

Brennpunkte in der Messtheorie

Peter Jaenecke

Messtheorie wird hier verstanden als Teil der Wissenschaftstheorie,
und diese wiederum als Teil der (theoretischen) Philosophie, allerdings mit dem Anspruch, methodische Hilfen für die wissenschaftliche Forschungspraxis zu liefern.

Es handelt sich also (1) bei der hier vorgestellten Untersuchungen nicht um welche philosophischen Reflexionen, etwa über die „Beobachtung“, sondern um Anleitungen für die messtheoretische Praxis aus meta-wissenschaftlicher Sicht.

Die Messtheorie ist ein Stiefkind in der Wissenschaftstheorie, wenn überhaupt, dann werden meist sehr einfache physikalische Messungen betrachtet: 'Einfach' heißt: Wechselwirkungen werden vernachlässigt.

Das führt auf die Frage:

Gibt es überhaupt eine einheitliche Messtheorie für alle Wissenschaften oder sind einzelwissenschaftliche Messtheorien erforderlich?

Antwort

Wenn Wissenschaften wirklich eigene Messprinzipien erforderten, hieße dies, sie verfügten über nur ihnen allein zukommende Erkenntnisweisen.

Das wiederum bedeutete: es gibt isolierte empirische Bereiche – eine unakzeptable Konsequenz.

Will man die Einheit der Welt aufrechterhalten,
kann man keine fachspezifischen
Messprinzipien zugestehen.

Es müssen also (2) Messprinzipien vorgestellt
werden, mit dem Anspruch, dass sie für *alle*
Wissenschaften verbindlich sind.

Die These von der Einheit der Welt ist eine meta-wissenschaftliche, philosophische (ontologische) Aussage; sie gehört zur „Meta“physik.

Gleichwohl lenkt sie die Überlegungen in eine bestimmte Richtung und beeinflusst dadurch oft unbemerkt unsere Handlungen.

Das Beispiel zeigt die Verzahnung von Wissenschaft und Philosophie: Jede Einzelwissenschaft stützt sich auf allgemeine Aussagen, die über die Einzelwissenschaft hinausgehen und dort auch nicht gerechtfertigt werden können.

Sie gehen in die Einzelwissenschaften als *stillschweigende Voraussetzungen* ein.

Stillschweigende Voraussetzungen können jedoch problematisch sein:

Es ist unter anderem Aufgabe der Wissenschaftstheorie, hier insbesondere der Messtheorie, solche versteckten Voraussetzungen sowie ihre Folgen ans Licht zu bringen.

Es gibt also bei wissenschaftstheoretischen Untersuchungen immer eine philosophische und eine technische Komponente; beide müssen miteinander vereinbar sein, d.h. die allgemeinen Voraussetzungen gehen zwar über die Einzelwissenschaften hinaus, dürfen aber dort nicht der Erfahrung widersprechen.

Sie müssen daher überprüft werden, und das gelingt nur, wenn sie ihren stillschweigenden Status verlieren: sie müssen – auch wenn sie als selbstverständlich erachtet und dadurch oft übersehen werden – explizit gemacht werden.

Dieser Aufgabe werden wir im Folgenden unsere besondere Aufmerksamkeit widmen.

- A Daten im Alltag
- B Messtheoretisches Basiswissen
- C Messtheorie
 - C₁ Messprozesse
 - C₂ Metrisierung
 - C₃ Fehlerrechnung
 - C₄ Interpretation
- D Resümee

- A Daten im Alltag
- B Messtheoretisches Basiswissen
- C Messtheorie
 - C₁ Messprozesse
 - C₂ Metrisierung
 - C₃ Fehlerrechnung
 - C₄ Interpretation
- D Resümee

Daten sind in unserer verwissenschaftlichen, technisierten Welt allgegenwärtig und erzeugen eine gewisse Datenhörigkeit, man denke etwa an

Share-holder Value

Arbeitslosenquote

Inflationsrate

Score (Schufa-Akte)

Ratings

Hochschulranking

Evaluierungsverfahren

Mit Datenhörigkeit ist gemeint:

Menschen (Privatpersonen, Politiker, Führungskräfte) richten ihr Verhalten ohne viel Überlegung und/oder oft ohne Kenntnis der Zusammenhänge allein danach aus, welcher Datenwert von einer bestimmte Größe medial oder auf andere Weise kommuniziert wird.

Betriebswirtschaftliche Entscheidungen

z.B. Maximierung bestimmter Größen

Gesellschaftspolitische Entscheidungen

z.B. Streben nach mehr Wirtschaftswachstum

z.B. Gesetze zur „Einsparung“ von Energie

Expertengutachten

z.B. vor Gericht oder bei Gesetzesvorhaben
(typisch etwa im Bildungsbereich)

Medizinische Therapierung

z.B. Verschreibung von blutdrucksenkenden Medikamenten

Persönliche Entscheidungen

z.B. Kauf von Wertpapieren in Abhängigkeit vom Kurs oder einem Aktien-Index

z.B. Quantified-Self-Bewegung
(Selbstbeobachtung oder Selbstoptimierung anhand von eigenverantwortlich ermittelten Biosignaldaten)

Betriebswirtschaftliche Fehlentscheidungen

Gesellschaftspolitische Fehlentscheidungen z.B. bei
Vergabe von Subventionen

Expertenunwesen

Persönliche (Fehl)kaufentscheidungen

Therapierung nach Zahlen ohne Anamnese

Selbstopтимierung des eigenen Körpers nach unklaren
Optimierungskriterien; Glaube an erhöhte Fitness?
„bessere“ Gesundheit? auf Kosten psychischer
Störungen und physischer Überbelastung?

