

TIMO GNAMBS / BARBARA STRASSNIG

## Experimentelle Online-Untersuchungen

### Internet-basierte Experimente

Ein Anliegen vieler Wissenschaftsdisziplinen besteht in der Ableitung von Gesetzmäßigkeiten zur Erklärung und Prognose von Phänomenen. Dazu ist es notwendig, anfangs unspezifische Vermutungen in Form eindeutiger, testbarer Hypothesen zu formulieren und sie mit geeigneten Methoden systematisch und kontrolliert auf ihren Wahrheitsgehalt zu überprüfen. Das Experiment nimmt zu diesem Zweck im Methodenrepertoire der sozial- und wirtschaftswissenschaftlichen Forschung (besonders seit Anfang des 19. Jahrhunderts) eine unverzichtbare Position als methodisches Hilfsmittel zur Untersuchung von Ursache-Wirkungs-Relationen, also Kausalbeziehungen, zwischen interessierenden Merkmalen ein, die anderen Forschungszugängen (z.B. korrelativen Studien) in der Regel verschlossen bleiben. Die Ableitung von kausalen Zusammenhängen zur Erklärung und Prognose von (Alltags-)Phänomenen wird durch die spezifische Versuchsanordnung experimenteller Studien ermöglicht. Der Versuchsleiter realisiert gezielt eine Untersuchungssituation, in der systematisch eine oder mehrere unabhängige Variablen, die unter seiner Kontrolle liegen, variiert und die Auswirkungen dieser Manipulation auf eine oder mehrere abhängige Variablen beobachtet werden können. Eine darauf gründende Ableitung und gegebenenfalls Quantifizierung kausaler Zusammenhänge ist jedoch nur dann statthaft, wenn nicht nur die unabhängigen Variablen systematisch variiert, sondern gleichzeitig die Wirkung aller anderen Einflussgrößen (Störvariablen), welche die Beziehung zwischen den zu untersuchenden Merkmalen beeinflussen können, kontrolliert werden. Eine derartige Kontrolle personengebundener

Störvariablen (z. B. Motivation, Interesse, Geschlecht o. Ä.) erfolgt häufig über den Prozess der Randomisierung, bei dem die Versuchspersonen zufällig auf die verschiedenen experimentellen Bedingungen aufgeteilt werden. Gelingt diese Zufallsaufteilung, was bei ausreichend großen Stichprobenumfängen der Fall sein sollte, wirken (unbekannte) potentiell störende Einflüsse anderer Variablen in allen Experimentalbedingungen gleichermaßen und entfalten keine systematischen Störeffekte. Schwieriger ist die Kontrolle umweltbedingter Störvariablen (DÖRING 2003), die durch variierende Situationsmerkmale (z.B. Lärm, Tageszeit o. Ä.) hervorgerufen werden. Dazu ist es notwendig, die Umgebungsbedingungen, unter denen die experimentelle Testung abläuft, genau zu kennen, und vor allem für alle Versuchspersonen konstant zu halten. Experimente, bei denen diese Kontrolle von Störvariablen gut gelingt, weisen eine hohe interne Validität auf; d.h. die Veränderungen in den abhängigen Variablen sind eindeutig auf die bewusste Manipulation der unabhängigen Variablen zurückzuführen und nicht auf störende Einflussfaktoren. Kausalinterpretationen können deshalb weitgehend zweifelsfrei gezogen werden. Als zweites Kriterium zur Beschreibung experimenteller Untersuchungen wird die externe Validität herangezogen. Diese bezeichnet den Grad der Generalisierbarkeit der Untersuchungsbefunde auf andere Situationen als jene, in der das Experiment durchgeführt wurde; also das Ausmaß der Übereinstimmung von tatsächlichem und intendiertem Untersuchungsgegenstand. Verschiedene ExperimentalfORMen unterscheiden sich nun hinsichtlich ihrer internen und externen Validität. Laborexperimente finden unter eindeutigen, vom Untersuchungsleiter vorab festgelegten, und weitgehend kontrollierbaren Rahmenbedingungen an einer bestimmten Örtlichkeit (meist ein Labor) statt und bieten allen Untersuchungspersonen weitgehend vergleichbare Umgebungsbedingungen. Die interne Validität derartiger Untersuchungen ist demnach groß. Allerdings finden diese Experimente unter eher künstlichen Bedingungen statt, die häufig mit Alltagssituationen nur bedingt vergleichbar sind. Somit ist bei Laborexperimenten zumeist von einer eher geringen externen Validität auszugehen. Im Gegensatz dazu finden Feldexperimente direkt im natürlichen Umfeld der Versuchspersonen (außerhalb des Labors) statt. Während die externe Validität bei Feldexperimenten bedeutend größer ist, wird durch die zumeist sehr eingeschränkte Möglichkeit zur Kontrolle von Störvariablen die interne Validität eingeschränkt.

Wie lassen sich nun Internet-basierte Experimente in diese Klassifikation traditioneller Untersuchungsmethoden einordnen? Generell werden im Bereich der Internet-Forschung zwei Typen von Experimentalstudien unterschieden: Web-Experimente und Online-Feldexperimente (DÖRING 2003). Web-Experimente stellen als Pendant zu traditionellen Laborexperimenten das Gros gängiger Experimentalbemühungen im Internet dar. Die Versuchspersonen greifen von einem beliebigen Ort aus über das Internet auf einen Experimental-Server zu, der ihnen (experimentell variiertes) Stimulusmaterial oder Web-Fragebögen über den Bildschirm zur Bearbeitung online zur Verfügung stellt. Der gesamte experimentelle Ablauf, von der Instruktion über die randomisierte Zuteilung zu den verschiedenen Untersuchungsbedingungen bis hin zum abschließenden Debriefing der Versuchspersonen, erfolgt online, vollkommen automatisiert und ohne Eingriff eines Versuchsleiters über den Computer. Die Untersuchungssituation ist in vielen Belangen vergleichbar mit Computertestungen in Experimentallabors, wie sie seit Einführung des Computers in den wissenschaftlichen Arbeitsalltag häufig eingesetzt werden (MUSCH/REIPS 2000). Lediglich die Standardisierung der Umgebungsbedingungen kann nicht gleichermaßen konsequent umgesetzt werden, da die Experimentalteilnahme eben nicht in den kontrollierten Räumlichkeiten des Untersuchungsleiters (Labor) erfolgt, sondern im privaten bzw. beruflichen Umfeld der Versuchspersonen, in dem sie auf das Internet-Experiment zugreifen.

