

Prozessdaten online erheben: Verschiedene Methoden im Überblick

Von Michael Schulte-Mecklenbeck & Ryan O. Murphy

Einführung

Die Online-Befragung bietet seit mehreren Jahren eine Ergänzung und oft sogar Alternative zu klassischen Befragungsmethoden. Der sich aufdrängende Vergleich von online- und offline erhobenen Daten wird in der Literatur breit diskutiert und zeichnet die Online-Befragung meist mit guten Noten aus.¹ Wie in diesem Band deutlich wird, sind die Einsatzmöglichkeiten dieser Methode weit gestreut und beinhalten qualitative wie quantitative Verfahren, die für sozialwissenschaftliche wie auch kommerzielle Fragestellungen eingesetzt werden. Online- wie Offline-Befragungen haben gemeinsam, dass die erhobenen Daten in die Kategorie der Input-Output-Daten fallen.² Dieser Beitrag möchte eine zusätzliche Datenebene beleuchten, die für Online-Befragungen gleichermaßen interessant ist wie für Online-Experimente: die Sammlung von Prozessdaten. Der Begriff Prozessdaten meint in diesem Zusammenhang, dass mehrere Datenpunkte pro Nutzer registriert und verarbeitet werden.

Welchen Mehrwert kann die Erhebung von zusätzlichen Daten bringen? Ist man beispielsweise daran interessiert, mehr über die Nutzbarkeit³ einer Webseite herauszufinden, so kann man die Nutzer nach der Verwendung der Seite befragen. Vieles wird allerdings zu diesem Zeitpunkt schon wieder vergessen sein. Ein genaueres Bild lässt sich z. B. mit der Prozessmethode des „lauten Denkens“ herausfinden, in der die Nutzer parallel zur Bearbeitung der Seite ihre Gedanken verbalisieren (siehe unten). Die daraus resultierenden verbalen Protokolle geben detailliert Auskunft über Probleme bei der Bearbeitung oder Hinweise auf offene Fragen des Nutzers. Wählt man einen noch höheren Auflösungsgrad, so ist die Registrierung von Augenbewegungen das Mittel der Wahl. Die Verweilzeiten auf verschiedenen Teilen der Webseite können mit dieser Prozessmethode im Detail analysiert werden.

Eine Vielzahl von Methoden erfüllt den Anspruch, Prozessdaten zu erheben. Die Liste der „klassischen“ Prozessmethoden enthält unter anderem *Informationstafeln*⁴, die Registrierung von *Augenbewegungen*⁵, die Methode der *aktiven Informationssuche*⁶ oder das *Laute Denken*.⁷ *Informationstafeln* sind, in den ersten Versionen, Tafeln, auf denen Briefumschläge mit Informationskärtchen befestigt sind. Der Teilnehmer nimmt ein Kärtchen aus einem Briefumschlag und steckt es nach dem Lesen wieder zurück. Der Forscher no-

¹ Vgl. Pötschke 2004; Li et al. 2004; Dillmann & Bowker 2001.

² Ein Input ist in diesem Zusammenhang die jeweilige Aufgabe, ein Output die Reaktion des Probanden, z. B. das Ankreuzen einer Antwort in einer Multiple-Choice-Aufgabe. Pro Aufgabe gibt es bei Input-Output-Daten also nur einen Datenpunkt bzw. eine Reaktion, die dann interpretiert wird.

³ Engl. „Usability“; siehe z.B. die Webseite von Jakob Nielsen, einem der führenden Usability-Experten: <http://www.useit.com/>.

⁴ Vgl. Payne 1976; engl.: „Information Boards“.

⁵ Vgl. Russo 1978; engl.: „Eye-Tracking“.

⁶ Vgl. Huber, Wider & Huber 1997; engl.: „Active Information Search“.

⁷ Vgl. Ericsson & Simon 1980; engl.: „Thinking Aloud“.

tiert dabei die Reihenfolge der betrachteten Informationen. Über die Anzahl und Art der betrachteten Informationen können Rückschlüsse auf die Suchstrategien getroffen werden. Eine Weiterentwicklung der Informationstafeln stellen die dynamischen Informationstafeln dar, die vor allem in der Politikwissenschaft eingesetzt werden.⁸ In dieser Methode entscheidet der Teilnehmer zuerst, welche Informationen auf der Informationstafel angezeigt werden sollen, um dann die einzelnen Zellen zu inspizieren. Bei der Registrierung von *Augenbewegungen* wird mittels einer Infrarotlichtquelle ein Reflex auf der Hornhaut erzeugt (Cornea Reflex), der mit einer Infrarotkamera aufgenommen wird. In frühen Versionen dieses Werkzeuges wurden die Augenbewegungsvideos dann von Hand ausgewertet. Mittlerweile gibt es ausgefeilte Computerprogramme, die diese Aufgabe übernehmen und eine Vielzahl von Parametern⁹ registrieren.¹⁰ Bei der *Methode der aktiven Informationssuche* bearbeitet der Nutzer ein Problem, zu dem er Fragen stellen kann. Die Fragen werden mit vorbereiteten Texten auf Karten beantwortet. Diese Methode ermöglicht dem Nutzer freies Explorieren einer Aufgabe. Dem Forscher wird ein unbeeinflusster Einblick in die Informationsselektion des Nutzers geboten.¹¹ *Lautes Denken* wurde Anfangs hauptsächlich in der Problemlöse-Forschung eingesetzt. In dieser Methode wird der Nutzer aufgefordert, Gedanken, die ihm während einer Aufgabe durch den Kopf gehen, zu verbalisieren. Dies kann retrospektiv (also nach dem Ende einer Aufgabe) oder parallel (gleichzeitig mit einer Aufgabe) durchgeführt werden. Lautes Denken ist auch außerhalb der Problemlöse-Forschung eine beliebte Methode, z.B. in der Expertenbefragung.¹²

Das Hauptaugenmerk wird in diesem Kapitel allerdings nicht auf die eben genannten, klassischen Prozessmethoden¹³ gelegt, sondern auf Methoden, die sich durch eine weitere Gemeinsamkeit auszeichnen: den Ort der Datenerhebung. Die im Folgenden vorgestellten Methoden erheben Daten online über das Internet und ermöglichen den Schritt aus dem Labor und der Universität hinaus in die (Online-)Welt. Um den Überblick zu erleichtern, werden wir kurz das Konzept der Informationsumgebung erklären, das die Einteilung der verschiedenen Methoden erleichtert.

