

Untersuchung von Designpräferenzen mit Hilfe von Skalierungsmethoden

MARTIN SCHREPP¹, THEO HELD¹ & PATRICK FISCHER²

¹⁾ SAP AG, Walldorf

²⁾ Universität Mannheim

Schlüsselwörter: Skalierung, Ästhetik, Formularlayout

Zusammenfassung

Wir untersuchen die Anwendbarkeit zweier klassischer Skalierungsverfahren (BTL-Skalierung und Conjoint-Analyse) auf die Untersuchung ästhetisch motivierter Präferenzen zwischen Benutzungsoberflächen. Die Anwendung beider Verfahren wird jeweils an einer Studie zum Design von Eingabefeldern illustriert. Die Ergebnisse dieser beiden Studien zeigen, dass beide Skalierungsmethoden als effektive Unterstützung zur Lösung von Entscheidungsproblemen im Bereich des Oberflächendesigns dienen können.

1. Einleitung

Messtheoretisch fundierte Skalierungsverfahren sind etablierte Methoden der empirischen Wissenschaften. Sie ermöglichen die Quantifizierung abstrakter Objektqualitäten, wie z.B. des ästhetischen Eindrucks den ein Objekt beim Betrachter hervorruft. Bei Fragestellungen in Bezug auf die Gestaltung von Benutzungsoberflächen werden diese Verfahren allerdings nur selten eingesetzt. Ein aktuelles Beispiel zum Einsatz eines Skalierungsverfahrens (Conjoint Analyse) zur Ableitung von Gestaltungsempfehlungen von e-Commerce Anwendungen findet sich in Silberer & Engelhardt (2003).

Wir stellen in dieser Arbeit mit der BTL-Skalierung und der Conjoint-Analyse zwei Skalierungsverfahren vor, die sich zur Beantwortung von Fragen im Bereich der Mensch-Maschine-Interaktion eignen. Speziell zur Untersuchung ästhetischer Präferenzen sind diese Verfahren gut geeignet.

Neben Effizienz und Effektivität spielen ästhetische Aspekte von Benutzungsoberflächen eine wichtige Rolle, wenn es um die Beurteilung der Usability eines Softwareproduktes geht. Aktuelle Arbeiten zeigen, dass solche Faktoren im unmittelbaren Zusammenhang mit Usability stehen: als Indikatoren für eine tatsächlich hohe Gebrauchstauglichkeit (Produkte mit guter Usability sind häufig auch attraktiv) und als Faktoren, die die wahrgenommene Gebrauchstauglichkeit von Produkten erhöhen (siehe z.B. Hassenzahl 2001; Laugwitz 2001; Lavie & Traktinsky 2004; Schrepp, Held & Laugwitz 2006).

Um für den Praktiker anwendbare Richtlinien in Bezug auf ästhetisches Design einer Benutzungsschnittstelle zu formulieren, müssen Erkenntnisse über die Faktoren gewonnen werden, die eine Software attraktiv, ästhetisch ansprechend oder begehrenswert machen (siehe hierzu auch Mahlke 2005).

Solche Erkenntnisse können sehr einfach mit Hilfe systematischer Variation (bzw. Konstanthalten) ausgewählter Eigenschaften von Benutzungsoberflächen und anschließender Beurteilung dieser erzeugten Alternativen durch Benutzer gewonnen werden. Skalierungsmethoden erlauben es, die Präferenzen der Beurteiler bzgl. der Alternativen auf einer psychometrischen Skala zu messen. Die Skalenwerte erlauben dann eine Entscheidung zwischen den Designalternativen bzw. eine Einschätzung, wie bedeutsam die Unterschiede in Bezug auf die gemessene Eigenschaft (z.B. Ästhetik oder wahrgenommene Benutzungsfreundlichkeit) wirklich sind.

Zunächst beschreiben wir die beiden Skalierungsverfahren *BTL-Skalierung* und *Conjoint Measurement* (Verbundene Messung). Anschließend werden wir zwei Anwendungen dieser Verfahren im Bereich der ästhetischen Gestaltung von Eingabefeldern vorstellen. Im letzten Abschnitt dieser Arbeit werden wir dann die Vorteile und Beschränkungen dieser beiden Verfahren diskutieren.

