

Überlegungen zur computergerechten Abfassung experimentalpsychologischer Publikationen¹

Dietrich Albert
Universität Graz

It is suggested to supplement the usual psychological experimental reports by a computer-readable description of the method section. Thus, an appropriate coding of the experimental method will be the basis of more straight forward computerized functions for searching, filtering and analyzing experimental reports. A definition of psychological experiments in general is presented which allows for a formal coding of specific experiments. The general coding scheme is based on concepts of discrete mathematics. It will be supplemented by a lexicon which establishes the relationship between the coded, formal description of an experiment and the reality of an experiment. The development of software tools is necessary for the aims of computer-based searching, filtering, classifying, comparing and interpreting of the formal descriptions of experiments. Superordinated tools will provide additional functions: An intelligent editor supports the generation of the formal description of an experiment, a verbalization tool transfers the formal description into a natural language description, a programming tool - in cases of computerized experiments - transfers the formal description automatically into a program code for running the experiment, an interface tool provides the format for storing the experimental data and for analyzing them by statistical software packages. The coding schema, the lexicon, and the software tools should be developed by an interdisciplinary team of experts under the auspices of the scientific societies.

1. Einleitung

Viele Wissenschaftler² meines Fachs – der Psychologie – klagen darüber, daß die Flut der für sie potentiell interessanten Publikationen kaum mehr überschaubar und rezipierbar ist und auch ihre eigenen Publikationen seltener zitiert und gelesen werden, als sie es verdient hätten. Gelegentlich heißt es sogar, die meisten Fachaufsätze würden von nur drei Personen rezipiert werden, nämlich von zwei Reviewern und dem Autor. Die Publikationsflut ist nicht zuletzt eine Folge der Erleichterungen der wissenschaftlichen Arbeit durch Computeranwendungen, ein Nachlassen ist nicht zu erwarten. Um sie zu bewältigen, erscheinen computerbasierte Erleichterungen und Hilfen für die Rezeption von Fachaufsätzen wünschenswert.

-
1. Herrn Prof. Dr. Ferdinand Merz (Universität Marburg) zum 71. Geburtstag gewidmet. Rainer Mausfeld (Kiel), Sylvie Molitor-Lübbert (Karlsruhe) und Norbert Tanzer (Graz) danke ich sehr für Überarbeitungshinweise zu einer früheren Fassung.
 2. Bei männlichen Ausdrücken denke ich den jeweils weiblichen oder geschlechtsneutralen Ausdruck mit; hiermit bitte ich den Leser, ebenso zu verfahren.

Die folgenden Überlegungen zur Überwindung von einigen der durch die Publikationsflut hervorgerufenen und mit ihr einhergehenden Probleme sind streng genommen Vorüberlegungen, die auf Wunsch der Herausgeberinnen zur Publikation zur Verfügung gestellt wurden. Als derzeitiger Vorsitzender des wissenschaftlichen Beirates der Zentralstelle für Psychologische Information und Dokumentation (ZPID) sehe ich meine Aufgabe auch darin, neue Formen und Inhalte psychologierelevanter Datenbanken zu fördern sowie Methoden und Techniken anzuregen, die das Recherchieren, die Selektion und die Rezeption von wissenschaftlichen Texten und Daten in der Psychologie erleichtern und verbessern helfen. Hinzu kommt, daß eine Lösung der Probleme und deren Ausarbeitung nicht von einem Einzelwissenschaftler zu erwarten ist, dafür erscheint vielmehr eine inter- bzw. multidisziplinäre Forschungskooperation von Experimentalpsychologen, Mathematikern, Modell- und Wissenschaftstheoretikern, Forschern der künstlichen Intelligenz und Computerlinguisten (im weiteren Sinne von Kommunikations- und Kognitionswissenschaftlern) erforderlich. Möglicherweise trägt also eine Publikation meiner ersten Überlegungen dazu bei, eine weiterführende Forschungskooperation zu initiieren.

Ich setze voraus, daß es demnächst – ein gewiß dehnbarer Begriff – Volltextdatenbanken für experimentalpsychologische Untersuchungsberichte und Faktendatenbanken für die in den Untersuchungen experimentell erhobenen Daten geben wird. Auf diese Situation sollten sich m. E. das Fach Psychologie und ihre Fachgesellschaften – wie z. B. die Deutsche Gesellschaft für Psychologie – vorbereiten.

Die grundsätzliche Problemstellung ist folgende: Wie müssen Publikationen über (theoretisch fundierte) psychologische Experimente und die Speicherung der Publikationen in Volltextdatenbanken sowie die Speicherung der zugehörigen Daten in Faktendatenbanken beschaffen sein, damit Aufgaben der unten genannten Art mittels Computer besser bewältigt werden können als bisher. Dabei beschränke ich mich ausdrücklich auf Probleme, die mit der Beschreibung der Versuchsmethodik und der Datenerhebung zu tun haben. Das Problem der Dokumentation von Daten wird nur kurz angesprochen. Die entsprechenden Probleme für die Darstellung von Theorien und Ergebnissen werden hier ausgeklammert.

Es wird davon ausgegangen, daß die herkömmlichen, in Umgangs- und Fachsprache abgefaßten Beschreibungen von Versuchsmethodik und Datenerhebung auch zukünftig ihre kommunikative Funktion haben werden, jedoch auf absehbare Zeit Computerprogramme für semantische Analysen derartiger Texte allenfalls für spezifische Zwecke zur Verfügung stehen werden. Deshalb wird vorgeschlagen, die herkömmlichen Versuchsbeschreibungen durch computergerechte Beschreibungen mittels geeigneter Kodiervorschriften zu ergänzen.

