

Skalierung von Designalternativen

Theo Held¹, Martin Schrepp¹, Patrick Fischer²

SAP AG Walldorf¹, Universität Mannheim²

Zusammenfassung

BTL-Skalierung und Conjoint-Analyse sind etablierte Skalierungsverfahren, die in verschiedenen Domänen (insbesondere in der Marktforschung) erfolgreich angewandt werden. Wir zeigen in dieser Arbeit, dass sich beide Verfahren auch sehr gut für die Skalierung von alternativen Varianten von Benutzungsschnittstellen einsetzen lassen. Die Verfahren sind robust und effizient und bieten in Bezug auf messtheoretische Eigenschaften und Berücksichtigung von Urteilsfehlern Vorteile gegenüber klassischen Rating-Skalen. Unsere Anwendungsbeispiele demonstrieren den Einsatz der Skalierungsverfahren in den Bereichen Formular- und farbliche Gestaltung von Benutzungsoberflächen.

1 Einleitung

In der täglichen Arbeit von User Interface Designern und Interaktionsdesignern stellen sich häufig Fragen zur Gestaltung von Detailbereichen der Benutzungsschnittstelle, die nicht eindeutig mit Hilfe allgemeiner Designprinzipien beantwortet werden können. Beispiele dafür sind Fragen der Anordnung von Bedienelementen oder Detailspekte des visuellen Designs.

Es gibt für Designer nun verschiedenen Möglichkeiten, zu validen Entscheidungen zu gelangen. Im günstigsten Fall hilft die Konsultation einschlägiger Fachliteratur weiter. Viele Designprobleme wurden bereits empirisch untersucht und für viele davon ist die Befundlage eindeutig. Häufig führt die nähere Betrachtung der Literatur aber zu dem Ergebnis, dass unterschiedliche, teils widersprüchliche Befunde mit gleichwertigem sachlogischen Gewicht existieren oder dass die bekannten Untersuchungsergebnisse nicht unmittelbar auf eine konkrete Problemstellung übertragen werden können.

Eine typische Schwierigkeit ergibt sich aus der Tatsache, dass Benutzungsschnittstellen im Allgemeinen multidimensionale Stimuli darstellen, wobei die Gewichtung und Interaktion der Dimensionen nicht evident sind. Ein weiterer Königsweg, derartige Fragestellungen zu beantworten, ist natürlich die experimentelle Untersuchung im Labor. Innerhalb der praktischen Arbeit von Interface Designern ist aber genau diese Alternative oft nur schwer zu realisieren, da zeitliche und kostenbezogene Restriktionen im Wege stehen.

Es sind also Verfahren gefragt, die einfach, schnell und kostengünstig durchzuführen, sowie bezüglich der Reliabilität der Ergebnisse hinreichend überprüfbar und robust sind.

In dieser Arbeit werden zu den inhaltlichen Gebieten des Formularlayouts und der farblichen Gestaltung von Benutzungsschnittstellen die Verfahren BTL-Skalierung (Bradley & Terry 1952; Luce 1959) und Conjoint-Analyse (Luce & Tuckey 1964) kurz vorgestellt, angewandt und bezüglich ihrer Brauchbarkeit im Tagesgeschäft von User Interface Designern diskutiert.

2 Skalierungsmethoden

BTL-Skalierung und Conjoint-Analyse sind Verfahren, die seit geraumer Zeit eine wichtige Rolle in der experimentellen Psychologie sowie in der empirischen Marktforschung spielen. Die Conjoint-Analyse stellt aktuell den „state of the art“ der empirischen Marktforschung dar (Wittink, Vriens, Burhene 1994; Klein 2002). Wir werden hier nur in Kürze die Grundideen beider Verfahren darstellen. Für vertiefende Informationen muss auf die angegebene Literatur verwiesen werden.

