

Weder die Autoren/innen, noch die Fachschaft Psychologie übernimmt irgendwelche Verantwortung für dieses Skript.

Das Skript soll nicht die Lektüre der Prüfungsliteratur ersetzen.
Verbesserungen und Korrekturen bitte an fs-psycho@uni-koeln.de mailen.

Die Fachschaft dankt den AutorInnen im Namen aller Studierenden!

Version 1.0 (03.04.2000)

Wissenschaftliche Hypothesen (Hussy/Möller)

WH sollten in Abgrenzung zu Alltagshypothesen mindestens zwei Merkmale erfüllen:

- sie sollen im Sinne einer synthetischen Aussage formuliert sein (\Rightarrow müssen sich als wahr oder falsch erweisen können)
- müssen überprüfbar sein.

WH sind theoretische Annahmen bzw. Vermutungen über Zusammenhänge von interessierenden Sachverhalten oder auch vorläufige (provisorische) Antworten für ein sich stellendes wissenschaftliches Problem. Vorläufiger Charakter wird ihnen unterstellt, da sie im Rahmen der wissenschaftlichen Forschung sowohl verworfen als auch beibehalten bzw. weiterentwickelt werden können. Von ihren Konzepten wird größtenteils angenommen, dass sie über die direkt beobachtbaren Sachverhalte hinausgehen und im Falle des Beibehaltens einer WH als Erklärung dienen können.

Arten von Hypothesen

singuläre Hypothese	Beschränkung in Personenkreis, der Situation und der Aktivität
idiographische Hypothese	keine Einschränkung in Raum und/oder Zeit
unbestimmte Existenzhypothese	enthält mindestens 1 Existenzquantor ohne weitere Einschränkung
bestimmte Existenzhypothese	Aussagen beschränkt auf bestimmte Räume und/oder Zeiten
quasi-universelle Hypothese	ähnlich wie echt universelle, mit Einschränkungen, die die generelle Aussage vage werden lassen
beschränkt-universelle Hypothese	alle Fälle einer Population von Individuen mit Beschränkung in Raum und Zeit
unbeschränkt universelle Hypothese	alle Fälle einer bestimmten Art ohne Einschränkung
statistische Hypothese	Hypothesen über die Verteilung von Werten

Weitere Merkmale bzw. Unterscheidungskriterien, die im Rahmen von Untersuchungen relevant sind: statistische vs. inhaltliche Hypothesen. Statistische Hypothesen sind auf Daten bezogen, inhaltliche Hypothesen auf konkret beobachtbare Sachverhalte oder auf theoretische Konstrukte.

Entstehung

Wissenschaft läßt sich als Problemlösevorgang verstehen. Hypothesen sind vorläufige (probabilistische) Antworten auf Fragen, die untersucht werden sollen. In diesem Problemlösevorgang kann man zwischen induktivem und deduktivem Vorgehen unterscheiden; das ist jeweils vom Stand der Forschung abhängig. Induktive Hypothesen tragen zur Schaffung einer soliden empirischen Datenbasis bei und bilden so die Grundlage zur Entwicklung von Theorien. Deduktive Hypothesen können zum Erkenntnisfortschritt auf Theorieebene führen.

(Beispielsweise kann der Geltungsgrad einer Theorie erweitert bzw. differenziert werden. Auch Hypothesen, die sich als „falsch“ erweisen, tragen jedoch zur Erkenntnis bei; sie lehren vielleicht sogar mehr als eine „bewiesene“ Hypothese.)

Bedeutung

Hypothesen stellen die „theoretische Brille“ eines Wissenschaftlers dar. Auf ihnen gründet sich seine Entscheidung, an einer bestimmten Stichprobe ein spezifisches treatment zu applizieren und genau definierte Maße zu erheben. Eine Hypothese basiert immer auf mehr oder weniger klaren Vorstellungen und Erwartungen über Zusammenhänge der interessierenden Phänomene, die einerseits Vermutungen sind, andererseits aber das Hintergrundwissen repräsentieren, das bislang in der Wissenschaft besteht und zur Untersuchung herangezogen wurde.