2. Problemstellung

Eine computergerechte Beschreibung der Versuchsmethodik und der Datenerhebung psychologischer Experimente kann und soll nicht ein Ziel an sich bzw. Selbstzweck sein; in diesem Falle wäre sie entbehrlich bzw. lediglich eine Arbeitsbeschaffungsmaßnahme. Andererseits wäre die Forderung, sie müsse allen Wissenschaftlern und akademisch ausgebildeten Praktikern der Psychologie nützen, zu weitreichend. Vielmehr ist eine an Problemen und Nutzergruppen orientierte Verbesserung und Weiterentwicklung computergerechter Methodenbeschreibung anzustreben. Zu unterscheiden sind Probleme der retrospektiven und der prospektiven Nutzung von experimenteller Fachliteratur durch unterschiedliche Nutzergruppen.

Generelles Ziel erfahrungswissenschaftlicher Forschung ist es, theoretische Annahmen zu finden, aus denen sich logische Folgerungen ableiten lassen, die zutreffende Aussagen über empirische Zustände und Phänomene sind bzw. sich als solche interpretieren lassen – also erfahrbare Realität theoretisch zu erklären.

Auch in der relativ jungen Wissenschaft Psychologie gibt es bereits gut erforschte Realitätsbereiche, wie z. B. Wahrnehmungsphänomene, in denen – außer für historiographische Ziele – eine retrospektive Betrachtung von Experimenten und eine Nachauswertung experimenteller Daten eher selten vorkommt und meist auch nicht erforderlich ist. In einem fortgeschrittenen theoretischen Entwicklungsstand ist nämlich eine Form der kumulativen Forschung möglich, für die lediglich das Problem besteht, den aktuellen theoretischen Erkenntnisstand durch eine bestimmte (vergleichsweise kleine und überschaubare) Auswahl von experimentellen Untersuchungen empirisch zu fundieren und diese Experimente zu dokumentieren. Es handelt sich dabei in erster Linie um Beschreibungen spezieller Experimente, in denen regelhaft stabile und theoretisch relevante Effekte bzw. Phänomene auftreten.

Außerhalb der Wahrnehmungsforschung und – mit Einschränkungen – der psychologischen Entscheidungsforschung läßt sich m. W. in der Psychologie dieser wünschenswerte Zustand bislang eher selten antreffen. Selbst in traditionellen Gebieten experimenteller psychologischer Forschung, wie z. B. der Gedächtnisforschung, ist kumulative Forschung im Sinne einer sukzessiven, aufeinander aufbauenden theoretischen Weiterentwicklung nur gelegentlich anzutreffen. Vielmehr gibt es zahlreiche Beispiele – auch aus jüngster Zeit – der Reinterpretation von in der älteren Fachliteratur dokumentierten experimentellen Daten und Befunden mittels neuer theoretischer Ansätze, wie dies beispielsweise die Reanalysen von zur Zweifaktoretheorie der Interferenz durchgeführten Experimenten mittels der „Search of Associative Memory“-Theorie (SAM) belegen. Während hierbei noch eine partiell kumulative Theorieentwicklung zu beobachten ist, trifft dies bei den neuesten Bemühungen, konnektionistische Modelle zur Erklärung gedächtnispsychologischer Phänomene zu entwickeln, weniger zu. Ein automatisiertes Auffinden von für derartige Reanalysen geeigneten Experimenten wäre nicht nur eine Arbeitser-

leichterung, sondern würde auch einem Auswahlbias entgegenwirken und die Reanalysen auf eine breitere, objektivere Basis stellen. Diese Form der Forschungsentwicklung ist zumindest insoweit eingeschränkt kumulativ, als neue theoretische Modelle verwendet werden, um bereits vorhandene experimentelle Befunde besser interpretieren zu können.

Eine automatisierte Analyse der Beschreibung von Experimenten würde auch dazu beitragen können, nichtkumulative Forschungstrends zu entdecken und sie in kumulative zu überführen. Nichtkumulative Forschungstrends sind u. a. dadurch gekennzeichnet, daß unter einem neuen Etikett – meist motiviert durch einen vermeintlich neuen, wenig weitreichenden theoretischen Ansatz (Mini-, Theorie‘) – alte Experimente dem Prinzip nach wiederholt werden („Alter Wein in neuen Schläuchen“), oder es werden in immer neuen Experimenten Variationen ohne erkennbaren theoretischen Fortschritt vorgenommen.

Der Beitrag einer automatisierten Auswertung der Beschreibung von Experimenten zur Überführung in kumulative Forschungstrends würde darin bestehen können, Experimente zu klassifizieren, Ähnlichkeiten, Analogien oder Beziehungen zwischen ihnen zu entdecken und Experimente zu ordnen. Oft können Beziehungen zwischen Experimenten nur dann erkannt werden, wenn ein bestimmter Abstraktionsgrad der Beschreibungen vorliegt – also Äquivalenzklassen der beschreibenden Eigenschaften anhand eines geeigneten Lexikons gebildet werden, welches einerseits der begrifflichen Ordnung dient und andererseits den Bezug zur Realität herstellt.

Für die Lösung angewandter, praktischer Probleme ist „nichts so gut wie eine gute Theorie“, nur ist diese häufig nicht verfügbar. Gleichwohl kann es nützliche empirische, erfahrungswissenschaftliche Informationen geben, die – ausgehend von der aktuellen praktischen Problemstellung – durch eine automatisierte Suche aufgrund von computergerechten Beschreibungen von Experimenten besser gefunden werden können, als mit herkömmlichen Methoden.

In einem fortgeschrittenen Stadium der Theorienbildung sind Berichte über neue Experimente nur unter dem Gesichtspunkt der Klärung theoretischer Probleme und theoretischer Weiterentwicklung bedeutsam. Ob dies im Einzelfall zutrifft, wird sich in der Regel kaum allein anhand der Methodenbeschreibung klären lassen. Auch ist in weit entwickelten oder hoch spezialisierten Teilbereichen das kommunikative Netzwerk meist dicht genug, um auf computerunterstütztes Ausfiltern spezifischer, interessanter Experimente verzichten zu können.

