

Wolfgang Neubarth

Präferenzdaten online

Herbert von Halem Verlag

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

Wolfgang Neubarth
Präferenzdaten online
Köln : Halem, 2008

Wolfgang Neubarth promovierte an der Friedrich-Alexander-Universität Erlangen-Nürnberg. Während dieser Zeit arbeitete er als wissenschaftlicher Mitarbeiter bei GESIS-ZUMA im Bereich Wissensvermittlung und Beratung. Dort war er für Grundlagenforschung und Consulting zum Thema Online-Umfragen zuständig. Heute unterstützt er verschiedene Bereiche von TNS, insbesondere TNS Infratest InCom (Technology Sector), TNS Infratest MediaResearch und die TNS Convergence Group. Er ist dort schwerpunktmäßig für das Themenfeld »Interactive« (Online-Forschung) im Bereich Business Development verantwortlich.

Alle Rechte, insbesondere das Recht der Vervielfältigung und Verbreitung sowie der Übersetzung, vorbehalten.
Kein Teil des Werkes darf in irgendeiner Form (durch Fotokopie, Mikrofilm oder ein anderes Verfahren) ohne schriftliche Genehmigung des Verlages reproduziert oder unter Verwendung elektronischer Systeme (inkl. Online-Netzwerken) gespeichert, verarbeitet, vervielfältigt oder verbreitet werden.

© 2008 by Herbert von Halem Verlag, Köln

ISBN 978-3-938258-69-9

Den Herbert von Halem Verlag erreichen Sie auch im Internet unter <http://www.halem-verlag.de>
E-Mail: info@halem-verlag.de

SATZ: Herbert von Halem Verlag
DRUCK: FINIDR, S.R.O. (Tschechische Republik)
GESTALTUNG: Claudia Ott Grafischer Entwurf, Düsseldorf
Copyright Lexicon ©1992 by The Enschedé Font Foundry.
Lexicon® is a Registered Trademark of The Enschedé Font Foundry.

Inaugural-Dissertation zur Erlangung des akademischen Grades
eines Doktors der Wirtschafts- und Sozialwissenschaften (Dr. rer. pol.)
der Friedrich-Alexander-Universität Erlangen-Nürnberg.

Erstreferent: Prof. Dr. Johann Bacher

Zweitreferent: Prof. Dr. Klaus Moser

Promotionstermin: 17.7.2007

Inhalt

Abbildungsverzeichnis	7
Tabellenverzeichnis	9
1. Herausforderung der Präferenzmessung online	13
2. Präferenzdatentheorie	18
2.1 Grundlagen	18
2.2 Definitionen	21
2.3 Formalisierung	24
2.4 Vor- und Nachteile der ipsativen Messung	33
2.5 Adäquate Anwendungen	42
2.6 Zusammenfassung und Diskussion	44
3. Präferenzdatenerhebung online	47
3.1 Auswahl der Beurteilungsobjekte	47
3.2 Exkurs: Das Internet und seine Geschichte	51
3.3 Besonderheiten online	53
3.4 Technische Möglichkeiten	59
3.5 Methodische Möglichkeiten	67
3.6 Erhebungsdesigns	96
3.7 Zusammenfassung und Diskussion	108
4. Präferenzdatenanalyse	110
4.1 Datenorganisation als Vorstufe der Analyse	112
4.2 Deskriptive Maße	123
4.3 Korrelations- und Distanzmaße	128
4.4 Testen von Unterschieden	139

4.5	Allgemeines lineares Modell	141
4.6	Skalierungsverfahren	154
4.7	Cluster-Analyse	177
4.8	Dekompositionelle Verfahren	179
4.9	Weitere Methoden	185
4.10	Zusammenfassung und Diskussion	189
5.	Präferenzdatenanalyse online	192
5.1	Selbsterstellte Auswertungen	192
5.2	Online-Auswertungen mit SPSS	193
5.3	Online-Auswertungen mit R	193
5.4	Zusammenfassung und Diskussion	194
6.	Präferenzdaten in empirischen Online-Studien	196
6.1	Studie I – Machbarkeitsstudie	198
6.2	Studie II – Kognitiver Pretest	205
6.3	Studie III – Experimentelle Vergleichsstudie	215
6.4	Studie IV – Paarvergleiche	232
6.5	Studie V – Allbus online	242
6.6	Zusammenfassung und Diskussion	252
7.	Zusammenfassung und Ausblick	255
	Literaturverzeichnis	259

Denken heißt vergleichen.
(Walter Rathenau)