Im Gegensatz dazu werden bei Online-Feldexperimenten Versuchspersonen nicht, wie bei Web-Experimenten üblich, eingeladen, ein bestimmtes Experiment auf einem vom Versuchsleiter aufgesetzten Experimentalserver zu bearbeiten, sondern sie werden vom Versuchsleiter aktiv aufgesucht und in ihrem natürlichen Umfeld studiert. Glaser, Dixit und Green (2002) untersuchten beispielsweise in einem Online-Feldexperiment die Bereitschaft, unter verschiedenen Bedrohungsszenarien rassistische Gewaltakte zu verüben. Da öffentliche Zustimmung zu derartigen Handlungen in der Regel auch von dafür prädestinierten Personengruppen nur selten zu erhalten ist, und dementsprechend klassische Web-Experimente zu wenig validen Ergebnissen führen, traten sie als Teilnehmer einschlägigen Diskussionsgruppen im Internet bei und führten als traditionelle Chat-Dialoge getarnte (experimentell variierte), semi-strukturierte Interviews durch. In vielen Fällen, wie in dieser Studie, ist den teilnehmenden Versuchspersonen gar nicht bewusst, dass

sie an einer experimentellen Untersuchung teilnehmen. Sie zeigen ausschließlich ihr natürliches Verhalten. Die experimentelle Manipulation wird ihnen dabei gar nicht mitgeteilt. Die aus derartigen Untersuchungsdesigns resultierenden ethischen Bedenken, wie z.B. fehlende Freiwilligkeit der Teilnahme oder nachträgliche Aufklärung über den Studienhintergrund, wurden bislang noch nicht befriedigend gelöst (vgl. NOSEK/BANAJI/GREENWALD 2002). Häufig steht bei Online-Feldexperimenten natürliches Verhalten in Online-Umgebungen im Zentrum des Interesses. Anstelle von Befragungen über hypothetische Verhaltensweisen, wie in Web-Experimenten, interessieren dabei primär Beobachtungen typischen Verhaltens (z.B. Navigieren auf Webseiten, Bezahlen von Einkäufen etc.), das jedoch den zu testenden Hypothesen entsprechend bestimmten Manipulationen unterworfen wird (z.B. SHOCHAT/MUSCH 2003). Da Online-Feldexperimente in der Regel relativ aufwendig zu realisieren sind, finden sie bislang nur in Einzelfällen Anwendung und stellen im Vergleich zu Web-Experimenten eine Randerscheinung Internet-basierter Forschungsansätze dar.

### Möglichkeiten und Grenzen von Web-Experimenten

Der Einsatz von Web-Experimenten wie ganz allgemein Online-Befragungen hat sich seit der explosionsartigen Verbreitung des Internets in Privathaushalten und der damit verbundenen zunehmenden Vernetzung breiter Bevölkerungsschichten seit Mitte der 1990er-Jahre stetig gesteigert (MUSCH/REIPS 2000) und findet insbesondere in angewandten Forschungsbereichen regen Zuspruch. Der ADM (Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute) gibt beispielsweise für das Jahr 2005<sup>1</sup> an, dass bereits 22 Prozent aller quantitativen Interviews<sup>2</sup> der Mitgliedsinstitute online durchgeführt wurden. Eine vergleichbare Verbreitung ist für den akademischen Sektor bislang nicht in Sicht. Skitka und Sargis (2006) schätzen den Anteil von in APA-Zeitschriften publizierten Studien, die Internet-basierte Erhebungsmethoden einsetzen, für die Jahre 2003/04 auf lediglich 1,6 Prozent. Ungeachtet dessen erfreut sich auch im wissenschaftlichen Bereich das Web-Experiment in den unter-

1 Siehe: [http://www.adm-ev.de/pdf/Jahresbericht\\_05.pdf](http://www.adm-ev.de/pdf/Jahresbericht_05.pdf) (2007-05-17).

2 Web-Experimente werden nicht explizit angeführt.

schiedlichsten Disziplinen (in der Persönlichkeits-, Kognitions- und Medienpsychologie ebenso wie in der Wahrnehmungs- und Entscheidungsforschung u.v.m.) einer steigenden Beliebtheit (REIPS 2003).

TABELLE 1

## Vorteile von Web-Experimenten

### VORTEILE

1. Auflösung geografischer Grenzen	Mit vergleichbar geringem Aufwand und in sehr kurzer Zeit können auch internationale und interkulturelle Studien durchgeführt werden (Brennan/Geuens/Weijters/Smith/Swinyard 2005).
2. Kurze Erhebungszeiträume	Der Großteil der Versuchspersonen nimmt innerhalb der ersten drei Tage nach Einladung teil oder gar nicht (Shannon/Bradshaw 2002).
3. Automatische Datenspeicherung	Durch die automatische elektronische Speicherung am Server entfällt die fehleranfällige und zeitintensive Eingabe per Hand.
4. Erhebung non-reaktiver Daten	Verhaltensdaten und technische Informationen in Form von Logfiles geben im Rahmen der Drop-Out-Analyse Aufschluss über Antwortverhalten und Verweildauer (Reips 2002).
5. Adaptive bzw. dynamische Gestaltung des Versuchsablaufs	Durch adaptive Versuchsabläufe und Variation der Items kann der Aufwand für potentielle Versuchspersonen gering gehalten werden.
6. Erreichbarkeit spezifischer Zielgruppen	Spezielle »stigmatisierte« Gruppen, die der Forschung traditionell nur sehr schwer zugänglich sind, können über das Internet leichter angesprochen werden (Rhodes/Bowie/Hergenrath 2003).

