

Zur Einführung

Die Psychologie ist ein Fach mit außerordentlich vielfältigen methodischen Herausforderungen. Selbst innerhalb der quantitativen Forschungstradition nutzen wir ein Sammelsurium der verschiedensten Arbeitsansätze. So bringt die Forschung ein buntes Füllhorn an Befunden nach dem anderen hervor. Welche inhaltlichen Aussagen aufeinander bezogen werden können, an welchen Stellen es Widersprüche gibt oder welche Ansätze schlicht unvergleichbare Ergebnisse liefern, bleibt dabei meist offen. Die Befunde der psychologischen Forschung darzustellen bedeutet in der Regel, sie *aufzuzählen*. Keineswegs fügen sich die Forschungsergebnisse wohlgefällig zu einem kohärenten Wissensgebäude. Besonders augenfällig ist dabei eine Bruchstelle zwischen der Klassischen Psychologie einerseits und der Biopsychologie oder Neurowissenschaft andererseits.

Schon aus der Presse weithin bekannt sind zum Beispiel die neurowissenschaftlichen Untersuchungen von Benjamin Libet (1985), der sich als Pionier in der Willensforschung einen Namen machte. Libet entwickelte ein Versuchsdesign, in dem es genau *eine Versuchsbedingung* gab. Alle Probanden erhielten dieselbe Instruktion. Der Versuchsleiter bat seine Studienteilnehmer, spontan die Hand zu bewegen, wenn ihnen der Sinn danach stünde. Dabei sollten sie sich den Zeitpunkt ihrer Willensbildung genau merken, wobei ihnen eine große, an die Wand projizierte Uhr mit rasch drehendem Zeiger eine Einschätzung im Millisekundenbereich ermöglichte. Über ein Elektromyogramm konnte der Versuchsleiter zudem zeitlich hochauflösend den Moment der Handbewegung erfassen. Parallel zum Geschehen wurden EEG-Daten von den Probanden aufgezeichnet.

Den Auskünften der Versuchspersonen zufolge bildete sich der Wille zum Handheben jeweils etwa 200 Millisekunden vor der ausgeführten Bewegung. Aber schon 550 Millisekunden vor der Handbewegung zeigte das EEG eine charakteristische Amplitudenveränderung. Der Wille selbst, so argumentierten Neurowissenschaftler in der Folge, bewirke gar keine Handlungen. Vielmehr bringe das Gehirn erst das subjektive Gefühl einer Willensbildung und dann die entsprechende Handlung hervor. "Philosophers studying conscious free will have discussed whether conscious intentions could cause actions, but modern neuroscience rejects the idea of mind-body causation" (Haggard, 2005, S. 289). In einer jüngeren Nachfolgestudie (Soon et al., 2008) konnten Handlungsentscheidungen sogar auf der Basis von Gehirnbild-Daten zutreffend vorhergesagt werden, die der bewussten Willensbildung nicht nur um 350 Millisekunden, sondern um etwa 10 Sekunden vorgelagert waren.

Studien in der Libet-Tradition werden meist unter inhaltlichen Gesichtspunkten besprochen. Aber auch und gerade die Methodik ist ein lohnender Diskussionsgegenstand. Am Beispiel der Willensforschung zeichnet sich paradigmatisch eine Diskrepanz zwischen der neurowissenschaftlichen Forschung einerseits und klassisch-psychologischen Untersuchungen andererseits ab: eine Diskrepanz, die sehr wohl ergebnisrelevant ist.

Neurowissenschaftler rekurrieren auf Studien in der Libet-Tradition als „Experimente“. Dem Vokabular der Psychologischen Methodenlehre entspricht das jedoch nicht. Zwar würden auch klassisch ausgebildete Psychologen „Experimente“ durchführen, um Kausalhypothesen zu überprüfen; geboten wäre damit in der Willensforschung aber ein ganz andersartiges Vorgehen. In jedem Falle müssten mindestens *zwei verschiedene Versuchsbedingungen* realisiert werden und nicht nur eine („Bedingungsvariation“). Der Versuchsleiter müsste die Intentionen seiner Versuchsteilnehmer gezielt beeinflussen und müsste erfassen, ob sich die Handlungen der Probanden entsprechend ändern. Eine Studie zum Verhältnis von Wille und Handlung könnte dann bspw. so aussehen: Der Versuchsleiter erklärt den Probanden einer ersten Versuchsbedingung, dass sie für jede Handbewegung innerhalb der nächsten fünf Minuten entlohnt werden – 50 Cent pro Handbewegung. Den Versuchspersonen der zweiten Bedingung wird mitgeteilt, dass sie

nach fünf Minuten 100 Euro erhalten, sofern sie ganz still sitzen; für jede Handbewegung in der Zwischenzeit werden jedoch fünf Euro abgezogen. Durch die verschiedenen Versuchsbedingungen sollten die Probanden zu gegensätzlichen Intentionen veranlasst werden. Die einen sollten ihre Hand möglichst oft heben wollen, die anderen niemals. In einem postexperimentellen Interview könnte man erfragen, ob die Willenslage der Probanden tatsächlich annahmengenau ausfiel. Geben die Versuchsteilnehmer der ersten Bedingung an, dass sie in der Tat den Willen ausbildeten, immer wieder ihre Hand zu heben, vielleicht, dass sie diesen Willen sogar angesichts markanter Ermüdungserscheinungen im Handgelenk aufrechterhielten, während die Teilnehmer der zweiten Bedingung ihre Hand erklärtermaßen absichtlich stillhielten, ja möglicherweise, dass sie angesichts eines Juckreizes mit Bedacht die Neigung zum Kratzen unterdrückten, so kann die experimentelle Bedingungsvariation als gelungen betrachtet werden. Findet sich im Anschluss beim Vergleich der Versuchsgruppen ein statistisch signifikanter Unterschied in der Häufigkeit von Handbewegungen, ist die Schlussfolgerung zu ziehen, dass eine Kausalbeziehung zwischen Wille und Handlung besteht. Ein solches Ergebnis wäre natürlich mehr als wahrscheinlich.

Wie aber stünden dann die Schlussfolgerungen der Forscher zu einander? Auf den ersten Blick schienen sich Klassische Psychologen und Neurowissenschaftler direkt zu widersprechen. Eine Kausalbeziehung zwischen Wille und Handlung würde von den einen affirmiert, von den anderen zurückgewiesen.

Andererseits ist es durchaus fraglich, ob die Forscher überhaupt dasselbe meinen, wenn sie von Kausalbeziehungen sprechen. Immerhin ist die methodische Vorgehensweise in den beiden (Sub-)Disziplinen ziemlich unterschiedlich, wenn über die Annahme oder Ablehnung einer Kausalaussage entschieden werden soll. Nur die Bezeichnungen sind deckungsgleich: Alle behaupten von sich, dass sie „Experimente“ durchführen.

Die methodische Praxis der Klassischen Psychologie könnte folgendermaßen zusammengefasst werden...

1 Bei Experimenten eine Bedingungsvariation voraussetzen: In der Klassischen Psychologie bezeichnet man heute nur Untersuchungen mit Bedingungsvariation als „Experimente“. Ein solches Versuchsdesign wird methodisch als das beste und eigentlich als das einzig durchweg überzeugende Mittel betrachtet, um Kausalhypothesen zu überprüfen.

Die Bedingungsvariation ist in der Klassischen Psychologie heute so selbstverständlich, dass man Studien ohne Bedingungsvariation gar nicht mehr *Experimente* nennt. Sogar bei vorhandener Bedingungsvariation spricht man nur noch von *Quasi-Experimenten*, wenn die Probanden den verschiedenen Versuchsbedingungen nicht per Zufall zugeordnet werden. Kausale Schlussfolgerungen werden dann bereits als hoch problematisch, streng genommen als unmöglich angesehen.

Das Experiment gilt als Königsweg psychologischer Forschung. [...] In der experimentellen Forschung wird untersucht, ob die gezielte Veränderung einer Variablen durch den Versuchsleiter die (erwartete) Veränderung einer anderen Variable zur Folge hat. [...] Ein weiteres wesentliches Kriterium der experimentellen Kontrolle ist die zufällige Aufteilung der Untersuchungsteilnehmer auf die Experimental- und Kontrollgruppe (Randomisierung). Nur Studien, in denen das Prinzip der Randomisierung Anwendung findet, sind echte Experimente. [...] Wenn ein Versuchsleiter die Probanden also nicht zufällig den Versuchsbedingungen zuordnet, oder wenn Variablen, die untersucht werden sollen, gar nicht manipuliert werden können, spricht man von einer quasi-experimentellen Untersuchung. Quasi-experimentelle Untersuchungen können im Gegensatz zu einem echten Experiment nicht so interpretiert werden, dass man Aussagen über klare Ursache-Wirkung Zusammenhänge begründen könnte[.]

(Imhof, 2010, S. 76ff.)

Augenfälligerweise umfassen nun aber Studien in der Libet-Tradition keine Bedingungsvariation. So kann natürlich auch keine Zufallszuweisung zu verschiedenen Versuchsgruppen erfolgen; es gibt ja nur eine einzige Versuchsgruppe. Trotzdem gelten diese Studien in der Biopsychologie nicht etwa als methodisch schwach oder gar inakzeptabel, sondern viel eher als brillant. Ganz offenkundig werden aus den Befunden auch kausale Schlussfolgerungen gezogen. So wird bspw. eine Kausalbeziehung zwischen Wille und Handlung verneint. In der Neurowissenschaft ist die methodische Praxis also eine andere als in der Klassischen Psychologie.

② **Bei Experimenten keine Bedingungsvariation voraussetzen:** In der neurowissenschaftlichen Forschungstradition bezeichnet man auch Untersuchungen ohne Bedingungsvariation als „Experimente“ und nutzt sie, um Kausalhypothesen zu überprüfen.

Im Prinzip liegt das Ansinnen wohl nahe, die Befunde verschiedener Forschergruppen zu vergleichen und zu verbinden, wann immer man am selben Thema arbeitet. Das könnte bspw. der Fall sein, wenn Klassische Psychologen einerseits und Neurowissenschaftler andererseits Studien durchführen, um das Verhältnis von Wille und Handlung zu untersuchen. Wie aber sollen verschiedene Forschungsergebnisse zu einem kohärenten Wissensgebäude zusammenwachsen, wenn schon zentrale methodologische Grundbegriffe – wie „Kausalität“ und „Experiment“ – hier dieses und dort jenes bedeuten?

Volker Gadenne (1994) wählt drastische Worte in der *Enzyklopädie der Psychologie* zum Themenbereich *Methodologie und Methoden*, um das Durcheinander innerhalb der Kausalforschung zu beschreiben.

