

**Das Konzept des "empirischen Gehaltes"
von Hypothesen nach POPPER
und eine praktikable Operationalisierung**

Georg Lind

1991

Kontakt:

apl. Prof. i.R. Georg Lind
Universität Konstanz
E-Mail: Georg.Lind@uni-konstanz.de
www.uni-konstanz.de/ag-moral/

Das Konzept des "empirischen Gehaltes"

von Hypothesen nach POPPER

und eine praktikable Operationalisierung*

Herr Dekan, meine Damen, meine Herren!

In der *Psychologischen Rundschau* hat Heinrich *Wottawa* kürzlich ein recht düsteres Bild von der psychologischen Methodenlehre gezeichnet: "Die gebräuchlichen statistischen Auswertungsverfahren", so Wottawa, "sind kein [...] Hilfsmittel für Erkenntnisfortschritt." Der Grund: "Welche Methoden zur Überprüfung von wissenschaftlichen Aussagen herangezogen werden, ist [...] eine Übereinkunft der entsprechenden Wissenschaftler-Gruppe; sie entzieht sich einer wissenschaftsinternen Gültigkeitsprüfung." Daher seien "grundlegende Verbesserungen mittelfristig kaum zu erwarten" (1990, S. 85).

Wottawa glaubt auch nicht, daß die Lage durch immer neue, komplexere statistische Auswertungsmethoden – wie Faktorenanalyse, Clusteranalyse, Lineare Gleichungssysteme, LISREL, Fuzzy-Logik usw. – verbessert werden könne. Die Suche nach einer besseren Methodologie, "darf nicht auf die Auswertungsverfahren konzentriert werden, da diese [...] ohne eine andere Forschungsstrategie keine zufriedenstellende Lösung bieten können" (S. 91). Noch schärfer formulierte es der Psychologe und Methodologe Paul *Meehl* (1967) bereits vor einigen Jahren: Statistische Verfahren würden in der Psychologie – im Gegensatz zur Physik – zumeist nur dazu benutzt, um gehaltlose Theorien zu bestätigen, was letztlich zur Behinderung des Fortschritts in der Psychologie führe.

Meehl meint jedoch, daß mit Hilfe des Popperschen Konzepts des Gehalts von Hypothesen eine Verbesserung der psychologischen Forschungsmethodologie möglich sei. Auch Volker *Gadenne* (1976; 1984) plädiert für die Entwicklung von Theorien und Hypothesen mit, er sagt, "möglichst hohem empirischen Gehalt" (1976, S. 54). Fortschritte in der psychologischen Forschung, so auch *Bredenkamp* (1980), könnten nur über falsifizierbarere, gehaltvollere Hypothesen erreicht werden.

*Überarbeitete Fassung meines Habilitationsvortrags vor der Philosophisch-Pädagogischen Fakultät der Katholischen Universität Eichstätt am 13.11.91. Die Anmerkungen und Literaturhinweise am Ende waren nicht Teil des Vortrags.

Diese Forderung nach Verbesserung des Gehalts von Hypothesen ist in der psychologischen Forschung bislang weitgehend ohne Resonanz geblieben. Die Lehrbücher zur Methodenlehre schweigen sich, soweit ich feststellen konnte, dazu fast alle aus. Für Studierende der Psychologie bleibt es daher unklar, wie diese Forderung methodisch umgesetzt werden kann.

Ich möchte hier das Konzept des empirischen Gehalts vorstellen und anhand von Beispielen zu zeigen versuchen, welche methodischen Konsequenzen es für die psychologische Forschung hat.

1.

“Das Konzept des “empirischen Gehalts” stammt von dem Wissenschaftstheoretiker Sir Karl *Popper*. Popper definiert den “empirischen Gehalt” (content) einer Hypothese als ihre “Unwahrscheinlichkeit”, d.h. als die Wahrscheinlichkeit, mit der sie sich falsifizieren läßt. Diese Falsifizierbarkeit ist bestimmt durch die *a-prior* Wahrscheinlichkeit, daß eine bestimmte Hypothese empirisch falsch ist. Nur wenn diese Wahrscheinlichkeit hoch und die Falsifikationswahrscheinlichkeit einer richtigen Vorhersage gering ist, läßt sich nach Popper von einer gehaltvollen oder informativen Hypothese sprechen.

Ein Beispiel: Die angebliche Bauernregel “Kräht' der Hahn auf dem Mist, ändert sich das Wetter, oder es bleibt wie es ist”, stellt eine Vorhersage oder Hypothese dar, die – aus rein logischen Gründen – immer wahr ist, weil sowohl die Aussage “das Wetter ändert sich”, wie auch ihr Gegenteil “das Wetter ändert sich nicht” vorhergesagt werden und somit die Vorhersage insgesamt durch kein denkbare Ereignis falsifiziert werden kann. Nichtfalsifizierbare, immer wahre Aussagen haben keinen Gehalt.

Für die Berechnung des Gehalts einer Hypothese hat *Popper* (1968b, S. 393 ff.) eine einfache Formel vorgeschlagen. Wenn wir aufgrund einer bestimmten Theorie den Sachverhalt *a* vorhersagen und erwarten, daß *a* mit der Wahrscheinlichkeit $p(a)$ anzutreffen ist, dann ist der Gehalt oder *Content* $Ct(a)$ der Hypothese *a* definiert als die zu $p(a)$ komplementäre Wahrscheinlichkeit $1-p(a)$ [s. Formel (1) im Anhang].