2.1 BTL Skalierung

Das BTL-Modell geht von einer Menge $A = \{a_1, a_2, \dots, a_n\}$ beliebiger zu beurteilender Objekte aus. Diese Objekte werden paarweise miteinander verglichen. Beispiele für solche Objekte sind einfache physikalische Stimuli wie z.B. Lichtreize unterschiedlicher Helligkeit, Gesichter oder Konsumprodukte. Bei dem paarweisen Vergleich haben die Beurteiler zu entscheiden, welchen Reiz des präsentierten Paares sie präferieren (also z.B. schöner, erstrebenswerter oder benutzbarer finden). Das BTL-Modell wird üblicherweise für Reize $a, b \in A$ folgendermaßen formuliert:

$$p(a, b) = \frac{\phi(a)}{\phi(a) + \phi(b)},$$

wobei $p(a, b)$ die Wahrscheinlichkeit bezeichnet, mit der Reiz a gegenüber Reiz b präferiert wird und $\phi: A \rightarrow [0, \infty]$ eine Funktion ist, die jedem Element aus A einen numerischen Wert größer oder gleich 0 zuordnet. Die Werte $\phi(a_i)$ nennt man die *Skalenwerte* der BTL-Skala, wobei die Wahrscheinlichkeit, dass ein Reiz a einem Reiz b vorgezogen wird, lediglich vom Verhältnis der Werte $\phi(a)$ und $\phi(b)$ abhängt.

Voraussetzung für die Anwendung des BTL-Modells sind Daten, die mit Hilfe der Methode des *vollständigen Paarvergleichs* gewonnen wurden. Üblicherweise werden alle möglichen Reizpaare wiederholt in den beiden möglichen Reihenfolgen dargeboten, also wird bei einer simultanen Darbietung der Reize a und b z.B. Reiz a einmal links und einmal rechts dargeboten. Es ist offensichtlich, dass der zeitliche Aufwand mit der steigenden Anzahl von Stimuli stark zunimmt. Für 10 Reize sind mindestens 80 paarweise Vergleiche erforderlich, bei 12 Reizen sind bereits 120 Paare zu beurteilen (bei doppelter Präsentation jedes Paares).

Da die Aufgabe für die Versuchsperson mit einer vergleichsweise geringen kognitiven Belastung verbunden ist, sind auch höhere Anzahlen von Vergleichen in einer Sitzung durchführbar.

Das BTL-Modell macht explizite und empirisch überprüfbare Annahmen über die Struktur der aus den Paarvergleichen resultierenden Daten. Insbesondere die Anzahl von Verletzungen der *Transitivitätsbedingung* (wenn Alternative a der Alternative b vorgezogen wird und Alternative b der Alternative c , dann sollte auch Alternative a gegenüber Alternative c präferiert werden) gibt klare Hinweise auf die innere Konsistenz der Urteile. Ist die Gültigkeit der Annahmen gegeben, so kann von einer Messung auf dem Niveau einer *Verhältnisskala* ausgegangen werden.

Für eine anschauliche Darstellung des Modells und der Vorgehensweise bei der Skalenkonstruktion sei z.B. auf (Gediga 1998) verwiesen. Eine Anwendung der BTL-Skalierung für die Beurteilung von Eingabefeldern wurde von Fischer, Held, Schrepp & Laugwitz (2006) vorgestellt.