1. Herausforderung der Präferenzmessung online

In dieser Arbeit werden die methodischen Aspekte der Erhebung und Auswertung von Präferenzdaten über das Internet diskutiert. Dabei entsteht eine Verknüpfung der beiden Forschungsfelder Online-Forschung und empirische Präferenzforschung. Eine Reihe von Argumenten spricht für dieses Forschungsvorhaben. Obwohl es sich beim Internet noch um ein sehr junges Medium handelt,¹ das bisher nicht flächendeckend verbreitet ist, argumentiert Hauptmann (1999) bereits früh, es sei sinnvoll sich jetzt mit dem Potenzial des Internets für wissenschaftliche Zwecke zu beschäftigen, um nicht wie beim Telefon zum Zeitpunkt der Nutzbarkeit über unzureichendes Wissen zu verfügen. In der kommerziellen Marktforschung hat das World Wide Web bereits einen größeren Anteil an der Gesamtzahl der durchgeführten Interviews als die schriftliche Befragung. Im Jahr 2005 unterschied sich der Anteil von persönlichen Interviews (24%) und Online-Interviews (22%) um lediglich zwei Prozent (ADM 2006). Die Verbreitung des Internets in der Bevölkerung nimmt weiter zu. Dafür sorgen unter anderem auch die sinkenden Preise für den Online-Zugang. In der Schweiz wurde z. B. ein Anschluss an das Internet zum Festpreis von der Regierung als Grundversorgung festgelegt.

Ein weiteres wichtiges Argument für die Online-Forschung ist die Weiterentwicklung der Hardware. Zum einen der Trend zu tragbaren Geräten und der Ausbau der Netzabdeckung für den mobilen Gebrauch,

1 Die Ursprünge des heutigen Internets reichen in die 1960er-Jahre zurück (vgl. MÜNZ/NEFZGER 2005).

zum anderen die Verschmelzung des Internets mit Telefon² und Fernseher geben Grund zu der Annahme, dass die Relevanz des Internets auch in den kommenden Jahren weiter zunehmen wird. Man kann vermuten, dass in Zukunft nicht mehr zwischen online vs. offline unterschieden wird, sondern lediglich zwischen elektronisch gestützten und nicht elektronisch gestützten Befragungen, da alle Endgeräte, die nicht einmal mehr als Computer erkennbar sein müssen, über eine Verbindung zum Netz verfügen. Alle Technologien dieser Arbeit funktionieren sogar dann problemlos³, wenn tragbare Geräte für CASI (Computer Administered Self Interviews) oder CAPI (Computer Administrated Personal Interviews) in die Haushalte eines Samples gebracht werden. CASI werden derzeit überwiegend dann eingesetzt, wenn Interviewer-Effekte zu erwarten sind. Individuen sind eher geneigt einem Computer heikle Fragen zu beantworten (vgl. TURNER et al. 1998). In der vorliegenden Arbeit werden empirische Daten aus Web-Befragungen vorgestellt. Jedes Erhebungsinstrument, das in Web-Befragungen eingesetzt und/oder erprobt wurde, kann aber zudem in computergestützten selbstadministrierten/persönlichen Interviews verwandt werden. Zusammengefasst zeigt sich die vielfältige Einsetzbarkeit web-basierter Erhebungsinstrumente.

Das dritte Argument für Online-Erhebungen ist die Beliebtheit des neuen Mediums. In diesem Zusammenhang ist zu erwähnen, dass im aktuellen kanadischen Zensus (2006) bereits ca. 22 Prozent der Befragten eine Beantwortung via Internet der schriftlichen Befragung vorzogen,⁴ obwohl sie den schriftlichen Fragebogen, auf dem auch die URL⁵ und ein Zugangscode abgedruckt war, mitsamt frankiertem Rückumschlag bereits in den Händen hielten. Die meisten klassischen Erhebungsmethoden und Fragetypen wurden bereits erfolgreich für die Nutzung im neuen Medium – Internet – adaptiert (für einen Überblick siehe z. B. BATINIC/WERNER/GRÄF/BANDILLA 1999; FRALEY 2004; DILLMAN 2007).

Für Präferenzdaten gibt es ebenfalls erste Ansätze, die überwiegend aus der kommerziellen Forschung stammen (vgl. SAWTOOTH SOFTWARE 2002;

2 Die Stilllegung des klassischen Festnetzes und dessen Substitution durch Datennetze ist von der Industrie bereits beschlossen.

3 Es ist lediglich notwendig einen lokalen Server auf den jeweiligen Rechnern zu installieren (z. B. www.apachefriends.org) und die Befragungssoftware auf jedes Endgerät zu spiegeln.