Hierzu ein weiteres Beispiel: Gegeben sei eine Theorie *T*, die besagt, daß bei einem Würfel jede Zahl dieselbe Chance hat, nach oben zu liegen. Auf dieser Grundlage machen wir die Vorhersage *a*, dass bei einem Wurf eine “3” oben liegt. Dann ist die logische Wahrscheinlichkeit $p(a)$ dieses Ereignisses gleich $1/6$, also ca. 0,17. Der empirische Gehalt dieser Hypothese ist $Ct(a) = 1-0,17 = 0.83$. Die Hypothese, eine “3” zu würfeln, ist also

relativ gewagt und damit ziemlich gehaltvoll.

Der Begriff der Wahrscheinlichkeit, wie er hier gebraucht wird, hat eine rein logische Bedeutung, er stützt sich nicht auf beobachtete Häufigkeiten. Die Wahrscheinlichkeit $p(a)$ ist definiert als das Verhältnis der von der Hypothese a vorhergesagten Ereignisse (VE) zu den unter der Hypothese a möglichen Ereignissen (ME). Im vorigen Beispiel gab es unter der Hypothese sechs mögliche Ereignisse, da der Würfel sechs Seiten hat. Mit anderen Worten: Die Hypothese a spannt einen *Möglichkeitsraum* auf, der sechs verschiedenen Ereignisse umfasst. Andere Ereignisse sind unter dieser Hypothese nicht möglich. Eines dieser möglichen Ereignisse, die Zahl "drei" ist das vorhergesagte Ereignis. Wenn von einem ungefälschten Würfel ausgegangen wird, hat jede Zahl dieselbe *a-priori*-Wahrscheinlichkeit, nämlich $p(a)=VE/ME=1/6$, und der Gehalt, wie gesehen, ist $Ct(a)=1-p(a)=.83$ [siehe Formel (2)]. Dieser Sachverhalt lässt sich auch anhand eines sogenannten Venn-Diagramms verdeutlichen. Die Wahrscheinlichkeit des Ereignisses a wird dabei durch den Sektor a in Abb. 2 dargestellt.

Popper bezeichnet den "empirischen Gehalt" von Hypothesen als den rationalen Kern einer Wissenschaft. Wissenschaftliche Hypothesen seien nicht allein danach zu beurteilen, ob sie wahr im Sinne von empirisch belegbar, sondern ob sie gehaltvoll oder informativ sind. Wahr können, so Popper, auch unwissenschaftliche Theorien sein. Solche Theorien zeichneten sich oft dadurch aus, dass ihre Vorhersagen fast immer empirisch wahr, und trotzdem oder gerade deswegen kaum gehaltvoll sind. "Das Kriterium für den wissenschaftlichen Status einer Theorie", so Popper, ist "seine Falsifizierbarkeit, Widerlegbarkeit oder Testbarkeit" – *falsifiability, refutability, or testability* (Popper 1968b, S. 37; Hervorhebung im Original).

2.

Diesem Kriterium, so Popper, würden weite Teile der Sozialwissenschaften – auch der Psychologie – nicht genügen. Er nennt hier vor allem den Marxismus und Sigmund Freuds Psychoanalyse. Aber auch manche Paradigmen der akademischen Psychologie gelten aus der Sicht der Falsifikationstheorie als pseudo- oder prä-wissenschaftlich. Hauptziel der Kritik ist – wie bei Wottawa, Meehl und Lakatos nachzulesen ist – die in der psychologischen Forschung weithin gebräuchliche Formulierung von *Alternativ-* und *Null-Hypothesen* und deren statistische Prüfung (Bortz 1984).

Worauf genau zielt diese Kritik?

Alternativ-Hypothesen haben gewöhnlich die Form: $H_1: \mu_1 \neq \mu_2$, wobei μ_1 und μ_2 die Mittelwerte zweier Messreihen bezeichnen. Dem entsprechen Formulierungen wie: “Personen mit internaler Kontrollwahrnehmung zeigen im Mittel einen *signifikant* anderen Partizipationsgrad als Personen mit externaler Kontrollwahrnehmung”.

Alternativ-Hypothesen werden in der Psychologie zumeist aus einer Theorie abgeleitet oder doch zumindest auf sie gestützt; so stützt sich das eben genannte Beispiel auf die Soziale Lerntheorie von Julian *Rotter* (Rotter 1971, Gore & Rotter 1972). Das Problem von Alternativ-Hypothesen liegt darin, dass mit diesen oft weder eine Richtung der Wirkung noch deren Größe spezifiziert wird. Im Fall der Sozialen Lerntheorie ist man sich zwar sicher, dass sich die Art der individuellen Kontrollwahrnehmung auf das Sozialverhalten von Personen auswirkt. Man ist sich jedoch nicht sicher, in welcher Richtung.¹ Die Literatur kennt zahlreiche Beispiele für solche unspezifischen Wirkungshypothesen.

