Götz Rohwer

Zahlen, Größen, Metriken

Diskussionsbeitrag zu einem Seminar zur Wissenschaftstheorie der empirischen Sozialforschung.

Einleitung

In der Statistik spielen Zahlen eine ziemlich große Rolle, aber das ist weder ein spezifisches Merkmal der Statistik noch der quantifizierenden Wissenschaften. Auch in vielen Bereichen des praktischen Lebens spielen Zahlen eine gro'se Rolle. Jeder lernt deshalb den Umgang mit Zahlen bereits in der Schule und durch das Vertrautwerden mit gesellschaftlichen Praktiken, in denen Zahlen eine Rolle spielen.

Aber das ist bereits ein wichtiger Unterschied. Im Mathematikunterricht lernt man den mathematischen Umgang mit Zahlen; hier haben die Zahlen keine spezifische Bedeutung, sondern es wird im wesentlichen nur festgelegt, nach welchen Regeln man mit ihnen rechnen kann. Anders verhält es sich, wenn Zahlen im Kontext gesellschaftlicher Praktiken eine Rolle spielen; zum Beispiel um Längen zu messen oder Preisgrößen auszuhandeln. In diesen Kontexten haben Zahlen eine Bedeutung (die sie in der Mathematik nicht haben); und den richtigen Umgang mit Zahlen zu lernen, besteht dann im wesentlichen darin, die Praktiken zu verstehen, in denen die Zahlen verwendet werden und dadurch eine Bedeutung erhalten.[1]

In der Statistik werden Zahlen primär dafür verwendet, Eigenschaften von Objekten (oder ihrer Beziehungen) zu repräsentieren. Es ist deshalb wichtig zu wissen, wie Zahlen zustande gekommen sind. Nur dann kann man ihre Bedeutung verstehen, und in einem zweiten Schritt dann auch

[1]Die Untersuchung gesellschaftlicher Praktiken unter der Fragestellung, *wie* in ihnen mit Zahlen umgegangen wird, ist seit einiger Zeit selbst zu einem sozialwissenschaftlichen Forschungsthema geworden; vgl. z.B. Lave [1986].

lernen, welche Art von Aussagen mithilfe statistischer Methoden gewonnen werden können. In diesem Beitrag sollen einige der unterschiedlichen Bedeutungen diskutiert werden, die Zahlen haben können.

1 Numerische Fakten

Der Leitgedanke ist: um numerische Aussagen verstehen und beurteilen zu können, ist es zunächst erforderlich zu verstehen, wie die verwendeten Zahlen zustande gekommen sind. Dementsprechend beginnen wir mit dem Versuch, hier einige Unterscheidungen zu treffen.

1. Die erste wichtige Unterscheidung ist die folgende: man kann einerseits numerische Fakten beobachten, und man kann andererseits selbst numerische Fakten erzeugen. Geht man zum Beispiel in einen Supermarkt, kann man vielleicht beobachten, daß auf einem gewissen Ding die Zahl 1.50 steht. Dann hat man eine numerische Tatsache beobachtet; und in diesem Fall versteht man vermutlich ihre Bedeutung: dieses Ding kostet 1.50 DM. Andererseits können wir z.B. die Länge eines Tisches mit einem Metermaß ausmessen. Dann gewinnen wir auch eine Zahl, in diesem Fall ein Meßergebnis, das wir selbst erzeugt haben.

Daß es sich um eine wichtige Unterscheidung handelt, wird deutlich, wenn man daran denkt, warum Sozialwissenschaftler Zahlen verwenden. Sie tun dies, um Aussagen über ihren Gegenstandsbereich – gewisse Aspekte gesellschaftlicher Verhältnisse – zu gewinnen. Dann gibt es zwei unterschiedliche Möglichkeiten, wie sie zu ihren Zahlen kommen können. Sie können die in ihrem Gegenstandsbereich erzeugten numerischen Fakten verwenden; zum Beispiel indem sie die Einkommensverteilung ermitteln. Sie können aber auch eigene Verfahren entwickeln, um mit deren Hilfe numerische Aussagen über gewisse Aspekte ihres Gegenstandsbereich zu formulieren; zum Beispiel können sie eine numerische Skala für „Prestige“ entwickeln und dann numerische Aussagen über Prestigeunterschiede machen.[2] Im ersten Fall muß man verstehen, wie die numerischen Fakten im Gegenstandsbereich zustande gekommen sind; im zweiten Fall muß man lernen, die von Sozialwissenschaftlern verwendeten Meßverfahren zu verstehen, mit deren Hilfe sie numerische Aussagen konstruieren.

Ein weiterer wichtiger Unterschied ist der folgende: Im ersten Fall kann man Zahlen wie Fakten behandeln; das Ding ist blau und trägt u.a. die Aufschriften „100 g“ und „1.50 DM“. Derjenige, der diese Fakten beobachtet, ist nur für ihre korrekte Beobachtung, nicht für ihr Zustandekommen verantwortlich. Im zweiten Fall werden dagegen (Meß-) Verfahren verwendet, um dem Anspruch nach wissenschaftliche Aussagen zu gewinnen. In diesem Fall hängt die Bedeutung der Aussage wesentlich auch von dem (Meß-) Verfahren ab; und zu ihrer Begründung muß – nicht in jedem Einzelfall, aber grundsätzlich – auch das Verfahren begründet werden, mit dem die Aussage gewonnen wurde.

[2]Vgl. zu diesem Beispiel Wegener [1988]; speziell zu meßtheoretischen Problemen von Statusskalen vgl. man Pawson [1982].

2. Unabhängig davon, ob Zahlen im Gegenstandsbereich der Sozialwissenschaftler oder von ihnen selbst produziert werden, sollte man versuchen, ihre Bedeutung zu verstehen, d.h. ein Verständnis der Prozesse bzw. Verfahren zu gewinnen, durch die die Zahlen zustande gekommen sind. Hier können mindestens drei unterschiedliche Arten solcher Prozesse bzw. Verfahren unterschieden werden.

a) Meßverfahren. Darunter verstehen wir zunächst sehr allgemein Verfahren, mit deren Hilfe numerische *Aussagen* über gewisse Aspekte der Realität gewonnen werden sollen. Zum Beispiel dienen die Verfahren zur Längenmessung dazu, quantitative Aussagen über die Länge von Gegenständen machen zu können (vorausgesetzt daß bei ihnen sinnvoll von einer Länge gesprochen werden kann).

b) Verfahren zur Ermittlung von Wertgrößen (Preise und Einkommen). Preise und Einkommen sind sicherlich numerische Größen, aber diese Größen entstehen nicht durch Meßverfahren, sondern zum Beispiel durch Aushandlungsprozesse. Der entscheidende Punkt betrifft jedoch nicht die Frage, wie Preis- und Einkommensgrößen gebildet werden, sondern daß die Zahlen in diesem Fall eine andere Bedeutung haben. Werden Zahlen mit einem Meßverfahren ermittelt, sollen sie dazu dienen, Aussagen über gewisse Eigenschaften von Objekten machen zu können. Werden (wie auch immer) Preis- und Einkommensgrößen gebildet, soll dies nicht dazu dienen, Aussagen über Objekte zu machen, sondern Austauschrelationen numerisch zu fixieren. Bei Meßverfahren kann man diskutieren, ob man mit ihrer Hilfe zu sinnvollen und verläßlichen Aussagen über die Beschaffenheit von Objekten kommen kann; im Hinblick auf Verfahren zur Ermittlung von Preis- und Wertgrößen hat *diese* Frage keinen Sinn (obwohl man auch für sie Beurteilungskriterien suchen kann).

c) Zeichenerzeugende Prozesse. Nicht alle Prozesse, durch die Zahlen erzeugt werden (können), sind entweder Meßverfahren oder Verfahren zur

Ermittlung von Wertgrößen. Wenn ich z.B. einen Spaziergang durch die Universität mache und auf die Dinge, die mir unterwegs begegnen, Zahlen schreibe, hat hinterher ein zahlenerzeugender Prozeß stattgefunden. Aber ich habe weder etwas gemessen noch eine Wertgröße ermittelt. Es hat einfach ein zeichenerzeugender Prozeß stattgefunden. Ob die Zeichen (Zahlen) in diesem Fall eine Bedeutung haben, kann nicht allgemein entschieden werden. Ich müßte ihnen eine Bedeutung gegeben haben, die von anderen verstanden werden kann. Das kann, muß aber nicht der Fall sein. Ein anderes Beispiel ist ein Buch, das 1955 von der Rand Corporation (USA) herausgegeben wurde: *A Million Random Digits with 100,000 Normal Deviates*. Die Zahlen in diesem Buch reflektieren weder Meßergebnisse noch Wertgrößen, sondern sind durch einen zeichenerzeugenden Prozeß zustande gekommen. Ihnen kann daher nicht ohne weiteres eine Bedeutung unterstellt werden.[3]

3. Wenn Zahlen durch Verfahren zustande kommen, die in der gesellschaftlichen Praxis (als Gegenstandsbereich der Sozialwissenschaften) institutionalisiert sind, kann man sie als numerische Fakten behandeln und im übrigen versuchen, ihre Bedeutung zu verstehen, indem man die Prozesse, durch die sie erzeugt werden und in denen sie verwendet werden, untersucht. Wir haben aber schon darauf hingewiesen, daß Sozialwissenschaftler auch selbst Meßverfahren entwickeln und anwenden, um zu numerischen Aussagen zu gelangen. Natürlich sind das nicht immer spezifisch sozialwissenschaftliche Meßverfahren. Wenn ein Sozialwissenschaftler zum Beispiel eine Uhr verwenden, um die Dauer eines sozialen Vorgangs zu bestimmen, verwendet er ein Verfahren zur Zeitmessung, das auch in anderen Wissenschaften und ganz generell in der gesellschaftlichen Praxis üblich ist. Wenn er aufgefordert wird, seine Verwendung dieses Meßverfahrens zu rechtfertigen, kann er darauf verweisen, daß das Verfahren allgemein üblich ist, und im übrigen auf die wissenschaftliche Literatur, die sich mit der Begründung von Zeitmeßverfahren beschäftigt.[4]

Anders verhält es sich, wenn Meßverfahren verwendet werden, die für sozialwissenschaftliche Anwendungen entwickelt worden sind, insbesondere wenn neue Verfahren dieser Art vorgeschlagen werden. Zum Beispiel, wenn Sozialwissenschaftler beanspruchen, Meßverfahren für „Prestige“, „Intelligenz“ oder „autoritären Charakter“ entwickelt zu haben. In diesen Fällen kann und sollte die Frage gestellt werden, ob es sich um sinnvolle Verfahren handelt, d.h. um Verfahren, mit denen sinnvolle und begründbare Aussagen gewonnen werden können.