4 Die Angabe stammt aus einer E-Mail von Graeme Gilmour (Manager, Internet Response Task). Die genaue Zahl war zum Zeitpunkt der Veröffentlichung noch unbekannt.

5 Engl. Unified Resource Locator. Dieses Akronym steht für die eindeutige Adresse einer Website im Internet und wird synonym mit dem Wort Link verwendet.

DAIBER/HEMSING 2006; THOMAS/JOHNSON/BEHNKE/SANDERS 2005). Bislang fehlt jedoch eine systematische Zusammenfassung der bestehenden Möglichkeiten. Auch werden grundlegend wichtige Erhebungsinstrumente zur Präferenzmessung im Internet erst im Zuge dieser Arbeit entwickelt. Eines der primären Ziele dieser Arbeit ist, ein Instrument zu entwickeln, mit dem es ohne großen zeitlichen und finanziellen Aufwand möglich ist, Rangreihen mit einer großen Anzahl von Bewertungsobjekten, selbstadministriert von vielen Befragten zu erfassen (z.B. die Werte des Rokeach Value Survey; ROKEACH 1968).

Das Internet wird derzeit als modernes, gefragtes und wichtiges Forschungsfeld angesehen (HAUPTMANN 1999; ILLINGWORTH 2001). Doch das zweite Thema dieser Arbeit – Präferenzdaten einschließlich ihrer empirischen Messung und Analyse – ist von grundlegenderer Bedeutung für die sozialwissenschaftliche Forschung. Präferenzentscheidungen durchdringen jedwedes menschliches Handeln im Sinne von Weber (1976). Dies setzt voraus, dass bewusst aus Handlungsalternativen gewählt wird. Es ist daher nicht verwunderlich, dass die Tradition der empirischen Präferenzdatenanalyse bereits ins Jahr 1871 (FECHNER 1871) zurückgeht. Nicht nur die lange Historie der Präferenzdatenanalyse, sondern auch die weit reichende Verbreitung von Forschungsfragen zur Präferenz in verschiedene Disziplinen unterstreicht die fundamentale Bedeutung von Präferenzdaten.

Disziplinübergreifend ist das derzeit häufigste Argument gegen die Erhebung von Präferenzdaten der enorme Arbeitsaufwand. Die einfachste Methode, diese Daten ›offline‹ zu erheben, ist das Ordnen eines Kartenspiels durch den Interviewten. Die Reihenfolge der Karten wird später vom Interviewer in den Fragebogen oder die entsprechenden Felder der CAPT-Software übertragen. Schon dies ist aufwendig. Versucht man jedoch Präferenzdaten in selbstadministrierter Form zu erheben, wurde letztlich nur der Ansatz von Rokeach (1968) erfolgreich erprobt. Hierbei erhält jede befragte Person mehrere bedruckte, gummierte Klebezettel, die von ihr, nach Wichtigkeit geordnet, in die persönliche Reihenfolge geklebt werden sollen. Die entstandene Wertordnung, gebildet durch Klebezettel, wird daraufhin zurückgeschickt.

Der zweite wichtige Kritikpunkt bei der Verwendung von Präferenzdaten ist deren komplexe Auswertbarkeit. Eine Vielzahl von Autoren weist auf negative Eigenschaften von Präferenzdaten hin. Sie können mit den falschen Analysemethoden zu Fehlschlüssen führen (vgl. BACHER 1987; VAN SCHUUR 1997; CORNWELL/DUNLAP 1991; DUNLAP/CORNWELL 1994), die Ver-

suchspersonen überfordern (vgl. KLEIN/DÜLMER/OHR/QUANDT/ROSAR 2003) oder nur sehr geringe Validität aufweisen (vgl. JOHNSON/SALLIS/HOVELL 1999; MERRITT/MARSHALL 1984).

Andere Autoren sehen überwiegend die Vorteile von ipsativen⁶ Daten: So wird erhöhte Varianz erzielt (z. B. MCCARTY/SHRUM 2000) und Antworttendenzen, wie z. B. die Ja-Sager-Tendenz unterbunden (ALWIN/KROSNICK 1985; CHEUNG/CHAN 2002; CHEUNG 2004; CUNNINGHAM/CUNNINGHAM/GREEN 1977; GURWITZ 1987). Auch geben sich die Befragten mehr Mühe bei der Beantwortung der Fragebögen (ROKEACH 1973). Letztlich wird betont, dass im realen Leben auch häufig Wahlentscheidungen getroffen werden müssen. Deshalb gäben Präferenzdaten ein realistischeres Abbild der Realität als Ratingdaten. Sogar einer Vielzahl von behavioristischen Theorien basiere auf einem ipsativen Messmodell (SAVILLE/WILSON 1991; BROVERMAN 1962).

