

Miriam Yom und Thorsten Wilhelm

Web Usability-Tests mit qualitativen Methoden am Beispiel von Online-Fokusgruppen

Web Usability-Testing and Qualitative Methods – A Case-Study with Online-Focus Groups

Web Usability_Testmethoden_Online-Fokusgruppen_Methodenvergleich

Zusammenfassung. Fokusgruppen und insbesondere ihre Online-Variante sind im Kontext der nutzerbasierten Usability-Beurteilung noch wenig untersucht. Es stellt sich die Frage, welche Effekte die spezielle Kommunikationssituation in Online-Fokusgruppen (im Weiteren OFGs) auf die Qualität und Verwertbarkeit der Daten hat. In diesem Artikel werden die konstituierenden Merkmale der synchronen computervermittelten Kommunikation (im Weiteren cvK) und somit relevanten Determinanten der Datenqualität von OFGs dargestellt. Des Weiteren stellen wir Ergebnisse eines Methodenvergleichs zwischen konventionellen Fokusgruppen im Labor und OFGs bei der Beurteilung eines Online-Reiseshops vor. Es zeigte sich, dass durch Kommunikationstraining und technische Unterstützung die Defizite der synchronen Online-Kommunikation ausgeglichen werden können. Anhand einer deskriptiven Themenanalyse konnte gezeigt werden, dass es starke Überschneidungen der Aussagen in den beiden Kommunikationsbedingungen gibt.

Summary. In the field of user-based usability evaluation focus groups and in particular their online alternative are still little examined. In this article we investigate the effects of the special communication condition in online focus groups (OFGs) on quality and validity of the data. First we present the characteristics of synchronous computer-mediated communication (CMC) and thus relevant determinants of data quality in OFGs. Further-more we report results of a method comparison between conventional focus groups in the laboratory and OFGs for evaluating an online travelshop. There is empirical evidence that deficits of synchronous online communication can be adjusted by communication training. The descriptive topic analysis revealed strong overlaps of the statements in both communication situations.

1. Einleitung

Bei Gruppendiskussionen oder sog. Fokusgruppen werden die verbalen Aussagen zu vorgegebenen Themen und Meinungsgegenständen im Kontext einer Gruppeninteraktion erhoben. Ursprünglich von Lewin in der Kleingruppenforschung eingesetzt, entwickelte sie sich aufgrund ihrer methodischen Flexibilität, vielfältigen Einsetzbarkeit und Ökonomie schnell zu einer der wichtigen qualitativen Methoden in der Marktforschung. Fokusgruppen werden für die unterschiedlichsten Fragestellungen ein-

gesetzt, z. B. zur Erhebung von Kaufentscheidungsmotiven, Produktpräferenzen im Rahmen von Produktinnovationsprozessen, Zielgruppennormen und -werten etc. Auch bei der nutzerorientierten Entwicklung von Internetapplikationen, mobilen Diensten und Websites haben sich heute Fokusgruppen als qualitative Datenerhebungsmethode etabliert.

Dabei liegt es bei diesen interaktiven Untersuchungsobjekten nahe, die klassische Variante aus dem Labor in den Online-Bereich zu übertragen, da hier eine Evaluation ohne Medienbruch möglich

ist. Bei diesen sog. Online-Fokusgruppen (im Weiteren OFGs) treffen sich die Teilnehmer mit einem Moderator in einem virtuellen Diskussionsraum, in dem zu vorgegebenen Themen synchron und textbasiert miteinander kommuniziert wird. Über den heimischen PC wählen sich die Teilnehmer per Internetverbindung in einen Chatraum ein, der an die speziellen Bedürfnisse des Usability-Forschers/Marktforschers technisch angepasst wurde (vgl. Bild 1).

Eine neue Erhebungsmethode wie die OFGs wirft jedoch eine Reihe von Fragen auf:

- Welche Vor- und welche Nachteile haben OFGs im Vergleich zu den Face-to-Face-Fokusgruppen?
- Werden bei OFGs Ergebnisse in gleicher Qualität erzeugt wie bei konventionellen Fokusgruppen?

Um diese Fragen zu beantworten, werden wir ausgewählte Ergebnisse eines Methodentests vorstellen, in dem wir die Ergebnisse von konventionellen vs. online erhobenen Fokusgruppen bei der Beurteilung eines prototypischen Online-Shops verglichen haben. Zunächst sollen jedoch die Möglichkeiten und Grenzen sowie die Einsatzfelder der OFGs im Detail vorgestellt werden.

2. Zur Methode der Online-Fokusgruppen

Es gibt verschiedene Meinungen dazu, für welche Fragestellungen Fokusgruppen im Allgemeinen und OFGs im Speziellen geeignet sind. So betont beispielsweise Görts (2001) die besondere Adäquanz von Fokusgruppen in explorativen Studienphasen, die quantitativen Studien vorgeschaltet sind. Morgan (1997) sieht den Sinn und Zweck von Fokusgruppen jedoch nicht nur auf die explorative Phase beschränkt. Fokusgruppen können seiner Meinung nach auch als eine singuläre Haupterhebungsmethode angewendet werden, die Antworten auf die Forschungsfrage gibt. Im praxisorientierten Usability-Test sind Online- und Offline-Fokusgruppen ähnlich wie Einzelinterviews oder Protokolle Lauten Denkens dazu geeignet, wahrgenommene Usability-Fehler, kognitives Verständnis und Einstellungen zu explorieren.

