

## Muster eines Algorithmus zur Realisation unterrichtswissenschaftlicher Falsifikationsexperimente

von Harald RIEDEL, Berlin (D)

aus dem Institut für Unterrichtsmethodenlehre im Bereich der Technischen Universität Berlin

### 1. Zum Problem der Falsifikation in der Unterrichtswissenschaft

Noch vor 10 bis 15 Jahren pflegten sich Unterrichtswissenschaftler, die (noch) nicht Anhänger der „kritisch-emanzipatorischen“ Richtung waren, dem wissenschaftstheoretischen Ansatz des „kritischen Rationalismus“ zuzuordnen. Als Konsequenz hätte man im vergangenen Jahrzehnt den systematischen Versuch erwarten können, die Poppersche Grundidee der Falsifikation wenigstens bei der Erforschung unterrichtswissenschaftlicher Gesetzmäßigkeiten zu verwirklichen.

Tatsächlich aber fand nicht die Methode, sondern lediglich der Terminus „Falsifikation“ Eingang in die Literatur. Die verhältnismäßig wenigen systematischen, theoriegeleiteten Experimente wurden nach wie vor als (im engeren Sinne positivistisch-induktivistische) Verifikationsexperimente durchgeführt. Erzielte man Ergebnisse, die - im Sinne der Theorie - positiv waren, so wurde befriedigt festgestellt, daß die Theorie „nicht falsifiziert“ worden war.

Ich erkläre diese unbefriedigende Entwicklung folgendermaßen:

- Mit der Verbreitung der fast übermächtig gewordenen „kritisch-emanzipatorischen“ Richtung wurde der „kritisch-rationale“ Ansatz von Popper - ungerechtfertigt pauschal - mit dem Positivismus identifiziert und demgemäß heftig abgelehnt.
- Als Poppers Ideen einer breiteren Öffentlichkeit zugänglich wurden (1973 erschien der Sammelband „Objektive Erkenntnis“), unterschied man noch nicht genügend zwischen den verschiedenen Absichten und Gegenständen theoretischer, technologischer und praxeologischer Forschung. Popper ging noch davon aus, daß Falsifikationsexperimente für theoretische wie für technologische Forschungsanliegen gleichermaßen geeignet seien (vgl. z.B. Popper 1948, S. 390).

Inzwischen hat sich diese Annahme jedoch als unhaltbar erwiesen. Die sog. „Strukturidentität“ der genannten Forschungsrichtungen ist nur scheinbar und formal. Gerade im unterrichts- und erziehungswissenschaftlichen Bereich dürfen sich Forscher glücklich schätzen, wenn sie überhaupt nachweisen können, daß sich *technologische* Verfahren und/oder Mittel als signifikant wirksam erweisen (vgl. dazu Riedel 1984). Ein solcher Nachweis gelingt angesichts der relativ kleinen Untersuchungsgruppen selbst in Verifikations-Experimenten oft nicht. Anders liegen die Verhältnisse in jenen Bereichen *theoretischer* Forschung, deren Anliegen es ist, unterrichtswissen-

schaftliche Gesetzmäßigkeiten zu produzieren. Hier hat sich das von Popper postulierte Falsifikationsprinzip nicht nur als praktikabel, sondern auch für weiterführende Forschungen als sehr fruchtbar erwiesen (vgl. dazu z.B. Riedel 1985a, Breyer u.a. 1986). So läßt sich die Ablehnung von Falsifikationsexperimenten wohl z.T. auf solche Versuche zurückführen, mit denen man - unangemessenerweise - technologische Fragen zu lösen versuchte und scheiterte.

Die wichtigste Ursache aber scheint mit in der Kompliziertheit des Verfahrens und in dem psychischen Aufwand zu liegen, der vom Forscher selbst getrieben werden muß, um sich aus den ihm meist unbewußten, aber starken Bindungen an seine eigene Theorie zu lösen.

Mit seiner pluralistischen Philosophie postulierte K.R. Popper (1968) die Existenz dreier ontologisch grundverschiedener Teilwelten:

Die Welt 1 der physikalischen Zustände

Die (subjektive) Welt 2 der Bewußtseinszustände

Die Welt 3 der objektiven Ideen.

Von den zwischen diesen Teilwelten bestehenden Beziehungen sind für den Forscher vor allem die folgenden wichtig:

- Als Träger subjektiven Bewußtseins kann jeder Mensch Gegenstände der 3. Welt, also objektive Denkinhalte, insbesondere Theorien, erfassen.
- Zwar sind die Gegenstände der 3. Welt von Menschen geschaffen, müssen aber von den Gegenständen der 2. Welt, den subjektiven Bewußtseinsinhalten, unterschieden werden.

Wenn ein Forscher eine Theorie entwickelt, so ist sie - weil Produkt seines subjektiven Forscher-Bewußtseins - zunächst noch Gegenstand der 2. Welt. Zum Gegenstand der 3. Welt kann die Theorie erst werden, wenn sie „kritisch überprüft“ worden ist. Dazu gehört für Popper vor allen Dingen eine intersubjektive Diskussion, deren Ziel darin besteht, die Theorie zu falsifizieren. Gelingt die Falsifikation nicht, so hat sich die Theorie vorläufig bewährt und wird damit zum „objektiven“ Gegenstand der 3. Welt.

Fehlen nun aber dem Produzenten einer Theorie entsprechend kritische und gleichzeitig sachverständige Gesprächspartner - und das ist bei Neuentwicklungen gerade im Bereich der Unterrichtswissenschaften gar nicht so selten der Fall -, dann ist der Forscher allein auf sich gestellt, wenn er daran geht, seine Theorie kritisch, d.h. mit ernsthaften Falsifikationsanstrengungen, zu überprüfen.