Informationsumgebungen

Befragungen und Experimente können als Exploration von Informationsumgebungen verstanden werden. In einer Befragung ist die Informationsumgebung die Liste von Fragen, die ein Teilnehmer bearbeitet. In einem Experiment, wie auch in einer Studie, in der Prozessdaten gesammelt werden, sind es die Aufgaben. Vor allem bei der Prozessdatenanalyse variieren solche Umgebungen in Strukturierungsgrad und Datenvolumen oft stark. Strukturierungsgrad meint in diesem Zusammenhang das Ausmaß der Informationsaufbereitung, die ein Forscher vor einem Experiment durchführen muss. In einer Studie mit Informationsta-

⁸ Vgl. Lau & Redlwasck 2006.

⁹ Fixationen: das Verweilen des Auges auf einer fixen Position des Stimulus für mindestens 200-300 Millisekunden. Während einer Fixation wird die meiste Information aufgenommen und verarbeitet (vgl. Rayner 1998), Sakkaden: willkürliche, schnelle Augenbewegungen. Pupillendilatation: die Größe der Pupille.

¹⁰ Wir stellen weiter unten das Programm „Flashlight“ vor, das die Aufnahme von Augenbewegungen in einer webbasierten Version nachbildet.

¹¹ Wir stellen weiter unten das Programm „WebDiP“ vor, das die online Informationssuche eines Nutzers aufzeichnet.

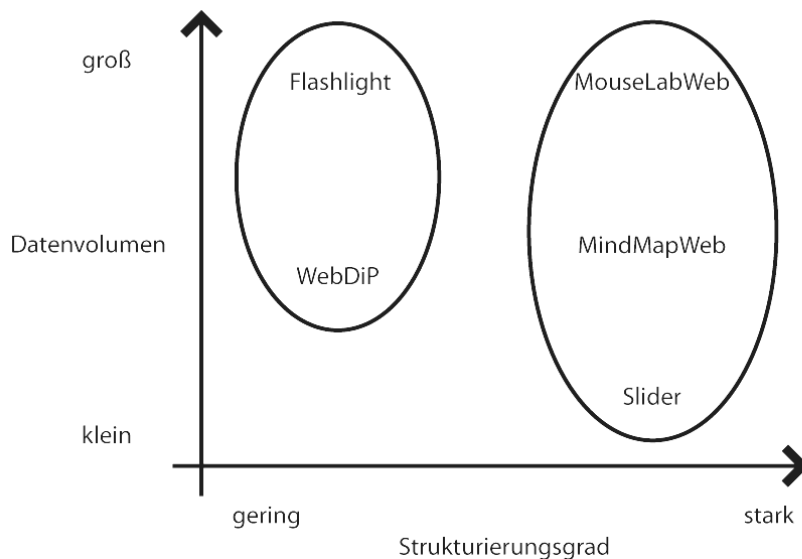
¹² Vgl. Ericsson 2006.

¹³ Vgl. Schulte-Mecklenbeck, Kühberger & Ranyard 2008.

fehlen müssen die vorhandenen Informationen zuerst in Alternativen und Attribute eingeteilt werden – es handelt sich also um eine stark strukturierte Informationsumgebung. In einer Studie, die Blickbewegungsdaten sammelt, können z.B. Bilder von Autos ohne strukturelle Veränderung verwendet werden.

Wir teilen die fünf später vorgestellten Online-Methoden (MouselabWeb, MindMapWeb, Slider, Flashlight und WebDiP) anhand ihrer Struktur ein, wobei die Menge der gesammelten Information die abhängige Variable darstellt (siehe Abbildung 1). Zwei Gruppen von Methoden lassen sich unterscheiden: In der *stark strukturierten Gruppe* benötigt MouseLabWeb klar strukturierte Stimuli. Der Forscher bietet diese meist in Form einer Informationsmatritze dar. MouseLabWeb hat gleichzeitig das höchste Datenvolumen. MindMapWeb ermöglicht die Konstruktion einer Mind Map zu verfolgen. Der Forscher stellt die zu organisierenden Informationen einzeln zur Verfügung und der Nutzer kann diese ordnen. Slider gibt Auskunft über Feinadjustierungen auf bi-polaren Skalen. Hier gibt der Forscher die Skalierung mit den Endpunkten vor, der Nutzer justiert die Skala dann nach Belieben. In der *zweiten Gruppe* finden sich Methoden, wie WebDip oder Flashlight, die *gering strukturierte Stimuli* verwenden. Flashlight hilft, den Suchpfad auf visuellen Stimuli zu registrieren und bildet somit die Datengewinnung einer Augenbewegungskamera nach. Die Stimuli müssen statisch und auf einem Bildschirm abbildbar sein. WebDiP liefert die Sequenz und Häufigkeit von Stichwörtern in einer Informationssuchaufgabe. Der Forscher bereitet hier eine Datenbank mit Informationen vor, die der Nutzer dann selbständig durchsuchen kann.