Die Bemühungen um *Kontrolle* in der Versuchsplanung und ein großer Teil der Entwicklungen in der Korrelationsstatistik dienen dem erklärten Ziel, Kausalbeziehungen nachzuweisen bzw. Kausalmodelle zu testen [...]. Hierbei liegt allerdings keine einheitliche und zum Teil überhaupt keine geklärte Auffassung von Kausalität zugrunde.
(Gadenne, 1994, S. 330, H.i.O.)¹

Angesichts derartiger Herausforderungen könnte es nun durchaus naheliegen, sich immer wieder und vielleicht sogar viel mehr als bislang um methodologische Klärungen zu bemühen. Auch aus der Sicht des empirisch tätigen Einzelwissenschaftlers dürfte es mehr als wünschenswert sein, dass die eigenen Arbeitsergebnisse zu einem inhaltlichen Gesamtverständnis beitragen. Sonst drohen die sorgsam produzierten Ergebnisse zusammen mit jenen der Kollegen letztlich nur eine fachliche Befundhalde zu speisen.

So scheint die Psychologie derzeit eigentlich einen besonders großen Bedarf an methodologischen Grundlagenarbeiten zu haben, wie sie traditionell der Wissenschaftstheorie zugeordnet werden. Obgleich aber das Erstarken der Biopsychologie gerade in jüngerer Zeit dringliche methodologische Fragen aufgeworfen hat, kann von einem parallelen Erstarken wissenschaftstheoretischer Diskurse innerhalb der Psychologie bislang keine Rede sein.

In seinem Artikel *Zum Aufstieg und Niedergang der Wissenschaftstheorie in der Psychologie: Eine persönliche Konstruktion* zeichnet Hans Westmeyer (2004) das Bild einer psychologiehistorischen Verfallsgeschichte. Noch in den 60er und 70er Jahren des 20. Jahrhunderts erwiesen sich zahlreiche Fachvertreter der Psychologie als wissenschaftstheoretisch außerordentlich wohlinformiert. Es gelang ihnen in vielen Bereichen, zeitaktuelle wissenschaftstheoretische Debatten zu intensiven fachinternen Diskussionen auszubauen. Wenn wir von den wissenschaftstheoretischen Früchten der 1960er und 70er Jahre noch heute zehren, dann auch deshalb, weil in der Nachfolgezeit wenig gereift ist. Aktuelle Auseinandersetzungen mit wissenschaftstheoretischen Themen können sich mit dem hohen

¹ Das Kürzel H.i.O. steht hier wie im Folgenden für „Hervorhebung im Original“.

Niveau der damaligen Diskussionen nicht messen. Die ehemals so ergiebige Beschäftigung mit wissenschaftstheoretischen Fragen droht, zu verebben.

Betrachten wir z.B. die Deutsche Gesellschaft für Psychologie, so fällt auf, dass zu wissenschaftstheoretischen Themen weniger Beiträge in mit dieser Gesellschaft verbundenen Fachzeitschriften erscheinen, kaum noch einschlägige Vorträge auf Kongressen und Tagungen dieser Gesellschaft gehalten werden, dass es an wissenschaftlichem Nachwuchs in diesem Bereich mangelt, dass sich zentrale Vertreterinnen und Vertreter wissenschaftstheoretischer Positionen in der Psychologie anderen Inhaltsbereichen zugewandt haben und dass diejenigen, die sich nach wie vor mit wissenschaftstheoretischen Themen beschäftigen, nicht selten den Anschluss an neuere Entwicklungen innerhalb der Wissenschaftstheorie verloren haben.

(Westmeyer, 2004, S. 15)

Während die reichhaltige Diskussionskultur seinerzeit den Blick für die Konsistenz wissenschaftlicher Entwürfe schärfte, begnügt man sich heute nicht selten mit zusammengestückelten Einzelideen verschiedener Provenienz. Charakteristisch für neuere Veröffentlichungen ist „oft ein unmotiviertes Nebeneinander nur partiell miteinander kompatibler Konstruktionen“ (S. 22).

Ob als Ursache oder als Wirkung einer niedergehenden Wissenschaftstheorie, nach meinem Eindruck hat sich seit den 1960er und 70er Jahren das Methodenverständnis der Fachgemeinschaft gewandelt. Während früher methodische Ansätze konzeptuell verhandelt wurden, geht es heute vor allem um statistische Rechenverfahren und ihre Verbesserung.

Ein gutes Beispiel für das alte Methodenverständnis liefert ein Artikel von Cronbach und Meehl aus dem Jahr 1955, der eine bis heute weitverbreitete methodische Praxis einführt, die *Konstruktvalidierung*. Allein schon in einer Fußnote, die dem Autor Paul Meehl im besagten Artikel beigelegt ist, spiegelt sich der Geist des damaligen Methodenverständnisses deutlich.

The second author worked on this problem in connection with his appointment to the Minnesota Center for Philosophy of Science. We are indebted to the other members of the Center (Herbert Feigl, Michael Scriven, Willfrid Sellars), and to D. L. Thistlethwaite of the University of Illinois, for their major contributions to our thinking and their suggestions for improving this paper.

(Cronbach & Meehl, 1955, S. 281)

Dem alten Verständnis zufolge geht es in der Methodenlehre um Probleme, die zu diskutieren sind. Es geht um das Ausarbeiten und Feilen an Lösungsvorschlägen, wobei Argumente dieser und jener Seite zu hören sind. Dabei versteht es sich von selbst, dass jeder für seine Position möglichst gute Argumente vorträgt, die Einwände der Gegenseite aufgreifend oder nach Möglichkeit vorwegnehmend. Nahe liegende Gesprächspartner des Methodikers sind hierbei Philosophen, insbesondere Wissenschaftstheoretiker, denn man arbeitet an gemeinsamen Problemen.

Das neue Methodenverständnis der Fachgemeinschaft ist ein ganz anderes. Als Beispiel kann etwa die folgende Äußerung eines bekannten Rationalitätsforschers dienen.

The methods of science protect an investigator from juggling the meaning of different results by deliberately making the investigator rigid and “unintelligent” in the same way that computers are rigid and unintelligent. Experimental results, like the input to a computer, must fall into certain pre-specified slots according to pre-specified rules or they are not processed at all. As scientists, we willingly sacrifice some “intelligence” and flexibility for the benefit of objectivity.

(Gilovich, 1993, S. 58)

In Stichworten ausgedrückt stehen auf der Seite der *Wissenschaftlichkeit* die Konzepte der *Methode*, der *Rigidität* und *Objektivität*. Die Gegenseite, vielleicht der Mensch im Alltag, wäre über *Flexibilität* und sogar *Intelligenz* zu charakterisieren (wenn auch letztere durch Anführungszeichen abgeschwächt).

Das neue Methodenverständnis mutet bescheiden an. Der Wissenschaftler nimmt nicht einmal „Intelligenz“ für sich in Anspruch. Doch so ehrbar die Bescheidenheit grundsätzlich ist – in dieser Form birgt sie Gefahren. Sie birgt Gefahren für die Qualität der Forschungsmethodik, Gefahren für die Qualität wissenschaftlicher Arbeiten. Denn wenn man von einem Wissenschaftler erwartet, rigide und ohne eigenen Intelligenzbeitrag vorgegebene methodische Richtlinien zu befolgen, verliert man allzuleicht einen Gesichtspunkt aus dem Blick, der wichtig ist: dass die Forschungsmethoden selbst auf kreative, intelligente und flexible Weise generiert, modifiziert und wohl im gesamten Forschungsprozess auch immer wieder mal verworfen werden müssen. Eben dabei gibt es keine rigiden Maßgaben, wie der optimale oder gar „richtige“ Weg zu finden sei.

Auf die Spitze getrieben bringt das neue Methodenverständnis eine geistige Haltung hervor, die man als „Deliberationsphobie“ bezeichnen könnte. Das Nachdenken selbst wird mit Argwohn bedacht. Wohl könne man es zum Gegenstand empirischer Untersuchungen machen. Aber das inhaltliche Ernstnehmen der Resultate des Nachdenkens, das Diskutieren der Gedanken, sei ein unwissenschaftliches oder zumindest ein zweifelhaftes Manöver. Wenn überhaupt etwas wissenschaftlich grundseriös sei, dann nur empirische Untersuchungen, die festgeschriebenen methodischen Richtlinien folgen.

Eine gewisse Irritation gegenüber dem „reinen Nachdenken“, die mir auf verwandte Weise janusköpfig erscheint, äußert der Hirnforscher Wolf Singer – hier im Gespräch mit dem Philosophen Lutz Wingert.

Wenn Neurobiologen Wahrnehmungsprozesse erforschen und erkennen, wie konstruktivistisch und zugleich wenig objektiv unsere Wahrnehmungen sind, und wenn sich ferner erweist, dass dies auch für die Prozesse gilt, die unserem Denken zugrunde liegen – dann muss das für jemanden, der davon ausgeht, dass man durch Nachdenken alleine zu verlässlicher Erkenntnis vorstoßen kann, irritierend wirken.

(Singer, 2003, S. 11)

Auch hier gilt wieder: Was auf der einen Seite bescheiden ist – das Hervorheben der menschlichen und damit auch der eigenen Fehlbarkeit – droht ausgerechnet dort einen blinden Fleck zu begünstigen, wo er besonders fatal ist, an den methodologischen Grundfesten der Wissenschaft. Empirisch forschen will man ja gerade deshalb, weil man dem Nachdenken allein nicht traut. Die empirisch-wissenschaftliche Arbeit soll eine Art von Korrektiv oder zumindest eine wertvolle Ergänzung für unsere fehlbaren Überlegungen darstellen.

Dabei zählt keineswegs jede beliebige Art von Beobachtung als „wissenschaftlich“. Zu trauen ist empirischen Arbeitsergebnissen nur, wenn die Untersuchungen „methodisch sauber“ durchgeführt wurden. Wie aber kommen wir dazu, ein bestimmtes Vorgehen als methodisch sauber einzustufen und gerade deshalb als wissenschaftliche Arbeitsgrundlage zu akzeptieren? Pointierter formuliert: Weshalb sind gängige Forschungsmethoden als wissenschaftlich seriös anzuerkennen, obwohl „das bloße Nachdenken“ mutmaßlich unzuverlässig ist? Als strategische Basis der empirischen Forschung sollten die Forschungsmethoden dann selbst eigentlich kein Resultat des bloßen Nachdenkens sein. So wären sie schließlich verdächtig, man dürfte ihnen nicht trauen. Andererseits wäre es sachlich falsch zu behaupten, gängige Forschungsmethoden würden durchweg unter (Mit-)Einsatz empirischer Strategien gewonnen.