Weil Alternativ-Hypothesen der Form $\mu_1 \neq \mu_2$ keine nähere Spezifikation der Wirkungsrichtung und -größe enthalten, haben sie eine extrem hohe *a-priori*-Wahrscheinlichkeit und damit einen extrem niedrigen Gehalt oder Informationswert. Wählt man eine genügend große Untersuchungsgruppe und/oder eine genügend hohe Präzision der Messwerte, dann können wir allein aufgrund logischer Überlegungen sicher sein, dass die Hypothese als statistisch “signifikant” bestätigt wird (vgl. Bortz 1984, S. 487 ff.). Mit anderen Worten: Wir haben es hier mit einer Sorte Hypothesen zu tun, die kaum oder nicht falsifiziert werden kann. Da die *a-priori*-Wahrscheinlichkeit p nicht-falsifizierbarer Hypothesen 1.0 ist, ist ihr Gehalt

$$Ct(\mu_1 \neq \mu_2) = 1 - p(\mu_1 \neq \mu_2) = 1 - 1 = 0 \text{ [s. Formel (3)].}$$

Viele Forscher meinen, wie Wottawa (1990, S. 89) feststellt, dass sich dieses Problem durch einen epistemologischen “Trick” lösen lässt: Man versucht, die Alternativ-Hypothese zu bestätigen, indem man ihr Gegenteil, die sogenannte “Null-Hypothese” H_0 widerlegt.² Null-Hypothesen haben die Form $H_0: \mu_1 = \mu_2$. In unserem Beispiel würde die Null-Hypothese lauten: “Personen mit internaler Kontrollwahrnehmung beteiligen sich genauso häufig am politischen Geschehen wie Personen mit externaler Wahrnehmung”. Wenn es gelänge, die Null-Hypothese, dass *keine* Unterschiede bestehen, zu widerlegen, dann sei damit automatisch die Alternativ-Hypothese H_1 bestätigt.

Wottawa verwirft diesen Lösungsweg als “unsinnig”, weil damit das Falsifikationsprinzip in sein Gegenteil verkehrt würde. Null-Hypothesen haben zwar eine verschwindend kleine *a-priori*-Wahrscheinlichkeit und sind daher leicht falsifizierbar und äußerst gehaltvoll. Die Bestätigung von Null-Hypothesen durch Beobachtungen trägt dennoch *nicht*, wie man aufgrund des Popperschen Falsifikationismus meinen könnte, zum Fortschritt der Wissenschaft bei. Aus der Widerlegung der Null-Hypothese folgt nämlich nur die völlig gehaltlose Alternativ-Hypothese, dass *irgend* ein Zusammenhang oder Unterschied besteht. Eine Theorie, aus der nur unspezifische, gehaltlose Hypothesen folgen, stellt aber, wie wir bereits sahen, kein Wissen dar. Sie ist ohne Bedeutung für praktische Anwendungen, und ihre Bestätigung in empirischen Untersuchungen ist daher nicht der Mühe wert. Damit ist gezeigt, dass die beiden klassischen Hypothesen-Typen der psychologischen Forschung keinen Beitrag zum Erkenntnisfortschritt in der Psychologie leisten.

3.

Der Trick mit der Null-Hypothese beruht tatsächlich auf Überlegungen Poppers, die sich inzwischen als unhaltbar erwiesen haben. Popper stellte in seiner *Logik der Forschung* die Behauptung auf, dass zwischen Theorien und Hypothesen auf der einen Seite und empirischen Sachverhalten auf der anderen Seite eine logische Beziehung bestehe, dass diese Beziehung aber, wie man sagen könnte, *a-symmetrisch* ist: empirische Sachverhalte, sogenannte Basissätze, könnten eine Theorie nicht logisch beweisen, aber sie könnten diese widerlegen. Popper führt hier das berühmte Rabenbeispiel an: Nehmen Sie an, Sie wollten die Theorie prüfen “Alle Raben sind schwarz”. Sie beobachten 100 Raben und stellen fest, dass jeder Rabe schwarz ist. Ist mit diesem Befund die Theorie “Alle Raben sind schwarz” logisch bewiesen? Popper sagt nein. Denn es sei nicht ausgeschlossen, dass der 101. Rabe, den man beobachtet, weiß ist. Eine endliche Zahl von Beobachtungen – und sei sie noch so groß – kann also eine allgemeine Theorie nie “beweisen”. Aber, so Popper, eine allgemeine Theorie kann durch Beobachtungen im logischen Sinne widerlegt werden. Die Aussage “Alle Raben sind schwarz” wäre widerlegt, wenn ein weißer Rabe beobachtet würde.

Dieser, wie Lakatos ihn nannte, “naive” Falsifikationismus Poppers ist jedoch nicht haltbar. Beobachtungen und Experimente sind, worauf auch Popper schon hinwies, immer “Theorie-impregniert”. Daraus folgt, dass auch Beobachtungen falsch sein können und eine Theorie fälschlicherweise widerlegt werden kann.³ Es gibt also keine *logische* Gewissheit, dass eine Beobachtung wahr und valide ist. Daher kann es auch keine *logische* Widerlegung einer Theorie durch Beobachtungen geben.⁴

Die Unhaltbarkeit von Poppers Falsifikationsprinzip lässt sich auch am Beispiel des weißen Raben zeigen: Der weiß erscheinende Rabe, der die Aussage “Alle Raben sind schwarz” angeblich widerlegt, könnte nur weiß eingefärbt sein oder aufgrund besonderer Lichtverhältnisse weiß erscheinen. Erst wenn diese und alle weiteren Einwände gegen die Feststellung, der Rabe sei weiß, widerlegt wären, könnte der Beobachtungssatz “Der Rabe ist weiß” als Widerlegung der Theorie “Alle Raben sind schwarz” gelten. Da nach Popper aber niemals *alle* Einwände widerlegt werden können – irgendwann könnte ein Einwand aufkommen, der nicht zu widerlegen ist –, muss die Beobachtung im Sinne des Falsifikationsprinzips als unbewiesen und die Theorie als unwiderlegt gelten.