Ausrichtung von Entscheidungen nach willkürlichen Scores

Irreführende Ratings

Nichtssagendes/irreführendes Hochschulranking

Nichtssagende/irreführende Evaluierungsverfahren

Fazit

Anhand von Daten werden – nahezu unbemerkt – oft weitreichende Entscheidungen getroffen, die unser Leben nachhaltig beeinflussen können.

Je größer der Dateneinfluss ist, desto wichtiger wird die Frage: was diese Daten eigentlich wert sind.

Anders gefragt:

Wurden die Daten korrekt ermittelt?

Um darauf antworten zu können, benötigen wir methodische Richtlinien; sie werden durch die Messtheorie zur Verfügung gestellt.

- A Daten im Alltag
- B Messtheoretisches Basiswissen
- C Messtheorie
 - C₁ Messprozesse
 - C₂ Metrisierung
 - C₃ Fehlerrechnung
 - C₄ Interpretation
- D Resümee

Durch eine Messung soll etwas über ein *Messobjekt* in Erfahrung gebracht werden.

In Erfahrung bringen setzt voraus, dass es jemanden gibt, der sich für das Messobjekt interessiert und die Messung ausführt; es sei dies der *Beobachter*. Da es bei einer Messung nicht darauf ankommt, wer sie durchführt, lassen wir als Beobachter auch Maschinen zu.

Bei einer Messung interessiert man sich nicht für alle am Messobjekt auftretenden empirischen Verhältnisse, sondern nur für die *Intensität* (oder Stärke oder Grad der Ausprägung) einer ganz bestimmten, eventuell erst durch den Messvorgang hervorgebrachten *Messeigenschaft*. Sie Messeigenschaft ist also diejenige Eigenschaft, deren Intensität durch eine Messung bestimmt werden soll.

Unter 'Intensität' ist hier eine relative, auf eine Einheitsintensität bezogene empirische Größe zu verstehen.

Ein *Standard* definiert empirisch die Einheitsintensität und die *Dimension*. So ist das Urmeter in Paris ein Standard für die Längenmessung. Es repräsentiert die Einheitsintensität 1 und die Dimension 'Meter'.

Messeigenschaften wie Zeitdauer, Länge oder Stromstärke, von denen bereits Messverfahren bekannt sind, bezeichnen wir als *Messgrößen*; ein *Messwert* ist das Ergebnis eines einzelnen Messvorgangs.

Um die Objektivität der Messergebnisse sicherzustellen, sind folgende Messprinzipien zu beachten:

Messprinzip 1

Reproduzierbarkeit der Messergebnisse

Am gleichen Messobjekt unter gleichen Bedingungen, aber zu einem anderen Zeitpunkt und eventuell an einem anderen Ort und von einem anderen Beobachter vorgenommene Messungen müssen innerhalb der erreichbaren Messgenauigkeit zu gleichen Ergebnissen führen.

Messprinzip 2

Vergleichbarkeit der Messergebnisse

Innerhalb der erreichbaren Messgenauigkeit gleiche, aber von verschiedenen Messobjekten stammende Ergebnisse müssen das gleiche bedeuten.