Die Vorteile der OFG im Vergleich zu ihrem Face-to-Face-Pendant sind insbesondere darin zu sehen, dass sie zeit- und kostenökonomischer, flexibler und falls notwendig über nationale Grenzen hinweg durchzuführen sind. Die Teilnehmer verbleiben in ihrer gewohnten Umgebung und fühlen sich deshalb freier, ihre Meinung kundzutun. Ein weiterer Vorteil ist, dass ansonsten schwierig zu akquirierende Zielgruppen beispielsweise aus dem B-to-B-Bereich eher für Online-Diskussionen zu gewinnen sind. Darüber hinaus sind OFGs in der Usability-Forschung von herausragendem Interesse, da eine Evaluation des Angebots



Bild 1: Ansicht des Moderators – Software zur Durchführung von Online-Fokusgruppen der eResult GmbH

ohne Medienbruch durchgeführt werden kann.

Die Prototypen-Entwicklung für Websites und Shops ist häufig hohen zeitlichen Restriktionen unterworfen, die sich in einem für den Usability-Test zur Verfügung stehenden kurzen Zeitraum niederschlägt. Diesem Umstand tragen OFGs Rechnung, da die ansonsten bei Labor-Fokusgruppen aufwendig zu erstellenden Transkripte der Videobänder durch die automatischen Diskussionsprotokolle entfallen. Dies erleichtert die anschließende quantitative und/oder qualitative Auswertung des verbalen Materials. Diese nutzbaren Rationalisierungseffekte ermöglichen ökonomisch attraktive OFG-Studien in dichteren Zeitintervallen und/oder mit größeren Stichproben.

Diesen Vorteilen stehen jedoch auch problematische Aspekte gegenüber. So können mögliche Störeinflüsse während der Datenerhebung nur bedingt oder gar nicht kontrolliert werden. Es kann beispielsweise nicht ausgeschlossen werden, dass die Kommentare eines Teilnehmers durch Dritte beeinflusst werden. Während es bei klassischen Diskussionsrunden in der Regel keinen Ausfall während der Datenerhebung gibt, besteht bei OFGs eine realistische Gefahr, dass Teilnehmer während der Diskussion im WWW rumsurfen, sich vom PC entfernen etc. und als Folge davon episodisch oder gar nicht mehr an der Diskussion teilnehmen.

Ein weit reichender Kritikpunkt betrifft die Ergebnisqualität. So wird an-

gezweifelt, dass die Güte der OFGs vergleichbar hoch sei wie die der konventionellen Fokusgruppen (vgl. z. B. Greenbaum 1995). Korrekt ist, dass die Effekte des Mediums bzw. genauer der spezifischen Kommunikationssituation auf die Ergebnisse unklar sind. Diese Forschungslücke impliziert jedoch nicht, dass die Validität dieser Erhebungsmethode per se geringer ist als bei ihrem Offline-Pendant. Es gilt vielmehr diese Forschungslücke zu schließen.

Zu diesem Zweck möchten wir zunächst kurz auf die besonderen Merkmale der Kommunikationssituation in OFGs eingehen, um dann ausgewählte Ergebnisse eines Methodenvergleichs (OFG vs. Labor-Fokusgruppen) vorzustellen.

3. Besonderheiten der Kommunikationssituation in Online-Fokusgruppen

Nach Herring (1999) kann die cvK in OFGs durch folgende Charakteristika gekennzeichnet werden:

- Audio-visuelle und psychologische Anonymität,
- synchrone, textbasierte Kommunikation mit zeitlicher Verzögerung sowie
- Wegfall von unmittelbarer Rückmeldung.