Dieses Paradoxon weist darauf hin, dass die Kluft zwischen Beobachtungen auf der einen Seite und Theorien auf der anderen nicht, wie Popper meinte, durch einen einfachen logischen Schluss überbrückt werden kann. Die Logik der Forschung ist, wie u.a. die Ausführungen von Bunge, Heisenberg, Lakatos oder Margenau zeigen, weit komplexer.⁵

4.

Der Poppersche Falsifikationismus ist gescheitert, aber nicht die Idee des empirischen Gehalts. Sie hat unser Augenmerk auf eine für den wissenschaftlichen Fortschritt entscheidende Frage gelenkt: Wie kann man Hypothesen gehaltvoller machen, ohne dass man auf sinnleere Hypothesen wie die Null-Hypothese zurückgreift?

Die wichtigste Bedingung für eine gehaltvolle Hypothese ist, dass die ihr zugrundeliegende Theorie so präzisiert wird, dass wenigstens die Richtung der vorhergesagten Wirkung daraus abgeleitet werden kann. Bei unserem Beispiel könnte man die Soziale Lerntheorie etwa so präzisieren: “Internale Kontrollkognition ist mit politischer Beteiligung *positive* korreliert.”⁶ Aus dieser Präzisierung folgt die Hypothese: “Internale Personen haben im Mittel eine höhere politische Beteiligungsrate als externale Personen”. Gerichtete Hypothesen dieser Art haben die Form $H_2: \mu_i \geq \mu_e$. Die *a-priori*-Wahrscheinlichkeit, dass H_2 zutrifft, ist $p(H_2)=0.5$; der Gehalt dieser Hypothese ist $Ct(H_2)=1-p=0.5$. Dies ist zwar kein großer Wert, aber er ist nicht gleich null, wie bei der klassischen Alternativ-Hypothese. Es kann durchaus interessant sein, diese Hypothese empirisch zu prüfen, da sie immerhin mit einer Wahrscheinlichkeit von 50% falsch sein kann [siehe Formel (3)].

Um Hypothesen noch gehaltvoller zu machen, müßte man weitere Spezifikationen vornehmen. Dazu zwei weitere Beispiele.

1. Beispiel: Wenn sich neben internal und external wahrnehmenden Personen auch noch

“Neutrale” finden lassen, also Personen, die weder internal noch external wahrnehmen, dann könnte man die Vorhersage machen, dass Internale höhere Partizipationschancen haben als Neutrale, und diese eine höhere als Externale. Die Hypothese hat dann die Form: $H_3: \mu_i \geq \mu_n \geq \mu_e$. Die Wahrscheinlichkeit dieser Hypothese $p(\mu_i \geq \mu_n \geq \mu_e)$ berechnet sich nach den Regeln der Wahrscheinlichkeitsrechnung als Produkt zweier Wahrscheinlichkeiten: $p(\mu_i \geq \mu_n) * p(\mu_n \geq \mu_e) = 0.5 * 0.5 = 0.25$; der empirische Gehalt dieser Hypothese ist demnach $Ct(H_3) = 1 - 0.25 = 0.75$ [s. Formel (3)].⁷

2. Beispiel: Wenn sich Personen auf einer kontinuierlichen Skala von internaler zu externaler Kontrollwahrnehmung anordnen ließen, dann könnte die Vorhersage noch präziser gefasst werden. Man kann dann die Hypothese aufstellen: “Internale Kontrollwahrnehmung und politische Partizipation korrelieren linear miteinander”. In diesem Fall liegt eine “Korrelations-Hypothese” der Form $H_4: r_{xy} = 1.0$ vor. Die Verifikationswahrscheinlichkeit dieser Hypothese ist nahezu Null, da ja erwartet wird, dass alle Wertepaare der Wahrnehmungs- und der Partizipationskala auf einer geraden Linie liegen sollen und nicht daneben. Der Gehalt solcher Hypothesen ist demnach nahezu 1, also maximal.

Aber eine solche Korrelations-Hypothese ist durch die Sozial Lerntheorie nicht impliziert und würde vermutlich auch nicht impliziert werden, wenn die Theorie weiter entwickelt wäre. Alle Untersuchungen auf diesem Gebiet deuten darauf hin, dass die Beziehung zwischen Kontrollwahrnehmungen und sozialem Verhalten komplexer ist (Gore & Rotter 1972; Nielsen 1979). So nimmt zum Beispiel Herring (1989) an, dass externale Personen nicht generell politisch apathischer sind als internale, sondern dass sie mehr unkonventionelle Formen der politischen Partizipation wählen. Damit wäre ein Zusammenhang zu erwarten, der mit dem linearen Korrelationskoeffizient nicht richtig abgebildet werden kann. Ein niedriger Korrelationswert würde in diesem Fall nicht eine große Abweichung der Beobachtungen von der Hypothese anzeigen, sondern die Inadäquatheit des Maßes selbst. Wir sehen also: Die Erhöhung des Informationsgehalts einer Hypothese ist nur *ein* Kriterium; ebenso wichtig ist die Adäquatheit des statistischen Verfahrens, mit der die empirische Bewertung der Hypothese durchgeführt wird.⁸

5.

Wottawa hat mit seiner Kritik an der psychologischen Methodenlehre teilweise recht. Die üblichen Hypothesen sind, wie im Fall der Alternativ-Hypothesen, inhaltlos oder, wie im Fall der Null-Hypothesen, ohne psychologische Bedeutung. Wenn sich die Methodenlehre

in diesem Punkt nicht ändert, sind tatsächlich keine Erkenntnisfortschritte zu erwarten.