Damit kommt die Frage auf – und möge ruhig erst mal im Raum stehenbleiben –, wie die Methodenentwicklung in der Wissenschaft vonstattengeht. Vor allem: Welche Rolle kann

und soll das Nachdenken oder Argumentieren in der Methodenentwicklung spielen? Auf welcher Grundlage werden Methoden als wissenschaftlich seriös oder unseriös beurteilt?

Wenn man Gary Heimans (2001) Auskünften in seinem Lehrbuch *Understanding research methods and statistics* folgt, dürfte jedenfalls die Logik bei der Methodenentwicklung keine maßgebliche Rolle spielen, denn sie wäre ebenfalls unzuverlässig. Seinem studentischen Leserkreis stellt er die rhetorische Frage: „Why should psychologists – including yourself – use the scientific method?“ (S. 6). Heimans eigene Antwort beginnt mit der Diskussion denkbarer Alternativen. Man könnte sich auf seine persönliche Intuition verlassen, auf den *Common Sense*, auf die Bekundungen von Autoritäten – oder auf die Logik. Es gebe aber gute Gründe, diese Quellen nicht zu wählen. Im Falle der Logik merkt Heiman an, „we do not trust logic because nature does not always conform to our logic“ (S. 6). Die vier erwogenen Quellen wiesen alle dasselbe Manko auf. „The problem with all of these sources of knowledge is that they ultimately rely on opinions or beliefs that may be created by someone who is biased or wrong or downright crazy“ (S. 6).

Wenn sogar die Grundsätze der Logik zu subjektiv sind, um als Basis der Wissenschaft in Betracht zu kommen, wenn es sich dabei sehr wohl um falsche Überzeugungen oder gar um die Konstruktionen von Verrückten handeln könnte, dann wird die Frage indes um so schwieriger zu klären, woher die Forschungsmethoden stammen, die den empirischen Disziplinen eine vermeintlich so solide Arbeitsgrundlage liefern. Gerade wer alle Hoffnungen auf die empirische Forschung setzt und meint, durch wissenschaftliche Untersuchungsverfahren könnten Irrtümer und Fehleinschätzungen des Alltags vermieden werden, muss sich diese Frage eigentlich mit Nachdruck stellen. Heiman bspw. hegt große Erwartungen gegenüber der von ihm im Singular angesprochenen „wissenschaftlichen Methode“.

The *scientific method* is a set of rules consisting of certain assumptions, attitudes, goals, and procedures for creating and answering questions about nature. [...] Psychology relies on the scientific method because it is the best approach for eliminating bias and opinion, for reaching a consensus about how the behavior truly operates, and for correcting errors. [...] The scientific method provides the most convincing evidence about nature because, instead of reflecting our own biases or intuition, scientific evidence is based on the events that occur in nature.

(Heiman, 2001, S. 6, H.i.O.)

Den Kern der sog. wissenschaftlichen Methode bilden für Heiman drei Grundannahmen. „These assumptions are that nature is lawful, deterministic, and understandable“ (Heiman, 2001, S. 6). Nicht thematisiert wird die Frage, wie denn diese drei Grundannahmen zum Herzstück der wissenschaftlichen Methode ausreifen konnten, ohne von Menschen ersonnen und akzeptiert zu werden, die vielleicht „biased or wrong or downright crazy“ (S. 6) sind.

Das neue, rigide Methodenverständnis wird offenkundig bereits in der universitären Ausbildung vorgebahnt und verfestigt. Schon bei den Lehrbüchern fällt bspw. im Vergleich mit anderen psychologischen Teildisziplinen auf, wie selten einflussreiche Theoretiker vorgestellt werden: Biographische Passagen zu einzelnen Methodikern oder Portraits sind für zeitgenössische Methodenlehrbücher mehr als untypisch. So scheint gemessen an anderen psychologischen Teildisziplinen in der Methodenlehre ein bemerkenswert geringes Theorie- und Theoretikerbewusstsein vorzuherrschen. Dass darüber hinaus auch die Fachhistorie in der Ausbildung kaum behandelt wird, trägt weiter zu dem Eindruck bei, Forschungsmethoden seien zwar wohl nicht gottgegeben, aber doch irgendwie vom Himmel gefallen.

Mit dem Rückzug aus methodologischen Kontroversen geht eine ausgeprägte Tendenz zur Statistisierung einher. In den Vordergrund der Lehre treten Formeln und Rechenverfahren, die als zeitlose Wege zur Wahrheit erscheinen. Die Verfahren haben scheinbar

keinen historischen Ursprung und besitzen auch keinerlei philosophischen Gehalt. Für ihre Prüfungen lernen Studierende verschiedenste Formeln auswendig und sind in der Regel froh, wenn sie die Zeichenketten dann fehlerfrei rezitieren oder niederschreiben können. Das kreative, bisweilen spielerische Moment des Nachdenkens, das ebenso wie eine logische Strenge aus meiner Sicht für die Methodenlehre ganz wesentlich ist, bleibt dabei außen vor. Die Forschungsmethoden wirken ätherisch und ungreifbar. Sie scheinen über jeden Zweifel erhaben, sind selbst – im Gegensatz zu den Grundsätzen der Logik – vermeintlich unstrittig und werden nicht als Resultat menschlichen Nachdenkens behandelt. Diese Arbeit wird sich weit von dem heute gängigen, statistiklastigen Methodendiskurs entfernen. Sie beginnt bei einer Feststellung, die oben am Beispiel der Libet-Studien eingeführt wurde: Verschiedene (Sub-)Disziplinen fällen auch in zentralen Fragen oft ganz unterschiedliche Urteile, wenn es um die Seriosität und wissenschaftliche Zulässigkeit von Methoden geht. So zählen innerhalb der Klassischen Psychologie bspw. nur Studien mit Bedingungsvariation heutzutage als wissenschaftlich grundseriöse Experimente. In den Neurowissenschaften werden hingegen auch Studien ohne Bedingungsvariation als wissenschaftlich seriös – oder gar geniale – Experimente akzeptiert. Hier drängt sich natürlich die Frage auf, wie solche unterschiedlichen Urteile zustande kommen. Ein wichtiges Ziel dieser Arbeit ist es, transparenter zu machen, wie über die Seriosität und wissenschaftliche Zulässigkeit von Methoden geurteilt wird. Soziologische Studien sind deshalb aber nicht anvisiert. Stattdessen sollen Mittel für die Forschungspraxis gefunden werden, um den Urteilsprozess als in sich stimmig und sachlich begründet (nach-)vollziehbar zu machen.

Wenn man davon ausgeht, dass keine (Sub-)Disziplin erratisch über die Seriosität von Methoden urteilt, dann dürfte es irgendwelche Unterschiede bei ihren Urteilsgrundlagen geben, die im Endeffekt zu verschiedenen Einschätzungen führen.

Das bedeutet zum einen: Forschungsmethoden sind tatsächlich alles andere als vom Himmel gefallene, zeitlose Wege zur Wahrheit. Sie werden akzeptiert und verworfen auf der Basis von Urteilsgrundlagen. Und es gibt keineswegs nur eine einzige Urteilsgrundlage: die richtige, die wissenschaftliche. Deshalb muss man also auch damit rechnen, dass die Urteilsgrundlagen selbst zum Gegenstand kritischer Diskussionen werden können.

Es bedeutet zum zweiten: Die Forschungsmethoden bilden zusammen kein „homogenes Set“ aus theorieneutralen Hilfsmitteln der Forschung. Überall dort, wo unterschiedliche Urteilsgrundlagen für die Zulässigkeit von Methoden gewählt werden, entstehen Brüche in der Forschungsmethodik. Was hier als wissenschaftlich seriös akzeptiert wird, wird dort verworfen. Es ist deshalb ein weiteres wichtiges Anliegen dieser Arbeit, Bruchstellen der Forschungsmethodik kenntlich zu machen.

Natürlich liegt dann auch die Frage nahe, worum es sich bei solchen „Urteilsgrundlagen“ der Methodenlehre handeln mag. Diese Arbeit wird gängige *philosophische Positionen* als Entscheidungsgrundlage für die Akzeptanz oder Ablehnung einzelner Methoden identifizieren. Dort, wo einzelne (Sub-)Disziplinen zu verschiedenen Urteilen kommen, arbeiten sie mit unterschiedlichen philosophischen Ausgangsannahmen. So wendet sich die Arbeit insgesamt gegen eine These, die in der Psychologie heutzutage weit verbreitet scheint.

Neutralitätsthese zur Forschungsmethodik:

Forschungsmethoden sind philosophisch neutral; sie bedürfen keiner philosophischen Rechtfertigung.

Das Gegenteil möchte ich nicht nur behaupten, sondern vor allem durch eine geeignete Analysemethodik zur Untersuchung von Forschungspraktiken im Detail transparent machen.

2 **Philosophiethese zur Forschungsmethodik:**

Forschungsmethoden sind philosophisch gehaltvoll; sie können nur über philosophische Zusatzannahmen gerechtfertigt werden.

So wird es der Arbeit im Weiteren auch darum gehen, eine Systematik von Forschungsmethoden zu ermöglichen, die an ihrem philosophischen Unterbau orientiert ist. Keineswegs sind rundweg alle Forschungsmethoden miteinander kompatibel. Es kommt durchaus vor, dass einzelne Methoden auf sich widersprechenden wissenschaftstheoretischen Ausgangsannahmen basieren. Dann ist die eine Methode nur als wissenschaftlich seriös zu akzeptieren, wenn man die andere verwirft. Sollen die Forschungsergebnisse der alternativen Forschungstraditionen am Ende zusammengeführt werden, bspw. in einem „Gesamtmodell“, ist besondere Umsicht geboten, um Nonsense zu vermeiden.

Um Forschungsmethoden philosophisch einordnen und systematisieren zu können, sollen Analysemittel entwickelt und erprobt werden, die helfen, den philosophischen Unterbau einzelner methodischer Praktiken zu beleuchten. Dabei wird sich vor allem die Frage als hilfreich erweisen, wie Urteile über die wissenschaftliche Zulässigkeit von Methoden sachlich zu rechtfertigen sind.

Im Kern werden die vorgeschlagenen Analysemittel meist nicht neu sein; vielmehr kommen ähnliche Verfahren vor allem in der *Analytischen Philosophie* längst schon zum Einsatz. Zur Würdigung dieser Quelle möchte ich den hier verfolgten Ansatz deshalb auch *Methodenanalyse* nennen.

Die wichtigsten Elemente der Methodenanalyse seien in hier in Kürze vorweggenommen.

Methodenpraxis: Unter dieser Überschrift werden konkrete Schritte und Maßnahmen aufgelistet, die bei der lehrbuchkonformen Anwendung einer Methode zu vollziehen sind.

