S 22170 F

Sonderdruck aus

ZEITSCHRIFT FÜR
EXPERIMENTELLE UND ANGEWANDTE
PSYCHOLOGIE

Organ der Deutschen Gesellschaft für Psychologie

Heft 2 / Band XXXIII 2. Quartal 1986

VERLAG FÜR PSYCHOLOGIE · DR. C. J. HOGREFE
GÖTTINGEN · TORONTO · ZÜRICH

Zeitschrift für experimentelle und angewandte Psychologie
1986, Band XXXIII, Heft 2, S. 208—253

Informationale Beschränkungen der Wissenserweiterung?[1])

August Fenk

Universität für Bildungswissenschaften Klagenfurt

Die für die kognitionspsychologische Theoriebildung zentrale Annahme einer aus Kapazitätslimits herrührenden Konstanz informational definierter Leistungen wird neu formuliert, vor allem mit Hilfe des Begriffes der „subjektiven Information", der sich — wie die Item-Schwierigkeit — auf eine Relation zwischen Nachricht und Verarbeitungssystem bezieht. Die Probanden mußten zwei unterschiedlich „schwierige" Lückentexte je zweimal buchstabenweise (Ratetest von Weltner, 1970) rekonstruieren; ein überdurchschnittlich günstiges Verhältnis von richtigen zu falschen Eingaben pro Zeit wurde mit einer Prämie honoriert. (Exp. 1: 48 Vpn; Exp. 2 mit pro Person breiter abgestufter Schwierigkeit: 16 Vpn.) Der Lerngewinn vom 1. zum 2. Ratedurchgang — also die Zunahme an Geläufigkeit oder die Ersparnis an bit oder bit × sec — sollte intraindividuell trotz wechselnden Ausgangsniveaus konstant bleiben. Obwohl hohe negative Korrelationen zwischen Rate-Geschwindigkeit und -Genauigkeit auf Kapazitätslimits hinwiesen und die gewählte Quantifizierung bit × sec rechtfertigten, widerlegten die entsprechenden Mittelwertsvergleiche die neue Konstanzannahme: Bei einem geringeren Score bit × sec im 1. Ratedurchgang, also bei höherem aufgabenrelevantem Vorwissen, fiel der Lerngewinn geringer aus.

Fragestellung

Die Informationshypothese

Die Vorstellung irgendeiner Obergrenze unserer Informationsverarbeitungskapazität erscheint, unabhängig von jeder empirischen Begründung, plausibel, wenn nicht trivial. Darüber hinaus lassen sich verschiedene kognitionspsychologische Befunde als Argument für dieses Konzept geltend machen, und umgekehrt bietet es den Rahmen für eine integrative Interpre-

1) Die Studie wurde vom „Theodor Körner-Stiftungsfonds" (Förderungspreis 1986) und vom Forschungs-Entwicklungs-Fonds der UBW Klagenfurt unterstützt.

tation eben dieser Befunde. Einige Beispiele, die längst in die Lehrbuch-Literatur Eingang gefunden haben:

— Sollen mehrere Aufgaben simultan durchgeführt werden, so kommt es pro Aufgabe zu Leistungseinbußen. Besonders dann, wenn die Aufgaben einander „ähnlich" sind. Die Interpretation: Diese Aufgaben (über)beanspruchen dieselben kognitiven Ressourcen (vgl. Eysenck, 1984; Wessels, 1984).
— Die Selektivität der Wahrnehmung, also das Prinzip der Auswahl bestimmter Wahrnehmungsangebote auf Kosten anderer, wird als Folge einer beschränkten Auffassungskapazität gedeutet (vgl. Wimmer & Perner, 1979).
— Die Steigerung der Qualität einer Aufgabenbewältigung (weniger Fehler, größere Genauigkeit) geht auf Kosten der Geschwindigkeit und vice versa („speed-accuracy tradeoff"), was unschwer mit Kapazitätslimits in Verbindung zu bringen ist (dazu und zu den oben erwähnten Punkten vgl. Mittenecker & Raab, 1973).

Mehr oder weniger explizit steht die Vorstellung von Begrenzungen derartiger Ressourcen immer im Hintergrund, wenn Phänomene der kognitiven Psychologie — aber auch der Linguistik (vgl. die Markiertheitstheorie von Mayerthaler, 1982) oder der Pädagogik (vgl. das „Lehrstoff-Zeit-Problem" bei Melezinek, 1977) — aus dem Blickpunkt eines ökonomischen Einsatzes von Ressourcen bzw. Aufwandes an kognitiven Kosten beschrieben werden.

Andererseits sind gerade jene Untersuchungen, die speziell darauf angelegt waren, die Grenzen unserer Informationsverarbeitungskapazität dingfest zu machen, als gescheitert anzusehen. Woran kann das liegen?

Wenn, so die Grundüberlegung dieser Untersuchungen, es einen Plafond der Kapazität der jeweils befaßten Instanz (oder Instanzen; vgl. Broadbent, 1977, Posner & McLeod, 1982) gibt, so müßte sich dieser — bei Auslastung oder Überforderung der Kapazität — als Konstanz der jeweiligen Leistung bemerkbar machen; und die Informationstheorie schien das geeignete dimensionslose Maß parat zu haben, welches die Prüfung dieser Konstanz in einer sehr universellen Weise erlaubt, also unabhängig von der Modalität der Nachricht als optisch oder akustisch, als analog oder digital, als codierend oder nicht codierend, als mehr oder weniger sinnlos oder sinnvoll. Diese Annahme einer informationalen Konstanz wird im folgenden als „Informationshypothese" bezeichnet.

„Typische" Experimente

Einige Resultate klassischer Experimente sollen kurz erwähnt werden, und zwar zusammen mit ihrer — recht unterschiedlichen — Bewertung in der Sekundär-Literatur: Während Frank (1971), Begründer der „Kybernetischen Pädagogik", die Informationshypothese als hinreichend gestützt ansieht, kommen eingehendere Literatur-Analysen (Attneave, 1959 bzw. 1969; Mittenecker & Raab, 1973) zu einer ganz anderen Einschätzung.

Howes & Solomon (1951, S. 410) berichten über ihre Versuche: „Data from two experiments, using 75 words, show that the visual duration threshold of a word, measured tachistoscopically by an ascending method of limits, is an approximately linear function of the logarithm of the relative frequency with which that word occurs in the Thorndike-Lorge word counts." Miller, Bruner & Postman (1954) variierten das Redundanzniveau (Approximationsstufen 0, 1, 2 und 4 nach Shannon, 1948) und die Präsentationsdauer von Buchstabenfolgen (10, 20, 40, 100, 200 und 500 msec.). Die Versuchspersonen sollten sich möglichst viele Buchstaben merken und unmittelbar nach jeder Exposition in jener Position, die der Buchstabe innerhalb der Folge eingenommen hatte, in ein vorgefertigtes Formblatt eintragen. Das Ergebnis: „When the lower informational value of the contextually constrained sequences is considered, the corrected identification scores for familiar and unfamiliar sequences of letters become essentially equivalent." (S. 139). Frank (1971, S. 173f.) interpretierte die Resultate der beiden zitierten Arbeiten dahingehend, daß die zum Bewußtwerden nötige Zeit („Apperzeptionszeit") der Information der Nachricht proportional sei. (Außerdem sei die Apperzeptionszeit länger als die kürzeren der im Experiment von Miller, Bruner & Postman verwendeten Expositionszeiten; eine Extrapolation der bei den längeren Expositionszeiten erzielten Scores ergebe eine Apperzeptionsgeschwindigkeit von 16 bit pro sec).

Rubenstein & Aborn (1954) variierten in einem Experiment nach dem Paradigma der freien Reproduktion den objektiven Informationsgehalt von Silben und die Einprägungsdauer. Die Reproduktionsleistung wurde in bit (objektive Information) ausgedrückt. Das Ergebnis widerspricht der Informationshypothese: „More information is recalled from material of lower degree of organization than from material of higher degree of organization. ... The amount of information recalled per unit of study time decreases as the length of study time increases." (S. 151f.)

Dieses Resultat fand Attneave (1969) äußerst verwunderlich. Und im Zusammenhang mit der Frage, wieviel bit pro sec übertragen werden, schildert er (S. 108) zwei Experimente von Sumby & Pollack (1954).[2]) Die

2) Meine Bemühungen, die Original-Arbeit zu erhalten, waren bisher erfolglos.

Ergebnisse: Die beim Lautlesen und beim Schreiben pro Zeit übertragene Information von Wortketten, deren objektiver Informationsgehalt durch sequentielle Beschränkungen des Wortmaterials variiert und deren subjektiver Informationsgehalt per Ratespiel-Technik gemessen worden war, blieb „für ein Material, das Unterschiede hinsichtlich der Redundanz aufwies, ziemlich konstant". Wurde der objektive Informationsgehalt hingegen durch die Anzahl alternativer Einheiten variiert, so stieg die Übertragungszeit pro Wort mit stark wachsendem Informationsgehalt nur sehr geringfügig an, und dies läuft nach Ansicht der Autoren auf ein starkes Anwachsen der Geschwindigkeit der Informationsübertragung hinaus. Ähnliche Prinzipien, meint Attneave, gelten auch für das Kurzzeitgedächtnis und eventuell sogar für das Langzeitgedächtnis: Einerseits wäre im Experiment von Miller & Selfridge (1950) der erinnerte Informationsbetrag bei Variation der sequentiellen Beschränkung des Wortmaterials annähernd konstant geblieben.[3]) „Andererseits fand Pollack (1954) in einer Reihe von Experimenten über unmittelbares Erinnern, daß die Gedächtnisspanne für sequentiell nicht beschränkte ‚Nachrichteneinheiten' praktisch konstant ist, gleichgültig aus welcher Zahl von Alternativen die Einheiten gewählt werden. Eine Versuchsperson war imstande, eine Folge von etwa sieben Einheiten zu erinnern, ob es sich dabei z.B. um zufällige Binärzahlen (mit je 1 bit), Dezimalzahlen (mit je 3,3 bit) oder Buchstaben des Alphabets (mit je 4,7 bit) handelte." (Attneave, 1969, S. 109f.)

Nach Miller (1956, S. 93) „müssen wir daher die Bedeutung der Gruppierung oder Zusammenfassung der Eingabe-Sequenz zu Einheiten oder chunks erkennen. Da die Gedächtnisspanne eine feststehende Anzahl solcher Einheiten darstellt, können wir die enthaltenen bits an Information dadurch erhöhen, daß wir immer größere Einheiten bilden, wobei jede neue Einheit mehr Information als die vorige enthält." (Deutsche Übersetzung aus Norman, 1973, S. 105). Die Einschätzung von Mittenecker & Raab (1973): Trotz der Relativierungen, die auch die „chunk-Konstanz" in späteren Untersuchungen erfahren hat, vermittelt die relative Konstanz an chunks bei drastischen Unterschieden im Informationsgehalt pro chunk zumindest den Anschein des „Fehlens einer Beschränkung der informationalen Kapazität" (S. 76) des Kurzzeitgedächtnisses; und aus einigen der zum Thema „Geschwindigkeit versus Genauigkeit" referierten Arbeiten läßt sich den Autoren zufolge „ersehen, daß die Transinformationsgeschwindigkeit keine konstante, von der Art der Aufgabe und von verschiedenen Randbedingungen unabhängige Größe ist". (S. 179)

3) Diese Hinterher-Interpretation kann sich allerdings auf keine Messungen des Informationsgehaltes der reproduzierten Wörter stützen.

Lernen als Informationsübertragung?
Objektive oder subjektive Information?

Die vorliegende Arbeit geht nun davon aus, daß die Informationshypothese ihre Berechtigung hat, und daß die ihr widersprechenden Resultate aus problematischen Operationalisierungen in den einschlägigen Untersuchungen herrühren. „Problematisch" deshalb, weil der zugrundegelegte Lernbegriff und das Konzept der objektiven Information inadäquat sind. (Problematisch sind übrigens auch „Bestätigungen" der Informationshypothese, weil diese aus forschungslogischer Sicht eine Null-Hypothese darstellt.)

Aus der Sicht Weltners (1970) ist Lernen „der Aufbau interner Modelle der Umwelt, die eine Transinformation auf eben diese Umwelt haben ..." (S. 115), und diese Beschreibung sei „mathematisch äquivalent" zu den Beschreibungen von Lernen als „Verringerung der subjektiven Information der Umwelt (Frank)" oder als „Erzeugung von Redundanz (v. Cube)" (S. 118).

Die behauptete Äquivalenz ist zwar bei genauerem Hinsehen nicht gegeben (s. unten), doch läßt sich „Lernen" mit Hilfe der Begriffe „Transinformation", „Redundanzausnutzung" und „Reduktion der subjektiven Information" tatsächlich auf eine Weise beschreiben, die für unsere Fragestellung — aber nicht nur für sie — interessante Perspektiven eröffnet: Aus der Vergangenheit für die Zukunft zu lernen heißt, Hypothesen (interne Modelle) über Regelhaftes — etwa: x tritt häufiger auf als y; x und y treten häufig gleichzeitig oder in dieser Reihenfolge auf — bei der Analyse des jeweils interessierenden Geschehens anzuwenden (Redundanzausnutzung) und nötigenfalls (bei zu geringer Transinformation) zu modifizieren. Ergebnis und „biologischer Zweck" des Lernens ist eine Steigerung der prognostischen Valenz der Hypothesen, so daß — trotz limitierter Ressourcen — die Analyse effektiver und das Geschehen besser antizipierbar wird (Reduktion der subjektiven Information) und das lernende System sich angepaßt verhalten bzw. rational planen und begründet handeln kann.

Die skizzierte Auffassung des ontogenetischen Lernens ist, ähnlich den Ausführungen der „evolutionären Erkenntnistheorie" über das phylogenetische Lernen (z.B. Lorenz, 1977, Vollmer, 1981), unleugbar von logischen Analysen des wissenschaftlichen Erfahrungsgewinns (vor allem Popper, 1976) inspiriert und — nach der von Estes (1976) vorgeschlagenen Klassifizierung von Lerntheorien — den „hypothesis-testing-models" zuzurechnen. Die informationstheoretische Terminologie, in der das Hypothesenprüfungsmodell oben formuliert wurde, eröffnet — wiederum vergleichbar mit Studien zur induktiven Logik (z.B. Greeno, 1970) — Wege zur Quantifizierung der Wissenserweiterung. Zumindest soll sie das ermög-

lichen, was Herzog (s. unten) für kaum möglich hält: nämlich eine sehr direkte Umsetzung des Hypothesenprüfungsmodells in prüfbare Situationen.

Aus behavioristischer Sicht wäre an diesem informationspsychologischen Hypothesenprüfungsmodell allenfalls zu beklagen, daß die (aus der Informationsreduktion resultierende) Verhaltensänderung nicht das Zentrum der Definition bildet, und mit der kognitivistischen Deutung steht es insofern in Einklang, als diese den Akzent auf die aktive Rolle eines explorierenden, hypothesenprüfenden, strategieauswählenden und rekonstruierenden Systems setzt. Gerade diese kognitiven Tätigkeiten werden im Ratetest — und das ist *die* informationspsychologische Methode zur Bestimmung der subjektiven Information — „sozusagen nach außen hin sichtbar. Die Ratefehlerzahl gibt Aufschluß über den erzielten Grad der Übereinstimmung zwischen Ereignisfolgen und den entsprechenden internen Modellen ..." (Lechner & Fenk, in Begutachtung). Und die Ratefehlerzahl kann nur durch ein hypothesen- und strategiegeleitetes, extrapolierendes Vorgehen unter Zufallsniveau gesenkt werden.