In den meisten anderen Fällen jedoch ergibt sich das Problem, aus einer großen Fülle von nahezu täglich erscheinenden Experimentalberichten die jeweils relevanten herauszufiltern. Dies automatisch durchführen zu können, wird in Zukunft möglicherweise noch bedeutsamer werden; durch die Möglichkeit, weltweit praktisch an jedem Computer-Arbeitsplatz wissenschaftliche Berichte der Fachöffentlichkeit über Netz anbieten zu können, werden viele der traditionellen Filter – wie z. B. Refereeverfahren – außer Kraft gesetzt.

Darüber hinaus kann es eine ganze Reihe von spezifischen Gründen für retrospektive und prospektive Literatursuchen geben, die im Falle einer geeigneten Kodierung von Methodenbeschreibungen automatisiert durchgeführt werden könnten. Dazu gehören beispielsweise

- das Auffinden von Vorgängerexperimenten oder Replikationen bei unvollständiger Zitierweise (z. B. bedingt durch Zitationskartelle oder bei Nichtberücksichtigung ausländischer Fachzeitschriften);
- das gezielte Recherchieren und Auffinden von Publikationen über Experimente, in denen bestimmte konkrete experimentelle Bedingungen realisiert worden sind (z. B. um festzustellen, ob für die Prüfung einer neuen theoretischen Aussage möglicherweise schon experimentelle Befunde vorliegen);
- das Auffinden von Publikationen über Experimente, die Realisierungen von bestimmten abstrakten Strukturen für Versuchspläne sind (z. B. für didaktische Zwecke in der Methodenlehre);
- das Auffinden und Vergleichen von Experimenten mit ‚überlappenden‘ oder einander einschließenden Versuchsplänen bzw. -bedingungen (z. B. im Hinblick auf die Stabilität eines Effektes).

Aus welchen Gründen auch immer eine Literaturrecherche angezeigt ist, bereits ein beschränktes Zeitbudget erzwingt ein hohes Maß an Zielgerichtetheit von Literaturrecherchen, d. h. sowohl die Anzahl der irrelevanten als auch die der nicht entdeckten Nachweise sollen gleichermaßen gering sein. Die heute zur Verfügung stehenden Verfahren (wie die Schlüsselwort-Methode, die Durchsicht von Titel- und Inhaltsverzeichnissen, wie z. B. in Current Contents, das Durchsuchen von Abstract-Sammlungen, wie z. B. PsycInfo und Psyn dex, oder von Fachzeitschriften, die Suche auf der Basis von Zitationen per (Social Science Citation Index) leisten dies trotz Computerunterstützung nicht in dem Maße, wie es aus der Sicht eines Experimentalpsychologen wünschenswert ist. Vielmehr erscheint eine gezieltere Suche von Experimenten realisierbar durch die Entwicklung geeigneter Codierungsvorschriften und Filter.

Oft ist nicht nur die Darstellung der theoretischen Überlegungen, sondern auch die Beschreibung der experimentellen Methodik zu ungenau, um nachvollzogen bzw. repliziert werden zu können. Insbesondere sind vollständige, partielle oder modifizierende Replikationen von Experimenten erforderlich, um die Stabilität von Ergebnissen und Effekten bzw. Phänomenen unter gleichartigen und variierten Bedingungen zu untersuchen. Möglicherweise könnte – sozusagen als erwünschter Nebeneffekt – eine computergerechte Kodierung der Methodik dazu beitragen, das Abfassen des Methodenteils einer experimentalpsychologischen Publikation samt ‚Vollständigkeitsprüfung‘ zu unterstützen.

Zu einer hinreichend vollständigen Beschreibung eines Experimentes gehört streng genommen auch die Veröffentlichung der Original- bzw. Rohdaten, durch die eine Reihe von Problemen zu lösen wären. So sind die heute üblichen, indirekten Formen von Metaanalysen experimenteller Ergebnisse für viele Methodiker und Theo-

retiker der Psychologie inakzeptabel; wünschenswert ist vielmehr ein leichter Zugang zu den Originaldaten von Untersuchungen, um Metaanalysen mehrerer Experimente anhand von Originaldaten einschließlich inferenzstatistischer Entscheidungen vornehmen zu können. Auch für theoriegeleitete Nachauswertungen von experimentellen Daten ist die, eine Methodenbeschreibung ergänzende, Verfügbarkeit der Originaldaten erforderlich (z. B. in Hinblick auf eine theoretische Neuinterpretation; mittels neuer Datenanalyse- und Auswertungsverfahren; mit einem anderen Grad der Vergrößerung).

Wie also sollen Experimente und die in ihnen erhobenen Daten beschrieben und dokumentiert werden, damit derartige Aufgaben und Probleme computerunterstützt bewältigt werden können? Hier wird vorgeschlagen, Experimente computerlesbar zu kodieren und darauf abgestimmte Filter zu verwenden. Als erster Schritt dazu wird nach der Definition eines psychologischen Experimentes gefragt.

3. Zur Definition eines psychologischen Experiments

Zunächst ist zu fragen, was ein (psychologisches) Experiment ist und welches die gemeinsamen Eigenschaften von Experimenten sind, die im Hinblick auf die oben formulierten Problemstellungen von Bedeutung sind. Ich versuche, im folgenden eine nichttechnisch formulierte, pragmatische Antwort zu geben, die m. E. unterschiedlichen Zielen von Experimenten genügt und die eine Grundlage für die Lösung der genannten Probleme liefern kann. Dabei verzichte ich sowohl darauf, die wissenschaftstheoretische Bedeutung des Experiments in der Psychologie abzuhandeln als auch die übliche Definition von Experiment vorwegzustellen, die in der einschlägigen experimentalpsychologischen Literatur zu finden ist. Einleitend sei lediglich daran erinnert, daß dabei Eigenschaften wie Kontrollierbarkeit, Replizierbarkeit, isolierte Bedingungsvariation und Objektivität genannt werden. Diese und andere Eigenschaften werden im Kontext einer eher abstrahierenden Darlegung an geeigneten Stellen (in runden Klammern) erwähnt werden.

