

Klauer, Karl Josef; Feger, Barbara

Die Konstruktion lehrzielvalider Aufgabenstichproben

Unterrichtswissenschaft 13 (1985) 3, S. 266-276



Quellenangabe/ Reference:

Klauer, Karl Josef; Feger, Barbara: Die Konstruktion lehrzielvalider Aufgabenstichproben - In: Unterrichtswissenschaft 13 (1985) 3, S. 266-276 - URN: urn:nbn:de:0111-pedocs-295760 - DOI: 10.25656/01:29576

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-295760>

<https://doi.org/10.25656/01:29576>

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

Allgemeiner Teil

Karl Josef Klauer, Barbara Feger

Die Konstruktion lehrzielvalider Aufgabenstichproben

Der Weg von der Definition eines kognitiven Lehrziels hin zur Erzeugung von Aufgabenstichproben, die das Lehrziel repräsentieren, wird im Detail beschrieben. Verfahrensmäßig ist zwischen zwei Arten von Lehrzielen zu unterscheiden. Ist der zugehörige Sachverhalt in Form allgemeiner Regeln oder aber in Form eines Lehrtextes dargestellt bzw. darstellbar, so sind jeweils spezifische Konkretisierungen des Verfahrens notwendig. Das wird im einzelnen dargestellt und an Beispielen erläutert.

The construction of objective valid task samples

This paper presents a detailed description of the successive stages leading from the definition of an objective in the cognitive domain to the generation of item samples representing the objective. With respect to the procedures two kinds of objectives ought to be distinguished. Objectives taking the form of general rules differ from those taking the form of an instructional text. This will be illustrated by examples.

Je nachdem, wie man es expliziert, ist *Hilbert L. Meyers* (1971) ungelöstes Deduktionsproblem gelöst – oder schlechthin unlösbar. Etwas Unlösbares nimmt man sich vor, wenn man aus einer nur ahnungsvoll-verschwommenen Zielvorstellung methodisch kontrolliert und zwingend herleiten will, was zu dieser Zielvorstellung gehört. Wer das versucht, der versucht, sich am eigenen Zopf aus dem Sumpf zu ziehen. Prinzipiell gelöst ist aber das Problem, aus einem präzise definierten Lehrziel durch methodisch kontrollierte Analyse herauszuarbeiten, was zu diesem Lehrziel gehört. Ein Versuch, aus unklaren Vorstellungen klare und abgegrenzte Begriffe zu gewinnen, zeigt hinlänglich, daß dies nicht ohne eine Folge mehr oder minder willkürlicher Entscheidungen möglich ist, die man der Natur der Sache nach auch anders hätte fällen können. Wir haben es aber mit einer völlig anderen Situation zu tun, wenn man von definierten Begriffen ausgeht. Steht das präzise definierte Lehrziel am Anfang, so ist es prinzipiell möglich, daraus konkrete Anweisungen für den Lehrgang wie für die Erfolgskontrolle abzuleiten. Der Weg vom Lehrziel zum Lehrgang wird an anderer Stelle beschrieben (*Klauer* 1985). Im vorliegenden Beitrag geht es darum darzustellen, wie aus dem Lehrziel Aufgabenmenge konstruiert werden können, die das Lehrziel repräsentieren. Solche Aufga-

benmengen dienen, wie *Schott, Neeb & Wieberg* (1981) ausführen und wie aus Abbildung 1 deutlich wird, entweder Aneignungs- oder Prüfzwecken. Wir beschränken uns im wesentlichen auf kognitive Lehrziele und verweisen bezüglich affektiver Ziele auf *Nußbaum, Feger & Leutner* (1984). Eine differenziertere und mit mehr Beispielen erläuterte Darstellung des Gesamtkonzepts findet man bei *Klauer* (1984a), *Feger* (1984) und *Klauer* (1984b).

