

**Monographien
zur
Prüfungsforschung**

Herausgegeben vom
Projekt Prüfungsforschung
am Institut für Pädagogik

Projekt Prüfungsforschung
am Institut für Pädagogik
der Christian-Albrechts-Universität zu Kiel (Hg.)

Monographien
zur Prüfungsforschung
Heft 19

Julia Buchheit, Peter Kroppe (Hg.)

Paradigmatische Auseinandersetzung mit Problemen der Testtheorie

Für Kontakte:

Prof. Dr. Peter Kroppe

D: 24118 Kiel

Olshausenstraße 75

Tel. (0431) 8801273

FAX (0431) 8801588

E-Mail: kroppe@paedagogik.uni-kiel.de

P: 24147 Kiel

Tauernweg 14

Tel. (0431) 7809230

Fax (0431) 7809232

E-Mail: pkroppe@kroppe.de

Kiel 2010

Inhalt

Vorwort	7
Julia Buchheit	
Das dialogische Moment einer sozialwissenschaftlichen Messung	9
Peter Kroppe	
Vom Repräsentationstheorem zur Prädikation auf empirischer Grundlage	19
Die Reihe „Monographien zur Prüfungsforschung“	29

Vorwort

Jedes Jahr im Frühjahr und im Herbst lädt das Team des Zentrums für Konstruktive Erziehungswissenschaft zu einem Workshop ein. In den Veranstaltungen werden Ergebnisse eines jeweils aktuellen Forschungsvorhabens vorgestellt und theoretische Grundsatzfragen und praktische Folgeprobleme erörtert.

Das Thema des Frühjahrsworkshop 2010 lautete: Paradigmatische Auseinandersetzung mit Problemen der Testtheorie. Im Workshop sollten die Beziehungen zwischen Tests, ihrer paradigmatischen Einbindung sowie ihrer Folgen für die pädagogische Praxis zur Sprache kommen.

Auf dem Programm standen die folgenden vier Vorträge:

Prof. Dr. Jürgen Rost	Graphische Testanalyse
Prof. Dr. Hans-Carl Jongbloed	Der innere Rahmen der Gütekriterien
Dr. Julia Buchheit	Der Messbereich Erfahrung
Prof. Dr. Peter Krope	Repräsentationstheorem vs. Prädikation auf empirischer Grundlage

Der dritte und der vierte Vortrag sind für das vorliegende Heft der Monographienreihe überarbeitet worden. Gemeinsam ist ihnen die Grundlage des Methodischen Konstruktivismus. Unterschiedlich sind die Herangehensweisen und die Darstellungsformen.

Julia Buchheit

Das dialogische Moment einer sozialwissenschaftlichen Messung

1. Einführung in die Thematik und Fragestellung

Manche Wissenschaftler sehen in unserer heutigen Zeit eine Kultur der Rechenschaftslegung, eine so genannte „Audit Society“¹ und zwar so gestaltet, dass Nummern und Zahlen zu Entscheidungsträgern stilisiert werden. Sozialwissenschaftliche Phänomene werden mit Nummern und Zahlen analysiert. Sie werden durchleuchtet und geprüft. Ihre Werte und ihre Maße werden ermittelt. Es wird gemessen.

Das Messen ist die Durchführung einer geplanten Tätigkeit, um numerische Aussagen über einen Sachverhalt zu erlangen. Messen ist das Vergleichen des Sachverhalts mit einer Einheit. Ziel ist es dabei, eine verlässliche und nachvollziehbare Aussage über den Sachverhalt zu erhalten.² Der Ausdruck SOZIALWISSENSCHAFTLICHE MESSUNG wird in diesem Aufsatz in seiner Extension sehr weit gefasst, d.h. er beinhaltet sowohl die Funktionen der numerischen Zuschreibung als auch des Vergleichs, desweiteren bezeichnet er die Funktionen des Urteilens und der Bewertung.³

Die heutige Situation ist ein Gemenge verschiedener Entwicklungsstränge. Historisch gesehen, kann der Beginn methodisch durchgeführter numerischer Beurteilungen mit der Zeit der Industrialisierung gleichgesetzt werden. In den ersten Jahrzehnten des zwanzigsten Jahrhunderts wurde in der amerikanischen Forschung hauptsächlich der Leitgedanke der Effizienzsteigerung verfolgt und um dies zu erreichen wurden gezielt Evaluationen durchgeführt. Die Entwicklung standardisierter Tests zur objektiven Kontrolle genormter Lernziele wurde in den USA bereits in den 50er Jahren regelmäßig umgesetzt.⁴ Theoretische Ansätze, die als Grundlage zur Ausföhrungen von Evaluationen dienten und die sich aus

¹ POWER, M. (1999) *The audit society: Rituals of Verification*, Oxford.

² Messen ist dann gelungen, wenn eine zweckerfüllende Statution und eine zweckerfüllende Prädikation des empirischen und numerischen Relativs erfolgten, sofern die Prädikationen dieselben relationalen Strukturen aufzeigen.

³ Vgl. dazu die Beschreibungen der vier Generationen der Evaluation in GUBA, E. G., & LINCOLN, Y.S. (1989). *Fourth generation evaluation*. Newbury Park, CA.

⁴ Vgl. MADAUS, G. F., STUFFLEBEAM, D. L. & WEBSTER, J.W. (1983). *Program Evaluation: A historical Overview*. In D. L. STUFFLEBEAM, G. F. MADAUS, & M. SCRIVEN (Hrsg.) *Evaluation models - viewpoints on educational and human services evaluation* (S. 3-22). Boston.

Durchführungen ergaben, entwickelten sich in unterschiedlichen Disziplinen. Im Besonderen trugen in den USA die theoretischen Ausführungen von LEWIN (Sozialpsychologie), TYLER (Erziehungswissenschaft) und LAZARSELD (Soziologie) dazu bei, numerische Evaluationen als gängige Methode einzuführen.⁵ In Deutschland ist der Einsatz von großflächigen Erhebungen im Bildungsbereich erst später – in den 70er und 80er Jahren – als übliche Evaluationsmethode zu beobachten.⁶ Als wissenschaftliche Disziplin, die sich kritisch mit den Grundlagen der bestehenden Methoden auseinandersetzt ist die Erziehungswissenschaft zu nennen.⁷ Bereits 1972 wurde auf der Arbeitstagung der Deutschen Gesellschaft für Erziehungswissenschaft das Thema „Zur Wissenschaftstheorie und Methodologie pädagogischer Begleitforschung“ detailliert besprochen. Unter anderem stellte DAHMER heraus, dass eine Theorie der angewandten erziehungswissenschaftlichen Forschung dringend erforderlich sei. In den Arbeitskreisen wurden die Problemfelder Theorie und Praxis, Repräsentativität, Validität und Operationalisierung diskutiert.⁸ Evaluationen zur Prüfung politischer Programme werden in Deutschland unter anderem in den Bereichen Bildung, Gesundheit, Arbeitsmarkt und Arbeitsleben durchgeführt. Die Ausführungen von LIENERT und FISCHER zur Testtheorie gelten in der empirischen Sozialwissenschaft bis heute als Grundlagentexte. Im Bereich der empirischen Bildungsforschung gehören unter anderem ROBINSOHN, ROTH, und BREZINKA zu den Initiatoren. "Die Diskussion über Evaluation im Bildungswesen, d.h. die systematische Beurteilung von Organisationsstrukturen, Lehr- und Lernprozessen und Leistungsmerkmalen mit der Zielsetzung der Qualitätsverbesserung, hat in der Bundesrepublik Deutschland erst Ende der 80er Jahre und damit später als in anderen europäischen Staaten eingesetzt."⁹

⁵ Vgl. SHADISH, W. R., COOK, T. D., & LEVITON, L. D. (1991). *Foundations of Program Evaluation: Theories of Practice*. Newbury Park, CA, S. 21.

⁶ Vgl. ROSSI, P. H., FREEMAN, H. E. & HOFMANN, G. (1988). Programm-Evaluation. *Einführung in die Methoden angewandter Sozialforschung*. Stuttgart, S. 7.

⁷ Vgl. LANGE, E. (1983) Zur Entwicklung und Methodik der Evaluationsforschung in der Bundesrepublik Deutschland, *Zeitschrift für Soziologie* 12 (3): S. 255-258. Mit den 1970 eingeführten Gemeinschaftsaufgaben von Bund und Ländern, festgehalten in den Gesetzesgrundlagen GG Artikel 91b, wurden Reformmaßnahmen und die dazugehörigen Evaluationen initiiert und durchgeführt.