2. Skalierungstechniken

Wir stellen die BTL-Skalierung und die Conjoint Analyse in dieser Arbeit nur in ihren Grundzügen dar. Eine anschauliche Darstellung des BTL-Modells findet sich z.B. in Gediga (1998). Eine übersichtliche Darstellung über die wesentlichen Aspekte der Conjoint Analyse gibt Klein (2002).

2.1 BTL-Skalierung

Die BTL-Skalierung (Bradley & Terry 1952; Luce 1959) ordnet Alternativen einer Menge $A = \{a_1, a_2, \dots, a_n\}$ Skalenwerte so zu, dass diese empirisch erhobene Präferenzdaten möglichst gut beschreiben. Die Alternativen aus A können dabei einfache physikalische Stimuli (z.B. Töne verschiedener Höhe, Farbkombinationen), Konsumprodukte, Geschmacksqualitäten (siehe dazu z.B. Lukas 1991) oder auch verschiedene Entwürfe einer Benutzerschnittstelle repräsentieren.

Die Daten für eine BTL-Skalierung werden in der Regel über einen *vollständigen Paarvergleich* erhoben. Dabei werden n Beurteilern jeweils alle Paare $(a, b) \in A \times A$ von Alternativen vorgelegt. Für jedes Paar von Alternativen (a, b) entscheidet der Beurteiler, ob er Alternative a oder Alternative b präferiert.

Die Alternativenpaare sollten dabei in zufälliger Reihenfolge dargeboten werden. Weiterhin sollten die Alternativen gleich häufig an den beiden Positionen auftauchen, um Reihenfolgeeffekte auszuschließen. Dies wird oft dadurch erreicht, dass jedes Paar zweimal mit vertauschten Positionen dargeboten wird. Aus den Daten der n Beurteiler kann dann die *Dominanzmatrix* berechnet werden. Diese listet für jedes Paar (a, b) von Alternativen auf, wie häufig a gegenüber b präferiert wurde.

Die Grundannahme des BTL-Modells ist, dass die Wahrscheinlichkeit $p(a, b)$, mit der eine Alternative a gegenüber einer Alternative b präferiert wird, nur von den Skalenwerten $\phi(a)$ und $\phi(b)$ abhängt. D.h. für beliebige Alternativen $a, b \in A$ gilt folgende Beziehung:

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

Die Skalenwerte $\phi(a)$ und $\phi(b)$ sind dabei reelle Zahlen ≥ 0 . Diese Grundannahme erlaubt es, die Skalenwerte der Alternativen aus A aus der beobachteten Dominanzmatrix des vollständigen Paarvergleichs zu schätzen.

Das BTL-Modell macht explizite und empirisch überprüfbare Annahmen über die Struktur der aus den Paarvergleichen resultierenden Daten. Sind diese Annahmen erfüllt, so kann von einer Messung auf dem Niveau einer *Verhältnisskala* ausgegangen werden.

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). Das heutige Hauptanwendungsgebiet der Conjoint Analyse ist die empirische Marktforschung (Wittink, Vriens, Burhene 1994; Klein 2002). Typische Untersuchungsobjekte sind hier Konsumgüter, die durch eine Anzahl von *Attributen* charakterisiert sind. Eine Grundannahme für die Anwendung der Conjoint Analyse ist hierbei, dass die einzelnen Attributausprägungen unabhängig voneinander variiert werden können.

Die Kaufentscheidung für ein bestimmtes Automodell hängt z.B. von Attributen wie Preis, Kraftstoffverbrauch, Komfort und Leistung ab. Jedes dieser Attribute kann unterschiedliche Ausprägungen besitzen. Zum Beispiel kann der Preis unter 15000€, zwischen 15000€ und 20000€, sowie über 20000€ liegen. Die Conjoint Analyse versucht die Frage zu klären, wie sich die Ausprägungen der Attribute auf die Präferenz von Käufern auswirken.