Nach Reips 2002; Rhodes/Bowie/Hergenrath 2003

Web-Experimente bieten gegenüber traditionellen Laborexperimenten eine Reihe von Vorteilen, aber natürlich auch manche Nachteile (vgl. Tab. 1). Großteils ergeben sich ähnliche Vorzüge und Schwierigkeiten, die auch anderen Online-Methoden wie zum Beispiel den inzwischen häufig verwendeten Online-Befragungen zu Eigen sind. Neben der Verfügbarkeit diversifizierter, also nicht nur wie in der Praxis häufig üblich studentischer Stichproben (GOSLING/VAZIRE/SRIVASTAVA/JOHN 2004) ermöglichen Web-Experimente durch geeignete Rekrutierungsstrategien auch Zugang zu speziellen Populationen, die auf herkömmlichem Weg nicht oder nur schwer zu erreichen wären (z. B. wie bei GLASER et al. 2002 zu deklarierten

Rassisten). Insbesondere erscheint die Möglichkeit, im Vergleich zu Laborexperimenten in kürzester Zeit umfangreiche Stichproben zu erreichen, für viele Forscher verlockend. Trotz dieser Aussicht darf dabei jedoch nicht vergessen werden, dass eine reine Erhöhung des Stichprobenumfangs nicht notwendigerweise zu besseren Ergebnissen führt und Repräsentativerhebungen über das Internet trotz Verfügbarkeit umfangreicher Online-Access-Panels (auch mit Personengruppen mit durchaus großer soziodemografischer Bandbreite) bislang und wohl auch in naher Zukunft nicht realisiert werden können, da nicht alle Bevölkerungsschichten und -gruppen gleichermaßen im Internet vertreten sind. Neben dem Einsatz potenziell heterogenerer Stichproben, die eine stärkere Generalisierung der Experimentalbefunde erlauben, weisen Web-Experimente vornehmlich aufgrund ihrer größeren Variabilität hinsichtlich der Untersuchungssituationen (z.B. unterschiedliche Tageszeiten, Umgebungen o. Ä.) eine gegenüber Laborexperimenten höhere externe Validität auf. Da Online-Untersuchungen in der Regel in einer vertrauten Umgebung (meist vor dem eigenen Computer in den eigenen Räumlichkeiten) bearbeitet werden, sind verzerrte Ergebnisse durch den Einfluss ungewohnter Umgebungsbedingungen, wie sie in Labor-Settings auftreten können, zumeist auszuschließen (REIPS 2002). Die automatische Randomisierung und dynamische Zuweisung (auch auf Basis von zuvor gestellten Fragen) von Versuchspersonen zu den Experimentalbedingungen erlauben elaboriertere Stichprobenziehungen (z.B. unterschiedliche Bedingungen auf Basis von Persönlichkeitseigenschaften) und komplexere Untersuchungsdesigns. Die potenziellen Vorteile von Web-Experimenten zeigen sich vornehmlich in Abhängigkeit des jeweiligen Untersuchungsziels. Zu den Feldern, in denen Web-Experimente ihre besondere Stärke entfalten können, zählen u.a. die interkulturelle Forschung (Aufhebung der räumlichen Trennung), anwendungsorientierte Forschung (hohe externe Validität) und Marktforschung (kurze Reaktionszeiten).

Allgemein muss davon ausgegangen werden, dass die interne Validität bei Web-Experimenten gegenüber Laborexperimenten geringer anzusiedeln ist (vgl. Tab. 2). Dies ist unter anderem auf die fehlende bzw. geringere Kontrollierbarkeit der Untersuchungssituation, in der die Experimentalmteilnahme erfolgt, zurückzuführen, da die Versuchspersonen unter beliebigen, vom Versuchsleiter nicht zu beeinflussenden, Umgebungsbedingungen auf das Web-Experiment zugreifen. Daher können Störungen während des Experimentalablaufs (z.B. durch Telefonanrufe oder andere

TABELLE 2

## Nachteile von Web-Experimenten

## NACHTEILE

1. Systematischer Drop-Out	Teilnahmeabbrüche sind bei Web-Experimenten generell häufiger als bei Laborexperimenten (Fricker/Galesic/Tourangeau/Yan 2005). Systematisch gehäufte Abbrüche in bestimmten Experimentalbedingungen schränken zudem die Interpretierbarkeit der Ergebnisse ein.
2. Selbst-selegierte Stichprobe	Um Verzerrungen durch die Freiwilligkeit der Experimental-Teilnahme auszuschließen, schlägt Reips (2002) die <i>Multiple Site Entry Technik</i> vor.
3. Mangelnde Repräsentativität	Obwohl auch bei Laborexperimenten nur selten auf repräsentative Stichproben zurückgegriffen wird, können bei Web-Experimenten nur Aussagen über die Grundgesamtheit der Internetnutzer getroffen werden (Hauptmann 1999).
4. Geringe Situationskontrolle	Die größere Varianz der Umgebungsvariablen (Ausstattung, auftretende technische Störungen etc.) beeinträchtigt die interne Validität des Experiments.