⁸ Vgl. dazu LÖFFELHOLZ, M. (1972) Pädagogische Begleitforschung: Arbeitstagung der Deutschen Gesellschaft für Erziehungswissenschaft vom 10. bis 12. April 1972 in Nürnberg. *Zeitschrift für Pädagogik*, 18 (3): S. 457-462.

⁹ Organisation des Bildungssystems in der Bundesrepublik Deutschland 2007/08 herausgegeben von The Education, Audiovisual and Culture Executive Agency (EACEA) digital abrufbar unter Eurybase - Descriptions of Education Systems, S. 254.

Das Verfahren der Messung wird als wertende und meist als entscheidungsgerichtete Verunfthanwendung oder Erkenntnisgewinnung beschrieben und steht im Dienste der Qualitätsermittlung. Im Rahmen von Optimierungssystemen zur Qualitätskontrolle bilden sie die Erkenntnisfunktion. Anhand von Evaluationen kann geprüft werden, ob das bestehende Vertrauen in die Institution oder Person weiter aufrecht zu erhalten ist, oder ob das Vertrauen anzuzweifeln oder gar entzogen werden soll. Schnell kann man hier ein Machtgefälle vermuten. Ein Machtgefälle zwischen Ausführenden und Untersuchten, welches die Teilhabe der zu Untersuchenden verneint und die Teilgabe an Ergebnisse methodischen Vorgehens knüpft.

Messungen werden durchgeführt, um Dinge oder Zustände exakt zu beschreiben. Sie erfüllen die unterschiedlichsten Zwecke. Sie werden unter anderem angewendet,

1. um Dinge oder Zustände zu verändern (Entscheidungsfunktion),
2. um etwas zu überprüfen und zu verbessern (Optimierungsfunktion),
3. oder um etwas bzw. dessen Funktionen Dritten gegenüber zu legitimieren (Legitimationsfunktion).¹⁰

In der Audit Society wird die letzt genannte Funktion hervorgehoben: die Funktion der Legitimation. Im Folgenden soll sie betrachtet werden. Es wird dargelegt, welche Prinzipien Legitimation in diesem Kontext beinhaltet. Dabei werden die Wechselbeziehungen innerhalb einer Messung in den Fokus genommen. Argumentiert wird am Beispiel einer Erfahrungsmessung, somit einer Messung, die sich auf Strukturen bezieht, welche sich der direkten Beobachtbarkeit entziehen.

Legitimationsvorgänge in und durch Evaluationen werfen die folgenden Fragen auf: Wer ist aktiv an einer sozialwissenschaftlichen Messung beteiligt? Oder besser gesagt, wie gestaltet sich Teilhabe an einer Erfahrungsmessung? Die Beantwortung der Frage nach Teilhabe an und innerhalb von sozialwissenschaftliche Messungen lenkt den Blick auf die inneren Relationen. Auf was bezieht sich der Messvorgang? Sind es Untersuchungsobjekte, die es zu zählen gilt? Oder misst man jemanden, der womöglich die Messung mitgestaltet?

Dass die Frage der Teilhabe an sozialwissenschaftlichen Messungen im Folgenden am Beispiel einer Erfahrungsmessung erörtert wird, liegt vor allem daran, dass Erfahrungen, betrachtet als Messobjekt, etwas Besonderes sind. Erfahrungen gehören 1.) zu den Messobjek-

¹⁰ Vgl. zu den angegebenen Funktionen BANK, V. und LAMES, M. (2000) Über Evaluation, Kiel. Zusätzlich wird von BANK die Funktion der Steuerung genannt. Die Funktion der Information, die ebenfalls aufgezählt wird, wird hier als basale Funktion der Messung betrachtet.

ten, die latent genannt werden, da sie nur „indirekt über ausgewählte Verhaltenskriterien“¹¹ darstellbar sind, und 2.) sind sie besonders, weil sie der Personalität¹² zugeordnet werden. Die Frage nach der Teilhabe der Versuchspersonen kann deshalb bei Erfahrungsmessungen augenscheinlicher herausgearbeitet werden.

Das Argumentationsschema ist strukturiert durch vier aufeinander folgende Handlungsschemata, die sich innerhalb eines Messprozesses abspielen. Über die Handlungen der STATUTION, PRÄDIKATION, KALKULATION und INTERPRETATION lassen sich spezifische Charakteristika der Methode und des Dialogs innerhalb eines Messprozesses darstellen.

Eine dialogische Handlung kann beschrieben werden als „eine sprachlich geführte Auseinandersetzung, zwischen zwei oder mehr Personen“. „Der Dialog ist ... eine zwischenmenschliche Beziehung, die durch Anerkennung des Anderen (Du) als Person, Verzicht auf Instrumentalisierung, Ernstnehmen der Freiheit des Anderen ausgezeichnet ist.“¹³ Eine methodische Handlung dagegen kann beschrieben werden als „ein nach Mittel und Zweck planmäßiges (=methodisches) Verfahren, das zu technischer Fertigkeit bei der Lösung theoretischer und praktischer Aufgaben führt“¹⁴.

¹¹ MITTELSTRAB, J. (Hrsg.) (2004) Enzyklopädie, Philosophie und Wissenschaftstheorie, Band 2, Sonderausgabe, Stuttgart, Abschnitt Messung.

¹² „**Person** (von lat. Persona), in der römischen Antike Bezeichnung für die Rolle des Schauspielers auf der Bühne (ursprünglich: Maske), aber auch für die Rolle des Individuums in der Gesellschaft ... als philosophischer Terminus in der Regel zu dem Zweck verwendet, den Menschen von anderen Lebewesen und materiellen Dingen zu unterscheiden, was sich noch in der gängigen rechtstheoretischen Unterscheidung von Person und Sache zeigt.“ (MITTELSTRAB, J. (Hrsg.) (2004) Enzyklopädie, Philosophie und Wissenschaftstheorie, Band 3, Sonderausgabe, Stuttgart, Abschnitt Person.) Personalität kommt dann einem Lebewesen zu, wenn es fähig ist, selbstbestimmt zu handeln und Verantwortung zu tragen. „Die ‚Wesensgleichheit‘ der Relationen, die den Redebereich der moralischen Personalität ausmachen, besteht darin, dass sie Resultate eines Reflexionsprozesses über Vorgänge in der Gemeinschaft sind.“ (PSARROS, N. (2000) Merkmal, Leistung oder Anerkennung? Drei Betrachtungen der Personalität. *Dialektik*, 2: S.50 f.) „Das Einräumen der Möglichkeit eigenverantwortlichen Handelns erfordert von den Individuen, dass sie ihre individuellen Zwecke in Bezug auf den gemeinschaftlichen Zweck auswählen, so dass das Resultat ihres Handelns als Beitrag zum gemeinschaftlichen Handeln zählt. Sie müssen in der Lage sein, nicht nur die gemeinsamen Zwecke nachzuvollziehen, sie müssen sie sich aneignen und in der Lage sein, das eigene Handeln am gemeinschaftlichen Maßstab zu beurteilen.“ (PSARROS, N. (2000) Merkmal, Leistung oder Anerkennung? Drei Betrachtungen der Personalität. *Dialektik*, 2: S. 54.)

¹³ MITTELSTRAB, J. (Hrsg.) (2004) Enzyklopädie, Philosophie und Wissenschaftstheorie, Band 1, Sonderausgabe, Stuttgart, Abschnitt Dialog.

¹⁴ MITTELSTRAB, J. (Hrsg.) (2004) Enzyklopädie, Philosophie und Wissenschaftstheorie, Band 2, Sonderausgabe, Stuttgart, Abschnitt Methode.

Einzuordnen ist dieser Artikel¹⁵ als ein erster Versuch soziologische Mikroprozesse innerhalb einer Erfahrungsmessung zu erklären. In diesem Sinne werden Voraussetzungen, Abläufe und Folgen des menschlichen Zusammenspiels während einer Erfahrungsmessung beleuchtet.

Es wird die Hypothese untersucht, ob Evaluationen aufgrund ihrer Funktion ein Machtgefälle darstellen, welche die Teilhabe der zu Untersuchenden verneint. Argumentativ werden einzelne Schritte des Messens betrachtet, um im Einzelnen die Prinzipien des Dialogs und der Methode herauszustellen. Ziel ist es dabei, allgemeine Aussagen festzustellen mit denen man zukünftige Messungen und Evaluationen bewerten und beurteilen kann.