Die Formulierung der Frage zeigt bereits, daß man meistens unterschiedlicher Meinung sein kann. Denn dafür, ob ein Meßverfahren sinnvoll und zweckmäßig ist, gibt es keine harten Kriterien. An dieser Stelle hilft auch ein Verweis auf die Empirie nicht weiter. Denn Meßverfahren können strenggenommen nicht empirisch geprüft werden, da sie ja dazu dienen sollen, empirische Daten zu gewinnen. Ich könnte z.B. irgendein Verfahren zur Längenmessung erfinden und damit die Länge von Gegenständen messen. Wenn dabei Meßergebnisse herauskämen, die von den Ergebnissen unserer üblichen Verfahren zur Längenmessung abweichen, würde das dennoch nicht zeigen können, daß mein Verfahren falsch mißt. Es würde nur zeigen, daß mein Verfahren etwas anderes mißt als das, was wir mit unseren üblichen Längenmeßverfahren messen. Mein Verfahren würde einfach nur andere empirische Daten erzeugen. Es hat deshalb keinen Sinn, danach zu fragen, ob ein Meßverfahren das, was es zu messen vorgibt, richtig mißt oder nicht. Denn es gibt keinerlei Möglichkeit, das durch Hinweise auf die Beschaffenheit der Realität zu beurteilen.[5] Die entscheidende Frage ist vielmehr, ob ein Meßverfahren sinnvolle, wissenswerte und verläßliche Aussagen über die Realität liefern kann.[6]

4. Die oben verwendete Formulierung, daß Meßverfahren dazu dienen, numerische Aussagen zu gewinnen, ist zwar nicht falsch, aber sicherlich zu allgemein und unbestimmt, um als Definition dienen zu können. Zwar gibt es prototypische Beispiele, etwa Längen- und Gewichtsmeßverfahren. Aber auch mit solchen Beispielen läßt sich noch keine klare Abgrenzung vornehmen. Ein Aspekt des Problems besteht darin, daß in der neueren Entwicklung der Sozialwissenschaften (ungefähr in den letz-

[3] Natürlich gab es Gründe, diese Zahlen zu erzeugen und zu publizieren; es wird sogar berichtet, daß beim Korrekturlesen einige Fehler gefunden worden sind. Die Publikation sollte in einer Zeit, als Computer noch nicht weit verbreitet waren, Zufallszahlen verfügbar machen. Rao [1995, S. 32ff] berichtet über einige ähnliche Publikationen, die bereits früher erschienen sind.

[4] Unabhängig davon bleibt allerdings die Aufgabe bestehen, die *Anwendung* eines Meßverfahrens für einen bestimmten wissenschaftlichen Zweck zu begründen.

[5] Diese Argumentation hat eine lange philosophische Tradition; in neuerer Zeit ist sie insbesondere von einigen Vertretern einer „konstruktivistischen“ Wissenschaftstheorie ausgearbeitet worden. Als Einstieg in die Diskussion vgl. Janich [1979].

[6] Dies impliziert, daß Begründungen von Meßverfahren sich an Zwecken orientieren müssen und insofern nicht „wertfrei“ sein können. Die Frage, ob diese Betrachtungsweise nicht nur für Meßverfahren im engeren Sinne, sondern für alle Prozesse sozialwissenschaftlicher Datenerhebung angewendet werden kann/sollte, wird oft diskutiert; vgl. z.B. Root [1993, S.124ff].

ten 50 Jahren) eine fast grenzenlose Ausweitung des Begriffs „Messen“ stattgefunden hat. In einer Einführung in die Methoden der empirischen Sozialforschung heißt es zum Beispiel:

> „Wenn Daten über gesellschaftliche Phänomene das Ergebnis begrifflich strukturierter, systematischer und kontrollierter Beobachtungen der Merkmalsausprägungen von Untersuchungseinheiten auf den Merkmalsdimensionen darstellen, dann ist Messen in einem sehr allgemeinen Sinn zunächst nichts anderes als der Prozeß dieses Beobachtens selbst; also die Datenerhebung.“ (Mayntz, Holm und Hübner [1969, S. 36])

Hier wird nicht einmal versucht, eine begriffliche Unterscheidung zwischen „beobachten“ und „messen“ festzuhalten. Die meisten neueren Autoren verlangen zumindest, daß von „Messen“ nur dann gesprochen werden sollte, wenn numerische Aussagen gemacht werden. Hier sind einige exemplarische Formulierungen:

> Measurement „may be defined, in general, as the assignment of numbers to represent properties.“ (Campbell [1921], S. 1797) — „Paraphrasing N. R. Campbell [...], we may say that measurement, in the broadest sense, is defined as the assignment of numerals to objects or events according to rules.“ (Stevens [1946], S. 677) — „Today, measurement in general is taken to be the assignment of numbers (numerals, say the nominalists) to entities and events to represent their properties.“ (Savage & Ehrlich [1992], S. 1) — „Messen ist die Bestimmung der Ausprägung einer Eigenschaft eines Dinges. Messen erfolgt durch eine Zuordnung von Zahlen zu Dingen, die Träger der zu messenden Eigenschaft sind.“ (Orth [1974], S. 13)

In gewisser Weise kann man sich mit diesen sehr allgemein gehaltenen Formulierungen zufrieden geben, denn sie klären zumindest einen möglichen Sprachgebrauch. Und vor allem verweisen sie auf das zentrale Problem: wie sich sinnvolle Meßverfahren begründen lassen. Denn es ist offensichtlich, daß sich beliebige unsinnige Verfahren ausdenken lassen, um Objekten und Sachverhalten nach gewissen Regeln Zahlen zuzuordnen.

2 Repräsentieren und Messen

Die Frage nach Sinn- und Gütekriterien für Meßverfahren ist allerdings kompliziert und umstritten. Orientiert man sich zunächst daran, welche Auffassungen in der Geschichte des Nachdenkens über Meßverfahren

vertreten worden sind, kann man etwas vereinfacht hauptsächlich zwei Leitideen unterscheiden. Erstens die „klassische“ Leitidee, daß „Messen“ als eine Verallgemeinerung von „Zählen“ (nämlich Zählen sinnvoll definierbarer Einheiten) verstanden werden sollte. Zweitens die Leitidee der „modernen“ Meßtheorie, daß Messen als *numerische Repräsentation* (der Beziehungen zwischen Objekten) verstanden werden kann. Um auf diese unterschiedlichen Leitideen zu verweisen, spreche ich im folgenden von der „traditionellen“ bzw. der „repräsentationalen“ Meßtheorie.[7]

Etwas quer zur Unterscheidung dieser beiden Leitideen steht die Frage, welche Rolle empirische Argumente bei der Begründung von Meßverfahren spielen können. Auf das grundsätzliche Problem wurde bereits hingewiesen: Meßverfahren sollen dazu dienen, empirische Aussagen zu gewinnen; wie soll es dann möglich sein, sie mithilfe empirischen Wissens zu prüfen? Dennoch wird von vielen Autoren in der Methodenliteratur zur empirischen Sozialforschung behauptet, daß Meßverfahren (in gewissen Grenzen) empirisch geprüft bzw. begründet werden können.[8]

Betrachten wir ein einfaches Beispiel: drei Objekte, die in der Realität in einer bestimmten Weise geordnet auftreten:

$$\omega_1 \prec \omega_2 \prec \omega_3$$

Zum Beispiel drei Dosen, die nebeneinander stehen, zuerst kommt ω_1, dann ω_2 und schließlich ω_3; das Zeichen „$\prec$“ soll diese empirische Relation symbolisieren. Man kann dann diese empirische Relation durch eine Variable repräsentieren, indem man den Objekten Zahlen zuordnet, zum Beispiel:

$$X(\omega_1) = 1,\ X(\omega_2) = 2,\ X(\omega_3) = 3$$

In diesem Fall entspricht die $<$-Relation zwischen den Werten der Variablen der $\prec$-Relation zwischen den Objekten. Aber man kann auch andere Zahlenwerte wählen, zum Beispiel

$$X(\omega_1) = 7,\ X(\omega_2) = 5,\ X(\omega_3) = 9$$

In diesem Fall entspricht die Relation zwischen den Zahlen nicht mehr der Relation zwischen den Objekten.

[7] Vgl. als Einführung in die repräsentationale Meßtheorie z.B. Orth [1974]. Einen historischen Überblick gibt Diez [1997]; kritische Überlegungen finden sich u.a. bei Adams [1966], Schönemann [1994].

[8] Dies sind in der Regel Autoren, die sich ein repräsentationales Verständnis von Messen zueigen gemacht haben, vgl. u.a. Schnell, Hill und Esser [1992, S. 142ff], Laatz [1993, S. 51ff], Kromrey [1994, S. 151ff], Diekmann [1995, S. 244ff].

Kann nun geprüft werden, und ggf. in welcher Weise, ob eine Repräsentation von Objekten durch Zahlen derart ist, daß die $<$-Relation zwischen den Zahlen der $\prec$-Relation zwischen den Objekten entspricht? Nun, die Frage kann sicherlich geprüft werden. In unserem kleinen Beispiel sieht man sofort, daß die erste Variablendefinition zu einander entsprechenden Relationen führt, die zweite jedoch nicht. Allerdings ist unklar, was damit gemeint sein könnte, daß dies *empirisch* geprüft werden kann. Denn der einzige empirische Sachverhalt in diesem Beispiel besteht darin, daß es drei Objekte gibt, die durch die $\prec$-Relation in bestimmter Weise geordnet sind. Dies prüfen wir hier nicht, sondern setzen es voraus. Wir setzen also das empirische Wissen voraus, um die genannte Frage beantworten zu können. Dann aber handelt es sich gar nicht mehr um eine empirische (durch Beobachtung zu klärende) Frage, sondern um eine rein mathematische Frage.

Die repräsentationale Meßtheorie beschäftigt sich im wesentlichen mit mathematischen Fragen dieser Art. Sie setzt Objekte und ihre empirischen Relationen als gegeben voraus und versucht dann zu klären, ob und ggf. wie den Objekten auf eine solche Weise Zahlen zugeordnet werden können, daß sich die empirischen Relationen zwischen den Objekten und die numerischen Relationen zwischen den Zahlen entsprechen. Das ist offensichtlich ein rein mathematisches Problem.

Was hat das nun mit unserer Frage zu tun, ob Meßverfahren empirisch geprüft werden können? Hier kommt es offenbar darauf an, was mit dem Wort „Meßverfahren“ gemeint sein soll. Folgen wir zunächst dem Ansatz der repräsentationalen Meßtheorie mit der Definition: Ein Meßverfahren ist ein Verfahren, mit dem einer Menge von Objekten Zahlen zugeordnet werden können, und zwar so, daß gewisse empirische Relationen zwischen den Objekten eine Entsprechung in den numerischen Relationen zwischen den ihnen zugeordneten Zahlen haben. Geht man von dieser Definition aus, sind zur Begründung eines Meßverfahrens zwei Schritte erforderlich:

- Zunächst müssen die Objekte und ihre empirischen Relationen fixiert werden, die durch das Meßverfahren erfaßt werden sollen.
- Dann muß eine numerische Repräsentation der Objekte, also eine statistische Variable mit einem numerischen Wertebereich, gefunden werden, so daß sich die empirischen Beziehungen zwischen den Objekten und die numerischen Beziehungen zwischen den ihnen zugeordneten Zahlen entsprechen.

Zu überlegen ist, ob mit empirischen Argumenten entschieden werden

kann, wie hierbei vorgegangen werden sollte. Beginnen wir mit dem zweiten Schritt. Hier sind die Objekte und ihre empirischen Beziehungen schon gegeben, und es bleibt – wie oben exemplarisch erläutert wurde – ein rein mathematisches Problem übrig. Um dieses Problem zu lösen, ist empirisches Wissen vollständig unerheblich.[9]

Wie verhält es sich beim ersten Schritt? Hier müssen die Objekte und ihre empirischen Relationen fixiert werden, die durch das Meßverfahren erfaßt werden sollen. Um was für eine Art von Problem handelt es sich? Jedenfalls nicht um ein mathematisches Problem; und es wird deshalb in der Regel von der modernen Meßtheorie (insoweit sie eine rein mathematische Theorie ist) auch gar nicht behandelt. Es handelt sich aber auch nicht um ein empirisches Problem, das durch Hinweise auf Beobachtungen gelöst werden könnte. Zwei Beispiele sollen verdeutlichen, was mit dieser Aussage gemeint ist.