Nach Reips 2002; Rhodes/Bowie/Hergenrath 2003

Personen) nicht ausgeschlossen werden und die Datenqualität beeinträchtigen. Technische Anomalien aufgrund unterschiedlicher Betriebssysteme oder Browser können zudem die Funktionsfähigkeit von Web-Experimenten beeinträchtigen und einen zusätzlichen Bias generieren, indem ungewollt bestimmte Personengruppen mit spezifischer Softwarekonstellation von der Teilnahme ausgeschlossen werden. Angesichts von Untersuchungsbefunden (BUCHANAN/REIPS 2001), die vermuten lassen, dass Personen, die unterschiedliche Technologien nutzen, sich auch hinsichtlich bestimmter Persönlichkeitseigenschaften unterscheiden können, ist derartigen unwillkürlichen Technologie-bedingten Selektionsprozessen so gut als möglich vorzubeugen. Obwohl in Web-Experimenten die Versuchsleiter eine weitgehend untergeordnete Rolle spielt, da das Experiment, einmal aufgesetzt, automatisch ohne weiteres Zutun von außen abläuft, und klassische Versuchsleitereffekte somit nicht zum Tragen kommen können, sind vergleichbare Einflussgrößen im Rahmen der Versuchspersonenrekrutierung nicht gänzlich auszuschließen. Bereits kleine Details, wie das Geschlecht des einladenden Forschenden, können massive Auswirkungen auf die Rücklaufquoten haben (ALTHOFF/GREIF/GRIEL/BATINIC 2006). Die

fehlende Präsenz der Versuchsleiter begünstigt, wie häufig befürchtet wird (DÖRING 2003), zudem mutwillig falsche oder unvollständige Angaben bzw. eine geringere Elaboriertheit der Antworten. Das größte Problem bei Web-Experimenten stellt jedoch systematischer Drop-Out dar; d.h. wenn in bestimmten Versuchsbedingungen systematisch mehr Teilnahmeabbrüche zu verzeichnen sind als in den übrigen. Dies kann beispielsweise in Lernexperimenten, in denen in einer Bedingung aus forschungsgeleiteten Gründen eher ein langweiliger Lerntext vorgegeben wird, in dieser Bedingung vermehrt Drop-Out zu verzeichnen sein, sodass nur noch der ›harte Kern‹ der motivierten Teilnehmer das Experiment beendet. Während die Teilnehmer dieser Bedingung somit nur noch aus hoch motivierten Personen bestehen, finden sich in den übrigen Experimentalbedingungen Personen mit unterschiedlicher Motivationslage. Derartige unwillkürliche Selektionen können Ergebnisse und Interpretationen massiv beeinträchtigen, wenn der Drop-Out nicht als zusätzlich Informationsquelle herangezogen wird.

### Zur Kontrolle von Drop-Out

Es liegen verschiedene Ansätze vor, systematischem Drop-Out entgegenzuwirken. Bei der *High-Hurdle-Technik* (REIPS 2002) werden am Anfang des Web-Experiments möglichst viele für die Versuchsperson demotivierende Faktoren gesetzt; wie z. B. eine absichtlich verlängerte Ladezeit der Startseite, eine Personalisierung der Befragung durch verpflichtende Angabe der E-Mail-Adresse oder ein Hinweis, dass durch die ebenfalls gespeicherte IP-Adresse die Versuchspersonen eindeutig identifizierbar sind. Das Ziel dieses Ansatzes ist es, Personen die Teilnahme an der Untersuchung als möglichst verbindlich darzustellen und mögliche ›Schnupperer‹, die gar nicht die Absicht haben, ernsthaft an der Untersuchung teilzunehmen, gleich zu Beginn abzuschrecken. Negativ ist hierbei der Umstand zu bewerten, dass auch Personen von der Teilnahme abgehalten werden, die möglicherweise erst im Laufe der Untersuchung, nachdem sie bereits Fragen beantwortet haben, eine höhere Motivation entwickeln würden. Die *Warm-Up-Technik* (REIPS 2002) beruht auf der Beobachtung, dass ein Großteil des Drop-Outs zu Beginn eines Experiments auftritt, bevor sich Personen ernsthaft dafür entschieden haben, am Experiment (nicht) teilzunehmen. Um den Drop-Out während der relevanten Phasen eines



ABBILDUNG 1

## Administrationsoberfläche von Globalpark zur Erstellung von Web-Experimenten

Frage Subjektives Wissen bearbeiten (Typ 311, Seite Subjektives Wissen)

Fragebogen | Übersicht Seite | Fragevorschau | Fragetyp ändern | Skalensystem | in Bibliothek eintragen | Platzhalter | + | -

Se bearbeiten den Fragebogen in der Sprache "Deutsch".

**Frageeigenschaften**

Frageid:

Frageinhalt:

Ausfüllanweisung:

Pflichtfrage (ja/ne):

Breite der linken Spalte mit den Einträgen (in Post):

Variables:  ID:  Anzahl:

Typ:  Eingabeformat:  Randomis. Anordnung:  Status:  Anzahl:  Anzahl:  Anzahl:

Massenexport Forme | In Liste aufnehmen | Gruppenübersicht

Neu:  Antwortkategorie:

Variablen	wert	ID	Anzahl	Typ	Eingabeformat	Randomis. Anordnung	Status	Anzahl	Anzahl	Anzahl
V_3	<input type="text"/>	<input type="text"/>	<input type="text"/>	A	<input type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V_4	<input type="text"/>	<input type="text"/>	<input type="text"/>	A	<input type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V_5	<input type="text"/>	<input type="text"/>	<input type="text"/>	A	<input type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
V_6	<input type="text"/>	<input type="text"/>	<input type="text"/>	A	<input type="text"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Legende:  wird von Plauschecks benutzt  in Ausfüllbedingungen verwendet  wird von Filten benutzt  in Triggern verwendet  in Listen verwendet

Gruppe

Experiments möglichst gering zu halten, wird bei der Warm-Up-Technik der Beginn der experimentellen Manipulation im Untersuchungsablauf möglichst weit nach hinten verschoben. Zu Untersuchungsbeginn werden allgemeine Themen wie Soziodemografie und Fragen zur Exploration der Teilnahmemotivation behandelt. Das eigentliche Web-Experiment mit den experimentell variierten Untersuchungsbedingungen folgt erst im Anschluss. Dadurch soll sichergestellt werden, dass nur noch die motivierten Personen das eigentliche Experiment in Angriff nehmen und kein systematischer Drop-Out zu verzeichnen ist. Schließlich können den Versuchspersonen noch *Incentives* (meist materielle Anreize wie Geld oder Gutscheine) für eine Experimentaltelnahme in Aussicht gestellt werden. Diese steigern nicht nur die Bereitschaft, Online-Untersuchungen in Angriff zu nehmen, sondern auch die Motivation, sie bis zum Ende fertig zu bearbeiten (GÖRTZ 2006). Allerdings sollten sie dennoch vorsichtig und wohlüberlegt eingesetzt werden, da Incentives auch eine potenzielle Störvariable darstellen können, indem sie die Motivation bzw. Motive

zur Teilnahme beeinflussen (REIPS 2000). Über die Seriosität bzw. Datenqualität der durch Incentives erkaufte zusätzlichen Versuchspersonen ist bislang wenig bekannt.