Jede Aussage, ob sie nun ausformuliert oder nur erzeugbar ist, bringt einen Teil des gemeinten Sachverhalts zum Ausdruck. (Siehe Abb. 1, Seite 274). Nun soll der Lernende aber nach erfolgreichem Lehrgang den Sachverhalt beherrschen, ihn anwenden können oder ähnliches. Das heißt, der Lernende soll, wenn ihm der Sachverhalt in bestimmter Weise präsentiert wird, darauf in einer näher charakterisierten Weise reagieren können. Eine *Aufgabenform* legt aber fest, wie etwas vorgelegt wird und welche Art von Antwort darauf erwartet wird. Eine Aufgabenform enthält eine Stimuluskomponente, die angibt, wie etwas vorgelegt wird, und eine Response-Komponente, die angibt, was als Antwort erwartet wird. Für Testzwecke haben sich eine Reihe von typischen Aufgabenformen herausgestellt (vgl. *Herbig* 1972; *Lienert* 1969; *Rütter* 1978, 1982), die auch pädagogisch interessant sind. Dennoch ist es unter Umständen sinnvoll, auch andere als diese Standardformen von Aufgaben zu verwenden. Hat man nun den Sachverhalt durch Aussagen oder eine Aussageform präzise festgelegt und sich für eine oder mehrere Aufgabenformen entschieden, so sind noch bestimmte Festlegungen zu treffen: Transformationsregeln geben an, wie der Sachverhalt eindeutig in Aufgaben zu transformieren ist, und Restriktionen (vgl. *Leutner* 1985) beschränken die Menge der zu erzeugenden Aufgaben. Es gibt viele Gründe für sinnvolle Restriktionen. Geht es im Fremdsprachenunterricht beispielsweise darum, nach bestimmten Satzgliedern fragen zu können, so wird man sich nur auf den Wortschatz beschränken, der den Schülern bekannt ist. Geht es in der Grundschule darum, schriftlich multiplizieren zu lernen, so wird man nicht daran denken, den Schülern Zahlen von je siebzehn Stellen zur Multiplikation vorzulegen; vielmehr wird man unterstellen, die Schüler hätten das Prinzip erkannt und hinreichend gelernt, wenn sie maximal drei- bis fünfstellige Zahlen multiplizieren können. Die Festlegung des Sachverhalts, der Aufgabenform und der Transformationsregeln mit den Restriktionen definiert eine *Aufgabenmenge*. Mit der Aufgabenmenge ist zugleich aber eine *Kompetenz* inhaltlich definiert, nämlich die Kompetenz, die sich in der Bewältigung der Aufgabenmenge erweist. Indirekt ist mit der Kompetenz zugleich auch das *Lehrziel* definiert, bei dem es um die Absicht geht, diese Kompetenz zu vermitteln. Mit diesen Festlegungen ist also, etwas verkürzt gesagt, das Lehrziel präzise definiert und insofern die erste Stufe der Operationalisierung abgeschlossen.

Nun ist es nicht immer möglich oder sinnvoll, den Schülern die ganze Aufgabenmenge zuzumuten. Würde man jedoch beliebige Aufgaben aus der Grundmenge vorlegen, so wäre nicht gewährleistet, daß die ausgewählten Aufgaben auch ein Abbild der Grundmenge darstellen. Wir definieren daher: Lehrzielvalide ist eine Aufgabenmenge dann, wenn sie entweder identisch mit der Grundmenge des Lehrziels ist oder wenn sie eine repräsentative Stichprobe daraus darstellt. Im folgenden

konzentrieren wir uns auf den letzteren Fall, weil er einige Probleme aufwirft. Zu beachten ist jedoch, daß nach der gegebenen Definition nicht einzelne Aufgaben lehrzielvalide sein können, sondern nur Aufgabenmengen. Lehrzielvalidität ist also nicht eine Eigenschaft von Aufgaben, sondern von Aufgabenmengen. Repräsentative Stichproben kann man durch reine Zufallsprozeduren gewinnen, etwa mit Hilfe des Loses. In dem Falle bestehen die Samplingvorschriften aus Angaben über die zu verwendende Zufallsauswahl. Dieses Vorgehen ist nur dann sinnvoll, wenn man relativ große Stichproben bildet, denn kleine Stichproben können sich in ihren Kennwerten sehr markant von der Grundgesamtheit unterscheiden. In vielen Fällen wird es daher zweckmäßig sein, geschichtet-zufällige (stratifiziert-zufällige) Stichproben zu bilden. Dabei wird die Grundmenge in sinnvolle Teilmengen zerlegt, und es wird festgelegt, mit welchen Anteilen die Teilmengen in der Stichprobe vertreten sein sollen. Innerhalb der Teilmengen wird unter Beachtung der innerhalb der Menge verbliebenen Merkmalsvariation zufällig ausgewählt. Alle diese Schritte sind in den Samplingvorschriften festzulegen. Wer immer sich an sie hält, erzeugt lehrzielvalide Aufgabenstichproben. Mit diesem Schritt ist die zweite Stufe des Operationalisierungsprozesses abgeschlossen.

Lehrzielvalide Aufgabenstichproben können zur Konstruktion lehrzielvalider Lehrgänge wie zur Konstruktion lehrzielvalider Tests herangezogen werden, wobei jeweils zusätzliche Anweisungen zu beachten sind. Solche Lehrgänge und Tests stellen die dritte und letzte Operationalisierungsstufe dar. Die Erzeugung lehrzielvalider Aufgabenstichproben ist nach allem ein zentraler Schritt des Verfahrens. Angenommen, mehrere Leute arbeiten unabhängig voneinander, um lehrzielvalide Aufgabenstichproben zu erzeugen, oder auch angenommen, eine Person bemühe sich darum zu verschiedenen Zeiten. Wenn die beschriebenen Festlegungen beachtet werden, so werden praktisch nie genau dieselben Stichproben erzeugt werden, sondern jedes Mal wird eine andere Stichprobe entstehen. Trotzdem wird jede Stichprobe für das Lehrziel valide sein. Daraus folgt zum Beispiel, daß solche Stichproben austauschbare Ergebnisse liefern müssen, wenn man sie Schülern zur Bearbeitung vorlegt: Sie müssen gleich schwer sein, gleiche Mittelwerte, Streuungen und Korrelationen haben. Kurz gesagt, lehrzielvalide Paralleltests sind unterschiedliche, aber einander äquivalente Tests. Diese Behauptung wurde inzwischen mehrfach empirisch belegt (Nußbaum 1980, Klauer & Dänecke 1981, Klauer 1984b, Leutner 1985).