Generelles Ziel psychologischer Forschung ist es – wie bereits gesagt –, Theorien zu finden, die zutreffende Aussagen über empirische Zustände und Phänomene machen. Ziel empirischer Forschungsarbeit ist es demnach, durch Beobachtungen in einem bestimmten, theoretisch interessanten Realitätsausschnitt festzustellen, welche der möglichen experimentellen Zustände bzw. Ereignisse tatsächlich auftreten und welche nicht oder wie häufig dies geschieht. Sodann ist zu prüfen, ob das Auftreten der beobachteten Ereignisse oder Phänomene mit den allgemeinen theoretischen Prinzipien bzw. den aus ihnen logisch abgeleiteten Aussagen übereinstimmt bzw. verträglich ist. So wurde beispielsweise schon vorexperimentell beobachtet, daß ein bewegter Körper, der sich in benachbarten Zeitintervallen befindet, auch benachbarte Ortskoordinaten hat – und sich nicht etwa an zwei beliebigen, ebenfalls existierenden anderen Orten befindet. Als ein in der Psychologie seit langem bekanntes theoretisch bedeutsames Phänomen sei folgendes Beispiel genannt: Betrachtet man beidäugig mit einem Graufilter vor einem Auge ein

in frontoparalleler Ebene schwingendes Pendel, so sieht man nicht nur ein Rechts-links-, sondern auch ein Vor-zurück-Schwingen, also eine elliptische Bewegung des Pendels. Üblicherweise wird die Theoriebezogenheit zwar zu den wünschenswerten, nicht aber zu den notwendigen, den definierenden Eigenschaften von Experimenten gerechnet, allerdings ist die – in den Beispielen ebenfalls angesprochene – Beobachtbarkeit eine definierende Eigenschaft von Experimenten.

Um von der speziellen Variante empirischer Forschungsmethoden, die Experiment genannt wird, sprechen zu können, müssen mindestens zwei weitere Eigenschaften zusätzlich zur Beobachtbarkeit eines (möglichst theoretisch) interessanten Realitätsausschnittes erfüllt sein.

Um die zweite definierende Eigenschaft von Experimenten herzustellen, wird vor Beginn der Beobachtungen nicht nur mehr oder weniger präzise bestimmt, welcher Realitätsausschnitt beobachtet werden soll, sondern es wird darüber hinaus zuvor eindeutig festgelegt, welche der beobachtbaren empirischen Ereignisse registriert werden sollen (und wie dies geschieht). So wird in einem Wahlreaktionsexperiment beispielsweise angegeben, welche Anzahlen von Reaktionsalternativen vorkommen, wie die räumliche Anordnung der Start- und Zielpositionen ist, zu welchen Zeitpunkten im Laufe eines Experiments welche Signale auftreten können, welche Latenzen gemessen werden und wie genau dies geschieht. Dabei gilt es im allgemeinen als unerheblich für die Definition des Experimentes, welche Gründe – seien sie theoretischer oder nichttheoretischer Art – ein Experimentator für eine spezifische Festlegung von zu registrierenden Ereignissen hatte.

Zu den zu registrierenden Ereignissen gehören – wie das einfache Beispiel eines Wahlreaktionszeitexperimentes bereits veranschaulichte – auch die vom Experimentator selbst hergestellten Ereignisse. Eine dritte, das Experiment kennzeichnende Eigenschaft ist nämlich das gezielte Herstellen von Ereignissen in der Realität durch den Experimentator (Herstellbarkeit, Experimentator-abhängige Ereignisse, Manipulierbarkeit, Kontrollierbarkeit, Realisierung und Registrierung der ‚unabhängigen Variablen‘ und der ‚konstanten Bedingungen‘ wären hier als Stichwörter zu nennen). Das gezielte Herstellen von Ereignissen geschieht einerseits aus Gründen der ökonomischen Versuchsdurchführung – man überläßt es nicht ‚dem Zufall‘ bzw. ‚Mutter Natur‘, ob bestimmte der zuvor als zu registrierende festgelegten Ereignisse auftreten oder nicht und wann und wie oft dies geschieht. Andererseits ist auf diese Weise sichergestellt, daß ein Experiment prinzipiell in wichtigen Ausschnitten wiederholt werden kann (Wiederholbarkeit, Replizierbarkeit). Ferner können auf diese Weise Ereignisse hergestellt werden, die normalerweise selten oder überhaupt nicht in der Realität auftreten würden, wie z. B. eine für jedes der beiden Augen separate, simultane Darbietung unterschiedlicher Objekte, wie es bei stereoskopischer Darbietung geschieht. Dieses Beispiel verdeutlicht bereits, daß die Kunst und der Einfallsreichtum des Experimentators darin besteht, interessante und theoretisch aussagekräftige Bedingungen herzustellen – ein Beispiel aus neuerer Zeit ist das Experiment von Sperling zur Speicherka-

pazität des visuellen Speichers. Das alleinige Beobachten und Registrieren der vom Experimentator selbst hergestellten Ereignisse bringt selbstverständlich keinen Erkenntnisgewinn – es sei denn für einen Studierenden, der Kenntnisse über Techniken des Experimentierens zu erwerben hat. Insofern wäre es ohne jeden wissenschaftlichen Erkenntniswert – und es wäre kein Experiment –, wenn alle in einer Untersuchung tatsächlich registrierten Ereignisse vom Experimentator hergestellt worden wären, wenn er alle ‚Freiheitsgrade‘ einschränken würde. Das Ergebnis einer solchen Untersuchung wäre trivial. Es muß also im Experiment Ereignisse geben, deren Auftreten zwar möglich – und aufgrund von (theoretischer) Vorüberlegungen als zu registrierend vorgesehen – aber nicht sicher ist. Im psychologischen Experiment werden diese Ereignisse in der Regel durch das Verhalten einer oder mehrerer Versuchspersonen hervorgerufen (Versuchsperson-abhängige Ereignisse, Registrierung der ‚abhängigen Variablen‘). Diese Ereignisse sind bedingte Ereignisse, treten sie doch unter der Bedingung der vom Experimentator hergestellten und sonstiger im Experiment ebenfalls zu registrierender Ereignisse (den experimentellen Bedingungen) auf.