2.2 Conjoint Analyse (verbundene Messung)

Die Conjoint Analyse oder verbundene Messung wurde innerhalb der experimentellen Psychologie zur Messung multidimensionaler (Objekt-)Eigenschaften entwickelt (siehe z.B. Luce & Tuckey 1964). Eine Anwendung der Conjoint Analyse auf die Gestaltung der Navigationsleiste bei Web-Shops findet sich in (Silberer & Engelhardt 2003). Das heutige Hauptanwendungsgebiet ist die empirische Marktforschung. Typische Untersuchungsobjekte sind Konsumgüter, die durch eine Anzahl von zusammenwirkenden *Attributen* charakterisiert sind. Beispielsweise kann die Kaufentscheidung für ein bestimmtes Automodell von Attributen wie Preis, Kraftstoffverbrauch, Komfort und Leistung abhängen. Jedes dieser Attribute kann unterschiedliche Ausprägungen besitzen, wie z.B. ein Preis unter 15000€, zwischen 15000€ und 20000€, sowie über 20000€. Die zentrale Frage ist, wie sich die Attribute in ihren jeweiligen Ausprägungen auf das ordinale Urteilsverhalten eines Beurteilers auswirken. Im Folgenden werden wir lediglich die Variante eines *additiven* Zusammenhangs zwischen den numerischen Skalenwerten der Attributausprägungen betrachten. Es gibt auch Anwendungsfälle, bei denen es geboten ist, z.B. von einem *multiplikativen* Zusammenhang auszugehen, der häufigste und üblichste Fall ist jedoch die additive Beziehung.

Formal geht man bei der additiv verbundenen Messung von den Attributmengen $A_1, A_2, \dots, A_n$ aus. Jedes der von einer Versuchsperson zu beurteilenden Objekte wird durch ein n -Tupel $(x_1, \dots, x_n) \in A_1 \times A_2 \times \dots \times A_n$ dargestellt.

Wenn wir den einfachen Fall eines Objektes mit zwei Attributen A_1 und A_2 betrachten, ist bei der additiv verbundenen Messung die Frage zu klären, unter welchen Bedingungen Ab-

bildungen $\phi_1 : A_1 \rightarrow \mathfrak{R}, \phi_2 : A_2 \rightarrow \mathfrak{R}^1$ existieren, so dass für alle $a, b \in A_1$ und alle $p, q \in A_2$ folgendes gilt:

$$(a, p) \succ (b, q) \Leftrightarrow \phi_1(a) + \phi_2(p) \geq \phi_1(b) + \phi_2(q).$$

Die Relation $\succ$ ist die empirische Präferenzrelation. In der Darstellung oben bezeichnet sie, dass das Objekt (a, p) gegenüber dem Objekt (b, q) präferiert wird.² Wenn diese Präferenz besteht, so müssen auch die Summen der jeweiligen Skalenwerte der Ausprägungen in der numerischen Relation $\geq$ stehen. Die Funktionen $\phi_1, \dots, \phi_n$ werden auch *Bewertungsfunktionen* genannt, sie sind *Intervallskalen* mit derselben Einheit (Orth 1974). Vor der Durchführung einer Conjoint Analyse muss für jedes Attribut eine Annahme bezüglich der Form der Bewertungsfunktion gemacht werden. In unserem Fall werden wir von einem linearen funktionalen Zusammenhang zwischen Ausprägung und Bewertung ausgehen. Zudem stellt die Unabhängigkeit der Attribute eine notwendige Voraussetzung dar. Außerdem muss jedes mögliche Paar von Attributausprägungen mindestens einmal in einer Alternative vorkommen.

Zur Erhebung der Präferenzen stehen unterschiedliche methodische Ansätze zur Verfügung. Weit verbreitet ist das einfache nicht-metrische *Ranking-Verfahren* in dem der Beurteiler alle Objekte aus einer gegebenen Menge entsprechend seiner Präferenz direkt anordnet (siehe dazu z.B. Alwin & Krosnick 1985). Dieses Verfahren hat den Vorteil, dass es bei einer überschaubaren Anzahl von zu beurteilenden Objekten (~ 10) sehr schnell durchzuführen ist. Es ist zu beachten, dass grundsätzlich die kognitive Belastung des Beurteilers mit steigender Anzahl von Alternative erheblich zunimmt. Prinzipiell ist aber auch der im vorhergehenden Abschnitt besprochene Paarvergleich einsetzbar.

Die Skalenwerte der einzelnen Attributausprägungen können neben varianzanalytischen Verfahren auch mit einer OLS (Ordinary Least Square) Regression bestimmt werden (Witink & Cattin 1981). Eine übersichtliche Darstellung verschiedener Auswertungsmethoden sowie eine Übersicht über die wesentlichen Aspekte der Conjoint Analyse gibt (Klein 2002).

