die Ergänzung des DL-KE für die Altersgruppen der Grundschule. Der DL-KG ist kein Konzentrationstest im herkömmlichen Sinne. Er liefert keinen Testwert als Maß für Konzentrationsfähigkeit, sondern drei Maßwerte für verschiedene Leistungsvariablen; darüber hinaus ist er als standardisierte Beobachtungssituation der Arbeitshaltung bei konzentrierter Tätigkeit und fremdgesetzter Aufgabenstellung geeignet. Es handelt sich hier um ein Verfahren, das die Analyse des Leistungs- und Arbeitsverhaltens bei konzentrierter Tätigkeit ermöglicht. Da diese Fähigkeit im schulischen Lernen eine wichtige Rolle spielt und das Arbeitsverhalten in erheblichem Maße den Schulerfolg mitbestimmt, ist er geeignet, Problemfälle früh zu erkennen und für sie differentielle pädagogische Maßnahmen einzuleiten.

Verlag für Psychologie · Dr. C. J. Hogrefe

Göttingen · Toronto · Zürich

3400 Göttingen · Rohnsweg 25 · Postfach 414

Unterrichtswissenschaft

Thema: Texte und Lernen

Norbert Groeben

Verstehen, Behalten, Interesse

Übereinstimmende Antworten und kontroverse Fragen zur Beziehung von Textstruktur, Textverständnis und Lerneffekt

Eine Zwischenbilanz der Forschungsbemühungen zum Problem „Textverständlichkeit“ zeigt, daß es (relativ) gesicherte Ergebnisse vor allem im Bereich der Dimensionen der Textstruktur, der Verständlichkeitsmessung und der textuellen Bedingungen für den kognitiven Lernerfolg gibt. Offen sind allerdings die meisten Fragen zur motivationspsychologischen Perspektive der Textgestaltung; hier ist das größte Problemlösungspotential von einer Verbindung der epistemischen Neugiertheorie (*Berlyne*) mit dem Modell des mathemagenen Verhaltens (*Rothkopf*) zu erwarten.

Comprehension, Retention, Motivation (Interest)

Concurring Answers and Controversial Questions on the Relationship between Structure and Comprehension of Texts and Efficiency of Learning
Interim results of research projects into the problem of “understandability of texts” show that some (relatively) safe results have been obtained, mainly regarding the dimensions of textual structure, testing of reading comprehension and the textual conditions of learning efficiency. Most questions on text formation under the perspective of motivation psychology remain open, however; in this field the greatest potential of solvable problems is to be expected from a combination of the epistemic theory of curiosity (*Berlyne*) and the concept of mathemagenic activities (*Rothkopf*).

1. Forschungsprogramm Textverständlichkeit: Ausgangspunkte

Die Frage nach der (kognitiven wie motivationalen) Wirksamkeit von Texten ist an sich ein Glücksfall für die Psychologie: denn die Beantwortung setzt die Nutzung einer Fülle psychologischer Theorien (psycholinguistische, kognitions-, motivations-, kommunikationspsychologische) voraus und drängt gleichzeitig geradezu nach einer möglichst direkten Anwendung in der (pädagogischen/didaktischen) Praxis. Diese Chance der Verbindung von Theorie- und Praxisrelevanz der Forschung ist sicherlich auch ein Grund dafür, daß diese instruktionspsychologische Forschungsperspektive in der amerikanischen (pädagogischen) Psychologie seit langer Zeit eine hervorragende Rolle spielt (vgl. *Hofer* in diesem Heft). Die parallele Forschung im deutschsprachigen Raum kann daraus den Vorteil ziehen, von vornherein auf einer breiten Grundlage methodologischer und theoretischer Modelle aufzubauen. Trotz der relativ kurzen Dauer dieser Forschungsbemühungen (vgl. wiederum *Hofer* in diesem Heft) lassen sich daher unter Einbeziehung der wichtigsten (übertragbaren) anglo-amerikanischen Ansätze doch eine ganze Menge gesicherter Ergebnisse feststellen – wenngleich auch das Ausmaß der offenen Fragen (wie bei jedem noch nicht degenerativen Forschungsprogramm) die Anzahl der Antworten übersteigt. Die Zusammenschau, welche Ergebnisse als (vorläufig) gesichert gelten können und welche Fragen als kontrovers noch offen sind, erfüllt daher zwei Funktionen: dem Praktiker Hinweise/Vorschläge zu geben, an welchen Stellen er möglichst valide (mehrfach abgesicherte), in seine Praxis umsetzbare Erklärungen/Technologien finden kann bzw. an welchen Punkten er sich (bisher) noch auf seine subjektive Instruktionsintuition verlassen muß; sowie dem Forscher (möglichst begründete) Anregungen zu geben, welche Probleme innerhalb des

thematischen Forschungsprogramms (dringend) als nächste anzugehen sind. Der Versuch, auf diese Weise innezuhalten und (auch sich selbst) Rechenschaft abzulegen über die derzeitige Problemlandschaft des Forschungsprogramms „Textverständlichkeit“, kann im Rahmen eines kürzeren Aufsatzes naturgemäß nur auf der relativ abstrakten Ebene der inklusivsten Konzepte (im Sinne *Ausubels* 1963) ablaufen – und soll zur Steigerung des Problemlösungspotentials die offenen Fragen differenzierter behandeln als die vorläufig akzeptablen Antworten.

(Advanced organizer:) Ich will bei diesem Versuch von der „readability“-Forschung ausgehen, deren objektivistischer Ansatz zwei Beschränkungen aufweist: die Konzentration nur auf Stilvariablen von Texten und die mangelnde Berücksichtigung der Rezeptionsprozesse beim Leser. Die Überwindung dieser Beschränkungen im Forschungsprogramm „Textverständlichkeit“ hat übereinstimmend auf mehr induktivem wie auf deduktivem Weg zu vier Dimensionen der Textstruktur geführt (2.), die neben der Stilistikebene auch noch die Ebenen der semantischen Redundanz, der kognitiven Gliederung und der konzeptuell-konflikthaften Stimulanz umfassen. Die Berücksichtigung der Rezipientenprozesse hat eine Explizierung des Konstrukts „Verständlichkeit“ (3.) als zwischen Textstruktur und Lernerfolg vermittelndes nötig gemacht; die Operationalisierung, d. h. das adäquate Meßverfahren stellt dabei das größte Problem dar: nach der Kritik des bisher am häufigsten verwandten Verfahrens (Mehrfach-Wahl-Antworten) gibt man heute (weiterentwickelten) Verfahren der Cloze procedure bzw. Rateverfahren (zur Schätzung der subjektiven Information) den Vorzug. Auf Seiten der Lernkriterien hat der Behaltenserfolg (4.1.) das größte theoretische wie auch empirische Gewicht; uneinheitlich sind die Positionen noch zur Frage, ob die Textverständlichkeit unbegrenzt gesteigert werden sollte, gesichert scheint aber die Vorordnung des kognitiven Lernerfolgs vor den Motivierungseffekt zu sein. Der Motivationserfolg (4.2.) ist in sich und in seiner Beziehung zur Textverständlichkeit das bisher am wenigsten geklärte Teilgebiet; es sind zwei z. T. konkurrierende, z. T. sich überlappende Perspektiven zu unterscheiden: der Ansatz des mathemagenen Verhaltens (*Rothkopf*) und der der epistemischen Neugier (*Berlyne*), von deren Verbindung ich mir eine Lösung der noch offenen Fragen verspreche (4.3.). Bezüglich der Anwendung (5.) der Forschungsergebnisse dürfte eine gegenseitige Ergänzung des zentral praxisorientierten Trainingsaspekts und der mehr erklärungsorientierten Forschungsperspektive zur größten Praxisrelevanz führen.

