

Norbert Froese

Messung und Systematisierung in der Soziologie*

Abstract: This essay is a systematic investigation of the problems concerning the measurement of theoretical qualities in the social sciences. If someone wants to measure theoretical qualities like attitudes, intelligence or aggressivity the first and most difficult task is to formulate an operationalisation. To show such operationalisations being correct is the central problem of empirical research. In this essay it is shown that no adequate methods exist to demonstrate operationalisations as true or false statements. As long as the problem of operationalisation is treated and analysed as a problem of isolated statements it seems to be absolutely unsolvable. The problem of the correctness of operationalisations must be interpreted as the problem of the correctness of the theories in which they are used. Whether an operationalisation is accepted or not depends only on the empirical success of the theory in which it is contained. This implies that we can only measure theoretical qualities if we have successful theories working with these theoretical qualities.

Obwohl sich die heutige Soziologie in all ihren Varianten und Schulbildungen als empirische Wissenschaft begreift, ist es bis jetzt nicht gelungen, eine allgemein anerkannte und befriedigende Deutung des Verhältnisses von soziologischer Theoriebildung und empirischer Sozialforschung zu entwickeln. Interpretation und Bewertung empirischer Studien stellt noch immer eher eine Kunst, als eine mittels wissenschaftlicher Kriterien überprüfbare Tätigkeit dar. Eine Schlüsselstellung in diesem Dilemma kommt dem Problem der Indikatorenbildung zu. Wie kann man von intersubjektiv kontrollierbaren Merkmalen wie Beruf, Einkommen, Antwortverhalten bezüglich eines vorgelegten Fragenkatalogs oder der Reaktion in einem Test auf Merkmale wie sozialer Status, Aggressivität, politische Einstellung oder Klassenbewußtsein schließen? Diese Frage ist einer der Hauptkristallisationspunkte für Schulbildung in den Sozialwissenschaften. Besonders deutlich wird dies beim Problem der Messung mentaler Eigenschaften.

Auf der einen Seite stehen die Positionen der hermeneutischen Tradition und die radikalen Varianten der Ethnomethodologie. Beide bestreiten, daß es möglich ist, von intersubjektiv kontrollierbaren Merkmalen (Indikatoren) auf mentale Eigenschaften zu schließen. Während jedoch in der hermeneutischen Tradition daraus die Konsequenz gezogen wird, daß in den Sozialwissenschaften neben dem intersubjektiv kontrollierbaren Protokollieren auch verstehendes Einfühlen systematisch unver-

* Die grundlegenden Ideen der hier vorgetragenen Argumentation wurden im DFG-Projekt „Theoretische Begriffe in den Sozialwissenschaften“ entwickelt. An diesem Projekt haben mitgearbeitet: Nils Heicke, Wolfgang R. Köhler, Dieter Mans und Heidrun Rheinfels-Immanns.

zichtbar ist, wird in der radikalen Ethnomethodologie hieraus die Schlußfolgerung gewonnen, daß mentale Begriffe höchstens Gegenstand aber nicht Mittel sozialwissenschaftlicher Untersuchungen sein können. Nicht „Wer ist klassenbewußt?“, sondern „Wer wird klassenbewußt genannt?“, ist hiernach die wissenschaftlich einschlägige Fragestellung.

Auf der anderen Seite stehen die Konzepte der Indikatorenbildung. Sie behaupten, daß es (unter geeigneten Umständen) möglich ist, von intersubjektiv kontrollierbaren Merkmalen auf mentale Eigenschaften zu schließen. Es sind zwei Varianten der Indikatorenbildung zu unterscheiden. Die erste Variante besteht in dem Versuch, mentale Prädikate (oder auch andere indikatorbedürftige Begriffe) durch geeignete (partielle) Definitionen oder Bedeutungspostulate (Reduktionsätze) mit als unproblematisch konstatierbar geltenden Merkmalen zu verknüpfen, so daß bereits diese Bedeutungsregeln Messungen der so eingeführten Größen gestatten. In der zweiten Variante wird die Formulierung einschlägiger Indikatorrelationen nicht als Bedeutungsfestlegung, sondern als das Aufstellen empirisch gehaltvoller Hypothesen gedeutet.

Die hermeneutische und die radikale ethnomethodologische Position beanspruchen für den Bereich der mentalen Prädikate jeweils einen Sonderstatus. Dies ist nur dann gerechtfertigt, wenn es sich tatsächlich als unmöglich erweist, solche Prädikate sinnvoll zu operationalisieren. Hier soll dafür argumentiert werden, daß eine derartige Möglichkeit der sinnvollen Operationalisierung auch für mentale Prädikate besteht. Es soll nicht bestritten werden, daß viele der verwendeten Operationalisierungen unangemessen sind. Es kann auch in vielen Fällen forschungspraktisch nützlicher sein, sich auf die hermeneutische Kompetenz der untersuchenden Sozialwissenschaftler zu verlassen, als sich an der Formulierung einer angemessenen Operationalisierung der einschlägigen Begriffe zu versuchen. Ebenso ist natürlich die Untersuchung der alltäglichen Verwendungsweisen mentaler Begriffe ein relevantes Thema für die Sozialwissenschaften. Ethnomethodologische wie hermeneutische Positionen werden von den folgenden Ausführungen nur insofern betroffen, als sie Behauptungen bezgl. der prinzipiellen Unmöglichkeit der Messung mentaler Eigenschaften aufstellen.

Wenn man sich mit der Frage beschäftigt, wie das Problem der Indikatorenbildung in den Sozialwissenschaften einer befriedigenden Lösung zugeführt werden kann, so muß man sich als erstes damit beschäftigen, was es heißen soll, wenn man eine Indikatorenbildung als korrekt bzw. eine auf deren Basis vorgenommene Messung als valide bezeichnet. Bei der Untersuchung des Verhältnisses von soziologischer Theoriebildung und empirischer Sozialforschung fällt der genauen begrifflichen Analyse des Validitätsproblems eine Schlüsselrolle zu. Die bisherigen Ansätze zur Lösung des Validitätsproblems gehen von einer inadäquaten Interpretation des Validitätsproblems aus. Als Folge davon verstricken sich alle diese Ansätze in Schwierigkeiten, die eine Lösung des Validitätsproblems als unmöglich erscheinen lassen. Da aber die Behauptung, daß eine Indikatorenbildung korrekt ist, nichts anderes bedeutet, als daß die auf ihrer Basis gewonnenen Resultate valide sind, schlägt sich die Fehlinterpretation des Validitätsproblems notwendig in einer inadäquaten

Interpretation des Problems der Indikatorenbildung nieder. Aus diesen Gründen wird der Untersuchung der Natur des Validitätsproblems in den folgenden Ausführungen die zentrale Rolle zufallen.

I. Die ontologische Interpretation des Validitätsproblems und ihre Unangemessenheit für theoretische Begriffe

Die Interpretation des Validitätsproblems, die den meisten Erörterungen in methodologischen Texten wie aber auch bei konkreten Validitätsdebatten bezüglich bestimmter Tests und Operationalisierungen zu Grunde liegt, läßt sich so rekonstruieren:

Gemessen werden soll ein Merkmal. Um uns nicht in technische Einzelheiten zu verlieren, unterstellen wir der Einfachheit halber, daß es sich hierbei um ein dichotomes Merkmal handelt. Das Messen des Merkmals ist dann damit identisch, festzustellen, welchen Objekten korrekterweise ein einstelliges Prädikat zugesprochen werden kann. Dabei wird unterstellt, daß das Prädikat eine Eigenschaft bezeichnet, mittels der die Menge der zu messenden Objekte in zwei Teilklassen zerlegt werden kann. Die eine Teilklassse wird von jenen Objekten gebildet, die die Eigenschaft aufweisen, sie bilden die Extension des Prädikates; die andere Teilklassse wird von jenen Objekten gebildet, die die Eigenschaft nicht aufweisen. Das Validitätsproblem besteht dann darin, sicherzustellen, daß die von einer bestimmten Operationalisierung erzeugten Resultate das Prädikat genau dann zusprechen, wenn das entsprechende Objekt tatsächlich zur Extension des Prädikats gehört. Bei der Validierung von Daten geht es also darum, zu gewährleisten, daß die vorgenommenen Zuschreibungen in Übereinstimmung mit der unabhängig von diesen Zuschreibungen festgelegten Extension des Prädikats geschehen. Das Problem des Auffindens einer korrekten Operationalisierung wird in dieser Interpretation des Validitätsproblems nicht als ein Problem der Formulierung sinnvoller Präzisierungen eines Begriffs aufgefaßt, sondern die zu messenden Eigenschaften werden als vorgegeben betrachtet. Es geht darum, eine bereits unabhängig von den Operationalisierungen und Zuschreibungen des Prädikats existierende Zerlegung der zu messenden Objekte in jene, die die Eigenschaft aufweisen und in jene, die sie nicht besitzen, durch Anwendung einer geeigneten Operationalisierung bei der Datenerhebung korrekt wiederzugeben.