3 Empirische Beispiele

In den folgenden Abschnitten stellen wir zwei Anwendungsbeispiele zur Skalierung von Designalternativen vor: zum einen die BTL-Skalierung von alternativen farblichen Gestaltungen von Benutzungsoberflächen und zum anderen eine Conjoint-Analyse systematisch variiert Eingabeformulare.

¹ Mit $\mathfrak{R}$ ist hier die *Menge der reellen Zahlen* bezeichnet.

² Das heißt, der Beurteiler findet Objekt (a, p) mindestens so gut, schön, benutzbar, etc. wie Objekt (b, q) .

3.1 BTL-Skalierung von Farbalgorithmen

Bei der Gestaltung von Benutzungsschnittstellen werden Farben sehr häufig zur visuellen Trennung verschiedener Bereiche eingesetzt. Für die Qualität des Designs ist es dabei wichtig, dass die verwendeten Farben einen ästhetischen Gesamteindruck hinterlassen. In der hier vorgestellten Untersuchung zur Farbharmonie werden mehrere Algorithmen untersucht, die Farbkombinationen ausgehend von einer beliebigen Ausgangsfarbe generieren. Wir beschränken uns in dieser Untersuchung auf die Auswahl von drei Farben. Generalisierungen der Algorithmen auf mehr als drei Farben sind aber möglich.

Die folgende Beschreibung der Algorithmen bezieht sich auf den CieLab-verwandten Farbraum CIELCh (siehe Commission Internationale de l'Eclairage 2004). In diesem Farbraum wird eine Farbe durch *Helligkeit* (Wertebereich von 0 bis 100), *Sättigung* (Wertebereich von 0 bis 120) und *Farbton* (Werte von 0° bis 360°) eindeutig beschrieben. Die Formulierung der Algorithmen 2-5 basiert auf etablierten Vorschlägen zur ästhetischen Auswahl von Farbkombinationen (siehe z.B. Birren 1969).

Es wurden folgende Algorithmen untersucht:

1. *Zufall*: Alle drei Farben werden zufällig im Farbraum gewählt.
2. *Benachbarte Helligkeit*: Die erste Farbe wird zufällig gewählt. Die beiden anderen Farben entstehen aus dieser durch Senkung bzw. Erhöhung des Helligkeitswerts um 10.
3. *Benachbarte Sättigung*: Die erste Farbe wird zufällig gewählt. Die beiden anderen Farben entstehen aus dieser durch Senkung bzw. Erhöhung des Sättigungswerts um 12.
4. *Nachbarfarben*: Die erste Farbe wird zufällig gewählt. Die beiden anderen Farben entstehen aus dieser durch Senkung bzw. Erhöhung der Werte für Helligkeit, Sättigung und Farbton um jeweils 10% der entsprechenden Dimensionslänge (d.h. um die Werte 10, 12 bzw. 36°).
5. *Gegenfarben*: Die erste Farbe wird zufällig gewählt. Die beiden anderen Farben entstehen aus dieser durch Senkung bzw. Erhöhung des Farbtons um 144°.
6. *Designer*: Die drei Farben wurden aus einer Farbpalette gewählt, die ein erfahrener visueller Designer für das Färben von Diagrammflächen erstellt hatte.

3.1.1 Teilnehmer

Insgesamt nahmen 20 Personen (7 männlich, 13 weiblich, Durchschnittsalter 27,4 Jahre) an der Untersuchung teil. Bei der Auswahl der Teilnehmer wurde sichergestellt, dass diese nicht an einer Farbfehlsichtigkeit litten. Die Teilnehmer wurden für die Teilnahme an der Untersuchung nicht entlohnt.

3.1.2 Reizmaterial

Pro Paarvergleich wurden die von jeweils zwei Algorithmen generierten Farbtupel verglichen. Die von einem Algorithmus generierten Farben wurden dabei auf 12 Quadrate gleich-

verteilt dargeboten. Die beiden Farbmuster waren umgeben von einem komplexen unbunten Hintergrund.