1. Wir beginnen mit dem bereits angeführten Beispiel: drei Dosen, die nebeneinander stehen. In diesem Fall können wir sicherlich voraussetzen, daß jeder weiß, was Dosen sind und was es bedeutet, daß Dosen nebeneinander oder nicht nebeneinander stehen. Deshalb können wir empirisch überprüfen, ob die drei Dosen wirklich nebeneinander stehen oder nicht. Aber dafür wollten wir unser Meßverfahren nicht haben, sondern es sollte die Reihenfolge, in der die Dosen nebeneinander stehen, erfassen. Zu diesem Zweck müssen wir zunächst den Dosen Namen geben, etwa D_1, D_2, D_3. Außerdem müssen wir verstehen, was es heißt, von der Reihenfolge von Dosen zu sprechen, und dafür bedarf es einiger Festlegungen (z.B. muß festgelegt werden, daß die Dosen von links nach rechts zu betrachten sind). Haben wir uns auf diese Weise mit der erforderlichen Sprache ausgerüstet, können wir schließlich (zum Beispiel) sagen: zuerst kommt D_1, dann D_2, und dann D_3.

Nun, unser Meßverfahren soll die Dosen und ihre Reihenfolge numerisch repräsentierten. Also verwenden wir folgende Abbildung (Variable): $X(D_1) = 1, X(D_2) = 2, X(D_3) = 3$. Warum? Die Antwort scheint selbstverständlich lauten zu müssen: weil dann die räumliche Ordnung der Dosen der $<$-Relation zwischen den sie repräsentierenden Zahlen entspricht. Aber ist diese Antwort wirklich ausreichend? Ich glaube, daß sie

[9] Es sei erwähnt, daß es für dies mathematische Problem nicht immer eine Lösung gibt, d.h. nicht jedes beliebige Geflecht empirischer Relationen läßt sich eindeutig in ein entsprechendes Geflecht numerischer Relationen abbilden. Wir können allerdings immer eine mathematische Sprache finden, um die Objekte und ihre empirischen Relationen zu repräsentieren und dann mit mathematischen Begriffen über sie zu sprechen.

vielleicht ausreichend ist, um das Wort „repräsentieren“ zu verstehen, daß sie aber nicht ausreichend ist, um zu verstehen, was „messen“ heißt. Jedenfalls dann nicht, wenn wir von dem Vorverständnis ausgehen, daß Messungen dazu dienen sollen, um sagen zu können, wie Objekte beschaffen sind und wie sie sich verändern. Betrachten wir nochmal die drei Dosen. Aber jetzt kommt ein Windstoß und bringt sie durcheinander. Wir können hingehen und sie wieder in einer Reihe aufstellen, zum Beispiel: $D_1 \prec D_3 \prec D_2$. Wird dann die numerische Repräsentation, die wir vor dem Windstoß vorgenommen haben, falsch? Sicherlich nicht; sie liefert sogar immer noch eine brauchbare Information (nämlich darüber, wie die Dosen vor dem Windstoß aufgestellt waren). Aber es kann aus dieser Repräsentation nicht ermittelt werden, in welcher Ordnung sie nach dem Windstoß stehen.

Müssen wir unser Meßverfahren also immer dann ändern, wenn in den zu repräsentierenden Relationen eine Veränderung eintritt? Natürlich nicht; ein Meßverfahren soll uns gerade dazu dienen, Unterschiede und Veränderungen zu erfassen. Wir müssen deshalb an ihnen festhalten und versuchen, sie gegenüber Veränderungen in der zu messenden Realität relativ konstant zu halten. Dies verlangt aber, daß sich die Repräsentation der Objekte durch Zahlen immer dann verändern muß, wenn sich deren Relationen (soweit sie durch das Meßverfahren erfaßt werden sollen) verändern.

Überlegen wir uns, wie das Meßverfahren beschaffen sein sollte. Das Meßverfahren soll uns zeigen, in welcher Reihenfolge zu zwei Zeitpunkten t_1 und t_2 die drei Dosen stehen. Es soll also Meßergebnisse liefern, die sich als eine Information über die Rangplätze der Dosen interpretieren lassen; in unserem Beispiel soll es die folgenden Rangzahlen liefern:

	D_1	D_2	D_3
t_1	1	2	3
t_2	1	3	2

Wenn wir annehmen können, daß das Meßverfahren so funktioniert, können wir die Zahlenwerte, die es uns liefert, verwenden, um Aussagen über die Reihenfolge der Dosen zu den Zeitpunkten t_1 und t_2 zu treffen. Das hat aber mit der Repräsentation der empirischen Relation $\prec$ durch die numerische Relation $<$ eigentlich nichts zu tun. Die numerische Relation $<$ besteht zwischen Zahlen, und an der Ordnung der Zahlen durch diese Relation ändert sich nichts, wie auch immer sich die Realität verändert. Die Relation $2 < 3$ gilt immer, egal wie die Dosen

aufgestellt werden. Das Meßverfahren muß dagegen gewährleisten, daß den Dosen die Zahlen stets so zugeordnet werden, daß aus dem Meßergebnis „$X(D_i) < X(D_j)$“ die Schlußfolgerung gezogen werden kann, daß in der zum Zeitpunkt der Messung bestehenden Dosenreihe die Dose D_i vor der Dose D_j kommt.

Können wir überprüfen, ob das Meßverfahren so funktioniert? Die Antwort hängt jetzt wieder davon ab, wie wir das Wort „Meßverfahren“ verwenden wollen. Stellen wir uns zunächst vor, daß der praktische Meßvorgang folgendermaßen abläuft: (a) In einem ersten Schritt beobachten wir die Dosen und stellen ihre Reihenfolge fest, etwa: Zuerst kommt D_1, dann D_3, dann D_2. Dabei folgen wir den Regeln, die für das Feststellen der Reihenfolge von Dosen vorab fixiert worden sind. (b) In einem zweiten Schritt repräsentieren wir diese Feststellung durch Zahlen, d.h. geben der Variablen X die Werte $X(D_1) = 1, X(D_2) = 3, X(D_3) = 2$.

Sagen wir nun, daß Messen darin besteht, empirisch gewonnene Beobachtungen numerisch zu repräsentieren, bezieht sich das Meßverfahren nur auf den zweiten Schritt. Und wenn das Ergebnis des ersten Schrittes schon feststeht, können wir selbstverständlich prüfen, ob der zweite Schritt richtig durchgeführt worden ist; wir brauchen nur zu prüfen, ob die Regeln zur numerischen Kodierung der Variablen X richtig angewendet worden sind. Das wäre dann allerdings ein merkwürdiger Gebrauch der Worte „messen“ und „Meßverfahren“, denn der zweite Schritt ist im Prinzip völlig überflüssig. Die gesamte Information, die unser „Meßverfahren“ liefern soll, steht bereits am Ende des ersten Schrittes fest. Die numerische Repräsentation, die der zweite Schritt hinzufügt, ändert nur die sprachliche Form, in der diese Information zum Ausdruck gebracht wird.

Wenn wir uns mit dem Wort „Meßverfahren“ jedoch auf beide Schritte beziehen, besteht das Verfahren aus zwei Arten von Regeln: erstens Regeln, die festlegen, wie die Reihenfolge von Dosen empirisch zu bestimmen ist, und zweitens die Kodierungsregeln für die numerische Repräsentation der Ergebnisse des ersten Schritts. Den zweiten Schritt haben wir gerade besprochen. Im ersten Schritt handelt es sich darum, die hierfür festgelegten Regeln richtig zu befolgen; und das kann wiederum in gewissen Grenzen (z.B. durch einen zweiten Beobachter) überprüft werden. Aber an dieser Stelle muß auf eine wichtige Unterscheidung geachtet werden. Geprüft werden kann, ob das Meßverfahren richtig angewendet worden ist. Das Meßverfahren selbst, d.h. die das Verfahren konstituierenden Regeln, werden hierdurch nicht geprüft. Es wäre auch ganz sinnlos, sie prüfen zu wollen, denn durch sie wird allererst festgelegt, was es bedeuten soll, die Reihenfolge der Dosen richtig festzustellen. Haben wir ein Meßverfahren definiert, können wir prüfen, ob damit richtig – nämlich entsprechend seiner Definition – umgegangen wird. Aber es bleibt immer noch die Frage übrig, ob unsere Definition des Meßverfahrens „richtig“ ist. Da die empirischen Daten, die wir schließlich gewinnen, von der Definition der Meßverfahren abhängen, kommt dieser Frage offenbar eine wichtige Bedeutung zu. Sie kann aber nicht empirisch geprüft werden und muß deshalb anders formuliert werden. Die Frage ist hier nicht, ob Meßverfahren richtig oder falsch sind, sondern ob sie sinnvoll sind und ob durch sie wissenswerte und verläßliche Daten gewonnen werden können.[10]

2. Um zu illustrieren, was „messen“ bedeutet, wird in der neueren Methodenliteratur gelegentlich die Frage behandelt, ob bzw. wie sich Präferenzen messen lassen.[11] Um dieser Diskussion zu folgen, beziehen wir uns auf eine Person P und drei Objekte: D_1, D_2, D_3 (nicht unbedingt Dosen). Wir unterstellen, daß P Präferenzen bzgl. dieser Objekte hat: $D_i \prec D_j$ soll jetzt bedeuteten, daß P das Objekt D_i dem Objekt D_j vorzieht. Folgen wir nun dem Ansatz der repräsentationalen Meßtheorie, muß versucht werden, die drei Objekte so durch Zahlen zu repräsentieren, daß aus „$X(D_i) < X(D_j)$“ die Schlußfolgerung „$D_i \prec D_j$“ gezogen werden kann, d.h. daß P das Objekt D_i dem Objekt D_j vorzieht. Wenn dies gelingt, sind insoweit die Präferenzen von P transitiv, d.h. wenn $D_1 \prec D_2$ und $D_2 \prec D_3$, kann daraus $D_1 \prec D_3$ abgeleitet werden.

Die interessante Frage ist nun, ob und ggf. wie festgestellt werden kann, ob P's Präferenzen transitiv sind. Einige Autoren vertreten die Ansicht, daß diese Frage – und infolgedessen auch die Frage der ordinalen Meßbarkeit von Präferenzen – empirisch geprüft und entschieden werden kann.[12] Aber wie kann das erreicht werden? Man kann zunächst die Person P für alle Paare von Objekten fragen, welches der Objekte sie

[10] Zum besseren Verständnis ist vielleicht folgende Überlegung hilfreich. Wir stellen uns vor, daß ein Meßapparat konstruiert werden soll, der die Ordnung der drei Dosen automatisch erfaßt. Wir könnten vielleicht einen Computer mit geeigneten Sensoren versehen und dann so programmieren, daß er stets drei Zahlen liefert, die der momentanen Ordnung der Dosen entsprechen. Können wir dann prüfen, ob dieser Meßapparat richtig funktioniert? Sicherlich können wir prüfen, ob sich der Apparat so verhält, wie er es aufgrund unserer Festlegungen tun sollte. Aber wie auch immer die Antwort ausfällt, sie sagt uns nur etwas über den Meßapparat, nicht jedoch darüber, ob unsere Festlegungen ein sinnvolles und zweckmäßiges Meßverfahren begründen.

[11] Vgl. Schnell, Hill und Esser [1992, S. 146], Diekmann [1995, S. 245ff].