Hiermit ist nun fast jeder Forscher überfordert. Die Theorie ist seines Geistes (liebster) Kind, da sie meist erst nach langer und mühsamer Entwicklungsarbeit zustande gekommen ist. Daher sind das subjektive Bewußtsein des Forschers und die (nur potentiell objektive) Theorie in einer so starken Weise miteinander verbunden, daß sich selbst bei positiver Einstellung des Forschers zur Falsifikation immer wieder *unbewußt* Fehler in die Prüfarbeit einschleichen, die die Theorie dann nicht - wie gefordert - belastet, sondern eher begünstigt. Ich selbst habe bei der Entwicklung früherer Falsifikationsexperimente oft mit Erstaunen feststellen müssen, welche Mängel anfangs auf diese unbewußte Weise in die experimentellen Überprüfungen von Theoremen eingingen.

Dies war der Anlaß, Kriterien und Verfahren zu entwickeln, die - zunächst für die Zwecke unserer Untersuchungen im Institut für Unterricht im allgemeinbildenden Bereich - vor der Einbeziehung solcher unbewußten und subjektiven Fehler schützen sollen. Sofern der Versuchsleiter eines Experiments nicht bereits selbst starke Vorbehalte gegenüber der zu überprüfenden Theorie hat, tritt der Effekt unbewußt subjektiver Begünstigung oder Verschleierung - mehr noch als bei der Planung - bei der Durchführung der Versuche auf. Das ist verständlich, da ja bei der Realisation komplexerer Experimente ohnehin die Bewußtseinskapazität des Experimentators vollständig durch die Organisation des Versuchsablaufs, die Beobachtung der Versuchsperson, das Notieren der Ergebnisse usw. beansprucht ist. Infolgedessen schien es sinnvoll und zweckmäßig, die erarbeiteten Kriterien in einen Algorithmus zu transformieren, durch den die gesamte Realisationsarbeit des Experiments gesteuert wird.

Ein Beispiel für solch einen Algorithmus will ich im folgenden darstellen.

## 2. Kriterien zur Steuerung eines Falsifikationsexperiments

Es wäre schwierig, den Algorithmus abstrakt, also ohne Bezug zu einem konkreten Experiment, noch verständlich darzustellen. Ich schicke daher einige Bemerkungen über ein Experiment voraus, für dessen Durchführung der Algorithmus angewendet wurde. Das Experiment beschäftigte sich mit den unterschiedlichen Schwierigkeitsgraden der beiden Internoperationen „Auswerten“ und „konvergent Denken“. Das Beispiel ist insofern typisch für die Anwendung des Algorithmus, als das Verfahren besonders geeignet ist, die als unterschiedlich prognostizierten Ausprägungen zweier (bei entsprechender Erweiterung: auch mehrerer) für den Unterricht wichtiger Operationen, Merkmale, Leistungen usw. falsifizierend zu überprüfen.

Zum Verständnis des folgenden genügt es, eine der drei aufgestellten Arbeitshypothesen zu nennen. Sie lautete: Da das auswertende Anwenden eines Sachverhalts leichter als das konvergent denkende Anwenden desselben Sachverhalts ist und da die Bewußtseinskapazität der Versuchspersonen im Untersuchungszeitraum konstant bleibt, werden die Versuchspersonen ... in Situationen, die für ... (sie) einen „kritischen Informationsgehalt“ besitzen, zwar noch fehlerlos auswerten, aber nicht mehr fehlerlos konvergent denken können (vgl. dazu Riedel 1985b, S. 115-117 und Breyer u.a. 1986, S. 61f). Überprüft wurde die Hypothese in Experimenten, die vom 17.9. 1984 bis 3.5.1985 an den musiktheoretischen Unterrichtsobjekten „Dreiklang“ und „Umkehrungen“ mit Schülern der 7. bis 10. Klassen durchgeführt wurden. Bild 1 zeigt die Voraussetzungsbeziehungen zwischen den einzelnen Teilobjekten des Unterrichtsobjekts. Die jeweils am Ursprung der Pfeile stehenden Unterrichtsobjekte werden als Voraussetzung zum Erwerb der an den Pfeilspitzen stehenden Objekte betrachtet. (Alle weiteren Einzelheiten über Anlage, Durchführung und Ergebnisse der Experimente können einem folgenden Beitrag, voraussichtlich im nächsten Heft, entnommen werden.)

### 2.1 Kriterium 1: Parallelisierung der Versuchspersonen entsprechend ihrem Anfangszustand

Sollen wie im geschilderten Fall die unterschiedlichen Merkmale oder Ausprägungen zweier Leistungen überprüft werden, so wird es in Falsifikationsexperimenten unum-

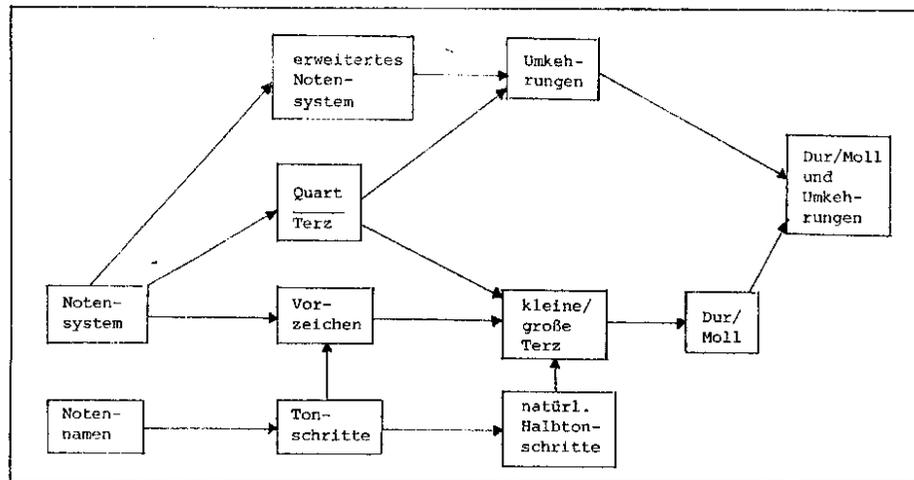


Bild 1: Voraussetzungs-Struktur der Unterrichtsobjekte „Dreiklang und Umkehrungen“