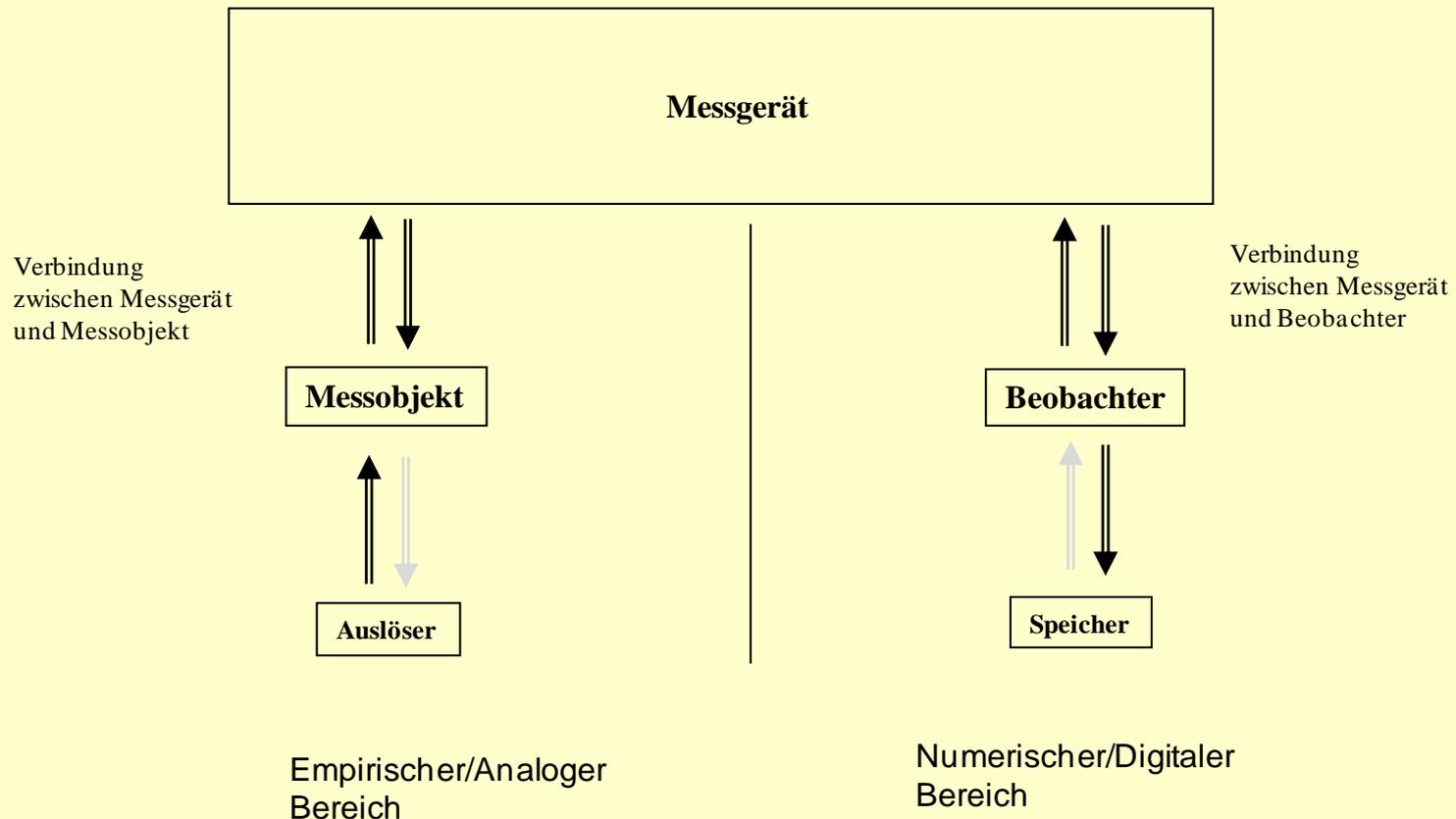


Abb. 1: Grundstruktur eines Messverfahrens. Die Pfeile zeigen die Einflussrichtung an; graue Pfeile deuten an, dass die Wirkung vernachlässigbar ist.

Physikalische Messprozesse sind gekennzeichnet

- durch einen Auslösevorgang (optional)
- durch eine Analog-Analog-Umwandlung
- durch eine Analog-Digital-Umwandlung

Zum Schluss wird das Ergebnis abgelesen und gespeichert.

Die Messtheorie gliedert sich in vier Teile

Experimenteller Teil

1 Messprozesse („Analogteil“)

2 Metrisierung („Digitalteil“)

Auswertung

3 Fehlerrechnung

4 Interpretation der Messergebnisse

- A Daten im Alltag
- B Messtheoretisches Basiswissen
- C Messtheorie**
 - C₁ Messprozesse
 - C₂ Metrisierung
 - C₃ Fehlerrechnung
 - C₄ Interpretation
- D Resümee

C₁ Messprozesse

Brennpunkt: Wechselwirkungseffekte

C₂ Metrisierung

C₃ Fehlerrechnung

C₄ Interpretation

Eine *Wechselwirkung* ist eine wechselseitige Einwirkung von mindestens zwei Systemen aufeinander.

Die Einwirkungen können sehr unterschiedlich sein; sie reichen von einer sehr starken bis hin zu einer gegen Null gehenden bzw. vernachlässigbaren Einwirkung. Es ist sinnvoll, auch Grenzfälle einheitlich als eine *Wechselwirkung* zu betrachten, selbst wenn man dies aufgrund ihrer rudimentären Erscheinung nicht täte.

Zwei wechselwirkende Systeme sind über einen Übertragungskanal miteinander gekoppelt:

- (1) Bei einem Übertragungskanal handelt es sich ausschließlich um ein physikalisches System, daher ist
- (2) jeder Übertragungskanal grundsätzlich physikalischen Störungen aus der Umwelt ausgesetzt und kann
- (3) auch wieder auf die Umwelt zurückwirken.

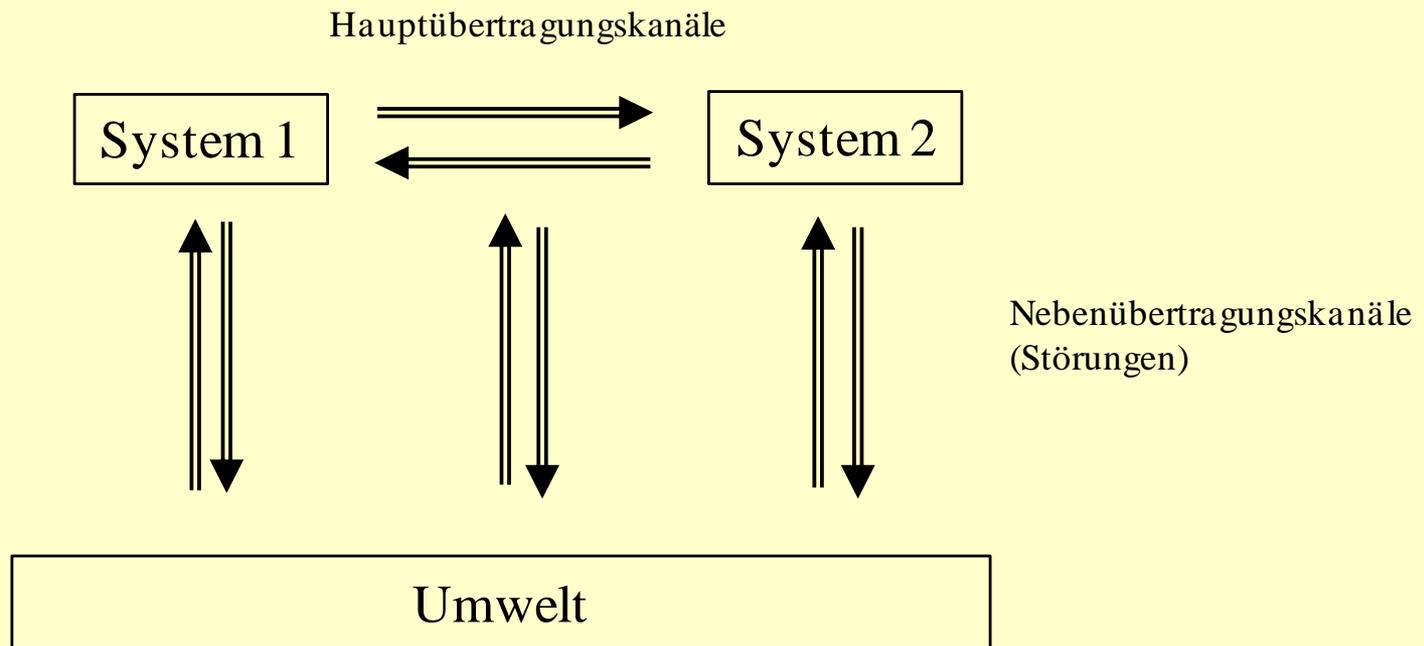


Abb. 2: Schematische Darstellung der Kopplung zweier Systeme.