Aber anders als Wottawa sehe ich durchaus eine Chance für “rationale Gültigkeitsprüfungen” auf dem Gebiet der psychologischen Methodenlehre und auch Chancen für einen Fortschritt, und zwar aufgrund des Popperschen Konzepts des empirischen Gehalts von Hypothesen. Dieses Konzept hilft uns, rational zu entscheiden, welche Hypothesen gehaltvoll sind und eine aufwendige Überprüfung lohnen und welche nicht. Dabei reicht die Erhöhung des Gehalts von Hypothesen allein natürlich nicht aus, um Erkenntnisfortschritte zu erzielen. Zum einen muss immer auch geprüft werden, wie gut Hypothesen mit empirischen Daten übereinstimmen. Hierfür werden adäquate Maße für den Vorhersageerfolg einer Hypothese benötigt, wie sie etwa durch die PRE-Maße (Nagl 1992) bereitgestellt werden. Die übliche Statistik ist dafür nur bedingt geeignet (Hildebrand et al. 1977). Zum anderen muss sich die Psychologie stärker als bisher der Entwicklung solcher Theorien zuwenden, aus denen gehaltvolle Vorhersagen abgeleitet werden können⁹. Zahlreiche solcher Theorien liegen zumindest in Rohform bereits zahlreich vor, und zwar sowohl in den Lehrbüchern der Psychologie wie auch im psychologischen Alltagswissen. Das Poppersche Kriterium des Gehalts kann uns, so wollte ich zeigen, dabei helfen zu entscheiden, welche dieser Theorien in gehaltvolle Forschungshypothesen übersetzt und empirisch geprüft werden können, und welche noch einer Präzisierung bedürfen.

Abb. 1 Wahrscheinlichkeit = VE / ME

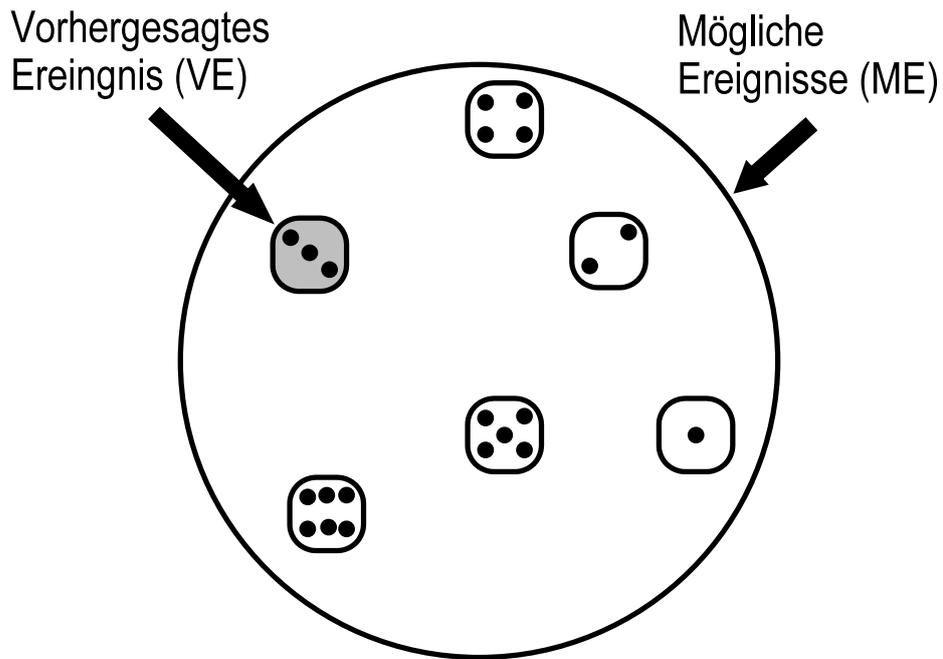


Abb. 2 Würfel-Hypothese

$$p(a) = .17 \quad Ct(a) = 1 - p = .83$$
$$p(\bar{a}) = .83 \quad Ct(\bar{a}) = 1 - p = .17$$

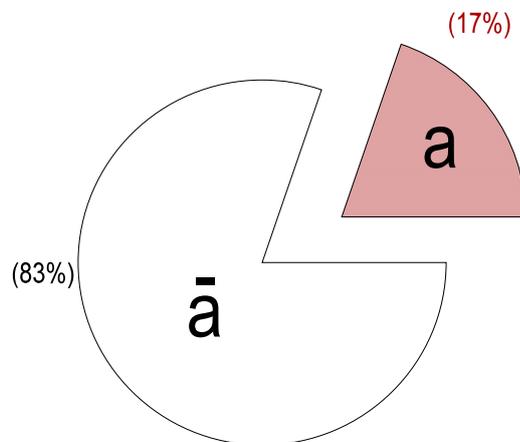
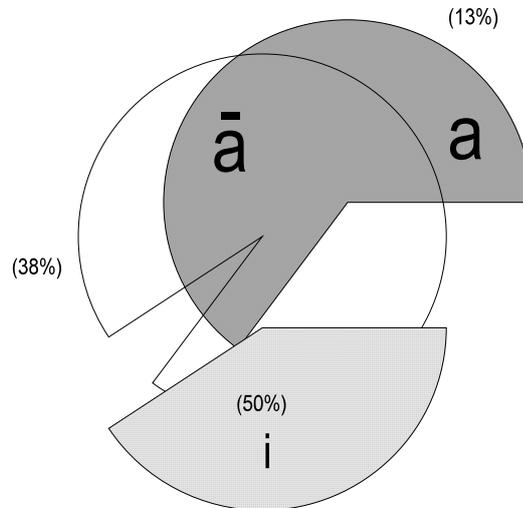