gänglich, beide Leistungen von ein und derselben Versuchsperson durchführen zu lassen und nicht etwa nur die statistischen Mittel oder Mediane zweier Gruppen miteinander zu vergleichen, von denen eine die Leistung X, die andere Gruppe die Leistung Y zu erbringen hat. Insofern muß also jede Versuchsperson als ihr eigener Kontrollpartner fungieren. Einzelheiten dazu werden noch ausgeführt.

Andererseits beeinflußt die Ausführung der jeweils ersten Operation auch die Fähigkeit, mit der eine Versuchsperson die zweite Operation durchführen kann. Um Verfälschungen der Ergebnisse zu vermeiden, müssen zwei Gruppen von Versuchspersonen gebildet werden, in unserem Fall:

- eine Gruppe A, die das jeweilige Unterrichtsobjekt immer erst auswertend und anschließend konvergent denkend anzuwenden hat
- eine Gruppe B, die das Unterrichtsobjekt erst konvergent denkend und anschließend auswertend anzuwenden hat.

Schon für Ergebnisse von Verifikations-Experimenten ist es ausschlaggebend, nach welchen Verfahren die Versuchspersonen den beiden Gruppen zugeordnet werden. Für Falsifikationsexperimente gilt dies in verschärfter Weise, denn die ohnehin in unterrichtswissenschaftlichen Experimenten vorhandenen Störungen durch nicht kontrollierbare Einflüsse fallen hier weitaus schärfer ins Gewicht. Entgegen Behauptungen in der einschlägigen Statistik-Literatur hatte sich in unseren vorangegangenen Experimenten herausgestellt, daß der Versuch, eine Parallelisierung der Versuchspersonen durch Zufallsziehung zu bewerkstelligen, völlig ungeeignet ist. Auch die Aufteilung der Versuchspersonen entsprechend ihren Schulnoten oder entsprechend

der Einschätzung ihrer Leistungen durch Lehrer genügen den Anforderungen eines Falsifikationsexperiments nicht. Parallelisierung heißt in unserem Falle: Die Versuchspersonen sollen auf die Gruppen A und B gleichmäßig verteilt werden, und zwar entsprechend ganz bestimmter Fähigkeiten, nämlich hinsichtlich der Beherrschung der Operationen „auswerten“ und „konvergent denken“ im Zusammenhang mit den musiktheoretischen Unterrichtsobjekten „Dreiklang“ und „Umkehrungen“.

Andere Fähigkeiten oder Merkmale der Versuchspersonen sind für den Versuch unwichtig, da ja jede Versuchsperson ihr eigener Kontrollpartner ist. Um dem oben beschriebenen Parallelisierungs-Ideal möglichst nahezukommen, wurden die Versuchspersonen in einem ersten Schritt vor Beginn des eigentlichen Experiments sog. „Filter-Aufgaben“ unterzogen, mittels derer ihr Anfangszustand nach den Gesichtspunkten „Bekanntheit des jeweiligen Unterrichtsobjekts“ und „Operationsfähigkeit in Verbindung mit diesem Unterrichtsobjekt“ festgestellt wurde. Filter-Aufgaben wurden zu den Objekten „Notensystem“, „erweitertes Notensystem“, „Vorzeichen“, „Umkehrungen“ und „Dur/Moll“ konstruiert. Jede Filter-Aufgabe bestand aus zwei Teilaufgaben: einer leichteren und einer schwereren bzw. einer Aufgabe zum Auswerten und einer zum konvergenten Denken. Es würde den Rahmen dieses Aufsatzes sprengen, darzustellen, wie der Versuchsablauf aufgrund der in den Filter-Aufgaben gezeigten Leistungen gesteuert wurde. Die Grundstruktur zeigt Bild 2. Den Versuchspersonen werden beide Teil-Filter-Aufgaben, die zu einem Objekt gehören, gegeben. Reagiert die Versuchsperson positiv auf beide Teilaufgaben, so kann angenommen werden, daß das behandelte Objekt für sie relativ informationsarm ist. Die Versuchsperson erhält dann Filter-Aufgaben zum nächst-schwereren Filter-Objekt. Im Extremfall würde die Versuchsperson alle Aufgaben lösen und damit für das weitere Experiment ausfallen. Lassen die (fehlerhaften) Reaktionen der Versuchsperson vermuten, daß das Objekt für sie relativ informationsreich ist, so folgt eine Unterrichtsphase.

Die Parallelisierung wurde nun so bewerkstelligt, daß jene Versuchspersonen, die nach derselben Filter-Aufgabe in die Unterrichtsphase gelangten, also demgemäß den gleichen Anfangszustand aufwiesen, abwechselnd den Gruppen A und B zugeordnet wurden.