2. Vier Handlungsschemata einer Erfahrungsmessung

Betrachtet man eine Erfahrungsmessung als eine Handlungsabfolge, stellt man fest, dass stets vier Handlungsschemata durchlaufen werden. Abbildung 1 zeigt eine grobe Schematisierung der aufeinanderfolgenden Schritte.

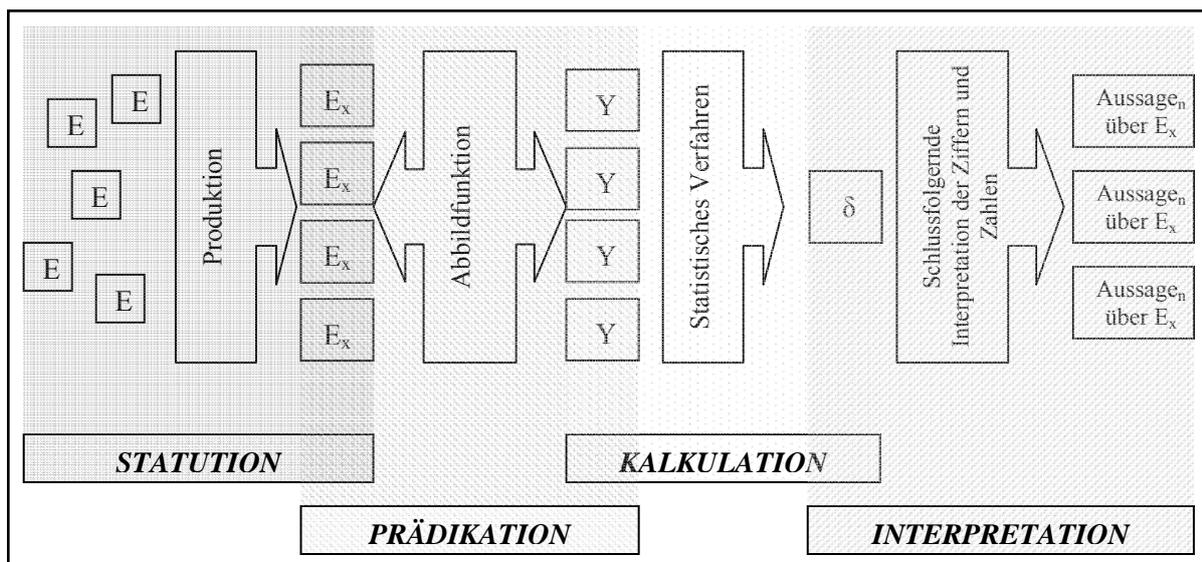


Abbildung 1: Ablauf einer experimentellen Forschung mit schlussfolgernder Interpretation¹⁶

¹⁵ Des Öfteren wurden Messungen von Erfahrung untersucht. Zum einen gibt es Abhandlungen, die das Thema aus testtheoretischer Sicht betrachten, zum anderen gibt es Ausführungen, die wissenschaftstheoretische Erkenntnisse ziehen. Vgl. zur testtheoretischen Auseinandersetzung dieses Feldes zum Beispiel FISCHER, G. (1974) Einführung in die Theorie psychologischer Tests, Grundlagen und Anwendungen, Bern, Stuttgart, Wien. Und ziehe zur wissenschaftstheoretischen Auseinandersetzung zum Beispiel KAMBARTEL, F. (1976) Erfahrung und Struktur – Bausteine zu einer Kritik des Empirismus und Formalismus, 2. Auflage, Frankfurt am Main, heran.

¹⁶ Grafik aus BUCHHEIT, J. (2009) Das Maß der Erfahrung: Ein Entwurf pädagogischen Handelns, Norderstedt, S. 215. „Legende: δ := exaktes und präzises Ergebnis der Kalkulation; E_x := statuierte Erfahrungen (empirisches Relativ); E := Erfahrungen; Y := formallogische Prädikationen (numerisches Relativ). Die Darstellung umfasst den gesamten Ablauf innerhalb eines Experiments. Die Pfeile stellen jeweils Arbeitsschritte dar. Einfachpfeile

Das erste Handlungsschema einer Erfahrungsmessung ist die STATUTION. Die STATUTION ist ein kreativer Prozess, der auf den vorhandenen Fundus zugreift und dem forschenden Interesse entspringt. Der Handelnde erzeugt ein EMPIRISCHES RELATIV¹⁷, indem er Zusammenhänge erkennt und Gegenstände bzw. Aussagen unterscheidet. Bei einer Erfahrungsmessung orientiert sich der Forscher, indem er bestimmt, was er untersuchen möchte. Aus der Mannigfaltigkeit der Erfahrungen – wie zum Beispiel die Erfahrung als Lebenserfahrung oder die Erfahrung als situatives Erlebnis – schöpft der Forscher eine Bestimmte heraus. Dabei kann es vorkommen, dass er sich auf eine bereits definierte Erfahrung konzentriert. Als Beispiel hierfür könnte sich das Forschungsinteresse auf folgende Definition beziehen:

Erfahrungen sind Reflektionen des Menschen über die kulturelle Welt mit Hilfe der Sprachsymbole.¹⁸ „Erfahrungen sind vorstrukturiert, transdirekt, aber immer noch unmittelbar.“¹⁹

Oder der Forscher konzentriert sich auf ein wiederholtes privates Erlebnis und beginnt dieses mit gegebenen Beschreibungsindices zu besetzen. Welchen Weg der Forscher oder die Forscherin wählt hat keine Auswirkungen auf die Funktion des Statuierens. Das Statuieren ist ein kreativer Vorgang.

Auf die STATUTION baut das zweite Handlungsschema einer Erfahrungsmessung auf, die PRÄDIKATION. Die PRÄDIKATION innerhalb einer Messung ist die sprachliche Festlegung des empirischen Relativs und dem dazugehörigen numerischen Relativ. Die Handlung der PRÄDIKATION ist eine genaue Beschreibung des Untersuchungsobjekts und die Angabe einer Handlungsanweisung, wie die unterschiedlichen Merkmalsausprägungen des Untersuchungsobjekts in einem empirischen Relativ erzeugt werden können. Zusätzlich werden Verfahren angegeben, wie die Darstellungen der Untersuchungsobjekte in einem formalen System kategorisiert werden. Damit wird die Überführung des empirischen Relativs in das numerische Relativ gesichert.

sind je ein Arbeitsschritt; Doppelpfeile sind je zwei Arbeitsschritte. Die Namen der einzelnen Arbeitsschritte sind am unteren Rand des Bildes fixiert und jeweils mit derselben Graustufe unterlegt, wie der dazugehörige Arbeitsschritt im oberen Feld des Bilds. Alle Rechtecke im oberen Teil des Bildes stellen strukturelle Stadien innerhalb eines Experiments dar.“ (Ebd. S. 215.)

¹⁷ Das EMPIRISCHE RELATIV beschreibt die relationale Menge von bezeichneten Objekten, welche untersucht werden sollen. Aus Sicht der Klassischen Testtheorie ist das Empirische Relativ etwas Gegebenes, welches es zu überführen gilt. Vgl. dazu LIENERT, G. A., RAATZ, U (1998) *Testaufbau und Testanalyse*. 6. Auflage. Weinheim.

¹⁸ Vgl. BUCHHEIT, J. (2009) *Das Maß der Erfahrung: Ein Entwurf pädagogischen Handelns*, Norderstedt, S. 82.

¹⁹ Ebd. S. 82.

Auf diesen Schritt, dem Schritt der PRÄDIKATION, folgt die KALKULATION. Dieses Handlungsschema ist keine kreative Tätigkeit, sondern sie ist eine Ableitung logischer Schlussfolgerungen. Ziel ist es hierbei, eindeutige Ergebnisse zu erzielen. Das Ergebnis wird meist dargestellt als Ziffer oder Zahl, welche anhand der KALKULATIONSFUNKTION hervorgerufen wird. Kalkulieren ist das methodische anwenden formallogischer Regeln.

Um die Messung zu vervollständigen folgt auf den Schritt der KALKULATION, die INTERPRETATION. Die Aufgabe dieses Handlungsschemas ist es, das formallogische Ergebnis mit dem Ausgangsinteresse, das heißt mit der Zwecksetzung der Untersuchung zu verbinden. In Bezug auf das Ergebnis werden dem Untersuchungsphänomen Attribute zu- und abgesprochen. Man kann sagen, es wird näher bestimmt.