Abbildung 1: Datenvolumen und Strukturierungsgrad von Informationsumgebungen



Quelle: Eigene Darstellung

Man kann unschwer erkennen, dass die eben genannten Methoden die Bearbeitung verschiedenster Fragestellungen zulassen. Die Selektion der richtigen Methode zu einer Fragestellung stellt dabei einen entscheidenden Schritt in der Planung einer Befragung oder eines Experimentes dar. Passen Frage und Methode nicht zusammen, werden oft zu viele Daten auf einem zu detaillierten Level erhoben – dies führt zu Problemen, die Ergebnisse aus der Datenflut herauszulösen. Im Gegensatz dazu führt das andere Extrem (zu wenige Daten auf einem zu oberflächlichen Level) dazu, dass Effekte unentdeckt bleiben. Wir werden den Punkt der Methodenselektion nochmals in der Diskussion aufgreifen, zuvor aber mehr Details zu den einzelnen Methoden.

Stark strukturierte Informationsumgebungen

Der Forscher bereitet die vorhandenen Informationen in dieser Gruppe von Methoden stark auf und stellt sie gezielt zur Verfügung. Das wohl älteste Tool zur Gewinnung von stark strukturierten Prozessdaten ist Mouselab.

Mouselab(Web)

Die Grundidee der Informationstafeln wurde in einer computerisierten Version erstmals von Payne, Bettman und Johnson¹⁴ verwendet und erhielt den Namen *Mouselab*. Die Namensgebung rührt daher, dass Mouselab einen Monitor und eine Maus für die Informationspräsentation bzw. -eingabe verwendet. Die Bewegung des Mauscursors in einen der Zellenbereiche (z.B. die dritte Zelle in der ersten Reihe in Abbildung 2) zeigt den vorher verdeckten Inhalt an. Nachdem die Maus aus dem Zellenbereich herausbewegt wurde, wird der Inhalt wieder verdeckt. Im Jahr 2004 wurde Mouselab für den Online-Einsatz neu programmiert und steht seither in verschiedenen Versionen zur Verfügung. Die wohl am meisten genutzten Versionen basieren entweder auf den Programmiersprachen Javascript¹⁵ oder Java.¹⁶ Mouselab(Web) ist ideal, um Informationen, die sich in Matrix-Form (2x2, 2x3, ...) organisieren lassen, darzustellen und zu untersuchen (siehe Abbildung 2 für eine 2x4 Matrix eines Spiels mit 2 Optionen). Es besteht natürlich auch die Möglichkeit, von diesem Standardformat abzuweichen.¹⁷ Folgende Daten werden in einem Mouselab(Web)-Experiment aufgezeichnet: Anzahl und Länge von Zellenöffnungen, Sequenz der Zellenöffnungen und gegebenenfalls die Entscheidung. Die Anzahl und Länge der Zellenöffnungen sind hoch korreliert, darum werden diese beiden Maße oft kombiniert und unter dem Begriff *Aufmerksamkeit* gegenüber einer Zelle zusammengefasst. Die Aufmerksamkeit ist umso größer, je größer die beiden Einzelkomponenten werden, und lässt sich somit grafisch als Fläche in sogenannten *attention plots*¹⁸ abbilden.

¹⁴ Vgl. Payne, Bettman & Johnson 1993.

¹⁵ Vgl. Willemsen & Johnson 2004: <http://www.mouselabweb.org/>.

¹⁶ Vgl. Payne 2005: <https://www.fuquaworld.duke.edu/blab/MouseLabExperimenter.jsp>.

¹⁷ Andere Formate werden z.B. in Willemsen, Böckenholt & Johnson (2006) oder Schulte-Mecklenbeck & Hirsch (2008) verwendet.

¹⁸ Ein Beispiel für solche Plots findet sich in Johnson, Schulte-Mecklenbeck & Willemsen (2008: 267). Die Größe der einzelnen Rechtecke korrespondiert mit der Aufmerksamkeit gegenüber der jeweiligen Zelle, die Länge der Pfeile zwischen den Zellen bildet die mittlere Anzahl der Übergänge zwischen zwei Zellen ab.

Abbildung 2: MouselabWeb-Matrix für ein Spiel mit 2 Optionen

	Amount to Win	Probability of that amount	Amount to Win	Probability of that amount
Gamble A:	<input type="text" value="Wa1"/>	<input type="text" value="Pa1"/>	<input type="text" value="\$4000"/>	<input type="text" value="Pa2"/>
Gamble B:	<input type="text" value="Wb1"/>	<input type="text" value="Pb1"/>	<input type="text" value="Wb2"/>	<input type="text" value="Pb2"/>

Quelle: Eigene Darstellung

In diesen Plots findet sich auch die aufsummierte und gemittelte Sequenz der Zellöffnungen wieder. Die gemittelten Abfolgen der Zellenöffnungen werden als Pfeile zwischen den Zellen dargestellt. Es ergibt sich somit ein Eindruck, gegenüber welchen Zellen die Aufmerksamkeit größer war und welche Reihenfolgen der Zellenöffnung am häufigsten waren.