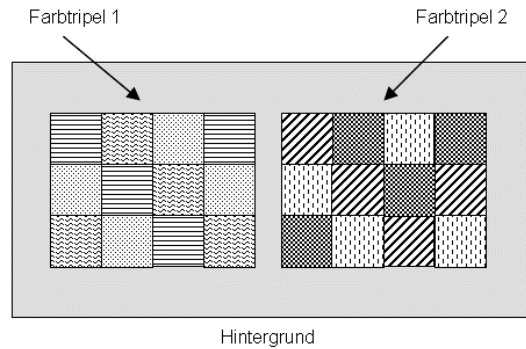


Abbildung 1: Skizze eines Paarvergleichs. Gleiche Farben sind dabei durch gleiche Schraffur angedeutet.

3.1.3 Durchführung

Die Teilnehmer wurden instruiert aus jedem dargebotenen Paar von Farbkombinationen die ihnen ansprechender bzw. schöner erscheinende auszuwählen. Ein Java-Applet, erstellt mit PXLab (Irtel 1997), zeigte jeweils zwei Farbkombinationen nebeneinander auf dem Bildschirm. Die Entscheidungen mussten über die Pfeiltasten „rechts“/„links“ auf der PC-Tastatur getätigt werden.

Da 6 verschiedenen Algorithmen skaliert werden sollten, standen 36 Reizkombinationen zur Verfügung. Im Rahmen eines vollständigen Paarvergleichs wurde jede Reizkombination 12-mal dargeboten, so dass jeder Teilnehmer 432 Vergleiche durchführen musste. Diese 432 Paarvergleiche wurden durch eine Pause unterbrochen. Die eigentliche Skalierung startete erst nach 10 Übungstrials, die nicht in die Auswertung eingingen.

3.1.4 Ergebnisse

Abbildung 2 zeigt die aus der Präferenzmatrix berechneten BTL-Skalenwerte für die 6 Algorithmen. Die Ergebnisse zeigen, dass die von einem erfahrenen Designer generierte Farbpalette bezüglich der wahrgenommenen Ästhetik präferiert wird. Der Unterschied zu Algorithmus 2 ist allerdings nicht signifikant ($p < .05$). Alle verwendeten Algorithmen zeigen signifikant höhere Skalenwerte als der Zufallsalgorithmus ($p < .05$). Die Ergebnisse zeigen, dass es durchaus möglich ist ästhetisch wirkende Farbkombinationen algorithmisch zu erzeugen. Insbesondere die Erzeugung von in der Helligkeit variierenden Farbtönen scheint hier ein guter Kandidat zu sein (siehe dazu z.B. auch Laugwitz 2001).

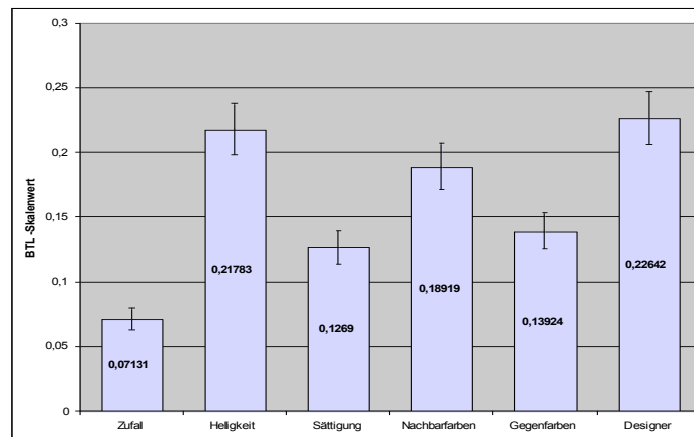


Abbildung 2: BTL Skalenwerte und 95%-Konfidenzintervalle für die 6 untersuchten Algorithmen.