Vor einer Conjoint Analyse muss der Zusammenhang der Attribute modelliert werden. Wir werden uns in diesem Artikel auf die einfachste Variante eines *additiven* Zusammenhangs zwischen den numerischen Skalenwerten der Attributausprägungen beschränken. Es gibt allerdings auch Anwendungsfälle, bei denen ein *multiplikativer* Zusammenhang angenommen werden muss.

Formal geht man bei der Conjoint Analyse also von Attributmengen A_1, A_2, \dots, A_n aus. Jedes von einer Versuchsperson zu beurteilende Objekte wird durch ein n -Tupel $(x_1, \dots, x_n) \in A_1 \times A_2 \times \dots \times A_n$ von Attributausprägungen dargestellt.

Betrachten wir als einfaches Beispiel ein Objekt mit zwei Attributen A_1 und A_2 . Die Conjoint Analyse konstruiert Abbildungen $\phi_1 : A_1 \rightarrow \mathfrak{R}, \phi_2 : A_2 \rightarrow \mathfrak{R}^1$, so dass für alle $a, b \in A_1$ und alle $p, q \in A_2$ folgende Beziehung gilt:

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

Dabei ist \succ die empirische Präferenzrelation, d.h. falls $(a, p) \succ (b, q)$ wird (a, p) gegenüber (b, q) präferiert. Die Funktionen ϕ_1, ϕ_2 werden auch *Bewertungsfunktionen* genannt. Die oben skizzierte Beziehung drückt also aus, dass sich die Präferenz für Objekte vollständig aus den Bewertungen ihrer Attributausprägungen erklären lässt. Die Bewertungsfunktionen sind dabei *Intervallskalen* mit derselben Einheit (Orth, 1974). Die Skalenwerte der einzelnen Attributausprägungen können z.B. über lineare Regression bestimmt werden (Wittink & Cattin 1981).

Die Daten für eine Conjoint Analyse werden in der Regel über *ein Ranking-Verfahren* ermittelt. Allerdings ist auch der im vorhergehenden Abschnitt besprochene vollständige Paarvergleich hier einsetzbar. Bei einem Ranking-Verfahren (z.B. Alwin & Krosnick 1985) ordnet ein Beurteiler alle Alternativen entsprechend seiner Präferenz direkt an. D.h. jeder Beurteiler produziert hierbei eine Rangreihe der Alternativen. Dieses Verfahren hat den Vorteil, dass es bei einer überschaubaren Anzahl von zu beurteilenden Objekten (~ 10) sehr schnell durchzuführen ist.

3. Anwendungen auf die Untersuchung von Layoutprinzipien

In vielen Softwareanwendungen erfolgt ein wesentlicher Teil der Benutzerinteraktion über Eingabeformulare. Das Layout solcher Formulare hat daher einen starken Einfluss auf den ästhetischen Eindruck, den die Anwendung beim Benutzer hervorruft. Das Layout eines Formulars sollte natürlich bekannte Prinzipien zur visuellen Gestaltung und Gruppierung von Informationen, z.B. die Gestaltgesetze, berücksichtigen. Allerdings gibt es auch dann noch viele Freiheitsgrade, z.B.:

- *Anordnung der Feldbezeichner*: Die Feldbezeichner können z.B. linksbündig, rechtsbündig oder über dem Eingabefeld angeordnet werden.
- *Visualisierung der Gruppierung*: Logische Gruppierungen von Feldern sollen visuell leicht zu erkennen sein. Bekannte Gruppierungsmechanismen sind z.B. die farbliche Hinterlegung der Gruppe, Gruppenboxen, oder Einfügen von Leerraum um die Gruppe.
- *Balance des Layouts*: Die Anordnung der Gruppen auf dem Formular sollte die Reihenfolge berücksichtigen, in der Benutzer die Felder in diesen Gruppen ausfüllen. Man hat aber oft noch Freiheitsgrade bei der Anordnung der Gruppen. Ein bekanntes Gestaltungsprinzip ist es, auf eine ausgewogene Verteilung der Formularelemente (Balance) zu achten.