Methodentheorie: In strenger Argumentform wird begründet, warum eine bestimmte Methodenpraxis wissenschaftlich zulässig oder gar geboten ist. Dies geschieht im Rekurs auf eine...

Methodologische These: Sie handelt davon, was man als Forscher (bei einer bestimmten wissenschaftlichen Ausgangslage) tun darf, kann oder soll. Bspw. können bestimmte Schlussfolgerungen zulässig oder geboten sein.

Aporetische Thesencluster: Es werden gängige Thesen verschiedener Arbeitstraditionen zusammengestellt, die miteinander inkompatibel sind. So wird besonders deutlich, dass bestimmte methodische Arbeitstraditionen sich logisch ausschließende Alternativen darstellen.

Methodenlandkarte: Hier werden Methoden gemäß ihrer philosophischen Grundlagen systematisiert. Das Format orientiert sich an gängigen Atlantendarstellungen; philosophische Bruchstellen erscheinen als Ländergrenzen.

Insgesamt möchte ich also eine neue methodische Praxis vorschlagen.

1 Methodenanalysen durchführen: Forschungsmethoden sind nicht aus sich selbst heraus sinnvoll. Dass eine bestimmte methodische Vorgehensweise „sinnvoll“ oder „wissenschaftlich seriös“ ist, kann stets nur über philosophische Zusatzannahmen begründet werden. Das Ziel der Methodenanalyse ist es, philosophische Grundannahmen hinter einzelnen methodischen Verfahrensweisen zu identifizieren. So werden ggf. auch Inkonsistenzen zwischen den Ausgangsannahmen verschiedener methodischer Praktiken deutlich.

Inhaltlich wird es speziell um Methoden der Kausalanalyse gehen. Für die Psychologie ist dieses methodische Terrain von zentraler Bedeutung. Man nimmt an, gängige Forschungsziele wie die Erklärung, Vorhersage und Kontrolle von Ereignissen im psychologischen Gegenstandsbereich gerade *durch Kausalanalysen* erreichen zu können.

Das erste Kapitel stellt dann auch Kausalität als ein zentrales Konzept der Psychologischen Methodenlehre vor. Im zweiten Kapitel werden Landmarken der philosophischen Kausalitätsdebatte eingeführt, die dem Methodendiskurs wertvolle Orientierungshilfen liefern. Den Hauptteil der Arbeit bildet im Anschluss die Auseinandersetzung mit drei einflussreichen philosophischen Denktraditionen: dem Nomologismus (in den Kapiteln 3 bis 6), dem Determinismus (in den Kapiteln 7 bis 10) sowie dem Materialismus (in den Kapiteln 11 bis 15). Alle drei Denktraditionen prägen die Psychologische Methodenlehre maßgeblich und verleihen dem kausalen Idiom der Disziplin charakteristische Facetten. Sie gehen einher mit spezifischen Annahmen darüber, wie man methodisch vorgehen sollte, um Kausalhypothesen zu prüfen oder Kausalerklärungen zu formulieren. Kontrastiert werden die drei traditionellen Positionen dabei immer wieder mit einer philosophischen Theorie jüngerer Datums, für die ich mich gelegentlich starkmachen werde: das kontrafaktisch-interventionistische Kausalverständnis. Diese jüngere Theorie legt andere methodische Praktiken nahe als die drei traditionellen „Ismen“ der Forschung. Deshalb inspiriert die kontrafaktisch-interventionistische Sichtweise auch ein paar neue methodische Vorschläge.

In der Auseinandersetzung mit verschiedenen philosophischen Basisannahmen werden schließlich auch Bruchstellen der Forschungslogik kenntlich, die verhindern, dass Ergebnisse verschiedener Arbeitstraditionen ohne weiteres in einem homogenen Wissensgebäude aufgehen. Eine solche Bruchstelle scheidet z. B. die Kausalforschung der Neurowissenschaft von jener der Klassischen Psychologie. Natürlich wird man versuchen, wo immer das möglich ist, Brücken zu schlagen und die Forschungsbefunde verschiedener Arbeitsgruppen doch irgendwie zusammenzubringen. Auch um diese Herausforderung wird es im Folgenden oft gehen. Letztlich soll dann vor allem das Modell der **Kausalniveaus** Disziplinen wie der Psychologie und der Neurowissenschaft helfen, trotz inkompatibler Ausgangsannahmen gemeinsame Wissensbestände soweit eben möglich aufzubauen.

Was diese Arbeit damit insgesamt vermitteln will, ist vor allem ein bestimmtes Bild von der Methodenlehre, das über die Statistik weit hinausreicht: Nur zusammen mit ihrem philosophischen Unterbau machen Forschungsmethoden überhaupt Sinn und können als wissenschaftlich seriös akzeptiert werden. Es wäre schön, wenn sich dieser Umstand auch in der Art, wie wir Methodenlehre betreiben, deutlicher widerspiegeln würde. Ob man also in der einen oder anderen inhaltlichen Frage der vorgetragenen Argumentation folgt oder nicht – am Ende würde ich mir vor allem einen geteilten Eindruck wünschen: dass der Einsatz philosophischer Analysemittel innerhalb der Methodenlehre naheliegend und ergiebig, ja eigentlich ganz selbstverständlich ist.

Entstanden ist diese Arbeit im Zuge einer Dissertation an der Freien Universität Berlin. Da ich mit einem möglichst allgemein verständlichen Text an die Öffentlichkeit treten wollte, danke ich allen Beteiligten für die Erlaubnis, hiermit eine überarbeitete Fassung der ursprünglichen Promotionsschrift publizieren zu dürfen.

Prof. Dr. Hans Westmeyer danke ich sehr für seine Bereitschaft, die durchaus unkonventionelle Herangehensweise dieser Arbeit an psychologische Forschungsmethoden grundsätzlich zu unterstützen. Sein scharfer Blick für Stimmigkeiten im Großen wie Kleinen war mir bis zuletzt Ansporn zu weiterer Präzision.

Der *Gesellschaft für Analytische Philosophie* bin ich sehr verbunden für eine Einladung zu ihrem Doktorandenworkshop im Jahr 2007. Insbesondere möchte ich in diesem Zusammenhang Prof. Dr. Geert Keil und Prof. Dr. Ralf Stoecker danken, die sich in großzügiger Weise für meine Argumentationsskizzen Zeit genommen haben. Prof. Dr. Geert Keil hat meine Arbeit zudem schon im Vorfeld beratend begleitet. In welchem Ausmaß diese Arbeit gerade durch Fragestellungen inspiriert wurde, mit denen sich Prof. Dr. Keil in verschiedenen Publikationen auseinandersetzt, wird auf den folgenden Seiten mit Sicherheit deutlich

werden. Daneben möchte ich Prof. Dr. Holm Tetens für seine – nicht nur an mich gerichtete – Ermunterung danken, auch „querulante“ Sichtweisen zu wagen.

Eine besondere Erwähnung verdienen Dr. Pascal Wallisch und engagierte Studenten der *University of Chicago*, die im Winterquartal 2007 an unserem Kurs *Methodological Challenges to Cognitive Neuroscience* teilnahmen. Immer wieder führte uns der Diskurs auf die Frage, wie ein Neurowissenschaftler seinen Gegenstand, den physischen Körper, mit klassisch-psychologischen Konzepten in Verbindung bringen könne. Der Gedanke zog uns in seinen Bann, dass psychologische Attribute nur auf Akteure wie den Menschen oder das Tier als Gesamtwesen sinnvoll anzuwenden seien, nicht aber auf ihre physischen Einzelteile, bspw. Gehirne oder gar Zellen. So verblieben wir bei der resignativen Mutmaßung, es gäbe wohl kein seriöses Verfahren, um die Kluft zwischen der klassisch-psychologischen Analyse einerseits und der neurowissenschaftlichen Forschungsperspektive andererseits zu überbrücken. Der Neurowissenschaftler müsse sich wohl entweder mit mathematischen Beschreibungen von physischen Vorgängen begnügen oder er müsse ein ganz neues Vokabular entwickeln, um „das Verhalten von Zellen“ zu analysieren.

Dieses Diskussionsergebnis war gewiss nicht nur für mich zutiefst unbefriedigend. Wäre es das letzte Wort in der Sache, müssten die Psychologie und die Neurowissenschaft getrennte Wege gehen. Dem steht eine ausgeprägte Intuition und Hoffnung entgegen, dass sich die Ergebnisse der Disziplinen doch gegenseitig ergänzen sollten, insofern beide den Menschen (oder das Tier) zum Untersuchungsgegenstand haben. Nicht selten nehmen Forscher die neurowissenschaftliche Arbeit ja auch gerade deshalb auf, weil sie sich in der Auseinandersetzung mit psychologischen Fragen noch ergänzende oder vertiefende Befunde wünschen. Die Intuition und Hoffnung, dass eine solche disziplinenübergreifende Befundintegration eigentlich doch möglich sein müsste, war für mich stets eine treibende Kraft in der Auseinandersetzung mit den kausalanalytischen Verfahren der Psychologie.

Jeder Entwurf profitiert von dem kritischen Gespür seiner potentiellen Rezipienten. Für Diskussionen und Anregungen danke ich Stéphane L'Hermite, Dr. Friederike Kendel, Tobias Koch, Felix Schirmann, David Schneider und meinen Eltern, Gabriele und Volker von Thienen. Birgit Jobst danke ich für Unterstützungen beim Grafikdesign und Layout.

Ermöglicht hat mir das Dissertationsvorhaben ein *NaFöG*-Stipendium, für das ich dem Land Berlin meinen Dank zolle.

I. Kausalität als zentrales Konzept der Methodenlehre

Viele zentrale Begriffe der Methodenlehre sind keine genuinen Neuschöpfungen der Wissenschaft, sondern sind der Alltagssprache entlehnt. Für „Kausalität“ gilt dies allemal. Nicht nur in der Wissenschaft, sondern schon im Alltag treffen wir eine Fülle von Kausalurteilen: Nachrichten verursachen Freude oder Bestürzung, ein Virus verursacht Computerabstürze, Staatsempfänge verursachen Kosten usw.

Wer die Alltagssprache mit ihrem kausalen Idiom gut beherrscht, ist jedoch nicht unbedingt auch qualifiziert, wissenschaftliche Kausalurteile zu formulieren und zu beurteilen. Das jedenfalls ist eine verbreitete Meinung unter Wissenschaftlern. Die Alltagssprache scheint das Maß einer noch vertretbaren Unschärfe deutlich zu sprengen – wie man hofft, im Gegensatz zur Wissenschaftssprache. Man hält es sehr wohl für möglich, dass alltäglichen Kausalurteilen überhaupt kein einheitliches, in sich konsistentes Kausalkonzept zugrunde liegt. Vielleicht meinen wir als Sprecher des Alltags mit einem Kausalausdruck in den verschiedenen Kontexten mal dieses und mal jenes.