[12] Vgl. die in der vorangegangenen Fußnote erwähnten Verweise.

vorzieht. Lautet die Antwort

$$D_1 \prec D_2,\ D_2 \prec D_3,\ D_1 \prec D_3$$

sind ihre Präferenzen transitiv, also ordinal meßbar. Antwortet sie jedoch

$$D_1 \prec D_2,\ D_2 \prec D_3,\ D_3 \prec D_1$$

sind sie nicht transitiv, also nicht ordinal meßbar. Aber müssen wir eine solche Antwort akzeptieren, oder können wir P darauf hinweisen, daß sie sich bei der Beantwortung unserer Frage geirrt hat? Wir könnten versuchen, das herauszufinden, indem wir sie nicht fragen, sondern ihr Verhalten in entsprechenden Wahlsituationen beobachten. Aber das liefert keine Lösung. Angenommen, wir beobachten zunächst $D_1 \prec D_2$, dann etwas später $D_2 \prec D_3$, und schließlich $D_3 \prec D_1$. Haben wir dann beobachtet, daß P nicht-transitive Präferenzen hat? Oder sollten wir stattdessen die Schlußfolgerung ziehen, daß P ihre Präferenzen geändert hat?

Diese Frage kann offensichtlich nicht durch Beobachtungen entschieden werden. Zunächst muß vielmehr geklärt werden, wie mit dem Wort „Präferenzen“ umgegangen werden soll. In handlungspraktischen Diskursen, in denen sich Akteure über Handlungsmöglichkeiten verständigen, ist meistens klar, daß sich das Wort nicht auf empirische Fakten bezieht (die beobachtet werden könnten), sondern auf reflektierbare und veränderbare Bewertungen von Handlungsalternativen. Die Frage der Transitivität spielt in diesen Diskursen in der Regel gar keine Rolle.[13] Wenn demgegenüber Sozialwissenschaftler Präferenzen empirisch ermitteln wollen, muß zunächst ein Bruch mit diesen handlungspraktischen Diskursen vollzogen werden; und stattdessen muß ein Verfahren festgelegt werden, durch das das Wort „Präferenzen“ als ein Name für empirische Fakten bestimmter Art *definiert* wird. Ein solches Verfahren muß dann insbesondere festlegen, ob die Transitivitätsannahme für das sozialwissenschaftliche Reden über Präferenzen konstitutiv sein soll oder nicht. Je nachdem erhält das Wort „Präferenzen“ eine unterschiedliche Bedeutung und folgen gewisse Implikationen für die praktische Ausgestaltung des Erhebungsverfahrens. Wird zum Beispiel festgelegt, daß Präferenzen transitiv sein sollen, und liefert dann eine Befragungsperson in dieser Hinsicht inkonsistente Auskünfte, dann impliziert das Verfahren, daß sie sich entweder geirrt oder ihre Präferenzen verändert hat.

[13] Da Präferenzen stets geändert werden können, erscheint es auch sinnlos, ihre Transitivität als eine Rationalitätsnorm zu fordern.

Dies Beispiel zeigt also erneut, daß der Glaube, Meßverfahren empirisch überprüfen zu können, problematisch ist. Ein Meßverfahren soll (das ist unstrittig) Zahlen liefern, so daß aus numerischen Relationen zwischen den Zahlen *empirische Aussagen* über das Bestehen oder Nichtbestehen von Relationen im Gegenstandsbereich des Meßverfahrens gewonnen werden können. Das Meßverfahren liefert jedoch nicht bloß eine numerische Repräsentation objektiv existierender empirischer Relationen, sondern seine wesentliche Bedeutung liegt zunächst darin, Regeln für die Ermittlung empirischer Fakten zu begründen.

3 Zählen und Messen

Es sollte deutlich geworden sein, daß die Idee, Messen als numerische Repräsentation empirisch *gegebener* Sachverhalte und Relationen aufzufassen, unzureichend ist und die konstruktiven Aspekte des Messens eher verschleiert als aufklärt. Es ist deshalb angebracht, auch noch einen Blick auf die traditionelle Meßtheorie zu werfen, die nicht von der Idee der Repräsentation, sondern vom Zählen ausgeht und darin den Prototyp des quantifizierenden Messens zu sehen versucht.[14]

Dieser Weg hat zugleich den Vorteil, von einem allgemein verständlichen Ausgangspunkt auszugehen. Jeder versteht, was es heißt, eine Menge von Dingen zu zählen, um schließlich sagen zu können: diese Menge besteht aus N Dingen (Einheiten, Elementen). Dadurch wird auf elementare Weise ein Größenbegriff begründet; genauer gesagt: eine Familie von Größenbegriffen, denn je nachdem, welche Art von Einheiten gezählt wird, entsteht ein spezifischer Größenbegriff. Wir werden im folgenden von *Zählgrößen* sprechen.

Zählgrößen. Im weiteren betrachten wir zunächst Objekte, die aus zählbaren Teilen bestehen; zum Beispiel Urnen, die Kugeln enthalten, oder Unternehmen, bei denen sich die Anzahl ihrer Beschäftigten zählen läßt.

[14] Die erste systematische Darstellung hat H. von Helmholtz [1887] gegeben. Sehr ähnliche Überlegungen finden sich bei Duhem [1908], Kap. 5. Die Formulierung „traditionelle Meßtheorie“ soll hier nur einen Kontrast zur modernen Meßtheorie bezeichnen, die im wesentlichen axiomatisch vorgeht, allerdings in der wissenschaftstheoretischen Diskussion meistens mit einem empiristischen Theorieverständnis verknüpft wird. Von Vertretern der modernen Meßtheorie wird Helmholtz oft als einer ihrer Ausgangspunkte genannt, dessen Ansatz durch die moderne Theorie „verallgemeinert“ worden sei; vgl. z.B. Narens und Luce [1986]. Ich finde es demgegenüber aufschlußreicher, den konzeptionellen Unterschied zwischen „Messen“ (als Verallgemeinerung von Zählen) und „Repräsentieren“ (einer irgendwie gegebenen Empirie) zu verfolgen.

Diese Objekte sind insoweit selbst Gesamtheiten. Wir repräsentieren sie wie bisher durch die Buchstaben $\omega_1, \omega_2, \ldots, \omega_N$. Die Menge Ω, die aus der Zusammenfassung dieser Objekte entsteht, ist dann eine Gesamtheit, deren Elemente selbst wieder Gesamtheiten sind. Mit $G_E(\omega)$ wird die Anzahl der Elemente bezeichnet, aus denen das Objekt ω besteht. Der Index E soll andeuten, daß wir uns auf eine bestimmte Zählgröße beziehen; z.B. Anzahl Kugeln oder Anzahl Beschaftigte.

Sobald für eine Menge von Objekten eine Zählgröße definiert worden ist, kann man mit ihrer Hilfe die Objekte vergleichen. Dabei kann man drei unterschiedliche Aspekte ins Auge fassen.

a) Im einfachsten Fall kann man sich auf die Frage beziehen, ob zwei Objekte gleich oder ungleich sind. Haben wir eine Zählgröße G_E, kann mit ihrer Hilfe definiert werden, was es heißen soll, zwei Objekte als gleich oder ungleich zu bezeichnen, nämlich:

$$\begin{array}{ll} \omega_i \equiv_E \omega_j & \text{wenn } G_E(\omega_i) = G_E(\omega_j) \\ \omega_i \not\equiv_E \omega_j & \text{wenn } G_E(\omega_i) \neq G_E(\omega_j) \end{array}$$

Wir verwenden hier das Zeichen „$\equiv_E$“, um anzudeuten, daß es sich nicht um eine Gleichheit zwischen mathematischen Objekten handelt, sondern um eine Gleichheit zwischen empirischen Objekten; und zwar im Hinblick auf die Eigenschaft E, die durch die Zählgröße G_E erfaßt wird. Durch $\equiv_E$ werden zwei Objekte dann als gleich behandelt, wenn sie die gleiche Anzahl von Elementen der Art E enthalten.

b) Eine andere Art des Vergleichens wird möglich, wenn es gelingt, die Objekte zu ordnen. Ist eine Zählgröße G_E gegeben, kann eine solche Ordnung leicht erreicht werden. Zunächst kann man durch Rückgriff auf die für Zählgrößen definierten Relationen $<$ und $\leq$ entsprechende empirische Relationen für die Objekte definieren, nämlich:

$$\begin{array}{ll} \omega_i \prec_E \omega_j & \text{wenn} \quad G_E(\omega_i) < G_E(\omega_j) \\ \omega_i \preceq_E \omega_j & \text{wenn} \quad G_E(\omega_i) \leq G_E(\omega_j) \end{array}$$

Die $\preceq_E$-Relation kann dann verwendet werden, um die Objekte in eine lineare Ordnung zu bringen:[15]

$$\omega_{(1)} \preceq_E \omega_{(2)} \preceq_E \omega_{(3)} \preceq_E \cdots \preceq_E \omega_{(N)}$$

[15] Durch die Verwendung von Klammern für die Indizes der Objekte soll angedeutet werden, daß wir im allgemeinen neue Indizes verwenden müssen, um die Objekte in die gewünschte Reihenfolge zu bringen.

Allerdings ist diese Ordnung nicht eindeutig, wenn es zwei im Hinblick auf E gleiche Objekte gibt. Für den Vergleich der Objekte ist es deshalb zweckmäßig, eine *Rangordnung* zu konstruieren. Zu diesem Zweck wird eine größte Teilmenge von Ω verwendet, die nur aus unterschiedlichen Objekten besteht. Die Objekte dieser Teilmenge können dann in eine strikte und eindeutige Rangordnung gebracht werden:

$$\omega_{[1]} \prec_E \omega_{[2]} \prec_E \omega_{[3]} \prec_E \cdots \prec_E \omega_{[M]}$$

(Hier haben wir angenommen, daß es M unterschiedliche Objekte gibt.) Hat man eine solche Rangordnung konstruiert, kann jedes Objekt aus Ω durch seinen Platz in der Rangordnung charakterisiert werden.

c) Schließlich kann man Objekte *quantitativ* vergleichen, indem man angibt, wie groß ihre Unterschiede sind. Hierfür dient das Konzept einer *Metrik*. Eine Metrik besteht aus einer *Distanzfunktion*, die für jeweils zwei Objekte angibt, wie groß ihr Unterschied (ihre Distanz) ist. Eine Distanzfunktion ist eine Abbildung

$$d : \Omega \times \Omega \longrightarrow \mathbf{R}$$

die jeweils zwei Objekten eine Zahl, $d(\omega_i, \omega_j)$, zuordnet, die als ihre Distanz interpretiert werden kann. Es muß natürlich noch definiert werden, welche Eigenschaften diese Distanzfunktion haben soll. Von einer metrischen Distanzfunktion (oder kurz: Metrik) spricht man, wenn folgende Bedingungen erfüllt sind:

- Für beliebige Objekte $\omega_i, \omega_j \in \Omega$ ist $d(\omega_i, \omega_j) \geq 0$, d.h. die Werte der Distanzfunktion sind stets nicht negativ.
- Es gilt $d(\omega_i, \omega_j) = 0$ genau dann, wenn $\omega_i \equiv_d \omega_j$.
- Die Funktion ist symmetrisch, also $d(\omega_i, \omega_j) = d(\omega_j, \omega_i)$.
- Es gilt die Dreiecksungleichung, d.h. für jeweils drei beliebige Objekte gilt: $d(\omega_i, \omega_j) + d(\omega_j, \omega_k) \geq d(\omega_i, \omega_k)$.

Ist eine Zählgröße G_E gegeben, können mit ihr im allgemeinen beliebig viele Metriken definiert werden. Orientiert man sich am üblichen räumlichen Denken, kann man insbesondere durch

$$d_E(\omega_i, \omega_j) = |\, G_E(\omega_i) - G_E(\omega_j)\,|$$

eine Metrik definieren. Man kann sich leicht überzeugen, daß die Distanzfunktion d_E alle für eine Metrik erforderlichen Eigenschaften aufweist.