## Umsetzung von Web-Experimenten

Die Realisierung experimenteller Designs im Internet erfordert die Entwicklung entsprechender Programmlogiken zur automatisierten und randomisierten Zuteilung der Versuchspersonen auf verschiedene experimentelle Bedingungen. Während für Online-Befragungen bereits eine Vielzahl unterschiedlicher Software-Lösungen zur Verfügung steht, ist die Software-Entwicklung im Bereich des Web-Experimentierens bislang noch nicht so weit fortgeschritten. Kommerzielle Programme zur Umsetzung von Online-Befragungen, die ohne technische Expertise über einfache Web-Formulare zu administrieren sind (vgl. Abb. 1), ermöglichen zwar über die Eingabe von Filterfragen sowie spezifischer Verzweigungsstrukturen in der Regel auch die Realisierung einfacher experimenteller Designs. Die Handhabung ist allerdings in den wenigsten Fällen einfach und intuitiv und erfordert meist eine längere Einarbeitungszeit. Zudem eignen sie sich weniger für die Umsetzung komplexerer experimenteller Designs mit einer größeren Zahl verschachtelter Experimentalbedingungen. Mit WEXTOR<sup>3</sup> (REIPS/NEUHAUS 2002) wurde an der Universität Zürich ein Web-basiertes Instrument zur Generierung komplexer Web-Experimente entwickelt, mit dem relativ einfach auch komplizierte Untersuchungsdesigns (unvollständige Pläne ebenso wie Inner-Subjekt-Designs und Quasi-Experimente) realisiert werden können. Während entsprechende Experimentalpläne relativ einfach konzipiert und auch visualisiert werden können, ist die allgemeine technische Funktionalität der Software relativ eingeschränkt und für professionelle Web-Experimente nur teilweise geeignet. Stößt man mit derartigen Software-Produkten an die Grenzen des Möglichen, was bei anspruchsvollen Projekten schnell der Fall sein kann, ist man auf selbst konstruierte Lösungen angewiesen, die allerdings ein gewisses Maß an technischer Expertise voraussetzen. Mit Hilfe mehr oder weniger leistungsfähiger Editoren sind die notwendigen Programmcodes selbst zu entwickeln (vgl. Abb. 2) und über

3 Verfügbar unter: [http://psych-wextor.unizh.ch/wextor/en/\(2007-05-17\)](http://psych-wextor.unizh.ch/wextor/en/(2007-05-17)).

ABBILDUNG 2

## Erstellung eines Web-Experiments mit einem Editor

### HTML-Quellcode

```

1 <!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN" "http://www.w3.org/TR/html4/strict.dtd">
2 <html><head><title>Online Umfrage</title>
3 <link rel="stylesheet" type="text/css" href="style.css"><style type="text/css"></style></head>
4
5 <body>
6 <form method="post" action="senden.php" name="F1">
7 <div id="header">Allgemeine Informationen </div><div id="inhalt">
8
9 <h2>Liebe TeilnehmerInnen!</h2><br>
10 <p>Vielen Dank für Ihr Interesse!</p>
11 <p>Es folgen drei Befragungen zu verschiedenen Themen.</p>
12 <p>Jede dieser Fragen folgen die drei Fragebögen, wobei jeder Fragebogen sehr kurz ist und in etwa 5 Minuten</p>
13 <p>Insgesamt dauert die Untersuchung ca. 15 Minuten.</p>
14 <p>Bitte lesen sie alle Anweisungen genau durch.</p>
15 <table border="1" width="100%" style="border-color: #FF0000; border-style: solid; border-width: 2px;>
16 <tr>
17 <td id="button"><input type="submit" name="button" value="START" name="F1"></td></tr>

```

---

### Grafische Umsetzung im Browser

## Allgemeine Informationen

## Liebe TeilnehmerInnen!

Vielen Dank für Ihr Interesse!

Es folgen drei Befragungen zu verschiedenen Themen.

Nach diesen Fragen folgen die drei Fragebögen, wobei jeder Fragebogen sehr kurz ist und in etwa 5 Minuten dauert. Insgesamt dauert die Untersuchung ca. 15 Minuten.

Bitte lesen sie alle Anweisungen genau durch.

START

eigene Webserver zu veröffentlichen. »wysiwyg«-Editoren können dabei die Erstellung der grafischen Oberfläche unterstützen; für experimentelle Funktionen wie Zufallszuweisungen ist allerdings »Handarbeit« und Kenntnis der notwendigen Programmbefehle notwendig. Verschiedene frei verfügbare Routinen wie z. B. der *Generic HTML Form Processor*<sup>5</sup> (GÖRITZ/BIRNBAUM 2005) erleichtern die Arbeit, indem sie für häufig verwendete Anwendungsfälle entsprechende Programmbibliotheken bereitstellen. Hat man sich zur technischen Eigenentwicklung von Web-Experimenten entschieden, sollte man sich der Unterscheidung zwischen *serverseitigen*