Eine Wissenschaft, freilich, sollte zumindest bei zentralen Konzepten sprachliche Unschärfen und Inkonsistenzen weitestgehend überwinden. Wie noch zu zeigen sein wird, ist *Kausalität* ein überaus zentraler Begriff für die Psychologie und ihre Methodenlehre. Wer Kausalforschung betreiben möchte, benötigt klare Kriterien dafür, wann ein Kausalzusammenhang vorliegt und wann nicht. Nur so kann man systematisch prüfen, ob die Kriterien in einem fraglichen Falle erfüllt sind, ob ein Kausalurteil gerechtfertigt ist.

Doch der Blick zur Forschungspraxis nötigt zum Innehalten. Konnte sich die Psychologie auch als eigenständige, wissenschaftliche Disziplin an den Universitäten etablieren, so wurde ihr doch offenkundig kein konsistentes Kausalkonzept mit in die fachliche Wiege gelegt. Wie Volker Gadenne (1994) zu Recht bemängelt, liegt psychologischen Forschungsmethoden „keine einheitliche und zum Teil überhaupt keine geklärte Auffassung von Kausalität zugrunde“ (S. 330).

Umgekehrt kann und soll „Wissenschaft betreiben“ natürlich nicht heißen: sich ausruhen und in mutmaßlichen Meriten sonnen. Das ist im Grunde klar. Deshalb bietet der Status quo auch nur bedingt Anlass zum Stirnrunzeln; er lädt vor allem ein, die Ärmel hochzukrempeln.

1.1 Das Billardspiel als paradigmatischer Anwendungsfall

Um sich die Bedeutung eines Begriffes klarzumachen, ist als erster Schritt der Blick auf paradigmatische Anwendungsfälle oft hilfreich. Ein klassisches Beispiel für Kausalzusammenhänge stammt aus dem Billardspiel. Die weiße Kugel stößt eine rote an, die rote Kugel rollt weg. Dass die rote Kugel genau in dem Moment beschleunigt wird, in dem die weiße zur Ruhe kommt, ist keine bloße Koinzidenz. Vielmehr kommt die rote Kugel *durch* den Anstoß der weißen in Bewegung. *Das* ist Kausalität.

Weshalb wir sowohl im Alltag wie auch in der Wissenschaft ein hohes Interesse für Kausalbeziehungen aufbringen, zeigt sich schon an diesem Beispiel. Kausalität ist für unser Argumentieren und Handeln von fundamentaler Bedeutung:

- Wir studieren Kausalzusammenhänge, um Geschehnisse **erklären** zu können.
(Warum begann die rote Kugel plötzlich zu rollen? Das war so, *weil* sie von der weißen angestoßen wurde.)
- Wir studieren Kausalzusammenhänge, um Geschehnisse **vorhersagen** zu können.
(Wohin wird sich die rote Kugel bewegen? Sie wird sich dort-und-dort-hin bewegen, *weil* die weiße Kugel sie in dem-und-dem Winkel mit der-und-der Geschwindigkeit angestoßen hat.)

- Wir studieren Kausalzusammenhänge, um Geschehnisse **manipulieren** zu können.
(Was muss ich tun, damit die rote Kugel ins Loch rollt? Du musst die weiße Kugel dazu bringen, die rote in dem-und-dem Winkel mit der-und-der Geschwindigkeit anzustoßen.)

Damit sind Stichworte gefallen, die jedem Psychologen wohlvertraut sind, denn es kommen grundlegende Forschungsziele der Disziplin zur Sprache.

Ziel der Psychologie ist die Erklärung, Vorhersage und Kontrolle der Ereignisse ihres Gegenstandsbereichs (BROWN & GHISELLI 1955, KERLINGER 1964, SELG 1971). Auf diese allgemeine Definition wird man sich am ehesten einigen können, zumal sie mit der Zielsetzung anderer empirischer Wissenschaften übereinstimmt.
(Westmeyer, 1973, S. 14, H.i.O.)

So nimmt es nicht Wunder, dass die psychologische Forschung sehr wesentlich eine Kausalforschung ist. „Empirische Kausalforschung ist in der Psychologie und ihren Nachbardisziplinen allgegenwärtig“ (Steyer, 2004, S. 170). Eindrücklich dokumentieren den zentralen Stellenwert der Kausalanalysen schon die in unübersehbarer Vielzahl durchgeführten Experimente, denn „Experimente werden durchgeführt, um wissenschaftliche Kausalhypothesen empirisch zu prüfen“ (Westermann, 2000, S. 270).

Nun mag das Billardspiel als paradigmatischer Anwendungsfall zwar verständlich machen, weshalb sich eine Wissenschaft für Kausalbeziehungen interessiert, wenn ihr an Erklärungen, Prognosen und Manipulationen gelegen ist. Offen bleibt aber noch, wie Psychologen in der Lage sein können, Kausalbeziehungen zu identifizieren, und wie genau das Experiment ihnen dabei weiterhelfen soll. Immerhin sind Billardkugeln nicht unbedingt der zentrale Gegenstand psychologischer Forschung. Dem gängigen Verständnis zufolge hat es die Psychologie vielmehr mit dem *Erleben und Verhalten von Personen* zu tun (vgl. Feger & Graumann, 1983; Métraux, 1981). Ob menschliche Erlebnisse und Verhaltensweisen auf eben dieselbe Weise zu analysieren und zu beschreiben sind wie das Stoßen und Rollen von Billardkugeln ist eine diskussionswürdige Frage. Man denke etwa an Humes einflussreiche Erörterung der Kausalthematik und seine Äußerung zum Billardspiel.

Here is a billiard-ball lying on the table, and another ball moving towards it with rapidity. They strike; and the ball, which was formerly at rest, now acquires a motion. This is as perfect an instance of the relation of cause and effect as any which we know, either by sensation or by reflection. Let us therefore examine it. [...] The first ball is in motion; touches the second; immediately the second is in motion: and when I try the experiment with the same or like balls, in the same or like circumstances, I find that upon the motion and touch of the one ball, motion always follows in the other.
(Hume, 1740/1978, S. 649f.)

Wenn wir nun den „Gegenstandsbereich Billardspiel“ mit dem Gegenstandsbereich der Psychologie vergleichen, fallen sogleich einige Unterschiede auf. Das gilt für die Eigenarten der untersuchten Phänomene ebenso wie für die Möglichkeiten ihrer Untersuchung. Im Billardspiel gibt es allem Augenschein nach verlässliche Regularitäten. Für einen Billardprofi dürfte sich die Unterstellung bewähren, dass die Kugeln ihre vorgeplanten Wege nehmen, sofern nur der Stoß gelingt. Man muss nicht damit rechnen, dass eine der Kugeln „einfach mal einen anderen Weg ausprobiert“, etwa um zu prüfen, wie das Publikum darauf reagiert. Verhielten sich Billardkugeln wie Menschen, sie würden auf Anstöße durch andere Kugeln gemäß ihren eigenen Belangen reagieren. Manchmal würden sie nicht verstehen, warum eine andere Kugel sie gerade anstößt, dann würden sie unschlüssig liegenbleiben, anstatt geradlinig in eine Richtung zu rollen – so wie Menschen nicht wissen, was sie tun sollen, wenn ihnen eine Situation oder Aufforderung unverständlich ist. Zudem können wir eine Billardkugel von einem Billardtisch auf den nächsten legen, um vorangegangene Versuche zu replizieren. Wie sollte dies mit Erlebnissen geschehen?

Man stelle sich etwa vor, einer Person A werde ihr Schmerzerlebnis aus dem Zahn „herausgegriffen“. Der Arzt legt den Schmerz auf einen Tisch, um seine charakteristischen Eigenschaften auszumessen. Anschließend schiebt er den Schmerz einem zweiten Patienten ins Gebiss, um zu prüfen, ob er dort dieselben Wirkungen entfaltet wie im ersten. Schließlich richten sich Menschen in ihrem Verhalten nach Geboten und Verboten, wie sie z. B. durch juristische Gesetze vorgegeben werden. Regularitäten im menschlichen Verhalten können erzeugt, modifiziert oder abgeschafft werden – indem man bspw. Gesetze oder Vertragsbedingungen ändert. Billardkugeln hingegen sind nicht auf dieses oder jenes Verhalten zu verpflichten. Wir können ihnen nicht abfordern, ein bestimmtes Verhalten – und stets nur dieses – zu zeigen, derweil sie sehr wohl in der Lage wären, etwas ganz anderes zu tun.

Gravierende Unterschiede zwischen psychologischen Phänomenen und Phänomenen im Bereich des Billardspiels scheinen also offenkundig. Was bedeutet das für die Suche nach Kausalzusammenhängen? Sind Kausalbeziehungen zwischen Erlebnissen und Verhaltensweisen dennoch analog zu Kausalverhältnissen im Billardspiel zu verstehen? Können sie vielleicht doch auf ähnliche Weise untersucht werden? Alles in allem: Mit welchem Kausalkonzept arbeitet die Psychologie – und wie angemessen ist es?

So lässt der erste Schritt in Richtung Kausaldebatte, der kurze Blick auf einen paradigmatischen Anwendungsfall, zunächst einige Fragen aufkommen. Der zweite Schritt soll näher an die Materie heranführen: mitten in die Psychologische Methodenlehre.

1.2 Methoden der Kausalanalyse

Wer Fachbücher zur Psychologischen Methodenlehre konsultiert, dem fällt vielleicht auf, dass man von „der Kausalität“ im Singular zu sprechen pflegt. Offenbar wird angenommen, dass der Terminus immer dasselbe bezeichnet – und zwar etwas, das in einem engen Zusammenhang mit Experimenten steht. Eine prototypische Auskunft zur Kausaltheematik formulieren bspw. Moore und McCabe (2001) in ihrem Lehrbuch *Introduction to the practice of statistics*.

How can a direct causal link between *x* and *y* be established? The best method – indeed, the only fully compelling method – of establishing causation is to conduct a carefully designed experiment in which the effects of possible lurking variables are controlled.
(Moore & McCabe, 2001, S. 210)

Ähnlich äußern sich auch Shadish, Cook und Campbell (2002) in *Experimental and quasi-experimental designs for generalized causal inference*. „Experiments are well-suited to studying causal relationships. No other scientific method regularly matches the characteristics of causal relationships so well“ (S. 7). Wenn in philosophischen Kausalitätstheorien das Experiment zunehmend an Zentralität gewinnt, ist das für Shadish et al. demgemäß ein wichtiger Fortschritt.