Das Produkt einer Wechselwirkung sind Zustandsänderungen an den beteiligten Systemen.

Die Systeme müssen daher einen Vorrat an möglichen Zuständen besitzen.

Der Zustandsvorrat ermöglicht
Gedächtnisfunktionen:

Ein System verfügt über ein *Gedächtnis*, wenn es in der Lage ist, Reste aus vergangenen Einwirkungen festzuhalten. Das ist ein sehr allgemeiner Gedächtnisbegriff, der sowohl das menschliche Gedächtnis als auch die „Erinnerungsleistungen“ von physikalischen Systemen umfasst.

Da die Reaktion eines Systems davon abhängt, in welchem aktuellen Zustand es sich gerade befindet, kann die gleiche Einwirkung unterschiedliche Reaktionen hervorrufen.

Da der aktuelle Zustand wiederum aus vergangenen Einwirkungen hervorgegangen ist, können diese auch noch in die Gegenwart und Zukunft hineinwirken:

Gedächtniseffekte treten als Störungen auf, von denen man oft nicht weiß, woher sie stammen. Man muss also vor allem bei nicht-physikalischen Messungen sorgfältig auf solche Gedächtniseffekte achten.

Jede Einwirkung auf ein gedächtnisfähiges System besitzt sowohl eine energetische als auch eine informetische Komponente.

Energetische Komponente:

Ist charakterisiert durch Energiewirkung.

Informetische Komponente:

Ist charakterisiert durch Informiewirkung.

Die informativische Komponente erfordert stets Übertragungsenergie, aber sie spielt bei der Wechselwirkung keine Rolle.

Informiewirkung heißt: Das System reagiert auf den *Inhalt* der Einwirkung: dabei sind Übertragungsenergie und Reaktionsenergie entkoppelt.

Beispiel:

Beleidigung über Schallwellen (Einwirkung) –
Faustschlag (Reaktion)

Welche Folgen eine informatische Einwirkung hat, hängt vom jeweiligen Zustand des Systems ab und dieser wiederum von dessen Vergangenheit.

Informatische Einwirkung plus Gedächtnis können einen beträchtlichen Störeinfluss auf Messungen haben.

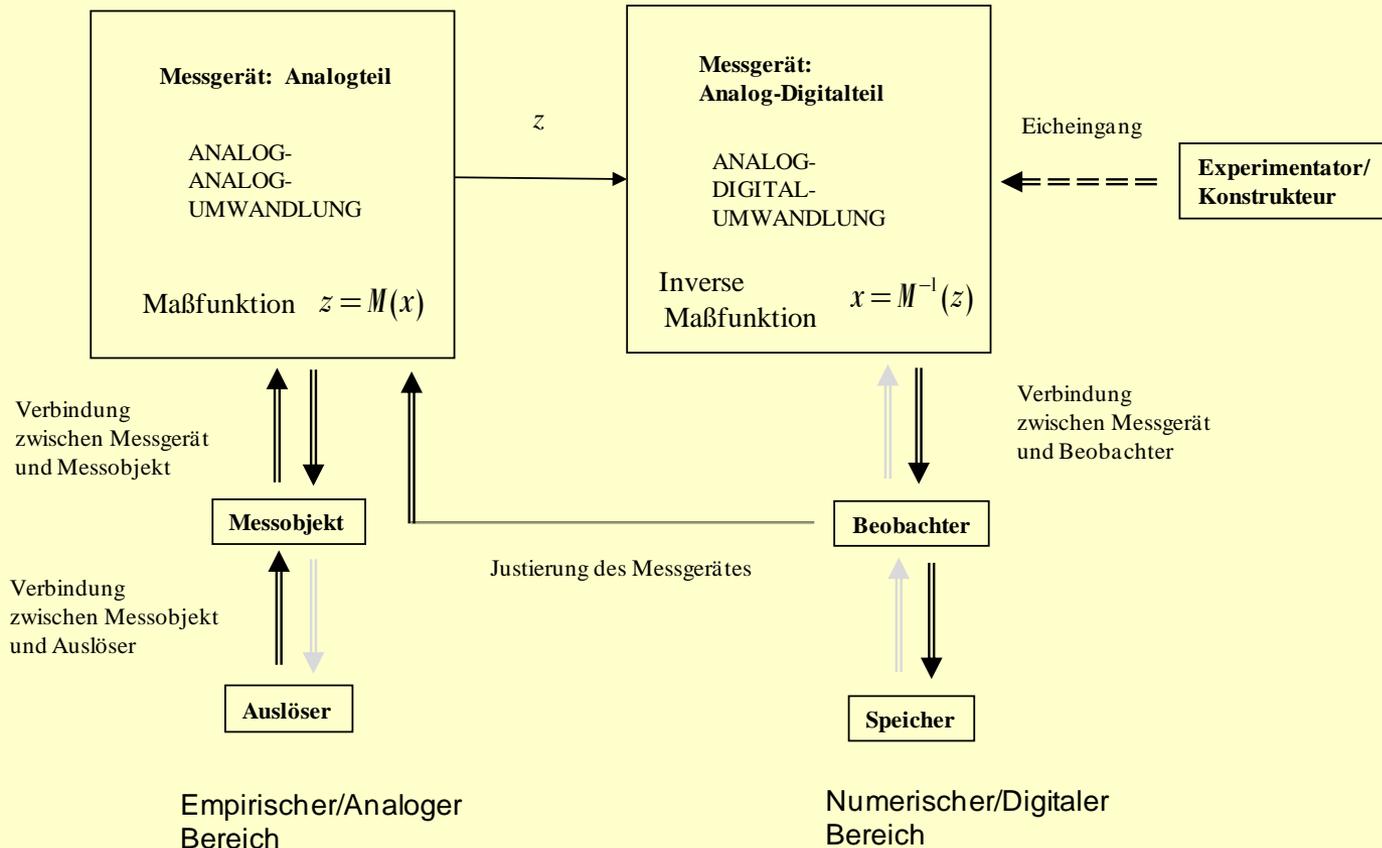


Abb. 3: Physikalische Messprozesse.

