

Saroyan, Alenoush

Formative Evaluation von Lehrmaterial. Unterschiede zwischen Instruktionsdesignern und Fachinhalteexperten

Unterrichtswissenschaft 19 (1991) 3, S. 238-259



Quellenangabe/ Reference:

Saroyan, Alenoush: Formative Evaluation von Lehrmaterial. Unterschiede zwischen Instruktionsdesignern und Fachinhalteexperten - In: Unterrichtswissenschaft 19 (1991) 3, S. 238-259 - URN: urn:nbn:de:01111-pedocs-298309 - DOI: 10.25656/01:29830

<https://nbn-resolving.org/urn:nbn:de:01111-pedocs-298309>

<https://doi.org/10.25656/01:29830>

in Kooperation mit / in cooperation with:

BELTZ JUVENTA

<http://www.juventa.de>

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Digitalisiert

Mitglied der


Leibniz-Gemeinschaft

Unterrichtswissenschaft

Zeitschrift für Lernforschung
19. Jahrgang/Heft 3/1991

Thema:

Instruktionstheorie und Instruktionsdesign

Verantwortlicher Herausgeber:
Franz Schott

- Franz Schott:
Instruktionstheorie und Instruktionsdesign: Eine Einführung 194
- Franz Schott:
Instruktionsdesign, Instruktionstheorie und Wissensdesign:
Aufgabenstellung, gegenwärtiger Stand und
zukünftige Herausforderungen 195
- Joost Lowyck, Jan Elen:
Wandel in der theoretischen Fundierung
des Instruktionsdesigns 218
- Alenoush Saroyan:
Evaluation von Lehrmaterial:
Unterschiede zwischen Instruktionsdesignern und
Fachinhalteexperten 238

Allgemeiner Teil

- Jo Kramis:
Eine Kombination mit hoher Effektivität:
Microteaching – Reflective Teaching – Unterrichtsbeobachtung 260

Berichte und Mitteilungen 278

Buchbesprechungen 286

193

Alenoush Saroyan

Formative Evaluation von Lehrmaterial: Unterschiede zwischen Instruktionsdesignern und Fachinhalteexperten

Formative Evaluation:
Problem Identification and Revision by Instructional
Designers and Subject Matter Experts¹

Diese Studie beschreibt den Grad an Konvergenz und Divergenz zwischen Instruktionsdesignern und Fachinhalteexperten bei der Problemidentifikation und der Revision im Rahmen von formativer Evaluation. Sechs Experten, drei Instruktionsdesigner und drei Mikrobiologen überarbeiteten einen Abschnitt eines Selbstinstruktionstextes zur Mikrobiologie mit der Methode des lauten Denkens. Die Daten wurden mit Hilfe eines dafür entwickelten Schemas kodiert. Für die einzelnen Äußerungen wurden absolute und relative Häufigkeiten berechnet und aufgelistet. Der am meisten hervorstechende Unterschied zwischen den zwei Gruppen ist die Vielfalt der hervorgebrachten Kommentare. Die Fachinhalteexperten gaben Kommentare zum Inhalt ab, während die Instruktionsdesigner sich auf Designparameter, Präsentation, Inhalt und Pädagogik bezogen.

This study describes the degree of convergence and divergence in problem identification and revision between instructional designers and subject matter experts in formative evaluation. Six experts, three instructional designers and three microbiologists reviewed a unit from a self-instructional microbiology module while thinking out loud. A coding scheme was developed and applied to the data. Frequencies and mean percentages were tabulated. The most salient difference between the two groups appears to be in the breadth of generated comments. Moreover, subject matter experts comment on issues relating to content while instructional designers address issues relating to design, presentation, content, and pedagogy.

Einleitung und Fragestellung

Formative Evaluation ist eine Komponente von Modellen des Instruktions-System-Designs (ISD). ISD-Modelle sind Heuristiken für die Planung, Entwicklung, Implementation und Evaluation von Unterrichtsmaßnahmen und -material. Bei Unterrichtsmaterialien gewährleistet die formative Evaluation, daß prototypisches Material einem iterativen Prozeß von Prüfen und Revidieren unterzogen wird, bis es für die Anwendung geeignet ist (Abedor, 1971; Baker, 1970; Dick & Carey, 1985; Gagné & Briggs, 1979; Markle, 1976).

Die Funktion des Instruktionsdesigns ist es, ein „perfektes“ Produkt zu liefern. Vor diesem Hintergrund mag formative Evaluation redundant erscheinen. Unbefriedigende Evaluationsergebnisse erinnern uns aber gleich an die heuristische Natur von ISD-Modellen. Dick (1977) nimmt

an, daß die fehlerhaften Ergebnisse der ISD-Modelle auf Mängel in der zugrundeliegenden Theorie zurückgeführt werden können. Gropper (1975) dagegen vermutet, daß die schlechten Ergebnisse auf einer unangemessenen Ausführung der Heuristiken der Modelle beruhen, was oft auf Mangel an Geld und Zeit zurückzuführen sei. Fehler, die auf diesen und anderen Gründen beruhen, können zum Teil durch die Durchführung von formativer Evaluation beseitigt werden. Darauf weisen die Analysen von Andrew & Goodson (1980) und Stolovitch (1982) hin. Sie empfehlen für über 98 % der 52 analysierten Modelle die Durchführung einer formativen Evaluation.

Formative Evaluation wird ausdrücklich als Prozedur verstanden, „... Aspekte des Kurses zu identifizieren, deren Revision wünschenswert ist... aber nicht kurz vor der Vollendung, wenn der Autor natürlich zögert, eine Gesamtheit von fertig gestelltem Material auseinander zu reißen, sondern in der Entwicklung...“ (Cronbach, 1963, p. 675). Scriven (1967), der den Begriff 'formative Evaluation' einführte, verband damit folgende Merkmale: (1) Formative Evaluation wird durchgeführt, nachdem ein erster Entwurf steht, d.h. bevor das Material seine endgültige Form hat und veröffentlicht wird. (2) Die Absicht ist, die Effektivität des Inhaltes zu verbessern. (3) Die Grundlage ist sowohl das Feedback von potentiellen Benutzern, als auch von Experten. (4) Die Ergebnisse können, müssen aber nicht vom Instruktionsdesigner umgesetzt werden. Insbesondere das dritte Merkmal war Gegenstand zahlreicher Untersuchungen. Aus ihnen geht hervor, daß es wichtig ist, sowohl die Lernenden als auch Experten (Instruktionsdesigner und andere) als Quelle des Feedbacks für die formative Evaluation heranzuziehen. Denn jede Gruppe kann unterschiedliche Beiträge leisten, um die Effektivität des Lehrmaterials zu erhöhen (Geis, 1987). So können verschiedene Experten die Qualität und Validität des Materials aus unterschiedlichen Perspektiven kommentieren, und die Lernenden können durch ihre Leistung belegen, ob die Annahmen über ihre Fähigkeiten richtig waren.

Empirische Untersuchungen zur formativen Evaluation haben sich vorwiegend auf Daten von den Lernenden konzentriert, wobei die Erhebung qualitativer und, wenn auch in geringerem Ausmaß, quantitativer Daten im Vordergrund stand (Baghdadi, 1980; Burt, 1989). In anderen Studien wurden die Auswirkungen untersucht, die der Experimentator in verschiedenen Rollen (aktiv oder passiv) auf die Lernenden sowie in Bezug auf das Hervorrufen von Feedback hat (Carroll, 1988; Dick, 1968; Geis, 1988). 'Learner Verification and Revision' (Komoski, 1983; Komoski & Woodward, 1985) und 'Developmental Testing' (Nathenson & Henderson, 1977; 1980) sind zwei Begriffe, die die formative Evaluation auf die ausschließliche Nutzung des Lernenden als Datenquelle begrenzen. Sie zeigen auch die Bedeutung, die dieser Vorgehensweise für die Revision von Lehrmaterial zugesprochen wird.

Das Urteil von Experten hat demgegenüber weniger Aufmerksamkeit auf sich gezogen, obwohl es kosten- und zeitsparend sein kann. Das hat vor allem Bedeutung für die Industrie, da dort Bildungsmaßnahmen während der Arbeitszeit durchgeführt werden (Foshay, 1984). Für MacDonal-Ross (1978) sind Urteile von Experten grundsätzlich die reichhaltigeren und bedeutsameren Daten im Vergleich zu Urteilen der Lernenden.