John Locke said [...]: „A *cause* is that which makes any other thing, either simple *idea*, substance, or mode, begin to be; and an *effect* is that, which had its beginning from some other thing“ (p. 325). Since then, other philosophers and scientists have given us useful definitions of the three key ideas – cause, effect, and causal relationship – that are more specific and that better illuminate how experiments work.
(Shadish et al., 2002, S. 4, H.i.O.)

Experimente gelten als beste Methode der Kausalanalyse und als wichtigstes Hilfsmittel, um den Zielen der Psychologie – Erklärung, Prognose und Manipulation – näherzukom-

men. „Wo das Experiment in keiner Weise weiter zu helfen vermag, ist die empirische Psychologie meist bald zu Ende“ (Selg, 1966, S. 25).

1.2.1 Das Vorgehen im Experiment

Zentral wie das Experiment in der Kausalforschung ist, lohnt es sich, den Grundgedanken der Methode noch einmal Revue passieren zu lassen. Dabei kann man ohne weiteres an einem Kausalverständnis anknüpfen, das schon aus dem Alltag bekannt ist.

Wer einen Virus als Ursache von Computerabstürzen bezeichnet, behauptet, dass sich die Computerabstürze nicht in dieser Weise ereignet hätten, wenn die Rechner von den Viren nicht infiziert worden wären. Hier wird offenkundig ein Abgleich zwischen zwei Verläufen vollzogen: zwischen dem Geschehen mit Virusinfektion und dem mutmaßlichen Geschehen ohne Virusinfektion. Man behauptet, dass die Virusinfektion einen Unterschied macht. Experimente sollen einen methodisch besonders sorgfältigen Abgleich dieser Art ermöglichen. Worauf der Forscher dabei zu achten hat, ist indes ein Gegenstand fortwährender Diskussion. Die methodische Verfahrensweise hat sich in den letzten 100 Jahren in der Psychologie deutlich gewandelt. Der Ausdruck „Experiment“ bezeichnet heute ein anderes Vorgehen als früher. Noch um 1920 war von einem „Experiment“ schon die Rede, wenn ein Geschehen nurmehr unter kontrollierten Bedingungen absichtlich in Gang gesetzt wurde; das Kapitel zum *Nomologismus* wird hierauf näher eingehen. Heute zählt in der sozialwissenschaftlichen Psychologie eine Untersuchung streng genommen nur dann als Experiment, wenn mindestens zwei Untersuchungsbedingungen realisiert werden und die Teilnehmer des Versuches *per Zufall* der einen oder der anderen Versuchsbedingung zugewiesen werden.

Untersuchungen beginnen mit einer *Hypothese*, derzufolge x möglicherweise y verursacht. Bspw. könnte untersucht werden, ob Koffeingaben zu einer Verlängerung der Konzentrationsspanne führen. Würden alle Versuchsteilnehmer dieselbe Menge an Koffein konsumieren, könnte der Experimentator später nicht beurteilen, wie lange die Konzentrationsspanne der Teilnehmer gewesen wäre, hätten sie kein Koffein zu sich genommen. Er könnte dann auch nicht beurteilen, ob sich die Konzentrationsspanne durch die Koffeingabe verlängert hat. Daher sollten mindestens zwei Versuchsbedingungen realisiert werden. Bspw. erhalten Teilnehmer der einen Bedingung eine Koffeingabe, Teilnehmer der anderen Bedingung nicht. Alternativ könnten die Versuchsgruppen auch verschieden hohe Dosen an Koffein erhalten.

Es ist üblich, den vom Versuchsleiter manipulierten Faktor, hier die Koffeingabe, als *unabhängige Variable* und die anschließend erhobene Messgröße, hier ein Maß für die Konzentrationsspanne, als *abhängige Variable* zu bezeichnen.² Wenn Koffein tatsächlich die Konzentrationsspanne verlängert, sollten Teilnehmer in der *Experimentalbedingung* (mit Koffeingabe) im Schnitt eine längere Konzentrationsspanne aufweisen als Teilnehmer in der *Kontrollbedingung* (ohne Koffeingabe).

Wird ein solcher Mittelwertsunterschied festgestellt, ist die Frage zu stellen, ob die untersuchte Kausalhypothese die einzige plausible Erklärung für den Gruppenunterschied liefert. Es könnte bspw. sein, dass die Teilnehmer selbst fest davon überzeugt sind, mit Koffein bessere Leistungen zu erzielen als ohne; ein *Placebo-Effekt* wäre denkbar. Allein die Überzeugung, Koffein erhalten oder nicht erhalten zu haben, könnte eine Steigerung bzw.

² Diese überaus verbreitete Redeweise ist genau genommen abkürzend und dabei reichlich schludrig: Natürlich sind die Koffeingabe und die Konzentrationsleistung empirische Gegebenheiten. Bei den Variablen handelt es sich hingegen um Symbole, die wir zwecks Dokumentation der empirischen Gegebenheiten einführen. Methodiker sprechen aber oft so, als ob die Variablen mit den kodierten empirischen Gegebenheiten identisch seien.

Minderung der Konzentrationsspanne bewirken. Die Überzeugung der Probanden stellt eine potentielle *Störvariable* dar, d. i. eine Variable, die in dem Experiment einen Einfluss auf die abhängige Messgröße hat, aber mit der unabhängigen Variablen nicht identisch ist. Die Annahme einer Kausalbeziehung zwischen Überzeugung und Konzentrationsleistung liefert eine *Alternativhypothese* zur Erklärung potentieller Gruppenunterschiede.

Der Experimentator kann in dieser Sache vorsorgen, indem er beide Versuchsbedingungen so ähnlich wie irgend möglich gestaltet. Die Teilnehmer der Experimental- und der Kontrollbedingung sollen jeweils zwei Tassen Kaffee trinken – und nur der Experimentator weiß, ob es sich dabei um koffeinhaltigen oder entkoffeinierten Kaffee handelt. Wenn sich die beiden Untersuchungsgruppen tatsächlich nur in dem einen interessierenden Faktor des Kaffeekonsums unterscheiden, dann sollte es keine plausible Alternativklärung für potentielle Gruppenunterschiede in der Konzentrationsleistung geben.

Alle Versuchspersonen Kaffee trinken zu lassen ist jedoch noch nicht hinreichend, um schon alle Unterschiede zwischen den Probandengruppen zu eliminieren, die über einen verschieden hohen Kaffeekonsum hinausgehen. Es ist ja ohne weiteres möglich, dass sich die Teilnehmer der beiden Versuchsgruppen systematisch voneinander unterscheiden: Vielleicht würden die Teilnehmer der Experimentalbedingung auch ohne Koffein bessere Leistungen erzielen als die Teilnehmer der Vergleichsbedingung, weil sie von vornherein besonders konzentrationsstark und die Teilnehmer der anderen Bedingung besonders konzentrations schwach sind. Gruppenunterschiede, die schon vor der Behandlung bestehen, liefern eine weitere Klasse von Alternativhypothesen zur Erklärung vermeintlicher Experimentaleffekte. Methodisch wird dieser Problematik Rechnung getragen, indem die Versuchspersonen *per Zufall* der einen oder anderen Versuchsbedingung zugewiesen werden (*Randomisierung*). Wird jeder Versuchsteilnehmer mit gleicher Wahrscheinlichkeit der einen oder anderen Bedingung zugeordnet, ist der *Erwartungswert* aller Personcharakteristika in den verschiedenen Versuchsbedingungen derselbe. Je größer die untersuchten Stichproben sind, desto eher kann angenommen werden, dass sich Unterschiede zwischen den Teilnehmern der verschiedenen Versuchsbedingungen „ausmitteln“. Es darf dann unterstellt werden, dass der Durchschnitt und die Varianz der Personeneigenschaften in den verschiedenen Versuchsbedingungen in etwa gleich ausfallen; das betrifft natürlich auch die Konzentrationsfähigkeit. Vor Untersuchungsbeginn bestehende Differenzen von Personen liefern so keine plausible Alternativhypothese zur Erklärung potentieller Gruppenunterschiede mehr.

Mit diesen methodischen Vorkehrungen wird ein Experiment realisiert, das den aktuellen Standards der Psychologischen Methodenlehre entspricht. Zeigen die Teilnehmer der Experimentalbedingung im Vergleich mit den Teilnehmern der Kontrollbedingung unter diesen Umständen in statistisch signifikantem Maße erhöhte Konzentrationsleistungen, wird eine Steigerung der Konzentrationsleistung als *Effekt* ausgemacht und die Kaffeingabe als ihre *Ursache* identifiziert.

Historisch gesehen ist übrigens die Randomisierung als zentrales Charakteristikum des Experiments vergleichsweise jung. Die Strategie der Zufallszuweisung wurde erst 1925 von dem Biologen und Statistiker Ronald A. Fisher (1890-1962) zur Verbesserung der Experimentalmethodik vorgeschlagen – und zwar zunächst in der Agrarforschung. Heute gilt die Randomisierung in den Sozialwissenschaften als ein wesentliches, wenn nicht das zentrale Bestimmungsstück des formvollendeten Experiments. Auch der vielleicht bekannteste Experimentalmethodiker der Psychologie Donald T. Campbell (1916-1996) findet deutliche Worte, um Fishers methodische Erfindung zu würdigen. „Experimental design in the Fisher tradition is the optimal setting for the study of interpretable change. The magic of randomization is that it attenuates the causal threads of the past as they might codetermine both exposure to the treatment and gain scores “ (Campbell, 1963, S. 213).

1.2.2 Nichtexperimentelle Kausalforschung

Wenn auch das Experiment unter methodischen Gesichtspunkten als beste Wahl gilt, ist es doch nicht immer möglich, das Verfahren einzusetzen. In einigen Fällen verbietet sich das experimentelle Vorgehen aus ethischen Gründen. So wäre es unangemessen, Personen absichtlich mit Viren zu infizieren, um die nachfolgende Beeinträchtigung ihrer Gesundheit zu erforschen. In anderen Fällen fehlen schlichtweg die Ressourcen; Experimente können natürlich aufwendig und teuer sein. Zudem interessieren bisweilen auch Kausalwirkungen von Faktoren, die nicht manipulierbar sind, wie Alter oder Geschlecht.

Kausalanalysen auf der Basis von Datensätzen, die nicht experimentell gewonnen wurden, gelten als problematisch. Welche Schwierigkeiten dabei auftreten können, mag ein Beispiel verdeutlichen, das weitestgehend analog zum obigen Fall aufgebaut ist. Die Fragestellung und die Datenauswertung sind gleich; nur werden die Probanden den Versuchsbedingungen nicht per Zufall zugewiesen (es fehlt die Randomisierung) und der Forscher verzichtet darauf, die Versuchsbedingungen selbst zu gestalten.