Schon diese Eigenheit der Ratespieltechniken, und noch mehr ihre konkrete Nutzung zur Messung des Lernfortschrittes (s. nächstes Kapitel), entkräftet die Kritik Herzogs am „Handlungsmodell" bzw. am „Mensch-als-Wissenschaftler-Modell", wonach ein derartiges Modell „methodologisch kaum in prüfbare Situationen umsetzbar ist". (Herzog, 1984, S. 162). Zum anderen wird hier deutlich, daß „subjektive Information" — ganz analog dem testtheoretischen Begriff der Aufgabenschwierigkeit — nicht nur eine Eigenschaft der Nachricht (der Aufgabe) oder eine Eigenschaft des verarbeitenden Systems (des Problemlösers) betrifft, sondern eine Relation zwischen beiden. So gesehen besteht ein meßbarer Lernfortschritt darin, daß die Analyse einer Situation (die Lösung eines Problems) dadurch leichter wird, daß mehr aufgabenrelevantes (deklaratives und prozedurales) „Vorwissen" verfügbar ist; wobei das Vorwissen immer hypothetischen Charakter hat — nur die jeweils verfügbaren Hypothesen, und nicht die „tatsächlichen Verhältnisse", können verhaltens- oder handlungsrelevant werden. (Wenn diese — z.B. bei Popper, 1976, begründete — Aussage über den grundsätzlich hypothetischen Charakter empirischen, also auch physikalischen Wissens zu Recht getroffen wurde, dann ist das Konzept der subjektiven Information keine Verwässerung des Informationsbegriffes, sondern umgekehrt das Konzept der objektiven Information eine Fiktion: „Objektive" Information ist dann nichts weiter als die subjektive Information, die eine Nachricht für den vermutlich besser Informierten enthält.)

Akte des Superierens, des Rekonstruierens und des Hypothesentestens setzen ein speicherndes Gedächtnis voraus, und sie erhöhen umgekehrt dessen „Leistung", sofern diese Leistung eben nicht in Termen subjektiver

Information bestimmt wird. Eine derartige Auffassung von Lernen macht einige Kontroversen zu kognitionswissenschaftlichen Problemen überflüssig. Indem sie etwa den Widerspruch zwischen informationaler Konstanz und chunk-Konstanz aufhebt: „From this point of view, cognitive mechanisms or strategies like chunking and semantic clustering are methods that enable us — in spite of strict limitations concerning the subjective information — to expand our capacity as defined in terms of objective information" (Fenk, 1985a, S. 362). Auch stellen sich diverse Befunde der letzten Jahre zum Thema „Wissen/Wissenserweiterung" aus dieser Perspektive bei weitem nicht so überraschend oder widersprüchlich dar wie etwa im Literaturüberblick von Wippich (1984). Und das Tempo des kindlichen Spracherwerbs wirkt angesichts der hohen Effektivität des aktiven Hypothesentestens bei der Aneignung von Regeln nicht mehr ganz so atemberaubend wie für Chomsky (1965), der zur Erklärung dieses Tempos sehr weitreichende Annahmen bezüglich der genetischen, sprachspezifischen Ausstattung treffen muß.

Die Auffassungen von v. Cube und von Frank hingegen sind, wie gesagt, mit unserer Auffassung — und auch untereinander — nicht vereinbar. Nach v. Cube (1967, S. 214) besagt das „Prinzip der Redundanzerzeugung", daß der lernfähige Organismus bestrebt ist, „die objektive Information der Außenwelt auf verschiedene Weise subjektiv zu verringern". Und für Frank (1971, S. 188) „liegt folgende informationspsychologische Definition des Lernens nahe: Ein System lernt etwas über seine Umwelt, indem es denjenigen Anteil an subjektiver Information über diese Umwelt verringert, der für ein möglichst zweckmäßiges Verhalten in dieser Umwelt verwertbar ist, aber dazu erst noch aufgenommen werden muß". v. Cube spricht also von der Reduktion objektiver Information, Frank dagegen — so wie wir — von der Reduktion subjektiver Information. Aber für Frank ist subjektive Information etwas, das ein Lernender aufnehmen kann; dies verträgt sich nicht mit unserer Auffassung der subjektiven Information als bloße Relation zwischen Nachricht und Subjekt bzw. als Inkongruenz zwischen Nachricht und Erwartung.

Natürlich kann man die „reduzierte Unsicherheit" gleichsetzen mit der durch eine Theorie übertragenen „positiven" Information (z.B. Salmon, 1971, S. 14). So dürfte Frank es aber nicht gemeint haben; es wäre ja (vor allem innerhalb einer Definition) unsinnig, die Reduktion der (negativen?) Information in die Gegenwart und die Aufnahme der (positiven?) Information in die Zukunft zu legen. Auch wird bei Frank die Information nicht durch eine — subjektive oder intersubjektive — Theorie „übertragen", sondern durch den „Kanal" oder die Kanäle des kognitiven Apparates. Seine Definition von Kanal: „Man bezeichnet mit ‚Kanal' jedes Medium, das eine Nachricht von einem Sender zu einem Empfänger zu übertragen gestattet,

unabhängig davon, ob die hierbei zu überbrückende Distanz vorwiegend räumlicher (‚Raumkanal', d.h. Übertragungskanal) oder vorwiegend zeitlicher (‚Zeitkanal', d.h. Gedächtnis oder Speicher) Art ist." (Frank, 1971, S. 227)

Laut Herrmann (1966, S. 20) aber hat man schon in den 50er Jahren die Einsicht gewonnen, „daß ein einfaches Kanalmodell mit festen Kapazitätsgrenzen zur Beschreibung des menschlichen Verhaltens nicht ausreicht", daß der Mensch also „nicht nur ein Informationsüberträger, sondern ein nach überaus komplexen Prinzipien arbeitender Informationsverarbeiter ist". Gleichwohl blieben die meisten Studien zur Informationshypothese dem nachrichtentechnischen Konzept der limitierten Kanalkapazität verhaftet. Ganz abgesehen von der Frage, ob das zugrundeliegende Menschenbild nicht zu „mechanistisch" sei (vgl. Herzog, 1984), resultieren daraus inadäquate Operationalisierungen. Denn untersucht wird in Experimenten dieser Art nicht die — nur indirekt, z.B. im Ratetest erschließbare — Transinformation zwischen internem Modell und externem Geschehen, sondern zwischen zwei Sorten von externem Geschehen, nämlich Input und Output. Zugegeben: Das Rateverhalten im Ratespielverfahren ist auch „Output"; aber eben nicht als jener Rest eines Inputs, welcher den verrauschten Kanal Mensch erfolgreich passiert hat, sondern als etwas, das erstens der „Kanal" (?) aktiv kreiert, und das zweitens — sehr zum Unterschied zu Reproduktions- und Wiedererkennenstets — direkte Rückschlüsse auf die Disparität zwischen internem Modell und externem Geschehen zuläßt.

Die neue Informationshypothese und das Prinzip der Operationalisierung

Die neue Informationshypothese besagt: Der zu erzielende Wissenszuwachs erweist sich als invariant gegenüber dem Vorwissen des jeweiligen Lerners, wenn Vorwissen und Wissenszuwachs in Einheiten subjektiver Information gemessen werden und der Zeitfaktor entsprechende Berücksichtigung findet.

Das Ausmaß der Wissenserweiterung bzw. des Lerngewinnes kann, wie im Zusammenhang mit einer ähnlichen Fragestellung vorgeschlagen (Fenk, 1980), durch Wiederholung eines Ratespiels bestimmt werden. Die „Einsparung" an Ratefehlern verrät uns, wieviel subjektive Unsicherheit (in bit) im Vergleich zu einem vorausgegangenen Ratedurchgang abgebaut werden konnte.

Aus dem Umstand, daß der Fortschritt im Laufe von Wiederholungen ein- und derselben Tätigkeit gemessen wird, ergeben sich zwei leistungstheoretisch interessante Vorteile gegenüber herkömmlichen Reproduktions- und Rekognitionstests: Zum einen wird das Vorwissen einkalkuliert,

welches jeder Lerner — wenn auch in unterschiedlich hohem Ausmaß — in jede experimentelle Situation mitbringt; der Lerneffekt wird in Form einer Veränderungsmessung erfaßt. Zum anderen werden keine Lerneffekte bezüglich der als relevant definierten Nachricht „übersehen". (Was man von Reproduktions- und Wiedererkennungstests nicht behaupten kann: Wer nicht mehr imstande ist, einen Satz frei zu reproduzieren, könnte vielleicht doch einige Wörter daraus wiedergeben, oder Synonyme zu diesen; oder er könnte, wenn er die Wörter erneut vorgegeben bekommt, noch Angaben zu ihrer Reihenfolge im Satz machen, oder sie innerhalb von Vexiermaterial wiedererkennen, oder hätte es doch vielleicht gekonnt, wären die Distraktoren in ihrer Bedeutung oder Form weniger ähnlich gewesen, etc. etc.)

Die Untersuchung unserer Frage erfordert nun eine Vergleichsmöglichkeit zwischen dem informationalen Lerngewinn ein- und desselben Probanden(kollektivs) einmal bei hohem, ein anderes Mal bei niederem Ausgangsniveau. Laut Informationshypothese müßte der Lerngewinn in beiden Fällen etwa gleich groß, also konstant sein; allenfalls feststellbare Schwankungen dürften sich zumindest nicht als von der Schwankung des Ausgangsniveaus abhängig erweisen.

Das Ausgangsniveau läßt sich aber, weil subjektive Information sich eben erst in der Auseinandersetzung eines Subjekts mit einer Nachricht konstituiert, experimentell nicht direkt manipulieren, sondern nur mittelbar über Maßnahmen (Auswahl, Veränderung) hinsichtlich der zu rekonstruierenden Nachricht, von denen eine Abstufung des Ausgangsniveaus erwartet werden kann.

Vorgeben ließe sich allenfalls die für das Ratespiel oder für einen „Rateplatz" zur Verfügung stehende Zeit. Dieses Zeitbudget wird dann aber nicht voll bzw. in unterschiedlichem Maße ausgenutzt, und in anderen Fällen müßte der Experimentator mit unvollständig erratenen Sequenzen vorliebnehmen. Aus diesem Grund sollte beim Ausloten des apostrophierten individuellen Leistungsplafonds auch der Zeitaufwand den individuellen Strategien überlassen werden.

Dies bringt aber auch Probleme mit sich. Wie soll man nun die beiden aus dem Ratespiel resultierenden Arten von Daten — Informationsgehalt und Zeitaufwand — zueinander in Beziehung setzen? Zwei Möglichkeiten bieten sich vor allem an:

Operationalisierung 1: Man stellt sich die Frage, innerhalb welcher Zeit bei der Wiederholung des Ratespiels (trial 2) wieviel der in trial 1 diagnostizierten Unsicherheit reduziert wird. Die Informationshypothese lautet dann:

$$\frac{i(1) - i(2)}{t(2)} \quad \text{variiert nicht in Abhängigkeit von } i(1)$$

i(1) = die in trial 1 ermittelte Unsicherheit (in bit) = Ausgangsniveau

i(2) = die in trial 2 ermittelte Unsicherheit
i(1) − i(2) = die Informationsreduktion (in bit) = Verbesserung
t(2) = die Ratezeit in trial 2 (in sec)

In dieser Operationalisierung wird die Formulierung von der — konstanten (?) — „Informationsreduktion pro Zeiteinheit" und die in ihr enthaltene Trennung des Arbeits-Aspektes vom zeitlichen Aspekt sozusagen wortgetreu in eine Formel übersetzt. Daß in diese Formel nur die Ratezeit von trial 2 eingeht, „may seem strange at first sight, but is in accordance with our question within what time the level of knowledge can be expanded (or the uncertainty reduced) to what extent beyond the level (of relevant ‚foreknowledge') the subjects have achieved pre-experimentally and in trial 1." (Fenk, 1985a, 368f.)

Operationalisierung 2: Man kann die Position beziehen, daß der Zeitaufwand notwendig in jede Bestimmung der Rateleistung bzw. des Wissensstandes eingeht; ein höheres Vorwissen (trial 1) ist aus einer geringeren Ratefehlerzahl nur dann abzulesen, wenn die geringere Fehlerzahl bei gleichem oder geringerem Zeitaufwand erzielt wurde. Aus dieser Überlegung heraus wird man i(1) und i(2) mit dem jeweiligen Zeitaufwand in Beziehung setzen, und zwar multiplikativ. Die Produktbildung basiert auf einer plausiblen Annahme: Je höher der getriebene Zeitaufwand, um so geringer die Fehlerzahl, und vice versa. (Hier soll ein Resultat der ersten Auswertungen vorweggenommen werden: Der ermittelte subjektive Informationsgehalt korrelierte tatsächlich hoch negativ mit dem Zeitaufwand. Vgl. Fenk 1984, 1985a.) Die Auswertung wird sich auf Operationalisierung 2 konzentrieren; diese entfernt sich zwar noch weiter als Operationalisierung 1 von der ursprünglichen, „alten" Informationshypothese, ist aber leistungstheoretisch besser begründbar. Die Informationshypothese lautet dann:

Die Differenz $i(1) \times t(1) - i(2) \times t(2)$ variiert nicht in Abhängigkeit von $i(1) \times t(1)$.
$i(1) \times t(1)$ = der in trial 1 ermittelte Geläufigkeitsgrad = Ausgangsniveau
$i(2) \times t(2)$ = der in trial 2 ermittelte Geläufigkeitsgrad
$i(1) \times t(1) - i(2) \times t(2)$ = Zunahme an Geläufigkeit = Wissenserweiterung = die Verbesserung

Das Problem der Null-Hypothese

Ein Manko hat die neue Informationshypothese mit alten Varianten gemeinsam. Sie ist aus forschungslogischer Sicht eine Null-Hypothese. (Keine Unterschiede des Lerngewinns in Abhängigkeit vom Ausgangsniveau!)

Welche Folge hat das für die Forschungsstrategie und für die Bewertung einschlägiger Resultate?

Man kann nur versuchen, durch möglichst taugliche Maßnahmen der Versuchsplanung (und Auswertung) eine Situation zu schaffen, in der die für richtig gehaltene Informationshypothese dadurch signifikant scheitert, daß die Konkurrenzhypothese („der Lerngewinn variiert in Abhängigkeit vom Ausgangsniveau") signifikant bestätigt wird. Eine Möglichkeit bestünde im Prinzip darin, den empirischen Gehalt der Nullhypothese — also ihre Chance zu scheitern — in die Höhe zu schrauben, indem man sehr enge Grenzen festsetzt, innerhalb derer man Schwankungen des Lerngewinns toleriert, innerhalb derer man also die geforderte „Konstanz" als unverletzt ansieht. Solche Grenzen sind aber aus unserer Theorie nicht ableitbar, der Ort der Grenzziehung wäre demnach sehr beliebig.

Unsere Situation ist also vergleichbar mit der von zwei Einsatzleitern eines Suchkommandos, das ein Waldstück nach einem Verbrecher durchkämmt. Einsatzleiter A glaubt, daß sich der Gesuchte nicht im Wald aufhält (= Nullhypothese), B glaubt, daß er sich dort aufhält. Wer definitiv recht hat, erfährt man nur, wenn der Gesuchte zur fraglichen Zeit im fraglichen Waldstück gefunden wird. Im umgekehrten Fall einer erfolglosen Suche kann A nicht behaupten, daß der Gesuchte nicht dort ist oder war. (Das würde unfehlbare Such- bzw. Diagnose-Methoden voraussetzen.) Aber seine Nullhypothese hat gegenüber dem Zeitpunkt vor der Suche an Gewicht gewonnen. Und sie gewinnt um so mehr an Gewicht, je „geeigneter" (systematischer, methodischer, gründlicher) die Suchbemühungen sind.