Es ist bisher wenig darüber bekannt, wie sich diese Gestaltungsmöglichkeiten auf die wahrgenommene Ästhetik bzw. Benutzungsfreundlichkeit eines Formulars auswirken. Das liegt an der Vielzahl von vorhandenen Gestaltungsmöglichkeiten und an möglichen Interaktionen zwischen diesen Gestaltungsmöglichkeiten. Eine ästhetische

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

Gestaltung ist von Bedeutung, da Personen offenbar von der Ästhetik einer Eingabemaske auf deren Gebrauchstauglichkeit schließen².

3.1 Beispiel 1: BTL-Skalierung von Formular-Layouts

Wir untersuchten in dieser Studie, wie sich die Anordnung der Feldbezeichner, die Gestaltung von Gruppenboxen und die Ausrichtung der Eingabefelder auf die wahrgenommene Ästhetik, Ordnung und Komplexität eines Formulars auswirken.

Ein weiteres Ziel der Untersuchung war es zu überprüfen, ob sich das von Birkhoff (1933) postulierte *Ästhetische Maß* (Aesthetic Measure) auf die Gestaltung von Formularen anwenden lässt. Der Kern von Birkhoffs Theorie lässt sich in der Formel $M = O / C$ ausdrücken. Dabei ist M der ästhetische Wert, O die Ordnung und C die Komplexität eines Reizes. Die Ästhetik eines Reizes sollte also mit zunehmender Ordnung zunehmen und mit zunehmender Komplexität abnehmen³.

3.1.1 Teilnehmer

Die Studie wurde als Online-Experiment durchgeführt. Die 28 Teilnehmer wurden über verschiedene Foren und Mailinglisten im Internet angeworben⁴. Eine Belohnung für die Teilnahme am Experiment wurde nicht angeboten.

3.1.2 Reizmaterial

Für die Studie wurden 8 Varianten eines Formulars konstruiert, die sich bezüglich der Ausrichtung der Felder (Flattersatz vs. Blocksatz), der Anordnung der Feldbezeichner (linksbündig vs. rechtsbündig) und der Gestaltung der Feldgruppierungen (farbliche Hinterlegung) unterschieden. Abbildung 1 zeigt zwei dieser Varianten.

Abbildung 1 zeigt zwei Varianten eines Formulars. Die linke Variante zeigt ein Formular mit linksbündigen Feldbezeichnern und Feldern im Blocksatz, ohne Gruppenboxen. Die rechte Variante zeigt ein Formular mit rechtsbündigen Feldbezeichnern und Feldern im Flattersatz, mit Gruppenboxen.

Abbildung 1: Feldbezeichner linksbündig, Felder im Blocksatz, ohne Gruppenbox (links) und Feldbezeichner rechtsbündig, Felder im Flattersatz, mit Gruppenbox (rechts).

3.1.3 Durchführung

Jeder Teilnehmer führte einen vollständigen Paarvergleich aller 8 Varianten des Formulars durch. Paare mit identischen Formularen wurden nicht gezeigt. Jedes Paar wurde zwei Mal präsentiert, wobei die Anordnung der Formulare getauscht wurde. Damit ergaben sich pro Teilnehmer 56 Paarvergleiche.

² Beispielsweise berichtet Tractinsky (1997) hohe Korrelationen zwischen der ästhetischen Wahrnehmung und der eingeschätzten Benutzungsfreundlichkeit einer einfachen Eingabemaske.

³ Es gibt eine Reihe von widersprüchlichen Befunden zur Gültigkeit von Birkhoffs Theorie mit verschiedensten Reizmaterialien, z.B. einfachen geometrischen Figuren (Davis 1936) oder Farbkombinationen (Moon & Spencer 1944; Granger 1955).

⁴ Die Daten von 3 Teilnehmern konnten wegen Antworttendenzen bzw. inkonsistentem Antwortverhalten nicht ausgewertet werden.