Ein Grund dafür, daß Expertenurteile nur in einigen Untersuchungen (Golas, 1982; Saroyan, 1989; Saroyan & Geis, 1988) zugrunde gelegt wurden, ist vielleicht, daß formative Evaluation kaum auf einen einzigen Typ von Experten beschränkt werden kann. So kann zum Beispiel die Einbeziehung von Experten in den Prozeß der Entwicklung von Instruktionstexten oder „Textzusammenstellungen“ (Ally, 1985; Nevo, 1977; O'Donnell, 1985; Wright, 1985) folgendermaßen beschrieben werden: Nachdem (a) ein erster Entwurf von Fachinhalteexperten (FIE) und/oder Instruktionsdesignern (ID) vorgelegt wird, kann (b) die Überarbeitung dieser ersten Fassung durch Experten wie z.B. Fachinhalteexperten, Pädagogen, Instruktionsdesigner, Dokumentendesigner, Angehörige der Regierung und von Bürgerinitiativen etc. erfolgen (nach Bowler (1978) können mehr als 17 Typen von Experten unterschieden werden); dann können (c) die verschiedenen Revisionsvorschläge in den Text durch den Autor, Instruktionsdesigner oder auch einen Außenstehenden (wie z.B. den Herausgeber) eingearbeitet werden. Der zuletztgenannte Arbeitsschritt baut auf der Fähigkeit, dem Wissen und dem Fingerspitzengefühl des Revisors auf und ist, ohne jede empirische Grundlage, völlig subjektiv. Flower, Hayes & Swarts (1980) untersuchten die Strategien von Experten in bezug auf die Revision von Gebrauchstexten. Sie nahmen an, daß es keine feststehenden Prinzipien gibt, die den Revisionsprozeß leiten, so daß die Dokumente verschiedenen Revisionsstilen ausgesetzt sind: „Jeder . . . revidiert . . . aufgrund eigener unabhängiger Kriterien . . .“.

Fast alle Studien aus diesem Bereich konzentrieren sich auf das Endergebnis der formativen Evaluation, wie es von einem bestimmten Typ von Experten ausgeführt wird: dem Instruktionsdesigner (cf. Dupont & Stolovitch, 1983; Golas, 1982; Martelli, 1979). Daraus resultiert ein schwerwiegender Mangel. Denn üblicherweise wird die Gestaltung und Überarbeitung von Instruktionstexten entweder von einem Fachinhalteexperten oder einem Instruktionsdesigner durchgeführt. Zwar ist auch eine Zusammenarbeit dieser Experten möglich, aber selten.

Auch in der Literatur wird die unterschiedliche Rolle von IDs und FIEs bei der formativen Evaluation bzw. die Wahrnehmung dieser Rolle nur selten behandelt. Berichtet wird aber, daß Instruktionsdesigner dazu neigen, eher Fragestellungen zu kommentieren, die sich mit Design, Präsentation und Inhalt befassen (Saroyan & Geis, 1988). Die Rolle des fachlichen Inhalts bei der formativen Evaluation wurde dagegen bis jetzt nicht systematisch untersucht.

Die im folgenden zu beschreibende Untersuchung wurde durchgeführt, um gerade diese Leerstelle zu füllen. Es geht darum, die Originalität oder auch die Redundanz von Beiträgen von Instruktionsdesignern und Fachinhalteexperten zur formativen Evaluation zu ergründen. Die spezifischen Ziele der Studie sind, (a) die von Instruktionsdesignern und Fachinhalteexperten bei der formativen Evaluation am häufigsten und mit größter Übereinstimmung identifizierten Probleme und durchgeführten Revisionen zu beschreiben und (b) den Grad von Konvergenz und Divergenz dieser Kommentare zu bestimmen.

Untersuchungsmethode

Um Daten von einem Detailliertheitsgrad zu erhalten, der eine tiefere Analyse der Evaluationsresultate von Instruktionsdesignern und Fachinhalteexperten erlauben, wurde die Methode des lauten Denkens verwendet (vgl. Ericsson & Simon, 1984; Hayes & Flower, 1983). Diese Technik wurde bereits häufig angewendet, z.B. in Untersuchungen zum Problemlöseverhalten von Experten und Novizen in gut definierten Bereichen wie Physik (Larkin, McDermott, Simon & Simon, 1979), Spielen wie Schach und GO (de Groot, 1966) und in schlecht definierten Bereichen wie dem Verfassen von Texten und deren Revision (Breuxleux, 1987; Flower & Hayes, 1977; Swarts, Flower & Hayes, 1984) sowie musikalischen Kompositionen.

Experimentelle Aufgabe

Instruktionsdesigner und Mikrobiologen (als FIEs) wurden gebeten, eine formative Evaluation einer Einheit eines umfassenden Instruktionstextes mit einer Einführung in die Mikrobiologie durchzuführen und dabei ihre Gedanken laut zu äußern. Die Probanden wurden schriftlich instruiert, den Text so zu revidieren, daß sie das Gefühl haben könnten, er wäre für eine Publikation und Anwendung im Unterricht geeignet. Des weiteren wurden die Probanden informiert, daß jede Veränderung des Textes akzeptiert würde, außer wenn sie glaubten, den Text völlig neu schreiben zu müssen. Alle verbalen Äußerungen während der Sitzung wurden aufgenommen. Die Dauer der Bearbeitung war nicht begrenzt. Die Probanden benötigten zwischen einer und eineinhalb Stunden.

Probanden

Sechs Experten nahmen an der Studie teil. Drei davon waren professionelle Instruktionsdesigner und die drei anderen Fachinhalteexperten (Mikrobiologen), die auch Einführungskurse in Mikrobiologie gaben.

Ein Einblick in die Literatur vermittelt den Eindruck, daß der wichtigste Faktor, der jemanden zum Experten macht, die Jahre an Erfahrung und beruflicher Praxis in einer bestimmten Domäne sind. Dies erlaube eine angemessene Entwicklung von deklarativem und prozeduralem Wissen (Anderson, 1982). Dieses Kriterium wurde in verschiedensten

Forschungsgebieten verwendet, wie z.B. Kartenlesen (Thorndike & Stasz, 1980), Elektronik (Egan & Schwartz, 1979), Problemlösen in der Physik (Simon & Simon, 1979, Revision (Beach, 1976; Faigley & Witte, 1981 Hayes, Flower, Schriver, Stratman & Carey, 1987; Sommers, 1980), und dem Planungsprozeß des Schreibens (Breuxleux, 1987). Es lag daher nahe, die Dauer der beruflichen Erfahrung in einem Bereich als Kriterium auch für die eigene Untersuchung zugrunde zu legen. Als weitere Indikatoren für die berufliche Erfahrung galten akademische Grade, Publikationen, Verfügbarkeit und die Bereitschaft, an der Studie teilzunehmen. Alle Experten hatten mindestens eine zehnjährige Berufserfahrung, die Mikrobiologen hatten zwischen 10 und 18 Jahren Lehrerfahrung auf College- und Universitätsniveau. Vier der Probanden, davon zwei in jeder Gruppe, hatten Dokortitel, während die übrigen das Diplom hatten. Einer der Mikrobiologen, ein Universitätsprofessor, war der Autor von mehr als 16 einführenden Textbüchern in Mikrobiologie. In Entsprechung dazu hatten die drei Instruktionsdesigner nach einer mindestens fünfjährigen Universitätskarriere erfolgreiche Trainings- und Beratungsfirmen aufgebaut. Eine der Routineaufgaben ihrer Tätigkeit war, mit verschiedensten Firmen zu arbeiten, um Instruktionsmaterial für vielfältige Trainingsanforderungen zu entwickeln und zu evaluieren. Keiner der Probanden in der einen Gruppe hatte spezifische Kenntnisse in den Qualifikationen der anderen Gruppe.