Versuchsanordnung

Unabhängige und abhängige Variablen: Da es auf die subjektive Information ankommt, entfällt die Notwendigkeit zur Konstruktion eines bezüglich der objektiven Information genau bekannten Lernmaterials. Das intraindividuell abgestufte Ausgangsniveau sollte durch zwei einfache Manipulationen an den verwendeten Lückentexten erreicht werden: Erstens wurde die Dichte der Aufeinanderfolge von Lücken variiert. (Es fehlte jedes achte, jedes vierte, jedes zweite, oder überhaupt jedes Wort.) Und zweitens variierte die Länge der mehr oder weniger unvollständigen und vom Ratenden fortzusetzenden Textfragmente, also die Dichte der Aufeinanderfolge jener Lücken, die er im Ratetext rekonstruieren mußte. Je unvollständiger der Text von vornherein ist und je unvollständiger er während des Ratens bleibt, um so schwieriger ist er vermutlich fortzusetzen, um so höher also voraussichtlich die subjektive Information, welche die Rateplätze für die Probanden enthalten. Die Anzahl der Rateplätze steigt natürlich ebenfalls,

wenn in einem Lückentext vorgegebenen Umfanges die zu füllenden Lükken dichter liegen. Aus der Kombination des in vier Stufen variierten Anteils der von vorneherein fehlenden Wörter mit dem verschieden hohen Anteil der zu füllenden Lücken resultierten neun verschiedene Lückentexte (Tab. 1).

Tabelle 1
Das Versuchsschema

VG	Lücken-Text	A	Fehlende Wörter	Zu ergänzende Wörter	VB
1	1	H	Jedes	Jedes zweite	I
	2	N	Jedes zweite	Jedes zweite	
2	3	H	Jedes	Jedes vierte	II
	4	N	Jedes zweite	Jedes vierte	
3	4	H	Jedes zweite	Jedes vierte	
	5	N	Jedes vierte	Jedes vierte	
4	6	H	Jedes	Jedes achte	III
	7	N	Jedes zweite	Jedes achte	
5	7	H	Jedes zweite	Jedes achte	
	8	N	Jedes vierte	Jedes achte	
6	8	H	Jedes vierte	Jedes achte	
	9	N	Jedes achte	Jedes achte	

VG = Versuchsgruppe VB = Versuchsbedingung
A = Grad der Unvollständigkeit (*H*ohes vs *N*iederes *A*usgangsniveau)

Die entscheidende unabhängige Variable — das Ausmaß der Unvollständigkeit des Textes bzw. der Grad der Kontextzerstörung — wurde also intraindividuell, soll heißen innerhalb je eines Probanden(kollektivs), abgestuft (siehe Tab. 1). Dagegen variierte die ebenfalls unabhängige Größe „Rateplatzzahl" nur zwischen „Versuchsbedingungen (= VBn)", die von verschiedenen Personen realisiert wurden; innerhalb dieser VBn wurde sie konstant gehalten.

VB I: 40 Wörter, 278 Rateplätze
VB II: 20 Wörter, 152 Rateplätze
VB III: 10 Wörter, 90 Rateplätze

Abhängige Variablen waren die Zahl der Ratefehler n(f) und die Ratezeit t.

Versuchsablauf: Nach der Instruktion beginnt für jede einzelne Versuchsperson ein in 2 × 3 Aufgaben gegliederter Versuch.

A) *Erraten:* Auf dem Bildschirm eines Kleincomputers (Apple II) wächst das erste Fragment einer mehr oder weniger lückenhaften Version (= „Lückentext") einer Textprobe (z.B. P), wobei an die Stelle der Buchstaben fehlender Wörter Punkte gesetzt sind. Sobald der Rechner innehält — in Versuchsbedingung (VB) I vor dem zweiten, in VB II vor dem vierten und in VB III vor dem achten Wort — ist die Versuchsperson an der Reihe: Sie soll das nächste Wort Buchstabe für Buchstabe erraten und ins Terminal eingeben. Pro Rateplatz wird nur einmal geraten (Weltner-Verfahren, 1970, S. 55—71), nach jeder Eingabe erscheint — unabhängig davon, ob die Eingabe richtig oder falsch war — der richtige Buchstabe. Sobald das zu erratende Wort komplett ist, wird vom Rechner das zweite Textfragment vorgesetzt, und so fort.

RT) *Reaktions-Tippen:* Als zweite Aufgabe folgen 100 sec. „Zwischentätigkeit". Auf dem Bildschirm tauchen, in einer vom Zufallsgenerator produzierten Reihenfolge, einzelne Buchstaben auf, und diese sollen möglichst schnell ins Terminal eingegeben werden. Durch die Eingabe wird die Präsentation des nächstfolgenden Buchstaben ausgelöst. Registriert wird die Zahl der falschen und der richtigen Eingaben. Das Reaktionstippen soll die Daten für eine Schätzung jener Zeit liefern, die (bei im Tippen unterschiedlich geübten Versuchspersonen) zwischen Absicht und Realisation der Eingabe eines bestimmten Buchstabens verstreicht (= Tippzeit).

B) *Wieder-Erraten:* Als dritte Aufgabe wird das unter „Erraten" beschriebene Ratespiel — an derselben Textprobe mit denselben Textlücken und derselben Rateplatzzahl — wiederholt.

Als vierte bis sechste Aufgabe wird dieselbe Prozedur — also A, RT, B — an der jeweils anderen Textprobe (z.B. K) durchgespielt, und zwar mit halb oder doppelt so vielen Wortlücken wie in der ersten Textprobe (vgl. Tab. 1 u. 2). Erst damit sind die Voraussetzungen zur Prüfung unserer Hypothese gegeben.

Textmaterial: Als Lernmaterial wurden eher „schwierige" Textstellen gewählt, von denen man erwarten durfte, daß sie auch beim Wieder-Erraten noch genügend bzw. genügend variable Information für den Ratenden enthalten. Schwieriger war eine zweite Bedingung zu erfüllen: Die beiden Texte durften sich hinsichtlich der Anzahl fehlender Buchstaben nicht unterscheiden, wenn jedes zweite — und damit jedes vierte oder achte — Wort weggelassen wurde. Um diese Bedingung zu erfüllen, mußte eine der schließlich ausgewählten Textproben, nämlich K, geringfügig abgeändert werden. Die beiden Textproben P (aus Popper, 1976, S. 225) und K (aus Klix, 1971, S. 713) sind im Anhang wiedergegeben, wobei die jeweils

zweiten, vierten und achten Wörter einfach, zweifach und dreifach unterstrichen wurden.

Versuchsschema, Versuchsbedingungen (VBn) und Versuchsgruppen (VGn): Aus einem 2 × 2 Design (2 Textproben; 2 Stufen der Unvollständigkeit) resultieren 4 Varianten, in denen der oben beschriebene Versuch ablaufen kann (Tab. 2). Je eine dieser 4 Möglichkeiten wird von einer Versuchsperson einer vierköpfigen „Versuchsgruppe" — eigentlich waren es ja Einzelversuche — realisiert. Das in Tabelle 1 gezeigte Schema des Gesamt-Experiments erfordert 6 solcher Vierergruppen, doch wurde, um die Zuverlässigkeit der Ergebnisse zu erhöhen, das gesamte Schema doppelt abgedeckt (also 2 × 6 VGn mit je 4 Vpn).

Tabelle 2

Das Versuchsschema pro Versuchsgruppe. (Links die vier Kombinationsmöglichkeiten aus 2 Texten und je 2 Stufen der Unvollständigkeit.)

		Unvollständigkeit									
		H	N	Vp	A1	RT1	B1	A2	RT2	B2
Text	P	PH	PN	1	PH	RT1	PH	KN	RT2	KN
				2	KN	RT1	KN	PH	RT2	PH
	K	KH	KN	3	KH	RT1	KH	PN	RT2	PN
				4	PN	RT1	PN	KH	RT2	KH

Instruktion: Sie umfaßt eine schriftliche Erklärung, deren genauer Wortlaut im Anhang wiedergegeben ist, sowie ein kurzes Trainingsprogramm und die Gelegenheit zu Rückfragen. Die schriftliche Erklärung enthält das Versprechen, daß jede Vp, unabhängig von ihrer Leistung, 50 S erhält, und daß — bei jeder einzelnen der 6 Aufgaben — die zwei besten jeder vierköpfigen VG eine Prämie von 50 S zusätzlich bekommen. Die Prämien sollten die Vpn dazu bewegen, in allen 6 Aufgaben nahe am persönlichen Leistungsplafond zu agieren. Im Leistungskriterium n(r)/n(f).t sollte ein möglichst hoher Wert erreicht werden: n(r) bzw. n(f) = Anzahl der Fälle, in denen ein Buchstabe richtig oder falsch getippt wurde; t ist die — beim Reaktionstippen konstante, bei A und B variable — Lösungszeit pro Aufgabe. Dieses Kriterium läßt dem Ratenden viel Freiraum bezüglich der Art und Weise, in welcher er im „Austausch" zwischen Geschwindigkeit (Ratezeit pro Zeichen) und Genauigkeit (relative Trefferfrequenz beim Raten) einen maximalen Score anpeilt.

Versuchspersonen: Bei den insgesamt 48 Vpn von Exp. 1 handelte es sich um 29 weibliche und 19 männliche Studenten verschiedener Fächer der

Universität Klagenfurt. Das Alter der Probanden bewegte sich, bei einem Mittelwert von 22,8 Jahren, zwischen 20 und 29 Jahren.

Präzisierung der Vorhersagen

Die neue Informationshypothese behauptet, daß die informational quantifizierte Verbesserung unabhängig ist vom ebenfalls informational quantifizierten Ausgangsniveau. Wenn sie und die von uns gewählte Form („Operationalisierung 2") der Berücksichtigung des Zeitaufwandes berechtigt sind, so müßte das oben beschriebene Design zu folgenden Ergebnissen führen.

1. Mittelwertsvergleiche: Es wird intraindividuell keine deutlichen (keine signifikanten) Unterschiede im Ausmaß der Verbesserung (V) geben; und zwar auch dann nicht, wenn das Ausgangsniveau intraindividuell sehr deutlich (signifikant) wechselt. (Hohes vs Niederes Ausgangsniveau)

$$V(H) \approx V(N) \; ; \; /V(H) - V(N)/ \approx 0$$

Dies ist die zentrale Vorhersage, auf ihre Prüfung hin ist das Design angelegt.

Die nachstehenden Vorhersagen beziehen sich auf flankierende Auswertungen, deren Ausgang wesentlich von den Streuungen über die Probanden(kollektive) hinweg bestimmt wird und nicht, wie bei der zentralen Vorhersage, durch die Streuung intraindividueller Differenzwerte. (Auf weitere Einschränkungen der Interpretierbarkeit dieser zusätzlichen Auswertungen wird bei der Darstellung der jeweiligen Resultate hingewiesen.)

2. Korrelationen:

2.1 Die Korrelation zwischen Ausgangsniveau und Verbesserung wird nicht wesentlich (nicht signifikant) von Null abweichen.

$$r_{AV} \approx 0$$

2.2 Der Zeitaufwand beim Raten wird (signifikant) negativ mit dem im Ratetest ermittelten Informationsgehalt der Nachricht korrelieren.

$$r_{bit\ sec} = -0,$$

3. Variabilitätsvergleiche: Das Ausmaß der Verbesserung wird eine geringere Variabilität aufweisen als das Ausgangsniveau. Für den Variabilitätskoeffizienten soll also gelten:

$$V < A$$

Und für die Vorhersagen 1, 2.1 und 3 gilt, daß abweichende Ergebnisse allenfalls dann noch tolerabel sind, wenn Ausgangsniveau und Verbesse-

rung nur in bit gemessen werden. Ergeben sich auch bei Berücksichtigung der Ratedauer „t" oder gar der geschätzten Netto-Denkzeit „D" signifikante Abweichungen, so ist die neue Informationshypothese als widerlegt anzusehen.

Ergebnisse

Anmerkungen zur Auswertung und zum Verständnis der Ergebnis-Tabellen

Berechnung der subjektiven Information: Aus den protokollierten Ratefehlerzahlen wurde nach der Weltner-Formel für buchstabenweises Raten (Weltner, 1970, S. 70) der subjektive Informationsgehalt bestimmt.

$$\text{subj. Inf.} = 0{,}27 + 4{,}93.C \qquad \text{mit } C = n(f)/n$$

Von Interesse war für unseren Zweck nicht eine Schätzung des Informationsgehaltes der gesamten Textprobe, sondern der subjektive Informationsgehalt pro Rateplatz; dieser sollte dann höher ausfallen, wenn der bis zum jeweiligen Rateplatz erfahrbare Kontext stärker zerstört war.

Versuchsgruppen (VGn) und Versuchspersonen (Vpn): Das Ausgangsniveau wurde, wie von vorneherein vermutet und im Versuchsschema pro Versuchsgruppe (Tab. 2) berücksichtigt, nicht nur vom Grad der Unvollständigkeit der Texte bestimmt. Vor allem die Reihenfolge der Texte und die Art des Texts war mit ausschlaggebend: Der zuerst präsentierte Text war im Durchschnitt „schwieriger" als der zweite, und Text K war im Durchschnitt „schwieriger" als Text P. Statistische Auswertungen wurden auch auf der Ebene einzelner Probanden durchgeführt; auf dieser Ebene war die Ausbalancierung der genannten Einflüsse nicht möglich, was aber keine nachteiligen Folgen für die Prüfung der zentralen Vorhersage hat.

Das Ausgangsniveau (H vs N, K vs P, Hreal vs Nreal): Der Unterschied zwischen den Ausgangsniveaus bei hoher und niedriger Unvollständigkeit der Texte lag, wie Tab. 4 zeigt, fast immer in der angepeilten Richtung, d.h. A(H) > A(N). Bei der Information i ohne Berücksichtigung der Zeit (t o. D) zeigt sich dieses Gefälle in ausnahmslos allen Versuchsgruppen. Eine vorweggenommene Interpretation des Umstands, daß dieses Gefälle unerwartet klein ausfiel und in einer Versuchsgruppe (VG 4) bei Berücksichtigung des Zeitaufwandes sogar umkippte: Die Vpn nutzten zum Niedrighalten der Ratefehlerzahlen bevorzugt die kontextuellen Restriktionen (möglicher Buchstaben) innerhalb der Wörter und nur sehr wenig die Restriktionen (möglicher Wörter) innerhalb des umfassenderen Satzkontextes. Und der subjektive Informationsgehalt hängt eben nicht nur vom Grad der Unvollständigkeit, sondern auch von der internen Redundanz des jeweiligen Wor-

tes ab sowie vom Beitrag zur Rekonstruierbarkeit des Kontextes, welchen die im konkreten Fall verbleibenden bzw. weggelassenen Wörter zu leisten imstande sind bzw. gewesen wären.

Bei unserem möglichst systematischen Bemühen um ein Scheitern der Informationshypothese lag der Gedanke nahe, aus der Not — relativ geringes Gefälle A(H) > A(N) — eine Tugend zu machen und auch den stärker unterschiedlichen Schwierigkeitsgrad der Texte P und K als experimentelle Variable zu behandeln; beim Vergleich zwischen P und K innerhalb der Versuchsgruppen sind die restlichen Variablen (H vs N; Reihenfolge der Texte) wiederum ausbalanciert. Eine zweite Maßnahme bestand darin, das Ausmaß der Verbesserung bei real höherem vs niedrigerem Ausgangsniveau zu vergleichen: Bei dieser Datenkonfiguration läßt sich zwar nicht mehr angeben, wodurch das Gefälle des Ausgangsniveaus zustandekommt, doch ist mit dem nun besonders deutlichen Gefälle des Ausgangsniveaus die Voraussetzung für eine besonders strenge Prüfung unserer Hypothese geschaffen.