Audio-visuelle und insbesondere psychologische Anonymität

Für die Kommunikation in OFGs steht ausschließlich der verbale, textbasierte

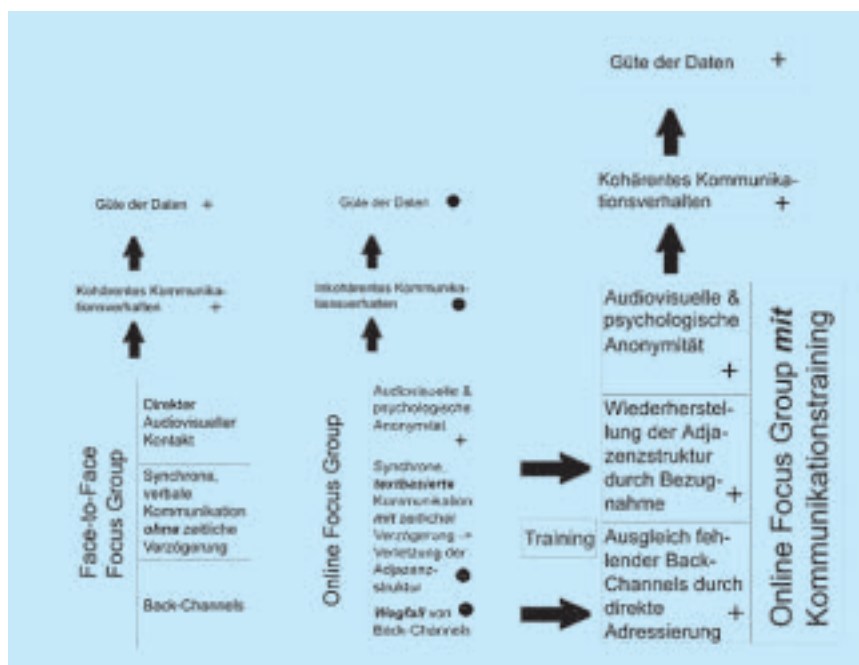


Bild 2: Kommunikationssituation in Face-to-Face- vs. Online-Fokusgruppen mit und ohne Training

Kanal zur Verfügung, da die Teilnehmer sich weder sehen noch hören können. Es wird davon ausgegangen, dass die anonyme Kommunikationssituation in OFGs im Vergleich zu Labor-Fokusgruppen zu mehr Offenheit im Antwortverhalten (vgl. Prickarz und Urbahn 2002), ausgeprägten Selbstoffenbarungen und geringerer sozialer Erwünschtheit führt (vgl. Joinson 2001).

Synchrone, textbasierte Kommunikation mit zeitlicher Verzögerung

Nachrichten müssen von den Teilnehmern zunächst geschrieben, eventuell editiert und an den Server abgeschickt werden. Auf diese Weise entstehen (a) spezifische Kommunikationsbarrieren, die durch die Umsetzung von Sprache in motorische Bewegungen bedingt sind, und (b) zeitliche Verzögerungen bei der Beantwortung von Nachrichten. Da alle Teilnehmer simultan Nachrichten abschicken können, ist eine Ordnung der Nachrichten in der von den Teilnehmern inhaltlich intendierten Reihenfolge nicht gewährleistet. Die Nachrichten erscheinen vielmehr nach ihrem Eingangszeitpunkt im zentralen Chat-Fenster. Diese technisch bedingten Eigenschaften der Kommunikationssituation führen zu einer Verletzung des so genannten *Adjazenzprinzips*. Dieses besagt, dass Gespräche in Sequenzen segmentiert wer-

den, z. B. indem auf Fragen Antworten, auf Aufforderungen Reaktionen etc. folgen (vgl. Cornelius 2001). Kommunikationssituationen, die dem Adjazenzprinzip in praktischer und semantischer Hinsicht folgen, werden von dem Teilnehmer als „geordnet“ wahrgenommen und erleichtern die Verarbeitung der Informationsinhalte.

Die zeitlichen Verzögerungen aufgrund des Schreib- und Editieraufwands in Kombination mit der Gleichzeitigkeit der Kommunikation in OFGs führen dazu, dass die Adjazenzstruktur der Diskussion unterbrochen ist. Die Verständlichkeit von thematischen Sequenzen wird dadurch erschwert, da die Zuordnung der Nachrichten aufeinander nicht mehr oder schwer nachvollziehbar ist. Es entsteht (der Eindruck) ein(es) „Konversations-Chaos“.

Fehlen unmittelbarer Rückmeldung (Back-Channels)

Rückmeldungen in Gesprächssituationen sind Äußerungen, die den Partner nicht unterbrechen, sondern Zustimmung und Ablehnung signalisieren, und damit den weiteren Verlauf eines Gesprächs beeinflussen. Sie dienen dazu, dem aktuellen Sprecher ein Feedback zu geben, dass das Gesagte aufgenommen und verstanden wurde, und unterstützen die Aufrechterhaltung des Ge-

sprächs (vgl. McLaughlin 1984). Dies kann in persönlichen Kommunikationssituationen über Gesten, Mimik, Körpersprache oder verbal erfolgen. In der cvK ist eine solche zeitgleiche Rückmeldung technisch bedingt nicht möglich. Das Fehlen bzw. die zeitliche Verzögerung von Rückmeldungen in der cvK führt zu einer zusätzlichen Reduzierung der Verständlichkeit der Nachrichten, da die Etablierung eines gemeinsamen Referenzsystems erschwert wird (vgl. Cornelius 2001).

Bei der FtF-Fokusgruppe adressieren die Teilnehmer ihre Aussagen über die Ausrichtung des Blicks oder der Körperhaltung eindeutig an die Gruppe oder an einzelne Personen. Diese nonverbale Adressierungsfunktion ist in textbasierten Online-Systemen nicht mehr gegeben. Gleichzeitig sind Personen mit geringer oder keiner Chat-Erfahrung nicht gewohnt, Adressaten explizit anzusprechen, da sie dies aus konventionellen Gesprächssituationen nicht kennen. Werden Nachrichten ohne expliziten Adressaten jedoch online artikuliert, sind diese zunächst unspezifisch an die Gruppe gerichtet. Als Folge dieses Kommunikationsverhaltens bleibt häufig unklar, an wen eine Nachricht gerichtet ist und auf welches Thema die Nachricht Bezug nimmt. Die dyadische Organisation des Gesprächsablaufs wird hierdurch erschwert. Die Dialogsequenzen zwischen einzelnen Teilnehmern sind jedoch notwendig und wichtig, um Themen zu etablieren und weiter diskutieren zu können.