Das möge als Überblick über das Verfahren genügen. Im Detail wären noch viele Einzelheiten zu erörtern. Das ist im vorliegenden Rahmen nicht möglich, weswegen auf die erwähnte Literatur verwiesen wird. Besonders wichtig sind die dort dargestellten Beispiele, die vieles klarer machen als dies allgemeine Beschreibungen vermögen. Deswegen sollen im folgenden auch einige Beispiele gegeben werden. Dazu ist es sinnvoll, eine bereits erwähnte Unterscheidung aufzugreifen.

Zwei Arten von Sachverhalten

Wie oben bereits dargelegt, ist es für Zwecke der Lehrzieldefinition sinnvoll, zwei Arten von Sachverhalten zu unterscheiden. Die eine Art läßt sich nur durch einen Text, das ist eine geordnete Menge von Aussagen, darstellen. Typisch für diese Klasse von Sachverhalten mag ein historisches Ereignis sein. Andere Sachverhalte können auch (oder nur) durch eine Aussagenform mit entsprechender Einsetzmenge dargestellt werden. Typisch für diese Klasse mag ein mathematischer Sachverhalt sein. In dem Fall wird die Grundmenge von Aufgaben durch Angabe der mengenstiftenden Merkmale gebildet. Andernfalls kann es notwendig werden, die Grundmenge durch vollständige Aufzählung der Aufgaben festzulegen.

Der Sachverhalt ist in einer Aussagenform dargestellt

Die folgenden Beispiele beginnen jeweils mit der Darstellung des Sachverhalts und enden mit den Samplingvorschriften.

Beispiel 1

Lehrziel:

Beherrschung der Multiplikation von Brüchen

Sachverhalt: $a \cdot b = c$

a , b und c sind Elemente von N_{100} , B_{100} oder G_{100} . N_{100} ist die Menge der natürlichen Zahlen von 1 bis 100. B_{100} ist die Menge der Brüche $\frac{n}{m}$, wobei n und m Elemente von N_{100} und teilerfremd sind. G_{100} ist die Menge der gemischten Zahlen $k \frac{n}{m}$, wobei für k gilt: $1 \leq k \leq 20$.

Aufgabenform:

Ergänzungsaufgabe des Typs $a \cdot b = \underline{\hspace{2cm}}$.

Transformationsregeln: 1) Ersetze a , b und c , so daß eine wahre Aussage entsteht.
2) Streiche c .

Restriktion:

Nur a oder b darf Element von N_{100} sein, beide dürfen es aber nicht gleichzeitig sein.

Samplingvorschriften:

Die fünf Teilmengen $B \cdot B$, $B \cdot G$ bzw. $G \cdot B$, $N \cdot B$ bzw. $B \cdot N$, $G \cdot N$ bzw. $N \cdot G$, $G \cdot G$ sind zu gleichen Anteilen zu berücksichtigen. In den Mischformen wie $B \cdot G$ und $G \cdot B$ sind beide Varianten gleich häufig zu verwenden.

Das Ergebnis der Multiplikation ist gleich häufig kürzbar und nicht kürzbar.

Das Ergebnis c ist (nach vollständigem Kürzen) gleich häufig ein Bruch, eine natürliche Zahl bzw. eine gemischte Zahl.

Die Festlegungen müßten im einzelnen näher erörtert werden. Manche davon sind notwendig, um im Lehrziel zu bleiben, andere könnten auch anders getroffen werden. Ersteres gilt für die als solche bezeichnete Restriktion, letzteres gilt für die Einschränkungen bei den Samplingvorschriften. Sie wurden hier so festgelegt, daß alle wichtigen Varianten gleich häufig auftreten.

Wer nach diesen Vorschriften arbeitet, erzeugt lehrzielvalide Aufgabenstichproben. Ferner kann mit den Vorschriften jede Aufgabenstichprobe daraufhin überprüft werden, ob sie lehrzielvalide ist.

Beispiel 2

Lehrziel:

Verstehen englischer Sätze, die mit der Form „There + be“ beginnen.