Schätzung der Netto-Denkzeit „D": Eine Schätzung der Zeit „D", die eine Vp aufwendet, bis sie sich für eine bestimmte Eingabe entscheidet, wurde auf folgende Weise versucht: Von der gesamten Ratezeit t pro Rateplatz wurde die im Reaktionstippen ermittelte „Tippzeit" pro Rateplatz subtrahiert, also jene Zeit, welche der jeweilige Proband zur Realisierung einer Eingabe braucht, für die er sich bereits entschieden hat (D = t − Tippzeit). Der Schätzung der Tippzeit wurde nicht die Anzahl der richtigen Eingaben beim Reaktionstippen zugrundegelegt, sondern der Mittelwert aus den richtigen und den insgesamt getätigten Eingaben — und zwar aus der Überlegung heraus, daß sich die Vpn auch im Ratetest manchmal richtig entschieden, dann aber „vertippt" haben mochten. (Zwei Versuchspersonen, nämlich die Vpn 7 und 37, hatten in RT 2 häufig nicht den zuletzt erschienenen, sondern den jeweils vorletzten Buchstaben eingetippt. Bei ihnen wurden die Leistungsverhältnisse von RT 1 in die Schätzung einbezogen; die Scores wurden von 20 auf 88 bzw. von 44 auf 69 „korrigiert". Die entsprechenden Positionen sind im Anhang kursiv gesetzt.)

Bei zwei anderen Versuchspersonen (Vpn 13 und 25, vgl. Anhang) ergab die Berechnung von D einen negativen Wert, und zwar jeweils bei Text K, trial 2. (Diese Fälle wurden bei der Auswertung auf Ebene der Einzelprobanden weggelassen; daher auch in einigen Tabellen Fußnoten bezüglich einer geringeren Anzahl verrechneter Datenpaare.) Mögliche Gründe: Bei einem der beiden Probanden vermerkt es das Versuchsprotokoll, und bei dem anderen ist es nicht auszuschließen, daß im Ratetest mehrmals die Taste REPT des Kleinrechners Apple II betätigt wurde, wodurch der gleichzeitig gedrückte Buchstabe in sehr rascher Folge immer wieder eingegeben wird und die Tippzeit im Ratetest sehr kurz ausfällt. Möglicherweise

ist aber auch das Reaktionstippen ein zu wenig valider Test für die Tippzeit: Denn gerade beim geübten Stenotypisten wird die Tippzeit bei den redundanteren Texten des eigentlich interessierenden Ratespiels kürzer ausfallen als bei der im Reaktionstippen verwendeten Zufallsfolge von Buchstaben. Und schließlich könnte ein Übungsgewinn durch das Reaktionstippen die Tippzeiten im nachfolgenden trial 2 verkürzt haben.

Das geometrische Mittel aus subjektiver Information und Zeitaufwand als Leistungsmaß: Der wechselnde Zeitaufwand (t bzw. D in sec pro Rateplatz) wird durch Produktbildung mit der subjektiven Information (i in bit pro Rateplatz) einkalkuliert. Aus diesem Produkt noch die Wurzel zu ziehen, bringt zwei Vorteile: Es erleichtert den Vergleich der zusammenfassend dargestellten Ergebnisse, weil das geometrische Mittel aus zwei Zahlenwerten in der Größenordnung eben dieser Werte angesiedelt ist. Und zweitens wird dadurch die statistische Prüfung der Hypothese noch strenger — das geometrische Mittel streut weniger als die Produkte, „hypothesenfeindliche" Unterschiede in der Leistungsverbesserung /V(H) — V(N)/ fallen eher signifikant aus. (Bei der flankierenden korrelationsstatistischen Auswertung 2.1 muß allerdings das Gegenteil in Kauf genommen werden.)

Bei der Auswertung auf Ebene der Versuchsgruppen wurde zuerst das arithmetische Mittel — über 8 Vpn — für bit und für sec berechnet, und hieraus das geometrische Mittel aus bit und sec. Aus dieser Berechnung resultieren aus rein mathematischen Gründen etwas höhere Werte als aus einer Aufsummierung der pro Vp kalkulierten geometrischen Mittelwerte.

Resultate zur zentralen Vorhersage 1: V (H) ≈ V (N)

In allen Versuchsgruppen kam es in trial 2 zu einer Verbesserung im Vergleich zu trial 1, und zwar nicht nur durch eine Einsparung an subjektiver Unsicherheit in bit — keine negativen Vorzeichen der Verbesserung V (= A—B) in den zwei Spalten unter i, Tabelle 4 — sondern auch durch eine Einsparung im Zeitaufwand (Tab. 3).

Wie Tab. 4 zeigt, ist das Ausmaß der Verbesserung — so wie das Ausgangsniveau — meist bei hochgradig zerstörtem Kontext „H" größer als bei „N". Die Ausnahmen: VG 3, aber nur dann, wenn die Zeit unberücksichtigt bleibt; und bei Berücksichtigung der Zeit VG 4, also jene Versuchsgruppe, in der auch das Ausgangsniveau der Leistung bei H geringer war als bei N. Mit einem Wort: Wird der Zeitaufwand einkalkuliert, so entfällt auf ein höheres Ausgangsniveau ausnahmslos eine stärkere Verbesserung, und auch wenn er nicht einkalkuliert wird, macht nur eine Versuchsgruppe eine Ausnahme.

Tabelle 5 gibt Auskunft über die Größe und statistische Bedeutsamkeit

Tabelle 3
Die Ratezeit t und die „Denkzeit D" pro Rateplatz in trial 1 (A) und trial 2 (B)

			N			H		
Maß	VB	VG	A	B	(A—B)	A	B	(A—B)
t	I	1	2.665	1.928	0.737	2.663	1.878	0.785
	II	2	3.078	2.330	0.748	2.977	2.100	0.877
		3	3.059	2.396	0.664	3.360	2.282	1.078
	III	4	3.722	2.307	1.415	2.993	2.113	0.881
		5	3.571	2.529	1.042	3.494	2.454	1.040
		6	3.601	2.593	1.008	4.671	2.864	1.807
D	I	1	1.384	0.647	0.737	1.405	0.620	0.785
	II	2	1.722	0.974	0.748	1.628	0.751	0.877
		3	1.680	1.016	0.664	2.048	0.970	1.078
	III	4	2.352	0.937	1.415	1.590	0.709	0.881
		5	2.107	1.065	1.042	2.108	1.068	1.040
		6	2.170	1.162	1.008	3.215	1.408	1.807

Tabelle 4
Das Ausgangsniveau (A) und die Verbesserung (V) in den 6 Versuchsgruppen. (Jeweils pro Rateplatz.)

VG	Lückentexte (H ; N)		i		$\sqrt{i \cdot t}$		$\sqrt{i \cdot D}$	
			H	N	H	N	H	N
1	(1 ; 2)	A	2.823	2.409	2.743	2.534	1.993	1.826
		V	0.563	0.286	0.682	0.511	0.808	0.654
2	(3 ; 4)	A	2.597	2.423	2.781	2.731	2.056	2.043
		V	0.616	0.369	0.741	0.543	0.836	0.628
3	(4 ; 5)	A	2.248	2.184	2.749	2.585	2.147	1.915
		V	0.414	0.450	0.702	0.547	0.812	0.588
4	(6 ; 7)	A	2.523	2.413	2.748	2.997	2.003	2.383
		V	0.535	0.411	0.698	0.848	0.815	1.013
5	(7 ; 8)	A	2.331	2.098	2.854	2.737	2.217	2.103
		V	0.603	0.466	0.794	0.705	0.858	0.784
6	(8 ; 9)	A	1.934	1.927	3.005	2.634	2.493	2.045
		V	0.698	0.692	1.124	0.845	1.175	0.847

Tabelle 5

Die Mittelwertsunterschiede pro Zeichen — bez. Ausgangsniveau „A" und Verbesserung „V" — zwischen den Versuchsbedingungen H vs N, zwischen den Texten K vs P und zwischen den Daten bei einem real höheren vs niedrigeren Ausgangsniveau.

				i				$\sqrt{i \cdot t}$				$\sqrt{i \cdot D}$			
Diff.		VB	n	A	p<	V	p<	A	p<	V	p<	A	p<	V	p<
	VGn		6	.168	5	.126	10	.110		.124	10	.099		.132	
H — N	Vpn	I	8*	.417	10	.277	5	.191		.135		.022		.076	
	Vpn	II	16	.120		.106		.090		.157	10	.092		.182	10
	Vpn	III	24*	.116		.089		.052		.044		.002		.039	
K — P	VGn		6	.347	1	—.043		.323	5	.019		.317	5	.036	
	VGn		6	.168	5	.126	10	.193	1	.174	1	.226	2	.198	1
H — N	Vpn	I	8*	.528	1	.122		.434	1	.322	1	.421	1	.394	1
(real)	Vpn	II	16	.379	1	.057		.434	1	.106		.489	1	.252	1
	Vpn	III	24*	.441	1	.011		.643	1	.306	2	.714	1	.397	1

*) In der Spalte $\sqrt{i \cdot D}$ sind die entsprechenden Datenpaare um je eines auf n = 7 bzw. n = 23 reduziert. (Erklärung im Text)

der Unterschiede zwischen H vs N und H(real) vs N(real). Gerade der letztgenannte Vergleich bringt die meisten signifikanten — und unserer Hypothese eindeutig widersprechenden — Unterschiede in der Verbesserung: V(Hreal) > V(Nreal). Aber sogar in Fällen, in denen die Abstufung des Ausgangsniveaus nicht signifikant ausfällt, entfällt auf H ein signifikant höherer Lerngewinn.

Etwas anders und sehr „hypothesenfreundlich" stellt sich der Vergleich K vs P dar: Trotz signifikanter Unterschiede im Ausgangsniveau (K > P) erweisen sich die Unterschiede zwischen den Texten bezüglich des Lerngewinns als denkbar gering und nicht signifikant. Die nach Versuchsgruppen getrennten Ergebnisse werden aus Platzgründen nicht in einer Tabelle (analog zu Tab. 4) dargestellt, lassen sich aber aus Abb. 2 einigermaßen ablesen: Die Unterschiede im Ausgangsniveau von K und P sind — vor allem ohne Berücksichtigung der Zeit — viel ausgeprägter als die zwischen H und N (vgl. mit Abb. 1). Das Ausgangsniveau liegt, unabhängig vom verwendeten

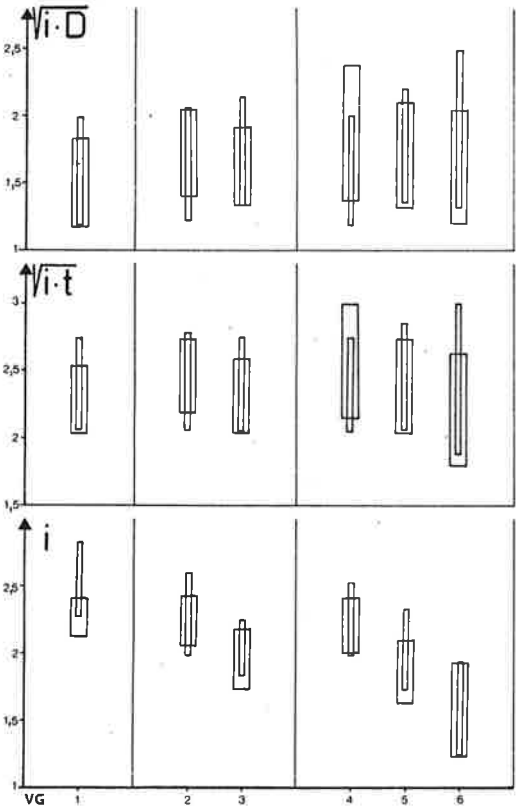

Abb. 1
Die Blöcke (dünn = H, dick = N) illustrieren für jede der 6 Versuchsgruppen das in trial 1 (oberes Blockende) und in trial 2 (unteres Blockende) erreichte Niveau in den Kriterien i, $\sqrt{i \cdot t}$ und $\sqrt{i \cdot D}$.

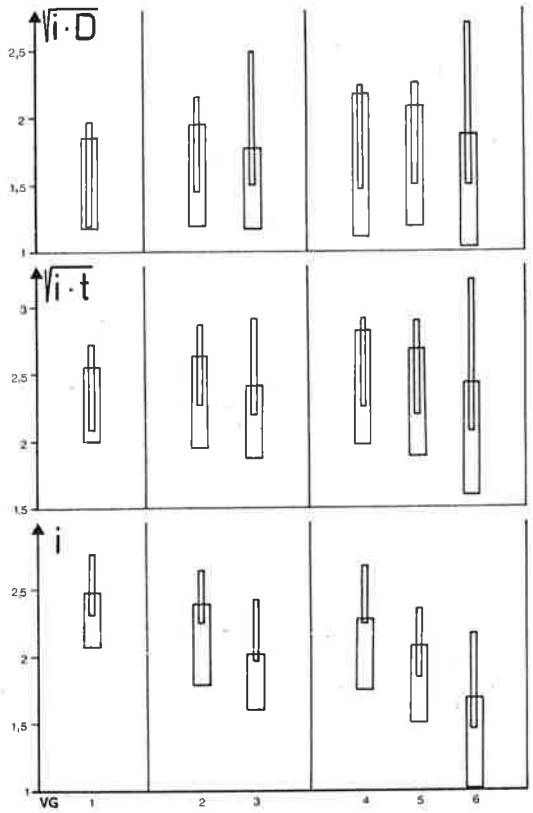

Abb. 2
Die Blöcke (dünn = K, dick = P) illustrieren für jede der 6 Versuchsgruppen das in trial 1 (oberes Blockende) und in trial 2 (unteres Blockende) erreichte Niveau in den Kriterien i, $\sqrt{i \cdot t}$ und $\sqrt{i \cdot D}$.

Maß, in jeder der sechs Versuchsgruppen in K höher als in P. Beim Lerngewinn hingegen ergibt sich wieder ein Unentschieden: In genau der Hälfte der Versuchsgruppen (VGn 2, 4 und 5) war die Verbesserung beim „leichten" Text P größer als beim „schweren" Text K.

Resultate zu den flankierenden Auswirkungen

Zur Vorhersage $r_{AV} \approx 0$: Die entsprechenden Ergebnisse sind in Tabelle 6 angeführt. Allerdings ist die Höhe der über die sechs Versuchsgruppen — und damit über drei verschiedene Rateplatzzahlen — hinweg berechneten Koeffizienten mit Vorsicht zu genießen; der Umstand, daß bei geringeren Rateplatzzahlen pro Rateplatz mehr gelernt wird, muß zu Verzerrungen geführt haben. Von dieser Einschränkung nicht betroffen sind die auf der Ebene der einzelnen Vpn einer Versuchsbedingung berechneten Koeffizien-

Tabelle 6

Die Korrelationen zwischen dem Ausgangsniveau „A" pro Zeichen und der Verbesserung „V" pro Zeichen.

			i				$\sqrt{i \cdot t}$				$\sqrt{i \cdot D}$			
		n	H	p<	N	p<	H	p<	N	p<	H	p<	N	p<
VGn		6	—.249		—.900	5	.983	0,1	.623		.923	1	.865	5
	I	8*	.56		—.216		.830	5	.603		.936	0,1	.803	5
Vpn	in II	16	.377		—.371		.736	1	.257		.754	0,1	.523	5
	III	24*	.011		—.087		.813	0,1	.675	0,1	.809	0,1	.768	0,1
			K		P		K		P		K		P	
VGn		6	—.831	5	—.535		.908	5	.550		.871	5	.925	1

* In der Spalte $\sqrt{i \cdot D}$ sind die entsprechenden Datenpaare um 1 auf n = 7 bzw. n = 23 reduziert. (Erklärung im Text)

ten sowie Vergleiche zwischen den Kriterien i, $\sqrt{i \cdot t}$ und $\sqrt{i \cdot D}$: Bei der bloßen Information i liegen die Koeffizienten nahe bei Null oder im negativen Bereich; bei Berücksichtigung der Zeit hingegen finden wir ausnahmslos positive Werte, wobei die statistisch bedeutsameren Werte auf H häufiger entfallen als auf N und auf $\sqrt{i \cdot D}$ häufiger als auf $\sqrt{i \cdot t}$. (Unserer Null-Hypothese zufolge dürften überhaupt keine signifikanten Korrelationen auftreten.) Die Relationen zwischen den Werten innerhalb eines jeden der drei Kriterien werden von dieser Verschiebung nur wenig berührt — je stärker negativ eine Position im Kriterium i, um so schwächer positiv ist sie im allgemeinen in den Kriterien $\sqrt{i \cdot t}$ und $\sqrt{i \cdot D}$.