Wir bezeichnen d_E als die zur Zählgröße G_E *korrespondierende Metrik*. Ihre Definition legt fest: der quantitative Unterschied zwischen zwei Objekten (bzgl. der Zählgröße G_E) besteht in der Anzahl der Elemente (der Art E), durch die sich die beiden Objekte unterscheiden.

Das Konzept einer Metrik ist sehr allgemein; und man kann – auch durch Rückgriff auf Zählgrößen – tatsächlich beliebig viele unterschiedliche Metriken definieren. Um mithilfe einer Metrik Objekte quantitativ vergleichen zu können, muß man jedoch in jedem Fall die Metrik definieren, d.h. im einzelnen sagen, was die durch sie quantifizierten Größenunterschiede bedeuten sollen. Bei Zählgrößen liegt es nahe, die korrespondierende Metrik zu verwenden; aber es sollte betont werden, daß sie tatsächlich nur eine von vielen möglichen Definitionen bildet.

Quantitative Größen. Es bleibt zu erörtern, in welcher Weise nicht nur von Zählgrößen, sondern allgemeiner von „quantitativen Größen“ gesprochen werden soll. Der traditionelle Ansatz geht von Zählgrößen aus und gelangt dadurch zu einer allgemeinen Idee meßbarer Größen. Dies Konzept soll im folgenden kurz erläutert werden. Ausgangspunkt ist die Bezugnahme auf eine gewisse Eigenschaft E, für die ein quantitatives Meßverfahren entwickelt werden soll. Der traditionelle Ansatz verlangt im wesentlichen zwei Bedingungen, die erfüllt sein müssen, um einen quantitativen Größenbegriff für E festlegen zu können.

a) Die erste Bedingung ist, daß es für jeweils zwei Objekte möglich sein muß festzustellen, ob sie im Hinblick auf E gleich oder ungleich sind. Verwenden wir die weiter oben eingeführte Terminologie, müssen also Definitionen für die empirischen Relationen $\equiv_E$ und $\not\equiv_E$ gegeben werden.

b) Die zweite Bedingung ist, daß sich ein Objekt mit der Eigenschaft E als Einheit zum Messen von E definieren läßt. Dies impliziert mehrere Anforderungen: Es sollte bei einem als Einheit brauchbaren Objekt gute (theoretische und praktische) Gründe geben, daß es im Zeitablauf bzgl. E gleich bleibt; es sollte möglich sein, von einer Einheit bzgl. E identische Kopien bilden zu können; und schließlich sollten sich die Kopien dieser Objekte zu neuen Objekten mit der Eigenschaft E verknüpfen lassen.

Nennen wir die Einheit ϵ und die Verknüpfungsoperation $\circ_E$. Dann sind also

$$\epsilon,\ \epsilon \circ_E \epsilon,\ \epsilon \circ_E \epsilon \circ_E \epsilon,\ \text{usw.}$$

wiederum Objekte mit der Eigenschaft E, und wir können die Abkürzung

$$n \cdot \epsilon \equiv_E \underbrace{\epsilon \circ_E \epsilon \circ_E \cdots \circ_E \epsilon}_{n-\text{mal}}$$

verwenden, um Objekte zu bezeichnen, die n-mal die Größe einer Einheit haben. Als Beispiel kann man an die Längenmessung denken. Als Einheit kann man einen Stab mit der Länge 1 m festlegen. Es ist möglich, von diesem Stab Kopien mit gleicher Länge anzufertigen, und die Verknüpfungsoperation $\circ_E$ kann als Hintereinanderlegen dieser Einheitsstäbe definiert werden.

Sind die beiden Bedingungen gegeben, kann die Größe von E bei einem beliebigen Objekt ω bestimmt werden. Wir müssen nur eine Anzahl n von Einheiten finden, so daß

$$\omega \equiv_E n \cdot \epsilon$$

ist. Wenn dies gelingt, können wir sagen, daß die Größe von E beim Objekt ω n-mal die Einheit ϵ beträgt, in einer kurzen Formulierung: $G_E(\omega) = n$.

Bei Zählgrößen gibt es offenbar eine natürliche Einheit: jedes der Elemente, die gezählt werden sollen, kann gleichermäßen als eine Einheit betrachtet werden, da sie sich per Definition im Hinblick auf die zu zählende Quantität nicht unterscheiden. Dadurch ist auch unmittelbar die Bedingung der Reproduzierbarkeit von Einheiten gewährleistet; es gibt immer schon so viele Einheiten, wie man zum Zählen benötigt. Soweit es sich nicht um Zählgrößen handelt, müssen Einheiten und Möglichkeiten zu ihrer Verknüpfung durch explizite Festlegungen definiert und institutionalisiert werden.

Bei Zählgrößen hat auch die o. a. Gleichung stets eine eindeutige Lösung. Im allgemeinen ist das jedoch nicht gewährleistet. Hat man z.B. als Einheit für die Längenmessung einen Stab mit der Länge 1 m definiert, kann man damit einen Stab mit der Länge 1.5 m nicht genau messen. Immerhin kann man ein Intervall angeben: dieser Stab ist länger als 1 m und kürzer als 2 m. Es ist klar, daß die praktisch erreichbare Meßgenauigkeit von der Größe der Einheiten abhängt, die jeweils definiert und praktisch anwendbar gemacht werden können.[16] Grundsätzlich kann man sagen, daß es von der Definition der kleinsten Einheit abhängt, in welcher Weise überhaupt sinnvoll von Größenunterschieden gesprochen werden kann. Wäre zum Beispiel die kleinste gegenwärtig verwendbare Einheit zur Längenmessung ein Stab mit der Länge 1 m, könnte von Stäben mit der Länge 1.50 m gar nicht sinnvoll gesprochen werden. Letztlich entscheidet die Festlegung des Meßverfahrens darüber, in welcher Weise von Größen und ihren Unterschieden empirisch gehaltvoll gesprochen werden kann.

[16] Wir verzichten hier bewußt auf die Idee, meßbare Größen durch ein Kontinuum reeller Zahlen zu repräsentieren. Denn dies würde die unter praktischen Meßgesichtspunkten unplausible Annahme voraussetzen, daß Objekte bzgl. E beliebig klein und beliebig groß gemacht werden können. Das Problem beginnt tatsächlich bereits bei der unter (a) genannten Bedingung, daß sich empirisch die Gleichheit bzw. Ungleichheit von Objekten bzgl. E feststellen läßt. Geht man realistischerweise davon aus, daß solche Feststellungen stets nur bis zu einem gewissen Genauigkeitsgrad möglich sind, kann die empirische Gleichheit $\equiv_E$ strenggenommen nicht durch eine transitive mathematische Gleichheit repräsentiert werden. Setzen wir dagegen eine kleinste Einheit als gegeben voraus (deren Größe sich natürlich durch die Entwicklung genauerer Meßverfahren verändern kann), treten diese Probleme nicht auf und es genügen grundsätzlich die natürlichen Zahlen zur Repräsentation meßbarer Größen. Bei der praktischen Einrichtung von Meßverfahren geht man allerdings meistens von einer Hierarchie von Einheiten aus und verwendet dann zweckmäßigerweise auch noch rationale Zahlen.

Quantifizierung. Es sollte zumindest skizzenhaft deutlich geworden sein, wie beim traditionellen Ansatz Größenbegriffe für meßbare Eigenschaften auf das Zählen von Einheiten zurückgeführt werden können, so daß sich der (traditionelle) Begriff der meßbaren Größe als eine Verallgemeinerung des elementaren Begriffs der Zählgröße verstehen läßt. Im wesentlichen kann dann auch alles, was oben über Zählgrößen gesagt worden ist, auf beliebige meßbare Größen übertragen werden. Insbesondere können Objekte im Hinblick auf eine meßbare Größe in eine lineare Ordnung gebracht werden und kann zu jeder meßbaren Größe eine korrespondierende Metrik definiert werden.

Eine wichtige Frage ist, ob sich dieser traditionelle Begriff quantitativer Größen auf sinnvolle Weise verallgemeinern läßt. Es liegt nahe, hierbei von der Idee einer Metrik auszugehen. Sobald für eine Menge von Objekten eine Metrik definiert worden ist, kann man die Objekte mit ihrer Hilfe quantitativ vergleichen: für jeweils zwei Objekte liefert die Metrik eine quantitative Distanz. Allerdings sieht man auch sofort, daß auf diese Weise im allgemeinen kein Begriff einer den Objekten individuell zurechenbaren Größe gebildet werden kann.

Etwas genauer. Sei, wie oben definiert, für die Objekte einer Menge Ω eine Metrik d gegeben. Die Frage, ob sie durch eine den Objekten zurechenbare Größe ausgedrückt werden kann, könnte zunächst sehr allgemein folgendermaßen formuliert werden: Gibt es eine Funktion (Variable)

$$X : \Omega \longrightarrow \tilde{X}$$

und irgendeine Funktion g, so daß dann für alle Objekte die Gleichung

$$d(\omega_i, \omega_j) = g(X(\omega_i), X(\omega_j))$$

gelten würde? Wird die Frage so formuliert, gibt es offensichtlich beliebig viele Lösungen. Wir können beliebige Zahlen zur Repräsentation der Objekte (zur Definition von X) verwenden und dann die Funktion g einfach durch die vorausgesetzte Metrik erklären. Das ist aber nicht unbedingt, was wir wollten. Denn es handelte sich dann nur um eine beliebige mathematische Reformulierung der vorausgesetzten Metrik. Die oben formulierte Frage gewinnt ihre Bedeutung jedoch nur aus einem inhaltlichen Kontrast der Begriffe „Metrik“ und „Größe“. Eine Metrik erlaubt uns, Objekte quantitativ zu vergleichen, sagt also etwas über ihre Beziehungen aus. Von Größen wollen wir jedoch üblicherweise nur dann sprechen, wenn man sie im Hinblick auf eine den Objekten individuell zurechenbare Eigenschaft interpretieren kann und wenn sich sinnvoll von Größenunterschieden sprechen läßt.

Übersetzen wir diese Überlegung in unsere formale Sprache. Wir suchen dann nicht nach irgendeiner Funktion X, sondern nach einer solchen Funktion, die es erlaubt, sinnvoll von Größenunterschieden, d.h. von $|X(\omega_i) - X(\omega_j)|$ zu sprechen und die vorausgesetzte Metrik mithilfe solcher Größenunterschiede zu reproduzieren. Die Frage ist dann, ob wir eine Lösung für die Gleichung

$$d(\omega_i, \omega_j) = |X(\omega_i) - X(\omega_j)|$$

finden können, so daß dann d die zur Größe X korrespondierende Metrik wird. Für dieses Problem gibt es jedoch keine generelle Lösung, wie man sich anhand eines einfachen Beispiels verdeutlichen kann.

Wir betrachten zunächst eine Metrik für drei Objekte, die durch folgende Tabelle definiert ist:

	ω_1	ω_2	ω_3
ω_1	0	1	2
ω_2	1	0	1
ω_3	2	1	0

In diesem Fall ist es offenbar leicht, eine Größe X zu definieren, die das gewünschte Resultat liefert, zum Beispiel

$$X(\omega_1) = 1, X(\omega_2) = 2, X(\omega_3) = 3$$

Dann ist die Gleichung $d(\omega_i, \omega_j) = |X(\omega_i) - X(\omega_j)|$ für alle Objekte erfüllt. Aber jetzt betrachten wir eine andere Metrik, etwa

	ω_1	ω_2	ω_3
ω_1	0	1	1
ω_2	1	0	1
ω_3	1	1	0

bei der der quantitative Unterschied zwischen allen Objekten gleich groß ist. Wie auch immer man die Werte von X (als eindimensionale Variable vorausgesetzt) auf einer Zahlengeraden wählt, ihr Abstand auf der Zahlengeraden kann diese Metrik nicht liefern.