Material

Das Material, das für die Untersuchung verwendet wurde, war aus einem ersten Entwurf eines Selbstinstruktions-Moduls entnommen, das für Studenten der Zahnmedizin im ersten Semester entworfen worden war (D.A.E. Project, 1978). Das für die Entwicklung dieses Materials verantwortliche Team bestand aus Fachinhalteexperten, Instruktionsdesignern und Evaluatoren. Von den sechs Unterrichtseinheiten mit dem Titel „*The Relationship of Microorganisms to Environmental Conditions. Disease and Immunity*“ wurde nur die dritte verwendet. (Die ersten beiden Einheiten bestanden aus einem Vorwort und einer Einleitung und waren nicht hinreichend spezialisiert, um den Inhalt für die Instruktionsdesigner fremd zu machen.) Die Einheit bestand aus acht Seiten Instruktionstext über Umweltbedingungen und die Rolle von Mikroorganismen für Krankheiten. Der Text war mit einem Vorwort versehen, das eine Menge von Lehrzielen umfaßte. Außerdem war am Ende ein Text mit vierzehn „Multiple Choice-“ und anderen Fragen zum Ausfüllen angefügt. Ebenso war ein Antwortbogen zu den Fragen beigelegt. Zwei Gründe sprachen für die Verwendung dieses Versuchsmaterials: Erstens stellte der spezialisierte Inhalt sicher, daß die Experten aus jeder Gruppe kein Wissen über die Domäne der anderen Gruppe hatten. Zweitens lag der Text als ein erster Entwurf vor und wies noch mehrere Mängel auf wie z.B. Tippfehler, falsche Abstände, schlechtes Layout und Inkonsistenzen im Format.

Vorgehensweise

Jede experimentelle Sitzung bestand aus (a) einer Einführung in die Methode des lauten Denkens, (b) einigen Aufwärmübungen im lauten Denken, (c) der formativen Evaluation des Versuchsmaterials, und (d) einem retrospektiven Interview.

Die Phase (a) war erforderlich, da das Verbalisieren von Gedanken während des Lösen einer Aufgabe für alle Probanden eine neue Aktivität war. Die Aufwärmübungen gaben den Probanden die nötige Praxis im lauten Denken. Nachdem die Probanden angaben, mit der Vorgehensweise des lauten Denkens genügend vertraut zu sein, wurde ihnen die Beschreibung der Aufgaben vorgelegt. Es wurde ihnen genug Zeit eingeräumt, um die Anweisungen zu lesen und Fragen zu klären. Dann folgte das eigentliche Experiment.

Während der Durchführung der Evaluationsaufgabe wurden die verbalen Äußerungen der Probanden aufgenommen. Die Kommunikation zwischen dem Experimentator und den Probanden war eingeeignet auf (a) Aufforderungen, wenn sie länger als zehn Sekunden eine Sprechpause machten, und (b) Antworten zu aufgabenbezogenen Fragen. Solche Unterbrechungen waren allerdings selten. Retrospektive Interviews wurden durchgeführt, nachdem die Aufgabe vollständig beendet war und nur im Falle, daß der Experimentator bestimmte Kommentare unklar oder zweideutig empfand. Die Daten aus diesen Interviews wurden lediglich als bestätigendes Material verwendet.

Analysemethode

Alle verbalen Äußerungen wurden wörtlich abgeschrieben, um Protokolle des lauten Denkens zu produzieren. Sodann wurde der Inhalt eines jeden Protokolls analysiert und mit Hilfe Winograds (1983) System der Satzanalyse in einzelne Satzeinheiten zerlegt. Diese Segmente wurden nach Kategorien sortiert, die die Elemente repräsentieren, welche die formative Evaluation konstituieren. Diese Elemente waren im Vorfeld der Studie mittels einer Aufgabenanalyse identifiziert und mit Pilotdaten getestet worden. Die Reliabilität zwischen zwei unabhängigen Kodierern lag bei 89 %.

Das gesamte Kodierschema bestand aus 13 Kategorien. Eine kurze Beschreibung jeder Kategorie und Beispiele sind im folgenden dargestellt:

Evaluative Feststellung: Eine Feststellung, die einen Vergleich zwischen einer internalen Repräsentation eines Zustands oder eines Ereignisses mit einer externalen Repräsentation desselben repräsentiert. Aufgrund dieser Definition repräsentiert diese Kategorie positive und negative Kommentare (die nicht explizit die Quelle des Problems benennen), Ausdrücke von Bevorzugung, Urteil und internen Gefühlen.

Problemidentifikation: Explizite Bezugnahme auf ein Problem im Text.

Revisionsvorschläge: Explizite, auf den Text bezogene Veränderungen, die darauf abzielen, den gegebenen Zustand in den Zielzustand zu überführen.

Feststellungen persönlichen Wissens: Ausdrücke, die persönliches Wissen widerspiegeln und sowohl deklaratives als auch prozedurales Wissen beinhalten. Diese Feststellungen werden häufig zusätzlich gemacht oder an Stelle von Problemidentifikationen oder als Angabe des Grundes für eine spezifische Veränderung.

Textwissen: Kommentare, die direkt vom Text erworbenes Wissen repräsentieren.

Wörtliche Feststellungen: Abschnitte, die wörtlich vom Text abgelesen werden.

Paraphrasierte Textwiedergabe: Bezüge zu verschiedenen Teilen des Texts, aber nicht wörtlich abgelesen.

Taktische Kommentare zur Aufgabe: Bezüge zur vorliegenden oder einer unmittelbar folgenden Aufgabe.

Strategische Kommentare zur Aufgabe: Hinweis auf Tätigkeit, die normalerweise durchgeführt wird und nicht unmittelbar mit der Aufgabe zu tun hat.

Dialog: Fragen an den Versuchsleiter.

Grenzmarkierungen: Lexikalische Abgrenzungen von Segmenten.

Falscher Anfang: Ein unvollständiges Segment, dem eine vervollständigte Version desselben folgt.

Beziehungslose Äußerungen: Kommentare, die nicht mit der Aufgabe oder dem Text in Beziehung stehen.

Aus den sortierten Segmenten jedes Expertenprotokolls wurde eine Datenbank geschaffen. Die Häufigkeiten wurden in einer Tabelle festgehalten und in relative Häufigkeiten umgerechnet. Diese Werte basierten ursprünglich auf der gesamten Anzahl der Segmente pro Protokoll und wurden später auch unter Auslassung irrelevanter Kategorien (Dialog, Falscher Anfang, Beziehungslose Äußerungen, Grenzmarkierungen) und Wiederholungen berechnet. Auf diese Daten wurden deskriptive statistische Operationen angewendet, um Gruppenmittelwerte vergleichen zu können.

Für die qualitative Analyse der Daten wurden mehrere kleine Kodierschemata angewendet. Ursprünglich wurden Problemidentifikation und Revisionsvorschläge als Inhalt, Design, Präsentation oder zur Pädagogik gehörend klassifiziert. (Zur Beschreibung dieser vier Kategorien, siehe in Saroyan & Geis, 1988). Die Konvergenz und Divergenz in der Problemidentifikation und in den Revisionsvorschlägen zwischen den zwei Gruppen von Experten wurde dadurch bestimmt, daß die aufgezeichneten Kommentare aus einem Abschnitt (Wort, Phrase, Segment, Satz, Paragraph) von zwei oder mehr Experten verglichen wurden. Diese Datenbasis machte es möglich, folgendes zu identifizieren:

- (a) Segmente, die bei mehr als einem Experten eine Problemidentifikation und/oder einen Revisionsvorschlag ausgelöst hatten;
- (b) Ähnlichkeiten und Unterschiede zwischen diesen Problemidentifikationen und/oder Revisionsvorschlägen über die Experten hinweg.

Ergebnisse

Im folgenden werden einige grundlegende Ergebnisse vorgestellt, um den Kontext zu schaffen für die qualitative Datenanalyse. Der Prozentsatz von Problemidentifikationen und Revisionsvorschlägen eines jeden Experten ist in Tabelle 1 dargestellt.

Tabelle 1:
Prozentsatz von Problemidentifikationen und Revisionsvorschlägen

Äußerung	FIE I	FIE II	FIE III	ID I	ID II	ID III
Problem-identifikation	5.9	9.3	7.6	11.5	17.7	12.3
Revision	12.5	11.3	10.4	6.4	9.5	12.0
Anzahl der Segmente	334	193	458	216	439	794

Es fällt auf, daß die FIEs weniger Problemidentifikationen und mehr Revisionsvorschläge äußerten, während es bei den Instruktionsdesignern genau umgekehrt war. Daraus kann abgeleitet werden, daß die durchschnittliche relative Häufigkeit der Problemidentifikationen bei den IDs größer war als bei den FIEs: $X = 13.8\%$ (SD 3.3) versus $X = 7.6\%$ (SD 1.7). Die durchschnittliche relative Häufigkeit für Revisionsvorschläge war bei den FIEs höher als bei den IDs ($X = 11,4\%$ (SD 1.0) versus $X = 9.3\%$ (SD 2.8)).