Diese Sicht kann man die ontologische Interpretation des Validitätsproblems nennen. Die Bezeichnung „ontologisch“ ist motiviert durch die bei dieser Interpretation benutzte Unterstellung einer unabhängig von Messung und Operationalisierung existierenden Zerlegung der zu messenden Objekte in jene, die die Eigenschaft besitzen, und in jene, denen sie nicht zukommt.

Wir wollen im folgenden untersuchen, welche Voraussetzungen erfüllt sein müssen, damit man ein solch ontologisch interpretiertes Validitätsproblem überhaupt lösen kann. Bevor wir in diese Diskussion eintreten, muß allerdings noch der Begriff „Operationalisierung“ erläutert werden. Unter der *Operationalisierung* eines Begriffs wird hier eine *nicht* auf Grund von expliziten Bedeutungsfestlegungen gel-

tende Zuschreibungsregel für den Begriff verstanden, die ihn mit unproblematisch konstatierbaren Indikatoren verknüpft. Diese Definition von „Operationalisierung“ weicht von der in der Wissenschaftstheorie üblichen Verwendung ab. Dort wird unter Operationalisierung die *Definition* eines Begriffs durch Meßverfahren verstanden. Obwohl auch in der methodologischen Literatur häufig von operationaler Definition gesprochen wird, werden dort jedoch Operationalisierungen meist nicht als Definitionen betrachtet. Dies wird daran deutlich, daß Operationalisierungen als korrekt oder inkorrekt gelten können, Definitionen aber höchstens nützlich oder nutzlos sind.

Wenn man nun überprüfen will, ob eine bestimmte Operationalisierung korrekt ist und damit die von ihr erzeugten Resultate valide sind, so muß man in der Lage sein, die auf der Basis der Operationalisierung vorgenommenen Zuschreibungen eines Prädikats mit seiner tatsächlichen Extension zu vergleichen. Man muß also zur Frage, ob einem Objekt ein Begriff korrekt zugesprochen werden kann, einen von der zur Debatte stehenden Operationalisierung unabhängigen Zugang haben. Nur unter dieser Voraussetzung kann im Sinne eines ontologisch interpretierten Validitätsproblems die Korrektheit einer Operationalisierung gesichert werden. Daß die Frage, ob eine Operationalisierung korrekt ist, auch beantwortbar ist, kann vor diesem Hintergrund also keineswegs als trivial vorausgesetzt werden. Entwickeln wir diese Argumente genauer. Es gibt zwei verschiedene Motive, die hinter der Konstruktion von Meßverfahren stehen können. Auf der einen Seite kann es uns darauf ankommen, neben den bereits zur Verfügung stehenden verlässlichen Methoden der Zuschreibung eines Begriffs, weitere und insbesondere praktisch interessantere Meßverfahren zu besitzen. Haben wir es z. B. mit einer Disposition des Typs: „x wird unter den Bedingungen A in der Weise B reagieren“ zu tun, so kann man, vorausgesetzt die Bedingungen A und eine Reaktion in der Weise B sind empirisch identifizierbare Sachverhalte, über das Vorliegen der Disposition unproblematisch entscheiden, solange die Bedingungen A erfüllt sind. Nun sind wir aber häufig gerade an Meßverfahren interessiert, die unabhängig von dem Vorliegen der die Disposition charakterisierenden Testbedingungen und Reaktionen funktionieren. So kann die Herstellung der Testbedingungen außerhalb der praktischen Möglichkeiten eines Wissenschaftlers liegen, oder der praktische Sinn der Identifikation solcher Dispositionen kann gerade darin bestehen, das Auftreten der charakteristischen Reaktion zu vermeiden. Beispiele für derartige Dispositionsprädikate sind: „x hat die Disposition unter den Bedingungen A Selbstmord zu verüben“, „x würde, wenn es kein Fernsehprogramm gäbe, häufiger ins Kino gehen“.

Die Probleme bei der Formulierung geeigneter Operationalisierungen und der Sicherung der Validität der von ihnen produzierten Daten sind in einem solchen Fall praktischer Natur. Dies bedeutet natürlich nicht, daß es sich hierbei um triviale Probleme handelt. Es ist keineswegs sichergestellt, daß man bei solchen Dispositionen stets auch in der Lage ist, interessante und brauchbare Operationalisierungen zu finden. Es kann sowohl an geeigneten Daten mangeln, an denen man eine Operationalisierung überprüfen kann, wie auch die getesteten Operationalisierungen sich als inkorrekt herausstellen können. Wenn also gesagt wird, daß die sich hier stellenden

Probleme praktischer Natur sind, so ist damit nur gemeint, daß die prinzipielle Logik der Überprüfung von Operationalisierungen und der Validität der von ihnen gelieferten Resultate keine besonderen Schwierigkeiten bereitet. Eine Operationalisierung gilt dann als korrekt, wenn die von ihr vorgenommenen Zuschreibungen in jenen Fällen, in denen die Testbedingungen erfüllt sind, mit den tatsächlich eintretenden Reaktionen übereinstimmen. Die Behauptung, daß eine Operationalisierung korrekt ist und die Daten valide sind, ist dann zwar hypothetisch und kann stets durch neue Erkenntnisse erschüttert werden, aber darin ist bei einer empirischen Wissenschaft nichts prinzipiell Problematisches zu sehen. Aus der Perspektive der hier verfolgten Probleme ist entscheidend, daß in einem solchen Fall klar ist, welche Argumente zur Stützung und zur Erschütterung einer Validitätsbehauptung verwendet werden können.

Auf der anderen Seite kann das Interesse an der Entwicklung von Operationalisierungen darauf beruhen, daß man ohne solche Operationalisierungen überhaupt nicht in der Lage ist, die Begriffe verlässlich zuzusprechen. Es geht dabei dann um die Operationalisierung von Begriffen, deren Gebrauch weder durch unsere alltägliche Verwendung und unser Vorwissen, noch durch die Bedeutungsfestlegungen für wissenschaftliche Zwecke hinreichend präzise geregelt ist. Dies dürfte der für die in der empirischen Sozialforschung auftretenden Operationalisierungsprobleme und Validitätsdebatten weitaus typischere Fall sein. Begriffe, bei denen es nicht zumindest prinzipiell möglich ist, die Korrektheit einer Operationalisierung dadurch zu überprüfen, daß man die zu messenden Eigenschaften unabhängig von Operationalisierungen verlässlich feststellen kann, werden *theoretische Begriffe* genannt. In empirischen Untersuchungen treten z. B. mentale Begriffe wie Einstellungen, Motive und Bedürfnisse meist in Form von theoretischen Begriffen auf. Die entsprechenden Eigenschaften werden dort nämlich meist nicht durch definierende Charakterisierungen als Dispositionen eingeführt, sondern die zu messenden Eigenschaften werden entweder direkt durch Bezug auf unsere alltägliche Verwendung der zugehörigen Begriffe bestimmt, oder man bezieht sich auf sogenannte theoretische Definitionen, die allerdings alleine auch nicht ausreichen, um über das Vorliegen der zu messenden Eigenschaft entscheiden zu können. Zwar wird man bei der Mehrzahl der theoretischen Begriffe über einige klare Fälle verfügen, wie „Einstein war intelligent“, „Franz Josef Strauß ist konservativ“, aber dies reicht zur Validierung der Daten nicht aus. Man kann nur jene Operationalisierungen ausscheiden, die eben mit diesen klaren Fällen kollidieren. Die eigentliche Leistung der Operationalisierungen soll doch aber darin bestehen, auch dort über die zu messenden Eigenschaften Auskunft zu geben, wo es sich nicht um klare Fälle handelt. In diesen Fällen steht aber kein Kriterium zur Verfügung, mittels dessen entschieden werden könnte, ob die von einer Operationalisierung gelieferten Zuschreibungen mit den tatsächlichen Extensionen der operationalisierten Prädikate übereinstimmen. Das ontologisch interpretierte Validitätsproblem ist hier nicht mehr lösbar. Man kann die Validität der Daten nicht mehr, wie bei den oben diskutierten Dispositionsprädikaten, zumindest prinzipiell dadurch überprüfen, daß man die einschlägigen Testbedingungen herstellt, bzw. ihr Eintreten abwartet, und dann die auftretenden Reaktionen

mit den gewonnenen Daten vergleicht. Es ist eben keineswegs vorab festgelegt, unter welchen Bedingungen eine Person welches Verhalten zeigen muß, damit man von ihr berechtigt behaupten kann, daß sie die Eigenschaft aufweist.