Welche Ereignisse im Sinne von Versuchsbedingungen – die hergestellten und die sonstigen – als zu registrierende vorgesehen werden, hängt – ebenso wie die Festlegung der Versuchspersonen-abhängigen, zu registrierenden Ereignisse – von der jeweiligen Zielsetzung eines Experimentes ab. Oft – aber nicht notwendigerweise – werden die Mengen der zu registrierenden Ereignisse (Variablen und ihre Abstufung) durch eine Theorie festgelegt. So müssen z. B. bei der Auswahl zu registrierender Ereignisse (Faktorstufen, Ausprägungen der unabhängigen Variablen) häufig Entscheidungen getroffen werden, für die die jeweilige Theorie nicht allein maßgebend ist. Hinzu kommen für alle psychologischen Experimente gültige (und zumeist sinnvolle) Konventionen. So ist es beispielsweise üblich, das Alter von Versuchspersonen auch dann – als sonstiges Ereignis – zu registrieren, wenn es sich nicht um ein entwicklungspsychologisches Experiment handelt.

Eine Definition des Experimentes wird jedoch auch Ziele einschließen müssen, die mit der Prüfung von theoretischen Aussagen nicht verbunden sind. So werden z. B. bei hinreichend hohem Entscheidungs- und Handlungsdruck Probleme der angewandten Forschung im Sinne der sogenannten ökologischen Validität relativ untheoretisch experimentell untersucht, weil eine geeignete Theorie (noch) nicht bekannt ist und ihre Entwicklung nicht abgewartet werden kann.

Aus einer Definition des Experimentes werden demnach jene Untersuchungen ausgeschlossen, in denen freie Antworten von Vpn ohne vorher festgelegtes Kodierungsschema beobachtet und notiert werden – selbst dann, wenn die Versuchsleiter-abhängigen und sonstigen Bedingungen zuvor präzise festgelegt worden sind. All diejenigen Untersuchungen, in denen nicht einmal dieses geschieht, sind selbstverständlich auch keine Experimente.

Die Mengen aller in einem bestimmten Experiment zu registrierenden Ereignisse sind vom Experimentator also vorher festgelegt worden. Damit ist auch die Menge

der mittels eines bestimmten Experimentes prinzipiell unterscheidbaren möglichen Kombinationen von zu registrierenden Ereignissen bzw. die Menge der Elementarereignisse (im Sinne einer Ereignisalgebra der Wahrscheinlichkeitstheorie) bekannt, im Extremfall also das Kartesische Produkt der Mengen der zu registrierenden Ereignisse.

Dabei werden durch die Produktbildung bereits auf Basis gesicherten Vorwissens alle Kombinationen ausgeschlossen, von denen a priori sicher ist, daß sie aufgrund von Unvereinbarkeitsbeziehungen nicht auftreten können, also Kombinationen von disjunkten, zu registrierenden Ereignissen. So wird beispielsweise aus der Menge der möglichen Reaktionszeiten genau eine in jedem Durchgang gemessen oder aus einer Menge von Reizkonstellationen genau eine auf einer bestimmten Fläche eines Bildschirms dargeboten. Die Kodierung als Kartesisches Produkt von Mengen einander ausschließender zu registrierender Ereignisse, den – wie man sagt – experimentellen Variablen, Konstanten usw., ist bereits angelegt in der Notation üblicher Beschreibungen von faktoriellen Versuchsplänen. Hierbei handelt es sich allerdings lediglich um einen rudimentären, allerersten Ansatz einer formalen Kodierung von Experimenten.

In der durch Produktbildung entstandenen Menge von Elementarereignissen, also in der Menge der geordneten Tupel (,Vektoren‘) von zu registrierenden Ereignissen, sind ferner alle Tupel bzw. Elementarereignisse nicht enthalten, deren Realisierung von vornherein durch Festlegungen des Experimentators ausgeschlossen sind, wie z. B. bei sogenannten ,genesteten Designs‘. Die verbleibende Menge von Elementarereignissen wird hier ,Registrierraum‘ genannt.

Der Registrierraum eines Experimentes besteht also aus allen prinzipiell möglichen Kombinationen von experimentellen Ereignissen, deren Registrierung vorgesehen ist. Diese Kombinationen der zu registrierenden Ereignisse werden hier (experimentelle) Elementarereignisse genannt.

Die bisher vorgenommene Unterscheidung der zu registrierenden Ereignisse in drei Kategorien (Versuchsleiter-, Versuchsperson-abhängige und sonstige) ist relativ undifferenziert. In jedem Experiment lassen sich nämlich zusätzlich die folgenden Klassen von zu registrierenden Ereignissen unterscheiden:

- zu registrierende Ereignisse, die vom Experimentator hergestellt oder manipuliert werden, die sogenannten Unabhängigen experimentellen Variablen, wie z. B. die Anzahl der Alternativen in einem Wahlreaktionsexperiment;
- zu registrierende Ereignisse, die von der Versuchsperson erzeugt werden (die sogenannten ,abhängigen Variablen‘, wie z. B. die Reaktionszeit);
- unkontrollierte oder anderweitig – d. h. nicht vom Experimentator – ,kontrollierte‘ zu registrierende Ereignisse (sogenannte ,Störvariable‘, z. B. die Höhe des Luftdrucks, Luftfeuchtigkeit im Versuchsraum, Blutzuckerspiegel der Vp);