Es muss in weiteren Studien untersucht werden, inwieweit sich die Ergebnisse auf Situationen übertragen lassen, bei denen eine größere Zahl von Farben benötigt wird.

3.2 Conjoint Analyse von Formular-Layouts

In Softwareanwendungen geschieht ein wesentlicher Teil der Dateneingabe über Formulare. Nutzer interagieren daher sehr oft mit Eingabefeldern. Das Layout solcher Formulare hat daher einen erheblichen Einfluss auf die Beurteilung der Anwendung durch Benutzer. In dieser Untersuchung wird der Einfluss der Verteilung der Informationen auf einem Formular auf die wahrgenommene Usability untersucht. Nach Ngo & Byrne (2001) bzw. Ngo, Samsudin & Abdullah (2000) sollten die Balance des Formular-Layouts, die Minimierung von Alignment-Punkten und eine klare Gruppierung durch Leerraum, einen positiven Einfluss auf die wahrgenommene Usability und Ästhetik eines Formulars haben. Ziel dieser Untersuchung ist es zu ermitteln, ob und wie stark diese einzelnen Gestaltungsmöglichkeiten die wahrgenommene Usability eines Eingabefeldes beeinflussen.

3.2.1 Teilnehmer

Insgesamt nahmen 42 Studenten der Universität Mannheim (10 männlich, 32 weiblich, Durchschnittsalter 21,6) an der Untersuchung teil. Die Teilnehmer wurden für die Teilnahme an der Untersuchung nicht entlohnt.

3.2.2 Reizmaterial

Für die Studie wurden 8 Varianten eines Eingabefeldes konstruiert. Diese Varianten unterschieden sich bezüglich der Balance des Layouts (gleiche oder unterschiedliche Zahl von Eingabefeldern auf der linken bzw. rechten Hälfte des Formulars), der Ausrichtung nebeneinanderstehender Gruppenboxen und der zusätzlichen Gruppierung durch eine Leerzeile zwischen Gruppen. Abbildung 2 zeigt zwei der konstruierten Varianten.

Abbildung 3 zeigt zwei Varianten eines Formulars mit dem Titel 'Kundendaten eingeben'. Die linke Variante (links) ist als 'balanciert, ausgerichtet' beschrieben und enthält Leerzeilen. Die rechte Variante (rechts) ist als 'unbalanciert, nicht ausgerichtet, keine Leerzeile' beschrieben.

Links (balanciert, ausgerichtet):

- Name:** Anrede, Titel, Vorname, Name.
- Bankverbindung:** Kreditinstitut, Bankleitzahl, Kontonummer.
- Kommunikation:** E-Mail, Telefon (gesch.), Telefon (priv.), Mobil, Fax.
- Adresse:** Strasse, Hausnummer, Postleitzahl, Stadt, Land, Region.

Rechts (unbalanciert, nicht ausgerichtet, keine Leerzeile):

- Name:** Anrede, Titel, Vorname, Name.
- Bankverbindung:** Kreditinstitut, Bankleitzahl, Kontonummer.
- Kommunikation:** E-Mail, Telefon (gesch.), Telefon (priv.), Mobil, Fax.
- Adresse:** Strasse, Hausnummer, Postleitzahl, Stadt, Land, Region.

Beide Varianten haben 'Sichern' und 'Zurück' Buttons am unteren Rand.

Abbildung 3: Balanciert, ausgerichtet, Leerzeile (links) und unbalanciert, nicht ausgerichtet, keine Leerzeile (rechts).

3.2.3 Durchführung

Die Teilnehmer wurden im Rahmen eines Ranking-Verfahrens instruiert, die 8 konstruierten Formulare in Bezug auf ihre Usability in eine Rangreihe zu bringen. Die Formulare standen den Teilnehmern dabei als Ausdrucke zur Verfügung. Jeder Teilnehmer konnte die Ausdrucke auf einem Tisch ausbreiten und vergleichen. Es bestand keine Zeitrestriktion, die Aufgabe wurde aber meist in ca. 10 Minuten erledigt.