Der Ansatz der „Readability“-Forschung sucht nach generellen stilistischen Merkmalen, die – weitgehend unabhängig vom Rezipienten – den Grad der Lesbarkeit eines Textes bestimmen. Möglichst viele solcher Merkmale werden mit entsprechenden „Lesbarkeitsmaßen“ (Lesegeschwindigkeit, Expertenurteile, „listenability“, „Verstehen“ qua Behalten usw.) korreliert und die dabei gewichtigsten durch eine Regressionsanalyse in einer „Lesbarkeits-Formel“ (readability-formulas) zusammengestellt. Es gibt mittlerweile eine Fülle solcher Formeln, die *Klare* (1963) in seinem hervorragenden Forschungsbericht zusammengestellt hat – die bekanntesten/gebräuchlichsten sind wohl die von *Flesh* („reading ease score“), *Dale* u. *Chall* sowie *Farr*, *Jenkins* u. *Paterson* (vgl. *Klare* 1963, S. 75ff.). Wenn man die stilistischen Merkmale (die schon bei *Gray* u. *Leary*, 1935, fast vollständig aufgeführt sind) nach fast 40jähriger Forschung überblickt und faktorenanalytisch zusammenfaßt, so ergeben sich zwei zentrale Faktoren, von denen aus man auf die Lesbarkeit eines Textes schließen kann: Wortschwierigkeit und Satzschwierigkeit (*Klare* 1971, S. 244). Die Messung der abhängigen Variable (Lesbarkeit) konzentriert sich (ebenfalls nach *Klare* 1971, S. 247) auf drei Maße: Lesegeschwindigkeit, Expertenurteile (über die Textschwierigkeit) und „Textverständnis“ (meist multiple-choice-Behaltenstests). Damit kann der Ansatz schon von der methodologischen Struktur her nur zu bestimmten grundlegenden Beziehungen führen, die vom Psychologischen her z. T. unbefriedigend sind, wie z. B. daß die Lesbarkeit von Texten gesteigert wird durch kurze (Silbenzahl), gebräuchliche, bekannte Worte, kurze und einfache Sätze usw.; das psychologisch Unbefriedigende liegt dabei in zwei Beschränkungen: zum einen auf der Bedingungsseite die mangelnde Berücksichtigung von über die Stilistik hinausgehenden Faktoren wie „Inhalt, Organisation“ usw., worauf auch *Klare* (1971, S. 251) hinweist; zum anderen die Beschränkung auf der Zielseite, daß die Lesbarkeit ohne direkte Berücksichtigung der kognitiven Verarbeitungsprozesse (Lese-

geschwindigkeit, Rating) erhoben wird oder aber zu weit weg vom Rezeptionsprozeß (Behaltenstests: Gedächtnisfaktoren, s. Kritik unter 3.) operationalisiert ist. Abgesehen von den umgrenzten Ergebnissen zur Textstilistik fungiert die Readability-Forschung daher heute vor allem als Folie, auf deren Hintergrund die Anforderungen an ein konstruktives, umfassendes Forschungsprogramm zur Textverständlichkeit ableitbar sind.

2. Die Dimensionen der Textstruktur

Die erste Anforderung besteht also in der möglichst umfassenden *Feststellung* der relevanten Dimensionen der Textstruktur. Diese Anforderung läßt sich auf einem mehr *empirisch-induktiven* sowie auf einem akzentuiert theoriegeleitet-deduktiven Weg erfüllen. Den eher induktiven Weg hat die Hamburger Forschergruppe (Langer 1971; Schulz von Thun 1973; Langer et al. 1974) mit ihrem *Rating-Ansatz* beschritten. Dabei werden Texte (mit möglichst unterschiedlichem Schwierigkeitsgrad) von Experten mit Hilfe von polaren Merkmalen, z. B. gegliedert . . . ungegliedert, auf einer fünfstufigen Skala eingeschätzt (geratet). Diese Ratings werden dann faktorenanalysiert und so die Einzelmerkmale (empirisch-induktiv) zu bestimmten Merkmalsdimensionen zusammengefaßt. Auf diese Weise haben sich mehrfach folgende Dimensionen der Textstruktur ergeben (vgl. Langer et al. 1974, S. 13ff.): 1. Einfachheit (charakterisiert durch die Merkmale: einfache Darstellung; kurze, einfache Sätze; geläufige Wörter; Fachwörter erklärt; konkret; anschaulich); 2. Gliederungs-Ordnung (gegliedert; folgerichtig; übersichtlich; gute Unterscheidung von Wesentlichem und Unwesentlichem; der rote Faden bleibt sichtbar; alles kommt schön der Reihe nach); 3. Kürze – Prägnanz (zu kurz; aufs Wesentliche beschränkt; gedrängt; aufs Lehrziel konzentriert; knapp; jedes Wort ist notwendig); 4. Zusätzliche Stimulanz (anregend; interessant; abwechslungsreich; persönlich).

Den mehr *deduktiven Weg*, ausgehend von theoretischen Ansätzen, habe ich selbst zu beschreiben versucht (Groeben 1972); ich habe dabei psycholinguistische Theorien zur Satzgestaltung/Stilistik, informationstheoretische Modelle zur semantischen Dichte/Redundanz und in bezug auf die kognitiv-konzeptuelle Inhaltsstruktur die kognitive Lerntheorie (sensu Ausubel 1963) sowie die motivationale Neugiertheorie (Berlyne 1960) herangezogen. Es ergaben sich ebenfalls vier wichtige Perspektiven der Textstruktur: 1. Grammatikalisch-stilistische Einfachheit (ausgehend besonders von transformationsgrammatikalischen Ansätzen: kurze Satzteile, aktive Verben, aktiv-positive Formulierung, keine Nominalisierung, persönliche Worte, keine Satzschachtelung usw., vgl. Groeben 1972, S. 17ff.); 2. Semantische Dichte (keine Weitschweifigkeit, besonders keine synonyme oder wörtliche Wiederholung von Satzgliedern mit gewichtiger Bedeutung: Inhaltsworte als Subjekt, Prädikat, Objekt usw. S. 23ff.); 3. Kognitive Strukturierung (Merkmale nach Ausubels Subsumtionstheorie des Lernens: advanced organizer; lineare Gedankenführung; Hervorhebung wichtiger Konzepte; Zusammenfassungen; Beispielgebung; Unterschiede/Ähnlichkeiten von Begriffen, Konzepten herausarbeiten usw., vgl. S. 28ff.); 4. Motivierender konzeptueller Konflikt (Merkmale nach Berlynes Theorie der epistemischen Neugier: Probleme mit alternativen Lösungsmöglichkeiten; Gleichwahrscheinlichkeit von Alternativen; Neuheit, Überraschung; inkongruente Konzepte; Einfügung von Fragen usw., vgl. S. 38ff.). Diese Dimensionen hatten in der empirischen Überprüfung (Groeben 1972, S. 104ff. – mit Ausnahme des Faktors der semantischen Dichte) auch einen bedeutsamen Einfluß auf die Verständlichkeit entsprechend gestalteter/modifizierter Versionen eines Standardtextes (über Grup-

penarbeit; die Verständlichkeit wurde über ein Rateverfahren zur subjektiven Information erhoben, vgl. 3.).

Wie leicht zu erkennen ist, entsprechen sich die Ergebnisse des mehr induktiven sowie deduktiven Ansatzes sehr gut: die Merkmale der Dimension „*sprachliche Einfachheit*“ überlappen sich weitgehend, ähnliches gilt für die Dimension „*kognitive Gliederung – Ordnung*“ und „*Kürze – Prägnanz*“ (wobei die letztere praktisch mit der theoretischen Dimension der semantischen Dichte identisch ist). Auch die vierte – motivationsgerichtete – Dimension ist weitgehend parallel, allerdings liegen die Merkmale z. T. auf verschiedenen Ebenen: während „interessant“, „anregend“ usw. die „stimulierende“ Wirkung einfach benennen, können die aus der Berlyneschen Neugiertheorie abgeleiteten Merkmale wie Neuheit, Überraschung, Inkongruenz usw. u. U. diese beschriebenen Effekte erklären. Einerseits sind Merkmale mit Erklärungswert sicherlich theoretisch interessanter und wertvoller, andererseits aber muß beim gegenwärtigen Forschungsstand offen bleiben, ob die stimulierende Wirkung nicht auch noch durch andere Textmerkmale erreicht werden kann (vgl. 4.3.). Immerhin ist die *Übereinstimmung der Ergebnisse* insgesamt so gut, wie es in der empirischen Forschung nicht häufig geschieht; daher darf man davon ausgehen, daß es sich bei den genannten *vier Dimensionen* um die relevantesten und relativ *umfassenden Merkmalsbereiche der Textstruktur* handelt. Leichte Gegenläufigkeiten in den Ergebnissen gibt es nur hinsichtlich der Gewichtung der Dimensionen: während „*sprachliche Einfachheit*“ und „*Kürze – Prägnanz*“ in den Hamburger Untersuchungen bei der Unterscheidung von schwierigeren vs. leichteren Texten immer eine bedeutsame (signifikante) Rolle spielen (Langer et al. 1974, S. 47ff.) wirkte sich in meiner Untersuchung der Faktor der Kürze (Redundanz) überhaupt nicht und der der stilistischen Einfachheit quantitativ nur untergeordnet (4% Varianzaufklärung im Gegensatz zu 86% für den Faktor „*kognitive Gliederung*“; Groeben 1972, S. 105) aus. Allerdings muß man berücksichtigen (worauf Schulz von Thun 1973, S. 105b berechtigterweise hinweist), daß durch die Umformung eines Standardtextes schon vom Ansatz her nur eine sehr eingeschränkte Verschiedenheit (Varianz) der Texte (im sprachlichen Bereich) erreicht werden kann; es kommt daher hier den Hamburger Ergebnissen (die mit einer Stichprobe völlig unterschiedlicher Texte von großer Streubreite erzielt wurden) mit Sicherheit ein sehr viel größerer Verallgemeinerungswert zu. Parallel allerdings sind *in beiden Ansätzen die in sich uneinheitlichen Ergebnisse zur Wirksamkeit der motivationsgerichteten Textstrukturierung*, die für die Notwendigkeit differenzierterer Forschung in diesem Bereich sprechen (vgl. 4.2.).