- außerdem gibt es Mengen von zu registrierenden Ereignissen, die möglicherweise nur ein Element enthalten, nämlich die vom Versuchsleiter kontrollierten konstanten Versuchsbedingungen, wie beispielsweise die Frequenz eines akustischen Vor- oder Ankündigungssignals oder eine für alle Versuchsbedingungen und Versuchspersonen gleiche Instruktion;
- unter den zu registrierenden Ereignissen spielt die Zeit eine zentrale Rolle, d. h. ein geordnetes Tupel von aneinander angrenzenden physikalischen Zeitintervallen. Jedes experimentelle Ereignis ist mindestens einem Zeitintervall zugeordnet. Durch diese Zuordnung wird z. B. festgelegt, wann ein Stimulus ein- und aussetzt. Es wird also seine Intensität als Funktion der Zeit angegeben. Ohne diese Zuordnung wäre eine (formale) Beschreibung des Versuchsablaufes unmöglich. Auch wird durch diese Zuordnung das Verhalten der Versuchsperson als ein in der Zeit ablaufender ‚Prozeß‘ darstellbar, oder es werden Reaktionslatenzen bestimmbar;
- ein psychologisches Experiment enthält regelhaft die Menge der Versuchspersonen (sie kann nur ein Element enthalten, $N=1$, aber auch aus mehreren Teilmengen oder Gruppen von Versuchspersonen bestehen, wie in vielen sozialpsychologischen Experimenten) und der oft einelementigen Menge der Versuchsleiter.

Auch die versuchspersonenabhängigen zu registrierenden Ereignisse müssen in manchen Experimenten differenziert betrachtet werden; so kann es z. B. sinnvoll sein, sowohl die Rohdaten als auch on-line bereits vorausgewertete Daten zu registrieren, wie dies beispielsweise in psychophysiologischen Untersuchungen vorkommt.

Werden – wie oben ausgeführt – die diesen Klassen zugeordneten Mengen von Ereignissen als zu registrierende für ein bestimmtes Experiment präzisiert und unter Berücksichtigung von Unverträglichkeiten miteinander kombiniert, so ergibt sich dadurch die Gesamtmenge der unterscheidbaren experimentellen Elementarereignisse, also der Registrierraum dieses Experimentes.

Es handelt sich bei dem Registrierraum – wie gesagt – um eine Menge von Vektoren bzw. geordneten Tupeln von zu registrierenden Ereignissen. Bei entsprechender Anordnung lassen sich die Vektoren in zwei Abschnitte – einen für die Experimentator-abhängigen, einen für die Versuchspersonen-abhängigen zu registrierenden Ereignisse – und in Unterabschnitte entsprechend den oben unterschiedenen Klassen unterteilt denken.

Die jeweilige Struktur eines Experimentes wird durch die Mengen der zu registrierenden Ereignisse, ihre Klassifikation und die Eigenschaften der laut Versuchsplan vorgesehenen Teilmengen aus dem Kartesischen Produkt dieser Mengen gebildet.

Das Interesse des Experimentators richtet sich in aller Regel nicht auf das Auftreten von experimentellen Elementarereignissen, sondern auf Teilmengen von Elementarereignissen, die hier – in Anlehnung an eine Ereignisalgebra der Wahr-

scheinlichkeitstheorie – experimentelle Ereignisse genannt werden.³ Schon praktischer Gründe wegen werden Elementarereignisse zu Ereignissen durch Vergrößerung bzw. Äquivalenzklassenbildung zusammengefaßt. Ein hoher Differenzierungsgrad von Elementarereignissen und ihre große Anzahl erzwingen (angesichts von vergleichsweise wenigen Realisierungschancen wegen begrenzter Zeit, geringerer Anzahl von Versuchspersonen usw.) Zusammenfassungen, um überhaupt Aussagen über die Häufigkeit von empirischen Ereignissen machen zu können. Diesen entsprechen theoretische Aussagen über Wahrscheinlichkeiten von theoretischen Ereignissen, die den experimentellen zugeordnet sind. Oder es wird eine deterministische Aussage über das Auftreten eines bestimmten Phänomens geprüft, welches durch genau eine Teilmenge von Elementarereignissen definiert worden ist. Tritt eines der Elementarereignisse dieser Teilmenge in einem Experiment auf, so ist das das Phänomen definierende experimentelle Ereignis beobachtet worden. Auf die gleiche Weise erfolgt die Prüfung einer deterministischen Aussage über das Auftreten von z. B. Wissens- bzw. Performanzzuständen im Sinne der Wissensraum-Theorie, durch welche die Menge der Elementarereignisse in zwei Teilmengen zerlegt wird.

Der Begriff des Experimentes wird also so gefaßt, daß ein Kodierungsschema entwickelt werden kann, welches für eine formale Methodenbeschreibung geeignet ist.

4. Das Ziel einer formalisierten Methodenbeschreibung

Mein ‚Plädoyer‘ zielt darauf ab, die gesamte Methoden-Beschreibung – und nicht nur die Beschreibung eines faktoriellen Versuchs-Designs – einer experimentellen Untersuchung zu formalisieren bzw. zu kodieren, also in computerlesbaren Symbolen der Mathematik auszudrücken. Die Mathematik, von der dazu nur ein kleiner Ausschnitt benötigt wird, eignet sich dazu in idealer Weise; sie besitzt – stark verkürzt ausgedrückt – eine normierte, international gebräuchliche Sprache, die es ermöglicht, Objekte auf unterschiedlichem Abstraktionsgrad präzise zu beschreiben und zu definieren, und deren Symbole und Zeichenketten von Computern verwendet werden können, wie z. B. durch Programme symbolischer Formelbearbeitung.