MindMapWeb

Als zweites Beispiel für eine stärker strukturierte Methode soll *MindMapWeb*¹⁹ vorgestellt werden, das die mittlerweile klassische Idee einer Mind Map²⁰ in einer Online-Version nutzt. In diesem Programm, das unter Zuhilfenahme von dynamischem HTML²¹ erstellt wurde, kann der Forscher sogenannte *Knoten*²² anlegen, die im eigentlichen Experiment vom Nutzer angeordnet werden. Im Beispiel in Abbildung 4 lautet die Aufgabe, Risikofaktoren der globalen Erwärmung in Relation zueinander zu setzen und diese Relationen dann zu bewerten. Der Nutzer kann dazu die vorgegebenen Knoten verschieben, uni- oder bidirektionale Pfeile (Relationen) zwischen den einzelnen Knoten setzen und dann diese Relationen einzeln, z. B. auf einer Skala von 1-100, gewichten. Das Programm zeichnet die Reihenfolge auf, in der Knoten bewegt werden, die Ausgangs- sowie Endposition und die Zeit, die der Nutzer für eine Aufgabe braucht. Weiterhin werden die Relationen zwischen den Knoten und deren Gewichte gespeichert. Dadurch ist es möglich, z.B. mittels Netzwerkanalyse, ein genaues Bild der strukturellen Zusammenhänge und der Repräsentation einer Aufgabe zu erhalten.

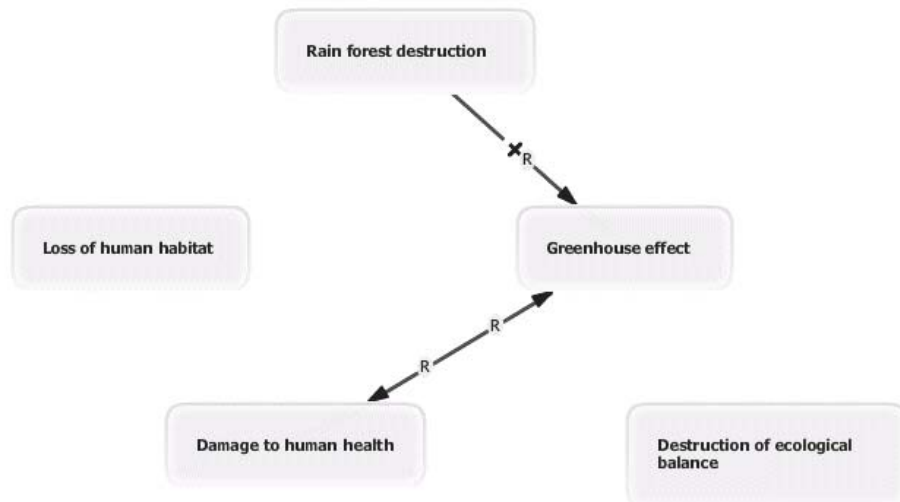
¹⁹ Vgl. Böhm, Pfister & Schulte-Mecklenbeck 2008.

²⁰ Vgl. Collins & Quillian 1969.

²¹ Vgl. http://de.wikipedia.org/wiki/Dynamisches_HTML.

²² Engl.: „Nodes“.

Abbildung 3: Verschiedene Knoten und Relationen in MindMapWeb



Quelle: Eigene Darstellung

MindMapWeb stellt ein interessantes Tool für die Erforschung von Inhalten dar, bei denen der Nutzer zwar Anhaltspunkte über die relevanten Informationen erhält, dann allerdings die Relationen zwischen den Informationen selbst definieren muss.

Slider

In der stark strukturierten Gruppe liefert *Slider* die geringste Datenmenge bei einem Strukturierungsgrad, der dem klassischen Fragebogen nahe kommt. Die gesammelte Datenmenge geht allerdings über eine simple Antwort auf einer Ja/Nein-Skala hinaus. Slider registriert, unter Zuhilfenahme von Javascript, Cursorbewegungen auf einer bi-polaren Skala sowie die Zeit, die ein Nutzer mit einer Aufgabe verbringt. Dies kann in mehrerer Hinsicht nützlich sein: In einem Fragebogen können Feinjustierungen des Nutzers analysiert werden, da jede Cursorposition gespeichert wird. In einem Experiment kann Slider z.B. als Eingabetool verwendet werden, um die Verteilung von verschiedenen Geldbeträgen in einer Entscheidungssituation zu testen (siehe Abbildung 4). Durch Verschieben des Cursors wird die Zuteilung der Geldbeträge automatisch aktualisiert und gespeichert.

Abbildung 4: Verteilungsspiel in Slider

Instructions

On the right you see a slider. You can change the slider to adjust the amount of money you and the other person will receive.

The numbers at the ends of the slider show the range of possible distributions for you and the other person.

Once you have moved the slider to the distribution you most prefer press the Submit button.

Slider 1 of 6

85 \$ for you 100 \$ for you

85 \$ for other 50 \$ for other

You will receive: 93 Dollars

The other will receive: 66 Dollars

Submit

Quelle: Eigene Darstellung

Der Nutzer kann die Verteilung der Geldbeträge nach verschiedenen Gesichtspunkten wie Fairness, Altruismus oder Egoismus durchführen. Fair wäre in diesem Zusammenhang z.B. eine gleichmäßige Aufteilung der Geldbeträge (weiter links auf der Skala in Abbildung 4), altruistisch wäre das Bevorzugen des Partners, egoistisch wäre das Bevorzugen der eigenen Person (weiter rechts auf der Skala in Abbildung 4). Dem Forscher ermöglichen die Prozessdaten (Cursorbewegungen, Verweilzeiten) einen genaueren Einblick in die Strategien von verschiedenen Nutzern und somit eine Kategorisierungsmöglichkeit.

Geringstrukturierte Informationsumgebungen

Geringstrukturierte Informationsumgebungen stellen flexiblere Möglichkeiten der Stimulusgestaltung zur Verfügung. Es kann hier leicht vom einfachen Matrix-Format abgewichen werden. Ein Vorteil der folgenden Methoden ist somit, dass die Vorstrukturierung der Informationsumgebungen durch den Forscher wegfällt. Dies ist vor allem bei Fragestellungen wichtig, in denen nicht das Verwenden von vorhandenen Informationen, sondern die Exploration der Informationsumgebung im Mittelpunkt steht.