Man stelle sich vor, ein junger Psychologiestudent sucht seine Universität in den Semesterferien auf und klopft an einige Seminartüren. Er findet zwei Räume belegt und die Kursteilnehmer erklären sich bereit, nach dem Ende ihrer Veranstaltung einen Fragebogen auszufüllen. Dabei sollen sie zum einen angeben, wie viele Kaffee (oder sonstige koffeinhaltige Nahrungsmittel) sie seit dem Aufwachen bereits zu sich genommen haben; zum anderen sollen sie einen Konzentrationsleistungstest durchführen.

In dem ersten Raum findet ein Mathematikkurs für hochbegabte Schüler statt, den die Teilnehmer freiwillig in ihren Schulferien besuchen. Kaum ein Schüler hat an diesem Tag schon Kaffee oder andere koffeinhaltige Nahrungsmittel zu sich genommen; alle zeigen außerordentlich hohe Konzentrationsleistungen.

Bei dem zweiten Kurs, der in den Semesterferien auf dem Campusgelände abgehalten wird, handelt es sich um ein Training für Personen mit Konzentrations- und Lernschwierigkeiten. Der Veranstaltungsleiter hat drei Kannen koffeinhaltigen Kaffee in die Veranstaltung mitgebracht; fast alle Teilnehmer haben sich bedient. Im Schnitt zeigen sie nur mäßige Konzentrationsleistungen.

Nehmen wir weiterhin an, dass eine Koffeingabe der Sache nach sowohl bei den hochbegabten Mathematikschülern als auch bei den Personen mit Lernschwierigkeiten die Konzentrationsspanne *verlängert*. Doch auch durch eine Koffeineinnahme erreichen Personen mit Lernschwierigkeiten nicht die außerordentlichen Konzentrationsleistungen der Mathematikschüler, ob diese nun Kaffee getrunken haben oder nicht.

Wenn der junge Psychologe nach der Erhebung in seinem Datensatz die Mittelwerte vergleicht, sieht es so aus, als ob Kaffee die Konzentrationsspanne *verringert*: Die durchschnittliche Konzentrationsleistung der Kaffeetrinker ist geringer als die der Nicht-Trinker.

Hier liegt ein sog. *Simpson-Paradox* vor. „In jeder der beiden Subpopulationen ein positiver Effekt, und in der Gesamtpopulation, die sich genau aus diesen beiden Subpopulationen zusammensetzt, ein negativer Effekt!“ (Steyer, 2004, S. 132). Sowohl die Mathematikschüler wie auch die Personen mit Lernschwierigkeiten profitieren im Grunde von einer Koffeineinnahme. In der Gesamtpopulation, die sich gerade aus eben diesen Mathematikschülern und Personen mit Lernschwierigkeiten zusammensetzt, zeigen aber im Schnitt diejenigen eine längere Konzentrationsspanne, die keinen Kaffee getrunken haben.

Der vermeintlich negative Effekt des Kaffeekonsums resultiert natürlich aus Gruppenunterschieden, die schon vor der Koffeineinnahme bestanden. Die beiden Probandenpopulationen des Psychologiestudenten sind nicht vergleichbar; unter den Kaffeetrinkern befinden sich fast nur Personen mit Lernschwierigkeiten, unter den Nicht-Kaffeetrinkern fast ausschließlich hochbegabte Mathematikschüler mit herausragender Konzentrationsfähigkeit.

Dass die Verhältnisse in der Gesamtpopulation dieser „Studie“ in Wahrheit keinen negativen Effekt einer Koffeineinnahme ausweisen, wird dann deutlich, wenn man die Daten der beiden Kurse getrennt analysiert. Zwar sind unter den Mathematikschülern nur wenige Kaffeetrinker und unter den Personen mit Lernschwäche nur wenige, die auf Kaffee verzichtet haben – aber ein paar finden sich doch, so dass der Abgleich möglich ist: In beiden Kursen zeigen die Kaffeetrinker eine längere Konzentrationsspanne als ihre Kurskollegen, die kein Koffein zu sich genommen haben.

1.3 Bloße Korrelation versus Kausalbeziehung

Korrelieren die unabhängige und die abhängige Variable einer *experimentellen* Studie miteinander, folgt in den Sozialwissenschaften für gewöhnlich das Kausalurteil auf dem Fuße. Hier kann eine weitere methodische Praxis der Fachgemeinschaft festgehalten werden, die später auch in einer Methodenlandkarte verzeichnet werden soll.

- ③ **Experimentbasierte Kausalurteile:** Korrelationen, die unter Experimentalbedingungen erhalten bleiben, werden kausal interpretiert.

Im Falle *nicht-experimenteller* Daten ist man mit kausalen Interpretationen indessen viel vorsichtiger. Der Psychologiestudent findet in seiner Gesamtpopulation eine negative Korrelation zwischen „Koffeineinnahme“ und „Konzentrationsleistung“. Verringt der Konsum koffeinhaltigen Kaffees die Konzentrationsleistung? Methodisch geschulte Sozialwissenschaftler betonen in solchen Fällen, dass keineswegs jede Korrelation kausal interpretiert werden dürfe. Zwar werden Korrelationen in der Psychologischen Methodenlehre meist als Voraussetzung für Kausalurteile gehandelt. Das Motto lautet dann: Keine Kausalität ohne Korrelation. Allerdings reichen dem gängigen Verständnis nach Korrelationen allein zur Stützung einer Kausalaussage nicht hin. (Wie schon in der Einleitung sollen auch hier und im Folgenden nummerierte Fahnen ► wichtige Thesen markieren.)

③ **Erste Interpretationsthese zu Korrelationen:**

Wenn die Variablen x und y (außerhalb eines Experiments) korrelieren, kann zwischen den untersuchten Faktoren eine Kausalbeziehung bestehen, muss aber nicht.

Der Fall des Psychologiestudenten ist so konstruiert, dass eine kausale Interpretation der gefundenen Korrelation gerade unangemessen wäre.

A well-known maxim in research is: *Correlation does not prove causation*. This is so because we may not know which variable came first nor whether alternative explanations for the presumed effect exist. For example, suppose income and education are correlated. Do you have to have a high income before you can afford to pay for education, or do you first have to get a good education before you can get a better paying job? [...] Correlations also do little to rule out alternative explanations for a relationship between two variables such as education and income. That relationship may not be causal at all but rather due to a third variable (often called a *confound*), such as intelligence or family socioeconomic status, that causes both high education and high income. (Shadish et al., 2002, S. 7, H.i.O.)

Einig ist man sich in der Psychologischen Methodenlehre darüber, dass nicht jede Korrelation zwischen zwei Variablen x und y unmittelbar kausal interpretiert werden dürfe. Einig scheinen sich viele Methodiker auch darin zu sein, dass *jede* Korrelation auf *irgend einem* Kausalzusammenhang basiert – wenn auch nicht unbedingt auf einer Kausalbeziehung zwischen x und y . Einig sind sich verschiedene Autoren schließlich sogar darüber, dass

beim Vorliegen einer Korrelation *vier* Interpretationen zu erwägen seien. Uneinig zeigt man sich nur in der Frage, um welche vier Interpretationsoptionen es sich dabei handelt. So benennt Jürgen Bortz (1999) in seinem Lehrbuch *Statistik für Sozialwissenschaftler* die folgenden Möglichkeiten.

1. x beeinflusst y kausal,
2. y beeinflusst x kausal,
3. x und y werden von einer dritten oder weiteren Variablen kausal beeinflusst,
4. x und y beeinflussen sich wechselseitig kausal.

Der Korrelationskoeffizient liefert keine Informationen darüber, welche der 4 Interpretationen richtig ist.

(Bortz, 1999, S. 226, H.i.O.)

Backhaus, Erichson, Plinke und Weiber (2000) besprechen in ihrem Lehrbuch *Multivariate Analysemethoden* im Kapitel zur LISREL-Analyse ebenfalls vier Interpretationsmöglichkeiten von Korrelationskoeffizienten – allerdings in Teilen andere.

Der Korrelationskoeffizient läßt [...] *keine* Aussage darüber zu, welche Variable als *verursachend* für eine andere Variable anzusehen ist.

Es sind vielmehr vier grundsätzliche Interpretationsmöglichkeiten einer Korrelation denkbar:

A: Die Variable x_1 ist verursachend für den Wert der Variablen x_2 :

$x_1 \rightarrow x_2$
[...]

B: Die Variable x_2 ist verursachend für den Wert der Variable x_1 :

$x_2 \rightarrow x_1$
[...]

C: Die Abhängigkeit der Variablen x_1 und x_2 ist teilweise bedingt durch den Einfluß einer exogenen (hypothetischen) Größe ξ (lies: Ksi), die hinter diesen Variablen steht:



[...]

D: Der Zusammenhang zwischen den Variablen x_1 und x_2 resultiert allein aus einer exogenen (hypothetischen) Größe ξ , die hinter den Variablen steht:



(Backhaus et al., 2000, S. 399f., H.i.O.)

Etwas überraschend ist der Entwurf von Backhaus et al. (2000) schon deshalb, weil erst streng zwischen „ $x_1 \rightarrow x_2$ “ und „ $x_2 \rightarrow x_1$ “ unterschieden, dann aber bei gleichzeitiger Einwirkung von ξ nur ein einziger Fall abgehandelt wird, nämlich

 und nicht .

Bortz erwägt im Gegensatz zu Backhaus et al. die Möglichkeit, dass sich die zwei untersuchten Variablen x und y (bzw. x_1 und x_2) gegenseitig beeinflussen. Backhaus et al. betonen stattdessen, die Korrelation zwischen den beiden Variablen könne zu einem Teil auf einen internen Kausalzusammenhang zurückgehen, gleichzeitig aber auch zum Teil von Kausalbeziehungen zu einer dritten Variablen bedingt sein. Bei Backhaus et al. ist von der dritten Variablen im Singular die Rede; Bortz erwähnt demgegenüber explizit, es könnten auch mehrere Variablen im Hintergrund wirksam sein.

Auflistungen von Interpretationsoptionen dieser Art sind wohl nicht als wörtlich zu nehmende, ausgefeilte Theoriefragmente zu lesen. Allein der Umstand, dass unterschiedliche Interpretationsoptionen für Korrelationen in der Literatur kursieren, ohne in ihrer Verschiedenartigkeit für Aufsehen zu sorgen, dürfte ein geringes Interesse der Fachgemeinschaft an derartigen Details dokumentieren. Auch die Behauptung, jede Korrelation basiere auf einem Kausalzusammenhang, ist im Grunde etwas wunderlich. Immerhin manifestieren sich sogar begriffliche Zusammenhänge als Korrelationen. Die Eigenschaften „Junggeselle sein“ und „männlich sein“ korrelieren bspw. allein aufgrund ihres definitionsgemäßen Zusammenhangs miteinander. Kausalbeziehungen sollen gleichsam empirische Relationen sein. Deshalb führen wir Untersuchungen durch, um Kausalaussagen zu überprüfen, anstatt in Lexika die Bedeutung von Worten nachzuschlagen.