Mit den anschließenden Unterrichtsphasen wurde ein zweiter Schritt zur Egalisierung der Anfangszustände realisiert. Durch den Unterricht sollten jene Leistungsunterschiede ausgeglichen werden, die durch die relativ groben Filter-Aufgaben noch nicht erfaßt werden konnten. In der Unterrichtsphase erhält daher jede Versuchsperson jene Information, die sie zur Bewältigung der (nur fehlerhaft gelösten) Filter-Aufgabe benötigt hätte. Außerdem aber werden noch Informationen über jenes Unterrichtsobjekt vermittelt, das in der Voraussetzungsstruktur nach Bild 1 dem gerade behandelten Unterrichtsobjekt folgt. Die Versuchsperson muß also zwei Unterrichtsobjekte nacheinander erkennen (ggf. teilweise erinnern). Neben der beabsichtigten Nivellierung des Anfangszustandes soll damit erreicht werden, daß die Versuchsperson genügend neue Informationen erhält, also an die für die Überprüfung der Hypothese wichtigen Schwelle der „kritischen Informationsmenge“ gebracht wird.

2.2. Kriterium 2: Individueller Verlauf des Experiments in Abhängigkeit bisheriger Leistungen

Nach der ersten Unterrichtsphase beginnt das eigentliche Experiment. Es setzt sich aus einer Reihe von Prüfaufgaben zusammen, die jeweils eine Teilaufgabe zum auswertenden und eine Teilaufgabe zum konvergent denkenden Anwenden der gerade gelernten Information enthält. Die Reihenfolge der Prüfaufgaben ist aber weder fest vorgegeben noch beliebig, sondern wird dadurch bestimmt, mit welchem Anfangszustand die Versuchsperson in das Experiment einstieg und welche Leistungen sie bisher in bezug auf das jeweilige Unterrichtsobjekt und die beiden Operationen Auswerten bzw. konvergent Denken erbracht hat.

Bild 2 zeigt ausschnittsweise, daß es nach jeder Prüfaufgabe grundsätzlich vier

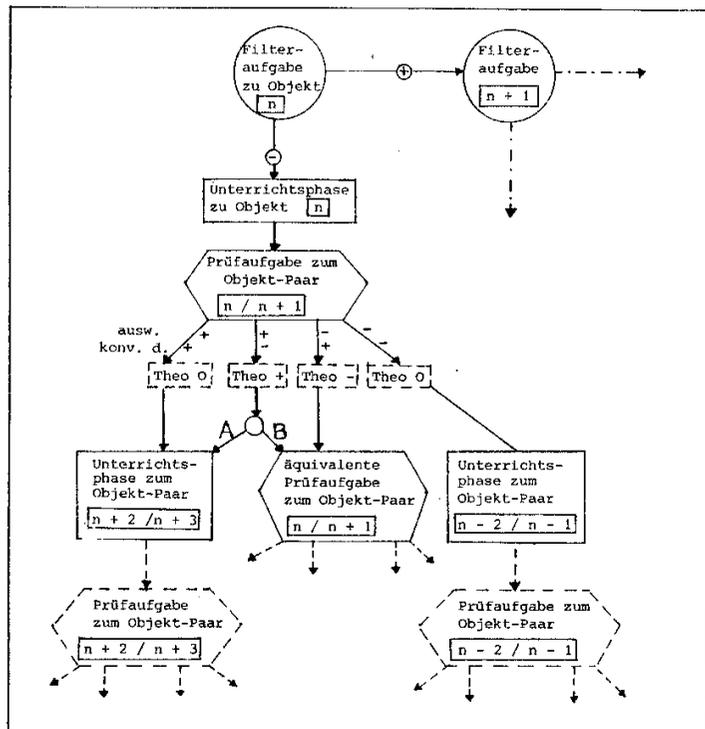


Bild 2: Grundstruktur des Algorithmus

Wegfortsetzungen gibt. Das gilt gleichermaßen für Versuchspersonen der Gruppe A, die zunächst die Teilaufgaben zum auswertenden, dann jene zum konvergent denkenden Anwenden erhalten, wie für die Versuchspersonen der Gruppe B, die die Teilaufgaben in umgekehrter Reihenfolge absolvieren.

Weg "+, +"

Die Versuchsperson führt im ersten Experiment beide Operationen fehlerlos durch. Der Test erbringt in diesem Fall keine Information über den Wahrheitswert der Hypothese (Theo-null). Außerdem ist davon auszugehen, daß die Prüfaufgaben für die Versuchsperson zu informationsarm waren. Um nochmals den Versuch zu unternehmen, den Probanden an die Schwelle der „kritischen Informationsmenge“ zu bringen, lernt er in der folgenden Unterrichtsphase die Informationen über das nächste Objektpaar (n+2/n+3) des Voraussetzungsnetzes aus Bild 1. Anschließend wird die Fähigkeit zum auswertenden und konvergent denkenden Anwenden dieser Unterrichtsobjekte durch eine weitere Prüfaufgabe zu diesen schwierigeren Unterrichtsobjekten getestet.

Weg "+, -"

Die Versuchsperson kann das erlernte Unterrichtsobjekt zwar fehlerlos auswerten, aber nur fehlerhaft konvergent denkend anwenden. Diese Reaktion stützt die Hypothese (Theo-plus). Gehört die Versuchsperson der Gruppe A an, so fährt sie auf dem Wege 1 fort, gehört sie dagegen zur Gruppe B, so wird sie mit einer äquivalenten Prüfaufgabe konfrontiert. Das ist eine Aufgabe, die sich auf dasselbe Unterrichtsobjekt bezieht wie die gerade vollzogene Prüfaufgabe und auch denselben Schwierigkeitsgrad besitzt. Warum sich die Wegfortsetzungen für die beiden Gruppen unterscheiden, wird mit Kriterium 2 erklärt.

Weg "-, +"

Die Versuchsperson kann umgekehrt wie im eben geschilderten Fall die Teilaufgabe zum konvergenten Denken fehlerlos lösen, nicht jedoch die Teilaufgabe zum Auswerten (-,+). Dieser Fall stellt einen krassen Widerspruch zur Hypothese dar (Theo-minus). Um den Fall - wegen des Bemühens um Falsifikation der Theorie nach Möglichkeit zu wiederholen, erhält die Versuchsperson sofort anschließend eine äquivalente, also im Schwierigkeitsgrad gleichartige, Prüfaufgabe.