3. Verständlichkeitsmessung

Auf der Grundlage der umfassenden Feststellung der Textdimensionen ist theoretisch konsequent auch die Beschränkung der Readability-Forschung *im Bereich der Zielgröße* zu überwinden: die *Ausweitung des Lesbarkeits- zum Verständlichkeitsaspekt*. Es kommt dann nicht mehr nur auf das Decodieren linguistischer Information (sozusagen Wort für Wort vom „Buchstabensinn“ bis hin zum Satzsinn) an – dieses wird vielmehr als eine für das Textverständnis notwendige Fähigkeit vorausgesetzt (Carroll 1972, S. 13; Groeben 1972, S. 69). Der zentrale Bedeutungsgehalt des Konstrukts Verständnis/Verständlichkeit liegt vielmehr in der Einordnung solcher (sprachlicher) Information in einen weiteren Kontext, d. i. sowohl im Hinblick auf die Textsemantik wie auf das Kognitionssystem des Rezipienten (Carroll 1972, S. 13ff.). Es besteht daher Einigkeit darüber, daß der Begriff der

Textverständlichkeit in einer Verbindungsfunktion zwischen Textstrukturen/-merkmalen und Lernerfolg theoretisch zu explizieren ist (vgl. Schulz von Thun 1973, S. 27f.). Daraus folgt (vgl. Carroll 1972, S. 5ff.) ausschließend, daß Textverstehen nicht in bloßem (echohaftem) Verbalisieren von Gelesenem bestehen kann, nur Kurzzeit-Gedächtnis-Prozesse umfassen kann (während das Involvieren des Langzeitgedächtnisses mehr das Lernkriterium des Behaltens Erfolgs thematisiert) und daß auch Inferenzprozesse über den vorliegenden Text (-abschnitt) hinaus („logical reasoning“) nicht mit Textverständnis gemeint sein können. Dies hat Konsequenzen für die Festlegung der adäquaten Meßmethoden (s. unten).

Die Frage nach der empirischen Validität dieser (deskriptiven) Konstruktexplikation wurde vor allem durch faktorenanalytische Untersuchungen der *Struktur der Fähigkeit „Textverständnis“* zu beantworten versucht. Dabei arbeitete Davis (1944) theoriegeleitet 9 Teilfähigkeiten (comprehension skills) heraus, die er z. T. (je nach Erhebung zwischen 4 und 7) in mehreren Untersuchungen als faktorenanalytische Dimensionen sichern konnte (ausführliches Überblicksreferat Davis 1972). Schon früh aber legte Thurstone (1946) für das gleiche Material (aus Davis 1944) eine Ein-Faktor-Lösung vor, die alle Davisschen Fähigkeitsdimensionen auf den Faktor „Kenntnis von Wortbedeutungen“ reduzierte. Diese unterschiedlichen Ergebnisse sind allerdings hauptsächlich Teil eines Methoden- und Modellstreits über die Faktorenanalyse. Die methodologischen Fragen und Probleme werden sehr differenziert und ausgewogen von Spearritt (1972) diskutiert und in seiner Re-analyse berücksichtigt, so daß man die durch ihn vorgelegten vier Faktoren als vorläufiges Endergebnis akzeptieren kann; danach ist die Fähigkeit „Textverständnis“ durch folgende Dimensionen bestimmt: 1. Schlußfolgerndes Denken während des Lesens (reasoning in reading); 2. Kenntnis von Wortbedeutungen; 3. Identifizierung von Intention, Einstellung, Stimmung des Textes/Autors; 4. Fähigkeit, der Textstruktur/-gliederung zu folgen (Spearritt 1972, S. 100f.). Die theoretische Bedeutungsfestlegung ist also insgesamt durch eine brauchbare (deskriptive) Konstruktvalidität abgesichert (z. B. wird ja das schlußfolgernde Denken auf den Lese-prozeß/-inhalt eingeschränkt).

Im praktischen Forschungsprozeß ist dies alles natürlich nur Vorarbeit für eine *adäquate Erfassung-/Meßmethode*. Dabei ergibt sich unter dem Aspekt der Textverständlichkeit die Schwierigkeit, daß – zumindest in der amerikanischen Forschung – zunächst immer die Überprüfung der individuellen Fähigkeit des Textverstehens im Vordergrund gestanden hat mit der technologischen Anwendungskonsequenz der Anpassung des Menschen an die Umwelt, hier den Text (durch z. B. Lesetraining usw., vgl. dazu Poulton 1961). Da die Erforschung der Textverständlichkeit aber gerade umgekehrt die Anpassung der Umwelt/ des Textes an den Menschen (Rezipienten) anstrebt, hat die Verständnismessung auch eine etwas modifizierte Funktion: es kommt nicht primär auf die interindividuellen Unterschiede an (vgl. aber in 4.1. die ATI-Perspektive), sondern es wird über die *durchschnittlichen Verständnismesswerte* einer möglichst großen Anzahl von Rezipienten (das Subjekt als Medium) auf die Verständlichkeit des Textes zurückgeschlossen. Das erfordert naturgemäß einer relativ großen empirischen Meßaufwand. Der Hamburger Rating-Ansatz versucht diesen zu vermeiden, indem er sich schlicht auf die Auskunft des Rezipienten über sein Textverstehen verläßt; diese Verbindung von Textmerkmalen und Rezeptionsprozeß (s. o. theoretische Bedeutungsexplikation) durch einen kognitiven Beurteilungsprozeß beim Rezipienten ist sicherlich sehr ökonomisch, allerdings auch sehr subjektiv. Diese Möglichkeit wird daher, wie aus der Aufstellung von Carroll (1972, S. 15ff.) hervorgeht, in der nach größtmöglicher Objektivität strebenden Forschung recht selten verwendet. Man sollte ver-

suchen, auch hochkomplexe Merkmale und ihre Auswirkungen auf den Rezeptionsprozeß möglichst exakt, differenziert und objektiv festzustellen, und sollte nicht zu früh aufgeben und diese Feststellung nur mehr dem „integrierten Urteil“ von Experten (Schulz von Thun 1973, S. 26) zuschreiben. Allerdings ist heute auch das am häufigsten verwendete Verfahren zur Messung des Textverständnisses: *Mehrfach-Wahl-Antworten*, die den Inhalt einer Textpassage abfragen, als *methodisch fragwürdig* nachgewiesen. Nicht nur, daß man unter dem Aspekt der theoretischen Explikation des Verständniskonstrukts hier zu viel Gedächtnisprozesse beteiligt sehen kann (und konsequenterweise diese Methode mehr für die Behaltensprüfung ansetzen müßte); auch die Reliabilität und Validität dieses Erhebungsverfahrens ist oft nicht befriedigend (vgl. Anderson 1972). Der Hauptgrund liegt darin, daß man die multiple-choice-Fragen z. T. auch ohne Verständnis, ja sogar ohne Kenntnis des zugrundeliegenden Textes richtig beantworten kann: zum einen sind die angebotenen Alternativen nicht immer gleichwahrscheinlich, so daß man durch allgemeines Kombinieren die richtige Antwort erraten kann (Pyrzack 1972, S. 64f.); zum anderen müssen die verschiedenen Items möglichst den gesamten Textinhalt abdecken, man kann daher durch In-Beziehung-Setzen der Items untereinander einen in sich kohärenten Textsinn konstruieren – so erreichten z. B. in einer Untersuchung von Weaver u. Bickley (1967) Versuchspersonen, ohne den thematischen Text je gesehen zu haben, 67% der richtigen Antworten der Textleser. Dies ist sicherlich ein Ausmaß von Inferenzprozessen, das nicht nur für die Operationalisierung des Verständlichkeitskonstrukts untragbar ist (sondern auch für die Einsetzung als Behaltentest; hier wären zumindest Trennschärfeuntersuchungen nötig, zu einem speziellen Trennschärfeindex für solche Mehrfach-Wahlantworten-Instrumente vgl. Pyrzack 1972, S. 66).

Es gibt aber Versionen von Rate- bzw. Satzergänzungsverfahren, die diese *methodischen Unzulänglichkeiten vermeiden*. Da ist zunächst die sehr einfallsreiche Methode des „*cunked comprehension test*“ nach Carver (1970) zu nennen, den auch Carroll (1972, S. 21f.) hervorhebt:

Dabei wird ein Text in Satzteile unterteilt, wobei für den Verständnistest ein Satzteil aus fünf (über Satzgrenzen hinweg) so verändert wird, daß der gesamte Text (-teil) einen in sich konsistenten, nur völlig unterschiedlichen Sinn ergibt – der Proband muß dann den ersetzten Satzteil ankreuzen, z. B.:
Gelesener Text: Voter apathy is almost a cliché in discussions of American politics. Yet, only a cursory look at voting and registration restrictions shows that many would-be voters do not cast ballots because they are prevented from doing so.