Jedem Teilnehmer wurde zufällig eine der folgenden Fragestellungen zugewiesen:

- *Welches Formular ist ansprechender bzw. schöner?*
- *Welches Formular ist ordentlicher?*
- *Welches Formular ist komplexer?*

Ein Java-Applet, erstellt mit PXLab (Irtel 1997), zeigte jeweils zwei Formulare nebeneinander auf dem Bildschirm. Die Entscheidungen mussten über die Pfeiltasten „rechts“/„links“ auf der PC-Tastatur getätigt werden.

3.1.4 Ergebnisse

Die Versuchspersonen konnten die vorgegebenen Reize im Paarvergleich hinreichend gut anordnen (geringe Häufigkeit von Transitivitätsverletzungen). Tabelle 1 zeigt die aus der Präferenzmatrix berechneten BTL Skalenwerte der Alternativen.

Tabelle 1: Die Skalenwerte der Formulare in Bezug auf Ästhetik, Ordnung und Komplexität. Die Werte in Klammern sind die Ränge der Alternativen in Bezug auf die Skala.

Alternative	Gruppierung	Feldbezeichner	Felder	BTL Skalenwerte		
				Ästhetik (N=9)	Ordnung (N=9)	Komplexität (N=7)
1	Box	Links	Blocksatz	0,132 (3)	0,126 (2)	0,07 (8)
2	Box	Links	Flattersatz	0,117 (4)	0,017 (6)	0,111 (5)
3	Box	Rechts	Blocksatz	0,169 (2)	0,373 (1)	0,079 (7)
4	Box	Rechts	Flattersatz	0,188 (1)	0,044 (4)	0,144 (3)
5	keine Box	Links	Blocksatz	0,026 (8)	0,03 (5)	0,081 (6)
6	keine Box	Links	Flattersatz	0,027 (7)	0,003 (8)	0,162 (2)
7	keine Box	Rechts	Blocksatz	0,041 (5)	0,064 (3)	0,115 (4)
8	keine Box	Rechts	Flattersatz	0,033 (6)	0,005 (7)	0,204 (1)

Die dominierende Eigenschaft bzgl. der ästhetischen Wahrnehmung der Formularvarianten ist offensichtlich die Verwendung von farblich hinterlegten Gruppenboxen. Alle Alternativen mit dieser Eigenschaft erzielten deutlich bessere Bewertungen als die Alternativen ohne diese Eigenschaft. Die Alternativen mit Gruppenbox variieren maximal um den Faktor 1,61 bzgl. ihrer ästhetischen Bewertung. Dagegen unterscheidet sich die beste Alternative mit Gruppenbox (Alt. 4) um den Faktor 4,59 von der besten Alternative ohne Gruppenbox (Alt. 7). Rechtsbündige Feldbezeichner haben ebenfalls einen positiven Einfluss auf die Einschätzung der wahrgenommenen Ästhetik, der Einfluss dieser Eigenschaft ist aber deutlich geringer.

Die dominierende Eigenschaft in Bezug auf die wahrgenommene Ordnung ist die Ausrichtung der Felder. Die bzgl. der Ordnung am besten bewertete Alternative (Alt. 3) mit Blocksatz unterscheidet sich von der am besten bewerteten Alternative ohne Blocksatz (Alt. 4) um den Faktor 8,48.

In Bezug auf Birkhoff's konkrete Formel $M = O / C$ lässt sich sagen, dass die Daten diese Hypothese nicht unterstützen. Zum Beispiel weist die Alternative 4 mit der höchsten Bewertung in Bezug auf ihre Ästhetik nur mittlere Werte in Bezug auf wahrgenommene Ordnung und wahrgenommene Komplexität auf. Allerdings zeigt

eine Betrachtung der Rangkorrelation⁵ der Skalen, daß die Ästhetik mit wachsender Ordnung zunimmt und mit wachsender Komplexität abnimmt.

3.2 Beispiel 2: Conjoint Analyse von Formular-Layouts

Arbeiten Ngo & Byrne (2001) bzw. Ngo, Samsudin & Abdullah (2000) legen nahe, dass die Balance des Formular-Layouts, die Minimierung von Ausrichtungspunkten und eine Gruppierung durch vertikalen Leerraum, die wahrgenommene Usability und Ästhetik eines Formulars positiv beeinflussen sollten. Wir untersuchen in dieser Arbeit den Einfluss dieser drei Faktoren auf die wahrgenommene Usability.