3. Prinzipien der Methode und des Dialogs

Nimmt man diese vier Handlungsschemata als Grundlage jeder Erfahrungsmessung und somit jeder Evaluation im Bereich der Erfahrung dann wird zuerst deutlich, dass Messungen einer methodischen Ordnung folgen, welche die Wiederholbarkeit der Untersuchung sichern. Sowohl die Wiederholbarkeit als auch die Möglichkeit der erneuten Produktion sind notwendige Bedingungen für wissenschaftlich exakte Beobachtungen, für experimentelle Beschreibungen für begründete Entscheidungen und die gegenseitige Einsicht in unterschiedliche Wertvorstellungen. Zeitgleich zu dieser methodischen Ordnung ist zusätzlich noch ein dialogisches Moment gegeben, welches sich als Basisprozess bei einer Erfahrungsmessung greifbar darstellen lässt.

Die Handlungsschemata STATUTION und PRÄDIKATION zeigen, dass eine Evaluation nicht nur dem methodischen Prinzip folgen muss, um Wiederholbarkeit zu sichern, sondern sie müssen auch dem Prinzip des Dialogs folgen, um gangbar zu sein.

STATUTION als ein kreativer Prozess, der auf den vorhandenen Fundus zugreift und dem forschenden Interesse entspringt, ist ein Prozess, der im Hier und Jetzt beginnt. Bezogen auf Erfahrungen bedeutet dies, er beginnt hier, hier wo ich als Forscher oder Forscherin bereits meine eigenen Erfahrungen kenne und hier, wo ich als Forscher oder Forscherin bereits erkannt habe, dass andere auch Erfahrungen machen. Für eine STATUTION bei der Erfahrungsmessung ist daher eine Anerkennung anderer Personen vorauszusetzen. Erst wenn der Handelnde personale Zusammenhänge erkannt hat, kann er oder sie ein Empirisches Relativ erzeugen. In unserem Beispiel müssten der Forscher und die Forscherin zumindest den Zusammenhang erkennen und anerkennen, dass andere Personen auch Erfahrungen machen bzw.

Erfahrungen haben. Im Handlungszusammenhang werden bei der STATUTION Gegenstände (Dinge oder Tatsachen) unterschieden, zum Beispiel mit dem Beschluss, dass die Forschenden im Speziellen ‚tiefgreifende Erfahrungen‘ untersuchen möchten. Dann bezieht sich die Untersuchung auf Personen und deren Traumata.

Auch beim zweiten Handlungsschema, bei der PRÄDIKATION, ist nicht nur das Prinzip der Methode maßgebend. Die PRÄDIKATION ist die sprachliche Festlegung des empirischen Relativs und dem dazugehörigen numerischen Relativ. Die Handlung der PRÄDIKATION ist eine genaue Beschreibung des Untersuchungsobjekts und die Angabe einer Handlungsanweisung, wie die unterschiedlichen Merkmalsausprägungen des Untersuchungsobjekts in einem empirischen Relativ erzeugt werden können. Dabei werden Verfahren angegeben, wie die Darstellungen der Untersuchungsobjekte in einem formalen System kategorisiert werden können. Hierbei wird die Überführung des empirischen Relativs in das numerische Relativ gesichert und die situative Merkmalsbeobachtung des Forschers in eine überdauernde Struktur überführt. Auch hier erkennt der Forscher die situative Begegnung zweier Personen und damit die Handlungsfreiheit der anderen Person an. Indem er ein Verfahren angibt, wie situative Begegnungen festgehalten und in einer Struktur konserviert werden können, bezieht er sich auf das dialogische Prinzip.

4. Schlussfolgerungen

Es hat sich gezeigt, dass in einer Erfahrungsmessung nicht nur das immer wieder genannte methodische Prinzip eine Basis darstellt, sondern auch das dialogische. So wie das methodische Prinzip die Wiederholbarkeit und Produktion der Ergebnisse sichert, so sichert das dialogische Prinzip die Verständlichkeit und Verlässlichkeit der Ergebnisse. Beide Prinzipien sind notwendig, um die Ergebnisse einer Erfahrungsmessung einer interessierten Gruppe zur Verfügung zu stellen. Methode und Dialog gemeinsam sichern die Transsubjektivität und Nachvollziehbarkeit der Ergebnisse.

Diese Ergebnisse lassen Zweifel aufkommen an der oben genannten Hypothese, Evaluationen stellten aufgrund ihrer Funktion ein Machtgefälle dar, welches die Teilhabe der zu Untersuchenden verneint und die Teilgabe – meist in Form von monetären Zuschüssen, die sogenannte Aufwandsentschädigung für Versuchspersonen – an Ergebnisse methodischen Vorgehens knüpft. Leitet man die Teilhabe vom faktischen Dasein, vom Interesse an der Person und vom Vorhandensein eines dialogischen Moments ab, so ist die Teilhabe der Versuchspersonen – zumindest bei einer Erfahrungsmessung – gegeben.

Inwieweit diese Teilhabe als aktiv zu gestaltendes Moment begriffen wird, kann hier leider nicht untersucht werden. Zu bemerken bleibt, dass sozialwissenschaftliche Messungen, die sich auf latente menschliche Merkmalsausprägungen beziehen, wohl in ihrer Ausführung kritisiert werden können. Eine Ablehnung jedoch, die sich auf die Mechanismen bezieht – wie zum Beispiel der Vorwurf: eine Messung würde sich ausschließlich auf den Menschen als reines Objekt beziehen – schlägt fehl. Um die mit dem Evaluationsboom einhergehenden Probleme und die zwischenmenschlichen und gesellschaftssteuernden Folgen beschreiben zu können, muss neben dem methodischen Moment vor allem das dialogische Moment als Grundlage sozialwissenschaftlicher Messung beachtet und rekonstruiert werden.

5. Literaturverzeichnis

- BANK, V., LAMES, M. (2000) *Über Evaluation*, Kiel.
- BREZINKA, W. (1972) *Von der Pädagogik zur Erziehungswissenschaft: eine Einführung in die Metatheorie der Erziehung*, 2. Auflage, Weinheim, Basel.
- BUCHHEIT, J. (2009) *Das Maß der Erfahrung: Ein Entwurf pädagogischen Handelns*. Reihe: Moderne der Tradition, Bd. 7, hgg. von JONGEBLOED, H. C. und BANK, V. Norderstedt.
- FISCHER, G. (1974) *Einführung in die Theorie psychologischer Tests, Grundlagen und Anwendungen*, Bern, Stuttgart, Wien.
- GUBA, E. G., LINCOLN, Y. S. (1989). *Fourth generation evaluation*. Newbury Park, CA.
- KAMBARTEL, F. (1976) *Erfahrung und Struktur – Bausteine zu einer Kritik des Empirismus und Formalismus*, 2. Auflage, Frankfurt am Main.
- LANGE, E. (1983) Zur Entwicklung und Methodik der Evaluationsforschung in der Bundesrepublik Deutschland, *Zeitschrift für Soziologie* 12 (3): S. 253 – 270.
- LIENERT, G. A., RAATZ, U. (1998) *Testaufbau und Testanalyse*. 6. Auflage. Weinheim.
- LÖFFELHOLZ, M. (1972) Pädagogische Begleitforschung: Arbeitstagung der Deutschen Gesellschaft für Erziehungswissenschaft vom 10. bis 12. April 1972 in Nürnberg. *Zeitschrift für Pädagogik*, 18 (3): S. 457-462.
- MADAUS, G. F., STUFFLEBEAM, D. L. & WEBSTER, J. W. (1983). *Program Evaluation: A historical Overview*. In D. L. STUFFLEBEAM, G. F. MADAUS, & M. SCRIVEN (Hrsg.) *Evaluation models - viewpoints on educational and human services evaluation* (S. 3-22). Boston.