- 4 wysiwyg ist die Abkürzung für What You See Is What You Get und bezeichnet die Bearbeitung von Dokumenten ohne direkte Programmcode-Eingabe, indem sie während der Bearbeitung genauso angezeigt werden, wie sie später ausgegeben werden sollen. Eine HTML-Seite kann also mit einfachen Arbeitsschritten in einer grafischen Oberfläche erstellt werden ohne dass dabei mit HTML-Befehlen gearbeitet werden muss.
- 5 Verfügbar unter: <http://www.goeritz.net/brmic/> (2007-05-17)

und *clientseitigen* Programmiersprachen bzw. Prozeduren bewusst sein. Diese haben verschiedene Vor- und Nachteile und ihr Einsatz sollte diese Besonderheiten Rechnung tragen. Clientseitige Anwendungen (z.B. in Java-Script), die ausschließlich auf dem Computer des jeweiligen Nutzers ablaufen, haben den Vorteil, dass sie auf bestimmte Angaben sehr schnell reagieren. Allerdings sind die Befehle (der Funktionen) nicht geschützt und von Unbefugten leicht einzusehen. Serverseitige Anwendungen (z.B. in PHP) gewährleisten eine größere Sicherheit gegenüber technischen Ausfällen und ›unauthorisierten‹ Zugriffen. Dafür sind mögliche Funktionen träger in ihrer Reaktion als clientseitige Programmiersprachen. Um möglichst hochwertige Experimente zu erstellen, sollten beide Arten entsprechend ihrer Vorteile eingesetzt werden. Die sensible Speicherung und Zuweisung zu Versuchsbedingungen kann optimal mit PHP umgesetzt werden, Routinen zur Eingabeprüfung, ob alle Items beantwortet wurden, lassen sich einfach mit Java-Script realisieren.

Unabhängig davon, wie das Web-Experiment letztendlich umgesetzt wurde, sollte das Endergebnis gängigen Kriterien des Interface-Designs entsprechen, um aussagekräftige Ergebnisse erzielen zu können. Vor allem der Heterogenität der Software-Ausstattung (Betriebssystem, Browser, Plugins etc.) der designierten Nutzer ist bei der Entwicklung Rechnung zu tragen, um nicht ungewollt durch fehlerhafte Programmierung oder zu hohe technische Hürden bestimmte Personengruppen von der Untersuchungsteilnahme systematisch auszuschließen. Um diese technischen Anforderungen an das (unbekannte) Computersystem der designierten Stichprobe möglichst gering zu halten und eine Technikbedingte Selektion zu vermeiden, sollten u.a. die Kriterien zur Gestaltung von Web-Experimenten beherzigt werden:

- *Konformität:*  
Ausschließliche Verwendung standardisierter Programmbefehle (in HTML etc.),
- *Kompatibilität:*  
Fehlerfreie und identische Darstellung in möglichst allen verfügbaren Browsern,  
Kein Ausschluss bestimmter Personengruppen durch Voraussetzung bestimmter Hardware oder bestimmter Bildschirmauflösungen,
- *Komplexität:*  
Möglichst geringe technische Anforderungen durch Verzicht auf exotische Plugins oder zusätzlich zu installierende Programme.

Falls die optische Darstellung von besonderer Bedeutung für das Experiment ist und die Ergebnisse beeinflussen kann, sollten automatisiert, zusätzliche Informationen über das Computersystem (wie z.B. Bildschirmauflösung) als Kontrollvariablen mit erhoben werden. Diese zusätzlichen Informationen über die Rahmenbedingungen des Web-Experiments können post-hoc im Rahmen explorativer Validitätsanalysen zur Kontrolle systematischer Verzerrungen herangezogen werden.

### Internet-basierte Experimente mit Verhaltensdaten

Wird von experimentellen Designs im Rahmen von Online-Untersuchungen gesprochen, beziehen sich diese primär auf Selbstbeschreibungsverfahren (Fragebögen), deren Inhalt anhand einer oder mehrerer Variablen manipuliert wird. In vielen Untersuchungssituationen stellt jedoch ein bestimmtes Verhalten die interessierende Forschungskonstante dar. Da tatsächliches Verhalten in der Praxis jedoch zumeist relativ aufwendig zu erfassen ist, wird häufig auf Befragungsverfahren zurückgegriffen, um *potenzielles* Verhalten zu erheben. Trotz der Beliebtheit (und in vielen Fällen auch ausreichenden prognostischen Validität) von Selbstbeobachtungen stellen diese eine verzerrte Methode zur Erfassung von *tatsächlichem* Verhalten dar, da derartige Selbstberichte durch eine Vielzahl von Einflussfaktoren unwillkürlich (z. B. Antworttendenzen) aber auch willkürlich (z.B. Lügen) verfälscht werden können. Auch wenn die Bandbreite an möglichen Verhaltensweisen, die über das Internet erhoben werden können, bislang noch nicht mit der Reichhaltigkeit traditioneller Laborexperimente vergleichbar ist, stehen bereits verschiedene Möglichkeiten der natürlichen Verhaltensbeobachtung in Online-Umgebungen zur Verfügung, die sich auch experimentellen Untersuchungsdesigns nicht verschließen. Exemplarisch seien im Folgenden drei Ansätze experimenteller Online-Verhaltensbeobachtungen kurz skizziert.

