

Warum „gutes Aussehen“ nicht immer von Vorteil ist

Über den Einfluss der optischen Gestaltung von Prototypen auf das Nutzerverhalten im Usability-Test

Andrea Struckmeier

eResult GmbH

Ludwig-Erhard-Str. 18

20459 Hamburg

andrea.struckmeier@eresult.de

Abstract

Der Beitrag stellt zunächst verschiedene Annahmen vor zum Einfluss der visuellen Verfeinerung (Aussehen) von Prototypen auf die Ergebnisse einer Usability-Evaluation. Im Anschluss werden die Ergebnisse einer Grundlagenstudie der eResult GmbH zum Einfluss visueller Verfeinerung von Prototypen dargelegt. In dieser Studie wurde die visuelle Verfeinerung als unabhängige Variable manipuliert, um die bisher bestehenden (heterogenen) Befunde zu überprüfen.

Die Ergebnisse der Studie legen nahe, dass Usability-Tests mit Prototypen von geringer Verfeinerung dazu führen, dass tendenziell mehr funktionale Probleme (z. B. hinsichtlich der Navigationsstruktur, des Wordings) von den Versuchspersonen erkannt werden als dies bei Prototypen der Fall ist, die visuell stark ausgereift sind. Zudem lässt sich vermuten, dass die Bewertung der subjektiv empfundenen Usability ebenfalls vom Grad der visuellen Verfeinerung abhängt.

Die Befunde aus der Studie werden abschließend kritisch diskutiert und es werden einige Hinweise gegeben, welche visuelle Verfeinerung bei der Konzeption von Prototypen zu beachten sind bzw. welche visuelle Verfeinerung vor dem Hintergrund welcher Fragestellungen empfehlenswert ist.

Keywords:

/// Prototyping
/// Konzeption
/// Fidelity
/// Usability-Testing
/// iteratives Testing

1.

Hintergrund und Fragestellung

Schlechte Usability kann teuer werden: Die Behebung von Usability-Problemen nach der Implementierung kostet bis zu zehn Mal mehr als in der davorliegenden Designphase (Bias & Mayhew, 1994). Eine Methode, mit der zentrale Probleme neuer Interaktionskonzepte bereits frühzeitig identifiziert werden können, ist das Prototyping. Mit Hilfe von Prototypen können Konzepte für z. B. Webseiten oder Software frühzeitig einem Nutzertest unterzogen werden. Dies kann einerseits in frühen Konzeptionsphasen geschehen, in denen zumeist erste Scribbles und Seitentemplates erstellt werden. Hier gibt es häufig mehrere Alternativen hinsichtlich des Konzeptes, die mit Hilfe von Prototypen gegeneinander getestet werden können. Doch auch während der Konkretisierung des Konzeptes ist es immer wieder sinnvoll, die Bedienbarkeit anhand von Prototypen

im Nutzertest zu überprüfen.

Aber: Die richtige Charakterisierung, Auswahl und Gestaltung von Prototypen für einen Usability-Test birgt einige Herausforderungen. Zu klären ist vor allem die Frage danach, ob der Prototyp bereits das vollständige Layout der späteren Anwendung abbilden muss, um wesentliche Usability-Probleme im Test zu identifizieren oder ob Nutzer qualitativ und quantitativ die gleichen Probleme identifizieren, wenn sie lediglich mit einem sehr groben graphischen Entwurf der Anwendung konfrontiert werden.

Dieser Frage wurde auch in der in diesem Beitrag skizzierten Untersuchung nachgegangen. Bevor näher auf diese eingegangen wird, sollen vorab Hintergrund und Ursprung der zu Grunde liegenden Fragestellung genauer beleuchtet werden.

1.1.

Prototypen und ihre Fidelity

In der Praxis lassen sich im Wesentlichen vier verschiedene Formen von Prototypen unterscheiden, anhand derer Anwendungen evaluiert werden können:

- Storyboard-Präsentationen, bei denen dem Nutzer verschiedene Screens einer Anwendung gezeigt werden.
- Papier-Prototypen, bei denen ein sog. Facilitator die Reaktion des Programmes auf die Aktionen der Nutzer simuliert und die Papierscreens entsprechend austauscht.
- Wizard-of-Oz-Simulationen, bei denen komplexere Funktionen (wie Spracherkennung) durch einen Mensch simuliert werden.
- Voll-funktionsfähige Computer-Prototypen, die bereits (in Teilen) eine vollständige Interaktion mit dem System ermöglichen (z. B. öffnen von Fenstern, Dateneingabe).

Die Charakterisierung von Prototypen erfolgt dabei anhand ihrer Fidelity. Diese wird in der Literatur häufig wie folgt definiert: "(...) the degree to which the prototype accurately represents the appearance and interaction of the product (...), not the degree to which the code and other attributes invisible to the user are accurate. (Rudd & Stern, 1996, S. 78). Mit anderen Worten: Je geringer der für den Nutzer augenscheinliche Unterschied zwischen einem Prototyp und der Originalanwendung, desto höher ist die Fidelity des Prototyps.