Zur Vorhersage $r_{it} = -0,\ldots$, $r_{iD} = -0,\ldots$: Aufgrund möglicher Verzerrungen durch einen größeren Lerneffekt pro Rateplatz bei geringerer Rateplatzzahl gibt es auch hier Einschränkungen der Interpretierbarkeit. Diese betreffen jedoch nur die Größe der über die VGn berechneten Koeffizienten, und auch hier nur trial 2, also die Spalten B in Tab. 7. 38 der 40 Korrelationskoeffizienten tragen ein negatives Vorzeichen. Fast durchwegs gilt also, wie erwartet, daß hoher Informationsgehalt mit geringer Ratezeit einhergeht. Bei H und K sind die Koeffizienten höher als bei N und P. In H sind z.T. auch die über die Einzelpersonen je einer Versuchsbedingung berechneten Werte signifikant und — zum Unterschied zu N — meist höher als in trial 2.

Tabelle 7

Die Korrelationen zwischen der subjektiven Textinformation „i" in bit pro Zeichen und dem Zeitaufwand „t" beim Raten bzw. der geschätzten Netto-Denkzeit „D", jeweils in sec pro Zeichen.

				H				N			
Korr.		VB	n	A	p<	B	p<	A	p<	B	p<
i, t	VGn		6	—.947	1	—.994	0,1	—.455		—.809	10
i, D	VGn		6	—.943	1	—.978	0,1	—.379		—.701	
	Vpn	I	8	—.747	5	—.349		—.099		—.376	
i, t	Vpn	II	16	—.437	10	—.330		—.129		—.097	
	Vpn	III	24	—.103		—.151		.137		—.204	
	Vpn	I	8	—.735	5	—.375		—.166		—.502	
i, D	Vpn	II	16	—.476	10	—.250		—.001		—.134	
	Vpn	III	24	—.090		—.145		.123		—.216	
						K				P	
i, t	Vgn		6	—.938	1	—.934	1	—.600		—.973	1
i, D	VGn		6	—.929	1	—.945	1	—.578		—.948	1

Abb. 3 illustriert zu den signifikanten Korrelationen aus Tab. 7 die entsprechenden Regressionsgeraden, definiert durch den Schnittpunkt mit der y-Achse und durch die Steigung. Jede Gerade wurde um ihren Mittelwert in y (t bzw. D) gekippt; so läßt sich aus der Abbildung ersehen, wie nahe am Koordinatenschnitt die Geraden vorbeigehen, in welcher Näherung wir es also mit einem — in Wahrheit umgekehrten — Proportionalitätsverhältnis zwischen Informationsgehalt und Ratezeit zu tun haben. Tatsächlich scharen sich die Kreuzungspunkte zwischen Regressionsgeraden und xy-Koordinaten rund um den Koordinatenschnittpunkt — wenn auch nicht so dicht am Schnittpunkt gebündelt, wie es Abb. 3 vielleicht deshalb suggeriert, weil die Regressionsgeraden der optischen Unterscheidbarkeit zuliebe über die Grenzen der Punktwolken hinaus verlängert wurden. Jedenfalls darf von einer Tendenz zur umgekehrten Proportionalität von Zeitaufwand und subjektivem Informationsgehalt ausgegangen werden: Der aus dem Ratetest ermittelte subjektive Informationsgehalt der Nachricht steigt um etwa denselben Faktor, um den sich der Ratende (im Vergleich zu anderen Ratenden) schneller entscheidet.

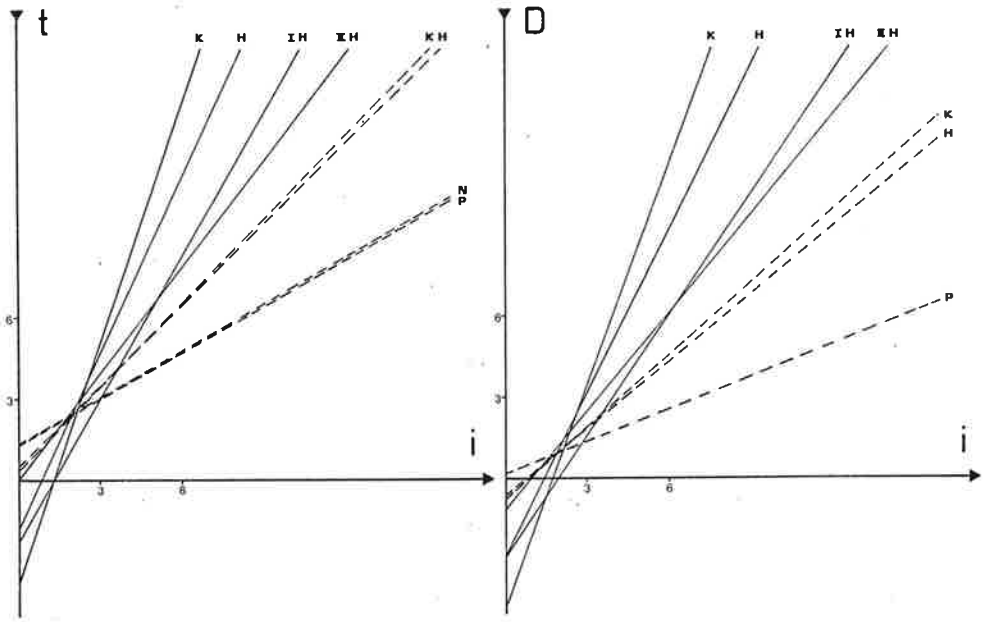

Abb. 3
Die den signifikanten Korrelationen (s. Tab. 7) zwischen i und t (links) bzw. D (rechts) entsprechenden Regressionsgeraden.

Zur Vorhersage 3, wonach der Variabilitätskoeffizient von V kleiner ist als jener von A: Die Berechnung des Variabilitätskoeffizienten ($= 100 \cdot \sigma/\mu$) trägt dem Umstand Rechnung, „daß durchschnittlich größere Maße auch größere mittlere Abweichungen zeigen werden..." (Mittenecker, 1979, S. 22). In unserem Fall ist das Maß V (= A — B) zwangsläufig kleiner als A. Die in Tab. 8 wiedergegebenen Resultate weisen, im Gegensatz zu unserer Vorhersage, der Verbesserung V die größte Variabilität zu. Die am häufigsten auffindbare Rangreihe heißt V > B > A, wobei sich unter den jeweils 36 Paarvergleichsmöglichkeiten nur wenige Ausnahmen finden:

V > B: Keine Ausnahme!
V > A: Nur drei Ausnahmen, und diese entfallen auf die bloße Information i.
B > A: Zwölf Ausnahmen; fünf entfallen auf $\sqrt{i \cdot t}$, sieben auf $\sqrt{i \cdot D}$.

Innerhalb der „besten" Leistungsschätzung $\sqrt{i \cdot D}$ (12 Paarvergleiche) sind also die „Ausnahmen" in der Überzahl, die Variabilität der Leistung ist hier in trial 2 geringer als in trial 1 und geringer als die der Leistungsverbesserung V (trial 1 — trial 2).

Eine zusätzliche Versuchsgruppe („Exp. 2")

Als sich herausstellte, daß die bisher pro VG gewählte Abstufung der Lückenhaftigkeit (in H doppelt so viele fehlende Wörter wie in N) für eine signifikante Abstufung des Ausgangsniveaus nicht immer ausreichte, wurde eine zusätzliche Versuchsgruppe gebildet. Wie in den einzelnen VGn aus „Exp. 1" wurde auch innerhalb der neuen VG die Rateplatzzahl konstantgehalten: Jedes achte Wort bzw. 90 Buchstaben war(en) zu erraten. (Dies entspricht VB III aus „Exp. 1".) Die Anzahl der Vpn wurde aber erhöht, und zwar von zwei Vierergruppen auf vier Vierergruppen; die 16 studentischen Probanden — 12 davon waren weiblich, 4 männlich — hatten ein Durchschnittsalter von 23,4 Jahren. Der wesentlichere Unterschied: Die Lückenhaftigkeit wurde breiter abgestuft; in H fehlte jedes zweite Wort, in N dagegen fehlte nur jedes achte — und damit nur das zu erratende — Wort (H entspricht dem Lückentext 7, N dem Lückentext 9. Vgl. Tab. 1! Das Schema pro Vierergruppe: Vp1 — P7, K9; Vp2 — K9, P7; Vp3 — K7, P9; Vp4 — P9, K7).

Die Ergebnisse sind in Tab. 9 dargestellt. Die statistisch bedeutsamsten Unterschiede im Ausgangsniveau entfallen auf die bloße Information „i". Differenzen (zwischen H und N, P und K, usw.) bezüglich des Verbesserungsbetrages erweisen sich nur bei Berücksichtigung des Zeitfaktors (t bzw. D) als signifikant; wie schon in Exp. 1 fallen auch hier höhere Verbes-

Tabelle 8

Die Variabilitätskoeffizienten des Ausgangsniveaus „A", des beim zweiten Ratedurchgang erreichten Scores „B" und der Verbesserung „V" (= A — B).

			i			$\sqrt{i \cdot t}$			$\sqrt{i \cdot D}$		
		n	A	B	V	A	B	V	A	B	V
H	VGn	6	12.85	18.82	16.70	3.66	3.46	21.33	8.77	6.21	16.24
N	VGn	6	9.21	18.64	30.69	6.10	6.82	23.23	9.29	7.37	21.41
K	VGn	6	8.99	16.14	23.22	5.49	3.95	27.01	10.44	8.32	21.56
P	VGn	6	13.27	22.00	20.83	5.99	7.67	20.44	7.60	5.41	20.24
H(real)	VGn	6	12.85	18.82	16.70	4.21	4.28	20.02	8.69	5.93	16.05
N(real)	VGn	6	9.21	18.64	30.69	3.40	6.35	20.23	5.08	7.73	15.14
	Vpn I	8*	11.40	16.55	37.21	11.97	8.86	44.65	29.75	17.89	64.10
H	Vpn II	16	14.05	18.06	51.91	13.85	12.83	33.34	21.21	23.54	35.58
	Vpn III	24*	17.97	29.00	42.05	23.50	20.11	52.55	34.38	36.02	53.76
	Vpn I	8	22.56	33.58	125.44	14.48	14.75	51.99	26.22	23.84	59.95
N	Vpn II	16	16.19	27.93	63.85	13.29	17.11	36.73	20.71	26.32	33.87
	Vpn III	24	19.90	35.57	67.58	17.37	18.21	46.28	24.24	26.55	48.88

* In der Spalte $\sqrt{i \cdot D}$ sind die entsprechenden Datenpaare um 1 auf n = 7 bzw. n = 23 reduziert. (Erklärung im Text)

Tabelle 9

Die für Exp. 2 berechneten Mittelwertsunterschiede pro Zeichen — bez. Ausgangsniveau „A" und Verbesserung „V" — zwischen den Versuchsbedingungen H vs N, zwischen den Texten K vs P (auch aufgeschlüsselt nach H vs N) und zwischen den Daten bei einem real höheren vs niedrigeren Ausgangsniveau.

		i				$\sqrt{i \cdot t}$				$\sqrt{i \cdot D}$			
Diff.	n	A	p<	V	p<	A	p<	V	p<	A	p<	V	p<
H — N	16	.175		.003		.142		—.007		.115		—.035	
K — P	16	.442	1	.072		.664	1	.264	10	.666	1	.270	
H(P) — N(K)	8	—.267	1	—.068		—.522	5	—.272	10	—.551	5	—.305	5
H(K) — N(P)	8	.616	1	.075		.806	10	.257		.781	10	.236	
H — N (real)	16	.462	1	.120		.733	1	.485	1	.794	1	.582	1

serungsbeträge mit höherem Ausgangsniveau zusammen. (Bei den signifikanten Differenzen haben A und V immer dasselbe Vorzeichen.)

Der Unterschied zwischen H und N (erste Zeile von Tab. 9) war bezüglich des Verbesserungsbetrages minimal, er machte sich erst in der zweiten oder dritten Kommastelle bemerkbar. (Noch hypothesenkonformer bzw. noch geringer stellt sich dieser Unterschied dar, wenn bit und sec über alle 16 Vpn summiert und dann erst das geometrische Mittel berechnet wird. Vgl. Fenk, 1985b, und Abb. 5: Berechnet man die Scores jener Lückentexte aus Exp. 1, die den Unvollständigkeitsstufen von Exp. 2 entsprechen — die Lückentexte 7 und 9 stehen also für H und N —, so nehmen H und N im „besten" Kriterium $\sqrt{\text{i.D}}$ sehr hypothesenkonforme und sehr ähnliche Positionen ein, obwohl wir es nun in Exp. 1 nicht mehr mit einem intraindividuellen Vergleich zu tun haben.) Allerdings lagen auch in der neuen VG mit der breiter abgestuften Lückenhaftigkeit die Ausgangsniveaus von H und N sehr dicht beisammen. Dieser Umstand war der Anlaß dafür, einerseits so wie schon in „Exp. 1" Vergleiche zwischen K und P bzw. zwischen H(real) und N(real) anzustellen, und andererseits durch noch differenziertere Auswertung (siehe Zeile 3 und 4 von Tab. 9 sowie die Abb. 4!) den Grund für diese enge Nachbarschaft aufzuspüren. Die obere Hälfte von Abb. 4 verrät diesen Grund. Nicht nur, daß die Texte P und K sich im Ausgangsniveau immer noch stärker unterscheiden als die zwei Stufen der Lückenhaftigkeit (7 vs 9), „reagieren" sie sozusagen auch noch gegensinnig: Bei Text P nimmt der Lerngewinn V mit wachsendem Ausgangsniveau zu, bei Text K dagegen nimmt er ab. Wie wenig reliabel das ganze Gefüge noch dazu ist, zeigt sich, wenn man nur eine einzige Versuchsperson (nämlich Vp 15, die durch einen enormen Zeitaufwand beim Raten in die Rolle eines statistischen Ausreißers gerät) aus der Auswertung nimmt: Das Gefälle zwischen K9 und K7 und zwischen N und H kippt, der vollständigere Text (N bzw. K9) provoziert plötzlich ein real höheres Ausgangsniveau, welches aber nun wieder — wie schon aus „Exp. 1" gewohnt — mit einer stärkeren Verbesserung zusammenfällt.