Sachverhalt:

	A	B	C
There	$\left\{ \begin{array}{l} a_{11} \text{ is} \\ a_{12} \text{ are} \\ a_{21} \text{ was} \\ a_{22} \text{ were} \\ a_3 \text{ will be} \end{array} \right\}$	$\left\{ \begin{array}{l} b_1 - \\ b_2 \text{ no} \end{array} \right\}$	$\left\{ \begin{array}{l} c_1 \text{ Substantiv sing.} \\ c_2 \text{ Substantiv plur.} \end{array} \right\}$
		D	
		$\left\{ \begin{array}{l} d_1 - \\ d_2 \text{ Ortsadverb} \end{array} \right\}$	

Aufgabenform:

Übersetzungsaufgabe. Vorgelegt wird ein englischer Satz, der mündlich/schriftlich ins Deutsche zu übersetzen ist.

Transformationsregeln:

Erzeuge englische Aussagensätze der Form

„There $a_j b_i c_i d_i$ “.

Restriktion:

a_1 kommt nur vor in Verbindung mit c_1 ,

a_2 kommt nur vor in Verbindung mit c_2 .

Samplingvorschriften:

Die $5 \times 2 \times 2 \times 2 - 2 = 38$ Teilmengen sind gleich häufig zu berücksichtigen. Alternative: Die vier Teilmengen entsprechend B und D sind gleich häufig zu berücksichtigen, wobei die Varianten von A zufällig variieren.

Die Mengen C und D enthalten nur Vokabeln, die im Schülerbuch bislang vorgekommen sind.

Die Vokabeln zu C und D streuen über verschiedene Bereiche, wobei sinnlose Sätze zu vermeiden sind.

Formal ergibt der Sachverhalt 40 Varianten bzw. Teilmengen erzeugbarer Sätze, von denen allerdings zwei herauszunehmen sind, weil sie der Rektion des Verbs widersprechen. Um auch nur jede Variante abzudecken, müßten mindestens 38 Sätze gebildet werden. Diese Zahl würde sich erhöhen, wenn alle Formen von *to be* berücksichtigt würden. Deshalb ist es hier zweckmäßig, sich auf die vier Teilmengen

$b_1 d_1$ („es gibt“),

$b_2 d_1$ („es gibt nicht/kein“),

$b_1 d_2$ („da ist/sind/liegen“) sowie

$b_2 d_2$ („da ist/sind/liegen nicht“)

zu beschränken und die Form von *to be* zufällig in vorgegebenen Anteilen zu variieren.

Beide Beispiele sind bewußt einfach gehalten, um das Prinzip deutlich werden zu lassen. Sie zeigen, daß es möglich ist, anhand solcher Festlegungen Aufgabenstichproben zu erzeugen, die für das Lehrziel gültig sind. Sicherlich werden komplexere Lehrziele komplexere Festlegungen erfordern, doch im Grunde kann man jedes Lehrziel, dessen Sachverhalt als Aussagenform darstellbar ist, relativ einfach in der gezeigten Weise operationalisieren. Ist der Lehrstoff jedoch nur durch Aussagen, bzw. durch einen Text darstellbar, so ist das bislang vorgestellte Verfahren nicht ohne wichtige Änderungen möglich.

Der Sachverhalt ist durch einen Text dargestellt

Den Ausgangspunkt in dem allgemeinen Schema in Abbildung 1 bildet der Sachverhalt, auf den sich das Lehrziel bezieht. Bei Lehrtexten, die in allen Bereichen der Unterweisung eine hervorragende Rolle spielen, steht man jedoch vor der Schwierigkeit, daß der Text so, wie er vorliegt, nur in seltenen Fällen mit dem Sachverhalt gleichgesetzt werden kann.

Wie *Klauer* (1979, S. 36/37) nämlich zutreffend feststellt, bieten Schulbücher, Lehrbücher und fachwissenschaftliche Abhandlungen nicht nur eine Darstellung des Sachverhaltes an; sie enthalten auch Wiederholungen, Zusammenfassungen, Ausblicke, Beispiele, eingeschobene Erklärungen, Textteile, die überwiegend motivieren und Interesse wecken sollen usw. Verschiedene Autoren, die denselben Sachverhalt darstellen wollen, gelangen so zu verschiedenartigen Texten; man kann also feststellen, daß darstellender Text und dargestellter Sachverhalt nicht identisch sind. Vorschläge, den Sachverhalt textfrei zu erfassen und darzustellen, legen als eine Möglichkeit die Verwendung einer Art Kunstsprache in Anlehnung an die moderne Logik nahe (Siehe Abb. 2, Seite 275).

