

AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität¹

Marc Hassenzahl
Technische Universität
Darmstadt

Michael Burmester
Hochschule der Medien
Stuttgart

Franz Koller
User Interface Design GmbH
Ludwigsburg

Zusammenfassung

Die Evaluation interaktiver Produkte ist eine wichtige Aktivität im Rahmen benutzerzentrierter Gestaltung. Eine Evaluationstechnik, die sich meist auf die Nutzungsqualität oder „Gebrauchstauglichkeit“ eines Produkts konzentriert, stellen Fragebögen dar. Zur Zeit werden allerdings weitere, sogenannte „hedonische“ Qualitätsaspekte diskutiert. Diese beruhen auf den menschlichen Bedürfnissen nach Stimulation und Identität, während bei Gebrauchstauglichkeit (bzw. „pragmatischer Qualität“) der Bedarf zur kontrollierten Manipulation der Umwelt im Vordergrund steht. In diesem Beitrag wird der „AttrakDiff 2“ Fragebogen vorgestellt, der sowohl wahrgenommene pragmatische als auch hedonische Qualität zu messen vermag. Ergebnisse zur Reliabilität und Validität werden vorgestellt und diskutiert. AttrakDiff 2 stellt einen ersten Beitrag zur Messung von Qualitätsaspekten dar, die über die reine Gebrauchstauglichkeit hinausgehen.

1 Einleitung

Menschen besitzen und benutzen interaktive Produkte, um – im weitesten Sinne – ihre Umwelt zu manipulieren. Das Produkt muss eine angemessene Funktionalität bieten (Nützlichkeit) und die Bedienbarkeit dieser Funktionalität sicher stellen (Benutzbarkeit). Ist beides gegeben, spricht man von Gebrauchstauglichkeit im Sinne der DIN EN ISO 9241-11. Gebrauchstauglichkeit ist mittlerweile ein anerkanntes und breit gefordertes Qualitätsmerkmal, das die Ansprüche der Benutzer auf effektive und effiziente Zielerreichung ohne psychische Belastung betont. Gebrauchstauglichkeit als alleinige Qualitätsanforderung zu verstehen ist allerdings eine eingeschränkte Sicht (Hassenzahl, Platz, Burmester & Lehner 2000; Burmester, Hassenzahl & Koller 2002). Personen verbinden mit einem Produkt auch noch die Bedürfnisse Stimulation und Identität² (Hassenzahl 2003).

- *Stimulation*: Menschen streben nach persönlicher Entwicklung, d.h. der Verbesserung von Kenntnissen und Fertigkeiten. Produkte können diese Entwicklung unterstützen, in dem sie stimulierend wirken. Neuartige, interessante und anregende Funktionalitäten, Inhalte, Interaktions- und Präsentationsstile können die Aufmerksamkeit erhöhen, Motivationsproble-

¹ Die diesem Beitrag zugrunde liegenden Arbeiten entstanden zum Teil im Verbundprojekt INVITE und wurden mit Mitteln des Bundesministeriums für Bildung, Wissenschaft, Forschung und Technologie (BMBF) unter dem Förderkennzeichen 01 IL 901 V 8 gefördert. Siehe auch www.attrakdiff.de.

² In Hassenzahl (2003) wird noch ein drittes Bedürfnis diskutiert: Symbolisieren (evocation). Allerdings spielt dieses Bedürfnis im Kontext interaktiver Produkte nur eine untergeordnete Rolle und soll aus diesem Grund hier nicht weiter vertieft werden.

me dämpfen oder das Finden neuer Lösungen für bestehende Probleme erleichtern. So kann Stimulation auch indirekt bei der Aufgabenerledigung helfen.

- *Identität*: Menschen bringen durch Objekte auch ihr Selbst zum Ausdruck (Prentice 1987). Sie wollen von relevanten Anderen in einer spezifischen Weise wahrgenommen werden. Ein Produkt kann dies unterstützen, indem es eine gewünschte Identität kommuniziert.

Ist ein interaktives Produkt zur Manipulation der Umwelt geeignet, und wird auch von seinen Benutzern so wahrgenommen, besitzt es „pragmatische“ Qualität. Erweitert ein interaktives Produkt hingegen durch neue Funktionen die Möglichkeiten des Benutzers, stellt neue Herausforderungen, stimuliert durch visuelle Gestaltung und neuartige Interaktionsformen oder kommuniziert eine gewünschte Identität (z.B., indem es professionell, cool, modern, anders wirkt) besitzt es „hedonische“ Qualität.