Hypothesen sind somit als Bindeglied zwischen Theorien und empirischen Daten zu verstehen. (Stichwort: Präzisierung der Fragestellung, Rückschluss auf Ausgangstheorie).

Logik der Prüfung

Wie läßt eine WH sich in der Untersuchung anwenden? Kann man ihr „richtig“ oder „falsch“ attestieren; und wenn ja, wie?

Vorbedingung der Prüfung

1. Standard bilden: Die Hypothese sollte im Sinne einer „je-desto“- oder „wenn-dann“-Aussage formuliert sein; das dokumentiert ihren vorläufigen Charakter und differenziert (auf empirischer Ebene) zwischen AV und UV (Kausalhypothese) oder Prädiktor und Kriterium (Zusammenhangshypothese).
2. Hypothesen müssen widerspruchsfrei formuliert werden.
3. Sie müssen widerlegbar sein. Dabei gilt: (1) je höher die Zahl der Falsifikatoren, desto höher ist der empirische Gehalt einer Hypothese; (2) die in ihr enthaltenen Begriffe müssen operationalisierbar sein.

Überprüfungsmöglichkeiten von Hypothesen grundsätzlich

unbestimmte Existenzhypothese	prinzipiell zu bestätigen (Nachweis 1 Fall)
bestimmte Existenzhypothese	zumindest prinzipiell belegbar oder widerlegbar
quasi-universelle Hypothese	durch die Vagheit ist eine strenge Prüfung nicht möglich
beschränkt-universelle Hypothese	prinzipiell untersuchbar (z. B. Meinungsumfragen)
unbeschränkt universelle Hypothese	streng genommen nur zu widerlegen, da Prozess nie abgeschlossen werden kann (kritischer Rationalismus!)

In der Psychologie wird die quasi-universellen Hypothese (Wahrscheinlichkeitshypothese) am häufigsten benutzt.

Prüfungsweg

Die TIH (theoretisch-inhaltliche Hypothese) ist in der Regel abstrakt formuliert. Die in ihr enthaltenen Begriffe sind nicht beobachtbar und müssen in eine EIH (empirisch-

inhaltliche Hypothese) überführt werden. Dort sind UV und AV auf Basis des vorliegenden Hintergrundwissens so festzulegen, dass sie einer Beobachtung, Messung und Erfassung zugänglich werden (\Rightarrow Operationalisierung). Denn:

Theoretische Konzepte entziehen sich der Beobachtung, empirische Konzepte müssen ihnen zugeordnet werden \Rightarrow Reduktion des semantischen Gehalts der theoretischen Konzepte. Das kann insofern zum Problem werden, als z. B. einzelne Forscher sehr unterschiedliche Vorstellungen darüber haben, was unter Aggression zu verstehen ist.

Zuordnungsbeziehungen zwischen den Hypothesen sollten idealerweise wissenschaftlich exakt definierbar sein; bislang existiert jedoch noch keine allgemein anerkannte Strategie.

Die Operationalisierung erfolgt unter Bezug auf Hintergrundwissen (das ungeprüft übernommen wird). Sie stellt den ersten, wichtigen Schritt im Forschungsprozess dar; hier entstandene Fehler lassen sich später oft kaum noch korrigieren.

Eine TIH kann in verschiedenen Experimenten ganz unterschiedlichen inhaltlichen Hypothesen zugeordnet werden, wodurch unterschiedliche Bewertungen möglich werden; also explizite Darstellung (auch im Forschungsbericht) erforderlich.

Auch die Relationen zwischen den Begriffen der TIH müssen auf EIH übertragen werden (z. B. „meistens“ in Form von „Primaten zeigen *m e h r* Beißverhalten als Schlangen“.)