Es sind aber auch eine ganze Reihe von Vorschlägen vorgelegt worden, wie ein Text auf seinen Kerngehalt reduziert werden kann, und zwar in Übereinstimmung mit den beim Lesen ablaufenden Verstehensprozessen. Verschiedene Varianten sind an anderer Stelle dargestellt worden (*Feger* 1984; vgl. auch *Hoppe* 1984). Beispielsweise über eine systematische Zusammenfassung erhält der Lehrer eine Folge jener Aussagen, die die meisten seiner Schüler nach dem Lesen behalten sollen. Bei einer systematisch durchgeführten Zusammenfassung sind auch die Lehrzielkategorien des Ausgangstextes – beispielsweise Wissen, Verständnis und Analyse – in den entsprechenden Proportionen noch vorhanden. Mit dem reduzierten Text haben wir eine Folge von Aussagen, die den Sachverhalt repräsentieren. Mit Hilfe einer Aufgabenform können die Aussagen in Items verwandelt werden; dies wird bei Lehrtexten – stärker noch als bei Aussageformen – durch Transformationsregeln bestimmt. Ehe dieser Komplex behandelt wird, sei jedoch auf die Zahl der Aussagen eingegangen, die ein reduzierter Text umfaßt; dies leitet unmittelbar über zur Stichprobenfrage bei einer Folge von Aussagen. Es sind eine ganze Reihe von Fällen denkbar, in denen der reduzierte Text nur eine relativ kleine Anzahl von Aussagen umfaßt; das könnte beispielsweise dann der Fall sein, wenn er auf dem Kapitel eines Buches basiert. Man erhält hier – beispielsweise – aus einer Menge von 30 Aussagen auch 30 Items. Diese Items würden in ihrer Gesamtheit den Text bilden. Ganz anders stellt sich jedoch die Situation dar, wenn ein umfassendes Lehrbuch die Grundlage für eine schriftliche Prüfung bilden soll. In diesem Fall ist es ohne weiteres möglich, daß sich die Zahl der zusammenfassenden Sätze, die den Kerngehalt des Buches enthalten sollen, auf 200 oder 300 beläuft. Hier stellt sich dann die Frage nach der Anzahl der für die Testendform auszuwählenden Items oder Sätze. (Zwischen Items und Sätzen soll an dieser Stelle noch nicht unterschieden werden.) Eine Formel, die auf der Stichprobentheorie basiert und die sich ausdrücklich auf textbezogene Tests bezieht, findet sich bei *Berk* (1980). An einem

Beispiel soll demonstriert werden, wieviele Aufgaben zufällig oder stratifiziert-zufällig gezogen werden müssen, damit der zugrunde gelegte Kompetenzgrad mit einer gewünschten Wahrscheinlichkeit bei einer vom Testautor festzusetzenden Fehlerschranke ermittelt wird.

Zunächst die Formel für die erforderliche Zahl N^* von Items:

$$N^* = \frac{\frac{z^2 \pi_z (1 - \pi_z)}{e^2}}{1 + \frac{1}{N} \left(\frac{z^2 \cdot \pi_z \cdot (1 - \pi_z)}{e^2} - 1 \right)}$$

Darin ist

z = z-Wert, der dem gewünschten Konfidenzniveau entspricht (z. B. $z = 1,64$ für das 90 %-Konfidenzniveau).

π_z = Zielmarke, theoretisch geforderter Kompetenzgrad (z. B. $\pi_z = 0,80$).

e = Fehlerschranke um π_z , die man in Kauf nehmen will (z. B. $e = 0,1$).

N = Größe der Grundgesamtheit von Aufgaben.

Beispiel 3

Beispiel für die Berechnung der erforderlichen Testlänge N^ eines lehrtextbezogenen Tests*

Der Text ist auf $N = 200$ Sätze (= 200 mögliche Items) reduziert worden; der Test soll mit $\pi_z = 0,90$ beherrscht werden; der zulässige Fehler sei $e = 0,10$. Das Konfidenzniveau liege bei 95 (zweiseitig; $z = 1,96$).

In die Formel werden dann die folgenden Werte eingesetzt:

$$N = 200 \quad \pi_z = 0,90 \quad e = 0,10 \quad z = 1,96$$

$$N^* = \frac{\frac{1,96^2 \cdot 0,90 \cdot 0,10}{0,10^2}}{1 + \frac{1}{200} \left(\frac{1,96^2 \cdot 0,90 \cdot 0,10}{0,10^2} - 1 \right)} = 29,6001$$

Entnimmt man der Grundgesamtheit 30 zufällig gezogene Items, dann wird in 95 von 100 Fällen das tatsächlich verwendete Lehrziel in dem Bereich von $\pi_z \pm e$, also zwischen 0,80 und 1,00 liegen.

Berk empfiehlt weiterhin an der Stelle von Zufallsstichproben die Technik des „systematischen Stichprobenziehens“. In einem ersten Schritt wird hier das Intervall k bestimmt, so daß jeder k -te Satz aus einer Grundgesamtheit entnommen wird. K erhält man, indem man die Gesamtstichprobe N durch N^* dividiert. In unserem Fall heißt das $\frac{200}{30} = 6,666 \dots$; aufgerundet wäre das also 7. Unter den ersten sieben Sätzen wird nun ein beliebiger Satz x ausgesucht, die als nächsten zu entnehmenden Sätze sind dann $x + 7$, $x + 14$, $x + 21$ (wobei im vorliegenden Fall zu bedenken ist, daß das Intervall kleiner als 7 ist, und deshalb auch Intervalle mit der Größe 6 berücksichtigt werden müssen).