Gängige Prinzipien, Regeln und Methoden des Usability Engineerings bzw. der Software-Ergonomie betonen meist einseitig pragmatische Qualität. Sicher ist Stimulation oder Identität nicht für alle interaktiven Produkte gleich wichtig. Es ist beispielsweise fraglich, ob Menschen durch die erfolgreiche Bedienung eines „anspruchsvoll“ gestalteten Bankautomaten stimuliert werden wollen. Hier ist die Philosophie des „Weniger ist Mehr“ sicher angemessen. Allerdings sollte man auch vorsichtig sein, diese Philosophie ungeprüft auf alle interaktiven Produkte zu übertragen. Auch ein rundherum gebrauchstaugliches Produkt kann an den tatsächlichen Bedürfnissen der Benutzer (z.B. Identität kommunizieren) vorbei gestaltet sein. Wir glauben, dass hedonische Qualität einen wichtigen, bisher weitestgehend unberücksichtigten Aspekt darstellt. Durch ihr Quantifizieren wird ein umfassenderer Blick auf interaktive Produkte möglich. Wie zentral dies ist, zeigen die momentan mannigfaltigen Versuche, den Begriff „Gebrauchstauglichkeit“ um nicht-utilitaristische Konzepte, wie z.B. Spaß (Draper 1999) oder Nutzungsfreude (Hatscher 2001) zu erweitern (vgl. Burmester, Hassenzahl & Koller 2002). Das ganzheitliche Berücksichtigen sowohl pragmatischer als auch hedonischer Bedürfnisse ist besonders dann wichtig, wenn sich das Selbstverständnis des Software-Ergonomen bzw. Usability Engineers von einem Spezialisten zu einem ganzheitlich denkenden Gestalter (vgl. Winograd 1996) wandeln soll.

Ansatzpunkte, wie hedonische Qualität systematisch bei der Gestaltung interaktiver Produkte berücksichtigt werden kann, liegen zur Zeit kaum vor. Im Rahmen benutzerzentrierter Gestaltung werden Verfahren zur Erhebung hedonischer Anforderungen bei der Nutzungskontextanalyse, Prinzipien hedonischer Gestaltung während der Entwurfsphase und Verfahren zur Erfassung der erreichten hedonischen Qualität eines interaktiven Produktes benötigt. Der vorliegende Beitrag stellt den Fragebogen AttrakDiff 2 vor. Er ermöglicht die Bewertung interaktiver Produkte hinsichtlich ihrer pragmatischen und hedonischen Qualität.

2 Grundannahmen und Vorarbeiten

Eine goldene Regel der benutzerzentrierten Gestaltung ist die empirische Überprüfung der resultierenden Produkte bzw. Prototypen (vgl. DIN EN ISO 13407). Dazu werden auch Fragebogen zur Bewertung der Produkte durch (zukünftige) Benutzer eingesetzt (vgl. Gediga & Hamborg 2002). Im Rahmen des Marketings und der Konsumentenpsychologie liegen bereits Fragebogen zur Messung von wahrgenommenen utilitaristischen (d.h., instrumentellen, funktionellen) und hedonischen (d.h., anregenden, erlebnisorientierten) Produktmerkmalen vor (z.B. Batra & Ahtola 1990). Allerdings eignen sich diese Ansätze nur bedingt zur Einschätzung interaktiver Produkte. Jordan (2000, 156ff) hat zwar einen Fragebogen zur Messung genereller „Freude“ mit einem Produkt vorgestellt. Dieser differenziert das Konstrukt aber nicht weiter. Diese insgesamt unbe-

riedigende Lage stellt den Ausgangspunkt für die Entwicklung des hier vorgestellten Fragebogens AttrakDiff 2 dar.

Zu einem Fragebogen gehört ein zugrunde liegendes Modell. Wie schon oben ausgeführt unterscheiden wir im Hinblick auf Software pragmatische (Manipulation) und hedonische Qualität (Stimulation und Identität). Diese Qualitäten sind subjektiv, d.h. jeder Befragte schätzt für sich persönlich ein, ob das Produkt seine Bedürfnisse befriedigt. Weiterhin sind diese Qualitäten unabhängig voneinander. Produkte, die als pragmatisch bewertet werden, werden nicht automatisch auch als hedonisch bewertet. Allerdings sind natürlich Produkte denkbar, die gleichzeitig sowohl als pragmatisch, als auch als hedonisch bewertet werden.