Eine Beurteilung der EIH erfolgt anhand empirischer Daten. Das könnte nach vorher festgelegten Kriterien erfolgen, die bislang jedoch kaum vorliegen.

Empirische Daten sind im allgemeinen fehlerhaft. Es besteht eine Diskrepanz zwischen empirisch erhobenen Daten und dem „wahren“ Wert.

Die Überprüfung der EIH erfolgt meist statistisch; inhaltlichen Hypothesen werden also statistische Hypothesen zugeordnet \Rightarrow statistische Vorhersage, über die in der Regel durch Signifikanztest entschieden wird \Rightarrow prüfen Testhypothesen H_0 , H_1 (= Alternativhypothese)

Beziehung zwischen inhaltlichen und statistischen Hypothesen

Inhaltliche Hypothesen beziehen sich auf empirische Sachverhalte, auf Populationen von Individuen; statistische Vorhersagen auf statistische Konzepte. Sie beinhalten Zahlenmaterial, das so gesehen ohne inhaltliche Bedeutung ist.

Die Relationen der TIH werden auf die EIH und ebenso die SV übertragen. Gelingt dies, spricht man von einer adäquaten Zuordnung der SV zur EIH.

Eine direkte Übertragung der Variablen(inhalte) ist jedoch nahezu unmöglich, so dass die Überprüfung inhaltlicher Hypothesen durch eine Entscheidung über statistische Vorhersagen als sehr indirekt gekennzeichnet wird. Diese sehr indirekte Beziehung gab häufig Anlass zu Kritik; die Funktion der Statistik in der Psychologie wurde in Frage gestellt. (Verständlich: mathematische Verfahren haben einen anderen Gegenstand.) Es wurde u. a. gefordert, zwischen EIH und SV soll eine logische Implikationsbeziehung bestehen oder die SV soll folgerichtig aus der EIH abgeleitet sein.

Implikationsbeziehung: Die EIH ist über das *modus tollens* falsifizierbar, wenn die SV sich als falsch herausstellt. Eine Bestätigung der EIH wäre danach nicht möglich.

Die Annahme einer solchen Beziehung zwischen EIH und SV jedoch kaum realistisch; wenn überhaupt, wäre nur die Verbindung aus Hintergrundwissen und inhaltlicher Hypothesen (logische Konjunktion) falsifizierbar, nicht die EIH allein. Davon abgesehen kann die EIH nicht falsifiziert werden, nur als gültig oder ungültig bestimmt. Auch diese Entscheidung lässt sich aufgrund der unterschiedlichen, nicht vergleichbaren Ebenen nicht direkt auf die EIH übertragen.

Eine Implikationsbeziehung wäre möglich, wenn die psychologische Hypothese statistische Konzepte enthielte, wie etwa in der Psychophysik. Allerdings läuft man dann Gefahr, dass die psychologischen Hypothesen dann gleich als SV formuliert werden. Eine Durchsicht der Literatur zeigt jedoch, dass Forscher in der Regel nicht an statistischen Parametern interessiert sind. (Sie wissen, dass sie mit der Statistik ein Hilfsmittel benutzen, und sie sind an Aussagen über Menschen, nicht an statistischen Werten interessiert.) Dennoch bleibt der Signifikanztest für indirekte Prüfungen in Frage gestellt.

WH im Bereich der Sozialforschung sind Wahrscheinlichkeitsaussagen, die sich durch konträre Einzelfälle prinzipiell nicht widerlegen lassen. Sie machen zudem verallgemeinernde Aussagen über Populationen, die in der Regel nur ausschnitthaft durch Stichproben untersucht werden können, so dass auch eine Verifikation nicht möglich ist. Daher hat die Scientific Community Kriterien festgelegt, um WH prüfbar zu machen; eines der Kriterien ist der Signifikanztest. (Bortz/Döring, 1. Kapitel)

Den psychologischen Daten werden Verteilungsfunktionen, Lokalisationsparameter, Streuungen usf. unterstellt. Das ermöglicht die Abschätzung von Fehlerrisiken, die mit den Entscheidungen verbunden sind. Das wiederum setzt spezifisches mathematisch-statistisches Datenmodell voraus sowie die Einhaltung der Forderungen des Modells.