Die Ausführungen zur Stichprobengewinnung wurden – entgegen dem Schema in Abbildung 1 – den Überlegungen zur Aufgabenform und den Transformationsregeln vorangestellt; außerdem wurde mehrfach keine Unterscheidung zwischen Sätzen und Items getroffen. Dies alles geschah deshalb, weil es bei einer großen Zahl von Sätzen in einem reduzierten Text sicherlich sinnvoll ist, aus dieser umfassenden

Grundgesamtheit *zunächst* eine *Stichprobe* zu entnehmen und diese erst *dann* in *Aufgaben* umzuwandeln, weil die Umwandlung in Items bei Aussagen in der Regel erheblich aufwendiger ist als bei Aussageformen. Andererseits gibt es einige sehr klare Transformationsregeln, die vorschreiben, wie ein Satz in ein Item umzuwandeln ist. Dabei gehen die Transformationsregeln und die Aufgabenformen hier eine besonders enge Verbindung ein. So bestimmt gelegentlich schon die Höhe des Aufwandes bei bestimmten Verfahren, wofür man sich entscheidet. Durchaus denkbar ist auch, daß bestimmte Verfahren als dem jeweiligen Text angemessener empfunden werden. Mit einer solchen Entscheidung ist aber auch gleich die Aufgabenform festgelegt. So mag man es bei einem sehr abstrakten soziologischen Text als angenehmer empfinden, Fragen nach den Vorschlägen von *Bormuth* (1970; vgl. auch *Rupprecht* 1972) oder *Finn* (1975) zu erstellen. Bei einem eher konkret greifbaren Sachverhalt und einem Test, der überwiegend Wissens- und Verstehenskategorien überprüfen soll, bietet sich oft das Verfahren der Gegensatz- oder Kontextstrategie an (*Klauer* 1979). Hält man es für wichtig, auf jeden Fall möglichst viele Distraktoren (bei einem durchweg höheren zeitlichen Aufwand) zu erhalten, dann entscheidet man sich möglicherweise für das Verfahren der semantischen Merkmale (*Feger* 1984). Diese Verfahren, die zugleich auch Transformationsregeln darstellen, sind an anderer Stelle ausführlich beschrieben worden (*Feger* 1984).

Auch bei Aussagen gibt es Restriktionen zu beachten. Dabei ist vor allem die wichtig, die besagt, daß aus einer Aufgabe nicht die Lösung zu einer anderen Aufgabe deutlich werden darf. Die Gefahr ist deshalb gegeben, weil ein wesentliches Merkmal des Textes in seiner Kohärenz besteht. Ein nachfolgender Satz nimmt etwas auf, was in einem vorausgegangenen bereits enthalten war. Ein kurzes Beispiel mag dies illustrieren. Angenommen, man habe einen reduzierten Text erhalten und wolle nun zu jeweils kursiv gesetzten Wörtern oder Wortgruppen Distraktoren finden, damit die Sätze in Items umgewandelt werden können. Dieser Text enthält nun zum einen den Satz: „Unter den Kulturpflanzen nimmt der *Weizen* den größten Raum ein.“ Und ein späterer Satz lautet: „Der *Weizen* wird vorzugsweise in *gemäßigten und subtropischen Gebieten* angebaut.“ Hier wird deutlich, daß jemand, der den Text zwar nicht kennt, die Items aber aufmerksam betrachtet, mit großer Wahrscheinlichkeit die richtige Lösung zu dem ersten Item finden wird. Als letzter Punkt sei noch die Testzusammenstellung erwähnt. Hier ist es angebracht, wegen der eben erwähnten Kohärenz die Items in eine Zufallsreihenfolge zu bringen. Das ist dann besonders wichtig, wenn alle Sätze eines reduzierten Textes in Items umgewandelt worden sind. Auch wenn kein Item ausdrücklich die Lösung zu einem anderen Item enthält, hat sich doch gezeigt, daß bei einer dem Text entsprechenden Abfolge der Items die Wahrscheinlichkeit des richtigen Ratens deutlich erhöht wird.