Ein entscheidender Vorzug von Verhaltensbeobachtungen im virtuellen Raum ist die automatische Aufzeichnung sämtlicher Nutzeraktivitäten. Jegliche Aktion, jeder Seitenaufruf, jeder Klick auf einen Link wird im Internet automatisch mitprotokolliert und generiert dadurch systematische Datenarchive, die auch zur wissenschaftlichen Analyse dienen können. Während derartige Online-Verhaltensspuren bereits

seit längerem im Rahmen von Logfile-Analysen post-hoc wissenschaftlichen Analysen unterzogen werden (vgl. ~~entsprechendes Kapitel in diesem Buch~~), d.h. im Sinne non-reaktiver Daten in der Regel ohne manipulativen Eingriff seitens des Forschers, bieten sich derartige Verhaltensaufzeichnungen auch für experimentelle Untersuchungsdesigns an. Insbesondere die Usability- und Marktforschung bedient sich bereits seit geraumer Zeit dieser Methodik (HOFACKER/MURPHY 2005). So lassen sich etwa durch gezielte Variation bestimmter Designelemente (z. B. die Gestaltung von Navigationselementen oder die Formulierung von Bannertexten) verschiedene Designstrategien von Web-Präsenzen ›live‹ im Web erproben. Weigend (2005) berichtet beispielsweise von einer experimentellen Variation der Platzierung des Warenkorbs (linke bzw. rechte Spalte) in einem populären Webshop, um die Auswirkung auf das tatsächliche Kaufverhalten der Kundschaft empirisch zu testen. Derartige Experimente auf aktiven Webseiten bieten eine Reihe von Vorteilen. Insbesondere die ökologische Validität dieser Untersuchungen ist groß, da die Untersuchungsstichprobe direkt aus den tatsächlichen Nutzern des Web-Angebots stammt und die Experimentalsituation für die Besucher nicht offensichtlich ist; d.h. in der Regel ist ihnen die experimentelle Manipulation gar nicht bewusst. Die Untersuchungen finden nicht in künstlichen Umgebungen statt, in denen sich Versuchspersonen in die jeweilige Praxissituation hineinversetzen sollen. Vielmehr wird das tatsächliche Verhalten in realen Situationen unter willkürlich variierten Situationen untersucht. Online-Feldexperimente lassen sich auch im Rahmen wissenschaftlicher Forschung auf recht unkonventionelle Weise realisieren. Shohat und Musch (2003) untersuchten beispielsweise in einem sozialpsychologischen Online-Feldexperiment ethnische Vorurteile in Online-Auktionshäusern. Sie stellten in Ebay vergleichbare Produkte zur Versteigerung und variierten lediglich die Ethnie der Verkäufer durch Variation ihrer Vornamen. Sie konnten deutliche Diskriminierungstendenzen nachweisen, da Verkäufer mit türkischem Namen deutlich länger auf gewinnbringende Gebote warten mussten als Personen mit deutschen Benutzernamen. Neuere Experimentalmethoden auf Basis von Verhaltenbeobachtungen zur Analyse von Nutzeraktionen finden sich in simulierten virtuellen Welten. Wiederum übernimmt hierbei die kommerzielle Marktforschung eine Vorreiterrolle (DUDA/SCHLIESSEL 2005; SCHREYEN/HERSTELL 2003). In virtuellen Supermärkten (vgl. Abb. 3) werden potenzielle Nutzer durch 3D-Simu-

lationen traditioneller Einkaufsmärkte geführt, in denen sie alltägliche Kaufroutrinen vornehmen sollen. Während reale Verkaufstests in der Praxis relativ aufwendig und kostspielig zu realisieren sind und demnach der experimentellen Manipulation aus ökonomischen Gründen starke Grenzen gesetzt sind, können virtuelle Kaufsituation relativ einfach systematisch variiert und unkompliziert verschiedene Produktgestaltungen oder auch -platzierungen durchgespielt werden. In der Regel stellen derartige Verhaltensbeobachtungen naturgemäß nur einen Teil der erhobenen Daten dar. Ihren Nutzen entfalten sie zumeist durch Kombination mit klassischen Befragungsdaten. Da bislang kaum unkompliziert zu bedienenden Simulationsinstrumente zur Verfügung stehen (vgl. SCHREYEN/HERSTELL 2003), um derartige virtuelle Umgebungen auch ohne ausgewiesene technische Expertise realisieren zu können, finden sich bislang nur wenige praktische Umsetzungen derartiger Experimentalstudien.

#### ABBILDUNG 3

### Virtueller Supermarkt des Lehrstuhls für Wirtschaftsinformatik der RWTH Aachen<sup>6</sup>



<sup>6</sup> Verfügbar unter <http://www.winfor.rwth-aachen.de/inhalte/forschung/vrmafo/supermarkt/ingabe.html> (2007-05-17).

Bislang ist die praktische Bedeutung von Internet-basierten Beobachtungsverfahren noch vergleichsweise gering. Während Online-Untersuchungen und auch Web-Experimente in der wissenschaftlichen und auch kommerziellen Forschung, nicht zuletzt dank einer Vielfalt zur Verfügung stehender Software-Produkte, in zahlreichen Disziplinen einen unverzichtbaren Stellenwert im alltäglichen Forschungsgeschehen erobert haben, ist die Realisierung von Beobachtungen noch weitgehend wenig standardisiert. Entsprechend hohe Anforderungen an das technische Können des Untersuchungsleiters werden gestellt, um online Beobachtungen in ausreichender Qualität realisieren zu können.

### Fazit

Anfängliche Anstrengungen, experimentelle Untersuchungen auf das Internet zu übertragen, waren vornehmlich von Befürchtungen um eine mindere Datenqualität geprägt. Wie jedoch zahlreiche Vergleichsstudien nahe legen (BATINIC 2004), führen unter Beachtung der medien-spezifischen Eigenheiten des Web-Experiments, aufgrund der Internet-basierten Realisierung und damit notwendigerweise verbundenen weniger strikten Kontrollmöglichkeit potenziell störender Rahmenbedingungen, diese zu vergleichbar validen und aussagekräftigen Resultaten wie traditionelle Labor-Experimente. Insbesondere der manipulations-spezifischen Drop-Out-Problematik ist jedoch in geeigneter Weise zu begegnen. Mittlerweile ist das Web-Experiment in eine Konsolidierungsphase eingetreten und zählt zum regulären Methodenkanon einer Vielzahl unterschiedlicher Disziplinen – von sozial- und wirtschaftswissenschaftlichen Forschungsrichtungen bis zu kommerziellen Anwendungen in der Usability- und Marktforschung. Während für experimentelle Web-Befragungen bereits zahlreiche Qualitätskriterien vorliegen (REIPS 2002), öffnen sich im Rahmen experimenteller Feldbeobachtung, insbesondere in virtuellen Simulationen, Anknüpfungspunkte für künftige Methodenentwicklungen.