Flashlight

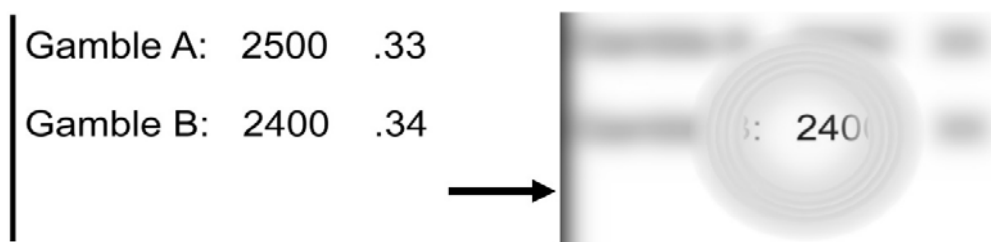
*Flashlight*²³ bietet von den hier vorgestellten Methoden die größte Freiheit in der Gestaltung von Stimuli. Jegliche Art von statischen Stimuli, die in Bildform dargeboten werden können, lassen sich verwenden. Als Beispiel soll hier die Wahl zwischen zwei einfachen Spielen dienen. Diese Art von Stimuli ist in der Entscheidungsforschung²⁴ beliebt, da sie eine kontextunabhängige, relativ reine Form einer Entscheidung darstellen. In Abbildung 5 (linke Hälfte) wird die Entscheidung zwischen zwei Spielen A und B dargestellt. In Spiel A

²³ Vgl. Schulte-Mecklenbeck, Murphy & Hutzler 2008.

²⁴ Vgl. Slovic & Lichtenstein 1968; Russo & Doshier 1983.

erhält man 2500 Punkte mit Wahrscheinlichkeit .33 und in Spiel B erhält man 2400 Punkte mit Wahrscheinlichkeit .34. Der Stimulus (siehe Abbildung 5, linke Hälfte) wird durch einen Weichzeichner unkenntlich gemacht (siehe Abbildung 5, rechte Hälfte). Nur ein runder Ausschnitt um den Mauscursor zeigt den Stimulus unverfälscht an. Durch Bewegen des Mausursors kann nun das Spiel vom Nutzer erkundet werden. Diese Bewegung bildet das Suchmuster der Augen, das normalerweise mit einer Augenbewegungskamera registriert wird, gut nach.

Abbildung 5: Spiel-Stimulus in Flashlight



Quelle: Eigene Darstellung

Die Bewegungen der Maus werden zehn Mal pro Sekunde (also mit einer Auflösung von 10 Hz) registriert. Dies ist zwar im Vergleich zu gängigen Augenbewegungskamera-Modellen eine relativ geringe Auflösung²⁵, für viele Fragestellungen ergibt sich allerdings eine ausreichende Datendichte. Die gesammelten Daten geben Auskunft über Fixationen (Anzahl und Länge) sowie die Sequenz der Informationsaufnahme. Verschiedene Stimuli-Arten werden gegenwärtig (2008) getestet: Dazu gehören Spiele (wie schon für Mouse-labWeb beschrieben), visuelle Suchaufgaben (auf einem Bild soll ein Ziel [roter Punkt] gesucht werden) und einfache arithmetische Aufgaben. Neben der schon erwähnten Einschränkung der Auflösung soll noch auf eine weitere Einschränkung der Methode hingewiesen werden: Es ist nur möglich, statische Stimuli zu verwenden, dynamische Stimuli wie Videos können nicht verarbeitet werden.

Web Decision Processes – WebDiP

In vielen Entscheidungssituationen suchen Menschen zuerst nach Informationen.²⁶ Diese Suche dient einerseits der Orientierung und andererseits dem Wunsch, eine gute Entscheidung zu treffen. *WebDiP*²⁷ registriert diesen Suchprozess in einem vorgefertigten Informationsraum. Der Forscher stellt im ersten Schritt Informationen, die eine Entscheidungssituation beschreiben, in einem Frage-Antwort-Format zusammen. Ein Beispiel für eine Entscheidung ist die Wahl zwischen zwei Impfungen. Zuerst wird man versuchen, mehr über

²⁵ Viele Modelle haben eine Mindestauflösung von 50-60 Hz, Highend-Modelle erreichen eine Maximalauflösung von über 1000 Hz.

²⁶ Die Suchmaschine Google (<http://www.google.com>) wird heute häufig dafür verwendet.

²⁷ Vgl. Schulte-Mecklenbeck & Neun 2005.

die Impfungen herauszufinden. Auf die Frage: „Gibt es eine Impfung gegen die Krankheit“ würde die Antwort z.B. lauten: „Es sind zwei Impfstoffe vorhanden.“ Dann kann man gezielt nach Informationen über diese beiden Impfstoffe suchen. Eine Vielzahl solcher Frage-Antwort-Paare wird vom Forscher in der WebDiP-Datenbank gespeichert. Der Nutzer bearbeitet das Problem, indem er einen Suchbegriff in eine Maske eingibt (siehe Abbildung 6). Der Suchbegriff, z.B. „Krankheit“, wird dann von WebDiP an die Datenbank weitergegeben und der Nutzer erhält eine Liste mit Fragen, die zum Stichwort „Krankheit“ gefunden wurden. Er klickt dann auf eine für ihn relevante Frage und erhält die Antwort angezeigt. Wichtig bei diesem Beispiel ist, dass der Forscher zwar die Informationen in der Datenbank vorgibt, der Nutzer aber frei in diesen Informationen navigiert. Dadurch erhält der Forscher Einblick in die Repräsentation der Informationsumgebung des Nutzers, ähnlich wie das auch bei MindMapWeb beschrieben wurde. Die verwendeten Suchwörter, die im zweiten Schritt geöffneten Fragen sowie die Zeit, die für eine Aufgabe verwendet wurde, werden gespeichert.