Aus der Kombination von hedonischen und pragmatischen Qualitäten können sich verschiedene Produktcharaktere ergeben. Erwünscht ist, aus unserer Sicht, ein Produkt bei dem *beide* Qualitäten stark ausgeprägt sind. Von ihm erhofft man sich, dass es nicht nur zufrieden stellt, sondern sogar Freude bei seinen Benutzern auslöst. Unerwünscht sind Produkte bei denen beide Qualitäten nur schwach ausgeprägt sind. Häufig findet man allerdings Softwareprodukte, die entweder schwach hedonisch und stark pragmatisch oder aber stark hedonisch und schwach pragmatisch sind. Ersteres ist ein handlungsorientiertes Produkt (act-product), letzteres ein selbstorientiertes Produkt (self-product) (Hassenzahl 2003). Handlungsorientierte Produkte sind effektive und effiziente Werkzeuge, allerdings geht der Benutzer keine starke Bindung mit ihnen ein. Benutzt man es erfolgreich, stellt sich Zufriedenheit als emotionale Reaktion ein. Selbstorientierte Produkte hingegen binden den Benutzer stärker, denn selbstbezogene Ziele sind meist persistenter und persönlich relevanter. Die emotionale Konsequenz eines selbstorientierten Produkts – Freude – ist ebenfalls stärker. Ob ein handlungsorientiertes oder ein selbstorientiertes Produkt das „Bessere“ ist, hängt von den Vorstellungen des Herstellers und der gewünschten Marktpositionierung ab.

Wir trennen die wahrgenommene pragmatische und hedonische Qualität von der Attraktivität eines Produktes. Das Attraktivitätsurteil („gut“, „sympathisch“, „motivierend“) ist eine globale Bewertung auf der Basis der wahrgenommenen Qualitäten. Es wird angenommen, dass die Wahrnehmung eines Produktes als pragmatisch oder hedonisch über verschiedene Situationen hinweg relativ stabil bleibt, während sich die globale Bewertung durchaus verändern kann (Hassenzahl, Kekez & Burmester 2002).

Das beschriebene Modell wurde von Hassenzahl und Kollegen in mehreren Studien untersucht (z.B. Hassenzahl, Platz, Burmester & Lehner 2000; Hassenzahl 2002). Zur Messung wurde das AttrakDiff 1 verwendet, ein eigens konstruierter Fragebogen im Format eines semantischen Differenzials. Er besteht aus 23 siebenstufigen Items, deren Endpunkte jeweils durch ein gegensätzliches Adjektiv gebildet werden (z.B. „verwirrend – übersichtlich“, „außergewöhnlich – üblich“, „gut – schlecht“). Jeweils mehrere Items werden zu einer Skala zusammengefasst. Der Mittelwert der Items bildet den Skalenwert für pragmatische Qualität (PQ), hedonische Qualität (HQ) und Attraktivität (ATT). Die zwei Studien zeigten, dass hedonische und pragmatische Qualitäten konsistente und unabhängig voneinander wahrgenommene Qualitäten sind. Beide trugen gleich stark zu dem Attraktivitätsurteil bei. Es zeigte sich auch, dass pragmatische Qualität signifikant mit einem Anstrengungsmaß korreliert. Je anstrengender die Aufgabenbearbeitung erlebt wurde, desto niedriger war die wahrgenommene pragmatische Qualität. Hedonische Qualität korrelierte, wie erwartet, nicht mit der Anstrengung. Sie ist eine nicht-aufgabenorientierte Qualität, deren Wahrnehmung durch Anstrengung nicht beeinflusst wird. Neben den Versuchen, die Annahmen des oben beschriebenen Modells zu überprüfen, wurde das AttrakDiff 1 auch schon mehrfach als Evaluationsmaßnahme im Rahmen benutzerzentrierter Produktentwicklung eingesetzt (z.B. Sandweg, Hassenzahl & Kuhn 2000; Kunze 2001).

Die bisherigen Ergebnisse sind vielversprechend. Allerdings hat der ursprüngliche Fragebogen – AttrakDiff 1 – einen entscheidenden Nachteil. Er vermag es nicht, die beiden Aspekte aus denen sich hedonische Qualität zusammensetzt, nämlich Stimulation und Identität, getrennt zu quantifi-

zieren. Ursprünglich wurden diese Aspekte gemeinsam als nicht-aufgabenbezogene Qualität definiert. Dementsprechend wenig Wert wurde auf die Trennschärfe der verwendeten einzelnen Adjektivpaare gelegt. Im Laufe der Erprobung und Anwendung des Fragebogens wurde allerdings deutlich, dass eine Trennung der beiden Aspekte wünschenswert wäre. Um dies zu verwirklichen, wurde ein neuer Fragebogen, das AttrakDiff 2, entwickelt und erprobt, welchem der Rest des vorliegenden Artikels gewidmet ist.

3 Konstruktion und Erprobung des AttrakDiff 2

3.1 Sammlung der Items

In einem ca. fünfstündigen Expertenworkshop (sechs Software-Ergonomen und ein Moderator) wurden zunächst in einer Kreativphase mögliche Items („Adjektivpaare“) gesammelt und dann in einer Bewertungsphase zur Aufnahme in eine Ausgangsversion des neuen Fragebogens ausgewählt. Als Einleitung wurden das zugrunde liegende, oben beschriebene Modell und die Konstrukte vorgestellt.