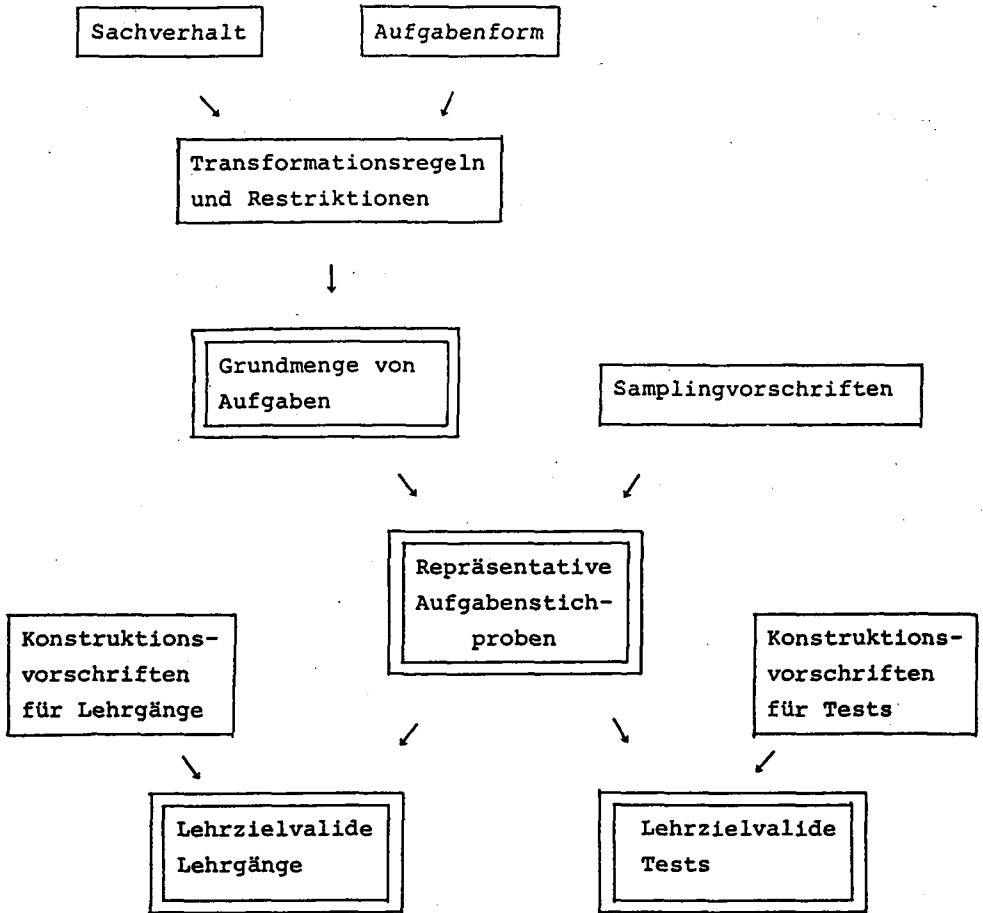


Abbildung 1: Lehrzieldefinition und Operationalisierungen des Lehrziels

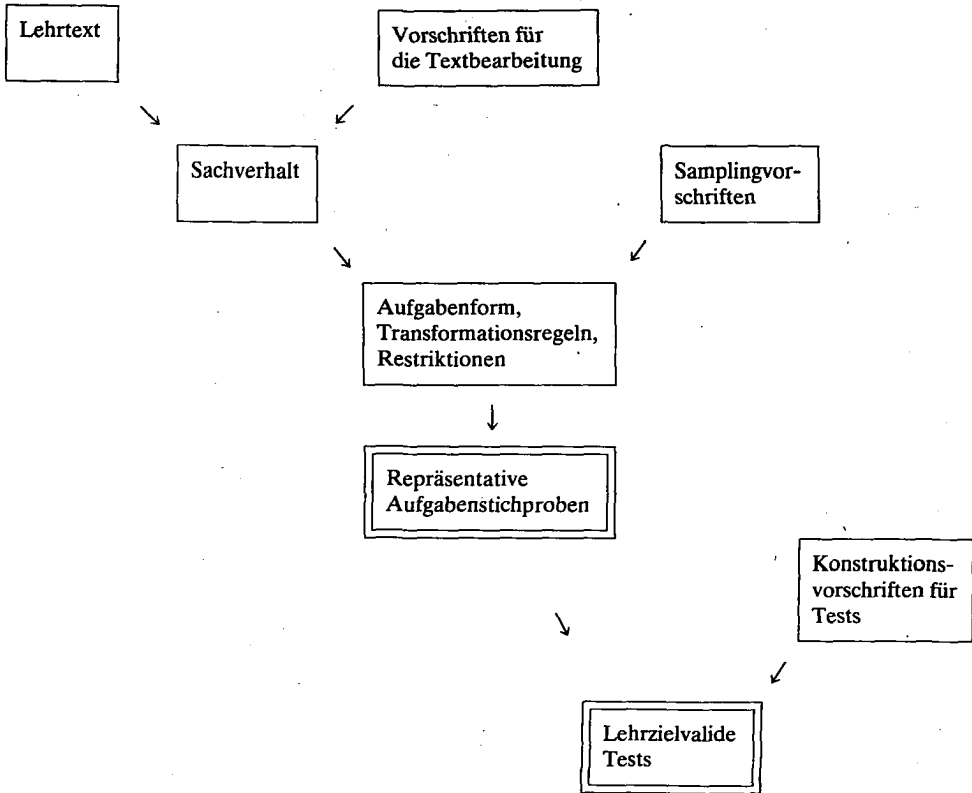


Abb. 2: Lehrzieldefinition und Operationalisierungen des Lehrziels bei lehrtextbezogenen Tests

Abbildung 2 zeigt das für Lehrtexte abgewandelte Schema.