Testtext: Voter apathy is almost a cliché in discussions of American politics. A recent poll directed / at voting and registration restrictions shows that many would-be voters seldom protest or demonstra- / because they are prevented from doing so. . . .

Der einzige Einwand, der gegen diese Meßoperation aufrechterhalten werden kann, ist der Hinweis darauf, daß auch durch dieses Erhebungsverfahren zuviel Gedächtnisprozesse impliziert werden (Carroll 1972, S. 21). Diesem Vorwurf kann man allerdings nur mehr durch *sukzessive Rateverfahren der subjektiven Informationsmessung* (im Amerikanischen „progressive cloze procedure“ genannt) entgegen: dabei wird (im Ausgangsmodell) ein Text Zeichen für Zeichen von links nach rechts erraten und aus den Rateversuchen der Informationsgehalt bestimmt (vgl. zum Modell und der Beziehung zum Verständlichkeitskonstrukt Groeben 1972, S. 71ff.). Shannon, der dieses „Ratespiel“ (1951) entwickelte, hat noch jeden Buchstaben (bis hin zum richtigen) durchraten lassen, was eine völlig unpraktikable Methode darstellt. Weltner hat in jahrelanger Forschungsarbeit diesen Grundansatz vereinfacht: zunächst, indem er nach dem ersten Rateversuch pro Zeichenabbruch (und die entsprechenden Redundanzwerte erster Ordnung durch eine Regressionsgleichung mit den subjektiven Informationswerten verband; Weltner 1967), sodann indem er von der

Einheit Buchstaben über Silben bis auf Worte hinaufging (Weltner 1970, S. 68ff.). Heute ist es dank Weltner möglich, Zufallsstichproben von Text (-teilen) mit dem vereinfachten Rateverfahren auf Wortebene zu untersuchen und mit Hilfe der Weltnerschen (Regressions-) Formeln die subjektive Information zu berechnen (Weltner 1970, S. 154ff.) – ein Verfahren, dem sicher keine Unpraktikabilität (Schulz von Thun 1973, S. 30) vorgeworfen werden kann und das m. E. auf die Dauer das Meßverfahren der Wahl für das Verständlichkeitskonstrukt darstellen sollte.

4. Lernerfolg: Behalten contra Interesse?

4.1. Behaltenskriterium

In bezug auf die Wirksamkeit von Texten (in Abhängigkeit von hoher Verständlichkeit) steht für die instruktionspsychologische Forschung wie für die pädagogische Praxis der kognitive Lernerfolg, d. h. konkret das Behalten von Wissensinformationen, im Vordergrund. Zum Behaltenserfolg beim Wissenserwerb aus schriftlichem Lernmaterial sind nicht nur die meisten Untersuchungen (vgl. Literatur in Freedle u. Carroll 1972) durchgeführt, sondern auch die gesichertsten, übereinstimmendsten Ergebnisse erzielt worden. So ist anhand von unterschiedlichen Probanden-Klassen, Textarten und -inhalten, Behaltenstests bestätigt, daß der wichtigste Faktor für den kognitiven Lernerfolg die kognitive Gliederungs-Ordnung des Textinhalts ist. In meiner Untersuchung von 1972 ergab sich mit Hilfe der klassischen cloze procedure als Behaltenstest (nach Taylor 1953; 1956: Auslassen von Inhaltsworten, insgesamt 20% des Gesamttextes in einem durchschnittlichen Abstand von fünf Worten), daß starke kognitive Strukturierung zu gutem konzeptuellen Lernerfolg führte, während kognitiver Konflikt mehr faktuelles Lernen verstärkte (Groeben 1972, S. 106ff.); die übrigen Faktoren hatten keinen Einfluß auf den Lernerfolg. In den Hamburger Untersuchungen war für den Behaltenserfolg (Langer et al. 1974, S. 65ff.) neben dem Faktor Gliederung auch die sprachliche Einfachheit sehr bedeutsam (96) – hier sind (wie unter 3. hinsichtlich der Verständlichkeit) die Hamburger Ergebnisse wegen der größeren Streubreite der Texte als valider anzusetzen! Die Behaltensprüfung erfolgte in diesen Untersuchungen zumeist durch offene Fragen (die man wegen der zur Beantwortung nötigen Gedächtnisprozesse – s. oben 3. – eindeutig als Behaltenstests klassifizieren sollte – Langer et al. sprechen etwas verschleiend immer von „Verständnis–Behalten“). Die Wirksamkeit der Faktoren kognitive Gliederungs-Ordnung und sprachliche Einfachheit für den kognitiven Lernerfolg kann daher, in Übereinstimmung mit dem Großteil der instruktionspsychologischen Forschung (vgl. Ausubel 1968), als gesichert gelten.

Offen ist dagegen bisher, ob generell zwischen der Verständlichkeit eines Textes und dem Lernerfolg eine lineare oder eine kurvilineare Beziehung anzunehmen ist, d. h. ob man die Textverständlichkeit (im Prinzip) unbegrenzt steigern sollte, um möglichst viel Lernen zu ermöglichen, oder ob von einer gewissen Vereinfachung an der kognitive Lehrerefolg (eventuell über die Interaktion mit verringerter Motivierung) absinkt. Die Hamburger Forschungsgruppe geht implizit immer von der maximalen Steigerung der Verständlichkeit von Texten aus, obwohl die maximal möglichen Verständlichkeitsgrade auch durch ihre optimierten Texte meistens nicht erreicht werden (vgl. Langer et al. 1974, S. 65ff.). Ich selbst habe, besonders nach Zusammenfassung des Behaltens- mit dem Interesssekriterium, eine kurvilineare Beziehung von Textverständlichkeit und Lernerfolg (über eine Regressi-

onskurve) herausgearbeitet (1972, S. 113ff.) – das kann aber keineswegs als eine ausreichende Sicherung dieser Beziehung gelten: denn zum einen ist der mittlere Bereich nur durch die relativ willkürlichen Standardversionen meines Ausgangstextes definiert (vgl. Kritik von Schulz von Thun 1973, S. 30), wengleich diese Standardversionen doch ein relativ repräsentatives Schreibverhalten von Wissenschaftlern darstellen (Groeben 1972, S. 88); zum zweiten aber (und das ist gewichtiger) leidet diese statistische Aufarbeitung an der mangelnden Validität der Erhebung des Interesses (s. unten 4.2.). Es gibt aber zumindest zwei theoretische Argumente, die es sinnvoll machen, weiter der Frage nachzugehen, ob eine Steigerung der Textverständlichkeit immer auch eine Verbesserung des Lernerfolgs bewirken muß:

a) Die Psychologie kennt eine Fülle von Adaptationsphänomenen/-prozessen; es ist nicht recht einzusehen, warum nicht auch im Bereich der Textrezeption (besonders der sprachlichen Einfachheit und der Kürze–Prägnanz) Gewöhnungserscheinungen (und damit eventuell Verschiebungen des Bezugssystems – Helson –) auftreten können, so daß, zumindest für bestimmte Leser, dann einzelne Texte „zu leicht“ werden. Entsprechende Anpassungsprozesse ließen sich mit kybernetischen Regelkreismodellen gut beschreiben – ein zusätzliches Argument für die oben (3.) vorgeschlagene informationstheoretische Messung der Verständlichkeit.

b) Der Rückgriff auf „bestimmte Leser mit Adaptationsphänomenen“ impliziert grundsätzlich schon das zweite Argument: die Notwendigkeit, interindividuell verschiedene sprachliche und wissensmäßige Rezeptionskompetenzen zu berücksichtigen – eine Perspektive, die in der pädagogischen Psychologie seit längerem als ATI-Ansatz („aptitude–treatment–interaction“: Wechselwirkung–Unterrichtsmethode–Schülermerkmal“; vgl. Flammer 1973; 1975) thematisiert wird und zur Individualisierung der Unterrichtsbe-mühungen beiträgt. Da sich das Ziel der Individualisierung in letzter Zeit als notwendige Weiterentwicklung und Differenzierung pädagogisch-psychologischer Theorien herausgestellt hat (vgl. Treiber et al. 1976), ist die Frage nach einer möglichen kurvilinearen Beziehung zwischen Textverständlichkeit und Lernerfolg mit ihren ATI-Implikationen als eines der wichtigsten zur Lösung anstehenden Probleme des Forschungsprogramms „Textverständlichkeit“ anzusehen.