Im Bereich der theoretischen Begriffe führt also die ontologische Interpretation der Validität in unauflösbare Probleme. Keine mit den Methoden der empirischen Wissenschaft erbringbaren Nachweise reichen aus, um die Validität von Daten sichern zu können. Dementsprechend ist dies auch das Feld, auf dem sich die allermeisten Validitätsdebatten abspielen. Die Argumente, die zum Nachweis der Korrektheit einer Operationalisierung und der Validität der Daten vorgebracht werden können, beruhen auf Prämissen, die genauso problematisch sind, wie die Thesen, die sie stützen sollen. Gehen wir ein paar einschlägige Verfahren der Validierung diesbezüglich durch.

a) Inhaltsvalidierung

Bei der Inhaltsvalidierung wird überprüft, ob von der in einer Operationalisierung vorgeschlagenen Zuschreibungsregel auf Grund unseres Vorwissens erwartet werden kann, daß sie zu korrekten Zuschreibungen führt. Es ist doch aber gerade die mangelnde Signifikanz und Verlässlichkeit unseres Vorwissens, die es als geboten erscheinen läßt, eine Operationalisierung vorzunehmen. Eine Inhaltsvalidierung kann zwar besonders absurde Operationalisierungen vermeiden, aber nicht den Kern des Validitätsproblems lösen. Im Zusammenhang mit der Inhaltsvalidierung wird besonders gern auf ein etwas dubioses Expertentum verwiesen. Experten sollen einen Test bezüglich seiner Eignung bewerten. Würde man aber ein wie auch immer im Detail ermitteltes durchschnittliches Expertenurteil als valide Messung akzeptieren? Dies doch wohl kaum. Jeder weiß nur zu genau, wie entscheidend es dabei auf die Auswahl der angeblichen Experten ankäme. Durch geschickte Auswahl der Experten könnte man die Messungen ganz erheblich manipulieren. Solange aber selbst unter sogenannten Experten kein Konsens darüber besteht, wann eine Eigenschaft vorliegt, können Expertenurteile auch nicht zur Lösung des Validitätsproblems genutzt werden. Natürlich kann man Expertenbewertung eines Tests auch so deuten, daß man sich von Experten seiner Wahl bei der Definition der zu erhebenden Eigenschaften beraten läßt. Dann ist aber das Resultat der Expertenkonsultation keine Validierung, sondern eine Bedeutungsfestlegung der zu messenden Eigenschaften. Durch solche Bedeutungsfestlegungen können Begriffe zwar enttheoretisiert werden, aber man gerät dabei leicht in den Umkreis von Definitionen a la „Intelligenz ist das, was der Intelligenztest mißt“. Auf diesem Wege kann man immer zur Definition eindeutig meßbarer Eigenschaften gelangen, es stellt sich dann nur die Frage nach der wissenschaftlichen Relevanz solcher Messungen.

b) Prognose- und Kriterienvvalidierung

Bei dieser Form der Validierung wird die Validität von Resultaten unter Voraussetzung einer empirischen Hypothese beurteilt. Sei A das zu messende Merkmal, so besitzt die Hypothese die Form: $A(x,t) \rightarrow B(x,t')$ bzw. $A(x,t) \leftrightarrow B(x,t')$. B stellt dabei die Kriterienvariable dar. Wird die Hypothese für alle $t=t'$ behauptet, so handelt es

sich um Kriteriervalidierung. Wird die Annahme hingegen für $t' = \Delta t + t$ für bestimmte Δt gemacht, so handelt es sich um Prognosevalidität.

Entscheidend bei der Prognose- und Kriteriervalidierung ist naturgemäß die Frage, welche Hypothesen man zur Validierung heranzieht. Wenn auf diese Weise die Validität von Resultaten *begründet* werden soll, so muß es Argumente für die Richtigkeit derartiger Hypothesen geben. Von einer Begründung kann ja nur dann gesprochen werden, wenn die verwendeten Prämissen sicherer sind als die erzielten Resultate. Wenn man aber über gute Argumente zur Stützung der verwendeten Hypothesen verfügen würde, könnte man die Hypothese ihrerseits als Operationalisierung für A verwenden. Die Frage, wie man Operationalisierungen als adäquat nachweisen kann, wird durch die Kriterien- und Prognosevalidierung nicht gelöst, sondern nur auf die zur Überprüfung verwendeten Hypothesen verschoben. Daß man, wenn man erst einmal eine adäquate Operationalisierung gefunden hat, andere Kandidaten an dieser überprüfen kann, ist klar und unstrittig. Im Kontext des grundsätzlichen Problems der Messung theoretischer Größen aber auch uninteressant.

c) Konstruktvalidierung

Dieser Begriff wird in der Literatur sehr uneinheitlich und häufig auch sehr unklar benutzt. Unter Konstruktvalidierung sollen hier alle Verfahren fallen, die die Validität von Resultaten mit Hilfe von Modellkonstruktionen nachweisen wollen. Zu denken ist dabei etwa an rekursive Modelle. Wie bei der Prognose- und Kriteriervalidierung soll hier die Überprüfung der Validität der vorliegenden Resultate dadurch ermöglicht werden, daß man die Menge der zur Validitätsprüfung verwendeten Prämissen verschärft. Im Unterschied zur Prognose- und Kriteriervalidierung besteht diese Verschärfung nicht darin, daß man eine bestimmte Operationalisierung bereits als adäquat voraussetzt, sondern man unterstellt, daß die empirisch erhobenen Daten und ihre Beziehung zu der zu messenden theoretischen Größe jeweils ein bestimmtes formales Modell erfüllen. Unter Voraussetzung des jeweiligen formalen Modells wird dann aus der Struktur der erhobenen Daten auf die Beziehung zwischen den erfaßten Merkmalen und der theoretischen Größe geschlossen. Genauso wie bei der Prognose- und Kriteriervalidität ist aber auch bei derartigen Verfahren nicht zu sehen, wodurch die Verwendung der Prämissen (hier in Form der Modellannahmen) zu begründen ist. Die im Rahmen des Modells vorgenommenen Unterstellungen über die Art der Beziehungen zwischen der zu messenden theoretischen Größe und den vorgefundenen Daten können nicht durch empirische Daten gestützt werden, ohne daß man bereits über eine als adäquat ausgewiesene Operationalisierung verfügt.

Weder die Inhaltsvalidierung, noch Kriterien- und Prognosevalidierung, noch Konstruktvalidierung sind geeignet, valide Resultate bei Messung theoretischer Größen zu garantieren.

Nun könnte man einwenden, diese Überlegungen seien nicht besonders relevant, da die Definition des Terminus „theoretischer Begriff“ bewußt so gewählt wurde,

daß eigentlich auch gar keine anderen Resultate zu erwarten waren. Die Probleme von Leuten, die sich mit der Messung so eigentümlicher Dinge wie theoretischer Größen beschäftigen, müssen aber nicht die Probleme der empirischen Sozialforschung sein. Daß jedoch die Beschäftigung mit dem Problem der validen Messung theoretischer Größen nicht nur ein abgehobenes akademisches Spiel ist, zeigt ein Blick auf die Praxis der empirischen Sozialforschung.