Inkohärenz in der computervermittelten Kommunikation

Sowohl der Wegfall der unmittelbaren Rückmeldung als auch die zeitliche Verzögerung und Simultaneität sind Ursachen für die Entstehung von *Inkohärenz* in der cvK. So haben Analysen von unmoderierten Chats gezeigt, dass durch die Verletzung des Adjazenzprinzips Themen nur kurz aufrechterhalten oder erst gar nicht etabliert werden (vgl. Herring 1999). Konversationale *Kohärenz* versteht Cornelius (2001 S. 28) „... als Fähigkeit und Motivation von Gesprächsteilnehmern ..., gegenseitig aufeinander einzugehen und gemeinsam Themen zu etablieren und weiterzuentwickeln“.

Inwieweit ist die Kohärenz für die Fragestellung nach der Güte der Daten aus OFGs im Zusammenhang mit der Beurteilung der Benutzerfreundlichkeit von

Websites von Relevanz? Zur Beantwortung dieser Frage muss die Aufgabenstellung von Fokusgruppen beim Website-Test berücksichtigt werden. Die Gruppe soll die Benutzerfreundlichkeit und das Gefallen von Shop-Prototypen *beurteilen* und *konkrete Verbesserungsvorschläge generieren*. Neben einer gemeinsamen Zieldefinition und Zielorientierung, müssen die Teilnehmer dafür problemrelevantes Wissen und Meinungen austauschen sowie vor dem Hintergrund der Problemstellung bewerten. Dabei ist der Grad der Unterschiedlichkeit der individuellen, subjektiven Bewertungen von alternativen Lösungsvorschlägen für den Usability-Praktiker von besonderem Wert. Auf diese Weise werden die zielgruppenspezifischen Bewertungsmechanismen und -kriterien sichtbar und stellen somit die Grundlage für die Herleitung von Handlungsempfehlungen dar.

Der Austausch von relevantem Wissen und die Bewertung von alternativen Verbesserungsvorschlägen setzt jedoch voraus, dass die Teilnehmer aufeinander eingehen, ein gemeinsames Thema etablieren (z. B. Beurteilung des Shops) und weiterentwickeln (z. B. Verbesserungsmöglichkeiten und -wünsche). Kohärenz in der Kommunikationssituation ist vor dem definierten Erkenntnisinteresse im Rahmen einer entsprechenden Evaluation also eine *notwendige Voraussetzung*, um zu befriedigenden Ergebnissen zu gelangen.

Als Zwischenfazit kann an dieser Stelle die Diskussion über die Güte von Daten aus OFGs spezifiziert werden. Die synchrone cvK in OFGs ist aufgrund fehlender unmittelbarer Rückmeldungen und der technisch bedingten Verletzung des Adjazenzprinzips durch Inkohärenz gekennzeichnet. Wenn die Evaluation und Generierung von Verbesserungsvorschlägen Zielsetzung der OFG ist, dann ist Kohärenz eine wichtige Determinante der Datenqualität. Werden in diesem Fall keine Maßnahmen zur Reduzierung der Inkohärenz in OFGs eingesetzt, so ist tatsächlich zu erwarten, dass die Güte der verbalen Daten in der Online-Bedingung im Vergleich zu konventionellen Fokusgruppen leidet. Umgekehrt heißt dies jedoch auch: Gelingt es, die Kohärenz in der OFG über geeignete Maßnahmen dem Niveau einer FtF-Diskussion anzugleichen, dann sollten

die online generierten verbalen Daten sich hinsichtlich der inhaltlichen Qualität nicht von denen der FtF-Bedingung unterscheiden (vgl. zusammenfassend Bild 2).

4. Ergebnisse des Methodenvergleichs: Face-to-Face- vs. Online-Fokusgruppen zur Beurteilung von Online-Shops

4.1 Untersuchungsanlage

Zur Überprüfung dieser Überlegungen wurde ein Methodenvergleich durchgeführt, in dem Daten von vier konventionellen Fokusgruppen und vier OFGs verglichen wurden. Zur Erhöhung der Kohärenz in den OFGs wurde ein Kommunikationstraining eingesetzt, dessen Wirksamkeit bei Entscheidungsfindungsgruppen bereits experimentell in anderen Studien geprüft wurde (vgl. CORNELIUS 2001). Bei dem Training handelt es sich um bild- und textbasiertes Trainingsmaterial, welches die Teilnehmer für die Probleme der computervermittelten Kommunikationssituation sensibilisiert. Es wurde des Weiteren mit den Teilnehmern trainiert, dass sie beim Versenden einer Nachricht explizit auf das referierte Thema Bezug nehmen (Wiederherstellung der Adjazenzstruktur) und ggf. den Adressaten der Nachricht benennen sollten (Kompensation des Wegfalls von unmittelbaren Rückmeldungen).