Diskussion

Die Ausgangsposition

Skizzieren wir noch einmal die Ausgangsposition der Untersuchung: Einerseits sprechen sehr plausible Argumente für die Informationskonstanzhypothese; andererseits widersprechen ihr die Ergebnisse einiger einschlägiger Experimente. (Wobei den hypothesenfeindlichen Resultaten wegen des Null-Hypothesen-Charakters der Konstanzannahme von vorne-

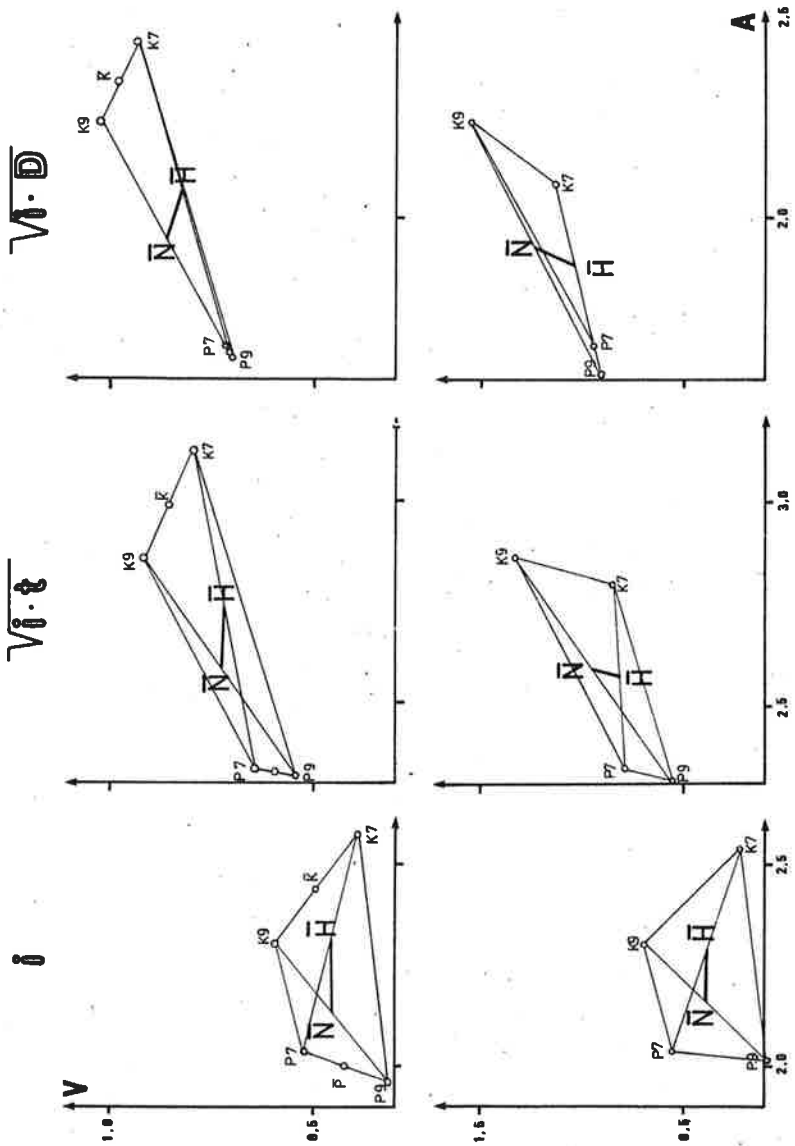

Abb. 4
Die Textproben K und P in den Unvollständigkeitsstufen H (=7) und N (=9) sowie die Durchschnittswerte von H und N im Koordinatensystem Ausgangsniveau/Verbesserung. (Unten: ohne Vp 15)

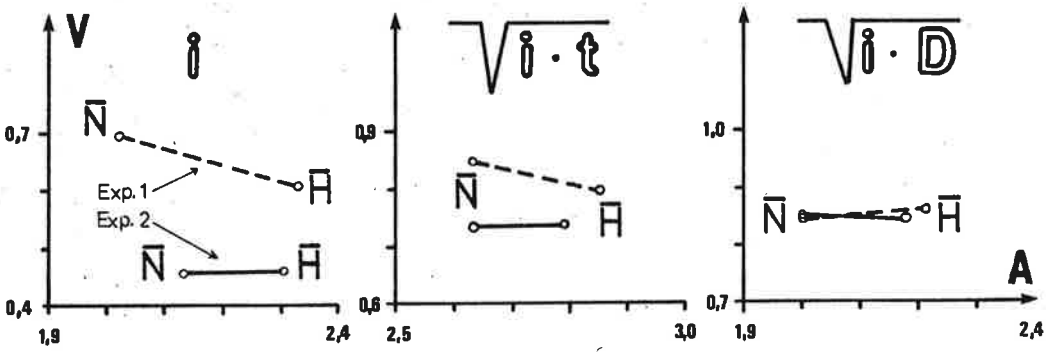

Abb. 5
Ein Vergleich zwischen Exp. 1 und Exp. 2 bezüglich der Lückentexte 7 (=H) und 9 (=N).

herein mehr Gewicht zukommt als den hypothesenfreundlichen.) Die Ursache für die hypothesenfeindlichen Resultate wurde nun in inadäquaten Operationalisierungen der einschlägigen Studien vermutet.

Die Sanierung der Konstanzannahme — soweit das bei einer Null-Hypothese überhaupt möglich ist — sollte eine Operationalisierung leisten, die sich konsequent innerhalb des Konzepts der subjektiven Information bewegt. Die „objektive" Information kann ja nur indirekt und in dem Maße relevant sein, als die ihr zugrundeliegende „objektive" Wahrscheinlichkeitsverteilung mit der subjektiven übereinstimmt. Trotz dieser nie (?) ernsthaft bestrittenen Auffassung hatten sich frühere Experimente vornehmlich an einem nachrichtentechnischen Modell der begrenzten Kanalkapazität orientiert, welches die Fixierung der Kapazitätsgrenzen in Termen objektiver Information vorsieht. Unsere methodische Ausrichtung nach dem Konzept der subjektiven Information rückt „information processing" in die Nähe von Konzepten wie „meaning and intuition", von deren Operationalisierung sich nicht nur Pribram (1985, S. 6) entscheidende Fortschritte der Psychologie erwartet. (Vielleicht ist die bei Pribram anklingende Unterscheidung ohnehin nur verständlich als Folge der oft vorschnellen Übertragungen nachrichtentechnischer Modelle auf ein Wesen, dessen „Informationsverarbeitung" ja gerade darin besteht, das relativ Neue durch entsprechende Hypothesenentwürfe seines Neuigkeitswertes zu berauben, es in bisheriges Wissen einzuordnen und dadurch mit Bedeutung zu versehen.)

Begreift man nun „Lernen" als Reduktion der subjektiven Information (einer Nachricht), so liegt der Gedanke nahe, das Ausmaß der Wissenserweiterung (des Lerneffektes = der Zunahme der Geläufigkeit von Nachrichten) als Ersparnis an bit subjektiver Information bei Wiederholung eines Ratespiels meßbar zu machen. In die (Bestimmung der) Geläufigkeit einer Nachricht geht aber auch die Entscheidungszeit (des Ratenden) mit

ein: Je größer die Geläufigkeit, um so kleiner das geometrische Mittel aus bit und sec. Die von einem Probanden(kollektiv) erzielte Reduktion dieser Scores von trial 1 zu trial 2 wird, so die umformulierte Informationshypothese, konstant bzw. invariant gegenüber intraindividuell wechselnden Ausgangswerten (Scores in trial 1) bleiben.

Vielleicht sind die „technischen" Vorteile (s. S. 215 f.) dieser Operationalisierung darin begründet, daß sie dem Begriff des „Lernens" und der „Wissenserweiterung" weniger Gewalt antut als das Konzept der limitierten Kanalkapazität. Für das zuletzt genannte Konzept gilt die Konstanzannahme als bestätigt, wenn die Menge an objektiver Information, die von unveränderlich gedachten (Komponenten von) informationsverarbeitenden Systemen pro Zeiteinheit „übertragen" — also aufgenommen und wieder abgegeben — wird, einen gleichbleibenden Maximalwert nicht übersteigt. In unserer neuen Hypothese dagegen bleibt „Lernen" die „Ausbildung oder Korrektur von individuellem Gedächtnisbesitz" (Klix, 1971, S. 348); Lernen bedeutet also, daß der Informationsverarbeiter sich bzw. seine internen Modelle ändert. Und das Ausmaß eben dieser Änderung wird als informational beschränkt angesehen. Etwas salopp formuliert: Es geht nicht um eine Veränderung (ein Verrauschen) der Nachricht durch den Informationsverarbeiter, sondern um die Veränderung des Informationsverarbeiters durch die Analyse der Nachricht. Wenn also irgendetwas nach oben hin begrenzt ist, dann ist das nicht der Informationsdurchsatz durch einen Kanal, durch eine „Enge des Bewußtseins" oder durch irgendwelche Flaschenhälse mit feststehendem Querschnitt, sondern dann sind das die dem Subjekt verfügbaren Möglichkeiten, seine Effizienz bei der Analyse einer Nachricht zu steigern. In dieser Leistungssteigerung (z. B. bei Wiederholung eines Ratespiels) dokumentiert sich die Revision der subjektiven Hypothesen (des Ratenden), also die Erweiterung des — in Ansätzen immer schon vorhandenen, aber immer hypothetisch bleibenden — (Vor-)Wissens in all seinen aufgabenrelevanten (deklarativen und prozeduralen) Aspekten. So gesehen und Ekman's (1972, S. 44) Klassifizierung folgend, beschäftigt sich unsere neue Hypothese nicht mehr mit S-R-Beziehungen bzw. mit Input/Output-Beziehungen, sondern mit „intrasubjektiven Beziehungen (Problemtypus III)"; sie trifft Aussagen über intraindividuelle Relationen zwischen Ausgangsniveau und Verbesserungsbetrag, welche beide in Termen subjektiver Information definiert sind.

Konfrontation der Ergebnisse mit den Erwartungen

Vor allem die uneingeschränkt interpretierbaren Mittelwertsvergleiche brachten Ergebnisse, die unsere Vorhersagen eindeutig widerlegen: Mit

einem (signifikant) höheren Ausgangsniveau war oft ein (signifikant) höherer Lerngewinn assoziiert; besonders ausgeprägt war diese Beziehung bei der vermutlich besten Operationalisierung $\sqrt{i \cdot D}$. Allerdings ergaben der Vergleich zwischen den Texten K vs P in Exp. 1 und der Vergleich zwischen verschieden hohen Graden der Kontextzerstörung (H vs N) in Exp. 2 ein durchaus hypothesenkonformes Bild. Diese hypothesenkonformen Resultate sind aber — angesichts der restlichen, signifikant widersprechenden Resultate — nicht als Bestätigung zu werten, sondern nur als das Fehlen signifikanter Abweichungen von der Nullhypothese.

Übrigens führt auch Operationalisierung 1 (s. S. 216 f.) zu einer Ablehnung der Konstanzannahme. Bei den unvollständigeren Texten, also in H, wird pro Zeiteinheit mehr Information reduziert als in N (Fenk, 1985a). Die im Detail bisher unveröffentlichten Resultate: Der Mittelwerts-Unterschied H > N ist bei Berechnung über die sechs Versuchsgruppen auf dem 10%-Niveau „signifikant", gleichgültig, ob man die Ratezeit brutto (t) oder netto (D) zugrundelegt. Bei Berechnung auf Ebene der Einzelpersonen, getrennt nach drei Versuchsbedingungen und nach t und D, ergeben sich sechs Vergleiche; einer davon, nämlich H > N in VB I bei Zugrundelegung von t, war signifikant ($p < 2\%$).

Argumente gegen die Informationskonstanz liefern schließlich auch die flankierenden Auswertungen. (Für sie gilt allerdings, daß die statistischen Kennwerte nicht aus Streuungen intraindividueller Differenzwerte berechnet wurden). Für den Verbesserungsbetrag, der laut Hypothese am allerwenigsten hätte variieren sollen, errechnete sich ein höherer Variabilitätskoeffizient als für das Ausgangsniveau. Und das Ausgangsniveau korrelierte positiv mit dem Verbesserungsbetrag.

Nur wenn der Zeitfaktor unberücksichtigt blieb, also bei Berechnung aus den bloßen bit-Scores, ergaben die Korrelationen zwischen Ausgangsniveau und Verbesserung ein anderes Bild: Bei Berechnung auf Ebene der Vpn je einer Versuchsbedingung standen den positiven Korrelationen in H negative Korrelationen in N gegenüber. (Aus der Berechnung über die sechs Versuchsgruppen hinweg resultierte sogar ein signifikant negativer Koeffizient. Da mögen die schon früher erwähnten Verzerrungen mitgespielt haben.) Wegen der negativen Korrelation der subjektiven Information mit dem Zeitaufwand (s. unten) muß man derartige Unterschiede zwischen H und N im Zusammenhang mit der Zeitinvestition (t, D) sehen, die sich pro Versuchsbedingung von Gruppe zu Gruppe (VGn 2—3, VGn 4—6) bei den schwieriger zu ergänzenden Texten (H, K) anders entwickelte als bei den leichteren (N, P): In trial 1 steht ein steiler Anstieg des Zeitaufwandes bei den schwierigen Texten einer eher abnehmenden Tendenz bei den leichten Texten gegenüber; und daraus resultiert, weil der Zeitaufwand in trial 2 bei den schwierigen wie bei den leichten Texten leicht ansteigt, bei den schwie-

rigen Texten eine Zunahme der Zeitersparnis, bei den leichten hingegen eine Abnahme.

Die einzige Vorhersage, die aus forschungslogischer Sicht eine signifikante Bestätigung finden konnte, fand diese auch: Der Zeitaufwand im Ratespiel korrelierte tatsächlich negativ mit der im Ratespiel ermittelten subjektiven Information in bit. Mehr noch: Die Lage der jeweiligen Regressionsgeraden vermittelt den Eindruck einer umgekehrten Proportionalität zwischen Entscheidungszeit und subjektiver Unsicherheit. Das heißt, der Informationsgehalt erhöht sich um etwa denselben Faktor, um den sich ein Ratender weniger Zeit nimmt (als ein anderer). (Und/oder: Der Ratende investiert um so weniger Zeit, je schwieriger ihm der Text bzw. je aussichtsloser ihm das interne Hypothesentesten erscheint — kurz, je weniger er sich von dieser Investition verspricht.) Eine Folgerung: Die Bestätigung der Vorhersage bedeutet der Tendenz nach eine empirische Hinterher-Rechtfertigung der von uns — wegen des von vorneherein vermuteten „speed-accuracy tradeoffs" — gewählten Leistungsscores in Form eines geometrischen Mittels aus bit und sec. Eine zweite Folgerung betrifft das Ergebnis, daß die negativen Korrelationen zwischen Zeitaufwand und Information in H (vor allem in trial 1) höher ausfielen als in N. (Nun wird natürlich diese höhere Korrelation schon rein rechnerisch dadurch begünstigt, daß der Zeitaufwand in H — vor allem über die Gruppen hinweg — stärker streut als in N. Als alleiniger Grund scheidet dies aber aus: Das zeigen Streuungsvergleiche H vs N bezüglich D in VBI ebenso wie die Beobachtung, daß wohl die Streuungen, nicht aber die Korrelationen generell in trial 1 höher liegen als in trial 2.) Demnach wirkten sich in H Unterschiede in dem von verschiedenen Vpn oder VGn getriebenen Zeitaufwand besonders massiv aus; im Vergleich zu N brachte jede Mehrinvestition an Zeit größere Vorteile in der Vermeidung von Ratefehlern.

Zusammenfassende Interpretation

Die Studie hat sich erstens die Aufgabe gestellt, eine adäquate Operationalisierung zur Analyse des Lernfortschrittes, der Beziehung Vorwissen/Wissenszuwachs, zu entwickeln. Die zweite Zielsetzung war es, entsprechend dieser Operationalisierung die Informations-(Konstanz-)Hypothese neu zu formulieren und zu prüfen. Wenn wir nun davon ausgehen, daß die Operationalisierung Hand und Fuß hat und das Weltner-Rateverfahren hinreichend präzise mißt, warum ist dann unsere zentrale Vorhersage gescheitert?