Abbildung 6: WebDiP Eingabemaske

Quelle: Schulte-Mecklenbeck & Neun 2005

Im nächsten Schritt möchten wir näher auf den Vergleich von verschiedenen Forschungsmethoden eingehen, um genaueres über deren Validität²⁸ herausfinden zu können.

Vergleich der verschiedenen Methoden

Wenn neue Forschungsmethoden vorgestellt werden, ist der Vergleich mit etablierten Methoden wichtig. Für diese Kapitel ist dabei vor allem der Vergleich zwischen im Labor und online erhobenen Daten wichtig. Die Erhebung von Daten über das Internet löste Ende der 1990er Jahre eine breite Diskussion aus, ob die online erhobenen Daten mit den klassisch in Labor, Hörsaal oder Feld gesammelten Daten vergleichbar sind.²⁹ Hier kamen viele Impul-

²⁸ Vgl. Huber 2005.

²⁹ Ein Überblick zu dieser Diskussion wird von Batinic 1999 gegeben.

se aus der eher kommerziellen Richtung (z.B. der Marktforschung) und der Fokus wurde zuerst vor allem auf Fragebogenstudien und E-Mail-Erhebungen gerichtet. Erst mit verbesserten Möglichkeiten der Online-Programmierung wurden Experimente interessanter und in die Diskussion aufgenommen.³⁰ Verschiedene Ansätze sind für den Vergleich von bestehenden Methoden und neuen Zugängen möglich. Wir wollen sowohl stark als auch gering strukturierte Methoden betrachten und anhand von zwei Beispielen Unterschiede und Gemeinsamkeiten diskutieren.

Ein Beispiel für eine Studie zu stark strukturierten Erhebungsverfahren stellt die Untersuchung von Lohse und Johnson dar³¹: Sie verglichen die Informationsaufnahme durch Mouselab mit Daten, die mit einer Augenbewegungskamera erhoben wurden. Der Vergleich zwischen den Verfahren basiert auf verschiedenen abhängigen Variablen, wie die Bearbeitungszeit einer Aufgabe, die Anzahl der Fixationen und dem Prozentsatz der betrachteten Information.³² Die Bearbeitungszeit der Aufgaben lag bei Mouselab um 67 Prozent höher als bei der Analyse der Augenbewegungen. Dieser Unterschied lässt sich vor allem durch den Umstand erklären, dass für Mausbewegungen immer Augen- und Handbewegung notwendig sind, während bei der Augenbewegungskamera Augenbewegungen ausreichen. Ein ähnliches Bild findet sich bei der Anzahl der Fixationen: Deutlich mehr Fixationen (im Durchschnitt 120) werden in der reinen Augenbewegungs-Bedingung gefunden. Im Gegensatz dazu kommt es in der Mouselab-Bedingung nur zu durchschnittlich 66 Fixationen. Haben die unterschiedlichen Bearbeitungszeiten und Fixationen nun eine Auswirkung auf die betrachteten Informationen? Der Prozentsatz der betrachteten Informationen war zwischen Mouselab und der Augenbewegungskamera praktisch der gleiche (90 Prozent versus 93 Prozent). Obwohl also die Zeit für eine Aufgabe unterschiedlich war, wurden trotzdem in beiden Methoden beinahe alle Informationen betrachtet.

Im Gegensatz zum ersten Beispiel standen im zweiten Beispiel, einer Studie von Schulte-Mecklenbeck und Huber, gering strukturierte Verfahren im Mittelpunkt der Analyse³³: Die Autoren boten den Probanden zweier Experimente die gleichen Entscheidungsaufgaben in WebDiP (siehe oben) einmal im Labor und einmal online dar. Das erste Experiment verwendete eine Zufallsstichprobe in der Online-Version und eine Studentenstichprobe in der Laborversion. Zwei Ergebnisse sollen hervorgehoben werden: In der Online-Version haben 36 Prozent der Teilnehmer das Experiment nicht beendet, in der Laborversion hingegen wurde das Experiment von allen Teilnehmern bis zum Ende bearbeitet. Was könnte diesen Unterschied erklären? Die Nutzer in einem Online-Experiment haben immer eine sehr einfache Möglichkeit, aus dem Experiment auszusteigen – sie schließen ihren Browser. In einem Laborexperiment kommt es hingegen sehr selten vor, dass ein Teilnehmer einfach aufsteht und den Raum verlässt. Man begegnet also sehr unterschiedlichen Dropout-Raten, wenn man Online-Experimente mit Laborexperimenten vergleicht. Ein zweites Ergebnis dieses Experimentes ist der Unterschied in der Menge der verwendeten Informationen zwischen dem Labor und dem Online-Experiment. In der Online-Stichprobe wurden weniger Informationen kürzer betrachtet, in der Labor-Stichprobe war sowohl der Umfang als auch die Länge der Betrachtungszeiten größer. Da dieser Effekt auch auf die beiden unterschied-

³⁰ Vor- und Nachteile von Online-Forschung sowie wichtige Punkte bei Online-Experimenten werden z.B. von Reips 2002 diskutiert.

³¹ Vgl. Lohse & Johnson 1996.

³² Weitere abhängige Variablen werden in Lohse & Johnson (1996: 31-33) detailliert erläutert.