Bei der untersuchten Website handelte es sich um einen HTML-Prototypen eines Online-Reiseshops, der über redaktionelle Inhalte, ausgewählte Produktangebote und Buchungstrecken verfügte. Vor der Diskussion in der Fokusgruppe nutzten die Teilnehmer das Angebot sowohl frei als auch szenariobasiert. Interviewleitfaden, Diskussionszeit (90 Minuten) und Moderatorin waren für beide Untersuchungsbedingungen konstant.

Die Gruppe à 6 Teilnehmer war heterogen im Bezug auf die personenbezogenen Merkmale *Alter*, *Geschlecht* und *Berufstätigkeit* zusammengesetzt. Eine zeitökonomische Akquise adäquater Teilnehmer war mit Hilfe des Online-Access-Panels der eResult GmbH möglich.

4.2 Methodische Vorgehensweise bei der Aufbereitung der Daten

Zur Transformation der verbalen Diskussionsdaten in quantitative Größen wurde das Verfahren der Inhaltsanalyse verwendet. Zwei ausgewiesene und trainierte Assistenten kodierten das Material nach vorgegebenen Kodierregeln. Die Kodierung durch die beiden Assistenten erfolgte unabhängig voneinander, so dass die Interkoder-Reliabilität berechnet werden konnte.

Zur Analyse der Interkoder-Reliabilität wurde sowohl die einfache prozentuale *Übereinstimmung* zwischen den beiden Kodierern als auch der *Kappa-Koeffizient* (κ) nach Cohen (1970 zit.n. Bortz and Döring 1995 S. 254) für die Variablen berechnet, die durch die Assistenten klassifiziert werden mussten. Es ergaben sich für die einfache prozentuale Übereinstimmung Werte zwischen 93,99 % und 98,88 % bzw. Kappa-Koeffizienten in Höhe von .71 bis .97. Dies sind gute bis sehr gute Reliabilitäten, so dass die Daten für eine weitere statistische Analyse verwendet werden konnten.

4.3 Ausgewählte Ergebnisse

Eine zentrale Überlegung bestand darin, die Inkohärenz in der Online-Bedingung zu reduzieren. Zur quantitativen Erfassung der Kohärenz wurde ein validiertes Kodierungsschema angewendet, bei dem die kohärenten und inkohärenten Themenkoordinationsmechanismen (TKM) durch zwei Versuchsleiter kodiert wurden (zur ausführlichen Darstellung der Operationalisierung des Konstrukts der Kohärenz vgl. Boos and Corneliuss 2001, Yom 2003). Als globales Maß der Kohärenz in der Fokusgruppe dient der *Kohärenzquotient*. Er wird berechnet aus der Differenz der Gesamtzahl der kohärenten TKM und der Gesamtzahl der inkohärenten TKM sowie dem Verhältnis dieser Differenz zur Gesamtzahl aller TKM. Das durch den Kohärenzkoeffizienten (K) abgebildete Kontinuum kann Werte zwischen 1 (extrem kohärent) und -1 (extrem inkohärent) annehmen. Für die Online-Bedingung ergab sich ein $K = .91$ und für die FtF-Bedingung ein $K = .94$. Beide Werte unterschieden sich nicht signifikant voneinander. Es ist also gelungen, die Kohärenz in der OFG dem Kohärenz-Niveau der konventionellen Fokusgruppe anzugleichen.

Anzahl der Wörter

Aufgrund des Tipp- und Editieraufwands in der Online-Bedingung zeigte sich erwartungsgemäß ein deutlicher Gruppenunterschied bei der Anzahl der Wörter. Dabei ist im Mittel ein Verhältnis von etwa 1 : 3 (Online : Labor) festzustellen. Die Teilnehmer der Offline-Diskussion produzierten also dreimal so viele Wörter wie die Teilnehmer der OFGs. Aber bedeutet dies auch, dass in der Online-Bedingung nur ein Drittel an Aussagen und Meinungen artikuliert wurde?

Anzahl der Statements

Bei einem Statement bzw. einer Nachricht handelt es sich um eine in sich geschlossene verbale Äußerung oder textbasierte Mitteilung eines Teilnehmers. Ein Sprecherwechsel bzw. das Eintreffen einer neuen Nachricht markiert das Ende des Statements bzw. der Nachricht. Betrachtet man die Anzahl der Statements, so wird deutlich, dass die Teilnehmer der OFGs (Mittelwert 41,55, SD 11,60) trotz des Schreib- und Editieraufwands im Durchschnitt sogar mehr Statements geäußert haben als die Teilnehmer der konventionellen Diskussion (Mittelwert 33,79, SD 18,50). Die Anzahl der Statements in den beiden Bedingungen unterscheidet sich jedoch nicht signifikant.

Anzahl der angesprochenen Themen im Vergleich

Auch wenn die Differenz der Statementanzahl bei beiden Kommunikationssituationen nicht bedeutsam ist, lässt dies noch keinen Schluss darüber zu, in welcher Variationsbreite auf die Benutzerfreundlichkeit bezogene Themen von den Teilnehmern in den beiden Kommunikationsbedingungen diskutiert wurden. Oder anders ausgedrückt: Inwieweit haben evtl. unterschiedliche Gruppendynamiken dazu geführt, dass sich je nach Kommunikationssituation unterschiedliche Themen etabliert haben? Denn letztendlich entscheiden die Teilnehmer, welche Themen sie für diskussionswürdig halten.