- MITTELSTRAß, J. (Hrsg.) (2004) *Enzyklopädie, Philosophie und Wissenschaftstheorie*, Band 1-4, Sonderausgabe, Stuttgart.
- PSARROS, N. (2000) Merkmal, Leistung oder Anerkennung? Drei Betrachtungen der Personalität. *Dialektik*, 2: S. 39 - 56.
- POWER, M. (1999) *The audit society: Rituals of Verification*, Oxford.
- ROBINSON, S. B. (1975): *Bildungsreform als Revision des Curriculum und ein Strukturkonzept für Curriculumentwicklung*. 5. Aufl., Darmstadt.
- ROSSI, P. H., FREEMAN, H. E., HOFMANN, G. (1988). *Programm-Evaluation. Einführung in die Methoden angewandter Sozialforschung*. Stuttgart.
- ROTH, H. (1966) *Pädagogische Anthropologie*. Band I: Bildsamkeit und Bestimmung. Hannover.
- SHADISH, W. R., COOK, T. D., LEVITON, L. D. (1991). *Foundations of Program Evaluation: Theories of Practice*. Newbury Park, CA.

Peter Krope

Vom Repräsentationstheorem zur Prädikation auf empirischer Grundlage

1 Die Fragestellung

In den Jahren von Mitte 2008 bis Anfang 2010 ist im Zentrum für Konstruktive Erziehungswissenschaft (ZKE) eine Studie durchgeführt worden, in der die Lebenssituation von Menschen ohne Aufenthaltsrecht in Schleswig-Holstein untersucht wurde (KROPE ET AL. 2010). Der Auftrag bestand darin, einen Beitrag zur Bestimmung der Anzahl von Menschen ohne Aufenthaltsrecht zu leisten, die Lebenssituation in zentralen Lebensbereichen zu beschreiben sowie (über den empirischen Ansatz hinaus) ein Angebot für Problemlösungen zu entwickeln.

Als Ausgangspunkt für die Entwicklung der Problemlösungen war in der Studie untersucht worden, wie Menschen in Schleswig-Holstein mit Menschen ohne Aufenthaltsrecht umgehen. Theoretische Grundlagen dieses Teils der Untersuchung war eine Arbeit von LORENZ (2005) mit dem Titel „Das Eigene und das Fremde im Dialog“. Für die Beschreibung der Begegnung zwischen Eigenem und Fremdem wurden in Anlehnung an LORENZ die vier Kategorien Abgrenzung, Dominanz, Submission und Dialog entwickelt. Zur Durchführung der Untersuchung war ein Fragebogen, der Migrationsfragebogen, konstruiert worden, bestehend aus fünf Multiple-Choice Items mit jeweils vier Distraktoren. Jeder der vier Distraktoren repräsentiert eine der vier Kategorien. In den Items werden Probleme aus Familienalltagen von Menschen ohne Aufenthaltsrecht beschrieben. In jedem Item wird ein anderes Problem geschildert. Die Konstruktion und der Einsatz des Fragebogens waren an der Klassischen Testtheorie (vgl. LIENERT 1967) orientiert. Dementsprechend war großer Wert auf die Einhaltung der Standards für die Gütekriterien eines Tests gelegt.

Die Personen, denen der Migrationsfragebogen zur Beantwortung vorgelegt wurde, waren aufgefordert anzunehmen, sie würden zu den Problemen aus den Familienalltagen der Menschen ohne Aufenthaltsrecht um Rat gefragt. Sie sollten angeben, auf welche Weise grundsätzlich mit den geschilderten Problemen umgegangen werden könnte. Die befragten Personen sollten nicht ihr aktuelles Tun oder ihr Tun in Akutsituationen zum Maßstab nehmen. Vielmehr waren sie nach Grundsätzen ihres Handelns gefragt. Dabei wurden jene Personen, die in der Migrationsberatung tätig sind, in einer ersten Befragung um Rat im Rahmen ihres

Berufs gefragt. Unmittelbar darauf wurden sie ein zweites Mal mit dem Migrationsfragebogen und identischen Items befragt und gebeten, als Privatpersonen zu antworten. Das Ergebnis war ein hochsignifikanter Unterschied zwischen den Antworten in beruflichem Rahmen und den Antworten vor privatem Hintergrund in diesem Personenkreis.

Bei der Analyse der Differenzen stieß das ZKE-Team auf die Behauptung einer der befragten Personen, sie habe privat anders geantwortet als beruflich, weil die Fragen des Migrationsfragebogens inhomogen seien. Diese Inhomogenitätsbehauptung einer Versuchsperson ist der Ausgangspunkt für die folgenden Ausführungen. In der klassischen Testtheorie wird die Annahme vertreten, dass die Kommentare von Versuchspersonen zu den Gütekriterien der ihnen vorgelegten Verfahren nicht nachvollziehbar seien. Im Allgemeinen könnten die fraglichen Kennwerte nicht sicher nach Augenschein beurteilt werden (DOMNICK und KROPE 1972, 79). Nun war als theoretische Grundlage für die Studie über Menschen ohne Aufenthaltsrecht in Schleswig-Holstein der methodische Konstruktivismus gewählt. Das methodisch-konstruktive Wissenschaftsverständnis hat die empirische Untersuchung mit dem Migrationsfragebogen bestimmt. Zu diesem Zweck wurde der Ansatz der klassischen Testtheorie in einer methodisch-konstruktiv re-interpretierten Version verwendet. Aus der Paradigmenwahl ergibt sich die Methodenfrage, die in den folgenden Ausführungen beantwortet werden soll. Sie lautet: Wie kann auf methodisch-konstruktiver Basis entschieden werden, wenn eine Versuchsperson mit ihrer Bemerkung über die Homogenität eines Verfahrens den Einzelfall wahrnimmt und ein Forschungsteam mit seinen Homogenitätsaussagen zum selben Verfahren über den Einzelfall hinauszugehen beansprucht?

In den folgenden Abschnitten 2 und 3 werden die grundsätzlichen Überlegungen ausgeführt, vor deren Hintergrund die Frage zu beantworten ist.

2 Das Repräsentationstheorem

Bei den Erkenntnisbemühungen um die Homogenität der Items des Migrationsfragebogens geht es um die Beziehungen zwischen empirischen Größen einerseits und Zahlen andererseits.

Diese Beziehungen werden in der klassischen Meßtheorie nach SUPPES und ZINNES (1967) im Repräsentationstheorem dargestellt. Formal wird dazu erstens ein empirisches Relativ beschrieben durch

$$(1) \quad \langle E, R_1, \dots, R_n \rangle.$$

In (1) stellt E die Menge der Objekte dar. Durch R_1, \dots, R_n werden die verschiedenartigen Relationen dargestellt. Eine auf zwei beliebige Objekte aus E bezogene Relation könnte darin

bestehen, dass ein Student b sich dogmatischer verhält als ein Student a . Zweitens wird ein numerisches Relativ

$$(2) \quad \langle R, S_1, \dots, S_n \rangle$$

definiert. In (2) wird die Menge aller reellen Zahlen als Objektmenge R betrachtet; S_1, \dots, S_n steht für unterschiedliche Typen von Relationen. Ein Beispiel für eine derartige Relation sei die Kleiner-größer-Relation, z.B. $3 < 4$. Unter bestimmten Bedingungen lässt sich das empirische Relativ in das numerische Relativ abbilden. Für zwei Objekte a und b mit der angenommenen Ordnungsrelation könnte die formale Darstellung folgendermaßen aussehen:

$$(3) \quad a \zeta b \leftrightarrow \varphi(a) < \varphi(b).$$

Dabei steht das Zeichen ζ für die oben angenommene Ordnungsrelation im empirischen Relativ, \leftrightarrow für „genau dann, wenn“ und der griechische Buchstabe φ für die Zuordnungsfunktion. In der Meßtheorie geht es darum, Regeln zu nennen, die im empirischen Relativ erfüllt sein müssen, damit es auf das numerische Relativ abgebildet werden kann. Diese Aufgabe soll durch das Repräsentationstheorem wahrgenommen werden.

Die Abbildtheorie wird in Frage gestellt durch die in der Mathematik geführte Grundlagendiskussion, die der niederländische Mathematiker BROUWER mit zwei Aufsätzen in den Jahren 1907 und 1908 eröffnet hat (BROUWER 1907, 1908). „Den Ausgangspunkt [der Überlegungen BROUWERS] bildet [...] der Unendlichkeitsbegriff. Nach BROUWER darf das Unendliche niemals als eine fertige Gesamtheit [...] betrachtet werden, sondern ist als eine bloße *Möglichkeit* des unbegrenzten Fortschreitens (als 'potentiell Unendliches') aufzufassen. Die Aussage, daß es unendlich viele Zahlen gibt, darf also z.B. nicht so aufgefaßt werden, als existierten alle unendlich vielen natürlichen Zahlen 'an sich' in einem Bereich idealer Objekte; vielmehr ist sie so zu interpretieren, daß man zu jeder natürlichen Zahl eine größere (etwa den Nachfolger) angeben kann“ (STEGMÜLLER 1969, 438).