Weg "-, -"

Die Versuchsperson kann keine der beiden Operationen fehlerlos durchführen (-,-). Wir erhalten damit keine Information über den Wahrheitsgehalt der Theorie (Theo-null). Die Informationsmenge der gerade gelernten Unterrichtsobjekte ist vermutlich zu groß. Daher erhält die Versuchsperson wieder Unterricht, und zwar über die beiden nächst-leichteren Objekte (n-2/n-1), die in der Voraussetzungsstruktur nach Bild 1 den gerade überprüften Unterrichtsobjekten vorangehen. Es folgen die Prüfaufgaben zum auswertenden und konvergent denkenden Anwenden dieser Unterrichtsobjekte.

Für den Weg 4 wurden an geeigneten Verzweigungspunkten des Organisations-schemas zusätzlich zwei alternative Fortsetzungsmöglichkeiten eingeführt, um dem bisherigen Lernweg der Versuchspersonen noch besser gerecht zu werden und damit auch die Wahrscheinlichkeit zu erhöhen, daß die Situation der „kritischen Informationsmenge“ erzeugt wird. Dabei war folgender Grundgedanke ausschlaggebend:

Bei jeder Prüfaufgabe besteht die Möglichkeit, null bis vier Fehler zu machen. Hat die Versuchsperson bei einer der Operationen alle vier Teilaufgaben falsch beantwortet, so ist davon auszugehen, daß sie mit dem Unterrichtsobjekt stark überfordert ist. Sie erhält daher Unterricht zu jenem Paar von Unterrichtsobjekten, das im Voraussetzungsnetz vor dem soeben überprüften Unterrichtsobjekt steht, und wird anschließend der Prüfaufgabe zu diesen Unterrichtsobjekten unterzogen. Hat die Versuchsperson dagegen wenigstens eine richtige Antwort gegeben, so folgt eine äquivalente Prüfaufgabe, um möglicherweise doch noch einen Fall „Theo-plus“ oder „Theo-minus“ zu erreichen.

Das Experiment wird - mit Ausnahmen, die noch beschrieben werden - solange fortgesetzt, bis die Versuchsperson den Versuch nicht mehr fortsetzen möchte oder die Operationsfähigkeiten bereits am letzten (schwierigsten oder leichtesten) Unterrichtsobjekt überprüft wurden.

### 2.3 Kriterium 3: Systematische Belastung der Hypothese

Will ein Forscher seine Theoreme bzw. die hiervon durch Operationalisierung abgeleiteten Hypothesen falsifizieren, genügt es nicht, die experimentellen Bedingungen so zu gestalten, daß die Chancen für und gegen das Eintreten der durch die Hypothesen vorausgesagten Ereignisse gleich groß sind. Vielmehr muß er die Hypothese belasten. Er muß also versuchen, die Wahrscheinlichkeit für das Eintreten der von der Hypothese postulierten Ereignisse systematisch zu reduzieren. Erst dadurch kann sich der Vorteil von Falsifikationsbemühungen auswirken, nämlich die Grenzen aufzuzeigen, innerhalb derer die Hypothesen (falls überhaupt) noch zutreffen.

Solche Falsifikationsbemühungen fallen besonders jenen Experimentatoren schwer, die von der Stichhaltigkeit ihrer Hypothesen überzeugt sind. Zusätzlich aber wird die Realisation durch zwei Momente erschwert: Zunächst ist es schwierig, vorherzusehen, bis zu welchem Ausmaß die Belastung sinnvoll ist. Dann aber erfordern systematische Falsifikationsbemühungen so unterschiedliche experimentelle Wegfortsetzungen, daß das gesamte Realisationsgeschehen den Experimentator stark überfordert, sofern er sich nicht durch externe Vorgaben, z.B. einen Algorithmus in Form von Flußdiagrammen, steuern läßt. In den erwähnten Experimenten konkretisierten sich die Falsifikationsbemühungen in vierfacher Weise:

#### Falsifikationsaspekt 1:

Die Versuchspersonen der Gruppe A müssen das jeweilige experimentelle Objekt zuerst auswertend und anschließend konvergent denkend anwenden, die Versuchspersonen der Gruppe B verfahren in umgekehrter Reihenfolge. Vorausgesetzt, daß Auswerten leichter als konvergentes Denken ist - was die Arbeitshypothese behauptet - läßt sich für jene Situationen, in denen der jeweils „kritische Informationsgehalt“ erreicht wird, folgendes voraussehen:

In Gruppe A führt die Versuchsperson zunächst die leichtere, dann die schwerere Operation durch.

Entsprechend der Theorie kann man davon ausgehen, daß durch die erste, leichtere Operation die Durchführung der zweiten Operation vorbereitet wird. Außerdem wird die zweite, schwierigere Operation dadurch erleichtert, daß während der Durchführung

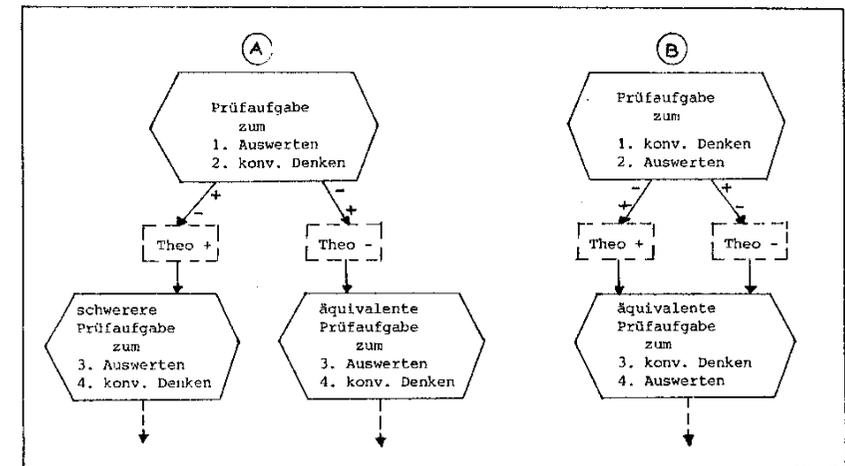


Bild 3: Wegfortsetzungen für Gruppe A und Gruppe B

der ersten Operation der subjektive Informationsgehalt des erlernten Unterrichtsobjekts verringert wird.