Diese Überlegung zeigt, daß die Idee einer Metrik für quantitative Vergleiche von Objekten viel allgemeiner ist als die Idee der Meßbarkeit von Objekten zurechenbaren Eigenschaften durch quantitative Größen. Beides sollte deshalb auch terminologisch unterschieden werden. Wie bereits erwähnt worden ist, wird gegenwärtig der Begriff des Messens in einer weitestgehend unspezifischen Bedeutung verwendet: Jede, im Prinzip beliebige Methode, um Objekte und ihre Beziehungen numerisch zu repräsentieren, gilt bereits als ein Meßverfahren. Bereits die bloße Klassifikation einer Menge von Objekten ist dann bereits ein Messen ihrer Eigenschaften.[17] Aus meiner Sicht ist dieser inflationäre Gebrauch des Begriffs „messen“ problematisch, da dadurch einige wichtige Unterscheidungen verdeckt werden. Terminologische Streitereien über die Worte „messen“ und „Quantifizierung“ sind zwar ziemlich müßig. Es ist jedoch sinnvoll, zumindest die folgenden Unterscheidungen im Auge zu behalten.

1. Die Verwendung mathematischer Konzepte und Strukturen zur Repräsentation empirisch ermittelbarer Sachverhalte. Dies kann, muß aber nicht mit einer Quantifizierung der Sachverhalte verbunden sein.

[17] Loether und McTavish [1974, S. 14] definieren dies sogar als den Prototyp des Messens: „Measurement, most simply stated, is a procedure for carefully classifying ‘cases’ [...] and putting them into previously defined categories of some variable.“

2. Die Verwendung von Metriken zum quantitativen Vergleich von Objekten. Dies kann man als eine minimale Bedingung dafür ansehen, daß eine quantifizierende Betrachtungsweise vorliegt.

3. Die Begründung empirisch gehaltvoller Größenbegriffe, durch die Objekten zurechenbare Eigenschaften quantitativ erfaßt werden können. Wenn und insoweit dies gelingt, kann man (i.e.S.) von „meßbaren Eigenschaften“ – und praktisch: von „messen“ – sprechen.

Man kann diese drei Stufen als eine Hierarchie betrachten. Am allgemeinsten ist der Begriff der Repräsentation. Das Konzept der Metrik umfaßt den Größenbegriff; denn mit einem Größenbegriff können im allgemeinen beliebig viele Metriken definiert werden. Jedoch kann nicht umgekehrt jede Metrik durch einen Größenbegriff ersetzt werden. Um für diesen Text terminologische Klarheit zu schaffen, werden wir von quantitativen Konzepten und Darstellungen sprechen, wenn eine Metrik vorausgesetzt wird, von „messen“ jedoch nur dann, wenn es sich im engeren Sinne um meßbare Größen handelt. „Quantifizierung“ wird also als Oberbegriff verwendet, die Begründung meßbarer Größen als ein spezieller Typ der Quantifizierung betrachtet.

4 Grenzen der Meßbarkeit

Gibt es irgendwelche Grenzen für quantifizierende Darstellungen der Realität? Versteht man unter Quantifizierung die Verwendung einer Metrik für quantitative Vergleiche von Objekten, muß diese Frage vermutlich verneint werden. Sozialwissenschaftler sind in der Lage, im Prinzip beliebige Metriken zu erfinden und mit ihrer Hilfe quantitative Vergleiche vorzunehmen. Zu fragen ist natürlich stets, ob mit einer Metrik *sinnvolle* Vergleiche angestellt werden können. Aber Sinngrenzen können nicht definitiv fixiert werden.[18]

Etwas anders verhält es sich vielleicht, wenn man nach Grenzen der Meßbarkeit fragt, womit hier (wie oben vorgeschlagen) gemeint ist: die Begründung von Größenbegriffen, mit denen Objekten zurechenbare Eigenschaften quantitativ erfaßt werden sollen. Einige kurze Anmerkungen sollen das Problem verdeutlichen.

[18]Kritische Überlegungen zur Frage der Quantifizierbarkeit sozialer Sachverhalte finden sich u.a. bei: Cicourel [1964]; Habermas [1966, S. 102 – 124]; Berger [1974]; Kreppner [1975]; Lorenzen [1975].

1. Auch in diesem Fall gibt es keine wirklich harten Kriterien. Zur Begründung meßbarer Größen sind in jedem Fall Festlegungen erforderlich. Diese Festlegungen beziehen sich grundsätzlich auf alle Aspekte des Meßverfahrens. Es muß festgelegt werden, wann zwei Objekte im Hinblick auf die zu messende Eigenschaft (E) als gleich, wann als ungleich anzusehen sind; es muß festgelegt werden, wie sie im Hinblick auf E in eine lineare Ordnung gebracht werden können; und es muß festgelegt werden, was als Einheit gelten soll und wie Einheiten verknüpft werden können. Begründungen für solche Festlegungen können sich schließlich nur an Zweckmäßigkeitsgesichtspunkten orientieren, und sie können deshalb bestritten und – wie die Geschichte der Institutionalisierung von Meßverfahren zeigt – auch immer wieder revidiert werden.

2. Daß zur Begründung meßbarer Größen (oder allgemeiner: von Metriken) Festlegungen getroffen werden müssen, kann man auch so ausdrücken, daß uns die Gegenstände, die wir in der Welt vorfinden, nicht von sich aus sagen, ob und ggf. wie gewisse ihrer Eigenschaften gemessen werden können. Insofern kann man A. Kaplan zustimmen, wenn er in seiner „Methodology for Behavioural Science“ [1964, S. 176] schreibt: „that whether we can measure something depends, not on that thing, but on how we have conceptualized it, on our knowledge of it, above all on the skill and ingenuity which we can bring to bear on the process of measurement which our inquiry can put to use [...].“ Allerdings geht es dann folgendermaßen weiter: „To say of something that it is uncapable of being measured is like saying of it that it is knowable only up to a point, that our ideas of it must inevitably remain indeterminate.“ Darüber kann sicherlich gestritten werden.[19] Die Feststellung qualitativer Unterschiede ist grundsätzlich nicht weniger legitim oder wissenswert als die Feststellung quantitativer Unterschiede. Gerade wenn man erkannt hat, daß Eigenschaften nicht an und für sich meßbar oder nicht meßbar sind, stellt sich nicht nur die Frage, wie Eigenschaften meßbar gemacht werden können, sondern primär wichtig wird die Frage, welche theoretischen und praktischen Interessen die Erfindung und Institutionalisierung neuer

[19]Eine oft wiederholte, aber auch bestrittene Formulierung dieser Position stammt von dem Physiker Kelvin aus dem Jahr 1889: „I often say that when you can measure what you are speaking about, you know something about it; when you cannot measure it [...] your knowledge is of a meagre and unsatisfactory kind.“ (Hier zitiert nach Hacking [1983, S. 242].) Das Programm eines rein quantitativen (physikalischen) Weltbilds ist jedoch wesentlich älter und geht insbesondere auf Descartes zurück. Vgl. hierzu die kurzen, aber sehr informativen Ausführungen von Duhem [1908, Kap. 5 und 6]; eine kritische historische Darstellung gibt auch Kline [1980].

Meßtechniken sinnvoll machen könnten.

3. Die Feststellung, daß Eigenschaften nicht an und für sich meßbar oder nicht meßbar sind, sondern erst durch Meßverfahren meßbar gemacht werden, hat einige wichtige Implikationen. Zunächst folgt, daß es zum Verständnis meßbarer Eigenschaften erforderlich ist, die Meßverfahren zu verstehen, durch die sie meßbar gemacht worden sind; und dazu gehört, die Begründungen und Zwecksetzungen zu verstehen, die zur Entwicklung der Meßverfahren geführt haben. Wittgenstein [1967, S. 4] hat das mit folgender Bemerkung verdeutlicht:

> „Wie würden wir mit der Wahrheit in Konflikt geraten, wenn unsere Zollstäbe aus sehr weichem Gummi wären, anstatt aus Holz oder Stahl? – 'Nun, wir würden nicht das richtige Maß des Tisches kennenlernen.' – Du meinst, wir würden nicht, oder nicht zuverlässig, *die* Maßzahl erhalten, die wir mit unsern harten Maßstäben erhalten. *Der* wäre also im Unrecht, der den Tisch mit dem dehnbaren Maßstab gemessen hätte und behauptete, er mäße 1.80 m nach unserer gewöhnlichen Meßart; sagte er aber, der Tisch mißt 1.80 m nach der seinen, so ist das richtig. – 'Aber das ist dann doch überhaupt kein Messen!' – Es ist unserem Messen ähnlich und kann unter Umständen 'praktische Zwecke' erfüllen. (Ein Kaufmann könnte auf diese Weise verschiedene Kunden verschieden behandeln.)"

Wittgenstein weist darauf hin, daß es zum Verständnis von Meßverfahren erforderlich ist, die Zwecke zu verstehen, die mit ihnen verfolgt werden sollen. Denkt man daran, wie sich unsere heutigen Meßverfahren erst in einem langen historischen, gelegentlich konfliktreichen Prozeß durchgesetzt haben, ist das sicherlich richtig.[20]

4. Eine zweite wichtige Implikation ist die folgende: Wenn wir uns zur Beschreibung von Objekten und ihrer Beziehungen auf meßbar *gemachte* Eigenschaften stützen, wird zunächst unklar, worüber wir dann sprechen. Denn der Sinn von auf diese Weise gebildeten Aussagen hängt dann sowohl von der Beschaffenheit der Objekte und ihrer Beziehungen als auch davon ab, wie ihre Eigenschaften meßbar gemacht worden sind. Denken wir an Wittgensteins Kaufmann, der die Länge eines Tisches mit einem Maßstab aus Gummi mißt und dann sagt, daß die Länge des Tisches 1.80 m beträgt. Es ist offensichtlich, daß diese Aussage nicht nur

[20] Über diesen Prozeß gibt es bereits eine ganze Reihe historischer Studien; einen guten Einstieg vermitteln die Beiträge in Wise [1995].

eine Eigenschaft des Tisches zum Ausdruck bringt, sondern sich in ihr auch ausdrückt, wie der Kaufmann die Länge des Tisches gemessen hat.

Das offenkundigste Beispiel sind Wertgrößen (Preise und Einkommen). Diese Größen genügen allen oben genannten Bedingungen für meßbare Größen. Alle Objekte, die einen Preis haben, lassen sich dadurch in eine lineare Ordnung bringen; es gibt eine Einheit (Geld), von der sich beliebig viele Kopien herstellen lassen; und diese Kopien lassen sich zu beliebigen Wertgrößen verknüpfen. Aber jeder weiß, daß diese Wertgrößen nicht durch ein objektives Meßverfahren ermittelt werden; sie werden vielmehr ausgehandelt, wobei das Ergebnis durchaus auch von den Präferenzen der beteiligten Akteure abhängt. Es ist klar, daß Aussagen über Wertgrößen nicht in erster Linie Aussagen über die Beschaffenheit von Objekten sind, sondern zum Ausdruck bringen, wie Objekte von den beteiligten Akteuren bewertet werden.