In bezug auf die Relation des kognitiven Lernkriteriums und des Ziels der Motivierung allerdings sind die Befunde wiederum völlig eindeutig und einheitlich, wenn die Untersuchungen auf der Grundlage der umfassenden Textdimensionen (s. oben 2.) beide Zielkriterien überprüfen (was in dieser Komplexität m. E. nur für Langer et al. 1974 und Groeben 1972 gilt): das Behaltenskriterium ist dem Motivierungsaspekt vorzuordnen, weil es hier, zumindest in den oberen Ausprägungsgraden der Stimulanz (durch z. B. konzeptuellen Konflikt), eine Gegenläufigkeit gibt: stimulierende Textstrukturen bewirken immer auch eine höhere Komplexität des Lernmaterials, die ohne eine übergeordnete kognitive Strukturierung/Ordnung nur zur Verwirrung des Rezipienten führt. Die Ergebnisse und Schlußfolgerungen von Langer et al. (1974, S. 24): „Die Dimension Zusätzliche Stimulanz war in Verbindung mit hoher Ausprägung in Gliederungs-Ordnung verständnisfördernd. Bei ungegliederten, ungeordneten Texten trug sie jedoch noch zur Verwirrung des Lesers bei.“ – und von mir (1972, S. 136): „Man sollte also die Komplexität des Lernmaterials zur Interessensteigerung soweit heraufsetzen, wie es die konzeptuelle Strukturierung des Lernenden nicht schädigt.“ – sind so übereinstimmend, daß man hier von einem (vorläufig) gesicherten Ergebnis sprechen kann.

Fast alles andere aber innerhalb der motivationspsychologischen Perspektive (der Textge-

staltung) ist noch ungesichert: „Diese Dimension bedarf weiterer Forschung, zumal sie vorerst ein Sammelbecken für ganz unterschiedliche Stimulierungsstrategien darstellt.“ (Langer et al. 1974, S. 24).

4.2. Motivierungseffekt: epistemische Neugier oder mathemagenes Verhalten?

Wie die Diskussion der Textdimensionen gezeigt hat (2.), setzt das Rating-Verfahren (bisher) auf einer zu abstrakten Beschreibungsebene (von Interessantheit) an, als daß damit konkrete Stimulierungstechniken erforscht werden könnten. Modelle, die durch Rückgriff/Angabe bestimmter Textmerkmale den (angezielten) Motivierungseffekt theoretisch zu erklären versuchen/beanspruchen, gibt es in der bisherigen Forschung (zur Textgestaltung) vor allem zwei: die *Theorie der epistemischen Neugier* (von Berlyne) und das *Modell des mathemagenen Verhaltens* (Rothkopf).

Berlynes Motivationstheorie geht zentral vom Konstrukt des konzeptuellen Konflikts aus (vgl. besonders 1960); er faßt „Konflikt“ lerntheoretisch auf: als widerstreitende Antworttendenzen im Individuum. Als Konflikt-Auslöser wirken die sog. „kollativen“ Gegenstandsvariablen (1960, S. 44): Neuheit, Inkongruität, Komplexität, Unsicherheit. Im Bereich der Wahrnehmung, von dem aus Berlyne seine Theorie entwickelte, resultiert aus solchen Konflikten Wahrnehmungsneugier (Berlyne 1963); in Ausweitung auf den kognitiven Bereich postuliert Berlyne (besonders 1965) als Folge konzeptueller Konflikte die epistemische Neugier, d. h. einen Wissensdrang, der sich erst nach Erreichung entsprechender Informationen legt. Die *Generierung konzeptueller Konflikte* ist damit als *eine Strategie zur Schaffung intrinsischer Motivation* anzusehen (Berlyne 1963; 1966). Die empirische Überprüfung dieses Ansatzes an sprachlichem Material (durch Berlyne und Mitarbeiter) führte zu folgenden Ergebnissen: epistemische Neugier wird hervorgerufen durch – Fragen mit unerwarteten Inhalten (Berlyne 1954); – möglichst gleichwahrscheinliche Problem-/Fragealternativen (Berlyne 1962); – Neuheit und Überraschung von Textinhalten (Berlyne u. Frommer 1966; Neuheit z. B. durch unbekannte Tiere als Protagonisten von Fabeln; Überraschung: ein kleiner Junge behauptet, eine Wand zu halten, die wirklich zusammenfällt, als er fortgeht); – insgesamt schon durch die einfache Vorschaltung von Fragen vor einen Lerntext (Berlyne 1954; 1963) mit der Folge auch eines (bedeutsam) höheren Behaltenseffekts (Berlyne 1960, S. 297). Das Neugieverhalten wurde dabei durch die Anzahl von Fragen bzw. angekreuzten Wünschen nach weiteren Informationen gemessen. Weitere Überprüfungen mit z. T. älteren Versuchspersonen und alltäglicheren (Lern-) Situationen haben allerdings auch *gegenteilige Ergebnisse* erbracht: in bezug auf das Neugieverhalten selbst konnte Brown (1972) zwar eine Korrelation zwischen Lesezeit von Textanfängen und späterer Textauswahl sichern, doch stellten Switzky et al. (1974) eine Gegenläufigkeit zwischen der Dauer des Explorationsverhalten und der Spieldauer (bei Kindern mit dreidimensionalen Polygonen) fest; hier ist zu fragen, ob bei der Rezeption von Texten die intensivere, weitere Durcharbeitung eines Textes nicht analog zu solchem „Spielverhalten“ anzusehen ist. Deswegen ist eine alltagsnähere Überprüfung mit komplexeren Texten unbedingt nötig: hier fanden Eisemann et al. (1973) lediglich Auswirkungen der Inkongruität von Illustrationen, nicht aber von textueller Inkongruität. Auch in meiner Untersuchung führte die konzeptuell-konfliktvolle Textgestaltung zu keinem (signifikant) stärkeren Neugieverhalten/Interesse (1972, S. 110ff.). Ich habe dieses Ergebnis damals durch die Annahme, daß die Überprüfung des Interesses durch Feststellung des Ausmaßes von

weiterer Aufnahmebereitschaft zu einzelnen Themen (der Gruppenarbeit) kein valides Beobachtungsverfahren darstellt, exhaustiert und an der Berlyneschen Theorie (auch für die Textgestaltung) festgehalten (1972, S. 132ff.). Ich würde diese Interpretation heute als zu weitreichend ansehen und die *Breite des Erklärungsanspruchs* (nicht der Operationalisierung/Textmerkmale) der epistemischen Neugiertheorie etwas *einschränken*: etwa im Sinn des mathemagenen Verhaltens-Modells von Rothkopf.

Rothkopf hat das Konzept des mathemagenen Verhaltens, ausgehend von der Aktivität des Lernenden beim Programmierten Unterricht (durch das Ausfüllen der Leerstellen), entwickelt (1963; 1965; 1969; und stützt sich aufgrund dieses Ausgangspunktes auf die Assoziationstheorie nach Cook u. Kendler 1956). Unter mathemagenem Verhalten werden die *Aktivitäten des Schülers* verstanden, die zu *Lernen* (im Sinn von Wissenserwerb aus schriftlichem Lernmaterial) *führen* (Wortbildung aus dem Griechischen: mathema [das zu Lernende] und gignesthai [geboren werden]). Für die kognitive Informationsverarbeitung von Texten sind die wichtigsten (hypothetischen) Prozesse die „Übersetzung“ der Zeichen in innere Repräsentation/Sprache, die syntaktische Decodierung („segmentation“) und die Verbindung von Wort-/Satzsinneinheiten zum Textsinn („processing“; vgl. Rothkopf 1970, S. 328ff.). Die wichtigste Forschungsfrage besteht in der Suche nach Bedingungen, die solches mathemagenes Verhalten ermöglichen/verstärken. Dabei wird die Schwierigkeit der adäquaten *Operationalisierung* der abhängigen Variable, der *Motivierung* zur Textrezeption/-verarbeitung, von vornherein vermieden: ob mathemagenes Verhalten erreicht wurde, wird *nur indirekt* erschlossen aus der (größeren) Behaltensleistung. Hinsichtlich der *Bedingungen für mathemagenes Verhalten* haben sich die *Forschungsbemühungen* praktisch ausschließlich auf das *Einfügen von Fragen* in den Text *konzentriert* (das allerdings in großer Quantität, so daß schon mehrere Sammelreferate vorliegen: Frase 1970; Rothkopf 1970; 1971). Der paradigmatische Versuchsplan (Rothkopf 1971, S. 289ff.) sieht bestimmte Fragen (an verschiedenen Stellen des Textes „interspersed“) vor, denen ein Teil der (Mehrfach-Wahl-) Fragen zur Behaltensprüfung am Schluß entspricht, ein anderer Teil nicht; auf diese Weise läßt sich die direkte Wirkung der eingeschalteten Fragen, wie auch ein (nur durch Intensivierung des mathemagenen Verhaltens erklärbarer) „Transfer“ auf andere Textinhalte (inzidentelles Lernen) feststellen. Die empirischen *Ergebnisse* hinsichtlich der direkten Wirkung von Fragen sind: *nach dem* entsprechenden *Textteil platzierte Fragen* führen zu *mehr Behaltenserfolg* als *davor* platzierte (z. B. Bruning 1968; Frase 1968a; 1970; Rothkopf 1963; 1966), auch für das Langzeitgedächtnis (Boker 1974) und besonders bei geringer Motivation (Frase, Patrick u. Schumer 1970). In bezug auf die generelle Wirkung (auch auf Inhalte, die nicht „befragt“ wurden) erwiesen sich vorangestellte Fragen als unwirksam, da sie die Aufmerksamkeit zu sehr auf die thematisierten Inhalte konzentrieren (Frase, Patrick u. Schumer 1970), während für nachgestellte Fragen bessere Behaltenswerte als bei Kontrollgruppen (ohne Fragen) festgestellt wurden (Bruning 1968; Boker 1974; Frase 1967; Patrick 1968; Rothkopf 1966), besonders wenn die Fragen nicht zu direkt/speziell waren (Rothkopf u. Bloom 1970; Rothkopf u. Billington 1974) und persönlich von einem Lehrer gestellt wurden (sozialer Kontakt: Rothkopf 1972). Die Wirkung der Fragen ist unabhängig davon, ob sie durch die Probanden beantwortet werden oder nicht (Rothkopf 1971, S. 294f.). Für die direkte Wirkung von Fragen (auf besseres Behalten der thematisierten Inhalte) ist allerdings die Erklärung durch mathemagenes Verhalten nicht nötig (worauf auch Rothkopf verweist: 1970, S. 335); nur bei der generellen (transferhaften) Intensivierung der Informationsverarbeitung ist die Annahme mathemagener Motivierungsprozesse sinnvoll. Nun hat aber Ladas (1973) gerade an den