Hingegen ist für die Versuchspersonen der Gruppe B folgendes zu erwarten: Die Versuchsperson muß mit dem Objekt, das ihre Bewußtseinskapazität fast völlig in Anspruch nimmt, sofort konvergent denkend (also auf höherer Schwierigkeitsstufe) operieren. So ist zu befürchten, daß die Versuchsperson nun einige Fehler macht und dadurch verwirrt wird. Damit wird die Durchführung der zweiten Internoperation keinesfalls erleichtert (was sich in den Versuchen immer wieder bestätigte).

Aus beiden Überlegungen folgt, daß in Gruppe A wegen der Vorbereitung der zweiten, schwierigeren Operation relativ wenig Fälle auftreten werden, die die Theorie stützen. In Gruppe B dagegen, in der ja das Auswerten nicht von einer leichteren Internoperation vorbereitet wird, ist anzunehmen, daß eher bestätigende Fälle eintreten werden, also fehlerhafte Leistungen beim konvergenten Denken, aber fehlerlose Leistungen beim Auswerten.

Für ein Verifikationsexperiment würde es nun reichen, wenn sich überhaupt signifikante Unterschiede zugunsten der Fälle „Theo-plus“ ergeben würden. Für unser Falsifikationsexperiment dagegen forderten wir, daß sich die Hypothesen sowohl in Gruppe A als auch in Gruppe B signifikant bewähren.

#### Falsifikationsaspekt 2:

Bild 3 zeigt die Wegfortsetzungen in den Gruppen A und B jeweils nach den Fällen „Theo-plus“ und „Theo-minus“. Nach den Theo-minus-Fällen wird das Experiment in beiden Gruppen mit einer äquivalenten Prüfaufgabe fortgesetzt. Die Versuchspersonen

werden also nach allen Fällen, in denen sie konvergent richtig und auswertend fehlerhaft gearbeitet haben, nochmals mit einer gleich-schweren Prüfaufgabe konfrontiert. Dahinter steht die Absicht, die Versuchsperson wiederum in eine Situation zu versetzen, in der sie die theoriebelastende Leistung erbringt. Unterschiedlich ist natürlich in Gruppe A und B die Reihenfolge der Internoperationen.

Die Fortsetzungen nach den Theo-plus-Fällen dagegen unterscheiden sich in Gruppe A und Gruppe B auch hinsichtlich des Schwierigkeitsgrads der Prüfaufgaben. In der Gruppe A folgt einem Theo-plus-Fall sofort eine schwerere Prüfaufgabe. Damit wird beabsichtigt, zu verhindern, daß die Versuchsperson mit einer eventuell äquivalenten Aufgabe nochmals den Fall „Theo-plus“ wiederholt, was der Absicht der Falsifikation ja widersprechen würde. Anders wird dagegen in Gruppe B verfahren. Zum Verständnis des unterschiedlichen Vorgehens ist es notwendig, sich die Reihenfolge der Internoperationen anzusehen. Eine Versuchsperson B hat vor dem ersten Theo-plus-Fall an einem bestimmten Objekt 1. konvergent denkend, 2. auswertend gearbeitet. Nach dem Theo-plus-Fall muß sie nun an einem äquivalenten Objekt 3. (wieder) konvergent denken. Damit ist aber die 3. (konvergente) Operation durch die 2. (auswertende) Operation vorbereitet. Dies kann aus den oben genannten Gründen als Erleichterung des konvergenten Denkens angesehen werden. Demzufolge wird die Wahrscheinlichkeit vergrößert, daß die Versuchsperson nach dem richtigen Auswerten des Objekts das äquivalente Objekt nun auch konvergent denkend richtig bewältigt. Jedenfalls ist diese Wahrscheinlichkeit bei einem gleich schweren Objekt größer als bei einem schwereren Objekt.

Insgesamt ist das Organisationsschema also derart angelegt, daß die Wiederholung von Theo-minus-Fällen wahrscheinlicher ist als jene von Theo-plus-Fällen.

#### Falsifikationsaspekt 3:

Die Eintrittswahrscheinlichkeit der Theo-minus-Fälle gegenüber den Theo-plus-Fällen wird durch eine weitere Maßnahme erhöht. Das zeigt Bild 4. Man betrachte zunächst den rechten Strang. Die Versuchsperson hat einen Theo-minus-Fall produziert, erhält also eine Prüfaufgabe 2 (die in diesem Fall äquivalent zur Prüfaufgabe 1 ist). Kommt wiederum ein Theo-minus-Fall zustande, wird der Versuch abgebrochen. Erreicht die Versuchsperson dagegen einen Theo-plus-Fall, so folgt eine dritte Prüfaufgabe in der Erwartung, daß sich möglicherweise nochmals ein Theo-minus-Fall ereignet. Geschieht dies, so wird das Experiment nach diesen (wiederum) zwei Theo-minus-Fällen abgebrochen.