In der *Kreativitätsphase* wurde jeder Teilnehmer zunächst gebeten, so viele Items wie möglich zu generieren. Dabei wurde jeder Aspekt (Manipulation, Stimulation, Identität) getrennt behandelt. Jeder Teilnehmer las dann seine Items der Gruppe vor. Dies wirkte wiederum als Anregung für neue oder umformulierte Items. In der *Bewertungsphase* wurden alle Items vom Moderator erneut vorgelesen. Jeder Teilnehmer wurde gebeten sein Veto einzulegen, wenn das Item als problematisch empfunden wurde (z.B. zu technisch, Umgangssprache etc.). Alle Probleme wurden diskutiert und protokolliert. Von den 133 Items aus der Kreativitätsphase erhielten 50 kein Veto. Diese 50 Items wurden mit den sieben Items zur Messung der Attraktivität aus dem AttrakDiff 1 kombiniert und bildeten so die Ausgangsversion des neuen Fragebogens. Dabei wurde die Reihenfolge und Polarität der Items zufällig gewählt. Die weitere Konstruktion des AttrakDiff 2 erfolgte empirisch im Rahmen einer Pilotanwendung.

3.2 Pilotanwendung

Es nahmen 22 Personen (9 Frauen, 13 Männer) an der Untersuchung teil. Alle Teilnehmer wurden über Anzeigen in lokalen Zeitungen geworben und erhielten für ihre Teilnahme finanzielle Kompensation. Das mittlere Alter betrug rund 38 Jahre (Minimum 23, Maximum 59 Jahre).

Als Bewertungsobjekte dienten drei Websites³ (Stand: Oktober/November 2001): (1) Löwenbräu (LB, URL: <http://www.loewenbraeu.de/>), (2) Becks Bier (BB, URL: <http://www.becks.de/>), und (3) Jägermeister (JM, URL: <http://www.jaegermeister.de>). Alle drei Websites hatten eine ähnliche Funktionalität, unterschieden sich allerdings erheblich in Gestaltung und Interaktionsstil.

Jede Website wurde von jedem Teilnehmer in einer zufälligen Reihenfolge benutzt. Um möglichst unterschiedliche Wahrnehmungen und damit Varianz zu erzeugen, wurden die Teilnehmer in zwei Gruppen geteilt. Die eine Gruppe der Teilnehmer wurde gebeten bestimmte Aufgaben (z.B. Informationssuche [„Gründungsjahr der Löwenbräu-Brauerei“] oder Einkauf im Onlineshop) zu

³ Im Rahmen der gleichen Untersuchung wurden auch die in Hassenzahl, Kekez und Burmester (2002) berichteten Daten erhoben. Allerdings wurden in dieser Untersuchungen Ergebnisse des AttrakDiff 1 berichtet. Dieser „alte“ Fragebogen ist nur bei zweien der drei Websites mit eingesetzt worden (Löwenbräu und Jägermeister).

formationssuche [„Gründungsjahr der Löwenbräu-Brauerei“] oder Einkauf im Onlineshop) zu erledigen. Die Teilnehmer der anderen Gruppe konnten sich selbst Ziele setzen oder einfach nur zum „Spaß“ surfen. Nach jeder Nutzung wurde die Website von den Teilnehmern mit Hilfe des Fragebogens bewertet. Die Studie dauerte im Mittel zwei Stunden pro Person.

Alle Items außer den sieben Attraktivitätsitems aus dem AttrakDiff 1 wurden zunächst mit Hilfe einer Hauptkomponentenanalyse analysiert. Da die zu konstruierende Version des Fragebogens die drei Aspekte „pragmatische Qualität“ (PQ), „hedonische Qualität – Stimulation“ (HQ-S) und „hedonische Qualität – Identität“ (HQ-I) möglichst unabhängig messen sollte, wurden drei Komponenten extrahiert und varimax-rotiert. Die Items für die endgültige Version des Fragebogens wurden nach folgender Methode ausgewählt:

- Jeder Aspekt (PQ, HQ-S, HQ-I) sollte durch eine Skala mit mindestens sechs Items repräsentiert sein.
- Jedes einzelne Item repräsentiert einen vorher festgelegten Aspekt (z.B. HQ-S). Dieser wurde schon bei der Sammlung der Items (Abschnitt 3.1) bestimmt. Eine extrahierte Komponente wurde nun als einen Aspekt repräsentierend identifiziert, wenn eine hohe Zahl von Items eines Aspekts hohe Komponentenladungen aufwies. Luden beispielsweise viele Items, die bei der Sammlung als HQ-I Items identifiziert wurden auf einer Komponente, so wurde diese Komponente als HQ-I bezeichnet.
- Nur die Items, die möglichst hohe Ladungen auf einer Komponente zeigten, nicht besonders hoch auf anderen Komponenten luden und inhaltlich zum Aspekt der Komponente passten, wurden ausgewählt.