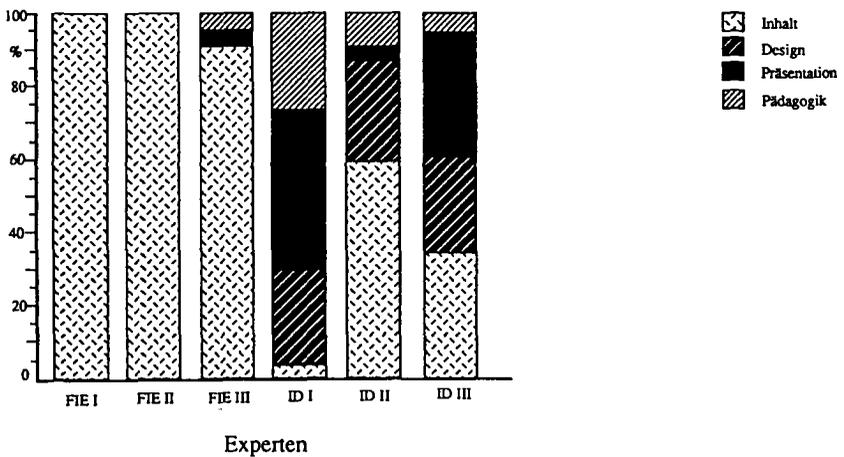


Abb. 1: Art der Problemidentifikation der Experten

Bei einer Einordnung der Problemidentifikationen in die Kategorien Inhalt, Design, Präsentation und Pädagogik zeigte sich, daß fast alle Äußerungen der FIEs auf den Inhalt bezogen waren, während die Kommentare der IDs sich auf alle möglichen Kategorien bezogen. Abbildung 1 zeigt diese Information.

Man sieht deutlich, daß Kommentare in bezug auf den pädagogischen Aspekt des Textes in den Protokollen der IDs häufiger waren als in denen der FIEs (15.8% > 1.9%).

Das Sortieren der Revisionsvorschläge in die vier Kategorien Inhalt, Design, Präsentation und Pädagogik ergab fast dieselben Ergebnisse. Die Mehrzahl der Kommentare der FIEs bezog sich auf die Inhaltskategorie, während die Empfehlungen der IDs sich sowohl auf den Inhalt, als auch auf die Kategorien Design und Präsentation bezogen. Empfehlungen für pädagogische Verbesserungen fehlten in allen Protokollen. Dies wird in Abbildung 2 dargestellt.

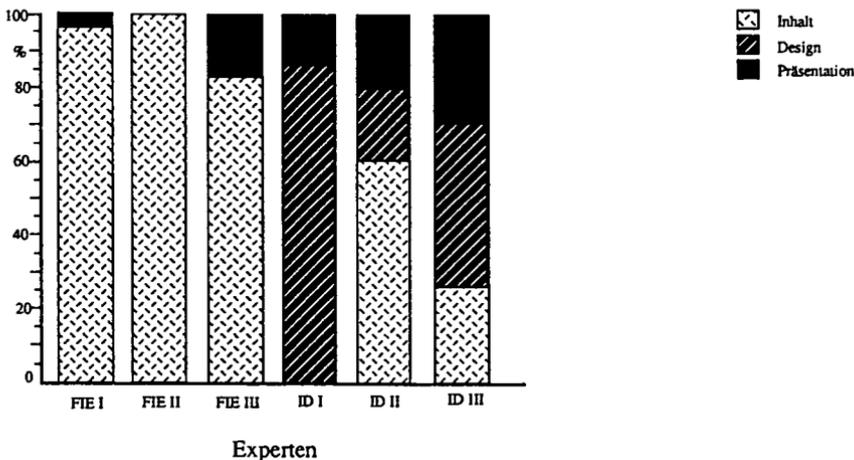


Abb. 2: Art der Revisionsvorschläge der Experten

Betrachtet man die Arten von Problemen, die die FIEs und die IDs identifizierten, sind zwei Aspekte von Interesse: (a) die Probleme, die die Aufmerksamkeit von allen Experten in einer Gruppe auf sich zogen (bei Vernachlässigung der Häufigkeit), und (b) die Probleme, die die höchste kumulative relative Häufigkeit in einer Gruppe aufwiesen. Die am hervorstechendsten Ergebnisse werden hier auf der Basis der kumulierten relativen Häufigkeiten innerhalb der Gruppen dargestellt:

Auf den Text bezogene Problemidentifikationen, die von allen drei FIEs angeführt wurden, waren beschränkt auf Angemessenheit des Inhalts (82%), Fehler in der Rechtschreibung (24,4%) und Redundanz im Inhalt (21,2%). Die auf den Text bezogenen Probleme, auf die die IDs am

häufigsten hinwies, beinhalten Angemessenheit des Inhalts (82,5%), Richtigkeit des Inhalts (47,2%), Klarheit der Beschreibungen (35%), der Grad der Spezifität des Inhalts (25%) und Rechtschreibung (24,4%). In bezug auf die IDs waren die Probleme, die von allen drei identifiziert wurden, die Angemessenheit des Inhalts aus der Sicht der Lernenden (44,9%), der pädagogische Wert des Textes (34,4%), Übereinstimmung zwischen den verschiedenen Teilen (wie z.B. Lehrziele mit dem Inhalt und dem Test: 30,7%), schlechtes Design (30,7%) und allgemeines Layout und Präsentation (22,4%). Mit einer Ausnahme, nämlich der Klarheit der Beschreibungen (21,9%), identifizierten die IDs am häufigsten dieselbe Art von Problemen.

Die Empfehlungen in bezug auf die Revisionsvorschläge bestanden auf seiten der Fachinhalteexperten darin, Inhalt hinzuzufügen (63,5%) und Rechtschreibfehler zu berichtigen (31,5%). Die am häufigsten vorgeschlagenen Änderungen beinhalteten, den Inhalt zu vervollständigen (63,5%) und den Inhalt relevanter zu machen (64,8%). Die drei Instruktionsdesigner gaben gleichfalls nur zwei Arten von Empfehlungen, nämlich eine Einleitung oder Struktur (20,6%) sowie eine Zusammenfassung (19,9%) hinzuzufügen. Zusätzlich dazu wurden Revisionen in bezug auf die relevanten Inhalte (28,5%) gefordert; des weiteren forderte man Ergänzungen des Inhalts (25,4%), Hinzufügung von Stichwörtern und Hervorhebungen (19,2%), am Ende eines jeden Abschnitts sollte ein kleines Quiz stehen (18,5%), der Text sollte logischer aufgebaut werden (18,3%), und Diagramme sollten in der größeren Nähe des entsprechenden Textes plaziert werden (16,2%).

Die dann erfolgenden Analysen legten fest, wie oft (a) Übereinstimmung (Konvergenz) und (b) unterschiedliche Meinungen (Divergenz) in der Problemidentifikation und in der Revision zwischen den IDs und den FIEs auftraten. Die Ergebnisse zeigen, daß es zwei verschiedene Arten von Problemidentifikationen und Revisionsvorschlägen gibt: Entweder waren sie spezifisch und auf einen bestimmten Abschnitt des Textes bezogen oder allgemein und auf den Text als Ganzes bezogen. Diese Ergebnisse werden nun etwas detaillierter beschrieben:

Problemidentifikationen und Revisionen

Spezifische Kommentare

Insgesamt gab es in dem Text 39 Abschnitte, die von *mehr als einem* Experten (beider Gruppen) mit Kommentaren zur Problemidentifikation oder zur Revision bedacht wurden. Diese Abschnitte stellen knapp 22% des gesamten Textes dar. Die Häufigkeit, mit der zwei oder mehr Experten einen bestimmten Abschnitt kommentierten, ist in Tabelle 2a dargestellt.