Stellungnahmen wie die von Bortz (1999) oder Backhaus et al. (2000) zeigen, dass nichtsdestotrotz eine Grundannahme der folgenden Art vertreten wird.

4 Zweite Interpretationsthese zu Korrelationen:

Wenn die Variablen x und y korrelieren, kodiert jede von ihnen entweder die Ursache oder die Wirkung einer Kausalbeziehung, oder beides.

Diese Annahme hat eine interessante Konsequenz: Die Welt ist voller Kausalzusammenhänge. Da es an Korrelationen nicht mangelt, ist die Welt offenkundig von Kausalzusammenhängen durchdrungen. Auch ist die Zahl verschiedenartiger Kausalbeziehungen schier unüberschaubar. Jede Variable, die mit einer anderen Variablen korreliert, soll entweder als Ausgangspunkt oder Endpunkt eines Kausalzusammenhangs verstanden werden, oder beides. Wird in einem Modell, das bislang fünf Variablen umfasst hat, eine weitere Variable aufgenommen, können sogleich fünf neue Korrelationen überprüft werden. Dabei kann jede der alten Variablen sowohl Ausgangs- als auch Endpunkt einer Kausalbeziehung mit der neuen Variablen sein. Mit jeder neu untersuchten Variablen wächst daher die Zahl der ggf. anzunehmenden Kausalzusammenhänge in beachtlichem Maße.

Dem entspricht in der sozialwissenschaftlichen Fachliteratur eine schier unübersehbare Fülle der verschiedenartigsten Kausalurteile. Welchem Phänomenbereich sich Psychologen auch immer zuwenden, innerhalb von kurzer Zeit machen sie dort in der Regel vielfältige Kausalbezüge aus. Dieser Umstand wird sich für das gängige Verständnis von Kausalität innerhalb der psychologischen Fachgemeinschaft durchaus noch als wichtig erweisen. Deshalb soll die bestehende Praxis schon hier eigens herausgestellt werden.

4 Zahllose Kausalurteile: In der Fachliteratur werden Kausalzusammenhänge *en masse* berichtet.

Auch pflegen Psychologen bei ihren theoretischen Modellen in auffallend großzügiger Zahl jene Pfeile einzuzeichnen, die Kausalbeziehungen anzeigen sollen.

5 Pfeilintensive Kausalmodelle: Es werden häufig Kausalmodelle mit sehr vielen Pfeilen publiziert, die jeweils (empirisch bestätigte) Kausalzusammenhänge anzeigen sollen.

Beliebt sind demgemäß auch Grundmodelle, die keinen Zweifel daran lassen, dass absolut alles mit absolut allem zusammenhängt. Ein gutes Beispiel hierfür liefert der *psychobiologische Systemansatz der Entwicklung* von Gottlieb (2001).

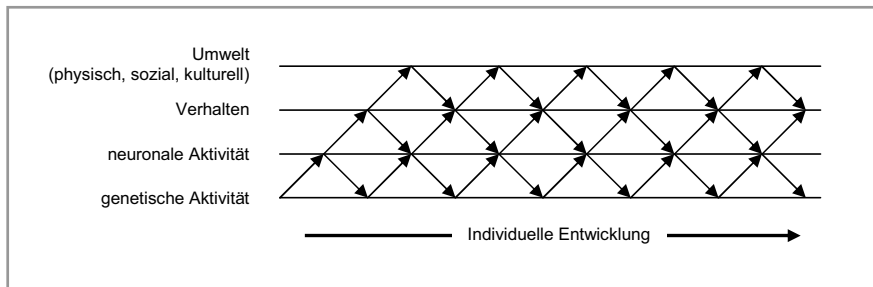


Abb. 1: Modell des psychobiologischen Systemansatzes der Entwicklung nach Gottlieb (2001, S. 186; dt. Niebank & Petermann, 2002, S. 68).

Die Annahme eines überaus komplexen Netzes von Kausalzusammenhängen speist sogar methodische Bedenken gegenüber dem Experiment. In diesem werden schließlich nur einige wenige Kausalbeziehungen isoliert überprüft.

Many researchers believe that the causal knowledge that results from this typical experimental structure fails to map the many causal forces that simultaneously affect any given outcome in complex and nonlinear ways (e.g., Cronbach et al., 1980; Magnusson, 2000). These critics assert that experiments prioritize on arrows connecting A to B when they should instead seek to describe an explanatory pretzel or set of intersecting pretzels, as it were. (Shadish et al., 2002, S. 457)

Die heute in der Psychologie gängige Auffassung von Korrelation und Kausalität vermittelt ein Weltbild, das etwas ironisch vielleicht so zu charakterisieren ist: Soweit die Welt des Psychologen Ordnung aufweist, besteht sie aus einem Wirrwarr an Kausalbeziehungen. Methodisch von Interesse ist natürlich, wie bei einer vorliegenden Korrelation zwischen x und y entschieden werden soll, ob x kausal auf y einwirkt oder nicht. Dem gängigen Verständnis nach ist ein solches Urteil im Grunde problemlos zu fällen, wenn man nur ein Experiment durchführen kann. Die zitierte Kritik von Cronbach et al. oder Magnusson tut dem keinen Abbruch. Bestritten werden schließlich nicht die methodischen Vorzüge des Experiments, wenn es darum geht, einzelne Kausalbeziehungen zu überprüfen. Allein, die Praxis wird in Frage gestellt, sich in der Forschung auf einzelne Kausalbeziehungen zu konzentrieren. Entschließt man sich aber zu einem Experiment, gilt die interessierende Entscheidung deshalb als leicht, weil eine trennscharfe Urteilsgrundlage gegeben scheint: *Nur kausal zu interpretierende Korrelationen bleiben unter Experimentalbedingungen erhalten.*

5 **Selektionsthese zum Experiment:**

Unter Experimentalbedingungen bleiben nur Kausalzusammenhänge systematisch erhalten, wohingegen bloße Korrelationen aufgelöst werden.

Eine rein-korrelative Beziehung zwischen x und y hört auf zu existieren, sobald man x als unabhängige Variable experimentell manipuliert und anschließend y erfasst. Außerhalb des Experiments könnten x und y aber bspw. deshalb korrelieren, weil sie beide von einem dritten Faktor z beeinflusst werden. So könnte bspw. in einer reinen Korrelationsstudie ein positiver Zusammenhang zwischen dem Konsum von Kondensmilch und der Kon-

zentrationsspanne von Personen gefunden werden, weil beide Faktoren vom Kaffee- und damit auch vom Koffeinkonsum der Betroffenen abhängen.

Bleibt eine Korrelation zwischen x und y jedoch auch unter Experimentalbedingungen erhalten, ist es offenkundig möglich, durch eine Variation von x unterschiedliche Werte von y herzustellen. Die Korrelation zwischen x und y wird dann kausal interpretiert.

So kann auch gleich eine erste *methodologische These* festgehalten werden, denn man pflegt eine spezifische Vorstellung davon, welche Handlung auf Seiten des Wissenschaftlers angemessen ist. Hier geht es dabei um die Frage einer angemessenen Dateninterpretation.

6 Methodologische These zum Experiment / Kausal interpretieren:

Bleibt eine Korrelation unter Experimentalbedingungen erhalten, ist sie kausal zu interpretieren.

In dem oben diskutierten Beispiel findet der junge Psychologiestudent eine negative Korrelation zwischen der Koffeineinnahme seiner Probanden und ihrer Konzentrationsleistung. Würde der Student ein Experiment durchführen, müsste er die Teilnehmer seiner Studie per Zufall verschiedenen Versuchsbedingungen zuweisen und selbst die Kontrolle darüber übernehmen, wieviel Koffein jeder Teilnehmer erhält. Bei einem solchen methodischen Vorgehen würde die negative Korrelation nicht erhalten bleiben, sie würde sich in dem Fall sogar zu einer positiven Korrelation verkehren. Der positive statistische Zusammenhang wäre dann, ganz wie im eingangs geschilderten Experimentalbeispiel, ohne weiteres kausal zu interpretieren: Die Einnahme von Koffein *verlängert* die Konzentrationsspanne.

1.4 Maßstäbe der Methodenentwicklung

Nach dem Blick auf einen paradigmatischen Anwendungsfall, das Billardspiel, hat die kurze Auseinandersetzung mit der Psychologischen Methodenlehre das Kausalverständnis der Fachgemeinschaft gewissermaßen ausschnitthaft herangezoomt: Abhängige und unabhängige Variable, Randomisierung, Alternativhypothese – unmittelbar taucht man ein in die Welt methodischer Details. Der dritte und abschließende Zug dieses Kapitels soll demgegenüber wieder etwas auf Abstand gehen.

Da gibt es zum einen das kausale Idiom des Alltags, in dem wir uns über Billardspiele und Computerviren verständigen. Zum anderen gibt es in der Psychologie eine ausgefeilte kausalanalytische Forschungsmethodik mit ureigenen Konzepten und Vokabularen. So liegt es nahe, abschließend der Frage nachzugehen, wie sich das Kausalverständnis der Psychologischen Methodenlehre zu alltäglichen Kausalurteilen verhält – und wie sich das Verhältnis gestalten *sollte*.

Zunächst bietet sich ein Blick zur Nachbardisziplin Philosophie an, deren Situation eine ähnliche ist. Auch Philosophen arbeiten zum Teil mit stark formalisierten Kausalitätskonzeptionen und ureigenen Fachvokabularen. Gleichwohl lebt die philosophische Debatte, wie Geert Keil (2006) betont, sehr wesentlich von einem steten Abgleich mit alltäglichen Kausalurteilen. Die Beteiligten der Debatte halten philosophischen Theorien zu Kritikzwecken alltagsweltliche Beispiele entgegen, in denen wir intuitiv einen Kausalzusammenhang ausmachen (oder umgekehrt gerade behaupten wollen, dass kein Kausalzusammenhang besteht), während die Kausalitätsdefinition der philosophischen Theorie uns zum jeweils gegenteiligen Urteil nötigt.

In der philosophischen Diskussion über die Vorzüge und Unzulänglichkeiten der verschiedenen Kausalitätstheorien spielen *Gegenbeispiele* eine prominente Rolle. [...] Ich möchte demgegenüber einen Schritt zurücktreten und fragen, was der Umstand uns lehrt, daß ein