5. Man könnte einwenden, daß in diesen Fällen (Verwendung von Maßstaben aus Gummi, Ermittlung von Wertgrößen durch Marktprozesse, Schätzung subjektiver Wahrscheinlichkeiten, usw.) eine wichtige Forderung nicht erfüllt ist: daß Meßverfahren *objektiv* sein sollten; womit gemeint ist: die durch sie ermittelbaren Größen sollen nicht von der Beschaffenheit der Subjekte abhängen, die diese Meßverfahren anwenden. Aber es sollte überlegt werden, was mit einer solchen Objektivitätsforderung erreicht werden kann.

Betrachten wir ein Beispiel: eine Menge von Objekten, die sich im Hinblick auf drei Farben (rot, grün, blau) unterscheiden lassen. In unserer gewöhnlichen Sprache handelt es sich um nur qualitativ unterscheidbare Farben. Folgen wir nun der Devise von Kelvin und Kaplan, diese Eigenschaft der Objekte, eine bestimmte Farbe zu haben, meßbar zu machen, und verwenden wir dafür den allgemeinen Ansatz der modernen Meßtheorie. D.h. wir müssen eine Regel aufstellen, mit deren Hilfe wir den Objekten im Hinblick auf die zu messende Eigenschaft Zahlen zuordnen können. Das ist leicht getan. Wir definieren eine Abbildung F, die jedem Objekt (ω) eine Zahl zuordnet, etwa folgendermaßen:

$$F(\omega) = \begin{cases} 1 & \text{wenn } \omega \text{ rot ist} \\ 2 & \text{wenn } \omega \text{ grün ist} \\ 3 & \text{wenn } \omega \text{ blau ist} \end{cases}$$

Wir haben dann die Farbeigenschaften der Objekte durch Zahlen repräsentiert, allerdings aus den drei unterschiedlichen Farben noch keine meßbare Eigenschaft gemacht. Es bleibt zu überlegen, wie in diesem

Fall von einer quantitativen Größe gesprochen werden könnte, die den in Abschnitt 3 diskutierten Bedingungen genügt. In gewisser Weise kann diesen Bedingungen jedoch durch eine Reihe von Festlegungen leicht entsprochen werden. Zunächst können wir festlegen, daß zwei Objekte als (bzgl. ihrer Farbe) gleich gelten sollen, wenn sie die gleiche Farbe haben, andernfalls als ungleich. Wir können die Objekte sodann im Hinblick auf ihre Farbe in eine lineare Ordnung bringen, indem wir einfach die Ordnung der sie repräsentierenden Zahlen verwenden. Und schließlich können wir eine Einheit festsetzen, um die Größe der Farbeigenschaft zu messen. Wir können zum Beispiel die roten Objekte als Einheiten verwenden (wenn es genügend viele von ihnen gibt); sie sollen durch das Symbol ϵ repräsentiert werden. Wir brauchen dann nur noch eine Festlegung darüber, wie die Verknüpfung von Einheiten zu interpretieren ist. Wir können z.B. vereinbaren: Eine Einheit gilt als rot ($F(\epsilon) = 1$), zwei Einheiten gelten als grün ($F(\epsilon \circ \epsilon) = 2$), drei Einheiten gelten als blau ($F(\epsilon \circ \epsilon \circ \epsilon) = 3$), vier Einheiten gelten wiederum als rot, usw.[21]

Wir haben dann formal allen Anforderungen an eine quantitative Größe (wie sie oben definiert worden ist) entsprochen; und insbesondere haben wir ein *objektives* Meßverfahren festgelegt. Man kann eindeutig sagen, wie mit diesem Verfahren Farbeigenschaften zu messen sind; wird das Verfahren vorausgesetzt, hängt das Meßergebnis nur von der Farbe der Objekte, nicht auch von irgendwelchen Eigenschaften der messenden Subjekte ab. Gleichwohl kann man sich natürlich darüber streiten, ob mit diesem Verfahren beansprucht werden kann, Farbeigenschaften meßbar gemacht zu haben. Man könnte zum Beispiel einwenden: Das Verfahren ist zwar objektiv (im Sinne der oben gegebenen Definition dieses Wortes), aber dennoch „willkürlich“, denn man könnte die drei Farben auch anders anordnen und man hätte auch ein anderes Objekt als Einheit auswählen können. Der Einwand ist sicherlich richtig, aber *diesen* Einwand kann man gegen jedes Meßverfahren richten, insofern es auf Festlegungen beruht, die man auch anders treffen könnte.

Das Beispiel soll natürlich nicht zeigen, wie Farbunterschiede sinnvoll gemessen werden können. Es soll aber deutlich machen, daß es zur Begründung von Meßverfahren nicht ausreicht, den Bedingungen für Meßbarkeit formal zu genügen, sondern daß es erforderlich ist, den theoretischen und/oder praktischen Sinn eines Meßverfahrens explizit darzulegen und dadurch einer Beurteilung zugänglich zu machen. Und es soll darüberhinaus zeigen, daß diese Aufgabe nicht schon dadurch erledigt werden kann, daß man ein objektives Meßverfahren vorschlägt. Das Beispiel zeigt schließlich auch, daß ein objektives Meßverfahren keineswegs garantiert, daß Meßergebnisse vom Meßverfahren unabhängig sind. Sie sind bei einem objektiven Meßverfahren zwar unabhängig von den subjekten Eigenschaften der Personen, die die Verfahren anwenden; aber in jedem Fall kommen die Meßergebnisse *durch das Verfahren* zustande. Insofern liefert auch die Entwicklung objektiver Meßverfahren keine Lösung für das grundsätzliche Problem, daß quantifizierende Aussagen über Objekte stets nicht nur Eigenschaften der Objekte, sondern auch Eigenschaften unserer Meßverfahren zum Ausdruck bringen.

[21] Allgemein: $F(n \cdot \epsilon) = 1 + (n-1) \,\%\, 3$, wobei „%“ die Modulus-Operation bezeichnet, d.h. $m \,\%\, n$ ist der ganzzahlige Rest bei einer Division von m durch n.

Man kann sicherlich eine Reihe von Gründen dafür anführen, daß Meßverfahren für wissenschaftliche und technische Zwecke objektiv sein sollten, d.h. die Meßergebnisse sollten nicht von der Verfassung der die Meßverfahren anwendenden Subjekte abhängen. Es ist jedoch eine offene und umstrittene Frage, ob wir darin ein grundsätzliches Ideal für alle Bereiche der gesellschaftlichen Praxis sehen sollten.[22] In zumindest einem Bereich, bei der Ermittlung von Wertgrößen, scheinen auch die meisten Menschen nicht bereit zu sein, dies Ideal anzuerkennen.

5 Variablen und Skalen

Bei allen elementaren Methoden der deskriptiven Statistik geht man von statistischen Variablen aus, also von (ein- oder mehrdimensionalen) Abbildungen

$$X : \Omega \longrightarrow \tilde{X}$$

die den Objekten einer Gesamtheit Ω Merkmale aus einem Merkmalsraum $\tilde{X}$ zuordnen. Merkmale können immer durch Zahlen repräsentiert werden, $\tilde{X}$ ist dann eine Menge von Zahlen, und der Merkmalsraum wird dann als eine *Skala* bezeichnet.

Zahlen müssen zwar nicht, aber sie können etwas bedeuten. Ob und ggf. was eine Zahl bedeutet, kann man ihr im allgemeinen nicht ansehen; es muß also explizit gesagt werden. Bei der Anwendung statistischer Methoden müssen dann zwei Aspekte geklärt werden. Erstens muß die Variable spezifiziert werden, als deren Werte die Zahlen verstanden werden

[22] Einen guten Einstieg in die neuere Diskussion vermitteln die Beiträge in Megill [1994].

sollen. Zweitens muß gesagt werden, welche Eigenschaften der Objekte durch die Zahlen repräsentiert werden sollen.

Obwohl dies selbstverständlich klingt, gibt es darüber eine Kontroverse. Einige Autoren sind der Auffassung, daß es bei vielen statistischen Methoden gleichgültig ist, ob die von ihnen verwendeten Zahlen überhaupt irgendeine Bedeutung haben. Sie betrachten dann statistische Methoden als rein mathematische Algorithmen, die aus einer gewissen Menge von Eingabezahlen eine andere, in der Regel wesentlich kleinere Menge von Ausgabezahlen erzeugen. Wir werden hier dieser Betrachtungsweise jedoch nicht folgen, sondern annehmen, daß mithilfe statistischer Methoden nicht Aussagen über Zahlen, sondern Aussagen über Gesamtheiten von Objekten gewonnen werden sollen. Diese Auffassung impliziert, daß statistische Aussagen eine Bedeutung haben sollen, die durch Bezugnahme auf Objekte und ihre Eigenschaften und Beziehungen expliziert werden muß.

Diese Auffassung erzeugt jedoch für die statistische Methodenlehre ein gewisses Dilemma. Die Bedeutung der durch statistische Variablen gelieferten Zahlenwerte ist außerordentlich vielfältig; und es ist weder möglich noch sinnvoll, jeweils unterschiedliche Methoden zu entwickeln. Es ist deshalb üblich, sich nicht auf spezifische Bedeutungen von Variablen zu beziehen, sondern eine gewisse schematische Unterscheidung unterschiedlicher Variablenarten vorauszusetzen. Die heutzutage gebräuchliche Einteilung wurde zuerst von S. S. Stevens [1946] vorgeschlagen und unterscheidet folgende Skalentypen.

- Nominalskalierte Variablen. Die Zahlen auf der Skala (im Wertebereich) dieser Variablen repräsentieren qualitativ unterschiedliche Merkmale der Objekte. Diese Variablen, auch „qualitative“, „klassifikatorische“ oder „kategoriale“ Variablen genannt, erlauben es, Objekte im Hinblick auf qualitativ unterschiedliche Merkmale zu unterscheiden.

- Ordinalskalierte Variable. Mit den Zahlenwerten dieser Variablen, oft kurz „ordinale Variablen“ genannt, können Objekte nicht nur unterschieden, sondern außerdem in eine lineare Reihenfolge gebracht werden. Sie setzen also die Existenz einer empirischen Relation $\prec$ zwischen den Objekten voraus, so daß aus der numerischen Relation $X(\omega_i) < X(\omega_j)$ die Schlußfolgerung $\omega_i \prec \omega_j$ gezogen werden kann.

- Absolutskalierte Variablen. In diesem Fall können die Zahlenwerte der Variablen als Zählgrößen interpretiert werden, wie in Abschnitt 3 besprochen worden ist. Es handelt sich also um quantitative Größen, zu deren genauer Bestimmung eine Zähleinheit angegeben werden muß.

- Intervallskalierte Variablen. Hierbei handelt es sich um Variablen, bei denen angenommen wird, daß eine korrepondierende Metrik sinnvoll interpretiert werden kann. Man kann dann $|X(\omega_i) - X(\omega_j)|$ als einen Größenunterschied interpretieren. Insbesondere sind alle absolutskalierten Variablen auch intervallskaliert.

- Ratioskalierte Variablen. Dieser Variablentyp wird häufig so erklärt, daß es sich um intervallskalierte Variablen handelt, bei denen sich außerdem ein „natürlicher Nullpunkt“ definieren läßt.[23] Die Unterscheidung zwischen einem „natürlichen“ und einem „nicht-natürlichen“ Nullpunkt ist allerdings fragwürdig. Eine bessere und praktisch vollständig ausreichende Definition kann folgendermaßen gegeben werden: daß es sich um intervallskalierte Variablen handelt, die nur positive Werte annehmen können. Das ist insbesondere bei allen quantitativen Größen, wie sie in Abschnitt 3 definiert worden sind, der Fall. Man beachte, daß durch diese Definition ausgeschlossen wird, daß eine ratioskalierte Variable den Wert Null annehmen kann.[24]

Wir werden häufig von *metrischen* oder auch *quantitativen* Variablen sprechen, wenn es sich um intervall-, ratio- oder absolutskalierte Variablen handelt, da in allen diesen Fällen eine korrespondierende Metrik für quantitative Vergleiche zur Verfügung steht.