3.2.4 Ergebnisse

Die Skalenwerte der Faktoren *Balance*, *Ausrichtung* und *Leerzeile* wurden durch eine multiple Regression der Faktoren auf die beobachteten Ränge der Alternativen ermittelt. Dabei ergaben sich folgende Skalenwerte: *Balance*: 0,22 (-0,213 – 0,652); *Ausrichtung der Gruppenboxen*: 2,171 (1,739 – 2,603); *zusätzliche Leerzeile*: 0,768 (0,336 – 1,2). Die Werte in Klammern beschreiben die entsprechenden 95%-Konfidenzintervalle der Skalenwerte.

Das Ergebnis zeigt, dass die Ausrichtung nebeneinander dargestellter Gruppierungen einen dominierenden Einfluss auf die wahrgenommene Usability eines Formulars hat. Das Einfügen einer zusätzlichen Leerzeile zwischen Gruppen hat dagegen nur einen moderaten Effekt auf die wahrgenommene Usability. Die Balance des Layouts hat offenbar nur einen schwachen bis keinen Einfluss auf die wahrgenommene Usability.

Tabelle 1 zeigt für alle 8 alternativen Formulardesigns den mittleren beobachteten Rang und den aus den Werten der Faktoren geschätzten Rang. Beide Ränge unterscheiden sich nicht signifikant ($p < .05$).

Alter-native	Balance	Ausrichtung	Leerzeile	mittlerer Rang	aus Skalenwerten berechneter Rang
1	ja	ja	ja	6,2	6,086
2	ja	nein	ja	3,93	3,915
3	nein	ja	ja	5,49	5,866
4	nein	nein	ja	3,95	3,695
5	ja	ja	nein	5,54	5,318
6	ja	nein	nein	2,8	3,147
7	nein	ja	nein	5,15	5,098
8	nein	nein	nein	3	2,927

Tabelle 1: Mittlerer Rang und aus den Skalenwerten vorhergesagter mittlerer Rang für die 8 alternativen Eingabeformulare.

4 Ausblick und Anwendbarkeit

Zentrales Anliegen dieser Arbeit ist es, zwei praktikable und etablierte Methoden vorzustellen, die es User Interface Designern ermöglichen, sowohl schnelle als auch wohl fundierte Entscheidungen zur Realisierung von Designalternativen zu treffen.

Paarvergleich und Ranking-Verfahren erfordern nur ein geringes Ausmaß an Instruktion und Training der Beurteiler. Paarvergleiche zeichnen sich dadurch aus, dass sie die kognitive Belastung der Versuchsperson sehr gering halten und die Aufgabe leicht verständlich ist. Das macht die Paarvergleichsmethode zu einem interessanten Kandidaten für die Verwendung in Online-Untersuchungen. In der Arbeit von (Fischer et al. 2006) wurde bereits gezeigt, dass die Methode hinreichend robust ist, um in einem Web-basierten Untersuchung angewandt zu werden. Ein weiterer wesentlicher Vorteil des Paarvergleichs liegt in der einfachen Möglichkeit zur Überprüfung der Datenqualität.

Das Ranking, wie wir es bei der Conjoint-Analyse angewandt haben, ist noch deutlich schneller durchzuführen und eignet sich ebenfalls für Online-Untersuchungen. Andererseits steigt die kognitive Belastung der Versuchsperson mit steigender Zahl der zu beurteilenden Objekte deutlich und es existiert auch keine Möglichkeit, die Datenkonsistenz innerhalb eines Beurteilers zuverlässig zu beurteilen. Eine im Zusammenhang mit dem Design von Benutzungsoberflächen sehr interessante Eigenschaft von Conjoint-Analysen ist die Möglichkeit, näheren Aufschluss über die Wirkung einzelner Objektattribute und deren Ausprägungen auf Präferenzurteile zu erlangen.