Es blieben 21 Items übrig, die erneut einer Hauptkomponentenanalyse unterzogen wurden. Diesmal wurden allerdings nicht gezielt drei Komponenten extrahiert und rotiert, sondern das Kaiser-Kriterium (Eigenwert >1) zur Bestimmung der Komponentenzahl angewendet. Die Komponentenlösung wurde wieder varimax-rotiert.

Die Hauptkomponentenanalyse extrahierte drei Komponenten mit einem Eigenwert größer als 1. Zusammen wurden durch die varimax-rotierte Lösung ca. 72% der Varianz erklärt. Tabelle 1 zeigt die 21 Items und die rotierte Lösung.

Die stärkste Komponente repräsentiert HQ-S. Sie erklärt ca. 29% der Gesamtvarianz. Die Komponentenladungen (d.h., die Stärke mit dem ein Item eine Komponente repräsentiert) sind im allgemeinen hoch (,758 - ,900). Die zweite Komponente ist HQ-I. Sie erklärt ca. 23% der Gesamtvarianz. Die Komponentenladungen sind etwas schwächer (,684 - ,831). Die dritte Komponente repräsentiert PQ. Sie erklärt ca. 20% der Gesamtvarianz. Die Höhe der Ladungen sind vergleichbar mit HQ-S (,642 - ,685). Ein offensichtlich problematisches Item ist „praktisch – unpraktisch“. Es sollte eigentlich alleine auf der PQ-Komponente laden. Es zeigt sich allerdings eine zweite, relative hohe Ladung auf HQ-I. Eine mögliche Erklärung ist, dass das Item der generellen Bewertung der Attraktivität semantisch näher ist, als der Wahrnehmung von pragmatischer Qualität. Die Items zur Messung der generellen Attraktivität tendieren dazu, auf allen Komponenten zu laden.

Tabelle 1: Hauptkomponentenanalyse mit Varimax-Rotation der endgültigen 21 Items

	Komponente		
	HQ - S	HQ - I	PQ
harmlos - herausfordernd	,758		
lahm - fesselnd	,831		
phantasielos - kreativ	,848		
originell - konventionell	,818		
neuartig - herkömmlich	,900		
innovativ - konservativ	,862		
mutig - vorsichtig	,890		
ausgrenzend - einbeziehend		,687	
bringt mich den Leuten näher - trennt mich von Leuten		,716	
isolierend - verbindend		,722	
nicht vorzeigbar - vorzeigbar	,408	,684	
minderwertig - wertvoll		,831	
stilvoll - stilllos		,728	
fachmännisch - laienhaft		,781	
praktisch - unpraktisch		,540	,644
widerspenstig - handhabbar			,865
voraussagbar - unberechenbar	-,418		,642
verwirrend - übersichtlich			,787
umständlich - direkt			,843
menschlich - technisch			,640
einfach - kompliziert			,789

Alles in allem weisen die ausgewählten 21 Items eine Struktur auf, die dem zugrunde liegenden Modell entspricht. Weiter kann man sich fragen, ob auch inhaltlich die gewünschten Konstrukte gemessen wurden (Konstruktvalidierung). Um dies zu prüfen, wurden anhand einer vorhergehenden analytischen Begutachtung folgende Hypothesen über die Qualitäten der einzelnen Websites formuliert:

- Löwenbräu ist betont pragmatisch. Die Website vermeidet neue Dialogelemente, Flash-Intros und Animationen.
- Jägermeister ist betont hedonisch. Die Website verwendet im starken Maße neue und ungewöhnliche Dialogelemente (z.B. horizontaler Bildschirmaufbau mit Scroll-Rad, animiertes Menü etc.). Fast jedes Element ist animiert. Die Website verwendet Cartoons, Hintergrundmusik und eine eigene Terminologie.
- Becks Bier ist balanciert. Diese Website ist nicht so ungewöhnlich wie Jägermeister, benutzt allerdings Flash-Intros, ein ungewöhnliches Hauptmenü und Hintergrundmusik.

Diese analytisch gewonnenen Unterschiede sollten sich auch in der Wahrnehmung der Websites durch die Teilnehmer widerspiegeln. Dabei beziehen sich die Aussagen zur hedonischen Qualität allerdings lediglich auf den Aspekt Stimulation (HQ-S). Über Identität konnten *a priori* keine Aussagen gemacht werden.

Abbildung 1 zeigt die mittleren Skalenwerte für pragmatische Qualität (PQ) und hedonische Qualität – Stimulation (HQ-S) pro Website.