Häufigkeit der Kommentare, die sich auf einen Abschnitt bezogen von FIEs und IDs:

Häufigkeit	FIE	ID
9	1	1
8	2	1
4	1	2
3	3	1
3	2	2
3	3	0
2	1	3
2	2	0
2	0	3
1	2	3
1	0	2
1	3	3

Tabelle 2a

Häufigkeit der Konvergenz in der Problemidentifikation und Revision zwischen den Experten:

Häufigkeit	FIE	ID
13	1	1
8	0	2
6	2	0
2	3	1
2	2	1
1	3	0
1	1	2

Tabelle 2b

Tabelle 2a kann man entnehmen, daß es nur eine Stelle im Text gab, die alle Experten kommentierten und ein Problem identifizierten oder eine Revision vorschlugen (letzte Zeile). Die Vielfalt der Kommentare für eben diesen Abschnitt kann über Ausschnitte aus den sechs Protokollen wie folgt belegt werden:

Ausschnitt aus dem Text:

Virus: Viren brauchen bestimmte Umweltbedingungen, um zu wachsen und sich zu vermehren. Verschiedene Viren brauchen verschiedene Bedingungen. Die Existenz lebenden Gewebes ist eine notwendige Bedingung für das Wachstum eines Virus.

Eine Übereinstimmung unter den Experten, daß ein Abschnitt einen Kommentar verdient, bedeutet nicht unbedingt, daß auch eine Übereinstimmung in bezug auf die Problemidentifikation oder die Revisionsnotwendigkeit vorliegen mußte. Solche Variationen in den Kommentaren wurden als Divergenz behandelt. Lag aber Übereinstimmung vor, wurden sie als Indikatoren für Konvergenz gedeutet.

Divergenz

Die Divergenz unter den Fachinhalteexperten war etwas geringer (3 versus 4 Fälle) als unter den Instruktionsdesignern. Mangelnde Übereinstimmung kam meistens aufgrund des Grades der Spezifität des identifizierten Problems zustande. So kam beispielsweise vor, daß ein FIE falsche Überschriften als das Problem ansah, während ein anderer auf allgemeine Inkonsistenzen in Struktur und Präsentation abhob. Nachstehend ist ein weiteres Beispiel für den Umfang und die Vielfalt der Kommentare wiedergegeben, die durch einen Abschnitt ausgelöst wurden.

EXPERTE	Ausschnitt aus dem Protokoll	Art des Kommentars
FIE I:	- Ohne lebendes Gewebe in einer lebenden Zelle würden Viren nicht wachsen und das sollte hier stehen/	Der Inhalt ist unangemessen Inhalt hinzufügen
FIE II:	- Das ist eine Beschreibung, als wären Viren extrazellulär/ - Das stimmt nicht/ - Was alle Viren brauchen, ist die Gegenwart einer lebenden Zelle/ - Diese lebende Zelle muß empfänglich sein für die Invasion aus der Umgebung und das ist die einzige Bedingung/	Inhalt ist unangemessen Ersetze den Inhalt*
FIE III:	- Dieser Abschnitt muß ergänzt werden/	Inhalt ist unangemessen Inhalt hinzufügen* Mit vorhandenem Wissen verbinden*
ID I:	- Ich finde, daß diese Überschrift "Virus" nicht aussagt, ob das etwas ist das ein Mikroorganismus nicht braucht, aber Viren brauchen/	Keine Übereinstimmung zwischen Text und Lehrzielen Unklarer Inhalt Beziehung zum vorhergehenden Inhalt unklar
ID II:	- Es ist eine Tatsache, daß Viren gar nicht in den Lehrzielen aufgeführt sind/ - Was ist der Unterschied (zwischen Mikroorganismen und Viren) und wie bezieht sich der Unterschied auf das, was ich gerade gelernt habe/ - Wahrscheinlich hätte ich eine kleine Zusammenfassung mit einer Tafel oder einem kleinen Test zum Üben eingefügt/	Falsch plazierter Inhalt Verbindung herstellen mit dem vorher aufgeführten Inhalt* Zusammenfassung/Tafeln und Übungsaufgaben hinzufügen*
ID III	- Ich frage mich, warum dieser Virus hier ist/ - es wird ein neuer Terminus eingeführt und ich bin nicht sicher wie er rein paßt/	Falsch plazierter Inhalt Die Beziehung zum vorherigen Inhalt ist unklar

● = Revisionsvorschläge

Tabelle 3

Ausschnitt aus dem Text:

Mikroorganismen im Körper: Manche Mikroorganismen sind für den Mann hilfreich und manche sind schädlich. Es ist möglich, daß manche Mikroorganismen in kleinen Mengen keinen Einfluß auf den Körper haben, aber in großen Mengen Krankheiten auslösen können. Die Antwort des Körpers auf Mikroorganismen ist ganz verschieden.

EXPERTE	Ausschnitt aus dem Text	Kommentar
FIE I:	- Ich finde das ein wenig oberflächlich/	Inhalt ist unangemessen
FIE II:	- Na, es wäre besser "Mensch" anstatt "Mann" zu nehmen/	Keine sexistische Sprache verwenden [●]
FIE II:	- du erwähnst hier nicht auf welche Weise die verschiedenen Mikroorganismen in den Körper gelangen, ob es über eine Verletzung geht oder ob die Person sie mit dem Essen aufnimmt/	Inhalt ist unangemessen Inhalt hinzufügen [●]
ID II:	- Ich habe eine Frage in Bezug auf was den Unterschied zwischen Mikroorganismen verursacht und warum der Unterschied wichtig ist?/ - sag mir warum es wichtig ist/	Inhalt ist unangemessen Inhalt hinzufügen [●]
ID III:	- Ich möchte eine Art Markierung, ein Zeichen, daß ich zu einem neuen Abschnitt übergeben/ etwas, daß mir jetzt sagt, daß wir wissen, daß es Umgebungsfaktoren gibt, laß uns die Rolle der Mikroorganismen in der Verursachung von Krankheit anschauen./	Schlechte Präsentation Hinweise verwenden um neue Konzepte einzuführen und sie mit vorherigen Konzepten zu verbinden [●]

● = Revisionsvorschläge

Tabelle 4

Konvergenz

Problemidentifikationen und Revisionsvorschläge, die auf dasselbe abzielten, sind in der Tabelle zusammengefaßt. Wenn beispielsweise ein Experte sagte — „*expecially*“ sollte „*especialy*“ heißen — und ein anderer sagte — *das ist ein Druckfehler*, wurden die beiden Aussagen als konvergent behandelt. Es gab 33 Fälle, in denen mehr als ein Experte dasselbe Problem identifizierte oder dieselbe Revision in bezug auf einen bestimmten Abschnitt empfahl. Tabelle 2b (siehe oben) zeigt die Häufigkeit der übereinstimmenden Kommentare über die Experten hinweg.

Die größte Häufigkeit (13) repräsentiert die Anzahl der Fälle, in denen ein FIE und ein ID, also Probanden aus unterschiedlichen Fachgebieten, in Problemidentifikation und/oder Revisionsvorschlag übereinstimmten. Die Aspekte, in denen sich Übereinstimmung zeigte, waren in absteigender Reihenfolge (Anzahl in Klammern): Klarheit des Inhalts (4), Angemessenheit des Inhalts (2), Rechtschreibfehler (2), Redundanz (1), Langatmigkeit (1), Klarheit eines Diagramms (1), Format (1), Angemessenheit eines Tests (1).

Experten derselben Herkunft stimmten häufiger überein in der Problemidentifikation und den Revisionsvorschlägen als Experten unterschiedlicher Herkunft. Übereinstimmungen zwischen Fachinhaltsexperten scheinen häufiger vorzukommen als zwischen Instruktionsdesignern. Zwischen beiden Gruppen gab es nur in zwei Fällen Konvergenz und zwar in bezug auf unangemessenen Inhalt und fehlerhafte Rechtschreibung. Eine detailliertere Beschreibung der Problemidentifikationen und Revisionsvorschläge ist in Tabelle 5 dargestellt.

Wie wir schon festgestellt haben, gab es keinen Fall, in dem mindestens zwei Experten von jeder Gruppe in Problemidentifikation und Revision übereinstimmten. Typischerweise war die Konvergenz zwischen den beiden Gruppen auf den Inhalt (Redundanz, Unangemessenheit, Fehler oder Doppeldeutigkeiten), auf die Präsentation (Format, Abbildungen, Ausrichtung des Textes), Rechtschreibfehler und die Menge an praktischen Übungen bezogen.

In den Fällen, in denen die FIEs und die IDs übereinstimmten, reflektierten ihre Kommentare häufig völlig verschiedene Perspektiven. Wurde zum Beispiel ein Abschnitt als unangemessen identifiziert, dann betrachteten die FIEs den Inhalt als unangemessen, weil ein bestimmtes Konzept vom Standpunkt des Fachexperten nicht genügend entwickelt worden war. Der ID beurteilte dagegen den Inhalt vom Standpunkt des Lernenden, der mit neuem Material konfrontiert wird.