Betrachten wir den linken Strang. Die Versuchsperson hat einen Theo-plus-Fall erreicht, gelangt zur nächst-schwierigeren Prüfaufgabe und wird, solange sie noch Theo-plus-Fälle erreicht, mit neuen Prüfaufgaben konfrontiert, deren Unterrichtsobjekte jeweils einen höheren Schwierigkeitsgrad aufweisen. Die Wahrscheinlichkeit also, daß sich überhaupt 3 Theo-plus-Fälle nacheinander ereignen, ist somit äußerst gering. (In der gesamten Untersuchung trat dieses Ereignis nur ein einziges Mal auf.)

Andererseits aber besteht die Möglichkeit, daß die Versuchsperson nach dem ersten Theo-plus-Fall auch einen Theo-minus-Fall zeitigt. In diesem Fall wäre die Prüfaufgabe 3 in ihrem Schwierigkeitsgrad äquivalent und damit die Wahrscheinlichkeit, daß der Theo-minus-Fall wiederholt wird, relativ groß. Nach zwei Theo-minus-Fällen aber wird

die Versuchsserie abgebrochen. Zusammengefaßt: Bewirkt die Versuchsperson einen Theo-plus-Fall, so wird das Experiment unter erschwerteren Bedingungen fortgesetzt, in der Hoffnung, daß sich noch Theo-minus-Fälle ereignen. Im Gegensatz hierzu wird das Experiment abgebrochen, wenn die Versuchsperson zwei Theo-minus-Fälle erreicht hat. Dies wird als absoluter Widerspruch zur Gültigkeit der Theorie bei dieser Versuchsperson gewertet.

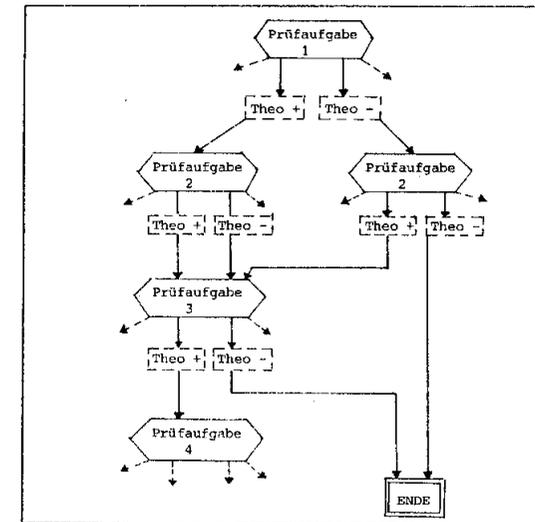


Bild 4: Ungleichgewicht zwischen Theo-plus- und Theo-minus-Fällen

#### Falsifikationsaspekt 4:

Die bisher dargestellten Maßnahmen zur Falsifikation waren alle formaler Art. Daneben gibt es inhaltliche Aspekte, nach denen die Hypothese belastet werden kann. Sie werden sich je nach den für die Untersuchung herangezogenen Objekten und Fragen unterscheiden. Das folgende Beispiel darf also lediglich als eine mögliche Konkretion verstanden werden.

Bei der Konstruktion von Prüfaufgaben ist es wichtig, den Schwierigkeitsgrad nach zwei Richtungen hin möglichst gleich zu halten:

- Zu jeder Teilaufgabe des konvergenten Denkens und des Auswertens muß eine äquivalente, also gleich-schwere Teilaufgabe erzeugt werden, da entsprechend dem Organisationsschema häufig äquivalente Prüfaufgaben zu vorhergegangenen benötigt werden.

- Die Aufgaben zum Auswerten einerseits und zum konvergent denkenden Anwenden desselben Unterrichtsobjekts andererseits müßten für Verifikationszwecke den gleichen Schwierigkeitsgrad haben. Zu jedem Unterrichtsobjekt müßten also - von der Internoperation abgesehen - vier Teilaufgaben gleichen Schwierigkeitsgrades erzeugt werden. Das läßt sich allerdings in der Praxis nicht einfach realisieren. Denn die Anzahl der nicht allzu schwierigen Kombinationen von beispielsweise Intervallen, Dreiklängen oder Umkehrungen ist ja begrenzt, und es muß verhindert werden, daß die Versuchspersonen sich an zuvor präsentierte Notenkombinationen erinnern können. Wir verfahren deshalb so, daß wir den Schwierigkeitsgrad der Teilaufgabe zum Auswerten immer höher als jenen der Teilaufgabe zum konvergenten Denken setzten, niemals aber umgekehrt, falls es nicht möglich war, äquivalente Aufgaben mit völlig gleichem Schwierigkeitsgrad zu erzeugen.

Ein konkretes Beispiel für eine extreme Belastung der Auswertaufgabe aufgrund der durchzuführenden Einzeloperationen zeigt Bild 5:

Bei der Aufgabe zum konvergenten Denken ist das „g“ vorgegeben. Es soll ein Ton erzeugt werden, der einen Halbton höher als „f“ klingt, es muß also ein „b“ eingetragen werden, da „g“ höher als „f“ liegt. Bei der Aufgabe zum Auswerten ist ein „dis“ vorgegeben. Zunächst ist zu bemerken, daß das „g“ in der Teilaufgabe zum konvergenten Denken innerhalb des einfachen Notensystems liegt, also in jener Tonleiter, die die Schüler zuerst lernen. Das „dis“ in der Aufgabe zum Auswerten dagegen liegt schon außerhalb dieses Tonbereichs, ist also schwerer zu erinnern.

Betrachtet man die Einzeloperationen, die die Versuchspersonen beim Auswerten oder beim konvergenten Denken ausführen muß, so erkennt man ein starkes - in diesem Fall beabsichtigtes - Mißverhältnis zu Lasten der Auswertaufgabe. Die einzelnen Detailoperationen sind in Bild 5 wiedergegeben. In der Auswertaufgabe muß die Versuchsperson achtmal kogneszierende Operationen durchführen (Erkennen, Erinnern), bei der Aufgabe zum konvergenten Denken dagegen nur vier. An produzierenden Operationen (auswerten, konvergent denken) muß die Versuchsperson bei der Auswertaufgabe vier, dagegen nur zwei bei der Aufgabe zum konvergenten Denken durchführen.