Der Auslösevorgang ist notwendig,

- wenn das Messobjekt in einen definierten Ausgangszustand gebracht oder
- wenn die zu untersuchende Messeigenschaft erst im Messobjekt erzeugt werden soll.

Der Auslöser kann sowohl energetischer (Lichtstrahl) als auch informatischer Natur (Fragebogen) sein.

Abgesehen von Spezialfällen (etwa Längenmessung mit einem Maßstab) besteht Messen stets in einer *indirekten* Messung:

Beispiel:

Gesuchte analoge Intensität der Messeigenschaft:
elektrische Spannung x
wird umgewandelt
in eine analoge Intensität am Messgerät:
Zeigerausschlag z

Damit kann eine zunächst „unsichtbare“, nicht direkt beobachtbare Messeigenschaft zugänglich und somit auch messfähig gemacht werden.

Dieser äußerst wichtige Schritt birgt aber eine Menge von Gefahren.

C₁ Messprozesse

C₂ Metrisierung

Brennpunkt: Maßfunktion

C₃ Fehlerrechnung

C₄ Interpretation

Die Metrisierung betrifft die Analog-Digital-Umwandlung.

Letztere vollzieht den Übergang vom empirischen zum mathematisch-sprachlichen Bereich und erfordert zwei Arbeitsschritte: Metrisieren und Ablesen.

Eine Analog-Analog-Umwandlung ist nur dann möglich, wenn zwischen beiden Eigenschaften ein naturgesetzlicher Zusammenhang besteht beschreibbar durch eine *Maßfunktion* der Form

$$z = M(x).$$

Die Maßfunktion beschreibt den gesetzmäßigen Zusammenhang zwischen der eigentlichen Messgröße x (z.B. elektrische Spannung) und der Größe z , die tatsächlich gemessen wird (z.B. Zeigerausschlag).

Um die gewünschte Intensität x zu erhalten, muss aus z das x berechnet werden gemäß

$$x = M^{-1}(z); \quad M^{-1} \text{ ist die inverse Maßfunktion.}$$

Aus einer bislang nicht direkt messbaren Intensität wird so eine messbare.

Hauptaufgabe der Metrisierung ist die Bestimmung einer gültigen Maßfunktion.

Am häufigsten gilt die lineare Beziehung

$$z = a + bx,$$

doch als Maßfunktion kommt

jede streng monotone, eine gesetzmäßige Beziehung zwischen den beteiligten Größen ausdrückende Funktion infrage.

Aber:

Die Maßfunktion bestimmt, was über die Messwerte ausgesagt werden darf:

Je weniger über sie bekannt ist, desto geringer ist der Gehalt der Messergebnisse und desto vorsichtiger müssen diese interpretiert werden:

Im Zweifelsfall müssen die Ansprüche an die Messergebnisse deutlich abgesenkt werden.

Stichwort für diese Problematik: „Skalentheorie“

Die Maßfunktion bestimmt den Skalentyp.

$$z = bx, b > 0 \text{ (Proportionalskala)}$$

zulässige Aussage: wenn z sich verdoppelt, verdoppelt sich auch x .

$$z = a + bx, b > 0 \text{ (Differenzenskala)}$$

zulässige Aussage: wenn sich bei zwei Messungen die *Differenz* $z_2 - z_1$ verdoppelt, verdoppelt sich auch die *Differenz* $x_2 - x_1$.

$z(x)$ ist streng monoton (Ordnungsskala)

zulässige Aussage: wenn z sich vergrößert hat, dann hat sich auch x vergrößert; der Wert der Vergrößerung bleibt offen.

Da für viele (physikalische) Messungen

$$z = bx$$

gilt und oft sogar $b = 1$ ist,

wird überhaupt die Existenz einer Maßfunktion übersehen.

Die Maßfunktion ist jedoch die Grundlage einer *jeden* Messung; es ist daher notwendig, sie etwa durch vorgeschaltete Messungen so genau wie möglich zu bestimmen (Stichwort Metrisierung).

Gegen diese Anforderung wird vor allem bei nicht-physikalischen Messungen häufig verstoßen – mit weitreichenden Folgen.

Dies soll an den folgenden Beispielen veranschaulicht werden.

Nicht-physikalische Messungen sind stets indirekt („abgeleitet“).

Während aber bei physikalischen Messungen die abgeleitete Größe physikalisch existent ist (z.B. die Spannung), wird sie bei den nicht-physikalischen oft erst im Zuge des Messvorhabens *definiert*; sie ist also nur eine *gedachte Größe*.

Als Maß für die Leistung einer Volkswirtschaft L_{VW} gilt der Verkaufswert W sämtlicher während eines Jahres hergestellten Güter und Dienstleistungen:

$$L_{VW} = c W, \quad c > 0$$

Was 'Leistung einer Volkswirtschaft' bedeutet, wird definatorisch festgelegt; andere Festlegungen sind möglich.

Definierte Messgrößen sind hochproblematisch:

Wie lässt sich eine Maßfunktion zwischen einer nur gedachten Messgröße und der bei der Messung tatsächlich verwendeten bestimmen?

(Zwischen beiden soll ja nach Voraussetzung ein gesetzmäßiger Zusammenhang bestehen!)