Abb. 3 Konfirmierende, falsifizierende und irrelevante Zustände

$$p(a) = .17; \quad Ct(a) = 1 - p(a) = .83$$
$$p(a,i) = p(a) / [1 - p(i)] = .34; \quad Ct(a,i) = .66$$



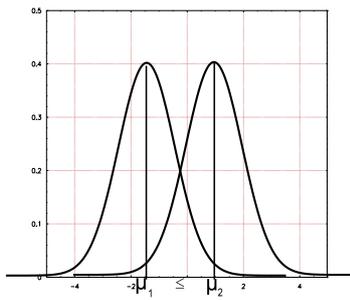
(1) Der Gehalt (content) von Hypothese a :

$$Ct(H) = 1 - p(H) = \frac{5}{6} = 0.83$$

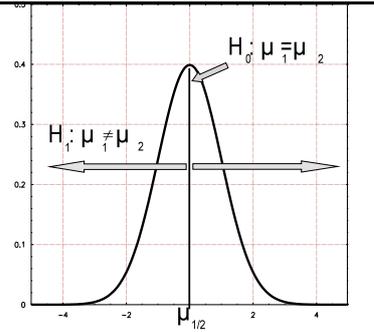
(2) Wahrscheinlichkeit: Mögliche (ME) und vorhergesagte (VE) Ereignisse

$$p(H) = \frac{VE}{ME} = \frac{1}{6} = 0.17$$

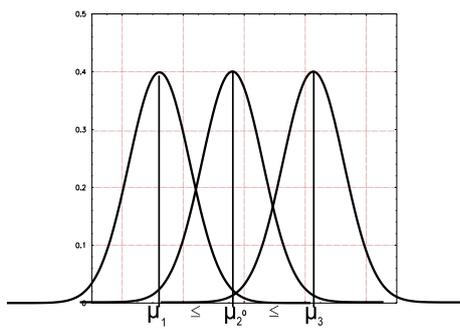
H_2



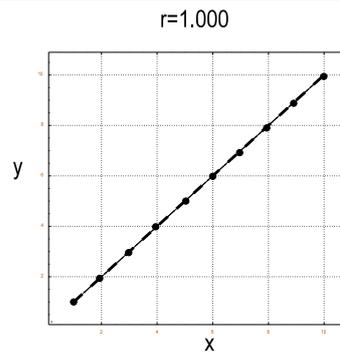
H_0 und H_1



H_3



H_4



(3) Der Gehalt von verschiedenen Hypothesen

$$H_0: my_i = my_e \quad Ct(H_0) = 1.0$$

$$H_1: my_i \neq my_e \quad Ct(H_1) \approx 0$$

$$H_2: my_i \geq my_e \quad Ct(H_2) = 0.5$$

$$H_3: my_i \geq my_n \geq my_e \quad Ct(H_3) = 1 - 0.5 * 0.5 = 0.75$$

$$H_4: r_{xy} = 1.0 \quad Ct(H_4) = 1.0$$

Anmerkungen

1. Rotter und seine Schüler nehmen einmal an, dass das Gefühl, *external* determiniert zu sein, ein andermal, dass das Gefühl, *internal* determiniert zu sein, notwendig für politische Beteiligung ist. Nach Gore & Rotter (1963) ist die politische Partizipation umso höher, desto mehr *internal* die Kontrollkognition ist: “Those who were more inclined to see themselves as the determiners of their own fate tended to commit themselves to more personal and decisive social action” (S. 62; zitiert nach 1972, S. 370).
Dagegen meint Rotter (1971), dass politische Partizipation umso höher ist, desto mehr *external* die Wahrnehmung einer Person ist: “Some social scientists believe that the impetus behind campus unrest is youth's impatient conviction that they can control their own destinies. My research over the past 12 years has led me to suspect that much of the protest, outcry and agitation occurs of the opposite reason - because students feel they cannot change the world.” (S. 37; Hervorhebung im Original).
2. Ein solcher “Trick” wird in Lehrbüchern zumindest impliziert:
E. Leiser (1980): “In der statistischen Testlogik wird immer nur gegen Hypothesen entschieden. Um für einen Hypothese – auch Alternativhypothesen – zu entscheiden, muss die gegenteilige Hypothese – auch Nullhypothese – als unplausibel verworfen werden. [...] An die Stelle der eigentlichen Hypothese, über deren Annahme/Ablehnung entschieden werden soll, ist formal deren Negation, die Nullhypothese, getreten. Erst über die Ablehnung der Nullhypothese ist die Annahme der Alternativhypothese möglich.” (S. 154).
J. Bortz (1984): “Aus der 'Unvereinbarkeit' von Nullhypothese und empirischem Ergebnis wird dann üblicherweise die Richtigkeit der Alternativhypothese erschlossen [...]” (S. 487)
3. Vgl. Gadenne (1976): “Obwohl Popper der Möglichkeit von Scheinfalsifikationen nicht sehr viel Aufmerksamkeit gewidmet hat, ist dieses Problem in der Falsifikationstheorie insofern enthalten, als sie zulässt, einen Basissatz, der mit einer Theorie in Widerspruch steht, selbst zu verwerfen” (S. 69). Trotz der Unhaltbarkeit des naiven Falsifikationismus ist er nach wie vor – auch in Lehrbüchern zur psychologischen Methodenlehre – populär. M. Lewin (1986): “Der Wissenschaftstheoretiker Sir Karl Popper hat [...] demonstriert, dass ein allgemeiner Satz wie 'Alle Schwäne sind weiß' durch die Auffindung eines einzigen nichtweißen Schwanes widerlegt, d.h. falsifiziert werden kann” (S. 216).
4. Obwohl Popper das Problem von Scheinfalsifikationen gesehen hat, geht er darauf nicht weiter ein. Lakatos verweist auf eine ganze Reihe von Beispielen in der Physik, bei denen Phänomene im Lichte neuer Theorien völlig neu interpretiert werden mußten. “Erst die Theorie”, sagt Albert Einstein einmal im Gespräch mit Werner Heisenberg (1969, S. 103), “entscheidet darüber, was man beobachten kann. [...] Die Beobachtung”, setzt er hinzu, “ist ja im allgemeinen ein sehr komplizierter Prozess.”
5. Siehe Bunge 1991a, 1991b; Heisenberg 1969; Lakatos 1972; Margenau 1950.
6. Siehe Herring (1989, S. 138).
7. Bei komplexeren Hypothesen sind Ausweitungen der Formel notwendig, wie zum Beispiel bei Kausalhypothesen der Art: Wenn das Ereignis *A* beobachtet oder experimentell ausgelöst wird, dann folgt Ereignis *B*. Oder: das Ereignis *A* ist die Ursache von *B*; kurz geschrieben: $A \rightarrow B$ (lies: wenn *A* dann *B*). In diesem Fall muss die Frage nach dem empirischen Gehalt für die *Wenn*- und die *Dann*-Komponente der Hypothese separat gestellt werden (Opp 1970). Die *Wenn*-Komponente beschreibt den “Anwendungsbereich” (scope) einer Hypothese, die *Dann*-Komponente ihre “Genauigkeit”