Jeder Einzelversuch wird – mittels eines Lexikons, welches die Verbindung zwischen den beschreibenden Symbolen und der Realität herstellt – abgebildet durch ein Elementarereignis aus der Menge aller durch den Registrierraum festgelegten Elementarereignisse bzw. durch die das Elementarereignis kodierenden Symbole. Jeder Einzelversuch ist disjunkt zu allen anderen, sofern die zu registrierenden Ereignisse hinreichend differenzierend festgelegt worden sind, also z. B. Zeitpunkt

3. Dadurch wird der Begriff des Ereignisses in diesem Aufsatz in zwei Bedeutungen verwendet: „Ereignis“ im umgangssprachlichen Sinne und „experimentelles Ereignis“ im Sinne einer Ereignisalgebra.

der Untersuchung und Name der Versuchsperson registriert werden.⁴ Einer experimentellen Untersuchung entspricht demnach eine Teilmenge des Registrierraumes, die durch Aufzählung von Elementen oder deren Eigenschaften symbolisch kodiert wird. Der Registrierraum ist als Kartesisches Produkt oder Teilmenge desselben ein einfaches mathematisches Objekt, welches sich mit den Ausdrucksmitteln bzw. Zeichen der Mathematik beschreiben läßt. Umgekehrt läßt sich auch eine experimentelle Untersuchung als eine spezielle Interpretation der formalen, symbolischen Beschreibung des einfachen mathematischen Objektes auffassen, wobei die Zuordnung von realen Versuchereignissen zu den Symbolen durch ein „Lexikon“ oder „Glossar“, welches für die spezielle Interpretation gilt, vorgenommen wird.

Als Ausdrucksmittel für die formale, symbolische Beschreibung stehen natürlich zunächst die üblichen Notationen der klassischen Mengenlehre zur Verfügung, also z. B. geschweifte Klammern, Zeichen für algebraische Operationen mit Mengen (für Produktbildung, Mengenvereinigung usw.) und der unterschiedlichen Formen, Mengen anzugeben (durch ihre Elemente, durch beschreibende Eigenschaften usw.). Ferner wird man den Begriff der Relation und den von Operationen mit Relationen (Produktbildung, Verkettung) und entsprechende Symbole benötigen, um den Registrierraum konstruieren und beschreiben zu können; auch Spezialfälle von Relationen wie die Begriffe der Abbildung und Funktion werden benötigt, um z. B. Vorverarbeitung von Daten und sogenannte adaptive Untersuchungsverfahren abbilden zu können. Bei adaptiven Verfahren wird ein Experimentator-kontrolliertes Versuchereignis, z. B. die Darbietung einer Aufgabe, von einem Versuchspersonen-kontrollierten Versuchereignis, z. B. von den Antworten bei vorherigen Aufgaben, abhängig gemacht.

Ein mit diesen Ausdrucksmitteln konstruierter und beschriebener Registrierraum stellt ein einfaches Objekt dar, welches zerlegt werden kann, mit anderen Registrierräumen verkettet oder hinsichtlich struktureller Eigenschaften mit ihnen verglichen werden kann.

Um eventuellen Mißverständnissen vorzubeugen, es geht mir nicht darum, eine Versuchsbeschreibung durch eines der üblichen Versuchssteuerungs- und Datenerfassungsprogramme vorzunehmen, welches von vielen Psychologen ebenfalls als eine formale Beschreibung eines Experimentes akzeptiert werden würde. Zwar lassen sich auf diese Weise im Prinzip alle vorkommenden Einzelversuche erzeugen, deren Registrierraum muß jedoch eigens beschrieben werden. Aus einer derartigen mathematisch-formalen Beschreibung läßt sich natürlich auf maschinell Wege automatisch (eine geeignete Software vorausgesetzt) ein Versuchssteuerungsprogramm erzeugen (nicht jedoch umgekehrt).

Es können allerdings Algorithmen, die in heute üblichen Versuchssteuerungsprogrammen verwendet werden, in die formalisierte Versuchsbeschreibung aufgenommen werden, und zwar zur Charakterisierung von Mengen und Funktionen. Ange-

4. Anderenfalls müßte im Folgenden zusätzlich der Begriff des ‚multi set‘ eingeführt werden.

nommen beispielsweise, die Menge der verwendeten Stimuli ist durch einen Mandelbroschen Algorithmus oder durch eine generative Grammatik erzeugt worden, dann wird man die Stimulusmenge nicht nur durch abstrakte Symbole $\{x, y, z, ..\}$ und die Stimuli selbst $\{.., .., .., ..\}$, sondern auch durch Angabe des Algorithmus oder der Grammatik beschreiben – ganz im Sinne unterschiedlicher Abstraktions- und Kodierungsniveaus.

Um den für die Beschreibung von Experimenten verwendeten formalen Symbolen die Fachausdrücke eines beschreibenden Textes und die realen Versuchereignisse eindeutig zuordnen zu können, bedarf es – wie bereits erwähnt – eines Lexikons. Die Symbole verweisen auf fachsprachliche Begriffe, die in einem Lexikon derart durch Texte, Abbildungen, Videos, Ausstellungsstücke eines virtuellen Museums usw. definiert sind, daß sie auch dann eindeutig auf Versuchereignisse verweisen, wenn der originale Versuchsaufbau selbst nicht mehr zur Verfügung steht. Das Lexikon enthält auch Informationen zu synonymen und homonymen Begriffen. Außerdem werden Ober-Unterbegriffs-Relationen dokumentiert, also die Ordnungsstruktur, um Vergrößerungen und Abstraktionen vornehmen zu können. Auch Begriffe für Vorverarbeitungsmethoden und für Funktionen bzw. Algorithmen adaptiver Versuchssteuerung werden ebenso definiert sein wie physikalische und psychologische Maßeinheiten.

Selbstverständlich müssen die symbolischen und fachsprachlichen Ausdrucksmittel für die Beschreibung von Experimenten sowie das Lexikon, welches den Bezug zur experimentellen Realität herstellt, ständig ergänzt werden. Sie sind dem wissenschaftlichen Erkenntnisstand anzupassen, dessen theoretischen und methodischen Entwicklungsstand sie widerspiegeln.

Die formale bzw. kodierte Beschreibung der Methode von Experimenten ist die Voraussetzung, um mittels geeigneter Software-Tools die oben genannten Aufgaben und Probleme bewältigen bzw. lösen zu können.