Dort werden nämlich keineswegs im Regelfall die zu messenden Eigenschaften über explizite Bedeutungsfestlegungen mit unproblematisch konstatierbaren Indikatoren verbunden, so daß die Qualität von Operationalisierungen nicht durch Vergleich mit dieser vorgegebenen Zuschreibungsregel bestimmt werden kann. Dementsprechend erweisen sich dort Validitätsdebatten, sind sie erst einmal ausgebrochen, als unerschlichtbar. Solange man Validität ontologisch interpretiert, gibt es keinen verbindlichen Bezugspunkt, mittels dessen Operationalisierungen beurteilt werden können. Aber selbst wenn faktisch in den Sozialwissenschaften häufig mit theoretischen Begriffen operiert wird, so stellt sich angesichts dieser Probleme die Frage, ob sie nicht doch prinzipiell verzichtbar sind. M. a. W. ob es nicht möglich ist, sich auf Begriffe zu beschränken, die bereits allein auf Grund ihrer Bedeutungsfestlegungen meßbar sind. Diese Frage soll im nächsten Abschnitt untersucht werden.

II. Die Unverzichtbarkeit theoretischer Begriffe für die Sozialwissenschaften

Es gibt Gründe, die es gerade für Sozialwissenschaften als zwingend erscheinen lassen, auf theoretische Größen nicht zu verzichten. Diese spezifisch auf die Sozialwissenschaften ausgerichteten Argumente für die Aufgabe der ontologischen Validitätsinterpretation ergeben sich aus der Struktur mentaler Prädikate. Hält man nämlich an der ontologischen Interpretation von Validität fest, so kann man nur dann mentale Größen messen, wenn man sie mittels Bedeutungsfestlegungen mit beobachtbarem Verhalten verbindet. Bei mentalen Prädikaten wie Einstellungen, Motiven und Bedürfnissen erweist sich aber ein solches Vorgehen als prinzipiell inadäquat.

Die nächstliegende Strategie, durch Bedeutungsfestlegungen meßbare mentale Prädikate einzuführen, besteht darin, sie als Dispositionen aufzufassen, unter bestimmten Bedingungen in charakteristischer Weise zu handeln. Einige Autoren schlagen auch in der Tat vor, Größen wie Einstellungen, Motive und Bedürfnisse als Dispositionen zu definieren. Gegen ein solches Vorgehen spricht aber, daß unsere alltägliche Verwendung derartiger mentaler Begriffe entscheidend von einer solchen Bedeutungsstruktur abweicht. Einstellungen, Motive und Bedürfnisse stellen zwar auch in anderen alltäglichen Verwendungen Größen dar, die auf das Verhalten einen bestimmenden Einfluß haben, aber eben nicht im Sinne einfacher Dispositionen. Jemand kann seine Einstellung zu verbergen suchen und ganz bewußt entgegen seinen Einstellungen handeln.

Die nächstliegende Modifikation dieser Vorgehensweise besteht darin, die Bedeutungsregeln so zu konstruieren, daß nur unter der Voraussetzung subjektiv redlichen Handelns mentale Eigenschaften mit entsprechendem Verhalten gekop-

pelt sind. Die allgemeine Form solcher Bedeutungsregeln wäre dann: unter der Voraussetzung, daß x subjektiv redlich ist und die Bedingungen A erfüllt sind, so besitzt x die mentale Eigenschaft M (genau) dann, wenn x in der Weise B handelt.

Dieses Schema scheint auf den ersten Blick wenn auch keine optimale, so doch eine befriedigende Möglichkeit der Einführung mentaler Begriffe zu sein. Zwar werden mentale Eigenschaften nur unter der Voraussetzung subjektiver Redlichkeit mit beobachtbarem Verhalten verbunden, doch verfügen wir ja auf Grund unserer sozialen Alltagskompetenz zumindest über eine grobe Möglichkeit, die Wahrscheinlichkeit für subjektiv redliches Verhalten abzuschätzen. Es sieht also so aus, als könnten wir auf diesem Wege sowohl Begriffe in Übereinstimmung mit den allgemeinen Charakteristika mentaler Prädikate einführen und gleichzeitig die theoretischen Voraussetzungen zur Lösung von Validitätsproblemen schaffen. Eine Operationalisierung gilt bei einem so eingeführten Begriff dann als valide, wenn in jenen Fällen, in denen wir aufgrund unserer sozialen Alltagskompetenz von subjektiv redlichem Verhalten ausgehen und in denen zusätzlich die einschlägigen Bedingungen vorliegen, die mentale Eigenschaft M (genau) dann zugesprochen wird, wenn dies von der entsprechenden Bedeutungsregel verlangt wird.

Wenn man sich im Wissenschaftsprozess auf seine Alltagsfähigkeiten beruft, so ist dies mit dem Anspruch verbunden, daß man mittels seiner Alltagskompetenz zu im wesentlichen richtigen Resultaten gelangt. Dieser Anspruch ist aber nur dann sinnvoll, wenn er wenigstens prinzipiell einer wissenschaftlichen Überprüfung zugeführt werden kann. Sehen wir uns dies im Fall der Annahme subjektiv redlichen Verhaltens an. Wie könnte eine Überprüfung unserer sozialen Alltagskompetenz in diesem Punkt aussehen? Subjektiv redliches Verhalten liegt genau dann vor, wenn sich jemand entsprechend seiner Einstellungen, Motive, Bedürfnisse etc. verhält. Genau dies könnte aber nur *unter der Voraussetzung* der Annahme subjektiv redlichen Verhaltens gemessen werden. Wenn man also die ontologische Interpretation des Validitätsproblems zugrunde legt, gelangt man in ein unlösbares Dilemma. Ob man die entsprechenden mentalen Größen gemessen hat, kann man aufgrund der Bedeutungsfestlegungen nur wissen, wenn man bereits weiß, in welchen Fällen subjektiv redliches Verhalten vorlag. Umgekehrt benötigt man die Kenntnis der mentalen Größen um festzustellen, ob subjektiv redliches Verhalten vorlag. Unsere soziale Alltagskompetenz läßt sich also vor dem Hintergrund der ontologischen Validitätsinterpretation keiner wissenschaftlichen Überprüfung zuführen. Jeder Überprüfungsversuch mündet direkt in einem Zirkel: Um feststellen zu können, welche Personen zur Extension des vorgegebenen Prädikats „verhält sich subjektiv redlich“ gehören, muß man wissen, welche mentale Eigenschaften diese Personen besitzen. Um aber entscheiden zu können, zur Extension welcher mentaler Prädikate eine Person gehört, muß man bereits voraussetzen können, daß sie sich subjektiv redlich verhält.

Um Mißverständnisse zu vermeiden: Es soll hier weder für die Untauglichkeit unserer sozialen Alltagskompetenz noch gar für die Unmöglichkeit einer mentalistisch orientierten Sozialwissenschaft argumentiert werden. Es geht vielmehr darum, zu zeigen, daß man sich im Rahmen einer bestimmten methodologischen Konzeption, nämlich der ontologischen Interpretation des Validitätsproblems, nicht

mehr sinnvoll auf bestimmte Alltagsfähigkeiten berufen kann, weil der damit implizit erhobene Anspruch der Triftigkeit dieser Alltagskompetenz nicht mehr eingeholt und überprüft werden kann.

Um verständlicher zu machen, wieso es trotz der hier aufgezeigten Probleme möglich ist, auch bei theoretischen Begriffen zu korrekten Operationalisierungen zu gelangen, muß man sich die Natur der Validitätsprobleme ansehen, die bei der Formulierung offensichtlich adäquater Operationalisierungen gelöst werden. Dies soll nun am Beispiel der Gewichtsmessung in Form eines Dialogspiels geschehen.

III. Die instrumentalistische Interpretation des Validitätsproblems

Um die Argumentation überschaubar zu halten und die Analogie zu sozialwissenschaftlichen Problemen deutlicher werden zu lassen, gehen wir davon aus, daß die Operationalisierung nicht aus schon bereits bestätigten Theorien hergeleitet werden kann.