„Lösung“:

Der gesetzmäßige Zusammenhang (die Maßfunktion) muss ebenfalls definiert werden.

Damit ist die Maßfunktion aber auch nur eine *gedachte* (nicht überprüfbare) Funktion.

Beim Ausdenken werden leicht Zusammenhänge übersehen, die dann nachträglich Korrekturen veranlassen, wenn sich die an das Messverfahren gestellte *Erwartungshaltung* nicht erfüllen lässt.

Beispiel

L_{VW} steigt/sinkt auch wenn die Inflationsrate steigt/sinkt; das gilt als Verfälschung der tatsächlichen Verhältnisse und wird durch eine nachträgliche „Inflationsbereinigung“ korrigiert. Aber auch die Inflationsrate muss gemessen werden, und zwar wie das BIP am Verkaufswert!

Scores sind stets definierte Messgrößen.

In einen „Score“ gehen stets $n > 1$ Messeigenschaften ein, deren Intensitäten z_1, \dots, z_n mit g_1, \dots, g_n gewichtet und dann linear verknüpft werden.

Die inverse Maßfunktion hat dann die Form

$$x = g_1 z_1 + g_2 z_2 + \dots + g_n z_n, \quad g_1, \dots, g_n > 0.$$

Ein Score bereitet als „multidimensional“ definierte Messgröße zusätzliche Probleme:

- (1) Die Komponenten wählen wenige, sich auf ihre Erfahrung berufene Personen aus; die Auswahl ist daher immer subjektiv und oft kultur- und/oder landesspezifisch.
- (2) Die Werte der Gewichte g_1, \dots, g_n werden vordergründig "erraten"; sie können aber bestenfalls statistisch abgeschätzt werden.

- (3) Naturgesetze sind nicht additiv verknüpft; es bleibt daher unklar, welchen Wert solch eine Verknüpfung besitzt.
- (4) Die Linearitätsannahme ist willkürlich; es könnten ja z.B. auch quadratische Terme mit in den Score eingehen.

- (5) Die einzelnen Institute, welche das Ranking vornehmen, legen jeweils ihre eigene Score-Gleichung zugrunde; das führt zu unterschiedlichen, nicht miteinander vergleichbaren Score-Werten.
- (6) Eine Fehlerabschätzung ist für die Score-Werte nicht vorgesehen; Messwerte ohne solch eine Abschätzung sind jedoch sinnlos.

Ein Score, allgemein: jede Maßzahl x , welche durch

$$x = g_1 z_1 + g_2 z_2 + \dots + g_n z_n$$

berechnet wird, hat keinen wissenschaftlichen Wert und ist daher auch kein objektives Maß (wie oft behauptet wird).

(Hoch)schul-Ranking

Beispiel: Kriterien für das Global MBA Ranking (für Business Schools) der *Financial Times* (FT):

Das durchschnittliche Gehalt und die Gehaltssteigerung im Vergleich zu der Zeit vor dem MBA Studium machen jeweils 20 Prozent der Bewertung aus.

Außerdem zählt die *FT* (als Maß für die Forschungsleistung), wie viele Fachartikel Professoren und wissenschaftliche Mitarbeiter in 45 ausgewählten Fachzeitschriften veröffentlicht haben.

Es handelt sich um ein dreidimensionales Modell mit den Komponenten

Durchschnittliches Gehalt (X_1).

Gehaltssteigerung im Vergleich zu der Zeit vor dem
MBA Studium (X_2)

Anzahl der Fachartikel in 45 ausgewählten
Fachzeitschriften (X_3).

Bleibt es bei diesen drei Komponenten und der erwähnten
20% - Bewertung, dann ergibt sich aus den
Einzelmesswerten x_1 , x_2 und x_3 die Score-Gleichung

$$S = 0,2 x_1 + 0,2 x_2 + 0,6 x_3.$$

Der Klout-Score „misst“ den Einfluss eines Menschen in der digitalen Welt, das Gewicht seiner Worte, seine Befähigung zum Meinungsführer.

Score = 0: Kein Einfluss

Score = 100: Maximaler Einfluss.

„Gemessen“ wird der Score mit einem geheimgehaltenen Algorithmus

In die Berechnung fließen ein:

1. Quantität, gemessen an der Anzahl der Facebook-Freunde, der Twitter-Follower oder der You Tube-Abonnenten sowie an der Intensität, mit der über diese Kanäle kommuniziert wird.
2. Mobilisierungsfähigkeit: die misst, wie viele Tweets oder Facebook-Einträge einer Person von anderen Nutzern kommentiert, gemocht oder weiterverbreitet werden.
3. Die Güte des eigenen sozialen Netzwerks, gemessen daran, wie einflussreich die eigenen digitalen Bekanntschaften sind.

Bei Weintests (und anderen Benotungen) wird angenommen:

Die Summe der Subjektivitäten ist der Beginn der Objektivität.

Hier gilt das arithmetische Mittel über Einzelurteile als „Messwert“.

Auch dieses Vorgehen ist fragwürdig.

Denn

- (1) Kann die Fragwürdigkeit der Einzelurteile, die auf einer unzulässigen Maßfunktion nicht „weggemittelt“ werden und
- (2) stellt sich spätestens hier die Frage nach den Messfehlern.

C₁ Messprozesse

C₂ Metrisierung

C₃ Fehlertheorie

Brennpunkt: Intervallansatz

C₄ Interpretation

Bei Zahlmesswerten x_n geht man davon aus, dass er sich aus dem wahren Wert x_0 , dem zufälligen Messfehler δx_n und systematischen Fehler δ_s additiv zusammensetzt gemäß

$$x_n = x_0 + \delta x_n + \delta_s$$

Als Richtwerte gelten

$$\bar{x} = \frac{1}{N} \sum_{n=1}^N (x_0 + \delta x_n + \delta_s) = x_0 + \delta_s + \frac{1}{N} \sum_{n=1}^N \delta x_n \quad \text{Mittelwert}$$

$$s^2 = \frac{\sum_{n=1}^N (x_n - \bar{x})^2}{N} \quad \text{Varianz}$$

Die Varianz ist unempfindlich gegenüber dem systematischen Fehler.