Innerhalb der Gruppen war die Übereinstimmung insgesamt bei den FIEs größer (10 versus 7 Fälle). Es gab 4 Fälle, in denen alle drei FIEs in ihren Kommentaren übereinstimmten, und 6 Fälle, in denen nur 2 übereinstimmten. Dagegen gab es keinen Fall, in dem alle drei IDs übereinstimmten, und 7 Fälle, in denen zwei von ihnen übereinstimmten.

KONVERGENZ IN PROBLEMINIDENTIFIKATION UND REVISION

Kommentar	FIE	ID	Häufigkeit
Falscher Inhalt	2	0	1
	3	0	1
	3	1	1
	0	1	1
Unangemessener Inhalt	1	1	3
	2	0	1
	2	1	1
	3	1	1
Zweideutiger Inhalt	1	1	3
	2	0	2
	2	1	1
Schlecht organisierter Inhalt	2	0	2
	0	2	2
Tipp- und Schreibfehler	1	1	2
	3	1	1
Fehlende Übungsaufgaben	1	1	1
	0	2	1
Herausgerichteter Text	1	2	1
	0	2	1
Inkonsistentes Format	1	1	1
	0	2	1
Übereinstimmung zwischen den Teilen	0	2	1
Schlechte Formatierung	0	2	1
Redundanter Inhalt	1	1	1
Unklares Diagramm	1	1	2

Tabelle 5

Unter den FIEs stimmten zwei in Problemen in bezug auf nebensächlichen und unangemessenen Inhalt, Rechtschreibfehler, Langatmigkeit und falschen Gebrauch von Termini überein. Zwei Instruktionsdesigner vertraten die gleiche Meinung in bezug auf Inkonsistenzen zwischen Text und Lehrzielen, nebensächlichem und unklarem Inhalt und schlechtem Format.

Der folgende Ausschnitt zeigt die Konvergenz unter den Experten an zwei Beispielen je eines FIE und eines ID:

Träger der Infektion: Ein „Träger“ ist eine Person, die den krankheitserregenden Mikroorganismus in sich trägt oder in sich hat, aber

EXPERTE	Ausschnitt aus dem Protokoll	Art des Kommentar
FIE I:	- "jemals" nicht "jedes"/ - das ist also ein Tippfehler/ - Die meisten Krankheiten werden durch Träger übertragen, also würde ich diese Aussage hineinbringen/	Tippfehler Inhalt ist unangemessen Inhalt hinzufügen*
FIE II:	- "epfänglich" sollte "empfänglich" heißen/ - das "m" fehlt/ - "jedes" soll "jemals" heißen	Tippfehler Tippfehler (FIE I)
FIE III:	- Ich verändere nur das "jedes" zu "jemals"/ - Es ist ein Tippfehler/ - später würde ich mehr ins Detail gehen: wo trägt der Träger den Virus im Körper? Was passiert im Körper der Träger, so daß sie die Krankheit nicht selber bekommen? Wie kommt es, daß der Virus in ihrem Körper leben kann, ohne von Antikörpern oder weißen Blutkörperchen getötet zu werden?/	Tippfehler verbessern* (FIE I & FIE II) Tippfehler Inhalt ist unangemessen Inhalt hinzufügen*
ID I:	- "Träger der Infektion" ist keine Infektionsart/ - wenn das verwendet wird, dann sollte es am Anfang vom Abschnitt Infektionsarten, genauso wie in der Einleitung verwendet werden/	Inhalt ist falsch Inhalt neu organisieren* Inhalt in der Einleitung hinzufügen*
ID II:	- "Träger der Infektion" ist keine Infektionsart/ - es muß nicht unbedingt in diesem Abschnitt über Infektionsarten sein/ - es sollte da sein, wo du über die Übertragung von Infektionen sprichst, zu einem gewissen Grad/	Inhalt ist falsch Inhalt neu organisieren*
ID III:	- gut, da ist ein Tippfehler/ - "jedes" sollte "jemals" sein/ - "Träger der Infektion" ist kein Begriff aus der Liste (der Lehrziele)/ - wieder verstehe ich nicht warum es unterstrichen ist/ - wenn es ein Begriff ist der wichtig ist, dann denke ich, daß ich ihn in meine (Lehrziel-) Liste aufnehme/	Tippfehler (FIE I, II, III) Tippfehler korrigieren* Keine Übereinstimmung zwischen Lehrzielen und Inhalt inkonsistente Präsentation Lehrziele ergänzen*

* = Revisionsvorschläge

Tabelle 6

keine beobachtbaren Symptome der Krankheit zeigt. Dennoch kann die Person den krankheitserregenden Mikroorganismus und die Krankheit auf ein anderes Individuum, das für diesen spezifischen Mikroorganismus empfänglich ist, übertragen. Zum Beispiel kann ein Individuum den Mikroorganismus, der Hepatitis verursacht, in sich tragen, aber keine Symptome der Krankheit zeigen, oder Hepatitis bekommen. Wenn der Träger in Kontakt kommt mit anderen, kann er/sie die Krankheit auf die anderen Individuen übertragen.

Allgemeine Kommentare

Neben den Problemidentifikationen und Revisionsvorschlägen, die sich auf identifizierbare Abschnitte des Textes bezogen, gab es auch Kommentare, die sich auf den Text als Ganzes bezogen. Diese Kommentare wurden ausschließlich von den IDs abgegeben, entweder nach einem flüchtigen Überblick über das gesamte Material oder am Ende, nachdem sie die spezifischen Bemerkungen gemacht hatten.

HEURISTIKEN	ART DES KOMMENTARS
Übereinstimmung und dem zwischen den Testitems	Übereinstimmung zwischen den Lehrzielen Text, Lehrzielen und Test, Text und
Teilen: Lehrziele:	in verhaltenstheoretischen Begriffen ausgedrückt; valide
Formatierung: Konzepte;	Layout; Typendichte; Abgrenzung der ineffektive Verwendung von
Überschriften und	Unterüberschriften;
Illustrationen:	Klarheit; Mangel an Klarheit;
Übungen:	unangemessen;
Abfolge:	Bezugnehmend auf vorhandenes Wissen; Organisation der Konzepte; sinnvolle
Struktur; Rationale:	Zweck des Materials; Relevanz;
Zusammenfassung:	zusammenfassende Paragraphen/Tafeln für Abschnitte;

Tabelle 7

Diskussion

Die Ergebnisse, die im vorherigen Abschnitt vorgestellt wurden, zeigen Unterschiede in der Problemidentifikation und Revision, wenn Fach-

inhaltsexperten oder Instruktionsdesigner formative Evaluation durchführen. In Anbetracht dessen, daß nur eine geringe Anzahl an Probanden in die Studie einbezogen war, sollte man die Ergebnisse mit Vorsicht interpretieren. Dennoch haben die deskriptive Natur dieser Untersuchung und die Methode geholfen, quantitative Daten zu generieren und zu sortieren. Damit eröffnet die Untersuchung eine neue Forschungsdimension in diesem Bereich.

Der wichtigste Unterschied zwischen den Gruppen scheint in der Vielfalt der hervorgebrachten Kommentare zu liegen. Die FIEs kommentieren auf lokaler Ebene und bezogen sich auf einen spezifischen Abschnitt des Textes. Die IDs gaben ihre Kommentare auf zwei Ebenen ab: (a) auf lokalem Niveau in bezug auf einen spezifischen Abschnitt des Textes, der evaluiert wird, und (b) auf einem globalen Niveau, d.h. in bezug auf den ganzen Text.

Der zweite Unterschied zwischen den zwei Gruppen besteht in den Merkmalen des Textes, auf die sie sich konzentrieren. Die FIEs kommentierten Merkmale des Inhalts: Richtigkeit, Angemessenheit und Rechtschreibung. IDs beurteilten, ob die Teile mit den Lehrzielen übereinstimmen, Schriftbild, Absätze, Illustrationen etc. Dies legt die Vermutung nahe, daß sich die IDs an die Heuristiken halten, die von den ISD-Modellen empfohlen werden. Die Merkmale des Textes, zu denen die IDs in dieser Studie Kommentare abgaben, entsprechen auch den Merkmalen, die in Studien als Attribute von effektivem Instruktionstext beschrieben werden. So haben beispielsweise Burkholder (1981-82), Choi, Merrill, Callahan, Hawkins & Norton (1979) und Merrill, Reigeluth & Faust (1979) die „Konsistenz“ und „Angemessenheit“ als zwei der wichtigsten Attribute identifiziert. Weiterhin wurde vorgeschlagen, sich an allgemeine Präsentationsprinzipien zu halten, zum Beispiel Konsistenz im Schriftbild zu bewahren und Diagramme im Text zu gebrauchen (Felker, Redish & Peterson, 1985; Hartley, 1978; Smillie, 1985; Wright, 1985).