Experimenten zur Überprüfung dieser indirekten Wirkung methodische Kritik geübt und den Nachweis eines entsprechenden inzidentellen Lernens völlig in Zweifel gezogen; auch gibt es gegenteilige (falsifizierende) Evidenzen (Walker 1974). Es ist daher zur Zeit trotz der beträchtlichen Forschungsbemühungen *keine sichere Entscheidung über den empirischen Bewährungsgrad* des Konzepts vom mathemagenen Verhalten möglich.

4.3. Integrationsmöglichkeiten

Die Gegenüberstellung der beiden wichtigsten motivationspsychologischen Forschungsperspektiven zur Textgestaltung zeigt einerseits, wie ungeklärt dieses Problem noch ist, weist aber gleichzeitig auch Entwicklungsmöglichkeiten und -notwendigkeiten auf. Die beachtliche Überlappung im Bereich der unabhängigen Variablen – bei Rothkopf spielen Fragen (als UV) eine praktisch ausschließliche, bei Berlyne eine doch recht gewichtige Rolle – läßt eine *Verbindung und damit gegenseitige Korrektur der beiden Ansätze* als möglich, sinnvoll und eventuell fruchtbar erscheinen. Es gibt auch bereits empirische Hinweise für diese Fruchtbarkeit: Bull (1973) kritisiert an den Experimenten der Rothkopf-Gruppe (Bell Telephone Laboratories), die im übrigen die Ergebnisse von Berlyne-Untersuchungen praktisch völlig vernachlässigt, die Einfachheit der in den Text eingestreuten Zwischenfragen und weist darauf hin, daß die motivierende (und damit auch Behaltens-) Wirkung der Fragen von deren Erregungspotential (sensu Berlyne) abhängen dürfte. Sie konnte in einer entsprechenden empirischen Untersuchung (Bull u. Dizney 1973) auch nachweisen, daß die typischen Rothkopf-Fragen mit ihrem geringen Erregungspotential einen viel schwächeren Lerneffekt, besonders in der Langzeitwirkung, erzielten als Fragen mit höherem Erregungspotential. Andererseits haben Frick u. Cofer (1972) gezeigt, daß die motivierende Wirkung von (Berlyne-) Fragen auch davon abhängt, daß sie nicht völlig ohne Bezug zum Textinhalt (allerdings auch nicht identisch mit diesem) sind.

Diese Ergebnisse weisen auf den ersten wichtigen Bereich der gegenseitigen Korrektur bei der Verbindung der epistemischen Neugiertheorie und des mathemagenen Verhaltenskonzepts hin: die *Ergänzung/Erweiterung des Modells mathemagenen Verhaltens im Bereich der empirischen (Antezedenz-) Variablen*. Es ist sowieso etwas befremdlich, daß ein Forschungsprogramm wie das Rothkopfsche so viel Beachtung erhält, obwohl es sich auf die z. T. doch recht triviale Antezedenzbedingung des Einschlebens von Fragen in den Text konzentriert und andere Möglichkeiten der Textgestaltung (zur Motivierung) praktisch gänzlich außer acht läßt. Hier könnte eine Berücksichtigung der einfallsreichen und originellen Operationalisierungen der kollativen Variablen (nach Berlyne) nur zu einer Bereicherung der motivationspsychologischen Forschungsperspektive (der Textgestaltung) führen! Mit einer solchen Erweiterung müßte aber auch die differenziertere, *präzisere Erfassung* der abhängigen Variable: *des Motivierungseffekts* einhergehen. Hier muß sich das Forschungsprogramm des mathemagenen Verhaltens m. E. der Anforderung stellen, den behaupteten Motivierungseffekt *möglichst direkt* (und nicht nur vermittelt über den – generalisierten – Behaltenserfolg) zu überprüfen; daß auch die Berlyneschen Operationalisierungen des „Interesses“ bzw. „Neuigerverhaltens“ hier noch nicht ausreichen, haben die bisherigen Ergebnisse gezeigt. Hier sollte eine stärkere Konzentrierung auf den Rezeptionsprozeß (und weniger auf den Textinhalt) weiterhelfen.

Damit ist der zweite Bereich der gegenseitigen Korrektur angesprochen: die *einschränkende Präzisierung der* (theoretischen) *Erklärungsbreite der Neugiertheorie durch den Roth-*

kopf-Ansatz. Zweifellos weist der Berlynesche Theorieansatz im Vergleich zu Rothkopfs Modell den größeren Erklärungsanspruch auf: nicht nur, daß er vom Wahrnehmungs- auf den Kognitionsbereich ausgreift und für beide Geltung beansprucht, er versucht auch im Bereich der Handlungs- und Erlebensrepräsentanz (Holzkamp 1964) weitere Dimensionen abzudecken; so ist das Konzept intrinsischer Motivierung und Interesses (zumindest tendenziell) in Zeit-, Inhalts- und Persönlichkeitsdimension auf eine relativ weitreichende Perspektive ausgerichtet: es wird idealtypisch als Endpunkt der Aufbau eines relativ zeitüberdauernden, inhalts- und persönlichkeitspezifischen Interesses (Interessensystems) mitgemeint. Demgegenüber beschränkt sich das Konzept des mathemagenen Verhaltens auf eine Erklärung sehr viel kürzerer Reichweite: es steht hier die sehr viel kurzfristigere (zeitvariante), auf den (Rezeptions-) Prozeß konzentrierte und damit nicht persönlichkeits-, sondern textspezifische Motivierung (zur weiteren Informationsaufnahme durch Lesen) im Vordergrund. Sowohl von der theoretischen Plausibilität wie von den empirischen Ergebnissen her scheint es mir sinnvoll, gerade für nichtlabilierte Alltagssituationen (Holzkamp 1970) *im Hinblick auf die Wirkung von Textmerkmalen/-gestaltung den Erklärungsanspruch auf diese prozeßorientierte Perspektive der Motivierung zu konzentrieren*. Eine entsprechende theoretische Explikation (sensu Rothkopf) mit einer Ergänzung durch konzeptuell-konfliktreiche Textmerkmale (sensu Berlyne) als Voraussetzung für die (Prozeß-) Motivierung und einer direkteren Überprüfung/Operationalisierung des Motivierungseffekts stellt m. E. derzeit die optimale Forschungsstrategie zur Klärung der offenen motivationspsychologischen Probleme der Textgestaltung dar. Zu diesen Problemen gehört dann selbstverständlich auch noch die Frage nach dem Aufbau spezifischer, möglichst intrinsischer Interessensysteme mit Hilfe des Lernens aus schriftlichem Lehrmaterial, deren Beantwortung aber nur als längerfristiges Forschungsteilprogramm in Interaktion mit sozial- und differentialpsychologischen Theorieansätzen aufgefaßt werden kann.