Unmittelbar betroffen von den Entdeckungen BROUWERS sind die Mathematik und die Logik. Aber die Folgen reichen über Mathematik und Logik hinaus. BROUWERS Ergebnisse stehen allen Wissenschaftsprogrammen entgegen, die von der Annahme ausgehen, dass die Realität, über die der Wissenschaftler und die Wissenschaftlerin arbeiten, unabhängig vom erkennenden Subjekt existiere und strukturiert sei.

3 Die Prädikation auf empirischer Grundlage

Der Methodische Konstruktivismus nimmt die Ergebnisse BROUWERS in einer Methodologie auf, die „realistische“ Aussagen zulässt, ohne „realistische“ Voraussetzungen zu machen (GETHMANN 1973, 43, unter Verweis auf LORENZEN 1962). Im Folgenden wird eine Methode

zur Lösung des Repräsentationsproblems skizziert, die vor diesem Hintergrund entwickelt worden ist. Die Methode wird als Prädikation auf empirischer Grundlage bezeichnet²⁰.

Im Mittelpunkt der Methode steht die Prädikation. Das ist der Vorgang, bei dem gewissen Gegenständen Prädikatoren zugesprochen oder abgesprochen werden. Ein Beispiel für eine Prädikation im Rahmen der Auseinandersetzung um die Homogenität des Migrationsfragebogens ist die Aussage über das Handeln einer Person im Umgang mit Menschen ohne Aufenthaltsrecht. Der Satz lässt sich symbolisch folgendermaßen darstellen:

$$(4) \quad N \pi p.$$

Dabei wird zur Bezeichnung des Eigennamens der befragten Person der Buchstabe N und für die (abgrenzende, dominante, submissive oder dialogische) Handlung der Tatprädikator mit dem Symbol p gewählt. Der griechische Buchstabe π steht für die Tatkopula. Alltagssprachlich könnte das Wort „tut“ dafür verwendet werden: „Die Person N tut (abgrenzend, dominant, submissiv oder dialogisch) handeln“. Durch den Asterisk in π^* sei eine vergangene Handlung ausgedrückt: „Die Person N tat (abgrenzend, dominant, submissiv oder dialogisch) handeln“. Bei der Prädikation wird das Wort „Gegenstand“ verwendet, wenn gesagt werden soll, wovon die Rede ist. Das, was einem Gegenstand mit einem Prädikator zugesprochen wird, wird als Eigenschaft bezeichnet. Im Beispiel ist der Person die Eigenschaft zugesprochen worden, im Umgang mit Menschen ohne Aufenthaltsrecht abgrenzend, dominant, submissiv oder dialogisch gewesen zu sein. Die Festlegung auf eine der vier Kategorien spielt bei den vorliegenden Überlegungen keine Rolle.

Eine der Bedingungen, die erfüllt sein müssen, damit gesagt werden kann, einer Person sei ein Prädikator zu Recht zugesprochen, ist die Untersuchung dieser Person nach dem Maßstab der empirischen Forschung. Gestützt auf empirische Befunde kann die Behauptung verteidigt werden, die Person N habe im Umgang mit Menschen ohne Aufenthaltsrecht abgrenzend, dominant, submissiv oder dialogisch gehandelt. Gelingt die Verteidigung, wird aus der Behauptung $N \pi^* p$ eine wahre Aussage, symbolisch dargestellt als

$$(5) \quad (N \pi^* p) \varepsilon w$$

mit w für „empirisch wahr“ und dem griechischen Buchstaben ε für das Zusprechen des Prädikators „empirisch wahr“. Wenn Übereinstimmung darin besteht, dass die Formulierung und die Testung von Hypothesen zum Standard einer empirischen Untersuchung gehört, dann kann für die konstruktive Methode der Prädikation auf empirischer Grundlage auch gesagt

²⁰ Zur ausführlichen Darstellung vgl. KROPE und LORENZ (1993, 14-22) und KROPE (2000, 34-42).

werden: Die Aussage über die Eigenschaft einer Person ist empirisch wahr, wenn die vorhergesagte Eigenschaft $((N \pi p)_v)$ und die empirisch ermittelte Eigenschaft $((N \pi^* p)_e)$ dieser Person gleichwertig sind. In symbolischer Darstellung:

$$(6) \quad (N \pi^* p) \varepsilon w \Leftrightarrow (N \pi p)_v \sim (N \pi^* p)_e$$

(mit \Leftrightarrow als Definitionszeichen und \sim als Zeichen für „äquivalent“)²¹. Ein Vergleich der Zeile (6) mit der Zeile (3) soll zeigen, dass die Methode der Prädikation auf empirischer Grundlage mit dem Repräsentationstheorem konkurriert. In beiden Fällen geht es um die Zuordnung von Eigenschaften zu Gegenständen. Im Unterschied zum Vorgehen nach dem Repräsentationstheorem nimmt der Methodische Konstruktivismus aber seinen Anfang in der Sprache.

Im folgenden vierten Abschnitt wird das Problem dargestellt, für das in den vorliegenden Ausführungen nach einer grundsätzlichen Lösung gesucht wird.

4 Homogen oder nicht homogen

In der klassischen Testtheorie wird ein Test in dem Maße „homogen“ genannt, in dem seine Items in eindimensionaler Weise genau eine Eigenschaft messen. Die Items eines homogenen Tests erfassen die gleiche Eigenschaft. In diesem Sinne schreibt LIENERT: „Die Aufgaben eines Tests können mehr oder weniger homogen und im Grenzfall heterogen sein. Homogenität bedeutet dabei inhaltliche Einheitlichkeit bei vollkommen erhaltener Unabhängigkeit der einzelnen Aufgaben voneinander“ (LIENERT 1967, 46).

LOEVINGER (1947) legt ein Konzept zur Berechnung für einen Homogenitätsindex vor, das auf der Anordnung der Items nach ansteigender Schwierigkeit und auf dem Vergleich dreier aus der Anordnung resultierender Varianzen beruht. Nach Lienert (1967, 254) hat der Homogenitätsindex für die Praxis der Testkonstruktion allerdings keine unmittelbare Bedeutung. Ein Konzept zur einfachen rechnerischen Bestimmung schlagen BORTZ und DÖRING (1995, 200 – 201) vor. Der Ansatz geht von dem Gedanken aus, dass die Items eines eindimensionalen Tests, die genau die gleiche Eigenschaft erfassen, miteinander korrelieren. Damit liefert der (nach einer Fisher-Z-Transformation berechnete) Durchschnitt $(\bar{r}_{ii'})$ der paarweise korrelierten Items eine quantitative Aussage über die Homogenität. Für die in der Tabelle aufgeführten Werte des Migrationsfragebogens ist $\bar{r}_{ii'} = 0,2727$ zu berechnen. Zur Bewertung der Höhe des Koeffizienten empfehlen BORTZ und DÖRING (1995, 201) einen Akzeptanzbereich

²¹ Der Ausdruck „Vorhersage“ wird im Sinne des von KROPE (2000) auf methodisch-konstruktiver Grundlage (ansatzweise) re-interpretierten Terminus „Hypothese“ verwendet. Es wird erwartet, dass durch (6) die Protokoll- bzw. Basissatzproblematik als gelöst gelten kann.

von 0,2 bis 0,4. Nach diesem Kriterium ist für den Migrationsfragebogen eine hinreichende Homogenität gewährleistet.

	Item 1	Item 2	Item 3	Item 4	Item 5
Item 1	1,0000				
Item 2	0,0370	1,0000			
Item 3	0,2371	0,3519	1,0000		
Item 4	0,3226	0,3147	0,0672	1,0000	
Item 5	0,2610	0,3651	0,2932	0,4774	1,0000

Tabelle: Korrelationskoeffizienten zur Homogenitätsberechnung für den Migrationsfragebogen.