Weitere Untersuchungen sollen zeigen, ob sich die vorgestellten Skalierungsverfahren auch für die Beurteilung der erlebten Nutzbarkeit von Software bei konkreten vom Beurteiler durchlaufenen *Interaktionsschritten* eignen. Dies würde unter bestimmten Bedingungen eine unmittelbare Vergleichbarkeit von Skalierungsergebnissen für verschiedene Benutzungsoberflächen ermöglichen.

Literatur

- Alwin, D. F.; Krosnick, J.A. (1985): The Measurement of Values in Surveys: A Comparison Ratings and Rankings. *Public Opinion Quarterly*, Vol. 49, S. 535-552.
- Birren, F. (Hrsg.) (1969): *A grammar of color*. Van Nostrand Reinhold Co.
- Bradley, R. A.; Terry, M. E. (1952): Rank Analysis of Incomplete Block Designs: I. The Method of Paired Comparisons. *Biometrika*, Vol. 39, S. 324-345.
- Commission Internationale de l'Eclairage (2004): *Colorimetry (3rd Edition)*. In CIE Publication 15. Wien: CIE.
- Fischer, P.; Held, Th.; Schrepp, M.; Laugwitz, B. (2006): Wahrgenommene Ästhetik, Ordnung und Komplexität von Formularen. In: A. M. Heinecke, H. Paul (Hrsg.): *Mensch & Computer 2006: Mensch und Computer im StrukturWandel*. München: Oldenbourg Verlag, S. 441-443.
- Gediga, G. (1998): *Skalierung*. Münster: Lit Verlag.
- Intel, H. (1997): PXL: A Library for Psychological Experiments on IBM PC Type Computers. *Spatial Vision*, Vol. 10, S. 467-469.
- Klein, M. (2002): Die Conjoint-Analyse. Eine Einführung in das Verfahren mit einem Ausblick auf mögliche sozialwissenschaftliche Anwendungen. In: *Zentralarchiv für empirische Sozialforschung* (Hrsg.): *ZA-Information 50*. Köln: Universität zu Köln, S. 7-45.
- Laugwitz, B. (2001): Experimentelle Untersuchung von Regeln der Ästhetik von Farbkombinationen und von Effekten auf den Benutzer bei ihrer Anwendung im Benutzungsoberflächendesign. Berlin: *dissertation.de* - Verlag im Internet.
- Luce, R. D. (1959): *Individual Choice Behavior: A Theoretical Analysis*. New York: Wiley.
- Luce, R. D.; Tuckey J. (1964): Simultaneous Conjoint Measurement: A New Type of Fundamental Measurement. *Journal of Mathematical Psychology*, Vol. 1, S. 1-27.
- Ngo, D.; Byrne, J. (2001): Another Look at a Model for Evaluating Interface Aesthetics. *International Journal of Applied Mathematics and Computer Science*, Vol. 11, Nr. 2; S. 515-535.
- Ngo, D.; Samsudin, A.; Abdullah, R. (2000): Aesthetic Measures for Assessing Graphic Screens. *Journal of Information Science and Engineering*, Vol. 16, S. 97-116.
- Orth, B. (1974): *Einführung in die Theorie des Messens*. Stuttgart: Kohlhammer.
- Silberer, G. & Engelhardt, J.F. (2003). Navigationsmerkmale und ihre Präferenzwirkung bei e-Shop-Nutzern. *iCom*, Vol. 2, Nr. 3, S. 11 – 17.
- Wittink, D. R.; Cattin, P. (1981). Alternative Estimation Methods for Conjoint Analysis: A Monte Carlo Study. *Journal of Marketing Research*, Vol. 18, S.101-106.
- Wittink, D. R.; Vriens, M.; Burhene, W. (1994): Commercial Use of Conjoint Analysis in Europe: Results and Critical Reflections. *International Journal of Research in Marketing*, Vol. 11, S. 41-52.

Kontaktinformationen:

Dr. Theo Held, Dr. Martin Schrepp: {theo.held | martin.schrepp} @sap.com - Patrick Fischer: fischerp@rumms.uni-mannheim.de