In Abb. 6 wird diese zentrale Vorhersage grafisch dargestellt: Die Verbesserung sollte immer etwa gleich groß sein, unabhängig vom Ausgangsni-

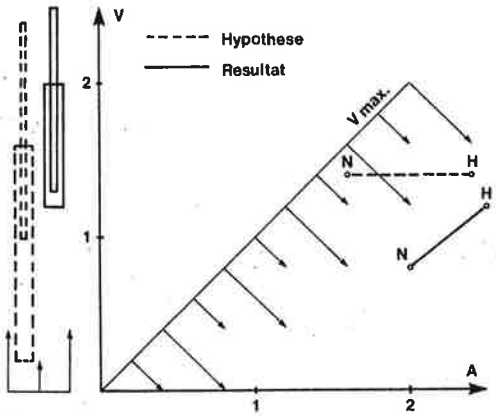

Abb. 6
Eine grafische Interpretation der Diskrepanz zwischen Hypothese und Resultat. (Die Blockdarstellung links ist so zu verstehen wie in Abb. 1.)

veau des jeweiligen Probanden(kollektivs). Der Score, in dem das Ausgangsniveau und die Verbesserung der Geläufigkeit gemessen werden — bit und vor allem bit × sec — ist ein Schlechtpunktescore. Das heißt, daß die Verbesserung in einer Reduktion des Scores besteht; und daß ein höheres Ausgangsniveau eine schlechtere Leistung bedeutet (zumindest absolut gesehen, also ohne Berücksichtigung des Umstandes, daß in H widrigere Umstände, nämlich stärkere Kontextzerstörungen, vorlagen). Die Tendenz unserer Ergebnisse, nämlich die von der Horizontale abweichende Lage der Verbindung zwischen H und N, wird in Abb. 5 ebenso illustriert wie die untenstehende, verbale Interpretation.

Wickelgren (1981, S. 42) über die Prinzipien menschlichen Lernens: „Of course, it would be wrong to go overboard on what could be called the reduction hypothesis, that we learn best when we have the least to add to what is already in memory. In the extreme, when everything in some input material is already encoded in memory, we often do further strengthen the existing associations, but this is not when the greatest amount of learning occurs. Learning curves always indicate diminishing returns after an earlier period of positive acceleration ...".

Wie haben wir uns dieses Verflachen der Lernkurve vorzustellen? Wenn (hohe) Leistung als (geringe) Fehlerzahl bei vorgegebenem Zeitbudget bestimmt wird, dann ist der theoretisch mögliche Leistungszuwachs eindeutig definiert und auch im (Prinzip) praktisch realisierbar: Der maximale Verbesserungsbetrag (die Zahl der reduzierbaren Fehler) ist gleich dem Ausgangsniveau (der Zahl der Fehler zu Beginn der beobachteten Lernphase).

Wenn dagegen die Zeit die freie Variable ist — bei Vorgabe einer minimalen Wegstrecke, einer maximalen Fehlerzahl, etc. — dann ist der „theoretisch mögliche" Leistungszuwachs niemals realisierbar: Kein Leichtathlet wird je die 100-m-Distanz ohne jeden Zeitaufwand zurücklegen. Jeder Verbesserung des Fehlerscores und besonders jeder Verbesserung des Zeitscores steht mit zunehmender Nähe der theoretisch möglichen Verbesserung ein wachsender „Widerstand" entgegen: Die Verbesserung der relativen Fehler (Nicht-Treffer) eines Schützen von 10% auf 5% oder die eines 100-m-Sprinters von 12 auf 11 sec ist leichter zu erzielen als die Verbesserung von 5% auf 0% oder von 11 auf 10 sec. Dasselbe gilt naturgemäß für Scores, in denen die Fehlerzahl sinnvoll mit dem — für jede Art von Aktivität nötigen — Zeitkonsum in Beziehung gesetzt wird. Die zusätzliche Lehre aus unserem Experiment: Es gilt auch dann, wenn die relative Fehlerzahl primär den jeweiligen Wissensstand widerspiegelt und in bit umgerechnet wird. Bei Einbeziehung des Zeitaufwandes fällt die Abweichung von der Konstanzannahme besonders deutlich aus, was nach dem oben Gesagten nicht mehr verwundern kann.

Hat man erst einmal den „Abflachungseffekt" mit unseren Ergebnissen in Verbindung gebracht, so erscheint das Scheitern der Konstanzannahme hinterher nicht nur einleuchtend, sondern trivial: Wenn von vornherein klar war, daß erstens die Reduktion eines niedrigen Fehler- oder Fehler × Zeit-Scores — im Vergleich zu einem höheren — auf stärkeren Widerstand stößt, und daß zweitens die nach der Weltner-Formel berechnete bit-Zahl eine lineare Funktion der relativen Ratefehlerzahl ist, war dann nicht auch das Scheitern der Konstanzannahme von vorneherein klar? So ganz logisch vorhersagbar war dieses Scheitern nicht. Denn der Score, auf den die Probanden ihre Strategien abstellen konnten (siehe Instruktion!) war nicht identisch mit dem für die Hypothesenprüfung berechneten, und die Strategien der Probanden bezüglich des speed-accuracy tradeoffs waren nicht vorgezeichnet. Und vor allem: Es hätte ohne weiteres sein können, daß die erhaltenen Leistungsdaten weit außerhalb einer „Abstoßungszone" des theoretischen Leistungsplafonds liegen. Und schließlich sind unsere Daten anders zustandegekommen als jene, die gewöhnlich einer Lernkurve zugrundeliegen: Wir haben nicht, um ein Beispiel zu bemühen, wieder und wieder ein- und denselben Text tippen lassen und mitverfolgt, wie von Mal zu Mal sich Fehler und/oder Zeit verringern. Vielmehr haben wir (je zweimal) zwei verschiedene Texte verwendet, deren einer von vorneherein schwieriger zu tippen war.

Aus diesen Gründen, und weil wir pro Probanden(kollektiv) nur zwei Punkte im Koordinatensystem Ausgangsniveau/Verbesserung kennen und deren Lage auf einer gekrümmten Kurve nur vermuten können, wurde nach weiteren Indizien für unsere Interpretation gefahndet. Dieser Interpretation

zufolge rührt der höhere Lerngewinn bzw. die stärkere Reduktion des Schlechtpunktescores bei höherem Anfangsscore daher, daß sich die Endscores (trial 2) bei hohem und niederem Anfangsscore weniger stark unterscheiden als die Anfangsscores.

Wenn diese Interpretation stimmt, dann sollte der Variabilitätskoeffizient beim trial-2-Score (B) niedriger ausfallen als beim trial-1-Score (A). Tabelle 8 bietet diesbezüglich ein uneinheitliches Bild; immerhin ist aber bei der „besten" Schätzung $\sqrt{i \cdot D}$ der Variabilitätskoeffizient von B meist (in 7 von 12 Vergleichsmöglichkeiten) kleiner als der von A.

Und wenn diese Interpretation stimmt, dann sollten die Mittelwertsdifferenzen zwischen hohem und niederem Anfangsscore geringer ausfallen als die zwischen den zugehörigen Endscores. Dazu gibt Tabelle 10 Auskunft. (Die Werte stammen nicht aus der Berechnung auf Ebene der Einzelpersonen. Vielmehr wurden pro VG zuerst bit und sec ermittelt, und dann erst das geometrische Mittel; die Vpn von Exp. 2 wurden wie eine VG behandelt.) In zwei der sechs Zeilen zeigen sich die bereits erwähnten, wunderschönen „Schein-Bestätigungen" der Konstanzannahme: Die Differenzen K − P in Exp. 1 und H − N in Exp. 2 sind bezüglich A und B praktisch gleich groß und daher bezüglich V (= A − B) praktisch gleich Null. Bei den restlichen Zeilen ist die Differenz bezüglich B (im Vergleich zur Differenz bezüglich A) sehr gering, was auf die vermutete Einengung des Spielraumes durch die Nähe des theoretischen Leistungsmaximums hinweist. Und genau in diesen Vergleichen finden sich auch die hypothesenwidrigen Differenzen bezüglich V.

Auf diese Einengung des Spielraumes weist schließlich auch das Ergebnis eines ähnlichen, aber leider ohne Zeitkontrolle konzipierten Experiments hin: „Ein extrem niedriges Ausgangsniveau bzw. ein sehr hohes Maß an ‚Vorwissen' bietet kaum Spielraum für einen drastischen Abbau von Unsicherheit oder für einen drastischen Lerngewinn — man kann dort nicht

Tabelle 10

Die für trial 1 (A), trial 2 (B) und die Verbesserung (V) berechneten Differenzwerte zwischen H und N, K und P sowie H(real) und N(real).

		i			$\sqrt{i \cdot t}$			$\sqrt{i \cdot D}$		
	Diff.	A	B	V	A	B	V	A	B	V
Exp. 1	H—N	.17	.04	.13	.11	—.01	.12	.10	—.03	.13
	K—P	.35	.39	—.04	.32	.30	.02	.32	.28	.04
	H—N (real)	.17	.04	.13	.19	.02	.17	.23	.03	.20
Exp. 2	H—N	.17	.17	.00	.16	.15	.00	.15	.15	—.00
	K—P	.44	.37	.07	.70	.44	.26	.75	.47	.28

mehr viel dazulernen." (Lechner & Fenk, in Begutachtung, vgl. auch Fenk & Lechner, 1985). Die vorliegende Studie verweist auf die Schlüsselrolle, die der Zeitaufwand dabei spielt: In trial 1, verglichen mit trial 2, sowie beim Lückentext H und der Textprobe K, verglichen mit P und N, operierten die Probanden in größerer Entfernung vom theoretischen Leistungsmaximum (0 bit, 0 sec). In dieser größeren Entfernung wirkte sich jede Zeitinvestition spürbarer in Richtung auf eine Minimierung der Ratefehlerzahlen aus; sie bot gerade deswegen auch mehr Anreiz für größere Zeitinvestitionen.

Es scheint so zu sein, daß wir eine langgehegte Hoffnung begraben müssen. Die Hoffnung nämlich, daß wir durch Zugrundelegung informationaler Maße zu sehr viel einfacheren Funktionen — etwa zwischen Anfangsleistung und Leistungszuwachs — bzw. zur Entkoppelung solcher Variablen gelangen könnten. Unser Versuch, den Lernfortschritt durch dessen Quantifizierung innerhalb des Konzeptes der subjektiven Information als unabhängig vom Ausgangslevel zu erweisen, ist gescheitert. Oder, positiv gewendet: Die gesetzesartige Aussage über einen geringeren Lernzuwachs in der Nähe eines theoretischen Leistungsmaximums hat sich als robust erwiesen gegenüber wechselnden Leistungsbereichen, wechselnden Maßen (informationalen, non-informationalen) und gegenüber unterschiedlich „liberalen" experimentellen Designs (z.B. vorgegebenes vs disponibles Zeitbudget).

Die Annahme einer strengen Proportionalität oder auch nur einer strengen Linearität der Beziehung zwischen dem informational gemessenen Erfolg (internen) Hypothesentestens und dem dafür nötigen Zeitaufwand wird man wohl aufgeben müssen. Hingegen bieten unsere Resultate keinen Anlaß, die gängige Auffassung bezüglich der Richtung dieses Zusammenhanges aufzugeben. Einschlägige Aussagen — nicht nur der Psychologie, sondern auch der Linguistik oder der Pädagogik — haben einen positiven Zusammenhang bzw. eine direkte Proportionalität von Zeit und Information zum Inhalt: So weisen etwa sprachvergleichende Studien darauf hin, daß sich natürliche Sprachen in Anpassung an unsere Limits entwickelt haben; mehr Information wird auf mehr Silben oder auf komplexere Silben verteilt (Fenk & Fenk, 1980; Fenk-Oczlon & Fenk, 1985). Und die (kybernetische?) Pädagogik fordert, daß dem Adressaten für die Verarbeitung informationsreicherer Botschaften mehr Zeit eingeräumt wird, wenn das Verarbeitungsniveau nicht beeinträchtigt werden soll. Sie kann sich dabei auf Beobachtungen (z.B. Kroebel, 1971) stützen, wonach wir unser Lesetempo der Informationsverteilung eines Textes in der Weise anpassen, daß wir uns für informationsreichere Passagen mehr Zeit nehmen. In unserem Experiment wurden Lückentexte gleichzeitig als „Lernmaterial" und als Diagnose-Instrument zur Bestimmung der subjektiven Information ein-

gesetzt, und Information wird ja im Laufe des jeweiligen Ratespiels abgebaut. Die dabei beobachteten negativen Korrelationen zwischen Zeitaufwand und subjektiver Information widersprechen den obigen Aussagen nur scheinbar; sie dokumentieren, sozusagen mit umgekehrtem Vorzeichen, denselben Sachverhalt: Je länger die Analysezeit, um so geringer die verbleibende Unsicherheit.

Summary

The assumption of constancy in informationally defined performance based on capacity limits, an assumption central to the understanding of cognitive processes, is reformulated, particularly with the assistance of the concept of subjective information, which, like item difficulty, refers to a relation between message and processing system. The subjects had to twice reconstruct letter by letter (guessing-game by Weltner, 1970) two fragmentary texts of different difficulty; an above average ratio of correct to wrong guesses per unit of time was rewarded with a cash payment (exp. 1, 48 subjects; exp. 2, with more broadly spread difficulty levels, 16 subjects). The learning increment from the first to the second trial – i.e. the increase of familiarity or the reduction of bits or bits × seconds – was expected to remain constant within subjects, despite different starting points. While high negative correlations between guessing speed and guessing accuracy point to capacity limits and justify the measure bit × sec, comparisons between the mean values refute the new constancy assumption: With lower scores bit × sec in trial 1, i.e. greater task-relevant knowledge, there was lower learning increment.

Literatur

Attneave, F.: Informationstheorie in der Psychologie. Huber, Bern 1965 (2. Aufl. 1969). Originalausgabe: Applications of Information Theory to Psychology. H. Holt and Company, New York 1959.

Broadbent, D.E.: Levels, hierarchies, and the locus of control. Quarterly Journal of Experimental Psychology, 1977, **29**, 181—201.

Chomsky, N.: Aspects of the theory of syntax. M.I.T. Press. Cambridge, Mass. 1965.

von Cube, F.: Was ist Kybernetik? C. Schünemann, Bremen 1967.

Ekman, G.: Psychophysik und psychologische Meßmethoden. In: Meili, R. & Rohracher, H.: Lehrbuch der experimentellen Psychologie. Huber, Bern 1963 (3. Aufl. 1972).

Estes, W.K.: Th cognitive side of probability learning. Psychological Review, 1976, **83**, 1, 37—64.

Eysenck, M.W.: A handbook of cognitive psychology. Lawrence Erlbaum Associates Ltd., London 1984.

Fenk, A.: „Ein Bild sagt mehr als tausend Worte...?" Lernleistungsunterschiede bei optischer, akustischer und optisch-akustischer Präsentation von Lehrmaterial. AV-Forschung, 1981, **23**, 5—50.

Fenk, A.: Der Zeitfaktor in Ratespielen zur Bestimmung der Informationsreduktion. In: Fleischmann, U.M. (ed.): Zusammenfassungen der Beiträge zur 26. Tagung exp. arbeitender Psychologen, Universität Nürnberg, 1984, S. 68.

Fenk, A.: Is the reduction of subjective information (per unit of time) independent from the starting level? In: d'Ydewalle, G. (ed.): Cognition, Information Processing, and Motivation. Elsevier Science Publishers B.V. (North-Holland), 1985a, S. 361—373.

Fenk, A.: Der Lerngewinn in bit als Funktion des Vorwissens und des Zeitaufwandes. Abstractband, 27. Tagung exp. arbeitender Psychologen, Universität Wuppertal, 1985b, S. 6.

Fenk, A. & Fenk, G.: Konstanz im Kurzzeitgedächtnis — Konstanz im sprachlichen Informationsfluß? Zeitschrift für experimentelle und angewandte Psychologie, 1980, **27**, 400—414.