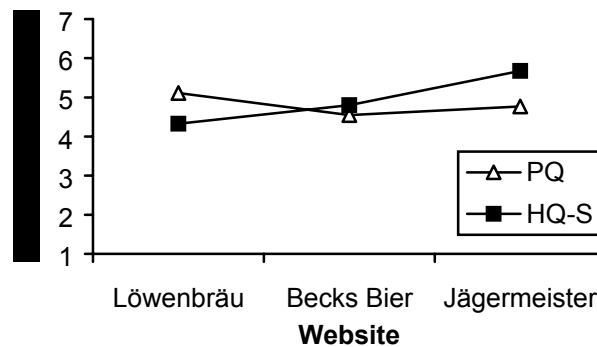


Abbildung 1: Mittlere Skalenwerte für pragmatische Qualität (PQ) und hedonische Qualität – Stimulation (HQ-S) pro Website

Eine 3×2-Messwiederholungs-Varianzanalyse mit den Faktoren Qualität (PQ, HQ-S) und Website (Löwenbräu, Becks Bier, Jägermeister) ergab eine hochsignifikante Interaktion zwischen Qualität und Website ($F=7,726$, $df=2$, $p=0,001$) und keine signifikanten Haupteffekte. Drei t-Tests für gepaarte Stichproben zeigen einen erwartungsgemäßen, marginal signifikanten Unterschied zwischen PQ und HQ-S bei Löwenbräu und Jägermeister und keinen bei Becks Bier (Löwenbräu: $t=2,27$, $df=22$, $p=0,033$; Jägermeister: $t=-2,51$, $df=22$, $p=0,020$; beide Signifikanzniveaus entsprechen einem alphasfehleradjustierten Niveau von $<.10$; Becks Bier: $t=-0,657$, $df=22$, n.s.). Alles in allem sind damit die aufgestellten Hypothesen bestätigt, was wiederum als ein erster Hinweis auf die Konstruktvalidität des AttrakDiff 2 gewertet werden kann. Im Rahmen einer zweiten Studie sollte das AttrakDiff 2 an einer unabhängigen Stichprobe weiter erprobt werden.

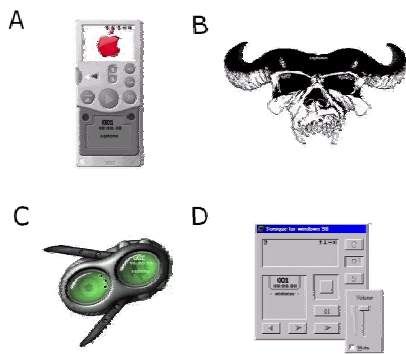


Abbildung 2: Die vier „skins“

3.3 Zweite Anwendung

Es nahmen 33 Psychologiestudenten (28 Frauen, 5 Männer) der Technischen Universität Darmstadt teil. Das mittlere Alter betrug rund 24 Jahre (Minimum 20, Maximum 40 Jahre). Insgesamt muss diese Stichprobe als eher homogen bezeichnet werden. Als Bewertungsobjekt diente eine Software zum Abspielen von MP3-Musikdateien (Sonique, URL: <http://sonique.lycos.com/>). Diese Software kann durch sogenannte „skins“ bei gleichbleibender Funktionalität optisch stark verändert werden. Es wurden für die Untersuchung vier, sehr unterschiedliche „skins“ ausgewählt (siehe Abbildung 2).

Die Teilnehmer wurden in zwei Gruppen aufgeteilt. Jeder Gruppe wurden nacheinander die Abbildungen der MP3-Player gezeigt (einer Gruppe in der Reihenfolge A-B-C-D, der anderen in der Reihenfolge D-C-B-A). Die Teilnehmer mussten dann die Software alleine auf der Basis ihres visuellen Eindrucks mit Hilfe des AttrakDiff 2 einschätzen. Die Studie dauerte ca. 20 Minuten.

Tabelle 2 zeigt die interne Konsistenz der einzelnen Skalen und die mittlere Interskalenkorrelation. Die interne Konsistenz ist für alle Skalen gut bis sehr gut. Die niedrigen mittleren Korrelationen zwischen den hedonischen und der pragmatischen Qualität unterstreichen erneut die Unabhängigkeit beider Konstrukte. Dies kann als Hinweis auf die Konstruktvalidität verstanden werden.

den. Die Korrelation zwischen den beiden Skalen HQ-S und HQ-I ist mit 0,55 recht deutlich. Diese beiden Konstrukte sind inhaltlich stärker miteinander verbunden als hedonische und pragmatische Qualität. Bedenkt man allerdings, dass das AttrakDiff 1 nicht zwischen beiden Konstrukten zu differenzieren vermochte, erscheinen 30% erklärte Varianz als akzeptabel.