Nach der Lektüre der vier Falsifikationsaspekte mag der Leser den Eindruck gewonnen haben, daß eine so systematische Belastung der Hypothese immer zur Falsifikation führen müsse. Tatsächlich aber hat sich in unseren Experimenten die oben dargestellte Hypothese (neben anderen) bewährt, allerdings mit Ausnahme jenes Objekts „Vorzeichen“, das zusätzlich durch die Art und Zahl der Detail-Operationen belastet wurde. Gerade die Falsifizierung der Hypothese an diesem einen Objekt liefert aber wichtige Zusatzinformationen: Wenn das Objekt hinsichtlich der Detail-Operationen doppelt so schwer beim auswertenden Anwenden ist wie jenes zum konvergent denkenden Anwenden, dann ist die Auswertaufgabe eben *nicht* mehr leichter als jene zum konvergenten Denken, sondern schwerer. Es müssen nun weitere Hypothesen aufgestellt werden, die genauere Informationen über die Grenze der inhaltlichen Belastung liefern. Genau in diesem Sinn aber wollte Popper das Falsifikationsverfahren eingesetzt wissen: Nicht nur um den allgemeinen Wahrheitsgehalt von Hypothesen zu klären, sondern auch um weiterführende und differenzierende Hypothesen abzuleiten.

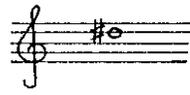
"Vorzeichen"	
auswertend anwenden	konvergent denkend anwenden
	
klingt <input type="checkbox"/> höher als (e) <input type="checkbox"/> tiefer als (c) <input type="checkbox"/> tiefer als (d)	Der Ton soll einen Halbton höher als (f) klingen. Trage das richtige Vorzeichen ein:
erkennen : Note d erkennen : # erinnern : dis höher d erkennen : (e) auswerten : d ≠ c erinnern : e höher d erkennen : (e) erinnern : d ≠ e auswerten : dis tiefer e erkennen : (d) auswerten : Note "d" ≠ Buchstabe d auswerten : # ≠ tiefer	erkennen : (f) erkennen : Note g auswerten : f ≠ g erinnern : g höher f konvergent d. → Halbton höher f = Halbton tiefer g erinnern : Halbton tiefer = b

Bild 5: Inhaltliche Belastung durch Art und Zahl der Detail-Operationen

### 3. Erfahrungen mit der Anwendung des Algorithmus

Alle beschriebenen Kriterien und Falsifikationsaspekte werden durch den Algorithmus zur Realisation unterrichtswissenschaftlicher Falsifikationsexperimente konkretisiert. Da der Algorithmus - wie in Punkt 1 ausgeführt - nicht nur für die Planung und Entwicklung von Falsifikationsexperimenten, sondern insbesondere als Hilfe für den Versuchsleiter gedacht ist, müssen darin alle Kriterien und Falsifikationsaspekte auf die einzelnen Unterrichtsobjekte, auf ihre Abhängigkeiten entsprechend dem Voraussetzungsnetz nach Bild 1 und auf die jeweiligen Reaktionen der Versuchspersonen

bezogen sein. Dadurch ist der Algorithmus zu differenziert und umfangreich, als daß er im Rahmen dieses Beitrags abgebildet werden könnte.\*

Der Algorithmus wurde zur Steuerung der Versuchsleiter-Tätigkeit in den in Abschnitt 2.0 erwähnten Experimenten eingesetzt. Dabei zeigte sich, daß die vom Experimentator zu leistende Informationsverarbeitung sehr stark reduziert werden konnte. Die daraus resultierende Entlastung führte zu einer drastischen Verringerung der Fehlerquote seitens des Versuchsleiters, insbesondere bei der Einhaltung der Reihenfolge, in der die Operationsobjekte für die Unterrichtsphasen und die Prüfaufgaben vorgelegt werden mußten.

### *Schrifttum*

- BREYER, I., H. RIEDEL, A. SIEGMUND: Kontrolle experimenteller Schwierigkeitsstufungen in der Internoperationalen Kybernetik. *Erkenn* 1986, S. 61-73
- POPPER, K. R.: Zur Theorie des objektiven Geistes. 1968. In: K. R. Popper: *Objektive Erkenntnis. Ein evolutionärer Entwurf*. Hoffmann und Campe, 1974, S. 172-212
- RIEDEL, H.: Aufbau und Ergebnisse eines Falsifikationsexperimentes zur Schwierigkeitsstufung in der Internoperationalen Kybernetik. *Erkenn* 1985a, S. 163-176
- RIEDEL, H.: Zur Methodologie der unterrichtswissenschaftlichen Experimente. In: K. Aurin, B. Schwarz (Hrsg.): *Die Erforschung pädagogischer Wirkungsfelder. Arbeitsgruppen empirischer pädagogischer Forschungen der DGfE Albert-Ludwig-Universität Freiburg/Breisgau* 1985b, S. 113-127

Eingegangen am 11. Juli 1986

### *An Example of an Algorithm for Executing Pedagogic Falsification Experiments (Summary)*

The planning and execution of falsification experiments causes researchers in education considerable psychological as well as organisational problems. To simplify the task for the researcher we have developed an algorithm that places a systematic strain upon the hypothesis in question. A justification for the algorithm is presented by means of an experiment already carried out.

\* Interessierte Leser können eine Kopie des entsprechenden Flußdiagramms beim Verfasser anfordern.