(precision) (Hildebrand et al. 1977).

Die *Wenn*-Komponente bestimmt den Anwendungsbereich einer Hypothese: Über je mehr Fälle die Hypothese eine Aussage macht, desto größer ist der Anwendungsbereich und auch der Gehalt einer Hypothese. Ein Beispiel: Wenn für alle, blonde und dunkelhaarige Menschen vorhergesagt werden kann, dass sie zu Fleisch Rotwein trinken, dann hat diese Aussage größeren Gehalt als wenn diese Hypothese zum Beispiel auf Blonde eingeschränkt wird. Allgemein gilt: Wenn in einer Kausalhypothese mehrere Bedingungen angesprochen werden, dann erhöhen *Oder*-Verknüpfungen den Gehalt der Hypothese (Wenn dieses *oder* jenes der Fall ist, dann ...); und *Und*-Verknüpfungen verringern ihn (Wenn dieses *und* jenes der Fall ist, dann...).

Die *Dann*-Komponente definiert die Genauigkeit einer Hypothese. Wenn statt eines einzelnen Ereignisses mehrere vorhergesagt werden, dann erhöht sich der Gehalt einer Hypothese bei logischen *und*-Verknüpfungen, und verringert sich bei *oder*-Verknüpfungen. Beispiel: In der einfachen Hypothese "Der Würfel zeigt die 3" lässt sich die Wahrscheinlichkeit ($p=.17$) einfach aus dem Verhältnis von möglichen und vermuteten Ereignissen berechnen. Bei der komplexeren Vorhersagen einer Ereignisklasse "Der Würfel zeigt die 3

8. Maße für die Güte einer Vorhersage müssen, wie Hildebrand et al. (1977) feststellen, "maßgeschneidert" für die zu untersuchende Hypothese sein. Nur wenn die Maße genau das beschreiben, was - im Sinne der zu prüfenden psychologischen Theorie - beschrieben werden soll, können sie als Kriterium für die empirische Bewertung der Hypothese gelten. PRE-Maße sind solche Maße. "PRE" steht dabei für den *Proportionate Reduction in Error* - die Reduktion des Vorhersagefehlers - durch die Formulierung einer Hypothese. PRE-Maße erlauben es, die Abweichung der empirischen Daten von einer theoretisch bedeutungsvollen Prognose adäquat zu messen. Sie stellen einen direkten Bezug zwischen empirischen Daten, Hypothesen und Theorien her - eine unverzichtbare Bedingung dafür, dass empirische Forschung zum Erkenntnisfortschritt beitragen kann. Der PRE-Ansatz wird bei Hildebrand et al. (1977) und bei Nagl (1991) ausführlich dargestellt.
9. Statt dessen dominieren in unserem Fach noch weithin "explorative Studien", bei denen eine Theorie gefunden werden soll. Explorative Studien haben, wie Popper sagt, eine entscheidende Schwäche. Sie beruhen auf dem Glauben, dass Theorien logische oder statistische Generalisierungen von Beobachtungen darstellen. Da aber, wie Popper zeigt, Beobachtungen immer schon eine Theorie voraussetzen, können mit Explorationsstudien keine neuen Theorien gefunden, sondern nur alte, implizite Theorien bekräftigt werden.
Empirische Generalisierungen beruhen, so Popper, nicht auf logisch-mathematischen Deduktionen, sondern sind immer theoriegeladen oder subjektiv. Da Beobachtungen, Messungen und dergleichen immer "theoriegetränkt", das heißt immer nur im-Lichte-einer-Theorie-interpretierte Wahrnehmungen sind, setzen sie immer schon eine Theorie voraus, auch wenn sich der Forscher dessen nicht bewusst ist (Popper 1968/1934, S. 27-30; 1979/1972, S. 7).
Ähnlich stellt Franz von Kutschera (1972) fest: "Hypothesen ergeben sich nicht aus dem Erfahrungswissen, aus Beobachtungen, sondern sie stellen Fragen an die Natur dar, die erst eigentlich systematische Erfahrung ermöglichen: Beobachtungen ohne Hypothesen sind blind, erst eine Hypothese ermöglicht sinnvoll organisierte Experimente" (S. 456).