Nehmen wir einmal an, A sei ein Wissenschaftler, der die Relationen „schwerer als“ und „gleich schwer“ mit einer Balkenwaage operationalisieren will. B sei ein konsequenter Methodologe, der nur dann bereit ist, dies als Etablierung einer ordinalen Gewichtsskala zu akzeptieren, wenn die Validität dieses Verfahrens nachgewiesen werden kann.

A: Mit Hilfe der Balkenwaage bin ich in der Lage, eine ordinale Gewichtsmessung vorzunehmen.

B: Wie kommst Du zu dieser Einschätzung?

A: Die mit Hilfe dieser Balkenwaage überprüfbareren Relationen erfüllen alle formalen Bedingungen einer ordinalen Gewichtsmessung. Halten zwei Körper die Waage im Gleichgewicht, so werden sie als gleich schwer eingestuft. Die Relation ist symmetrisch, denn, lege ich die Körper in jeweils eine andere Waagschale, so befindet sich die Balkenwaage genau dann im Gleichgewicht, wenn sie dies auch vor der Vertauschung war. Die Relation ist transitiv: halten zwei Körper jeweils mit einem dritten die Waage im Gleichgewicht, so werden sie auch im direkten Vergleich als gleich schwer ausgewogen. Die Reflexivität, daß ein Körper genau so schwer wie er selbst ist, versteht sich ja von selbst und bedarf keiner empirischen Überprüfung. Empirisch nachweisbar hingegen ist die Transitivität der Operationalisierung von „schwerer als“. Neigt sich nämlich bei zwei Körpern a und b der Balken auf der Seite des Körpers a herab und tut er desgleichen bei dem Vergleich der Körper b und c auf der Seite von b, so wird a mit Hilfe der Balkenwaage auch schwerer als c ausgewogen. Mehr ist nicht zu zeigen, also ist diese Operationalisierung angemessen.

B: Nein. So einfach geht es nicht. Was Du bis jetzt gezeigt hast, ist nichts weiter als daß Du bestimmte formale Merkmale der durch die Balkenwaage operationalisierten Eigenschaft nachgewiesen hast. Diese formalen Merkmale können aber

viele Eigenschaften aufweisen. Woher bist Du Dir sicher, daß Du genau die Eigenschaft „Gewicht“ mißt?

- A: Nun, ich kann z. B. zeigen, daß die Länge eines Körpers keine Rolle spielt. Ich kann einen Körper beliebig deformieren, auseinanderziehen und wieder zusammendrücken, ohne daß dies die an der Balkenwaage ablesbaren Relationen auch nur im geringsten verändert.
- B: Mag sein, daß Du nachweisen kannst, daß die von Dir vorgeschlagene Operationalisierung nichts mit Längenmessung zu tun hat, aber vielleicht spricht Dein Instrument in Wirklichkeit auf eine Eigenschaft an, an die wir im Augenblick gar nicht denken, vielleicht auch noch gar nicht kennen.
- A: Nun, wir brauchen uns ja nicht unwissender zu stellen als wir sind. Wir verfügen doch über ein ganz alltägliches Wissen darüber, welche Körper schwerer als welche anderen sind. In vielen Fällen ist dieses Wissen über jeden vernünftigen Zweifel erhaben. Nehmen wir solche Paare von Körpern, bei denen aufgrund unseres alltäglichen Wissens feststeht, welcher von beiden schwerer ist und legen sie auf die Waage. Dann wirst Du sehen, daß die Waage mit unserem Vorwissen vollkommen übereinstimmende Urteile liefert und Du wirst endgültig beruhigt sein.
- B: Das glaube ich nicht. Denn wenn Deine Waage irgendeinen praktischen Zweck erfüllen soll, dann muß sie uns auch dort Auskunft geben, wo wir aufgrund unseres Vorwissens uns nicht zutrauen, eindeutige Urteile zu fällen.
- A: Da liegst Du vollkommen richtig. Diese Balkenwaage arbeitet sehr genau. Sie vermag auch dort noch Gewichtsunterschiede festzustellen, wo wir aufgrund des Wägens mit unseren bloßen Händen kein sicheres Urteil mehr abzugeben vermögen.
- B: Da beginnt aber auch genau das Problem. Im Bereich, in dem die Balkenwaage ihre eigentliche Nützlichkeit entfalten soll, vermögen wir auch nicht die Validität dieses Meßverfahrens aufgrund unseres Vorwissens zu sichern. Also sind wir im Endeffekt kein bißchen weiter gekommen. Gut, Deine Balkenwaage zeigt keinen groben Unsinn an. Dies könnten wir aufgrund unseres Vorwissens ausschließen. Wenn diese Balkenwaage aber ein Instrument zur Erweiterung unseres Wissens über Gewicht und die Beziehung dieser Größe zu anderen Eigenschaften sein soll, dann muß ihre Validität auch gerade in jenen Bereichen gesichert sein, in denen unser Vorwissen versagt. Ansonsten könnten wir ja von vornherein uns mit diesem Vorwissen bescheiden und bräuchten keine Arbeit in die Konstruktion von Balkenwaagen zu verschwenden.
- A: Nun gut, dann will ich es mal ganz anders probieren. Sieh Dir die Konstruktion der Balkenwaage genau an. Auf was anderes als auf Gewicht soll sie ansprechen!
- B: Nun, ich gebe zu, daß unser Vorwissen über mechanisches Verhalten von Körpern uns erwarten läßt, daß die Resultate dieser Waage nicht mit unseren alltäglichen Urteilen über Gewichtsrelationen kollidieren werden. Jedoch, wie schon

gesagt, unser einbringbares Vorwissen hat Grenzen. Es gibt viele Gewichtsrelationen, die wir auf seiner Basis nicht sicher beurteilen vermögen. Auch unser Vorwissen über das Verhalten eines solchen Gerätes, wie Deine Balkenwaage, hat Grenzen. Vielleicht gibt es ja Kräfte, die das Verhalten Deiner Balkenwaage dirigieren, die nichts mit dem wahren Gewicht der Körper zu tun haben, aber nur so schwache Abweichungen verursachen, daß sie es nicht ermöglichen, Deine Balkenwaage aufgrund unseres Vorwissens als invalide einzustufen. Wenn dem aber so wäre, so wäre Deine Waage wertlos. Dort, wo sie valide Resultate liefert, könnten wir auch ohne die Waage zu Rande kommen, dort aber, wo wir uns auf sie verlassen müßten und nicht durch unser Vorwissen kontrollieren könnten, würde sie nicht valide Resultate liefern. Es ist genau jener Zweifel an der Validität in den nicht mehr kontrollierbaren Bereichen, den Du ausräumen müßtest.

A: Mir scheint jetzt langsam, Du willst mir eine Falle stellen. Konstruiere ich die Balkenwaage so unempfindlich, daß sie nur dort zu Resultaten führt, wo wir aufgrund unseres Alltagswissens sowieso schon sicher sind, nennst Du sie nutzlos, konstruiere ich sie sehr empfindlich, so daß sie zu einer echten Verbesserung unserer Möglichkeiten der Beurteilung von Gewichten führt, lautet Deine Reaktion: „Wie kann die Validität gesichert werden?“ Mir ist letztlich egal, ob die Balkenwaage so etwas mißt wie das, das Du „wahres Gewicht“ nennst. Mit Hilfe dieser Balkenwaage kann ich Gesetze formulieren und gute Prognosen über das Verhalten von Körpern machen. Letztlich bin ich als Wissenschaftler daran interessiert, und nicht an einem „wahren Gewicht“. Solange mir die Balkenwaage bei der Formulierung von Gesetzen gute Dienste leistet, werde ich sie zur Gewichtsmessung benutzen.

Dieses Dialogspiel hat hoffentlich deutlich gemacht, daß die konsequente Anwendung der ontologischen Interpretation von Validität schon bereits bei so unverdächtigen Größen wie Gewicht zu unlösbaren Schwierigkeiten führt. Die einzige Möglichkeit, die Messung mit einer Balkenwaage als im ontologischen Sinne korrekt auszuweisen, besteht darin, die Gewichtsrelationen durch Bedeutungspostulate mit dem Verhalten der Balkenwaage zu verbinden. Dies ist aber offensichtlich inadäquat. Jedes Verfahren der Gewichtsmessung hat Grenzen seiner Genauigkeit. Unterschiede des Gewichts können mit einer Balkenwaage erst ab einer bestimmten Größe wahrgenommen werden. Ab welchem Punkt aber Gewichtsunterschiede auch mit einer Balkenwaage festgestellt werden können, ist eine empirische Frage, die nicht auf dem Wege der Bedeutungsfestlegung vorentschieden werden kann.