Im Verlauf der Erhebungen hat eine der befragten Personen die Behauptung aufgestellt, der Migrationsfragebogen sei eben nicht homogen. Bei einer Nachbefragung zu einem zweiten Termin gab der Befragte an, als Leiter einer Ausländerbehörde darauf angewiesen zu sein, primär juristische Sachverhalte zu beurteilen. In vier der vorgelegten fünf Fragen habe dies kein Problem dargestellt. Aber die fünfte Frage beziehe sich auf die rechtliche Entwicklung in einem kleineren osteuropäischen Land, deren aktuellen Stand er aus dem Stegreif nicht sicher beurteilen könne. Deswegen habe er von Inhomogenität gesprochen und behauptet, dass die fünf Fragen unterschiedliche Eigenschaften erfassen, und er nicht vor seinem beruflichen Hintergrund, wohl aber im privaten Rahmen habe antworten können²².

5 Die Auseinandersetzung

Mit der Frage nach der Gleichheit von Eigenschaften wird ein grundsätzliches Problem aufgeworfen. Wie können Wissenschaftler und Wissenschaftlerinnen sicherstellen, dass sie mit den Items eines Fragebogens Vergleichbares untersuchen? Dieser Frage wird im vorliegenden Abschnitt 5 nachgegangen.

Im Methodischen Konstruktivismus gilt Gleiches als Handlungsschema, das durch Abstraktion aus Ungleichem gewonnen wurde. Dabei wird als Abstraktion jener Vorgang bezeichnet, bei dem man von allem absieht, was zwei Ausdrücke unterscheidet, und nur das ins Auge fasst, worin sie gleichzusetzen sind. SEIFFERT (1973, 38) beschreibt die Abstraktion als einen „Übergang von einer ‚Ungleichheit‘ zu einer ‚Gleichheit‘: zunächst ungleiche Gegenstände werden gleichgesetzt, indem man nur auf bestimmte ihrer Merkmale achtet“. Und er folgert: „Gleichheit und Ungleichheit gibt es immer nur in Bezug auf bestimmte Eigenschaften der gegebenen Gegenstände. Zwei Gegenstände sind niemals absolut gleich oder ungleich, sondern immer nur gleich oder ungleich in Bezug auf bestimmte Eigenschaften“ (SEIFFERT 1973, 40). Das über Gleichheit und Ungleichheit Gesagte gilt für einzelne Wörter ebenso wie für

²² Die Frage, ob Homogenität oder Validität gemeint ist, wird im Weiteren nicht untersucht. Eine Klärung bedürfte einer eigenen Analyse entsprechend dem im Folgenden dargelegten Argumentationsmuster.

einzelne Aussagen. So kann in den Sätzen „Wir werden der Philippina Frau S. bei der Rückkehr in ihre Heimat behilflich sein“ und „Wir werden der Philippina Frau S. helfen, sich in die deutsche Arbeitswelt einzupassen“ sowohl die Gleichheit wie auch die Ungleichheit ins Auge gefasst werden. Insofern im ersten Satz Abgrenzung und im zweiten Dominanz im Vordergrund steht, wird Unterschiedliches ausgesagt. Sieht man aber von diesem Unterschied ab, kann gesagt werden, dass Satz 1 und Satz 2 gleich sind. In beiden Sätzen wird beschrieben, wie Menschen in Schleswig-Holstein mit Menschen ohne Aufenthaltsrecht umgehen. Gleichheit ist „Gleichheit in Bezug auf etwas“. Zunächst Ungleiches wurde gleichgesetzt, indem nur auf bestimmte Inhalte geachtet wurde.

Wenn mehrere Einzeldinge immer nur in gewisser Hinsicht gleich sind, dann setzt Gleichheit die Annahme eines Standpunktes voraus, von dem aus ein Einzelnes als Wiederholung eines anderen Einzelnen bestimmt werden kann. Damit wird die Frage nach Gleichheit zur Frage nach dem präzifizierenden Subjekt.

Der Nachweis dieses Zusammenhangs geht von einem Aufsatz aus, in dem LORENZEN (1962) zur Fundierung der konstruktiven Theorie die Existenz abstrakter Objekte mit Hilfe von Äquivalenzrelationen beweist²³. Als „Äquivalenzrelationen“ werden dabei Beziehungen zwischen Aussagen verstanden werden, die das Merkmal der Reflexivität, der Symmetrie und der Transitivität haben. In seinen Ausführungen nimmt LORENZEN an keiner Stelle explizit auf einen Handelnden Bezug. Vielmehr hat es zunächst den Anschein, dass eine Gleichung wie „ $x = x$ “ selbst, gleichsam „von sich aus“, auf Reflexivität hinweist. Aber bereits die Übertragung auf einfache sprachliche Wendungen verdeutlicht die Unhaltbarkeit dieser Annahme, wie ein Beispiel (nach SEIFFERT 1973, 97) zeigt. Die Aussage „A kennt B“ kann sowohl nichtreflexiv als auch reflexiv sein. Sie ist nichtreflexiv, wenn $A \neq B$, wenn beispielsweise mit „A“ Herr A. und mit „B“ Herr B. gemeint ist; und sie ist reflexiv unter der Bedingung $A = B$, was etwa der Fall sein kann, wenn mit „A“ Herr A. und mit „B“ Herr A. zu einem früheren Zeitpunkt bezeichnet wird. Indes: Ist dann noch von demselben Herrn A. die Rede? Überträgt man die Aussage „A kennt B“ zurück auf die Gleichung „ $x = x$ “, dann stellt sich ebenfalls die Frage, in welcher Hinsicht die Gleichheit behauptet wird. Denn wie gesagt gibt es Gleichheit immer nur im Hinblick auf einen bestimmten Gesichtspunkt. Absolut gleich wäre ein Buchstabe x nur in „ein und derselben Realisierung“ des „Buchstabens zum gleichen Zeitpunkt“ (SEIFFERT 1973, 40). Wie also sollte die Frage der Gleichheit entschieden werden, wenn ein Subjekt der Entscheidung ausgeklammert bleibt?

²³ Ausführlich vgl. KROPE (1986).

In der Logischen Propädeutik haben KAMLAH und LORENZEN (1973, 45f.) die Träger der Handlungen dort eingeführt, wo es um explizite Vereinbarung und um exemplarische Einübung der sprachlichen Handlungs-Schemata geht. Die Frage ist von daher: Wer anders soll und kann beurteilen, ob die Wiederholung eines Einzelnen als Gleichheit bestimmt werden kann, als diejenigen, die sich verbindlich auf ein Handlungs-Schema verpflichtet und eingeübt haben. Folgt man der Argumentation, dann ist – Legitimität vorausgesetzt – der Zweck, den die prädzierenden Personen verfolgen, ein Maßstab für die Beurteilung von Gleichheit und Ungleichheit. Bei den beiden Sätzen über die Philippina Frau S. handelt es sich um Ungleiches, wenn der Zweck verfolgt wird, vor dem gewählten theoretischen Hintergrund nach LORENZ (2005) auf den Unterschied zwischen abgrenzendem und dominantem Handeln aufmerksam zu machen. Damit wird in der Auseinandersetzung um Gleichheit oder Ungleichheit der von den Items erfassten Eigenschaften die Darstellung der Zwecke zur Grundlage einer methodisch geregelten Auseinandersetzung. Während das wissenschaftliche Team der empirischen Studie den Zweck gesetzt hat, mit allen fünf Items zu untersuchen, wie Menschen in Schleswig-Holstein mit Menschen ohne Aufenthaltsrecht umgehen, geht es dem Abteilungsleiter bei einem der fünf Items um Informationen über die rechtliche Situation in einem der Herkunftsländer.

6 Fazit

Ausgangspunkt der vorliegenden Ausführungen ist ein Dissens in einer aktuellen empirischen Studie über die Itemhomogenität eines Fragebogenverfahrens, ausgelöst durch den Einspruch einer Versuchsperson. Die daran anschließende Fragestellung lautet: Wie kann der Dissens auf methodisch-konstruktiver Basis entschieden werden, wenn eine Versuchsperson mit ihrer Bemerkung über die Homogenität eines Verfahrens den Einzelfall wahrnimmt und ein Forschungsteam mit seinen Homogenitätsaussagen zum selben Verfahren über den Einzelfall hinauszugehen beansprucht?

Der Einzelfall besteht darin, dass eine befragte Person auf ein Item als Beispiel und auf vier Items als Gegenbeispiele verweist und damit exemplarisch ihre Erfahrung vorführt, die sie im vorwissenschaftlichen Alltagsbereich gewonnen hat. Mit der auf methodisch-konstruktiver Basis empirisch erzeugten Erfahrung dagegen verfügt das ZKE-Team in wissenschaftlicher Rede über kontextunabhängige Sätze, die unabhängig vom Alltagsgebrauch sind.