Die Wahrscheinlichkeiten des Zutreffens der Bewertungen (inhaltliche Hypothesen) bzw. der Entscheidungen (SV und Signifikanztest) können definiert werden. Fehlerrisiken: α und β

H_0 gilt und wurde angenommen: $1 - \alpha$

H_0 gilt und wurde abgelehnt: α

H_1 gilt und wurde angenommen: $1 - \beta$

H_1 gilt und wurde abgelehnt: β

e und f (für die TIH) sind numerisch nicht bestimmbar, hängen aber mit ε und φ (der EIH) und letztlich mit α und β zusammen. Der Signifikanztest macht eine genaue Abschätzung von α und β möglich und erlaubt dadurch eine gewisse Kontrolle von e und f:

Fehlerkontrolle von α und β

EIH: $A_1 > A_2 \Rightarrow SV: \mu_{A1} > \mu_{A2}$

$\Rightarrow SV: \mu_1 > \mu_2$

(a) nicht-signifikantes Ergebnis des t-Tests: Ablehnung von $H_1 \Rightarrow SV$ unzutreffend \Rightarrow EIH nicht bewährt \Rightarrow je größer β , desto größer φ und f.

(b) signifikantes Ergebnis des t-Tests: Annahme von $H_1 \Rightarrow SV$ zutreffend \Rightarrow EIH bewährt \Rightarrow je größer α , desto größer ε und e.

Kontrolle von α - und β -Fehler notwendig \Rightarrow Simultankontrolle der Determinanten des Signifikanztests (sog. Testplanung nach Hager).

Bewertung

Warum ist es günstiger, von Bewertung bzw. Beurteilung der inhaltlichen Hypothesen zu sprechen? Der Prüfungsprozess führt nicht zu Verifikation oder Falsifikation der inhaltlichen Hypothesen. (Auf statistischer Ebene Zahlenmaterial – Häufigkeiten – vor, das aus sich heraus gar nichts sagt. Dieses Zahlenmaterial wird unter Einbeziehung des Hintergrundwissens auf Ebene der EIH interpretiert; das sagt aber nur etwas darüber aus, ob die Annahmen des Forschers sich bewährt ha-

ben oder nicht. Von einem „falsch“ oder „richtig“ zu sprechen – so verstehe ich es – wäre nicht angemessen; damit würde ein vom Raum-Zeit-Kontinuum unabhängiges Gesetz behauptet, und soviel Sicherheit werden wir im Bereich sozialwissenschaftlicher Forschung vermutlich nie erreichen.)

Rückweg

von Testergebnissen zu EIH und TIH bzw. deren Bewertung

Unter jeder Testhypothese sind zunächst alle statistisch-empirischen Ergebnisse möglich, wenn auch unterschiedlich wahrscheinlich. Ein spezifisches empirisches Datum kann nicht gegen oder für eine statistische Hypothese entscheiden. Der Forscher entschließt sich, H_0 oder H_1 anzunehmen; das ist eine konventionalistische Entscheidung, die von α bzw. β abhängt (vgl. oben Text aus Bortz).

Sind SV und Testhypothese identisch, entscheidet der Forscher unmittelbar über beide; ist die Beziehung komplexer, gilt dies auch für die Entscheidung. Das Verfahren ist im Grundsatz jedoch relativ statisch.