Das ontologische Validitätsverständnis gestattet es nicht, zu rekonstruieren, warum es sinnvoll ist, eine Balkenwaage an Stelle des Auswiegens mit Händen und Gefühl einzusetzen. Wenn aber ein methodologisches Konzept sich im Widerspruch zu relevanten Aspekten der Wissenschaftsentwicklung befindet, ist dies ein guter Grund für die Preisgabe des methodologischen Konzepts. Da die Preisgabe nicht in der Form geschehen kann, einfach davon auszugehen, daß es bei der Messung überhaupt kein Validitätsproblem gibt, muß man nach einer angemessenen Interpreta-

tion für das Validitätsproblem suchen. Die Situation, von der wir dabei auszugehen haben, ist scheinbar paradox. Einerseits sollen Operationalisierungen den Status von empirischen Behauptungen haben, andererseits können sie isoliert betrachtet keinen empirischen Überprüfungen zugänglich gemacht werden. Es stehen weder Bedeutungsregeln zur Verfügung, an Hand deren sie auf ihre Korrektheit hin überprüft werden können, noch kann man unproblematisch auf andere Operationalisierungen des gleichen Begriffs zurückgreifen, da sich bei diesen ja das gleiche Problem stellt. Die Sicherung der Validität muß also darin bestehen, die Richtigkeit empirischer Behauptungen nachzuweisen, die isoliert keiner Überprüfung fähig sind. Die Auflösung dieser paradoxen Situation kann nur dann gelingen, wenn es möglich ist, Operationalisierungen im Zusammenhang mit anderen Aussagen einer empirischen Testung zuzuführen.

Eine solche Möglichkeit besteht in der Tat. In der letzten Passage des Dialogspiels wird darauf verwiesen, daß die mit einer Balkenwaage vorgenommenen Messungen zu guten Prognosen über das Verhalten der Körper führe. Dies ist nicht nur ein pragmatisches Argument, um sich gegenüber einem überzogenen methodologischen Anspruch zur Wehr zu setzen, sondern liefert auch den Schlüssel zur Auflösung der paradoxen Situation.

Das entscheidende Ziel jeder empirischen Wissenschaft ist es, zu Systematisierungen ihres Gegenstandsbereichs zu gelangen. Erreicht werden kann das Ziel nur auf dem Wege der Konstruktion von Theorien und Erklärungsmodellen. Operationalisierungen theoretischer Begriffe können, wenn überhaupt, *nur als Bestandteile solcher Theorien und Modelle* auf ihre Korrektheit hin überprüft werden. Ergibt sich aber, daß eine Operationalisierung wesentlicher Bestandteil einer leistungsfähigen Systematisierung ist, so ist damit ihre Relevanz nachgewiesen. Letztlich ist dieser Nachweis der Relevanz auch Beweis für die Korrektheit der Operationalisierung.

Wenn man im Rahmen eines Modells Personen bestimmte Einstellungen zuschreibt und mit Hilfe dieser Zuschreibungen in dem Modell Verhalten erklären und prognostizieren kann, so ist dieses Modell ein erfolgreiches Instrument der Systematisierung eines bestimmten Gegenstands- und Verhaltensbereichs. Die Operationalisierungen, über die die Einstellungen gemessen werden, sind dann wesentliches *Element* einer leistungsfähigen Systematisierung. Ohne diese Operationalisierungen könnten die Erklärungen und Prognosen nicht vorgenommen werden. Daß die Verwendung bestimmter Operationalisierungen und nicht irgendwelcher anderen zu den guten Prognosen und Erklärungen führen, ist der Grund, warum man an diesen Operationalisierungen festhält. Sie zeichnen sich durch ihren Beitrag zu gelungenen Systematisierungen vor anderen Operationalisierungen aus. Man wird sie erst preisgeben, wenn man Operationalisierungen gefunden hat, die zu noch besseren Prognosen und Erklärungen führen. Der Umstand, daß man jederzeit bereit wäre, eine Operationalisierung aufzugeben, wenn man eine bessere Alternative zur Verfügung hat, macht klar, daß die Operationalisierungen den Status empirischer Behauptungen haben. Obwohl sie isoliert keiner empirischen Überprüfung zugänglich sind, können sie doch im Rahmen eines Modells oder einer Theorie auf ihre Güte hin getestet werden.

Wenn man in den Sozialwissenschaften mit solchen Größen wie Einstellungen, Motiven, Bedürfnissen etc. arbeitet, so doch nicht deswegen, um Etikettierungen vornehmen zu können, sondern um mit ihrer Hilfe zu möglichst guten Erklärungen und Prognosen gelangen zu können. Folglich sind jene Operationalisierungen gute Operationalisierungen, mit denen wir solche Leistungen vollbringen können. Diese Leistungen können aber natürlich nicht von den Operationalisierungen allein erbracht werden, sondern es müssen Modelle und Theorien vorliegen, *innerhalb* deren die Operationalisierungen ihren Beitrag zur Systematisierung liefern. Es ist also ganz und gar sinnlos, sich nach der Güte einer bestimmten Operationalisierung zu fragen, solange keine Modelle und Theorien vorliegen, in deren Rahmen sie verwendet wird. Isoliert betrachtet, ist eine Operationalisierung eines theoretischen Begriffs weder wahr noch falsch. Es können aus ihr weder wahre noch falsche Schlußfolgerungen gezogen werden. Erst im Zusammenhang mit anderen Behauptungen lassen sich aus einer Operationalisierung testbare Schlußfolgerungen ziehen. Die Tatsache, daß eine Operationalisierung allein keine überprüfbaren Schlußfolgerungen herzuleiten gestattet, verträgt sich durchaus damit, daß sie *als Bestandteil eines Aussagensystems* wesentlichen Einfluß auf die Menge der testbaren Schlußfolgerungen dieses Systems von Aussagen hat. Dies hängt aus logischer Sicht damit zusammen, daß die Menge der Schlußfolgerungen aus einer Konjunktion von Aussagen i.a. größer ist als die Vereinigung der Mengen von Schlußfolgerungen der einzelnen Aussagen. Es kann eine Konjunktion von Aussagen testbare Schlußfolgerungen aufweisen, obwohl jede Aussage für sich betrachtet keine überprüfbaren Schlußfolgerungen aufweist.

Daß eine Operationalisierung eines theoretischen Begriffs korrekt ist, kann also sinnvoller Weise nichts anderes bedeuten, als daß sie *Bestandteil* einer korrekten Theorie ist. Die Korrektheit einer Theorie erweist sich aber an der Richtigkeit der aus ihr herleitbaren testbaren Schlußfolgerungen. Also besagt die Validität von Daten, daß man mit Hilfe der gewonnenen Resultate zu guten Prognosen und Erklärungen kommen kann. Kehren wir nochmals zu dem Beispiel der Gewichtsmessung zurück. Würden wir die von einer Waage gelieferten Resultate nur im Sinne von Etikettierungen der gewogenen Körper verwenden, so hätten wir keinen Grund, die von der Waage gelieferten Werte als korrekte Gewichtsmessung zu betrachten. Erst der Umstand, daß die von der Waage gelieferten Daten es uns erlauben, mit Hilfe von Gesetzen und Theorien zu guten Prognosen zu kommen, rechtfertigt es, die Balkenwaage als funktionierendes Instrument zu betrachten, das auch dort noch valide Resultate liefert, wo wir aufgrund unseres Vorwissens nicht mehr in der Lage sind Gewichtsrelationen festzustellen. Ganz analog sind die Probleme bei der Zuschreibung theoretischer Begriffe in den Sozialwissenschaften. Solange es nur darum geht, welche Personen wir politisch konservativ nennen, sind alle nicht vollständig absurden Operationalisierungen gleich gut bzw. gleich schlecht. Keine Operationalisierung weist gegenüber den anderen aus wissenschaftlicher Sicht Vorteile auf. Man kann sie allenfalls dazu verwenden, den noch immer wissenschaftsgläubigen Teil der Öffentlichkeit durch scheinbar wissenschaftliche Untersuchungen in der einen oder anderen Richtung zu beeinflussen. Dementsprechend gibt es auch keine Möglichkeit, Debatten darüber, welche Operationalisierung die bessere ist, auf rationale Weise zu

entscheiden. Die Sachlage ändert sich erst, wenn man bestimmte Operationalisierungen im Kontext einer Theorie zu erfolgreichen Systematisierungen nutzen kann. Dann sind eben nicht mehr alle denkbaren Operationalisierungen gleich gut. Dann sind wir in der Lage, zwischen den verschiedenen Operationalisierungen auf Grund wissenschaftlicher Kriterien eine Selektion vorzunehmen. Bei theoretischen Begriffen besteht zwischen der Etablierung von Meßverfahren und der Entwicklung von Theorien nicht nur ein pragmatischer, sondern ein systematischer Zusammenhang.