Die Ergebnisse der Untersuchung sind bedeutsam, weil sie die Schlußfolgerung nahelegen, daß Instruktionsdesigner eher auf der Grundlage eines Systems als auf der Grundlage von Intuition operieren, und daß sie bis zu einem gewissen Grad einen hohen Leistungsstandard halten. Gleichwohl zeigt sich hier auch ein Nachteil: Stellt man die Evaluation auf Konsistenz und Angemessenheit ab, muß man davon ausgehen können, daß der Inhalt richtig und relevant ist. Die Daten haben gezeigt, daß die Wahrscheinlichkeit sehr gering ist, daß IDs falsche Inhalte entdecken. Darüber hinaus kommt es vor, daß der Inhalt, den IDs als unangemessen identifizieren, von den FIEs nicht als solcher beurteilt wird. Umgekehrt ist es wahrscheinlicher, daß FIEs falschen Inhalt identifizieren und einstimmig in die Problemidentifikation und/oder Revision aufnehmen. Auf der anderen Seite schenken die FIEs der allgemeinen Instruktionsqualität, der Präsentation, Organisation, Kohärenz des Materials und der Konsistenz zwischen den einzelnen Teilen nur wenig Aufmerksamkeit. Die wichtigste Implikation dieser

Befunde ist, daß sie die bedeutsame Rolle der FIEs und IDs in der Entwicklung von Instruktionsmaterial und die einzigartigen Beiträge jeder Gruppe zu diesem Prozeß betonen.

Bei einer detaillierteren Analyse, die über die Einteilung in generelle Kategorien wie Präsentation und Design hinausgeht, zeigt sich ein geringes Maß an Übereinstimmung zwischen den Experten einer jeden Bezugsgruppe. Allgemein scheint es eine höhere Variabilität zwischen den IDs als zwischen den FIEs zu geben. Auf einem allgemeineren Niveau jedoch neigen die FIEs dazu, sich auf die Genauigkeit und Angemessenheit des Inhaltes zu konzentrieren, während die IDs die Übereinstimmung zwischen den Teilen, die Präsentation, das Format und den pädagogischen Wert der Lehrziele und des Inhalts im allgemeinen beachten.

Dieser Befund hat zumindest eine praktische Implikation: Häufig wird die formative Evaluation unter zeitlichen und finanziellen Einschränkungen durchgeführt. Deshalb ist es wichtig festzulegen, welche und wieviele Experten diese Aufgabe durchführen sollten. Da auf allgemeinem Niveau die Experten aus einer Bezugsgruppe sich auf bestimmte Aspekte des Textes konzentrieren, ist es empfehlenswert, mindestens einen FIE und einen ID zu beauftragen, um eine sinnvolle Überarbeitung von Lehrmaterialien zu gewährleisten.

Schließlich verzichteten die FIEs trotz ihrer pädagogischen Erfahrung und ihrer Vertrautheit mit der Zielgruppe auf Kommentare zum pädagogischen Aspekt des Textes. Im Gegensatz dazu betonen die IDs gerade die pädagogische Seite und die Bedürfnisse der potentiellen Benutzer. Tatsächlich basiert ein Teil ihrer Evaluation auf dem Wissen, das sie als naive Lernende eines neuen Inhalts von dem Text gewinnen können. Die Feststellung, daß FIEs mit beträchtlicher Lehrerfahrung nicht unbedingt den pädagogischen Wert des Textes kommentieren, ist ein ernüchternder Befund, der weitreichende erzieherische und praktische Implikationen hat. In der Vergangenheit und in der gegenwärtigen Praxis wurden viele Lehrbücher nur von FIEs geschrieben und revidiert. Deshalb wird es ausgesprochen wichtig, auch andere Experten, die besonders auf die Bedürfnisse der Schüler achten, vor der Veröffentlichung von Lehrmaterialien einzubeziehen.

An der McGill University gibt es zur Zeit zwei Projekte von Studenten in höheren Semestern, die man teilweise als Replikationen der vorliegenden Studie ansehen kann. In ihnen werden die gleiche Methode, das gleiche Kodierschema und die gleiche Vorgehensweise verwendet, um die Strategien und die Ergebnisse der formativen Evaluation zu umreißen. Sie unterscheiden sich dadurch, daß als Probanden Instruktionsdesigner eingesetzt werden und die Zielpopulation Fachinhalteexperten sind. Außer diesen sind mir keine Studien bekannt, die dem hier vorgestellten Untersuchungsansatz ähneln. So gibt es denn in bezug auf die Fragestellung, ob Revisionen von IDs und FIEs einen Lehrtext tatsächlich effektiver machen, nur wenige empirische Belege. Im allgemeinen wird betont, daß jede Form der Revision einen Text effektiver macht (Abedor, 1971; Baker, 1970; Kandaswamy, Stoltovitch

& Thiagarajan, 1976). Macdonald-Ross (1978) sieht das Hauptproblem darin, daß die Reliabilität zwischen den Revisoren fragwürdig ist, und daß es keine detaillierte Auflistung gibt, wie die identifizierten Probleme in der Revision umgesetzt werden. In diesem Zusammenhang betont Kandaswamy (1980), daß radikal unterschiedliche Modifikationen von den verschiedenen Revisoren auf der Grundlage derselben Daten gemacht werden können. Deswegen ist es von größter Wichtigkeit, die Variabilität der Revisoren und die Validität der Tests (wenn man die Effektivität an der Leistung der Studenten im Test messen will) zu kontrollieren, um überzeugende Anmerkungen über die Effektivität eines Textes machen zu können.

Zusammenfassend kann man sagen, daß die vorliegende Studie die qualitativen Unterschiede zwischen FIEs und IDs beschreibt und die Bereiche und Niveaus nachzeichnet, auf denen sie in ihren Kommentaren konvergieren oder divergieren.

Anmerkung

¹ Übersetzung von Caroline Schröder.

Literatur

- ABEDOR, A. (1971): Development and validation of a model of explicating formative evaluation process for multi-media self-instructional learning systems. *Dissertation Abstracts International*, 32, 9-10. (University Microfilms No. 72-8621)
- ALLY, M. (1985): A team approach to computer courseware design. *Educational Technology*, 25 (6), 2-15.
- ANDERSON, J. (1982): Acquisition of cognitive skill. *Psychological Review*, 89, 369-406.
- ANDREWS, D.H. & GOODSON, L.A. (1980): A comparative analysis of models of instructional design. *Journal of Instructional Development*, 3 (4), 2-16.
- BAKER, E.L. (1970): Generalizability of rules for an empirical revision. *Audiovisual Communication Review*, 18, 3, 300-305.
- BEACH, R. (1976): Self evaluation strategies of extensive revisers and non reviser. *College Composition and Communication*, 27, 160-162.
- BOWLER, M. (1978): The making of a textbook. *Learning*, 6 (March), 38-42.
- BRANSON, R. (1973): *Analysis and assessment of the state of the art in instructional technology*. Fort Monroe, VA: Army Training and Doctrine Command. (NTIS Document Reproduction Service No. ADA010 394; ERIC Document Reproduction Service No. ED 088-436.
- BREUXLEUX, A. (1987, July): *Discourse and the investigation of cognitive skills in complex tasks*. Paper presented at the meeting of the Cognitive Science Society, Seattle, WA.
- BURKHOLDER, B. (1981-82): The effectiveness of using the instructional strategy diagnostic profile to prescribe improvements in self-instructional materials teaching abstract concepts. *Journal of Instructional Development*, 5 (2), 2-9.
- BURT, C. (1989): *Identification of critical variables in developmental testing and the experimental examination of the number and roles of participants in testing sessions*. Unpublished doctoral dissertation, McGill University, Montreal, Canada.
- CARROLL, M.J. (1988, April): *Differences in quantitative and qualitative data from four developmental testing conditions*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
- CHOI, S., MERRILL, M., CALLAHAN, E., HAWKINS, R. & NORTON, R. (1979): The relationship of test performance to ISDP rating in organic chemistry texts. *Journal of Instructional Development*, 3 (1), 16-25.