Als Messergebnis wird meist

$$x = \bar{x} \pm u$$

angegeben, wobei als Unschärfe u die (doppelte) Standardabweichung bzw. Streuung s verwendet wird.

Auch u ist gegenüber dem systematischen Fehler unempfindlich.

" \pm " ist keine mathematische Operation.

Als Alternative bietet sich die auf *Zahlenintervallen* beruhende verteilungsfreie Fehlerrechnung an:

Anstelle einen Zahlenwert möglichst dicht am wahren Wert zu ermitteln, wird bei der Intervallmethode eine möglichst gute untere und obere Schranke für diesen Wert anzugeben versucht.

Sei

$$\mathbb{I}_1 = [a_1, b_1], \dots, \mathbb{I}_N = [a_N, b_N]$$

eine aus Intervallmesswerten bestehende
Messreihe.

Der wahre Wert liege jeweils im Inneren der
Intervalle – wo er dort liegt, ist gleichgültig.

Dann definieren die Längen der Intervalle

$$l_n = b_n - a_n \quad n = 1, \dots, N$$

ein direktes Fehlermaß und der
mengen theoretische Durchschnitt

$$\bar{\mathbb{I}} = \bigcap_{n=1}^N \mathbb{I}_n = [\bar{a}, \bar{b}]$$

den Mittelwert dieser Reihe.

Es gilt bei nicht-leerem Durchschnitt stets

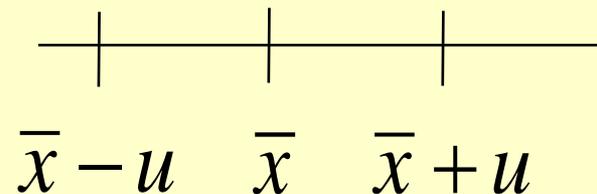
$$\bar{l} = \bar{b} - \bar{a} \leq l_n \quad n = 1, 2, \dots, N$$

d.h. der Mittelwert ist stets genauer, mindestens aber gleich genau wie der genaueste Messwert in der Messreihe.

Ist der Durchschnitt leer, muss mindestens 1 Intervall nicht den wahren Wert enthalten.

Es gibt auch andere Definitionen für die Unschärfe u , welche den systematischen Fehler besser berücksichtigen. Man könnte daher auf die Idee kommen, als Intervallergebnis

$$x = [\bar{x} - u, \bar{x} + u]$$



anzugeben.

Mit dem alternativen Ansatz ist jedoch etwas anderes gemeint:

Bei den Unschärfen u handelt es sich um *Schätzwerte, die aus den Messwerten berechnet werden.*

Beim alternativen Ansatz sind jedoch die untere und obere Schranke experimentell zu bestimmen: sie müssen *gemessen* werden.

Dies erfordert ein Umdenken:

Um die Schranken zu messen, ist meist ein anderer experimenteller Ansatz erforderlich als zur Messung des wahren Wertes.

Bei der Messung der Schranken tritt ebenfalls eine Fehlerproblematik auf; sie lässt sich jedoch leicht durch eine großzügige Fehlerabschätzung lösen.

Bei nahezu allen nicht-physikalischen Messungen fehlen die Fehlerangaben; sie sind daher wertlos.

In unserer Gesellschaft fehlt das Bewusstsein, dass Daten fehlerbehaftet sind. Es werden Entscheidungen getroffen anhand von „nackten“ Zahlen, ohne dabei ihre Unschärfe in Erwägung zu ziehen.

Beispiele:

Wirtschaftswachstum

gemessen am Bruttosozialprodukt:

Angaben des Statistischen Bundesamtes für 2001:

Anfang 2002: Erhöhung beträgt 0,6 %

Später korrigiert: Erhöhung beträgt 1,2 %

Fragwürdig sind ferner alle Score-Werte.

Dennoch werden aus solchen fragwürdigen Daten Schlussfolgerungen gezogen und aus diesen Entscheidungen abgeleitet:

Die Daten werden interpretiert.

C₁ Messprozesse

C₂ Metrisierung

C₃ Fehlertheorie

C₄ Interpretation

Brennpunkt: Schlussfolgerung

Eine Interpretation ist formal eine Schlussfolgerung der Form:

Allgemeines Wissen/Theorie
Messergebnis(se)

Konklusion

Schlussfolgerungen erfordern Prämissen inform von wahren Aussagen (allgemeines Wissen).

Beispiel:

- P₁ Wenn Objekt A höher ist als Objekt B, kann man das Objekt B in das Objekt A hineinstellen.
- P₂ Ein Objekt ändert seine Höhe nicht, wenn es von einem Ort zu einem anderen transportiert wird.
- P₃ Das Zimmer A ist höher als die Schrankwand B.
-
- K Also passt die Schrankwand B in das Zimmer A.

Die ersten beiden Prämissen werden als trivial empfunden und daher nicht eigens erwähnt:

Man setzt sie stillschweigend als gültig voraus.

Stillschweigende Voraussetzungen, aber auch plausibel erscheinende Hypothesen und ideologisch bedingte Aussagen können zu Fehlinterpretationen und diese wiederum zu Fehlentscheidungen mit weitreichenden Folgen führen.

Eine Fehlinterpretation ist formal eine Schlussfolgerung der Form:

Ausgewähltes Wissen
stillschweigende Vermutungen
plausible erscheinende Hypothesen
ideologisch bedingte Annahmen
Messergebnis(se) (ohne Fehlerangaben)

Fehlinterpretation

Beispiel: Beweis der Willensfreiheit

In der Philosophie sorgten in jüngster Zeit neurophysiologische Experimente für Aufregung, weil aus ihnen zu folgen scheint, dass Menschen über keinen freien Willen verfügten.