Auf EIH-Ebene wird der Prozess umfassender: Hier folgt eine Beurteilung aufgrund der Daten und in Abhängigkeit von verschiedenen Gütekriterien. Eine mögliche Konvention zur Beurteilung ist die hinreichende Erfüllung der Gütekriterien und die gelungene Kontrolle von α , β , ε und φ . Dabei geht man nicht nur von reiner Signifikanz aus, sondern gewichtet statistische Effekte (Effektstärken), die vor Datenerhebung auf Ebene der SV festgesetzt werden. Gestützt auf die Testplanung läßt sich der Beurteilungsraum der EIH dann in 4 bzw. 3 Bereiche aufteilen: EIH „bewährt – bedingt bewährt – bedingt nicht bewährt – gar nicht bewährt“. (Die relevante SV wird zwar angenommen, aber die Beurteilung der inhaltlichen Hypothesen richtet sich danach, ob der statistische Effekt unter- oder überschritten wurde.) Dieses Vorgehen erlaubt, inhaltliche Überlegungen in den Signifikanztest einzubeziehen und ein Ergebnis gegebenenfalls als nur bedingt relevant zu bewerten, wenn die Effektstärke zu niedrig ausfällt, um der EIH praktische Bedeutsamkeit beizumessen.

Die Bewertung der TIH erfolgt dann nahezu ausschließlich aufgrund nicht-statistischer Überlegungen. Hier sind Fragen der Validität relevant: interner Validität und der Operationalisierung. e und f sind numerisch nicht bestimmbar; und auf dieser Ebene fließen subjektive Kriterien ein. Somit hat die Untersuchungsvalidität einen entscheidenden Anteil an der Beurteilung der inhaltlichen Hypothese.

(Der Text behandelt im Weiteren nur Gütekriterien, die sich direkt auf die Untersuchung und deren Hypothesenbewertung beziehen.) Die Gewichtung der jeweiligen Gütekriterien bemißt sich im Einzelfall.

Zuordnung der EIH zu den TIH

1. Operationalisierung der Variablen der TIH \Rightarrow Variablenvalidität, kann für AV und UV beeinträchtigt sein. Operationalisierung läßt sich nicht über Koeffizienten validieren.
2. Übertragung der Relationen von TIH nach EIH \Rightarrow Ableitungsvalidität, kann fehlerhaft sein.

Zuordnung SV zu EIH

Keine Übersetzung, nur Zuordnung möglich; die Güte der Zuordnungsbeziehung hängt u.a. vom Skalenniveau der erhobenen Daten ab. Alle relevanten Infos der EIH müssen auf statistischer Ebene wiederzufinden sein \Rightarrow suffiziente Zuordnung.

Die Relationen der Variablen der EIH müssen korrekt auf statistische Ebene übertragen

werden \Rightarrow adäquate Zuordnung. Beide Kriterien beeinträchtigen als Störungen der Ableitungsvalidität der statistischen Vorhersage.

Zuordnung Signifikanztest zu SV

Beeinträchtigungen der Ableitungsvalidität der Testhypothese, wenn Signifikanztests gewählt werden, die zur Prüfung der SV nur eingeschränkt tauglich sind.

Validität des statistischen Schlusses

Die Verwendung von Signifikanztests beruht auf spezifischen Voraussetzungen, deren Nicht-Erfüllung zur Beeinträchtigung der statistischen Validität an sich führen kann. Das geschieht meist durch unkontrollierte Veränderungen der Fehlerwahrscheinlichkeiten für α und β .

Anmerkung

Das Skript wurde am 27. März 2000 von mir erstellt. Inhaltlich folgt es der Logik des Originals. An zwei Stellen, die ich nicht bis ins kleinste Detail verstehen wollte, habe ich Kürzungen vorgenommen: Fehlerkontrolle von α und β bis hin zur Ebene der TIH (e und f) sowie Feinheiten gegen Ende. Also bitte den Text vergleichen. Meine Lernempfehlung wäre, den Originaltext zu lesen und diesen hier zu korrigieren; so übt man gleichzeitig das Gelernte ein und stellt ein auf das eigene Verständnis zugeschnittene Skript her.

Viel Spaß, und: viel Erfolg!