³³ Vgl. Schulte-Mecklenbeck & Huber 2003.

lichen Stichproben zurückgeführt werden kann³⁴, wurde in einem zweiten Experiment aus der gleichen Population (Studenten) rekrutiert und eine Zufallszuteilung zur Online- und Laborversion vorgenommen. Die Ergebnisse aus dem ersten Experiment wurden hier repliziert. Trotz der beschriebenen Unterschiede war die Struktur der betrachteten Information³⁵ in beiden Experimenten vergleichbar. Dies deutet, wie auch schon in der Studie von Lohse und Johnson³⁶, darauf hin, dass zwar Unterschiede im Umfang der aufgenommenen Information zwischen Methoden und Orten der Datenerhebung bestehen, allerdings strukturell durchaus vergleichbare Ergebnisse gefunden werden.

Diskussion

Wir haben in diesem Beitrag Methoden zur Online-Prozessdatenerhebung vorgestellt. Die Auflistung erhebt nicht den Anspruch vollständig zu sein. Vielmehr haben wir Methoden ausgesucht, die frei verfügbar und leicht anzuwenden sind. Zwei Dimensionen wurden zur Einteilung der verschiedenen Methoden gewählt: Die Menge der resultierenden Daten und den Strukturierungsgrad der Stimuli.

Da die Wahl einer passenden Methode einen wichtigen Punkt in einem Forschungsprojekt darstellt, soll sie im Folgenden genauer diskutiert werden. Bei der Auswahl einer Methode in Bezug auf die Forschungsfrage gilt die Devise, dass man besser „nicht mit Kanonen auf Spatzen schießen“ sollte. Kann man die Forschungsfrage relativ genau eingrenzen, ist es sinnvoll, strukturierte Verfahren zu verwenden, da die resultierenden Daten klare Aussagen über die Verwendung einzelner Informationseinheiten zulassen. MouseLabWeb, MindMapWeb oder Slider liefern genaue Betrachtungszeiten und Navigationsmuster für ein breites Spektrum von Stimuli. Je weniger strukturiert eine Informationsumgebung ist, je mehr man also an der Frage interessiert ist, wie sich ein Nutzer durch die Informationen bewegt, desto mehr sollte man sich in den Bereich der wenig strukturierten Methoden begeben. WebDiP oder Flashlight helfen hier, ein genaueres Bild der Suchmuster zu erhalten. Der recht simple Ratschlag, eine passende Methode zu einer Fragestellung zu wählen, wird oft ignoriert. Dies kann bei Prozessmethoden gerade wegen des großen Datenvolumens fatale Folgen bei der Datenauswertung mit sich bringen. Es ist viel schwieriger, aus einer großen Datenmenge die relevanten Daten herauszufiltern, als in einer kleinen Datenmenge eine einfache Zuordnung der vorhandenen Daten zu den Forschungsfragen zu treffen.

Weiterhin sollte auf das Zusammenspiel von Online- und Labor-Methoden eingegangen werden. Diskutiert man Online- versus Offline-Methoden, trifft man oft auf die Aussage, dass die Leichtigkeit, mit der Daten online erhoben werden könnten, mit – im Vergleich zum Labor – geringer Kontrolle bezahlt werden muss. Natürlich liegt darin etwas Wahres. Der Forscher legt im Labor Wert darauf, dass der Nutzer ungestört ist, dass das Experiment verstanden wird, dass ohne Unterbrechung gearbeitet wird. Man kann diese Punkte online zwar nur schwer beeinflussen, allerdings nach der Datenerhebung durchaus, z.B. durch Analyse von Aufgabenzeiten, erschließen. Etwas schwieriger wird es, den Status, den ein Forscher im Labor hat, online nachzubilden. Dieser Status garantiert, dass Nutzer Experimente in einem Labor zu einem hohen Prozentsatz abschließen. Fällt die Anwesenheit des

³⁴ Oder aber auch auf die mangelnde externe Validität von Laborexperimenten. Auf diese soll allerdings in diesem Zusammenhang nicht eingegangen werden.

³⁵ Also wie die Probanden die Aufgaben exploriert haben.

³⁶ Vgl. Lohse & Johnson 1996.

Forschers weg, verändert sich die Dropout-Rate stark. Unser Vorschlag ist, aus der Not eine Tugend zu machen und die Vorteile der jeweiligen Forschungsorte zu nutzen und zu kombinieren. Ein Beispiel wäre eine Augenbewegungs-Laborstudie mit einer Online-Flashlight-Studie zu kombinieren. Im ersten Schritt werden Daten im Labor erhoben (hohe Kontrolle, kein Dropout). Dies ist aufwändig, führt allerdings zu exzellenter Datenqualität in Bezug auf die Auflösung. Im zweiten Schritt wird das gleiche Experiment mit Flashlight online durchgeführt (geringere Kontrolle, höherer Dropout). Die geringere Auflösung in Flashlight (siehe oben) wird hier mit einer viel einfacheren Datenerhebung wettgemacht, es ist zeitlich egal ob man 10 oder 100 Nutzer in ein Online-Experiment aufnimmt. Durch die größere Anzahl an Nutzern erhöht sich allerdings die statistische Aussagekraft einer Untersuchung und auch kleine Effekte, die eventuell im Labor nicht gezeigt werden können, werden leichter nachgewiesen.

Die Betrachtung des Informationsaufnahme-Prozesses in einem Experiment oder in einer Befragung gibt Auskunft über Vorgänge, die sonst verborgen bleiben würden. Diese zweite Ebene der Datenerhebung kann in vieler Hinsicht hilfreich sein, eine Fragestellung detaillierter zu beantworten. Vor allem die Online-Datenerhebung stellt einen einfachen Weg dar, um schnell an große Datenmengen zu gelangen, die dann für sich stehend oder in Kombination mit Labordaten analysiert werden können.