Tabelle 2: Interne Konsistenz der einzelnen Skalen (Cronbachs α wurde für jede „skin“ ermittelt und dann als Bereich angegeben), die mittlere Interskalenkorrelation, sowie die erklärte Varianz.

Skala	Reliabilität (Cronbachs α)	Mittlere Interskalenkorrelation	
		HQ-S	HQ-I
hedonische Qualität–Stimulation (HQ-S)	,76 - ,90		
hedonische Qualität–Identität (HQ-I)	,73 - ,83	0,55 (30,25%)	
pragmatische Qualität (PQ)	,83 - ,85	0,18 (3,24%)	- 0,13 (1,69%)

Abbildung 3 zeigt die mittleren Skalenwerte für hedonische Qualität – Stimulation (HQ-S), hedonische Qualität – Identität (HQ-I) und pragmatische Qualität je „skin“. Da im Vorfeld keine unabhängigen Experteneinschätzungen oder andere Bewertungen vorgenommen wurden, dürfen die Ergebnisse streng genommen nicht als Hinweis auf Konstruktvalidität interpretiert werden. Dazu hätten *a priori* spezifische Hypothesen aufgestellt werden müssen. Trotzdem möchten wir auf die Ergebnisse im Sinne einer Plausibilitätsprüfung kurz eingehen.

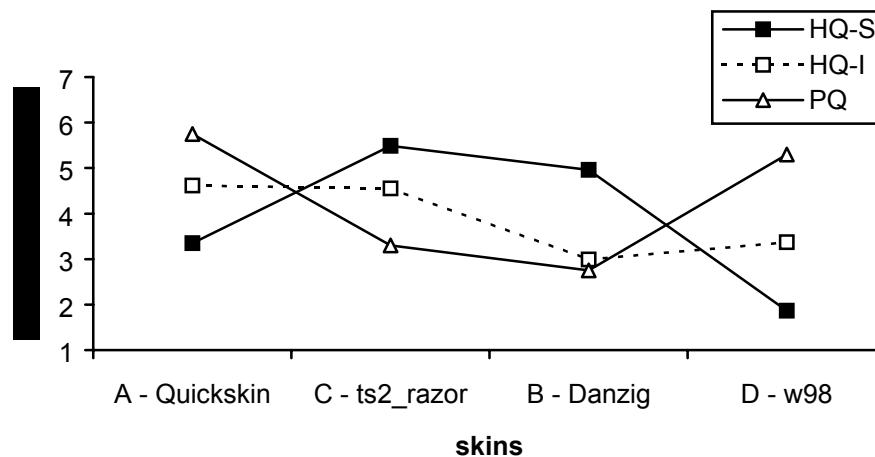


Abbildung 3: Mittlere Skalenwerte für hedonische Qualität – Stimulation (HQ-S), hedonische Qualität – Identität (HQ-I) und pragmatische Qualität je „skin“

Beim eingehenderen Begutachten der in Abbildung 3 gezeigten „skins“ kamen wir zu dem Schluss, dass sowohl „skin“ A als auch „skin“ D betont pragmatisch (handlungsorientiert) und „skin“ C und B als betont hedonisch (selbstorientiert) wahrgenommen werden sollten. Dies zeigt sich auch in den Daten. Auch sollte „skin“ B wegen seines martialischen Aussehens im Schnitt einen eher geringen Wert auf der Identitätsskala (HQ-I) bekommen. Es sind sicher Anhänger von Subkulturen denkbar, denen „skin“ B einen besonders hohen Grad an Identifikation erlaubt. Solange man allerdings eine solche Subkultur nicht gezielt als Teilnehmer gewinnt, ist mit einem eher niedrigen Mittelwert auf der HQ-I-Skala zu rechnen. Auch diese Annahme bestätigt sich. Insgesamt scheinen diese Ergebnisse plausibel. In Verbindung mit den guten bis sehr guten internen Konsistenzen und den plausiblen mittleren Interskalenkorrelationen kann dies als Hinweis auf eine akzeptable Qualität des Fragebogens verstanden werden.