Bisher haben wir so argumentiert als wäre der Systematisierungsbeitrag, den Operationalisierung liefert, der einzig relevante Bezugspunkt für ihre Beurteilung. Dies stimmt so nicht ganz. Wir erwarten von einer Operationalisierung eines theoretischen Begriffes natürlich auch, daß die von ihr erzeugten Zusprechungen sich zumindest im weiteren Umfeld unserer alltäglichen Verwendung aufhält und mit den eventuell vorliegenden theoretischen Definitionen verträglich ist. Wenn wir diese Seite bisher unterschlagen haben, so deswegen, um den eigentlich relevanten Kern des Problems der Messung theoretischer Größen deutlicher herausarbeiten zu können. Im Endeffekt handelt es sich hierbei ja nicht um ein Problem der Signifikanz theoretischer Begriffe, sondern um ein Problem der Namensgebung. Ob ich bereit bin, einen bestimmten Namen für eine theoretische Größe zu akzeptieren, ändert nichts an den Leistungen, die man mit Hilfe dieser Größen vollbringen kann. So gesehen bedeutet also die Absicht, eine bestimmte Größe durch Etablierung eines Meßverfahrens wissenschaftlich nutzbar zu machen nichts anderes, als ausgehend von einem bestimmten Begriff wissenschaftliche Theorien zu formulieren, die a) erfolgreiche Systematisierungen darstellen und in denen b) der theoretische Begriff im Umfeld seiner vorgegebenen Bedeutung verwendet wird.

Es kann keineswegs als trivial vorausgesetzt werden, daß alle theoretischen Begriffe unserer Alltagssprache so zu Kristallisationspunkten erfolgreicher Systematisierungen dienen können. Es ist auch nicht ausgeschlossen, daß sich bei Entwicklung wissenschaftlicher Theorien theoretische Größen als nützlich erweisen, die nicht mehr als Präzisierung eines in unserer Alltagstheorie auftauchenden Begriffs aufgefaßt werden können. Will man z. B. Einstellungsmessung betreiben, so kann sich unter Umständen herausstellen, daß man erfolgreicher arbeitet, wenn man Typen von Einstellungen definiert, für die es so in unserer Alltagssprache keinen passenden Namen gibt. Hier stellt man sich vielleicht die Frage, worin dann die Bedeutung solcher Begriffe besteht. Die einzige verbindliche Charakterisierung eines solchen Begriffs stellen seine Operationalisierung und die Aussagen der entsprechenden Theorie dar. All diese Beziehungen sollen aber als empirisch betrachtet werden. Was legt aber dann seine Bedeutung fest? Ganz ähnlich kann man auch schon bei den der Alltagssprache entlehnten theoretischen Begriffen argumentieren. Dort werden ja ebenso die zur Präzisierung eingesetzten Charakterisierungen durchweg als empirisch betrachtet.

Um diese Frage beantworten zu können muß man sorgfältig zwischen zwei Problemen unterscheiden:

1. Wann hat ein Begriff Bedeutung?
2. Wann haben zwei Begriffe die gleiche Bedeutung?

Was die erste Frage angeht, so können wir uns mit der Feststellung begnügen, daß ein Begriff auf jeden Fall dann eine Bedeutung hat, wenn man zwischen korrekten und inkorrekten Verwendungen unterscheiden kann. Dies ist besonders dann der Fall, wenn man zwischen korrekten und inkorrekten Zusperechungen unterscheiden kann. Dafür reicht aber ein System gut bestätigter empirischer Annahmen durchaus aus. Man benötigt nicht eine Reihe expliziter Bedeutungsfestlegungen, um mit einem Begriff vernünftig umgehen zu können. Niemand würde z.B. bestreiten, daß „Zitrone“ eine Bedeutung hat, ohne daß wir jemals eine explizite Bedeutungsfestlegung für „Zitrone“ vorgenommen haben. Wir haben ganz einfach als Kinder gelernt, diesen Begriff erfolgreich (in Übereinstimmung mit Sprachgebrauch anderer) zu verwenden. Auch wenn man in einem Lexikon nachschlägt, wird man einfach eine Aufzählung der Eigenschaften von Zitronen finden, ohne daß ein Unterschied zwischen bedeutungsfestlegenden und nur empirischen Merkmalen gemacht wird.

Das zweite Problem ist anderer Natur. Um beim Beispiel der Zitrone zu bleiben: Welche Abweichungen von dem jetzigen Gebrauch des Worts „Zitrone“ dürfen wir machen und trotzdem berechtigt behaupten, daß es sich noch um den gleichen Begriff handelt? Wenn sich eines Tages herausstellt, daß die bisher als Zitronen bezeichneten Objekte gar nicht von Natur aus gelb sind, sondern industriell gefärbt werden und zudem nicht an Zitronensträuchern wachsen, sondern wie Kartoffeln unter der Erde gedeihen, können wir dann noch behaupten, wir verwendeten den gleichen Begriff von „Zitrone“, wenn wir die Objekte wie bisher so nennen, oder benutzen wir einen anderen Begriff, der nur den gleichen Namen hat? Um diese Frage entscheiden zu können, benötigt man in der Tat eine explizite Bedeutungsfestlegung unseres Jetztbegriffs von Zitrone. Eine Sprache kann aber offensichtlich ganz gut funktionieren, ohne daß man solche expliziten Bedeutungsfestlegungen vornimmt. Ob das gleiche Wort in zwei verschiedenen Theorien verwendet die gleiche Bedeutung hat, mag durch dieses Konzept theoretischer Begriffe unbeantwortbar werden, aber bei der Entwicklung von Theorien geht es ja nicht darum, Sprachnormierung zu betreiben, sondern leistungsfähige Systematisierungen zu entwickeln. Um ein in der Wissenschaftstheorie viel diskutiertes Problem aufzugreifen: Ob die relativistische Mechanik den gleichen Massenbegriff wie die klassische Mechanik verwendet ist für die Beurteilung der Leistungsfähigkeit der beiden Theorien gleichgültig. Ebenso irrelevant ist die Frage, ob die in einem bestimmten Einstellungskonzept verwendete Größe „konservative Einstellung“ noch die gleiche Bedeutung wie in der Umgangssprache hat. Die Unlösbarkeit des Problems der Bedeutungsgleichheit kann bei der hier vorgeschlagenen Konzeption von Validierung ohne fatale Folgen in Kauf genommen werden.