- CRONBACH, L. (1963): Course improvement through evaluation. *Teachers College Record*, 64, 672-683.
- D.A.E. PROJECT (1978): *Microbiology: A submodule to sterilization and disinfection*. Seattle, WA: Office of Research in Medical Education, School of Medicine, University of Washington.
- DE GROOT, A. (1966): Perception and memory versus thought: some old ideas and recent findings. In: B. Kleinmuntz (Ed.), *Problem solving*. New York: Wiley.
- DICK, W. (1968): A methodology for the formative evaluation of instructional materials. *Journal of Educational Measurement*, 5, 99-102.
- DICK, W. (1977): Formative evaluation. In: L.J. Briggs (Eds.) *Instructional design: Principles and applications* (pp. 311-333). Englewood Cliffs, NJ: Educational Technology.
- DICK, W. & CAREY, L. (1985): *The systematic design of instruction* (2nd ed.). Glenview, IL: Scott, Foresman.
- DUPONT, D. & STOLOVITCH, H. (1983): The effects of a systematic revision model on revisers in terms of student outcome. *National Society for Performance & Instruction Journal*, 22 (2), 33-37.
- EGAN, D. & SCHWARTZ, B. (1979): Chunking in recall of symbolic drawings. *Memory and Cognition*, 7 (2), 149-158.
- ERICSSON, K. & SIMON, H. (1984): *Protocol analysis*. Cambridge, MA: MIT.
- FAIGLEY, L. & WITTE, S. (1981): Analyzing revision. *College Composition and Communication*, 32, 400-414.
- FELKER, D., REDISH, J. & PETERSON, J. (1985): Training authors of informative documents. In: T. Duffy & R. Waller (Eds.). *Designing usable text* (pp. 43-61). Orlando, FL: Academic.
- FLOWER, L. & HAYES, J. (1977): Problem solving strategies and the writing process. *College English*, 39, 449-461.
- FLOWER, L., HAYES, J. & SWARTS, H. (1980): *Revising functional documents: The Scenario Principle*. (Tech. Rep. No. 10). Pittsburgh: Carnegie-Mellon University, Document Design Project.
- FOSHAY, W. (1984): QA and QC: A training vendor's vision of the formative/summative evaluation distinction. *Performance and Instruction Journal*, 23 (10), 15-17.
- GAGNE, R. & BRIGGS, L. (1979): *Principles of instructional design* (2nd ed.). New York: Holt Rinehart 5 Winston.
- GEIS, G. (1987): Formative evaluation: Developmental testing and expert review. *Performance and Instruction*, 26 (4), 1-8.
- GEIS, G. (1988, April). *Profiles of the actors in formative evaluation*. Paper presented at the annual conference of the American Educational Research association, New Orleans, LA.
- GEIS, G., BURT, C. & WESTEN, C. (1984): *Instructional development: Developmental testing*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
- GOLAS, K. (1982): The effectiveness and cost of alternate models of formative evaluation for printed instructional materials (Doctoral dissertation, Florida State University). *Dissertation Abstracts International*, 43, 2873A.
- GROPPER, G. (1975): *Diagnosis and revision in the development of instructional materials*. Englewood Cliffs, NJ: Educational Technology.
- HARTLEY, J. (1978): *Designing instructional text*. London: Kogan Page.
- HAYES, J. & FLOWER, L. (1983): Uncovering cognitive processes in writing: An introduction to protocol analysis. In P. Mosenhal, L. Tamor & S. Walmsley (Eds.). *Research on writing: Principles and methods* (pp. 207-220). New York: Longman.
- HAYES, J., FLOWER, L., SCHRIVER, K., STRATMAN, J. & CAREY, L. (1987): Cognitive processes in revision. In: S. Rosenberg (Ed.). *Advances in applied psycholinguistics: Vol. 2* (pp. 176-240). Cambridge, England: Press Syndicate of the University of Cambridge.
- KANDASWAMY, S. (1980): Evaluation of instructional materials. A synthesis of models and methods. *Educational Technology*, 20 (6), 19-26.
- KANDASWAMY, S., STOLOVITCH, H. & THIAGARAJAN, S. (1976): Learner verification and revision: An experimental comparison of the two methods. *Audio Visual Communications Review*, 24 (3), 316-338.

- KOMOSKI, K. (1983): Formative evaluation. *Performance and Instruction Journal*, 22 (5), 3-4.
- KOMOSKI, K. & WOODWARD, A. (1985): The continuing need for Learner Verification and Revision of textual material. In: D.H. Jonassen (Ed.). *The technology of text: Vol: 2* (pp. 396-417).
- LARKIN, J., McDERMOTT, J., SIMON, D. & SIMON, H. (1980): Models of competence in solving physics problems. *Cognitive Science*, 4, 317-345.
- MACDONALD-ROSS, M. (1978): Language in texts. *Review of Research in Education*, 6, 229-275.
- MARKLE, S., (1976): Evaluating instructional programs. In: *Programmed Instruction: Part II* (104-138). The 66th annual yearbook of the National Society for the study of Educational Research. Chicago, L.
- MARTELLI, M. (1976): A study of a theory-based model of formative evaluation (Doctoral dissertation, Florida State University). *Dissertation Abstracts International*, 40, 6, 3089-A.
- MERRILL, M., REIGELUTH, C. & FAUST, G. (1979): The instructional quality profile: A curriculum evaluation and design tool. In H.F. O'Neil, Jr. (Ed.). *Procedures for functional systems development 1* (pp. 165-204). New York: Academic.
- MERRILL, M., RICHARDS, R., SCHMIDT, R. & WOOD, N. (1977): *The instructional strategy diagnostic profile training manual*. San Diego, CA: Coruseware, Inc.
- NATHENSON, M. & HENDERSON, E. (1977): Problem and issues in developmental testing. *National Society for Performance and Instruction Journal*, XVI (1), 9-10.
- NATHENSON, M. & HENDERSON, E. (1980): *Using student feedback to improve learning materials*. London: Croom Helm.
- NEVO, D. (1977): A model for the utilization of formative evaluation in the process of developing instructional materials. *Programmed Learning and Educational Technology*, 14, 2, 127-133.
- O'DONNELL, H. (1985): Improving textbooks — Who is responsible? *Journal of Reading*, 29, 3, 268-270.
- SAROYAN, A. (1989): *The review process of formative evaluation: The role of the content expert and the instructional designer*. Unpublished doctoral dissertation, McGill University, Montreal, Canada.
- SAROYAN, A. & GEIS, G. (1988): An analysis of guidelines for expert reviewers. *Instructional Science*, 17, 101-128.
- SCRIVEN, M. (1967): The methodology of evaluation. *AERA Monograph Series on Curriculum Evaluation*, 1, (39-83).
- SIMON, D. & SIMON, H. (1979): A tale of two protocols. In: J. Lochhead & J. Clement (Eds.) *Cognitive process instruction* (pp. 119-132). Philadelphia, PA: Franklin Institute.
- SMILLIE, R. (1985): Design strategies for job performance aids. In: T. Duffy & R. Waller (Eds.). *Designing usable texts* (pp. 213-243). Orlando, FL: Academic.
- SOMMERS, N. (1980): Revision strategies of studies student writers and experienced adult writers. *College Composition and communication*, 31, 378-388.
- STOLOVITCH, H. (1982): Application of the intermediate technology of Learner Verification and Revision (LVR) for adapting international instructional resources to meet local needs. *National Society for Performance and Instruction Journal*, XXI (7), 16-22.
- SWARTS, H., FLOWER, L. & HAYES, J. (1984): Designing protocol studies of the writing process. In: R. Beach & L. Bridwell (Eds.). *New directions in composition research* (pp. 53-71). New York: Gilford.
- THORNDYKE, P. & STASZ, C. (1980): Individual differences in procedures for knowledge acquisition from maps. *Cognitive Psychology*, 12, 137-175.
- WINOGRAD, T. (1983): *Language as a cognitive process: Vol 1*. New York: Addison Wesley.
- WRIGHT, P. (1985): Editing: Policies and processes. In: T.M. Duffy & r. Waller (Eds.). *Designing usable text* (pp. 63-96). Orlando, FL: Academic.

Anschrift der Autorin:

Alenoush Saroyan, McGill University, 3700 McTavish, Montreal, Quebec H3A 1Y2, Canada.