4 Abschließende Diskussion

Friedmann (zitiert nach Ulich 1994, 118) betont, dass die eigentliche Arbeitsfreude (im Gegensatz zur Arbeitszufriedenheit) auf einer „tiefbegründeten Bejahung der Arbeit durch die Persönlichkeit beruht“, die „durch die Arbeit bereichert wird und selbst Entfaltung erfährt“. Ähnliches fordern wir auch für interaktive Produkte. Dem etwas blassen Anspruch auf Effizienz, Effektivität und Zufriedenstellung, steht der eigentliche Wunsch nach einer tiefen Bejahung, einer positiven Einstellung zum Produkt und der Hoffnung auf Freude bei der Nutzung gegenüber. Letzteres wird erst durch die Berücksichtigung auch anderer menschlicher Bedürfnisse neben der Manipulation der Umwelt im Rahmen benutzerzentrierter Produktgestaltung möglich. Der hier vorgestellte Fragebogen AttrakDiff 2 versteht sich als ein kleiner Beitrag in diese Richtung. Er ist eher als ein Forschungsinstrument, denn als eine voll entwickelte Evaluationsmethodik zu verstehen. Wir erhoffen uns, dass er von Wissenschaftlern und Praktikern gleichermaßen eingesetzt, überprüft und weiter verbessert wird. Im Vordergrund sollte dabei allerdings nicht die weitere Validierung des Fragebogens stehen, sondern ein besseres Verständnis des zugrunde liegenden Modells. Der Ansatz, hedonische Qualitäten von pragmatischen zu trennen und die weitere Differenzierung in Stimulation und Identität scheint uns als ein erster Schritt in die richtige Richtung. Weitere müssen folgen.

5 Literatur

- Batra, R.; Ahtola, O. T. (1990): Measuring the hedonic and utilitarian sources of consumer choice. *Marketing Letters*, Vol. 2, S. 159 – 170.
- Burmester, M., Hassenzahl, M.; Koller, F. (2002): Usability ist nicht alles – Wege zu attraktiven interaktiven Produkten. *I-Com*, Vol. 1, Nr. 1, S. 32 – 40.
- Draper, S. W. (1999): Analysing fun as a candidate software requirement. *Personal Technology*, Vol. 3, S. 1 – 6.
- Gediga, G.; Hamborg, K.-C. (2002): Ergonomische Evaluation von Software: Methoden und Modelle im Software-Entwicklungsprozess. *Zeitschrift für Psychologie*, Vol. 210, S.40 – 57.
- Hassenzahl, M. (2002): The effect of perceived hedonic quality on product appealingness. *International Journal of Human-Computer Interaction*, Vol. 13, S. 479-497.
- Hassenzahl, M. (2003): The thing and I: Understanding the relationship between user and product. In: Blythe, M.; Overbeeke, C.; Monk, A. F.; Wright, P. C. (Hrsg.): *Funology: From Usability to Enjoyment*. Dordrecht: Kluwer Academic Publishers, S. 31 – 42.
- Hassenzahl, M., Kekez, R.; Burmester, M. (2002): The importance of a software's pragmatic quality depends on usage modes. In: Luczak, H.; Cakir, A. E.; Cakir, G. (Hrsg.): *Proceedings of the 6th international conference on Work With Display Units (WWDU 2002)*. Berlin: ERGONOMIC Institut für Arbeits- und Sozialforschung, S. 275 – 276.
- Hassenzahl, M., Platz, A., Burmester, M., Lehner, K. (2000): Hedonic and Ergonomic Quality Aspects Determine a Software's Appeal. In: Turner, T.; Szwillus, G. (Hrsg.), *Proceedings of the CHI 2000 Conference on Human Factors in Computing*. New York: ACM, Addison-Wesley, S. 201 – 208.

- Hatscher, M. (2001): Joy of use - Determinanten der Freude bei der Softwarenutzung. *MMI-Interaktiv* [On-line]. Available: http://www.mmi-interaktiv.de/ausgaben/07_01. Letzter Zugriff 25.01.2003.
- Jordan, P. (2000): *Designing pleasurable products. An introduction to the new human factors*. London: Taylor & Francis.
- Kunze, E.-N. (2001): How to get rid of boredom in waiting-time-gaps of terminal-systems. In: Helander, M.G.; Khalid, H. M.; Ming Po, T. (Hrsg.), *Proceedings of the International Conference on Affective Human Factors Design*. London: Asean Academic Press.
- Prentice, D. A. (1987): Psychological correspondence of possessions, attitudes, and values. *Journal of Personality and Social Psychology*, Vol. 53, S. 993 – 1003.
- Sandweg, N., Hassenzahl, M.; Kuhn, K. (2000): Designing a telephone-based interface for a home automation system. *International Journal of Human-Computer Interaction*, Vol. 12, S. 401 – 414.
- Ulich, E. (1994): *Arbeitspsychologie, 3. Auflage*. Stuttgart: Schäffer-Poeschel.
- Winograd, T. (1996): *Bringing design to software*. New York: ACM Press.

Kontaktinformationen

Marc Hassenzahl

TU Darmstadt
Institut für Psychologie
Steubenplatz 2, 64293 Darmstadt
hassenzahl@psychologie.tu-darmstadt.de

Dr. Michael Burmester

Hochschule der Medien
Wolframstraße 32
70191 Stuttgart
burmester@hdm-stuttgart.de

Franz Koller

User Interface Design GmbH
Teinacherstraße 38
71634 Ludwigsburg
franz.koller@uidesign.de