5. Anwendung

Die praktische Anwendung von Forschungsergebnissen zeigt vom pragmatischen Ablauf her die umgekehrte Struktur wie deren Gewinnung; während man bei der Aufstellung und Überprüfung wissenschaftlicher Gesetzmäßigkeiten von der Variation/Realisation bestimmter Bedingungen (hier Textmerkmale) ausgeht und davon abhängige Effekte (hier Verständlichkeit und Lernerfolg) feststellt, setzt man für die Anwendung bei den erforschten Effekten – als Ziele – ein und geht (mit Hilfe der bewährten Hypothesen/Gesetze) auf die Bedingungen dieser Ziele zurück: solche Bedingungen (Textmerkmale) sind dann in der pädagogischen Praxis zur Erreichung des angestrebten Ziels (Lernerfolg) zu schaffen. Diese wissenschaftstheoretische Struktur der (Anwendungs-) Technologie (vgl. Groeben u. Westmeyer 1975, S. 213f.) ist der Grund dafür, daß die Forschung auch und gerade in bezug auf die Praxisrelevanz möglichst konkrete und objektive, unabhängige Variablen realisieren sollte; *die praktische Ergiebigkeit* (z. B.) des Forschungsprogramms „Textverständlichkeit“ *hängt daher davon ab, wieviel begründete Hinweise auf konkrete Bedingungen/Mittel der Textgestaltung am Ende gegeben werden können*. Ich habe daher am Ende meiner Untersuchung von 1972 (S. 147ff.) diese aus dem Rückschluß vom Lernerfolg über die Verständlichkeitsoptimierung abgeleiteten Handlungsanweisungen für die anzustrebenden Textmerkmale zusammengestellt. Allerdings gilt auch für die Wirksamkeit solcher Ratschläge (wie überall bei Lernprozessen), daß ein aktives Üben erfolgreicher ist als blo-

bes Rezipieren. Dies haben für das Beispiel der Textoptimierung überzeugend Langer et al. (1974) nachgewiesen, die ein *Trainingsprogramm zur Textverbesserung* entwickelt haben (1974, S. 101ff.), dessen Durcharbeitung – empirisch gesichert – zu besseren Resultaten führt als lediglich das Rezipieren einschlägiger Forschungsergebnisse (1974, S. 173ff.). Allerdings wird sich die Wirksamkeit dieses Programms sicherlich noch steigern lassen, wenn man bei der zu lernenden Textgestaltung mehr auf objektive Merkmale als auf die subjektiven Eindrucks-kategorien des Hamburger Rating-Ansatzes zurückgreift; das zeigt sich auch schon darin, daß die Autoren bei der Erläuterung ihrer Textverbesserungen innerhalb des Programms selbst auf Merkmale zurückgreifen, die mit dem Rating-Ansatz gar nicht herauszufinden sind (und auch nicht mit diesem gefunden wurden wie „advanced organizers“ – nach Ausubel –, überschiftete Unterabschnitte usw.; vgl. Langer et al. 1974, S. 148; S. 152). Dem Praktiker ist hier schon jetzt die Zusammenlegung von Arbeiten zur mehr objektiven Erforschung von Textmerkmalen (vgl. Hofer in diesem Heft) mit dem Hamburger Trainingsprogramm zu empfehlen, für den Wissenschaftsbereich resultiert die Anforderung, an der Differenzierung und damit Verbesserung solcher Trainingsmodelle zu arbeiten.

6. Zusammenfassung

Abschließend lassen sich die vorläufig gesicherten, übereinstimmenden Antworten sowie die offenen Fragen des Forschungsbereichs „Textverständlichkeit“ stichwortartig zusammenstellen:

Antworten: – die vier Dimensionen der Textstruktur; – die (mediale) Explikation des Verständlichkeitskonstrukts; – die Faktoren des Textverständnisses; – Verständlichkeitsmessung durch cloze prozedure bzw. Rateverfahren; – die Bedeutung der kognitiven Gliederung und sprachlichen Einfachheit für den Behaltenseffekt; – die Vorordnung des Behaltenskriteriums vor den Motivationseffekt.

Fragen: – Linearität oder Kurvilinearität in der Beziehung zwischen Verständlichkeit und Lernerfolg; – die adäquate Operationalisierung des Motivierungseffekts; – die theoretische Beschreibung/Erklärung von Interesse/Motivierung (Berlyne oder Rothkopf); – die (umgreifenden, differenzierten) Bedingungen der Motivierung in der Textgestaltung; – die Ergänzung vorliegender Trainingsprogramme durch objektivitätsgerichtete Forschungs-/Technologieansätze.

Literatur

- Anderson, R. C.: How to construct achievement tests to assess comprehension. *Rev. educ. Res.* 42 (1972), 145–170.
- Ausubel, D. P.: The psychology of meaningful verbal learning. New York 1963.
- Ausubel, D. P.: Educational psychology: a cognitive view. New York 1968.
- Berlyne, D. E.: A theory of human curiosity. *Brit. J. Psychol.* 45 (1954), 180–191.
- Berlyne, D. E.: Conflict, arousal and curiosity. New York 1960 (dtsh. Stuttgart 1974).
- Berlyne, D. E.: Motivational problems raised by exploratory and epistemic behavior. In: Koch, S. (ed.): *Psychology: a study of science*, Vol. 5., New York 1963.
- Berlyne, D. E.: Structure and direction in thinking. New York 1965.
- Berlyne, D. E.: Notes on intrinsic motivation and intrinsic reward in relation to instruction. In: Bruner, J. S. (Ed.): *Learning about learning*. Washington 1966.
- Berlyne, D. E., Frommer, F.: Some determinants of the incidence and content of childrens questions, *Child Develop.* 37 (1966), 177–189.
- Boker, J. R.: Immediate and delayed retention effects of interspersing questions in written instructional passages. *J. educational Psychol.* 66 (1974), 96–98.
- Brown, L. T.: A behavioral index of the exploratory value of prose materials. *J. educ. Psychol.* 63 (1972), 437–445.
- Bruning, R. H.: Effects of review and test-like events within the learning of prose material. *J. educ. Psychol.* 59 (1968), 16–19.
- Bull, S. G.: The role of questions in maintaining attention to textual material. *Rev. educ. Res.* 43 (1973), 83–87.
- Bull, S. G., Dizney, H. F.: Epistemic curiosity-arousing prequestions: Their effect on longterm retention. *J. educ. Psychol.* 65 (1973), 45–49.
- Carroll, J. B.: Defining language comprehension: some speculations. In: Freedle, R. O., K. B. Carroll (Ed.): *Language comprehension and the acquisition of knowledge*, p. 1–29. Washington 1972.
- Carver, R. P.: Analysis of "chunked" test items as measures of reading and listening comprehension. *J. educ. Measur.* 7 (1970), 141–150.
- Cook, J. O., T. S. Kendler: A theoretical model to explain some paired-associate learning data. In: Finch, G. F., Cameron, (Ed.): *Symposium on Air Force human engineering, personnel, and training research*, p. 90–98. Washington 1965.
- Davis, F. B.: Fundamental factors of comprehension in reading. *Psychometrika* 9 (1944), 185–197.
- Davis, F. B.: Psychometric research on comprehension in reading. *Read. Res. Quart.* 7 (1972), 628–678.
- Eisemann, M., N. Sieveking, M. E. Binkley: Use of incongruous materials in classroom learning, *J. genet. Psychol.* 123 (1973), 227–230.
- Flammer, A.: Wechselwirkung zwischen Schülermerkmal und Unterrichtsmethode. *Z. Entwickl.- u. päd. Psych.* V (1973), 130–147.
- Flammer, A.: *Individuelle Unterschiede im Lernen*, Weinheim 1975.
- Frase, L. T.: Learning from prose material: length of passage, knowledge of results, and position of questions. *J. educ. Psychol.* 58 (1967), 266–272.
- Frase, L. T.: Questions as aids to reading: some research and theory. *Amer. educ. Res. J.* 5 (1968a), 319–332.
- Frase, L. T.: Some data concerning the mathemagenic hypothesis. *Amer. educ. Res. J.* 5 (1968b), 181–189.
- Frase, L. T.: Boundary conditions for mathemagenic behavior. *Rev. educ. Res.* 40 (1970), 337–347.
- Frase, L. T., E. M. Patrick, H. Schumer: Effect of question position and frequency on learning from text under different levels of incentive. *J. educ. Psychol.* 71 (1970), 52–56.
- Freedle, R. O., J. B. Carroll: *Language comprehension and the acquisition of knowledge*. Washington 1972.
- Frick, J. W., Ch. N. Cofer: Berlyne's demonstration of epistemic curiosity: An experimental re-evaluation. *Brit. J. Psychol.* 63 (1972), 221–228.
- Gray, W. S. B. Leary: *What makes a book readable*. Chicago 1935.
- Groeben, N.: Die Verständlichkeit von Unterrichtstexten. Dimensionen und Kriterien rezeptiver Lernstadien. Münster 1972.
- Groeben, N., H. Westmeyer: *Kriterien psychologischer Forschung*. München 1975.
- Holzkamp, K.: *Theorie und Experiment in der Psychologie*, Berlin 1964.
- Holzkamp, K.: Zum Problem der Relevanz psychologischer Forschung für die Praxis. *Psychol. Rundsch.* XXXI (1970), 1–22.
- Klare, G. R.: The measurement of readability. Ames 1963.
- Klare, G. R.: Some empirical predictors of readability. In: Rothkopf, E. Z., P. E. Johnson (Ed.): *Verbal learning research and the technology of written instruction* p. 241–254. Columbia 1971.
- Ladas, H.: The mathemagenic effects of factual review questions on the learning of incidental information: A critical review. *Rev. educ. Res.* 43 (1973), 71–82.
- Langer, I.: Verständnisfördernde Merkmale der sprachlichen Gestaltung von Unterrichtstexten. *Phil. Diss.* Hamburg 1971.
- Langer, K., F. Schulz von Thung, R. Tausch: *Verständlichkeit in Schule, Verwaltung, Politik und Wissenschaft*. München-Basel 1974.
- Patrick, E. M.: Prose learning: induced question and response rehearsal, and question repetition. University of Massachusetts 1968 (unpublished manuscript).
- Poulton, E. C.: British courses for adults on effective reading. *Brit. J. educ. Psychol.* 31 (1961), 128–137.
- Pyrzrak, F.: Objective evaluation of the quality of multiple-choice test items designed to measure comprehension of reading passages. *Read. Res. Quart.* VIII (1972), 62–71.
- Rothkopf, E. Z.: Some conjectures about inspection behavior in learning from written sentences and the response mode problem in programmed self-instruction. *J. program. Instr.* 2 (1963), 31–46.
- Rothkopf, E. Z.: Some theoretical and experimental approaches to problems in written instruction. In: Krumboltz, J. D. (Ed.): *Learning and the educational process*, p. 193–221. Chicago 1965.
- Rothkopf, E. Z.: Learning from written materials: an exploration of the control of inspection behavior by test-like events. *Amer. educ. Res. J.* 3 (1966), 241–249.
- Rothkopf, E. Z.: Concerning parallels between adaptive processes in thinking and self-instruction. In: Voss, J. (Ed.): *Approaches to thought*, 229–316. Columbus 1969.
- Rothkopf, E. Z.: The concept of mathemagenic activities. *Rev. educ. Res.* 40 (1970), 325–336.
- Rothkopf, E. Z.: Experiments on mathemagenic behavior and the technology of written instruction. In: Rothkopf, E. Z., P. E. Johnson (Ed.): *Verbal learning research and the technology of written instruction*, p. 284–303. Columbia 1971.