5. Ziel der Entwicklung von Software-Tools

Um über derartige Vorüberlegungen hinauszukommen, muß der erste Schritt darin bestehen, ein Kodierungsschema für die Beschreibung von psychologischen Experimenten zu entwickeln und an konkreten Experimenten zu erproben.

- Im nächsten Schritt müssen, um die oben genannten Ziele erreichen zu können, Software-Tools entwickelt werden, mit denen die zur Versuchsbeschreibung verwendeten Zeichenfolgen
 - interpretiert,
 - gezielt gesucht,
 - miteinander verglichen und
 - ausgewertet werden können.

Einschlägige Vorbilder finden sich beispielsweise in Programmen symbolischer Formelbearbeitung und relationaler Datenbanken.

Mit derartigen Software-Tools würden sich bereits die meisten der oben genannten Aufgaben, wie gezielte Suche von Experimenten, Erkennen von nichtkumulativer experimenteller Forschung usw. bewältigen lassen.

Auf diese Software-Tools lassen sich eine Reihe von weiteren Tools aufsetzen, die darüber hinaus einzeln oder in Kombination die Bewältigung der folgend genannten Aufgaben unterstützen können.

- Mit einer formalen Versuchsbeschreibung werden zwar die oben genannten Aufgaben computergestützt bewältigt werden können, ein unmittelbares Verständnis des Untersuchungsgeschehens ist bei den meisten Interessenten jedoch in der Regel leichter zu erreichen durch eine gut gegliederte umgangs- und fachsprachliche Beschreibung eines Experimentes, wie wir sie heutzutage in guten Fachzeitschriften finden. Da eine gewisse Normierung dieser Beschreibungen für das Verständnis und die Archivierung vorteilhaft ist, wäre ein Software-Tool wünschenswert, welches derartige umgangs- und fachsprachliche Beschreibungen automatisch erzeugt, also ein Verbalisierungs-Tool bzw. ein Tool zur automatischen Textproduktion. Die erzeugten verbalen Beschreibungen wären – da sie Redundanz erzeugen – sicher auch eine effektive Hilfe und Kontrolle während des Prozesses der Erstellung der formalen Versuchsbeschreibung. Um das Verbalisierungs-Tool zu entwickeln, ist natürlich die interdisziplinäre Kooperation durch Sprachwissenschaftler und Informatiker anzureichern.
- Das Verbalisierungs-Tool könnte ein Bestandteil eines ‚Intelligenten Editors‘ sein, welcher die Erzeugung von formalisierten Versuchsbeschreibungen unterstützt. ‚Intelligent‘ würde der Editor genannt werden, wenn er Prüfungen auf Syntax, Vollständigkeit usw. automatisch vornimmt und ‚lernfähig‘ ist.

Da mittels des Intelligenten Editors formale Beschreibungen von Experimenten erzeugt werden, können diese neu interpretiert werden, um ein neues konkretes Experiment zu erhalten, indem den abstrakten Symbolen andere Ereignisse in der Realität zugewiesen werden. So läßt sich eine Reihe von abstrakten Standardversuchsbeschreibungen angeben, die für die jeweilige Untersuchung nur neu interpretiert zu werden brauchen.

Der Intelligente Editor sollte benutzerfreundlich sein. So könnte beispielsweise für Standardversuchsbeschreibungen, also für Experimente mit wiederkehrender abstrakter Struktur, ein Formular-Editor verwendet werden. Mittels des Verbalisierungs-Tool würden die mit dem Intelligenten Editor erhaltenen Arbeitsergebnisse kontrolliert werden können.

Auch die Entwicklung eines Intelligenten Editors ist eine interdisziplinäre Aufgabe.

- Wie bereits erwähnt, läßt sich aus einer vollständigen, formalen Beschreibung eines Experimentes automatisch ein Versuchssteuerungsprogramm erzeugen – vorausgesetzt, ein entsprechendes Software-Generierungs-Tool wurde entwickelt. Für Computer-gestützte Experimente würde damit neben dem Verbalisie-

rungs-Tool eine weitere redundante Hilfe und Kontrolle für die Arbeit zur Erstellung der formalen Versuchsbeschreibung mittels des Intelligenten Editors bereitstehen.

- Da mit dem Intelligenten Editor eine formal-sprachlich normierte Beschreibung des Registrierraumes – also des gesamten Experimentes – erzeugt wird, wird gleichzeitig die Struktur für die Registrierung, Speicherung und Archivierung der im Experiment anfallenden Daten erzeugt. Ein entsprechendes Software-Tool würde dem Nutzer eine für sein Experiment speziell zugeschnittene und gleichzeitig für experimentelle Faktendatenbanken normierte Schnittstelle bereitstellen.

Für derart beschriebene Experimente wäre demnach die Erfassung der Daten in einer Faktendatenbank leicht zu realisieren.

- Außerdem wird durch die formale Beschreibung des Experimentes eine normierte „Schnittstelle“ für Auswertungsprogramme und die Beschreibung des Ergebnisteils eines experimentalpsychologischen Manuskripts bereitgestellt und in einem Software-Tool realisiert werden.

Die Entwicklung dieser Software-Tools kann sukzessive und geordnet erfolgen. Dadurch wird gewährleistet, daß die jeweils entwickelten Tools bereits genutzt werden können, bevor das Entwicklungsprogramm abgeschlossen ist.

6. Schluß

Die vorgeschlagenen Tools für den Methodenteil von Experimenten und die Speicherung der experimentellen Daten sind m. E. in interdisziplinärer Kooperation vergleichsweise leicht zu realisieren. Es wird von der Initiative Einzelner und der Fachgesellschaften wie auch von den Forschungsförderungsorganisationen abhängen, wann angesichts zunehmender Publikationsflut derartige Hilfsmittel zur Verfügung stehen und dazu beitragen werden, den Zugang zu experimentalpsychologischen Publikationen und möglicherweise auch deren Qualität zu sichern und zu verbessern.