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

Es wurden acht Varianten eines Formulars konstruiert, die 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 unterscheiden. Abbildung 2 zeigt zwei dieser Varianten.

Abbildung 2 zeigt zwei Varianten eines Formulars zur Kundendateneingabe. Die linke Variante ist unbalanciert und nicht ausgerichtet, die rechte Variante ist balanciert und ausgerichtet.

Links (unbalanciert, nicht ausgerichtet, keine Leerzeile):

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

Rechts (balanciert, ausgerichtet, Leerzeile):

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

Abbildung 2: Unbalanciert, nicht ausgerichtet, keine Leerzeile (links) und balanciert, ausgerichtet, Leerzeile (rechts).

3.2.3 Durchführung

Die Teilnehmer wurden gebeten, die 8 Formularvarianten in Bezug auf ihre wahrgenommene Benutzungsfreundlichkeit in eine Rangreihe zu bringen. Jede der Varianten lag den Teilnehmern als separater Ausdruck vor. Die Teilnehmer konnten die Ausdrücke vor sich ausbreiten, vergleichen und anordnen. Die Durchführungszeit war nicht limitiert. In fast allen Fällen beendeten die Teilnehmer die Aufgabe in weniger als 10 Minuten.

⁵ Ästhetik und Ordnung zeigen eine hohe positive Rangkorrelation (0,64), Ästhetik und Komplexität eine negative Rangkorrelation im mittleren Bereich (-0,31) und Ordnung und Komplexität eine stark negative Rangkorrelation (-0,76).

3.2.4 Ergebnisse

Über eine multiple Regression wurden folgende Skalenwerte für die Faktoren *Balance*, *Ausrichtung* und *Leerzeile* ermittelt (Werte in eckigen Klammern beschreiben die 95%-Konfidenzintervalle): Balance: 0,22 [-0,213 – 0,652], Ausrichtung: 2,171 [1,739 – 2,603], Leerzeile: 0,768 [0,336 – 1,2]. Diese Skalenwerte zeigen, dass sich die von Ngo & Byrne (2001) postulierten Faktoren positiv auf die wahrgenommene Benutzungsfreundlichkeit von Formularen auswirken.

Den bedeutsamsten Einfluss auf die wahrgenommene Benutzungsfreundlichkeit hat die Ausrichtung nebeneinander dargestellter Gruppierungen. Der Effekt der zusätzlichen Leerzeile zwischen Gruppierungen ist dagegen moderat. Letztendlich hat die Balance des Layouts offenbar einen vernachlässigbaren Einfluss auf die wahrgenommene Usability. Allerdings ist hier zu beachten, dass die Balance nur gering variierte (7 versus 11 Felder auf der linken und rechten Seite des Formulars). Es ist anzunehmen, dass der Einfluss dieses Faktors bei größerem Ungleichgewicht steigt. Dies muss in weiteren Untersuchungen geklärt werden.

In Tabelle 2 finden sich für alle 8 Designalternativen der mittlere beobachtete Rang und der aus den Werten der Faktoren geschätzten Rang. Beide Ränge unterscheiden sich jeweils nicht signifikant ($p < .05$). Die aus dem Conjoint Measurement hervorgehende Skala beschreibt die beobachtete Rangreihung also gut.

Tabelle 2: Mittlerer Rang und aus den Skalenwerten vorhergesagter mittlerer Rang für die 8 alternativen Eingabeformulare. Die jeweils beste Alternative hatte dabei den höchsten Rang, d.h. höhere Ränge stehen für eine bessere Bewertung.

Alternative	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

4. Zusammenfassung

Wir haben in dieser Arbeit zwei Skalierungsverfahren vorgestellt, die sich sehr gut eignen, Erkenntnisse im Hinblick auf ästhetisch motivierte Präferenzen von Designalternativen zu erfassen. Wir wollen nun abschließend noch die Vorteile, aber auch die existierenden Beschränkungen dieser Verfahren, diskutieren.