Die Unterscheidung verschiedener Skalentypen wurde von Stevens [1946] mit einer weitergehenden Überlegung verknüpft: daß es zu jedem Skalentyp eine Menge zulässiger Skalentransformationen gibt. Die Idee ist, daß es mehr oder weniger willkürlich ist, welche Zahlenwerte zur numerischen Repräsentation von Eigenschaften und Beziehungen von Objekten verwendet werden. Um zum Beispiel drei unterschiedliche Farben numerisch zu repräsentieren, kann man die Zahlen $1, 2, 3$ verwenden. Aber ebensogut könnte man drei andere Zahlen nehmen. Man muß in jedem Fall nur festlegen, welche Farbe durch eine Zahl repräsentiert werden soll. Folgt man diesem Gedanken, scheint es für jeden Skalentyp zulässige und unzulässige Skalentransformationen zu geben.

- Bei nominalskalierten Variablen kann man beliebige Zahlen verwen-

[23] Vgl. z.B. Schnell, Hill, Esser [1992, S. 150]; Diekmann [1995, S. 253].

[24] Sie erspart deshalb die Frage, was damit gemeint sein könnte, daß eine ratioskalierte Variable den Wert Null tatsächlich annimmt. Schnell, Hill und Esser [1992, S. 150], die von der zuerstgenannten Definition ausgehen, sagen hierzu: „Der Meßwert ‘Null’ entspricht der tatsächlichen Abwesenheit des gemessenen Merkmals.“ Aber wie kann man nichts messen? Wie kann man z.B. messen, daß eine schlechte Definition das Gewicht Null hat?

den, um die Merkmale zu unterscheiden.

• Bei ordinalskalierten Variablen kann man jede beliebige Skala verwenden, die es erlaubt, aus der numerischen Relation $X(\omega_i) < X(\omega_j)$ die Schlußfolgerung $\omega_i \prec \omega_j$ zu ziehen.

• Bei absolutskalierten Variablen ist keinerlei Skalentransformation zulässig. Da es sich um Zählgrößen handelt, müssen stets die natürlichen Zahlen verwendet werden.

• Bei intervallskalierten Variablen wird angenommen, daß alle linearen Skalentransformationen zulässig sind, also alle Abbildungen

$$g : \tilde{X} \longrightarrow \tilde{X}^*$$

der Art

$$g(\tilde{x}) = a + b\,\tilde{x}$$

wobei a eine beliebige und b eine beliebige positive Zahl ist. Dann ist

$$|\, g(X(\omega_i)) - g(X(\omega_j))\,| = b\,|\, X(\omega_i) - X(\omega_j)\,|$$

und die korrespondierenden Metriken sind proportional. Als Begründung wird meistens angeführt, daß bei intervallskalierten Variablen die Wahl der Einheit willkürlich ist. Aber diese Begründung ist nicht vollständig befriedigend. Wenn es eine Einheit gibt (was bei quantitativen Größen vorausgesetzt werden sollte), muß man sie kennen, um die Bedeutung der Zahlenwerte zu verstehen. Sobald man aber eine Einheit fixiert hat, sind nur noch additive Skalentransformationen ($b = 1$) zulässig. Dann erhält man auch identische Metriken. Die gleiche Überlegung kann für ratioskalierte Variablen angestellt werden. Hier sind dann ebenfalls alle additiven Skalentransformationen zulässig.

Die Idee „zulässiger“ Skalentransformationen ist in der statistischen Methodenlehre umstritten, insoweit sie oft mit der Auffassung verknüpft wird, daß es vom Skalenniveau der Daten abhänge, welche statistischen Methoden zu ihrer Analyse verwendet werden dürfen.[25] Der Kern der Kritik kann etwa folgendermaßen zusammengefaßt werden: Die für statistische Methoden vorauszusetzenden Daten haben kein „intrinsisches“ Skalenniveau, sondern dies sei abhängig von der Bedeutung, *die den Daten gegeben wird*; deshalb schreiben uns die Daten nicht vor, was wir mit ihnen tun dürfen. So wird zum Beispiel gesagt, daß ordinalskalierte Daten häufig auch als intervallskalierte Daten betrachtet werden können und dann auch alle statistischen Verfahren, die eine Metrik voraussetzen, angewendet werden können.

[25] Eine Zusammenstellung kritischer Positionen haben Velleman und Wilkinson [1994] gegeben; eine gründliche systematische Diskussion findet sich bei Adams et al. [1965].

Der entscheidende Punkt liegt in der Feststellung, daß das Skalenniveau von der Bedeutung abhängt, die den Daten gegeben wird. Das ist in gewisser Weise offensichtlich. Trotzdem ist die Formulierung problematisch, da sie unterstellt, daß von Daten gesprochen werden kann, die keine bestimmte Bedeutung haben, sondern denen erst nachträglich – im Prozeß ihrer statistischen Bearbeitung – die eine oder andere Bedeutung gegeben werden kann.[26] Demgegenüber kann man durchaus die Position vertreten, daß Daten sich gerade durch ihre Bedeutung von reinen Zahlen unterscheiden, und daß daher Daten und ihre Bedeutung nicht voneinander getrennt werden sollten und die Fixierung eines Skalenniveaus ein wesentlicher Bestandteil des Prozesses sein sollte, durch den Daten – d.h. ihre Bedeutung – definiert werden. Diese Position könnte nicht nur dazu dienen, die Beliebigkeit der Dateninterpretation *während* des Prozesses ihrer statistischen Analyse einzuschränken, sondern könnte auch nochmal betonen, daß Annahmen über Metriken und insbesondere über quantitative Größen einer Begründung bedürfen.

[26] Die Kritik an der Idee zulässiger Skalentransformationen stammt in der Regel von „reinen“ Statistikern, die Daten in erster Linie nur als numerischen Input für statistische Methoden betrachten.

Literaturverzeichnis

Adams, E. W. [1966]. On the Nature and Purpose of Measurement. Synthese 16, 125 – 169.

Adams, E. W., Fagot, R. F., Robinson, R. E. [1965]. A Theory of Appropriate Statistics. Psychometrika 30, 99 – 127.

Berger, H. [1974]. Untersuchungsmethode und soziale Wirklichkeit. Eine Kritik an Interview und Einstellungsmessung in der Sozialforschung. Frankfurt: Suhrkamp.

Campbell, N. [1921]. Measurement. In: J. R. Newman (ed.), The World of Mathematics, Vol. 3, 1797 – 1813. New York: Simon and Schuster 1956.

Cicourel, A. V. [1964]. Method and Measurement in Sociology. New York: Free Press. Dt. Übers.: Methode und Messung in der Soziologie. Frankfurt: Suhrkamp 1970.

Diekmann, A. [1995]. Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen. Reinbek: Rohwolt.

Diez, J. A. [1997]. A Hundred Years of Numbers. An Historical Introduction to Measurement Theory 1887 – 1990. Part I: The Formation Period. Studies in History and Philosophy of Science 28, 167 – 185. Part II: Suppes and the Mature Theory. Studies in History and Philosophy of Science 28 237 – 265.

Duhem, P. [1908]. Ziel und Struktur der physikalischen Theorien. Hamburg: Felix Meiner 1978.

Habermas, J. [1966]. Zur Logik der Sozialwissenschaften. Sonderheft der Philosophischen Rundschau.

Hacking, I. [1983]. Representing and Intervening. Cambridge: Cambridge University Press.

Helmholtz, H. von [1887]. Zählen und Messen erkenntnistheoretisch betrachtet. Neudruck Darmstadt: Wiss. Buchgesellschaft 1959.

Janich, P. [1979]. Möglichkeiten und Grenzen quantitativer Methoden. In: J. Mittelstraß (Hg.), Methodenprobleme der Wissenschaften vom gesellschaftlichen Handeln, S. 370 – 383. Frankfurt: Suhrkamp.

Kline, M. [1980]. Mathematics. The Loss of Certainty. Oxford: Oxford University Press.

Kreppner, K. [1975]. Zur Problematik des Messens in den Sozialwissenschaften. Stuttgart: Klett.

Kromrey, H. [1994]. Empirische Sozialforschung (6. Auflage). Opladen: Leske und Budrich.

Laatz, W. [1993]. Empirische Methoden. Ein Lehrbuch für Sozialwissenschaftler. Thun/Frankfurt: Harri Deutsch.

Lave, J. [1986]. The Values of Quantification. In: J. Law (ed.), Power, Action and Belief. A New Sociology of Knowledge? pp. 99 – 111. London: Routledge & Kegan Paul.

Loether, H. J., McTavish, D. G. [1974]. Descriptive Statistics for Sociologists. Boston: Allyn and Bacon.

Lorenzen, P. [1975]. Autonomie und empirische Sozialforschung. In: J. Mittelstraß (Hg.), Methodologische Probleme einer normativ-kritischen Gesellschaftstheorie, S. 244 – 265. Frankfurt: Suhrkamp.

Mayntz, R., Holm, K., Hübner, P. [1969]. Einführung in die Methoden der empirischen Soziologie. Opladen: Westdeutscher Verlag.

Megill, A. (ed.) [1994]. Rethinking Objectivity. Durham and London: Duke University Press.

Narens, L., Luce, R. D. [1986]. Measurement: The Theory of Numerical Assignments. Psychological Bulletin 99, 166 – 180.

Newman, J. R. (ed.) [1956]. The World of Mathematics (4 vol.). New York: Simon and Schuster.

Orth, B. [1974]. Einführung in die Theorie des Messens. Stuttgart: Kohlhammer.

Pawson, R. [1982]. Desperate Measures. British Journal of Sociology 33, 35 – 63.

Rao, C. R. [1995]. Was ist Zufall? Statistik und Wahrheit. München: Prentice Hall.

Root, M. [1993]. Philosophy of Social Science. The Methods, Ideals, and Politics of Social Inquiry. Oxford: Blackwell.

Savage, C. W., Ehrlich, P. (eds.) [1992]. Philosophical and Foundational Issues in Measurement Theory. Hillsdale: Lawrence Erlbaum.

Schnell, R., Hill, P. B., Esser, E. [1992]. Methoden der empirischen Sozialforschung (3. Aufl.). München: Oldenbourg.

Schönemann, P. H. [1994]. Measurement: The Reasonable Ineffectiveness of Mathematics in the Social Sciences. In: I. Borg, P. P. Mohler (Hg.), Trends and Perspectives in Empirical Social Research, S. 149 – 160. Berlin: de Gruyter.

Stevens, S. S. [1946]. On the Theory of Scales of Measurement. Science 103 (No. 2684), 677 – 680.

Velleman, P., Wilkinson, L. [1994]. Nominal, Ordinal, Interval and Ratio Typologies are Misleading. In: I. Borg, P. P. Mohler (Hg.), Trends and Perspectives in Empirical Social Research, S. 161 – 177. Berlin: de Gruyter.

Wegener, B. [1988]. Kritik des Prestiges. Opladen: Westdeutscher Verlag.

Wise, M. N. (ed.) [1995]. The Values of Precision. Princeton: Princeton University Press.

Wittgenstein, L. [1967]. Bemerkungen über die Grundlagen der Mathematik. Dt.-engl. Ausgabe, hrsg. von G. H. von Wright, R. Rhees, G. E. M. Anscombe. Oxford: Basil Blackwell.