Abschließend soll ein Beispiel die Vorgehensweise bei Aussagen illustrieren. Es geht dabei um einen längeren Text über Brasilien. Eine Gruppe von Lehrerstudenten reduzierte den Text auf eine Zusammenfassung von 30 Sätzen.

Beispiel 4

Lehrziel:

Kenntnis der wirtschaftlichen, geographischen und klimatischen Gegebenheiten Brasiliens und ein Verständnis für die Zusammenhänge zwischen diesen Faktoren

Lehrtext:

Ein Lehrtext über Brasilien von gut acht Druckseiten

Vorschriften für die Textbearbeitung:

Erstellung einer systematischen Zusammenfassung so, daß die Sätze als Grundlage für Items geeignet sind

Sachverhalt:

Dargestellt in einer Folge von 30 Aussagen

Samplingvorschriften:

Keine, da alle Aussagen in Items umgewandelt werden sollen

Transformationsregeln:

Vorgehen nach dem Verfahren der semantischen Merkmale

Aufgabenform:

Das Verfahren der semantischen Merkmale führt automatisch zu Antwortauswahlaufgaben

Restriktionen:

wie oben erwähnt

Konstruktionsvorschriften:

Zufallsreihenfolge der Items

Literatur

- Berk, R. A.: Estimation of test length for domain-referenced reading comprehension tests. *Journal of Experimental Education* 1980, 48, 188–193.
- Bormuth, J. R.: On the theory of achievement test items. The University of Chicago Press, Chicago 1970.
- Feger, B.: Die Generierung von Testitems zu Lehrtexten. *Diagnostica* 1984, 30, 24–46.
- Finn, P. J.: A question writing algorithm. *Journal of Reading Behavior* 1975, 7, 341–367.
- Herbig, M.: Aufgabentypen zur Leistungsüberprüfung. In: Klauer, K. J., Fricke, R., Herbig, M., Rupperecht, M. & Schott, F.: *Lehrzielorientierte Tests*. Pädagogischer Verlag Schwann, Düsseldorf 1972, 74–100.
- Hoppe, F.: Die Extraktion von Bedeutung aus Geschichten: Ein Produktionssystem zur Erzeugung von Makro- aus Mikrostrukturen. *Archiv für Psychologie* 1984, 136, 63–76.
- Klauer, K. J.: Lehrtextbezogene Tests. Transformation von Lehrtexten in Universa von Testaufgaben. In: K. J. Klauer & H.-J. Kornadt (Hrsg.): *Jahrbuch für Empirische Erziehungswissenschaft* 1979. Pädagogischer Verlag Schwann, Düsseldorf 1979, 33–52.
- Klauer, K. J.: Kontenvalidität. *Diagnostica* 1984, 30, 1–23 (a).
- Klauer, K. J.: Über Parallelität, Reliabilität und Validität kontenvalider Paralleltests. *Diagnostica* 1984, 30, 67–80 (b).
- Klauer, K. J.: Framework for a theory of teaching. *Teaching and Teacher Education* 1985, 1, 5–17.
- Klauer, K. J. & Dänecke, K.: Wie parallel sind lehrzielvalide Paralleltest? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 1981, 13, 181–189.
- Leutner, D.: *Lehrstoffstruktur und Leistung*. Dissertation Philosophische Fakultät der RWTH Aachen, Aachen 1985.
- Lienert, G. A.: *Testaufbau und Testanalyse*. Verlag Julius Beltz, Weinheim 1969.
- Meyer, H. L.: Das ungelöste Deduktionsproblem in der Curriculumforschung. In: F. Achtenhagen & H. L. Meyer (Hrsg.): *Curriculumrevision. Möglichkeiten und Grenzen*. Kösel, München 1971.
- Nußbaum, A.: *Konstruktion, Planung und Analyse lehrzielorientierter Tests auf der Grundlage der Generalisierbarkeitstheorie*. Dissertation Philosophische Fakultät der RWTH Aachen, Aachen 1980.
- Nußbaum, A., Feger, B. & Leutner, D.: Definition und Messung affektiver Lehrziele. In: K. Ingenkamp (Hrsg.): *Sozial-emotionales Verhalten in Lehr- und Lernsituationen*. Erziehungswissenschaftliche Hochschule Rheinland-Pfalz, Landau 1984, 211–222.
- Rütter, T.: *Formen der Testaufgabe*. Verlag C. H. Beck, München 1973.
- Rupperecht, H.: *Konstruktion von Testaufgaben nach einem Verfahren von Bormuth*. In: K. J. Klauer, R. Fricke, M. Herbig, H. Rupperecht & F. Schott: *Lehrzielorientierte Tests*. Pädagogischer Verlag Schwann, Düsseldorf 1972, 101–115.

Verfasser:

Prof. Dr. K. J. Klauer, Dr. Barbara Feger, Institut für Erziehungswissenschaft der Technischen Hochschule, Eilfschornsteinstr. 7, 5100 Aachen