In der Literatur findet man z.B. bei LIBET (1985) etwa folgende Argumentationsstruktur:

- P₁ Liegt ein Ereignis A zeitlich vor einem anderen Ereignis B, so kann Ereignis A nicht von Ereignis B veranlasst worden sein.
- P₂ Das Bereitschaftspotential erscheint einige 100 ms früher als der Tastendruck der Versuchspersonen, mit dem sie anzeigen, dass sie sich bewusst für eine Handlung entschieden haben.
-
- K Menschen verfügen über keinen freien Willen.

Dieser Schluss ist ohne zusätzliche Prämissen nicht nachvollziehbar.

Es fehlt (A) die Verbindung zwischen dem experimentell beobachteten Bereitschaftspotential und dem Ausführen einer bewusst gewollten Handlung.

Die Lücke ließe sich etwa durch folgende Prämissen schließen:

- P₃ Bevor eine Handlung ausgeführt werden kann, muss stets ein Bereitschaftspotential aktiviert werden.
- P₄ Ein Bereitschaftspotential gehört genau dann zu einer Handlung, wenn es kurz vor ihrer Ausführung erscheint.
- P₅ Beim Abwägen einer Handlung und bei anderen bewussten Überlegungen entsteht kein Bereitschaftspotential.
- P₆ Bereitschaftspotentiale werden „vom Gehirn“ und nicht vom Bewusstsein der Versuchspersonen ausgelöst.
- P₇ Was die Versuchspersonen als bewusste Handlungsabsicht wahrnehmen, ist lediglich die Wirkung des Bereitschaftspotentials.

Weiter bezieht sich die obige Konklusion nicht etwa bloß auf die Versuchspersonen, sondern auf alle Menschen.

Es fehlen also noch (B) die Prämissen

- (1) für die Relevanz der in den Experimenten als ‚bewusst‘ deklarierten Handlungen und
- (2) für den Brückenschlag von wenigen ausgewählten Personen zu den Menschen

Zu ergänzenden sind also noch die Prämissen

- P₈ Die Handlungen, für die sich die Versuchspersonen in den Experimenten entscheiden mussten, sind repräsentativ für alle bewussten Handlungen.
- P₉ Die bei den Versuchspersonen beobachteten Phänomene treffen auf alle Menschen zu.

Die Prämissen P₃ – P₉ sind zum Teil hochproblematisch:

Von einem Beweis der Willensfreiheit kann deshalb keine Rede sein.

Folge 1:

Dennoch wurde von Juristen der Unfreiheit des Willens Rechnung tragende Gesetzesänderungen vorgeschlagen.

Folge 2:

Was sind die gesellschaftliche Folgen solcher Gesetzesänderungen?

Eine gültige Schlussfolgerung setzt voraus, dass alle Prämissen wahr sind.

Um sich vor Fehlinterpretationen zu schützen, ist dringend zu raten

- (1) sich über die verwendeten Prämissen Klarheit zu verschaffen und
- (2) ihre Wahrheit zu überprüfen.

Bei Nachweisexperimenten liegt die Konklusion bereits fest; sie lautet: also existiert ...

Der experimentelle Aufbau wird daher so ausgerichtet, dass sich bei entsprechenden Messergebnissen automatisch die Konklusion ergibt.

Gefährlich, denn diese Vorgehensweise kann die Form einer selbsterfüllenden Prophezeiung annehmen.

„Durch systematisches kombiniertes Suchen nach diesen Zerfällen konnten an zwei unabhängigen Detektoren des LHC deutliche Hinweise auf die Existenz eines entsprechenden Teilchens gefunden werden“ (WIKIPEDIA).

Es wurde also nicht das HIGGS-Boson selbst nachgewiesen, sondern nur bestimmte Zerfälle, von denen man auf die Existenz solch eines Bosons *schließt*.

- A Daten im Alltag
- B Messtheoretisches Basiswissen
- C Messtheorie
 - C₁ Messprozesse
 - C₂ Metrisierung
 - C₃ Fehlerrechnung
 - C₄ Interpretation
- D Resümee

D Resümee

Die Messtheorie umfasst die vier Teilgebiete
Messprozesse
Metrisierung
Fehlerbehandlung
Interpretation.

In allen Teilgebieten lauern Gefahren, denen bei ernsthaften Messverfahren begegnet werden muss:

D Resümee

Messprozesse:

Messprozesse können das Messergebnis verfälschen und sogar unbrauchbar machen; zugleich sind sie für eine Kopplung von Messobjekt und Messgerät unentbehrlich.

Metrisierung:

Die Maßfunktion muss bestimmt und darf nicht geraten werden; zumindest die mathematischen Mindestanforderungen müssen bekannt sein, um eine sinnvolle Interpretation zu gestatten.

D Resümee

Fehlertheorie:

Mit der Abschätzung einer unteren und oberen Schranke steht eine alternative, verteilungsunabhängige Methode zur Behandlung der Fehlerproblematik zur Verfügung.

Interpretation:

Interpretationen sind als Schlussfolgerungen im strengen logischen Sinn zu behandeln, wobei deren Prämissen mit größter Sorgfalt behandelt werden müssen.

JAENECKE, PETER (2014):
Einführung in die Messtheorie.
Messtheorie Bibliografie

Verfügbar unter:

<http://www.peterjaenecke.de/messtheorie.html>

JAENECKE, PETER (2013):

Offenes Kodieren in der Qualitativen Forschung

JAENECKE, PETER (2014):

Desorientierung durch falsche Maße

Verfügbar unter:

<http://www.peterjaenecke.de/soziologie.html>

E-Mail: peter@jaenecke.com