Literatur**

- Agassi, J. (1975). *Science in Flux*. Dordrecht.
- Bortz, J. (1984). *Lehrbuch der empirischen Forschung*. Heidelberg: Springer.
- Bredenkamp, J. (1972). *Der Signifikanztest in der psychologischen Forschung*. Frankfurt: Akademische Verlagsanstalt.
- Bredenkamp, J. (1980). *Theorie und Planung psychologischer Experimente*. Darmstadt: Steinkopff.
- Bunge, M. (1991a). What is science? Does it matter to distinguish it from pseudoscience? a reply to my commentators. *New Ideas in Psychology* 9, 245-283.
- Bunge, M. (1991b). A sceptic's beliefs and disbeliefs. *New Ideas in Psychology* 9, 131-149.
- Cohen, J. (1977). *Statistical Power Analysis*. (2nd edition) New York: Academic Press.
- Gadenne, V. (1976). *Die Gültigkeit psychologischer Untersuchungen*. Stuttgart: Kohlhammer.
- Gadenne, V. (1984). *Theorie und Erfahrung in der psychologischen Forschung*. Tübingen: Mohr.
- Gadenne, V. (1990). Methoden als Hilfsmittel für heuristische Forschungsentscheidungen - Bemerkungen zu Wottawas Abhandlung der psychologischen Methodenlehre. *Psychologische Rundschau* 41, 98-100.
- Gore, P.M. & Rotter, J.B. (1972). Personality correlates of social action. In J.B. Rotter, J.E. Chance & E.J. Phares, Hrg., *Applications of a Social Learning Theory of Personality*. New York: Holt, Rinehart and Winston, S. 366-370.
- Heisenberg, W. (1969). *Der Teil und das Ganze. Gespräche im Umkreis der Atomphysik*. München: Piper.
- Herring, C. (1989). Acquiescence or activism? Political behavior among the politically alienated. *Political Psychology* 10, 135-153.
- Hildebrand, D.K., Laing, J.D. & Rosenthal, H. (1977). *Prediction Analysis of Crossclassifications*. New York: Wiley.
- Keuth, H. (1976). Verisimilitude or the approach to the whole truth. *Philosophy of Science* 43, 311-336.
- Kutschera, F.v. (1972). *Wissenschaftstheorie II. Grundzüge der allgemeinen Methodologie der empirischen Wissenschaften*. München: W. Fink Verlag.
- Lakatos, I. (1972). Falsification and the methodology of scientific research programmes. I. Lakatos & A. Musgrave, Hrg., *Criticism and the Growth of Knowledge*. Cambridge University Press, S. 91-196.
- Leiser, E. (1980). *Einführung in die statistische Methoden der Erkenntnisgewinnung*. Köln: Pahl-Rugenstein.
- Lewin, M. (1986). *Psychologische Forschung im Umriß*. Heidelberg: Springer.
- Margenau, H. (1950). *The Nature of Physical Reality*. New York: McGraw-Hill.
- Meehl, P.E. (1967). Theory-testing in psychology and physics - a methodological paradox. *Philosophy of Science* 34, 103-115.
- Meehl, P.E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft Psychology. *Journal of Consulting and Clinical Psychology* 46, 806-834.
- Nagl, W. (1992). *Statistische Analysen mit SAS*. Wiesbaden: Campus (im Druck).
- Nielsen, A. (1979). *Hinderungsgründe für politisches Engagement bei Studenten*. Diplomarbeit im Fach Psychologie, Universität Konstanz.
- Opp, K.-D. (1970). *Methodologie der Sozialwissenschaften*. Reinbek b. Hamburg: Rowohlt.
- Popper, K.R. (1968a). *The Logic of Scientific Discovery*. London: Hutchinson (Original 1934).
- Popper, K.R. (1968b). *Conjectures and Refutations*. New York: Harper & Row (Original 1962).
- Popper, K.R. (1979). *Objective Knowledge. An Evolutionary Approach*. Oxford: Clarendon Press (Original 1972).
- Rotter, J.B. (1971). External control and internal control. *Psychology today* 5, 37-62.
- Westmeyer, H. (1987). Zum Problem des empirischen Gehalts psychologischer Theorien - Eine Analyse am

** Nur auf einen Teil der hier verwendeten Literatur wird im Text explizit hingewiesen.

- Beispiel der Theorie der Verhaltensinteraktion in Dyaden. In J. Brandstädter, Hrg., *Struktur und Erfahrung in der psychologischen Forschung*. Berlin: deGruyter, S. 35-70.
- Wottawa, H. (1990). Einige Überlegungen zu (Fehl-)Entwicklungen der psychologischen Methodenlehre. *Psychologische Rundschau* 41, 84-107.