Über die Gegenüberstellung der Anzahl der angesprochenen Themen kann die inhaltliche Dimension der Diskussionen in den Kommunikationsbedingungen stärker berücksichtigt werden. Betrachtet man die angesprochenen Themen in den Online- und Offline-Gruppen gemeinsam, dann sind insgesamt

65 Themenaspekte diskutiert worden. Dabei zeigte es sich, dass insgesamt von den Teilnehmern der OFG (n = 50) etwa gleich viele Aspekte der Benutzerfreundlichkeit angesprochen und diskutiert wurden wie in der konventionellen Fokusgruppe (n = 55). Die Ergebnisse zur Anzahl der angesprochenen Themen sind bemerkenswert, bedenkt man, dass aufgrund des zusätzlichen kognitiven Aufwandes durch das Schreiben und Editieren der Nachrichten für die Teilnehmer der OFGs eine geringere Produktionsgeschwindigkeit zu erwarten gewesen wäre. Wie sind diese Befunde zu erklären?

Zum einen wirkt sich die Möglichkeit zur Teilnahme an parallel laufenden Diskussionssträngen positiv auf die Zeiteffizienz aus. Des Weiteren kann in den Logfiles und Transskripten abgelesen werden, dass in den OFGs wesentlich kürzer und prägnanter von den Teilnehmern formuliert wird als in den konventionellen Gruppen (vgl. hierzu auch Prickarz and Urbahn 2002). Dieses Ergebnis spricht auch für die Überlegungen von Walther (2000), der davon ausgeht, dass die Botschaften in cvK mit mehr Bedacht formuliert werden. So beansprucht der Schreib- und Editieraufwand in der cvK zwar kognitive Ressourcen. Gleichzeitig müssen jedoch für die konventionelle Kommunikationssituation notwendige Ressourcen wie z. B. die Umweltbeobachtung, Kontrolle des eigenen nonverbalen Verhaltens etc. nicht aktiviert werden. Er schließt daraus, dass diese Energie vom Sender dazu genutzt werden kann, die inhaltliche und formale Konstruktion der geschriebenen Botschaften im Sinne des intendierten Eindrucks beim Kommunikationspartner zu nutzen. Diese „überschüssigen“ Ressourcen können u. E. jedoch auch der aufgabenbezogenen Dimension zugute kommen und sich in knappen, aber auch prägnanten, themenbezogenen Äußerungen niederschlagen.

Anzahl der Meinungen und Urteile

Diese Vermutung bestätigt sich auch in der quantitativen Analyse der Statements. So wurden Statements, die eine bewertende Komponente enthalten, nochmals nach dem Kriterium Bezugsobjekt in drei Kategorien unterschieden. Bezog sich das Statement

auf einen grundsätzlichen Zusammenhang zum Thema Websites/Internet allgemein, dann handelte es sich um eine allgemeine Meinung. War das Statement beschreibend auf die zu untersuchende Website (Prototyp oder Liveangebot) bezogen, dann wurde diese Aussage als objektbezogene Meinung kodiert. Bezog sich die Aussage auf ein konkretes Gestaltungsmittel des Prototypen oder Liveangebots, dann wurde dies als Urteil registriert. Um Verzerrungen bei den Gruppenunterschieden aufgrund der absoluten Anzahl an evaluativen Statements zu vermeiden, wurden die absoluten Individualwerte in relative Werte transformiert. Es zeigte sich, dass die relative Anzahl der allgemeinen Meinungen und die relative Anzahl der objektbezogenen Meinungen sich zwischen den beiden Kommunikationsbedingungen nicht signifikant unterscheiden. Ein deutlicher und signifikanter Unterschied zeigte sich jedoch bei der relativen Anzahl der Urteile. Teilnehmer der OFGs äußerten signifikant mehr Urteile als Teilnehmer der Laborbedingung ($T = 2.83$, $DF = 42$, $p < .05$). Das heißt, sie machten mehr Aussagen, deren Inhalt sich auf ganz konkrete Gestaltungsmittel des Webangebots bezog. Diese Ergebnisse sprechen also dafür, dass die Teilnehmer der OFGs prägnantere Aussagen gemacht haben als die Teilnehmer der Face-to-Face (FtF)-Fokusgruppen.

Gruppenspezifische Themen

Betrachtet man die ausschließlich in einer Versuchsbedingung diskutierten Themen, also jene, die nur von den Teilnehmern der einen oder der anderen Gruppe angesprochen wurden, so zeigt sich folgende Verteilung: 15 Themenfelder der insgesamt 65 angesprochenen Themen (23,1 %) wurden nur von Teilnehmern der Laborgruppe angesprochen, während sich zu 10 Themen nur Teilnehmer der OFGs äußerten (15,4 %). Es ist auffällig, dass beide Methoden in ähnlichem Maße gruppenspezifische Themen generieren.