Ein Nachteil der beschriebenen Skalierungsverfahren gegenüber klassischen Rating-Verfahren ist sicherlich, dass der Durchführung eines Paarvergleichs aber auch eines Ranking-Verfahrens zeitlich aufwändiger ist. Insbesondere beim vollständigen Paarvergleich steigt dieser Aufwand mit der Zahl der Alternativen natürlich stark an. Für die Skalierung einer sehr großen Zahl von Alternativen ist dieses Verfahren daher

nur eingeschränkt geeignet⁶. Für die Anwendung eines Ranking-Verfahrens ist die Zahl der Alternativen nicht ganz so kritisch, aber auch hier steigen die kognitive Belastung und der zeitliche Aufwand mit der Zahl der Alternativen stark an.

Dem höheren Aufwand bei der Durchführung steht allerdings gegenüber, dass die beschriebenen Skalierungsmethoden Messwerte auf hohem Skalenniveau generieren. Unterschiede zwischen Alternativen können auf einer Verhältnisskala gemessen werden. D.h. der Einfluss bestimmter Eigenschaften eines Designs auf den dadurch erzeugten ästhetischen Eindruck beim Benutzer kann damit sehr genau bestimmt werden. Dies kann insbesondere bei Fragen relevant werden, bei denen eine Abwägung zwischen verschiedenen Faktoren wichtig ist. Zum Beispiel stellt sich bei der Gestaltung von Web-Anwendung oft die Frage, ob der durch Verwendung eines komplexeren Style-Sheets verbesserte ästhetische Eindruck die dadurch entstehenden Verschlechterungen bei der zum Seitenaufbau notwendigen Zeit aufwiegt. Solche Fragen können mit Hilfe der beschriebenen Skalierungsmethoden objektiv untersucht werden. Selbstverständlich entfallen auch die typischen Urteilsfehler, die Rating-Verfahren mit sich bringen. Exemplarisch seien hier die Tendenz zur Mitte, Primacy-Recency-Effekt und Milde-Härtefehler genannt.

Bei der Interpretation der durch die BTL-Skalierung bzw. die Conjoint-Analyse erzeugten Skalenwerte ist zu beachten, dass der Wert einer Alternative nur relativ zu den Werten der anderen Alternativen interpretiert werden kann. D.h. aus dem Skalenwert kann abgeleitet werden, welche Alternativen besser als andere sind und wie groß der Unterschied ist. Es ist also möglich, die ermittelten Skalenwerte unmittelbar in ein quantitatives Verhältnis zu setzen (z.B. „Alternative A wird als doppelt so ästhetisch wie Alternative B beurteilt“). Natürlich ist es nicht zulässig, aus den Skalenwerten eine Beurteilung der absoluten Qualität einer Alternative abzuleiten, was für praktische Anwendungen in der Regel aber auch nicht notwendig ist.

Unsere beiden Studien zeigen, dass sich etablierte Skalierungsmethoden gut eignen, um Präferenzordnungen systematisch variiertes Designalternativen herzustellen. BTL-Skalierung und vollständiger Paarvergleich sind insbesondere die Methoden der Wahl, wenn untersucht werden soll, ob Designalternativen von Beurteilern zuverlässig und homogen angeordnet werden können. Vorteile der indirekten Skalierung liegen darin, dass widersprüchliche und inkonsistente Urteile leicht identifiziert werden können und dass die Aufgabe, die die Versuchsperson zu bewältigen hat, nur mit einer sehr geringen kognitiven Belastung einhergeht.

Die Kombination Conjoint Analyse und direktes Ranking zeichnet sich durch eine hohe Effizienz in der Durchführung aus und ist insbesondere dann relevant, wenn untersucht werden soll, wie stark sich die variierten Eigenschaften der Designs jeweils auf das Urteil auswirken. Wie bei allen Direktskalierungsverfahren sind die Ansatzpunkte für eine Identifikation inkonsistenten Urteilsverhaltens eher gering.