Literatur

- Batinic, B., Werner, A., Gräf, L. & Bandilla, W. (1999): Online Research. Göttingen.
- Böhm, G., Pfister, H. R. & Schulte-Mecklenbeck, M. (2008): Mapping the Mind in an Environmental Decision Task. Manuskript in Vorbereitung.
- Collins, A. & Quillian, R. (1969): Retrieval Time from Semantic Memory. In: *Journal of Verbal Learning and Verbal Behavior*, 8, S. 240-247.
- Dillman, D. A. & Bowker, D. K. (2001): The Web Questionnaire Challenge to Survey Methodologies. In: Reips, U.-D. & Bosnjak, M. (Hrsg.): *Dimensions of Internet Science*. Lengerich, S. 159- 177.
- Ericsson, K. A. (2008): Protocol Analysis and Expert Thought: Concurrent Verbalizations of Thinking During Experts' Performance on Representative Tasks. In: Ericsson, K. A., Charness, N., Feltovich, P. J. & Hoffman, R. R. (Hrsg.): *The Cambridge Handbook of Expertise and Expert Performance*. Oxford, S. 223-241.
- Ericsson, K. A. & Simon, H. A. (1980): Verbal Reports as Data. In: *Psychological Review*, 87, S. 215-251.
- Huber, O. (2005): *Das psychologische Experiment* (3. Auflage). Bern.
- Huber, O., Wider, R. & Huber, O. W. (1997): Active Information Search and Complete Information Presentation in Naturalistic Risky Decision Tasks. In: *Acta Psychologica*, 95, S. 15-29.
- Johnson, E. J., Schulte-Mecklenbeck, M. & Willemsen, M. C. (2008): Process Models Deserve Process Data: Comment on Brandstätter, Gigerenzer & Hertwig (2006). In: *Psychological Review*, 115/1, S. 263-272.
- Lau, R.R. & Redlwask, D. (2006): *How Voters Decide: Information Processing during an Election Campaign*. New York: Cambridge University Press.
- Li, H., Berrens, R. P., Bohara, A. K., Jenkins-Smith, H. C., Silva, C. L. & Weimer, D. L. (2004): Telephone versus Internet Samples for a National Advisory Referendum: Are the Underlying Stated Preferences the Same? In: *Source Applied Economics Letters*, 11/3, S. 173-176.
- Lohse, G. & Johnson, E. J. (1996): A Comparison of Two Process Tracing Methods for Choice Tasks. In: *Organizational Behavior and Human Decision Processes*, 68/1, S. 28-43.
- Payne, J. W. (1976): Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis. In: *Organizational Behavior and Human Performance*, 16, S. 366-387.
- Payne, J. W. (2005): Mouselab Experimenter. Available from <https://www.fuquaworld.duke.edu/blas/MouseLabExperimenter.jsp>.
- Payne, J. W., Bettman, J. R. & Johnson, E. J. (1993): *The Adaptive Decision Maker*. New York.
- Pötschke, M. (2004): Schriftlich oder online? Methodische Erfahrungen aus einer vergleichenden Mitarbeiterbefragung. Vortrag auf der German Online Research, 2004, Essen.

- Rayner, K. (1998): Eye Movements in Reading and Information Processing: 20 Years of Research. In: Psychological Bulletin, 124/3, S. 372-422.
- Reips, U.-D. (2002): Standards for Internet-Based Experimenting. In: Experimental Psychology, 49/4, S. 243-256.
- Russo, J. E. & Hunt, H. K. (1978): Eye Fixations Can Save the World: A Critical Evaluation and Comparison with Other Information Processing Methodologies. In: Advances in Consumer Research, 5, S. 561-570.
- Russo, J. E. & Doshier, B. A. (1983): Strategies for Multiattribute Binary Choice. In: Journal of Experimental Psychology: Learning, Memory, and Cognition, 9, S. 676-696.
- Slovic, P. & Lichtenstein, S. (1968): Relative Importance of Probabilities and Payoffs in Risk Taking. In: Journal of Experimental Psychology Monograph, 78/3, S. 1-18.
- Schulte-Mecklenbeck, M., Kühberger, A. & Ranyard, R. (2008): The Handbook of Process Tracing Methods in Decision Making. Manuskript in Vorbereitung.
- Schulte-Mecklenbeck, M., Murphy, R. O. & Hutzler, F. (2008): Flashlight – An Online Eye-Tracking Addition. Manuskript in Vorbereitung.
- Schulte-Mecklenbeck, M. & Hirsch, B. (2008): The Influence of Different Presentation Formats on Controlling Decisions of Managers. Manuskript in Vorbereitung.
- Schulte-Mecklenbeck, M. & Neun, M. (2005): WebDiP – A Tool for Information Search Experiments on the World-Wide-Web. In: Behavior Research Methods, 37/2, S. 293-300.
- Schulte-Mecklenbeck, M. & Huber, O. (2003): With or Without the Experimenter: Information Search in the Laboratory and on the Web. In: Behavior Research Methods, Instruments & Computers, 35/2, S. 227-235.
- Skitka, L. J. & Sargis, E. G. (2006): The Internet as Psychological Laboratory. In: Annual Review of Psychology, 57, S. 529-555.
- Willemsen, M. C., Böckenholt, U. & Johnson, E. J. (2006): How Losses Affect Choice: Loss Encoding, Gain Enhancement and a Comparison of Three Theories. Unveröffentlichtes Manuskript.
- Willemsen, M. C. & Johnson, E. J. (2005): MouselabWEB: Performing Sophisticated Process Tracing Experiments in the Participants Home! Vortrag: Society for Computers in Psychology, 2005, Minneapolis.