Wenden wir uns nun der inhaltlichen Dimension der gruppenspezifischen Themen zu. Gibt es möglicherweise einen charakteristischen Unterschied, welche Art von Themen eher von Online- oder Laborgruppen ange-

sprochen werden? Es zeigte sich, dass sich die beiden Kommunikationsbedingungen insbesondere bei der Generierung von Themen im Bereich „Buchung“ unterscheiden. Dieser Themenblock ist allerdings eine Ausnahme, da er ursprünglich keinen zentralen Aspekt der Diskussion darstellte. Aufgrund der insbesondere durch die Teilnehmer der Laborbedingung initiierten Etablierung dieses Themas wurde diese neue Hauptkategorie ex post notwendig. In den OFGs trat dieses Diskussionsverhalten so gut wie nicht auf, d.h. die Teilnehmer antworteten und diskutierten eng im Rahmen der vorgegebenen Themen.

Überlappende Themen

40 (61,5 %) von insgesamt 65 inhaltlichen Aspekten wurden in beiden Kommunikationsbedingungen diskutiert. Auffällig ist dabei, dass bei den überlappenden Themen die Richtung der Beurteilung bzw. Aussagen in 31 (78 %) der 40 Themen unabhängig von der Kommunikationssituation übereinstimmten. Lediglich bei neun Themen differierten die Meinungen zwischen den beiden Kommunikationssituationen. Bei acht Themen wurden in den OFGs die Meinungen und Urteile nicht gegenläufig zur Laborbedingung, sondern lediglich kontroverser diskutiert. Eine kontroverse Diskussion kann die Ergiebigkeit der Ergebnisse durchaus positiv beeinflussen. Einmütige Urteile der Gruppe können durch die Beeinflussung eines charismatischen Wortführers zustande gekommen sein, die eine vertiefende Diskussion von Pro und Contra zu den einzelnen Themen im Keim erstickt oder sehr stark in eine Richtung gelenkt hat.

Nur ein Thema wurde tatsächlich gegensätzlich beurteilt. Beim Thema „Guided Tours“ waren die Teilnehmer der Laborbedingung der Meinung, dass diese eher unnötig sei, während die Online-Teilnehmer dieses Hilfeangebot positiv hervorhoben. Für die Handlungsempfehlungen bezüglich einer benutzerfreundlichen Gestaltung wären diese gegenläufigen Meinungsbilder jedoch nicht von großer Tragweite, da „Guided Tours“ eine optional angebotene Hilfsfunktion darstellen, die nicht zu den zentralen Funktionsbereichen des Shops gehört.

Anzahl der Verbesserungsvorschläge

Statements, die einen konkreten Vorschlag zur gestalterischen oder strukturellen Veränderung der Website oder deren Inhalte beinhalteten, wurden als Verbesserungsvorschlag kodiert. Auch hier wurden zur Berechnung von Gruppenunterschieden mittels t-Tests die absoluten Individualwerte in relative Werte transformiert. Es zeigte sich, dass die Teilnehmer der OFGs (MW .16, SD .13) tendenziell mehr konkrete Verbesserungsvorschläge formulierten als die Teilnehmer der Offline-Fokusgruppen (MW .13, SD .09). Der t-Test für unabhängige Stichproben mit der normalverteilten Variable „Relative Anzahl der Verbesserungsvorschläge“ ergab jedoch, dass dieser Unterschied nicht signifikant ist. In beiden Kommunikationsbedingungen wurden also ähnlich viele konkrete Verbesserungsvorschläge generiert.

Partizipation

Betrachten wir nun die Partizipationsrate, die für jeden Teilnehmer berechnet wurde, in dem die absolute Anzahl der Statements in Relation zur Gesamtanzahl der Statements in der jeweiligen Gruppe berechnet wurde. In der Online-Bedingung (MW .20, SD .05) zeigten die Teilnehmer im Durchschnitt eine höhere Partizipationsrate als die Teilnehmer der Labordiskussion (MW .17, SD .08). Dabei sind die Gruppenunterschiede der normalverteilten Variable tendenziell signifikant ($T = 1.81$, $df 42$, $p < .10$).

5. Zusammenfassende Beurteilung der Ergebnisse

Die Ergebnisse dieser Grundlagenstudie für die Güte und Verwertbarkeit der verbalen Daten aus OFGs sind vielversprechend. Identifizierte Störvariablen wie Inkohärenz in der cvK können durch geeignete Maßnahmen neutralisiert werden. Es zeigte sich trotz des Tipp- und Editieraufwandes für die Teilnehmer der OFGs keine generelle Überlegenheit der konventionellen Fokusgruppen bei der Variationsbreite der etablierten Themen. Auch bestätigte sich, dass die Teilnehmer der OFGs den Tipp- und Editieraufwand durch Reduzierung von Redundanzen

und prägnante Aussagen kompensieren können. So formulierten Online-Teilnehmer mehr Urteile, also Aussagen zu ganz konkreten Gestaltungsmitteln des Webangebots, als die Teilnehmer der konventionellen Gruppen, während sich die Anzahl der allgemeinen Meinungen und der objektbezogenen Meinungen nicht unterschieden. Auch wiesen die Diskussionsinhalte in beiden Bedingungen einen hohen Anteil gemeinsamer Themen auf, die zum überwiegenden Teil in gleicher Richtung bewertet bzw. von den OFG-Teilnehmern kontroverser diskutiert wurden. Ebenso wie in den konventionellen Fokusgruppen wurde auch in den OFGs eine äquivalente Anzahl an Verbesserungsvorschlägen generiert. Für die Online-Variante spricht, dass die Partizipationsraten online höher waren als bei den Fokusgruppen im Labor.