⁶ Es ist allerdings möglich, dies über eine hohe Zahl von Versuchspersonen zu kompensieren. In diesem Fall führt eine Person keinen vollständigen Paarvergleich durch, sondern vergleicht nur eine geringe Anzahl zufällig gewählter Paare. Aus den Daten ausreichend vieler Personen, lässt sich dann eine für die Skalierung verwendbare Dominanzmatrix herleiten (Lukas, 1991).

Literatur

- Alwin, D. F. & Krosnick, J.A. (1985). The Measurement of Values in Surveys: A Comparison Ratings and Rankings. *Public Opinion Quarterly*, 49, 535–552.
- Birkhoff G. D. (1933). *Aesthetic Measure*. Cambridge, MA: Harvard University Press.
- Bradley, R. A. & Terry, M. E. (1952). Rank Analysis of Incomplete Block Designs: I. The Method of Paired Comparisons. *Biometrika*, 39, 324–345.
- Davis, R.C. (1936). An evaluation and test of Birkhoff's aesthetic measurement formula. *Journal of General Psychology*, 15, 231–240.
- Granger, G. W. (1955). Aesthetic Measure applied to colour harmony: An experimental test. *Journal of General Psychology*, 52, 205–212.
- Gediga, G. (1998). *Skalierung*. Münster: Lit Verlag.
- Hassenzahl, M. (2001). The effect of perceived hedonic quality on product appeal- ingness. *International Journal of Human-Computer Interaction*, 13, 481–499.
- Irtel, H. (1997). PXL: A Library for Psychological Experiments on IBM PC Type Computers. *Spatial Vision*, 10, 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, 7–45. Köln: Universität zu Köln.
- Laugwitz, B., Schrepp, M. & Held, T. (2006). Konstruktion eines Fragebogens zur Messung der User Experience von Softwareprodukten. In: A.M. Heinecke & H. Paul (Hrsg.): *Mensch & Computer 2006*, 125–134. München: Oldenbourg Verlag.
- Lavie, T. & Tractinsky, N. (2004). Assessing dimensions of perceived visual aesthetics of web sites. *International Journal of Human-Computer Studies*, 60, 269–298.
- 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.
- 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*, 1, 1–27.
- Lukas, J. (1991). BTL-Skalierung verschiedener Geschmacksqualitäten von Sekt. *Zeitschrift für experimentelle und angewandte Psychologie*, 38, 605–619.
- Mahlke, S. (2005). An integrative model on web user experience. In P. Isaiás & M. B. Nunes (Hrsg.), *Proceeding of ICWI2005 (Vol. 2)*, 91–95. Lisbon, Portugal: IADIS.

- Moon P. & Spencer D.E. (1944). Aesthetic Measure applied to colour harmony. *Journal of the Optical Society of America*, 34, 234–242.
- Ngo, D. & Byrne, J. (2001). Another Look at a Model for Evaluating Interface Aesthetics. *International Journal of Applied Mathematics and Computer Science*, 11, 515–535.
- Ngo, D., Samsudin, A. & Abdullah, R. (2000). Aesthetic Measures for Assessing Graphic Screens. *Journal of Information Science and Engineering*, 16, 97–116.
- Orth, B. (1974). *Einführung in die Theorie des Messens*. Stuttgart: Kohlhammer.
- Schrepp, M., Held, T. & Laugwitz, B. (2006). On the influence of hedonic quality on the attractiveness of user interfaces of business management software. *Interacting with Computers*, 18, 1055–1069.
- Silberer, G. & Engelhardt, J.F. (2003). Navigationsmerkmale und ihre Präferenzwirkung bei e-Shop-Nutzern. *i-com*, 2, (3) 11–17.
- Tractinski, N. (1997). Aesthetics and Apparent Usability: Empirically Assessing Cultural and Methodological Issues. In: *CHI '97 Proceedings*, 115–122. New York: ACM Press.
- Wittink, D. R. & Cattin, P. (1981). Alternative Estimation Methods for Conjoint Analysis: A Monte Carlo Study. *Journal of Marketing Research*, 18, 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*, 11, 41–52.