- Rothkopf, E. Z.: Variable adjunct question schedules, interpersonal interaction, and incidental learning from written material. *J. educ. Psychol.* 63 (1972), 87–92.
- Rothkopf, E. Z., M. J. Billington: Indirect review and priming through questions. *J. educ. Psychol.* 66 (1974), 669–679.
- Rothkopf, E. Z., E. E. Bisbicos: Selective facilitative effects of interspersed questions on learning from written materials. *J. educ. Psychol.* 58 (1967), 56–61.
- Rothkopf, E. Z., R. D. Bloom: Effects of interpersonal interaction on the instructional value of adjunct questions in learning from written material. *J. educ. Psychol.* 61 (1970), 417–422.
- Schulz von Thun, F.: Effektivität und Trainierbarkeit von Verständlichkeit bei der schriftlichen Informationsvermittlung. *Phil. Diss. Hamburg* 1973.
- Shannon, C. E.: Prediction and entropy of printed English. *Bell Syst. techn. J.* 30 (1951), 50–64.
- Spearritt, D.: Identification of subskills of reading comprehension by maximum likelihood factor analysis. *Read. Res. Quart.* VIII (1972), 92–111.
- Switzky, H. N., H. C. Haywood, F. Isett: Exploration, curiosity, and play in young children: Effects of stimulus complexity. *Developm. Psychol.* 10 (1974), 321–329.
- Taylor, W. L.: Cloze procedure: A new tool for measuring reasoning. *Journ. Quart.* 30 (1953), 415–433.
- Taylor, W. L.: Recent development in the use of cloze procedure. *Journ. Quart.* 33 (1956), 42–48.
- Thurstone, L. L.: Note on a reanalysis of Davis' reading tests. *Psychometrika* 11 (1946), 185–188.
- Treiber, B., F. E. Weinert, N. Groeben: Bedingungen individuellen Unterrichtserfolgs. *Z. Päd.* (im Druck).
- Walker, B.: Effects of inserted questions on retroactive inhibition in meaningful verbal learning. *J. educ. Psychol.* 66 (1974), 486–490.
- Weaver, W. W., A. C. Bickley: Sources of information for responses to reading test items. *Proc. 75th Ann. Conv. Amer. Psychol. Ass.* 1967.
- Weltner, K.: Subjektive Information von deutschen Texten und didaktische Transinformation. In: Merz, F. (Ed.), *Ber. 25. Kongr. Dtsch. Ges. Psychologie*, Göttingen 1967
- Weltner, K.: *Informationstheorie und Erziehungswissenschaften*. Quickborn 1970

Verfasser: Norbert Groeben, Psychologisches Institut, Psycholinguistische Abteilung, Hauptstraße 242, 69 Heidelberg 1.

Manfred Hofer

Textverständlichkeit: Zwischen Theorie und Praxeologie

Nach einem kurzen historischen Abriss werden drei Forschungsrichtungen, die mit den Namen Teigeler, Groeben und Langer/Schulz von Thun/Tausch gekennzeichnet sind, auf Grund der vorliegenden Veröffentlichungen nach ihren theoretischen Grundlagen und Intentionen analysiert und verglichen. Der Beitrag ist eine Einführung in den heutigen Diskussionsstand.

Textual Intelligibility: Between Theory and Praxeology

Following on a short historical survey three schools of research – marked out by the names of Teigeler, Groeben, and Langer/Schulz von Thun/Tausch – are analysed and compared, on the basis of their available publications, with regard to their theoretical foundations and intentions. The paper is an introduction into the present state of discussion.

Man könnte glauben, die extensive Diskussion über unser Bildungssystem mit ihren mannigfachen Impulsen für bildungsbezogene Forschungsbereiche in den vergangenen Jahren hätte auch die Bedeutung von schriftlichem Material erkannt und eine Reihe von Ansätzen und bedeutsamen Bemühungen zum Problem der Rezeption und Gestaltung von Texten nach sich gezogen. Immerhin spielte das Lehrbuch als ältestes Unterrichtsmedium – die „Elemente der Geometrie“ von *Euklid* waren von 300 v. Chr. bis in die Neuzeit hinein in Gebrauch – im Hochschulbereich eine führende und auch im primären Bildungsbereich nicht zu unterschätzende Rolle. Bezogen auf den tertiären Bildungsbereich werden viele reflektierende Zeitgenossen nicht müde, den quadratischen bis höher-exponentiellen Anstieg zu rezipierenden und rezipierenswerten Wissens zu beschwören und die Frage zu stellen, wie man der Flut informierender Gedanken ordnend, selektierend bis zusammenfassend Herr werden könnte. Da es nur teilweise sinnvoll sein kann, die Nützlichkeit des geschriebenen Wortes zu verleugnen, wäre es verständlich, wenn (zumindest bezogen auf Rezipierenswertes) die Frage nach der Optimierung schriftlichen Materials gestellt worden wäre. Immerhin existiert im englischen Sprachbereich bereits seit den 30er Jahren eine intensive Forschungstätigkeit über Probleme der Textgestaltung und -aufnahme. Und Amerikaner fordern und entwickeln seit zwei Jahrzehnten Manuale, die Autoren und Verlage über die Textgestaltung instruieren.

Im deutschen Sprachbereich rückt dagegen das Thema erst in jüngster Zeit in den Blickwinkel von Wissenschaftlern und textproduzierenden Organisationen. Die im Zitat:

„... in westlichen Ländern ermöglichten wissenschaftliche Untersuchungen zum Bereich der Lesetechniken im Grundschulbereich die Konstruktion von Schulbüchern, bei denen die Seitengestaltung genauso sorgfältig erfolgte wie die Wahl des Vokabulars. Wortzählungen bestimmen den Text. Visuelle Experimente determinieren die Drucktype. Von Illustrationen wird ausgiebig Gebrauch gemacht“ (*Encyclopaedia Britannica* 1962, S. 991)

zum Ausdruck kommende, quasi selbstverständliche Berücksichtigung jener Variablen, deren Einfluß auf Rezipiervariablen zum jeweiligen Zeitpunkt für wesentlich gehalten wird, bleibt auf den englischsprachigen Raum beschränkt. Von dem vor einigen Jahren in Tübingen gegründeten Deutschen Institut für Fernstudien ist das Thema unter empirischem Aspekt bislang auch leider völlig ausgeklammert worden.

Die vier Publikationen von Teigeler (1968 und 1972), Groeben (1972) und Langer et al. (1974), die zu Fragen nach dem Lernen und der Optimierung schriftlicher Sprache gleichzeitig drei verschiedene Forschungsrichtungen repräsentieren, lassen auf eine Beendigung