Insgesamt zeigen die Ergebnisse, dass bei der inhaltlichen Qualität der Daten keine wesentlichen Unterschiede bestehen und OFGs im Rahmen von Web Usability-Tests eine interessante Ergänzung zum Methodenarsenal des Usability-Praktikers sein können. OFGs erlauben eine zeiteffiziente empirische Überprüfung von Website-Konzepten und werden damit der Schnelllebigkeit im Webprojekt-Geschäft besonders gerecht. Unsere praktische Erfahrung hat auch gezeigt, dass die Akquise der OFG-Teilnehmer bei internetspezifischen Fragestellungen sowohl zeit- als auch kostenökonomisch mittels Online-Access-Panels erfolgen kann. Ein gepflegtes Access-Panel kann so gut wie alle Quotierungswünsche des Forschers erfüllen. Insbesondere bei klar definierten Zielgruppen (z. B. regelmäßige Nutzer von Reise-Websites o. ä.) bietet ein nationales Access-Panel eine gute Ausgangsbasis für eine ökonomisch durchführbare Quotierung, bei der auch das Auftreten eines „urbanen Bias“ vermieden werden kann (vgl. Eppler und Hahn 2001 S. 252).

Für die Zukunft wünschenswert sind weitere Methodentests, um die Anwendbarkeit bei verschiedenen Fragestellungen, Zielgruppen und Anwendungen zu prüfen.

Literatur

- Boos, M.; Cornelius, C.: Zur Erfassung konversationaler Kohärenz in direkter und computervermittelter Kommunikation. In: *Partizipation und Interaktion im virtuellen Seminar* (Hrsg. Hesse, F.W. and Fiedrich, H.F.) Münster: Waxmann, 2001, S.55-80.
- Bortz, J.; Döring, N.: *Forschungsmethoden und Evaluation für Sozialwissenschaftler*. 2.Aufl. Berlin: Springer, 1995.
- Cornelius, C.: *Gegenseitiges Verständnis in Computerkonferenzen. Voraussetzungen und Folgen konversationaler Kohärenz in Entscheidungsfindungsgruppen im Medienvergleich*. Münster: Waxmann, 2001.
- Görts, T.: Gruppendiskussionen - Ein Vergleich von Online- und Offline-Focus-Groups. In: *Online-Marktforschung – Theoretische Grundlagen und praktische Erfahrungen* (Hrsg. Theobald, A.; Dreyer, M.; Starsetki, T.). Wiesbaden: Gabler, 2001, S.149-164.
- Greenbaum, T.: Focus Groups on the Internet: An Interesting Idea but not a Good One (1995). URL: <http://www.quirks.com/CGI-BIN/SM40i.exe?docid=3000:58911&%7OassArticleID=136,Stand:30.11.99,3Seiten>.
- Herring, S.C.: International Coherence in CMC. *Journal of Computer Mediated Communication* Vol. **4**, No.4 (1999). URL: <http://www.ascus.org/jcmd/vol4/issue4/herring.html>, Stand: 17.06.01, 25 Seiten.
- Joinson, A.: Self-Disclosure in Computer-Mediated Communication: The Role of Self-Awareness and Visual Anonymity. *European Journal of Social Psychology* Vol. **31**, No.2 (2001).177-192.
- McLaughlin, M.L.: *Conversation: How Talk is Organized*. Beverly Hills, 1984.
- Morgan, D.L.: *Focus Groups as Qualitative Research*. 2. Aufl. Newsbury Park: Sage Publications, 1997.
- Prickarz, H.; Urbahn, J.: Qualitative Datenerhebung mit Online-Fokusgruppen. *Planung & Analyse* **28** Jg., Nr.1 (2002) S.63-70.
- Walther, J.B.: Die Beziehungsdynamik in virtuellen Teams. In: *Computervermittelte Kommunikation in Organisationen* (Hrsg. Boos, M.; Jonas, K.; Sassenberg, K.) Göttingen: Hogrefe, 2000, S.11-25.
- Yom, M.: *Web Usability von Online-Shops*. Göttingen: Better Solutions, 2003.



1



2

1 Dr. Miriam Yom, Mitglied der Geschäftsführung eResult GmbH (Göttingen), Hauptarbeitsgebiete: (Web) Usability, Zielgruppenmarketing in interaktiven Medien, qualitative Methoden der Online-Forschung, Branchenfokus: Frauenmärkte, Medien- & Tourismusbranche E-Mail: miriam.yom@eresult.de

2 Dipl.-Kfm. Thorsten Wilhelm, Mitglied der Geschäftsführung eResult GmbH (Göttingen) Hauptarbeitsgebiete: Online-Werbung, Web Usability-Tests, Blickregistrierung, Branchenfokus: Werbung, Medien und Seniorenmärkte E-Mail: thorsten.wilhelm@eresult.de URLs: www.eresult.de, www.usability-competence.de, www.usability-contest.de