Die, wie wir sie nennen wollen, instrumentalistische Interpretation von Validität gestattet es, die Einbettung eines theoretischen Begriffs in eine Theorie gleichzeitig als Präzisierung seiner Bedeutung aufzufassen. Die Fixierung der Zuschreibungsregeln in Form von Operationalisierungen kann die Bedeutung eines theoretischen Begriffs, zumindest was die wissenschaftlich relevanten Aspekte betrifft, hinreichend eindeutig festlegen. Nachdem die theoretischen Begriffe durch Integration

in eine Theorie präzisiert worden sind, kann man die Messung so interpretieren, daß durch sie festgestellt wird, ob einem Objekt die theoretische Eigenschaft zukommt oder nicht. Deswegen findet man auch in Naturwissenschaften bei Validitätsproblemen häufig eine ontologisch orientierte Sprechweise. Dies ist solange ungefährlich, wie man nicht versucht, die Präzisierung eines theoretischen Begriffs durch eine Theorie auch ontologisch zu interpretieren. Hat man aber bereits eine funktionierende Theorie, die einen theoretischen Begriff meßbar macht und sucht nur weitere Meßverfahren für die theoretischen Eigenschaften, so sind dadurch keine Komplikationen zu befürchten. Man besitzt durch die Theorie und deren Operationalisierungen einen Bezugspunkt, an dem weitere Meßverfahren und Operationalisierungen geeicht werden können. Der grundlegende Unterschied zwischen ontologischer und instrumentalistischer Interpretation von Validität wird dadurch nicht berührt. Während es bei der ontologischen Interpretation darum geht, einen vorgegebenen Begriff durch eine Operationalisierung korrekt zu erfassen, geht es im instrumentalistischen Verständnis darum, einen Begriff durch Integration in eine Theorie für die wissenschaftliche Systematisierung fruchtbar zu machen. Im ersten Fall will man sicherstellen, daß die Zuschreibungen eines Begriffs in Übereinstimmung mit einer vorgegebenen Extension geschehen, im zweiten Fall, daß die Zuschreibung zu guten Prognosen und Erklärungen führt. Solange es nur darum geht, zu bereits vorhandenen Möglichkeiten der Zuschreibung andere, pragmatisch interessantere Alternativen zu finden, kann man sich auf das ontologische Verständnis von Validität beziehen. Man verfügt bereits über Kriterien an denen neue Operationalisierungen getestet werden können. Will man hingegen durch Operationalisierung die Präzision, mit der ein Begriff zugeschrieben werden kann, über die vorher mögliche Genauigkeit hinaus verbessern, so führt die ontologische Interpretation in eine Sackgasse.

Es soll hier wenigstens am Rande vermerkt werden, daß es bei der Entwicklung von Operationalisierungen, neben der beschriebenen Form der Rückkopplung zwischen Theorien und Meßverfahren, noch zu einer weiteren Variante des Zusammenspiels zwischen Theorien und Validierungstechniken kommen kann. Die Meßinstrumente und ihre Funktionsweise können selbst zum Gegenstandsbereich gehören, der durch die Theorie systematisiert wird. Dies ist z.B. für naturwissenschaftliche Theorien geradezu typisch. Wir wollen hier aber nicht versuchen, die allgemeine Logik der sich dadurch eröffnenden Möglichkeiten der Verbesserung von Meßverfahren zu analysieren. Die Sozialwissenschaften sind von derartigen Rückkopplungen noch sehr weit entfernt.

Wenden wir uns lieber der Frage zu, welche praktischen Konsequenzen die dargelegte Notwendigkeit einer instrumentalistischen Interpretation von Validität für die Sozialwissenschaften hat. Als erstes fällt auf, daß die Arbeitsteilung zwischen empirischer Sozialforschung und theoretischer Sozialwissenschaft in ihrer gegenwärtig praktizierten Form nicht in der Lage ist, die notwendige Vermittlung zwischen Theoriebildung und Operationalisierung zu gewährleisten. Wenn man das Auffinden geeigneter Operationalisierungen nicht mehr als Problem der korrekten Anwendung von Theorien und statistischen Methoden deuten kann, so ist die gegenwärtig

praktizierte Trennung zwischen theoretischer Analyse und empirischer Anwendung nicht zu vertreten. Die Frage, wie man einen Begriff operationalisiert, ist, i.a. integraler Bestandteil der Theoriebildung und kann nicht einfach von der theoretischen Analyse abgekoppelt werden. *Erst wenn die vorliegenden Theorieansätze durch Aufnahme von Operationalisierungen zu testbaren Konzepten werden, kann man von Theorien sprechen.* Welche Operationalisierungen man geeigneter Weise benutzt, läßt sich aber aus den Theorieansätzen allein nicht ableiten. Erst eine testbare Theorie als Ganzes legt die Verwendungsregeln der in ihr enthaltenen theoretischen Begriffe fest. Die Korrektheit einer Operationalisierung zeigt sich nur im Erfolg dieser Theorie. Wenn also empirische Sozialforscher erfolgreich messen wollen, müssen sie erfolgreiche Theorien zu den entsprechenden theoretischen Begriffen formulieren, und wenn Theoretiker überprüfbare Theorien formulieren wollen, so müssen die notwendigen Operationalisierungen integraler Bestandteil der Theorien sein.

Die gegenwärtige Teilung der Sozialwissenschaften in theoretisch und empirisch Arbeitende führt, zusammen mit der vorherrschenden ontologischen Interpretation, zu einer fatalen Entkoppelung zwischen empirischen Daten und Theoriebildung. Wenn der empirische Sozialforscher die Validität der von ihm erhobenen Daten als eine von der Theoriebildung unabhängige Frage klären soll, gerät er bei theoretischen Begriffen in eine hoffnungslose Lage. Ihm stehen keine unproblematischen Kriterien zur Verfügung anhand deren er die Validität nachweisen könnte. Man kann die produzierten Resultate stets mit dem Vorwurf der mangelnden Validität abschmettern. Die mangelnde Validität fällt aber bei dieser Verteilung von Legitimationslasten auf den empirischen Sozialforscher zurück. Der besitzt aber, solange an der ontologischen Interpretation festgehalten wird (und keine empirisch gehaltvollen Theorien vorliegen), gar keine Möglichkeit mit Hilfe seiner Untersuchungsmethoden die Validität zu gewährleisten. Die einzige Möglichkeit der Legitimation von Operationalisierungen besteht darin zu zeigen, daß empirisch gehaltvolle Theorien, die dann eben auch die Operationalisierungen der theoretischen Begriffe enthalten, sich bei der Konfrontation mit Daten als triftig erweisen.

Solange nicht klar ist, in welcher Weise in welchem Gegenstandsbereich ein theoretischer Begriff wie z.B. konservative Einstellung zur Systematisierung eingesetzt werden soll, ist vollkommen sinnlos sich danach zu fragen, wie man konservative Einstellungen valide erfassen kann. Ist aber klar, wo und wie dieser Begriff zur Systematisierung verwendet werden soll und es liegt eine empirisch testbare Theorie vor, die auch die benötigten Operationalisierungen enthält, so können diese durch den Erfolg der Theorie gerechtfertigt werden. Die von der Theorie implizit vorgenommene Spezifizierung der Bedeutung des theoretischen Begriffs legitimiert sich durch den Erfolg. Dies deckt sich auch mit dem Muster der erfolgreichen Meßpraxis der Naturwissenschaften. Wie anders, als über den durch eine Theorie vermittelten Erfolg könnten die in der Mechanik verwendeten Meßverfahren zur Bestimmung z.B. von Kraft verteidigt werden. Wir haben keinen Grund anzunehmen, daß in der Physik irgendein vorgegebener Kraftbegriff in seiner Extension besonders genau getroffen wird, hingegen gute Gründe, an einer erfolgreichen Mechanik und ihrer Kraftmessung festzuhalten. Liegt aber eine erfolgreiche Theorie vor, die mit

theoretischen Begriffen arbeitet, so vergessen wir schnell, daß es die eigentlich entscheidende Leistung der Theorie war, die theoretischen Begriffe in einer für Systematisierungen fruchtbaren Weise zu präzisieren. Wir sehen nur noch die erfolgreichen Messungen und Anwendungen, die sich nach der Entwicklung einer solchen Theorie auch durchaus mit einer ontologisch orientierten Interpretation von Validität vertragen. Die zu messende Eigenschaft ist eben die durch eine erfolgreiche Theorie etablierte Größe. Versucht man aber diese Interpretation von Validität auf theoretische Begriffe vor der Formulierung erfolgreicher Theorien anzuwenden, so gelangt man zwangsläufig in die geschilderten Schwierigkeiten. Aus diesen Schwierigkeiten gelangt man nur dann heraus, wenn man zu einer instrumentalistischen Interpretation von Validität übergeht und die in diesem Stadium unangemessene Aufgabenverteilung zwischen empirischer Sozialforschung und theoretischer Soziologie aufhebt. Man kann das Problem der Formulierung geeigneter Operationalisierungen nicht von dem Problem der Konstruktion erfolgreicher Theorien trennen.