Aussagen über die Zwecke können als Brücke zwischen vorwissenschaftlichem und wissenschaftlichem Bereich betrachtet werden, indem auf ihrer Grundlage die Alltagserfahrungen für die Belange der Studie wie folgt aufgenommen, modifiziert oder verworfen werden. Ihre

Kenntnis kann dazu beitragen, die Differenz zwischen den Antworten im beruflichen und privaten Rahmen zu erklären. Oder sie kann der Überarbeitung der Items des Migrationsfragebogens dienen. Oder die Kenntnis, gewonnen am Besonderen, kann darauf verweisen, dass die im Alltag und in der Wissenschaft verfolgten Zwecke unverträglich sind. Die Zwecke können nicht zugleich verfolgt werden. Die Ausländerbehörde ist ein ausführendes Organ der Legislative. Ihr Zweck besteht darin, geltendes Recht im Umgang mit Menschen ohne Aufenthaltsrecht so anzuwenden, wie es vom Gesetzgeber gewollt ist. Der Zweck der Studie dagegen bestand darin, für die Beantwortung der Frage, wie die Begegnung mit Menschen ohne Aufenthaltsrecht zu gestalten ist, überhaupt erst eine wissenschaftliche Grundlage zu schaffen.

7 Literatur

- BORTZ, JÜRGEN, DÖRING, NICOLA: Forschungsmethoden und Evaluation. Berlin usw. ²1995.
- BROUWER, LUITZEN EGBERTUS JAN: Over de Grondslagen der Wiskunde. Amsterdam/Leipzig 1907.
- BROUWER, LUITZEN EGBERTUS JAN: De onbetrouwbaarheid der logische principes. In: Tijdschrift voor wijsbegeerte. 2 (1908), 152 – 158.
- DOMNICK, JOACHIM, KROPE, PETER: Student und Latinum. Untersuchung zum Bestand und Bedarf an Lateinkenntnissen bei Studenten. Weinheim/Basel 1972.
- GETHMANN, CARL FRIEDRICH: Allgemeinheit. In: KRINGS, HERMANN, BAUMGARTNER, HANS MICHAEL, WILD, CHRISTOPH: Handbuch philosophischer Grundbegriffe. Studienausgabe Band 1. München 1973, 32 – 51.
- KAMLAH, WILHELM, LORENZEN, PAUL: Logische Propädeutik. Vorschule des vernünftigen Redens. Mannheim usw. ²1973.
- KROPE, PETER: Der Beitrag der Prüfungsforschung zur erziehungswissenschaftlichen Theoriebildung. Dargestellt am Problem des Allgemeinen und des Einzelnen. In: Hochschulausbildung. 4/1986, 29 – 41.
- KROPE, PETER: Konstruktive Pädagogische Diagnostik. Münster usw. ²2000.
- KROPE, PETER, LORENZ, PAUL (Hg.): Pädagogik zwischen Dogmatismus und Autonomie. Untersuchungen zum pädagogischen Selbstverständnis im Osten und im Westen der Bundesrepublik Deutschland. Münster/New York 1993.

- KROPE, PETER, VOLK, RUTH, unter Mitwirkung von ARNDT, GERRIT, BUCHHEIT, JULIA, GÖRNER, JENNY, GÖSER, TORBEN, KOHRS, ARNE, LATUS, KNUT, PETERSEN, JOHANNES PETER, PRIEB, WOLF, WILDER, NICOLAUS: Einladung zum Dialog. Eine methodisch-konstruktive Studie über Menschen ohne Aufenthaltsrecht in Schleswig-Holstein. Kiel 2010 (Forschungsbericht, im Druck).
- LIENERT, GUSTAV A.: Testaufbau und Testanalyse. Weinheim/Berlin ²1967.
- LOEVINGER, JANE: A systematic approach to the construction and evaluation of tests of ability. In: Psychol. Monogr. 61, 1947.
- LORENZ, KUNO: Das Eigene und das Fremde im Dialog. In: RUSTERHOLZ, PETER, MOSER, RUPERT (Hg.): Wie verstehen wir Fremdes? Bern usw. 2005, 137 – 154.
- LORENZEN, PAUL: Gleichheit und Abstraktion. In: Ratio 4 (1962), 77 – 81.
- SEIFFERT, HELMUT: Einführung in die Logik. Logische Propädeutik und formale Logik. München 1973.
- STEGMÜLLER, WOLFGANG: Hauptströmungen der Gegenwartsphilosophie. Eine kritische Einführung. Stuttgart ⁴1969.
- SUPPES, PATRICK, ZINNES, JOSEPH L.: Basic Measurement Theory. In: LUCE, ROBERT D., BUSH, ROBERT, GALANTER, EUGENE (eds.): Handbook of Mathematical Psychology. Vol. I. New York etc. ²1967, 1 - 76.

Folgende Monographien zur Prüfungsforschung sind bisher erschienen:

- Nr. 1 Peter Kroppe, Hans-Werner Zumach
Leistungsmessung als Deformation der Wirklichkeit
- Nr. 2 Peter Kroppe, Lothar Viehöfer
Prüfungsangst und Aufgabenkomplexität
- Nr. 3 Arne Kohrs
Bibliographie zur Angst unter besonderer Berücksichtigung der Prüfungsangst
- Nr. 4 Lothar Viehöfer
Entwicklung eines Instruments zur Messung motorischer Reaktionen
- Nr. 5 Peter Kroppe
Die Kieler Affekt-Adjektiv-Liste (KAAL) zur Messung situativer Prüfungsangst
- Nr. 6 Jochen Hestermann
Kooperative Gruppenprüfungen als Alternative zu Einzelprüfungen an Hochschulen
- Nr. 7 Olaf Albers
Das Problem der Ausgangswerte bei der Physiologischen Messung von Prüfungsangst
- Nr. 8 Peter Kroppe, Arne Kohrs
Prüfungsangst und kooperative Gruppenprüfung
- Nr. 9 Cornelia Klink
Mündliche Hochschulprüfungen und Labeling Approach
- Nr. 10 Peter Kroppe, Arne Kohrs
Zur medikamentösen Behandlung von Prüfungsangst bei Studenten - eine Methodenstudie
- Nr. 11 Burkhard Seidler
Inhaltsanalyse und Leistungsmessung - eine Methodenstudie
- Nr. 12 Peter Kroppe, Arne Kohrs, Karin Recklies, Maria Peters, Jutta Schneider
Über kriterienorientierte, nicht-quantitative Leistungsmessung - Ein Untersuchungsbericht
- Nr. 13 Olaf Albers, Beate Conzen, Axel Klimek, Cornelia Klink, Wolfgang König, Peter Kroppe, Annette Krüger, Karin Lüdders (Hg.)
Das Medizinerwahlgespräch im Dialog

- Nr. 14 Peter Kroppe (Hg.)
Prüfungsforschung in der Bundesrepublik - Zwischenbilanz und Perspektiven
- Nr. 15 Peter Kroppe, Boleslaw Niemierko (Ed.)
Proceedings of the 1992 Workshop on School Examinations
- Nr. 16 Peter Kroppe
Berichtszeugnisse: Die Quelle der Mißverständnisse
- Nr. 17 Peter Kroppe, Boleslaw Niemierko
Educational Measurement in Germany and Poland
- Nr. 18 Alina Karaškiewicz, Peter Kroppe (Hg.)
Pädagogische Diagnostik zwischen Standardisierung und Individualisierung II – ein deutsch-polnischer Werkstattbericht

Die Monographien zur Prüfungsforschung sind, soweit nicht vergriffen, zu beziehen direkt über die Verfasser oder über das Projekt Prüfungsforschung am Institut für Pädagogik der Universität, 24118 Kiel, Olshausenstraße 75.

Die Monographien zur Prüfungsforschung werden ab Heft 16, 3. Auflage, unter der folgenden Adresse als Word-Dateien elektronisch publiziert und kostenlos zur Verfügung gestellt:

<http://www.zke-kiel.de>

Die Monographien werden inhaltsgleich in geringer Auflage in traditioneller Weise gedruckt und vorrätig gehalten. Auf diese Weise soll in Zweifelsfällen die Autorenschaft gesichert und der Originalzustand der Textdateien überprüft werden können.