### Literatur

ALTHOFF, S.; V. GREIF; B. GRIEL; B. BATINIC: *Determinants of Response Rates of Online Surveys – The Anita Effect – Results of a Joint Project*. Vortrag auf der



- General Online Research Conference (GOR), Bielefeld, 2006
- BATINIC, B.: Online-Research. In: BENTE, G.; R. MANGOLD; P. VORDERER (Hrsg.): *Lehrbuch der Medienpsychologie*. Göttingen [Hogrefe] 2004, S. 252-270
- BIRNBAUM, M. H.: Human Research and Data Collection via the Internet. In: *Annual Review of Psychology*, 55, 2003, S. 803-832
- BRENGMAN, M.; M. GEUENS; B. WEIJTERS; S. M. SMITH; W. R. SWINYARD: Segmenting Internet shoppers based on their Web-usage-related lifestyle: a cross-cultural validation. In: *Journal of Business Research*, 58, 2005, S. 79-88
- BUCHANAN, T.; U.-D. REIPS: Platform-dependent biases in Online Research: Do Mac users really think different? In: JONAS, K. J.; P. BREUER; B. SCHAUBURG; M. BOOS (Hrsg.): *Perspectives on Internet Research: Concepts and Methods*. Verfügbar unter: <http://www.psych.uni-goettingen.de/congress/gor-2001/contrib/buchanan-tom> [2007-05-17] 2001.
- DÖRING, N.: *Sozialpsychologie des Internet*. Göttingen [Hogrefe] 2003
- DUDA, S.; M. SCHIESSL: Vorsprung im Supermarkt. Reale und virtuelle Shop-Studien für Hersteller und Handel. In: *Research & Results*, 34, 2005
- FRICKER, S.; M. GALESIC; R. TOURANGEAU; T. YAN: An Experimental Comparison of Web and Telephone Surveys. In: *Public Opinion Quarterly*, 69 (3), 2005, S. 370-392
- GLASER, J.; J. DIXIT; D. P. GREEN: Studying hate crime with the Internet: What makes racists advocate racial violence? In: *Journal of Social Issues*, 58 (1), 2002, S. 177-193
- GÖRITZ, A.: Incentives in Web Studies: Methodological Issues and a Review. In: *Journal of Internet Studies*, 1, o. S. Verfügbar unter: <http://www.ijis.net> [2007-05-17] 2006
- GÖRITZ, A. S.; M. H. BIRNBAUM: Generic HTML Form Processor: A versatile PHP Script to save Web-collected data into a MySQL database. In: *Behavior Research Methods*, 37 (4), 2005, S. 703-710
- GOSLING, S. D.; S. VAZIRE; S. SRIVASTAVA; O. P. JOHN: Should we trust web-based studies?: A comparative analysis of six preconceptions about internet questionnaires. In: *American Psychologist*, 59, 2004, S. 93-104
- HAUPTMANN, P.: Grenzen und Chancen von quantitativen Befragungen mit Hilfe des Internets. In: BATINIC, B.; A. WERNER, L. GRÄF; W. BANDILLA (Hrsg.): *Online Research: Methoden, Anwendungen und Ergebnisse*. Göttingen [Hogrefe] 1999, S. 277-295

- HOFACKER, C. F.; J. MURPHY: Using Server Log Files and Online Experiments to Enhance Internet Marketing. In: KRISHNAMURTHY, S. (Hrsg.): *Contemporary Research in E-Marketing*. Hershey [Idea Group] 2005, S. 226-249
- MUSCH, J.; U.-D. REIPS: A Brief History of Web Experimenting. In: BIRNBAUM, M. H. (Hrsg.): *Psychological Experiments on the Internet*. San Diego [Academic Press] 2000, S. 61-88
- NOSEK, B. A.; M. R. BANAJI; A. G. GREENWALD: E-Research: Ethics, Security, Design and Control in Psychological Research on the Internet. In: *Journal of Social Issues*, 58 (1), 2002, S. 161-176
- REIPS, U.-D.: Standards for Internet-based experimenting. In: *Experimental Psychology*, 49 (4), 2002, S. 243-256
- REIPS, U.-D.: Web-Experimente – Eckpfeiler der Online Forschung. In: THEOBALD, A.; M. DREYER; T. STARSETZKI (Hrsg.): *Online-Marktforschung – Theoretische Grundlagen und praktische Erfahrungen*. Wiesbaden [Gabler] 2003, S. 73-89
- REIPS, U.-D.; C. NEUHAUS; WEXTOR: A Web-based tool for generating and visualizing experimental designs and procedures. In: *Behavior Research Methods, Instruments, and Computers*, 34, 2002, S. 234-240
- RHODES, S. D.; D. A. BOWIE; K. C. HERGENRATHER: Collecting behavioural data using the world wide web: considerations for researchers. In: *Journal of Epidemiology and Community Health*, 57, 2003, S. 68-73
- SCHRYEN, G.; J. HERSTELL: Online-Marktforschung – Analyse von Konsumentenverhalten in virtuellen Umgebungen. In: *Information Management & Consulting*, 18 (2), 2003, S. 74-77
- SHANNON D.; C. BRADSHAW: A comparison of response rate, response time, and costs of mail and electronic surveys. In: *The Journal of Experimental Education*, 70 (2), 2002, S. 179-187
- SHOHAT, M.; J. MUSCH: Online auctions as a research tool: a field experiment on ethnic discrimination. In: *Swiss Journal of Social Psychology*, 62, 2003, S. 193-1945
- SKITKA, L. J.; E. SARGIS: The Internet as Psychological Laboratory. In: *Annual Review of Psychology*, 57, 2006, S. 529-555
- WEIGEND, A.: *Data Mining & E-Business*. Vortrag auf der General Online Research Conference (GOR), Bielefeld, 2005