Fenk-Oczlon, G. & Fenk, A.: The mean length of propositions is 7±2 syllables — but the position of languages within this range is not accidental. In: d'Ydewalle, G. (ed.): Cognition, Information Processing, and Motivation. Elsevier Science Publishers B.V. (North-Holland), 1985, 355—359.

Fenk, A. & Lechner, B.: Zwei Experimente zum Redundanzgewinn bei wechselndem Ausgangsniveau. In: Albert, D. (ed.): Ber. üb. d. 34. Kongr. d. DGfP in Wien 1984, S. 250—251. Verl. f. Psychologie, Hogrefe, Göttingen 1985.

Frank, H.: Kybernetische Grundlagen der Pädagogik. Kohlhammer, Stuttgart 1971.

Greeno, J.G.: Evaluation of Statistical Hypotheses Using Information Transmitted. Philosophy of Science, 1970, **37**, 279—293.

Herrmann, T.: Psychologie und Kybernetik. Psychologische Rundschau, 1966, **17**, 13—25.

Herzog, W.: Modell und Theorie in der Psychologie. Verl. für Psychologie, Hogrefe, Göttingen 1984.

Howes, D.H. & Solomon, R.L.: Visual duration threshold as a function of word-probability. Journal of Experimental Psychology, 1951, **41**, 401—410.

Klix, F.: Information und Verhalten. Huber, Bern 1971.

Kroebel, W.: Ein neues Verfahren zur Bestimmung der relativen subjektiven Information aus Messungen des Lesegeschwindigkeitsverlaufs. In: Rollett, B. & Weltner, K. (ed.): Fortschritte und Ergebnisse der Unterrichtstechnologie. Ehrenwirt Verlag, München 1971.

Lechner, B. & Fenk, A.: Die Ersparnis an subjektiver Textinformation bei Wiederholung eines Ratespiels. In Begutachtung.

Lorenz, K.: Die Rückseite des Spiegels. Versuch einer Naturgeschichte menschlichen Erkennens. dtv, München 1977.

Mayerthaler, W.: Markiertheit in der Phonologie. In: Vennemann, Th. (ed.): Silben, Segmente, Akzente. Niemeyer, Tübingen 1982, 205—246.

Melezinek, A.: Ingenieurpädagogik. Springer-Verlag, Wien 1977.

Miller, G.A.: The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychological Review, 1956, **63**, 81—97.

Miller, G.A., Bruner, J.S. & Postman, L.: Familiarity of letter sequences and tachistoscopic identification. The Journal of General Psychology, 1954, **50**, 129—139.

Miller, G.A. & Selfridge, J.A.: Verbal context and the recall of meaningful material. American Journal of Psychology, 1950, **63**, 176—185.

Mittenecker, E.: Planung und statistische Auswertung von Experimenten. Franz Deuticke, Wien 1979 (9. Aufl.).

Mittenecker, E. & Raab, E.: Informationstheorie für Psychologen. Verl. f. Psychologie, Hogrefe, Göttingen 1973.

Norman, D.A.: Aufmerksamkeit und Gedächtnis. Beltz, Weinheim—Basel 1973.
Popper, K.R.: Logik der Forschung. J.C.B. Mohr (Paul Siebeck), Tübingen 1976 (6. Aufl.).
Posner, M.I. & Mc Leod, P.: Information processing models — in search of elementary operations. Ann. Rev. Psychol., 1982, **33**, 477—514.
Pribram, K.H.: ‚Holism' could close cognition era. American Psychological Association Monitor, 1985, **16**, **9**, 5—6.
Rubenstein, H. & Aborn, M.: Immediate recall as a function of degree of organization and length of study period. Journal of Experimental Psychology, 1954, **48**, **2**, 146—152.
Salmon, W.C.: Statistical explanation and statistical relevance. University of Pittsburgh Press, 1971.
Shannon, C.E.: A mathematical theory of communication. Bell Syst. Techn. J., 1948, **27**, 379—423, 623—656.
Sumby, W.H. & Pollack, I.: Short-time processing of information. HFORL-Report, TR—54—6, 1954.
Vollmer, G.: Evolutionäre Erkenntnistheorie. S. Hirzel, Stuttgart 1981.
Weltner, K.: Informationstheorie und Erziehungswissenschaft. Schnelle, Quickborn 1970.
Wessels, M.G.: Kognitive Psychologie. Harper & Row, New York 1984.
Wickelgren, W.A.: Human learning and memory. Ann. Rev. Psychol., 1981, **32**, 21—52.
Wimmer, H. & Perner, J.: Kognitionspsychologie. Kohlhammer, Stuttgart 1979.
Wippich, W.: Lehrbuch der angewandten Gedächtnispsychologie. Kohlhammer, Stuttgart 1984.

Anschrift des Verfassers: Univ.-Doz. Dr. August Fenk, Abteilung für Kognitionsforschung am Inst. für Unterrichtstechnologie und Medienpädagogik, Universität für Bildungswissenschaften Klagenfurt, Universitätsstraße 65—67, A-9010 Klagenfurt.

Anhang

(Phantasie) Name:
Alter: . . Jahre
Geschlecht: männl. ☐
weibl. ☐

Nr.

Texte:

Liebe Kollegin, lieber Kollege,

auf dem Bildschirm erscheint der Anfang eines mehr oder weniger lückenhaften Textes. Die Buchstaben von Wörtern, die in dem Textfragment fehlen, sind durch eine entsprechend große Anzahl von Punkten ersetzt. *Sie sollen* nicht die Lücken schließen, sondern *die Ihrer Meinung nach wahrscheinlichste Fortsetzung des Textes Buchstabe für Buchstabe eintippen.* Sie raten pro Buchstabe nur einmal, dann antwortet das Gerät mit dem richtigen Buchstaben. Sobald Sie ein Wort zu Ende getippt haben, liefert das Gerät eine weiteres Textfragment, und dann sind wieder Sie dran, etc.

Hinweise:

- *Nach jedem eingetippten Buchstaben ist die* [Return] *-Taste zu drücken.* (Ein Tip: Lassen Sie die rechte Hand immer in der Nähe dieser Taste und schreiben Sie nur mit der linken Hand!) Wenn Sie glauben, einen falschen Buchstaben getippt zu haben und Sie die Return-Taste noch nicht gedrückt haben, können Sie die *Korrekturtaste* [←] betätigen und den gewünschten Buchstaben eingeben.
- Anstelle der Umlaute werden vom Gerät — und bitte auch von Ihnen — nur die jeweiligen Vokale geschrieben, ohne ein E anzuhängen (Also A, O und U statt Ä, Ö und Ü). Zwischen zwei Sätzen setzt das Gerät einen Schrägstrich statt des Punktes.

Wir machen jetzt eine *Probe* mit einem kurzen Text, in welchem jedes zweite Wort fehlt und jedes vierte Wort zu erraten ist.

Nach dieser Probe kommt das eigentliche Experiment, bestehend aus 6 Aufgaben:

1) Text A (jedes . Wort fehlt, jedes . Wort ist zu erraten)
2) Reaktionstippen (100 sec lang sollen möglichst schnell die auf dem Bildschirm erscheinenden Buchstaben und sofort darauf die Return-Taste gedrückt werden)
3) Text A noch einmal
4) Text B (jedes . Wort fehlt, jedes . Wort ist zu erraten)

5) Reaktionstippen (siehe oben)
6) Text B noch einmal

Worauf es ankommt: Möglichst oft den richtigen Buchstaben zu erraten bzw. einzugeben und möglichst schnell zu arbeiten. Sie sollen im Leistungskriterium $\left(\frac{\text{Anzahl richtiger Buchstaben}}{\text{Anzahl falscher Buchstaben} \times \text{Zeit}}\right)$ einen möglichst hohen Wert erreichen. Je 4 Teilnehmer haben vergleichbare Aufgaben zu lösen. Nach Abschluß der Untersuchung erhalten Sie für jede der 6 Aufgaben, in der Sie zur besseren Hälfte dieser 4 gehören, S 50,— Prämie, insgesamt also maximal S 300,—. Weitere S 50,— erhalten Sie sofort nach dem Experiment und unabhängig von Ihrer Leistung.

Bitte keine Wörter bzw. Textstellen kolportieren!
Im voraus vielen Dank für Ihre Mühe.

Text K

Sosehr die Betrachtung der lernabhängigen Verbesserung von Denkleistungen geeignet ist, bestimmte allgemeine Verlaufsgesetze zu formulieren, sowenig taugt sie doch auch, die subjektiven, individuell wirksamen Faktoren der Entscheidungsbildung über internen kognitiven Strukturen transparent zu machen. Werden hingegen beide Wege unabhängig und gleichzeitig gegangen, so kann es möglich werden, das Individuelle als Komponente genereller Verlaufsgesetze zu erkennen und umgekehrt, die generellen, für die Population getroffenen Aussagen in individuellen kognitiven Mikroprozessen wiederzufinden. Um kognitive Strukturkomponenten in individuellen Problemlösungsansätzen zu identifizieren, betreibt man beschreibende Verhaltensanalyse.

Text P

Spricht aus unserer Auffassung Resignation? Kann die Wissenschaft nur ihre biologische Aufgabe, sich in praktischer Anwendung zu

bewähren, erfüllen — ist ihre intellektuelle Aufgabe unlösbar? Ich glaube nicht. Niemals setzt sich die Wissenschaft das Phantom zum Ziel, endgültige Antworten zu geben oder auch nur wahrscheinlich zu machen; sondern ihr Weg wird bestimmt durch ihre unendliche, aber keineswegs unlösbare Aufgabe, immer wieder neue, vertiefte und verallgemeinerte Fragen aufzufinden und die immer nur vorläufigen Antworten immer von neuem und immer strenger zu prüfen.

				A 1		RT 1		B 1			A 2		RT 2		B 2	
VB	VG	Vp	Text	n(f)	t	Σ	n(r)	n(f)	t	Text	n(f)	t	Σ	n(r)	n(f)	t
I	1	1	P1	129	942	58	55	93	641	K2	109	689	76	74	89	606
		2	K2	125	1079	74	73	106	587	P1	138	664	86	84	116	532
		3	K1	141	943	103	100	105	557	P2	140	829	109	109	109	593
		4	P2	91	748	79	79	54	498	K1	128	690	88	87	87	450
II	2	5	P3	73	386	76	76	48	236	K4	80	359	81	79	63	249
		6	K4	67	505	89	84	61	395	P3	66	335	102	97	45	259
		7	K3	71	693	73	73	53	435	P4	60	497	88	20	44	388
		8	P4	53	418	61	61	29	316	K3	74	364	72	72	63	297
	3	9	P4	46	718	55	55	30	495	K5	57	669	65	65	39	541
		10	K5	59	612	86	75	36	410	P4	61	334	100	97	38	230
		11	K4	57	839	54	53	38	517	P5	44	477	56	55	33	417
		12	P5	50	412	65	62	38	380	K4	59	406	78	71	56	322
III	4	13	P6	34	330	65	64	30	182	K7	49	221	70	63	49	135
		14	K7	43	294	94	88	36	225	P6	43	185	93	86	34	123
		15	K6	45	244	74	74	38	142	P7	36	180	77	76	26	161
		16	P7	50	522	59	57	36	233	K6	44	275	68	68	33	229
	5	17	P7	28	365	67	67	18	185	K8	26	218	78	78	20	164
		18	K8	38	390	73	67	37	251	P7	33	200	91	86	34	183
		19	K7	35	471	49	48	19	323	P8	24	345	60	59	16	251
		20	P8	39	350	52	48	31	218	K7	51	269	71	75	35	194
	6	21	P8	29	617	46	43	19	357	K9	32	468	62	58	17	303
		22	K9	31	304	64	64	13	233	P8	23	247	71	71	11	214
		23	K8	38	1061	60	59	23	466	P9	22	293	60	57	15	283
		24	P9	29	287	70	69	13	238	K8	35	279	71	68	24	225

VB	VG	Vp	Text	A 1		RT 1		B 1		Text	A 2		RT 2		B 2	
				n(f)	t	Σ	n(r)	n(f)	t		n(f)	t	Σ	n(r)	n(f)	t
I	1	25	P1	134	991	59	58	126	594	K2	159	636	63	62	186	405
		26	K2	162	591	74	66	131	576	P1	184	372	86	85	155	412
		27	K1	153	728	68	67	109	549	P2	94	600	77	76	83	520
		28	P2	85	755	80	78	78	502	K1	146	592	95	93	108	441
II	2	29	P3	68	569	74	73	62	338	K4	80	392	81	81	73	346
		30	K4	67	676	40	38	71	500	P3	78	455	48	45	48	415
		31	K3	66	419	70	68	42	279	P4	54	344	79	76	37	288
		32	P4	70	552	79	75	62	351	K3	78	399	89	78	61	295
	3	33	P4	44	530	80	79	43	318	K5	65	388	86	86	40	287
		34	K5	84	390	83	77	82	297	P4	72	295	98	93	54	255
		35	K4	73	514	74	72	65	290	P5	60	319	79	76	51	245
		36	P5	53	453	72	72	42	336	K4	76	450	80	80	62	348
III	4	37	P6	38	366	57	54	28	242	K7	43	281	72	44	45	229
		38	K7	37	438	83	83	20	207	P6	36	207	73	69	21	133
		39	K6	42	312	61	58	38	291	P7	31	318	69	68	15	254
		40	P7	24	426	72	68	26	217	K6	47	236	89	87	29	179
	5	41	P7	32	418	77	76	21	249	K8	27	294	86	86	20	258
		42	K8	38	393	74	69	30	264	P7	33	259	84	82	16	206
		43	K7	44	265	67	63	39	188	P8	36	225	72	70	19	167
		44	P8	39	356	62	59	26	248	K7	45	269	73	70	31	239
	6	45	P8	26	297	73	70	14	205	K9	35	315	86	84	33	226
		46	K9	40	317	68	68	25	217	P8	26	182	77	77	11	154
		47	K8	32	377	76	76	19	268	P9	28	290	82	81	10	200
		48	P9	25	319	74	71	15	167	K8	34	303	81	80	20	173

VG	Vp	Text	A 1		RT 1		B 1		Text	A 2		RT 2		B 2	
			n(f)	t	Σ	n(r)	n(f)	t		n(f)	t	Σ	n(r)	n(f)	t
1	1	P7	30	290	71	*15*	18	195	K9	33	239	72	71	15	186
	2	K9	37	419	58	57	12	217	P7	32	278	67	66	20	170
	3	K7	39	398	53	52	22	265	P9	26	246	63	63	29	193
	4	P9	22	249	61	60	24	214	K7	37	239	71	68	35	192
2	5	P7	37	188	102	98	24	133	K9	39	179	107	106	34	111
	6	K9	36	463	68	67	33	268	P7	31	219	81	80	25	174
	7	K7	44	376	69	68	45	179	P9	47	184	83	81	41	174
	8	P9	31	329	71	68	20	217	K7	40	239	82	82	39	204
3	9	P7	30	238	89	87	29	178	K9	37	252	83	78	33	142
	10	K9	36	452	55	53	32	345	P7	27	260	61	61	21	217
	11	K7	50	270	89	76	34	157	P9	38	155	99	91	26	126
	12	P9	27	270	88	84	27	140	K7	37	270	101	100	35	159
4	13	P7	32	260	85	83	16	155	K9	37	227	95	92	17	156
	14	K9	42	400	80	79	34	245	P7	39	221	82	78	28	164
	15	K7	47	937	52	52	36	570	P9	24	330	81	78	16	240
	16	P9	32	292	74	74	18	124	K7	43	179	84	78	34	174

Rohdaten aus Exp. 2

Bei Vp 1 wurde n(r) in RT 1 von 15 auf 70 abgeändert. Der Anlaß für diese „Korrektur" war derselbe wie in Exp. 1 bei den Vpn 7 und 37 (vgl. „Schätzung der Netto-Denkzeit ‚D'").