Die Verknüpfung des klassischen Forschungsfeldes der Präferenzdaten mit den neuen Möglichkeiten des Internets birgt Potenzial für neuen Erkenntnisgewinn. Besonders vielversprechend ist die Verbindung von Präferenzforschung und Online-Umfrageforschung aus dem Grund, dass häufig multifaktorielle Designs zur experimentellen Analyse von Präferenzen notwendig sind, die große Fallzahlen in randomisierter Zuordnung benötigen. Demnach sind vorrangig viele Versuchspersonen notwendig, durch den experimentellen Versuchsaufbau ist die Repräsentativität der Versuchspersonen hingegen von geringer Bedeutung. Genau diese Anforderungen können Online-Befragungen erfüllen.

Der erste Teil (Kapitel 2) der Arbeit erörtert den theoretischen Hintergrund von Präferenzdaten. Hier werden die Grundlagen der Präferenz erklärt, Begriffe definiert und erläutert, wann es sinnvoll ist, Präferenzdaten zu erheben sowie welche Grenzen im Umgang mit Präferenzdaten gesetzt sind. Ausgehend von Grundlagen, Definitionen und Formalisierungen relevanter Begriffe im Forschungsfeld der Präferenzdaten, erfolgt die Analyse der Vor- und Nachteile von Präferenzdaten. Dabei wird herausgearbeitet, worauf die unterschiedliche und teilweise widersprüchliche Bewertung des Nutzens von Präferenzdaten zurückzuführen ist. Durch die Diffe-

6 Ipsative Daten sind »nachempfundene« Präferenzdaten. Sie lassen sich daran erkennen, dass die Summe der gegebenen Antworten eines Individuums genau gleich denjenigen aller anderen Individuen ist (vgl. DUNLAP/CORNWELL 1994: 115). Komplette Rangreihen fallen, da jede Person dieselbe Anzahl von Objekten (m) mit Werten von 1 bis m reihen muss und sich daraus die Konstante ($c=1+\dots+m$) ergibt, unter die Bezeichnung ipsativer Daten.

renzierung zwischen der Analyse von Objekten im Gegensatz zur Analyse von Personen können diese Widersprüche aufgelöst werden.

Der zweite Teil (Kapitel 3) thematisiert zu Beginn die Auswahl der Bewertungsobjekte. Nach einem Exkurs zur kurzen Geschichte des Internets und einer groben Erläuterung der Besonderheiten von Online-Befragungen im Vergleich zu klassischen Erhebungsmedien wird ausführlich auf die Operationalisierung und Datenerhebung von Präferenzdaten im Internet eingegangen. Dabei wird detailliert darlegt, in welcher Form sich Präferenzdaten sinnvoll erheben lassen. Daraus leiten sich adäquate Auswertungsmöglichkeiten für Präferenzdaten ab.

Der dritte Teil (Kapitel 4) behandelt sehr unterschiedliche Methoden der Datenauswertung von Präferenzdaten. Die einzelnen Verfahren werden systematisch gegliedert und bewusst anwendungsbezogen beschrieben. Besonderes Augenmerk gilt dabei den Effekten der linearen Abhängigkeit von Präferenzdaten. Eines der Ziele dieser Arbeit ist zu zeigen, welche Verfahren sich gut zur Analyse von Präferenzdaten eignen und welche als eher ungeeignet erscheinen. Der vierte Teil (Kapitel 5) beschreibt Verfahren, die es ermöglichen die Analysen direkt im Internet durchzuführen bzw. zu veröffentlichen. Im fünften Teil (Kapitel 6) werden Ergebnisse eigener Erhebungen vorgestellt und ausgewählte Forschungsfragen zum Thema ›Präferenzforschung online‹ empirisch beantwortet. Eines der Ziele dieser Arbeit ist dabei die Bereitstellung erster Kennzahlen zur Beurteilung verschiedener Präferenzmessverfahren für den Einsatz im Internet. Die empirischen Analysen im letzten Teil dieser Arbeit bearbeiten Fragestellungen zu den Operationalisierungsformen Rangreihung, Paarvergleich, Rating und Idealskala.

Abschließend folgt eine Zusammenfassung der Ergebnisse mit einem Ausblick auf viel versprechende Weiterentwicklungen zum Thema ›Präferenzdaten online‹ (Kapitel 7).