Dementsprechend werden Papier-Prototypen und Storyboard-Präsentationen häufig als low-Fidelity-Prototypen bezeichnet und voll funktionsfähige Computer-Prototypen als high-Fidelity-Prototypen.

Doch welche Prototyp-Fidelity ist für welche Phase im Designprozess geeignet? Hierzu gibt es äußerst unterschiedliche Befunde und Ansichten.

1.2. Studien zum Einfluss der Fidelity

In der Literatur wird häufig empfohlen, in frühen Entwicklungsphasen eher mit low-Fidelity-Prototypen zu arbeiten und weit fortgeschrittene Konzepte mit high-Fidelity-Prototypen zu evaluieren (Rudd, Stern, & Isensee, 1996). Dieser Empfehlung liegt die Annahme zu Grunde, dass Usability-Probleme mit dem späteren System besser vorausgesagt werden können, wenn ein Prototyp diese Anwendung hinsichtlich der Optik und der Interaktionsmöglichkeiten möglichst genau abbildet.

Einige Autoren gehen allerdings davon aus, dass low-Fidelity-Prototypen auch in späteren Entwicklungsphasen sehr gut für die Systemevaluation geeignet sind (z. B. Sefelin, Tscheligi, & Giller, 2003; Snyder, 2003). Diese Autoren kommen in ihren Studien zu dem Ergebnis, dass Prototypen von geringer und hoher Fidelity sich nicht signifikant in der Anzahl und Qualität der Usability-Probleme unterscheiden, die sie aufdecken.

Als eine Ursache dieser unterschiedlichen Annahmen kann auch die Tatsache gesehen werden, dass die Fidelity in den in den verschiedenen Studien häufig sehr unterschiedlich definiert wurde. Manche Studien zählten Papier-Prototypen grundsätzlich zu den low-Fidelity-Prototypen (Nielsen, 1990; Rudd et al., 1996; Lim, 2006). Andere Studien machten hingegen deutlich, dass allein das verwendete Medium (Papier vs. Computer) nicht ausschlaggebend für die Höhe der Realitätsnähe sein muss (Walker, Takayama, & Landay, 2002; Sefelin et al., 2003). Sie machten die Einstufung eines Prototyps als high oder low Fidelity von der visuellen Ausarbeitung und den Interaktionsmöglichkeiten abhängig.

1.3. Von high und low Fidelity zum Konzept der mixed Fidelity

Die o.g. Sachverhalte verdeutlichen, dass eine einfache Unterscheidung in high und low Fidelity nicht ausreicht, um die Vielzahl verschiedener Prototypen zu charakterisieren und auszuwählen. Je nach System oder Fragestellung der Evaluation ergeben sich sehr unterschiedliche Anforderungen an den Aufbau des Prototyps: Möchte man z. B. nur eine einzige Funktion eines gesamten Programmes untersuchen oder soll das gesamte Navigationskonzept evaluiert werden? Geht es eher um das Layout oder die Verständlichkeit von Eingabemöglichkeiten? Hier wird sehr schnell deutlich, dass die Frage nach der angemessenen Fidelity nicht pauschal mit „high“ oder „low“ beantwortet werden kann. Vielmehr gibt es eine Vielzahl von Aspekten, über deren Fidelity im einzelnen und je nach Fragestellung entschieden werden muss.

Vor diesem Hintergrund entwickelte McCurdy (2006) das sogenannte Konzept der mixed Fidelity. Dieses trägt der o. g. Vielzahl von Aspekten Rechnung, die bei der Konzeption von Prototypen berücksichtigt werden müssen. Bei diesem Konzept wird die Fidelity eines Prototyps auf mehreren Dimensionen definiert, welche unabhängig voneinander in Bezug auf high bzw. low Fidelity variiert werden können. Diese Dimensionen sind:

- Visuelle Verfeinerung (Aussehen)
- Breite der Funktionalität
- Tiefe der Funktionalität
- Umfang der Interaktionsmöglichkeiten
- Vollständigkeit des Datenmodells

Ein mixed-Fidelity-Prototyp ist nach dieser Definition ein Prototyp, der in Bezug auf einige (oder eine) Dimensionen über hohe Fidelity verfügt und während die Fidelity der übrigen Dimensionen mittel bis gering ausgeprägt ist (McCurdy, 2006, S. 1233).

Möchte man also, wie oben genannt, nur eine einzige Funktion eines Systems untersuchen, sollte demnach die Tiefe der Funktionalität stark ausgearbeitet sein, damit der Nutzer im Test den gesamten Ablauf erproben kann. Die Breite der Funktionalität kann hier über geringe Fidelity verfügen, da andere Funktionen nicht relevant sind und daher zwar auf oberster Ebene abgebildet werden können – sie müssen jedoch nicht klickbar sein.

Es zeigt sich, dass das Konzept der mixed Fidelity es ermöglicht, sehr differenziert festzulegen, welche Dimensionen eines Prototyps für eine bestimmte Fragestellung ausgearbeitet sein sollten und welche vernachlässigt werden können. [Abb. 1] So wird auch vermieden, dass unnötige zeitliche und finanzielle Ressourcen in

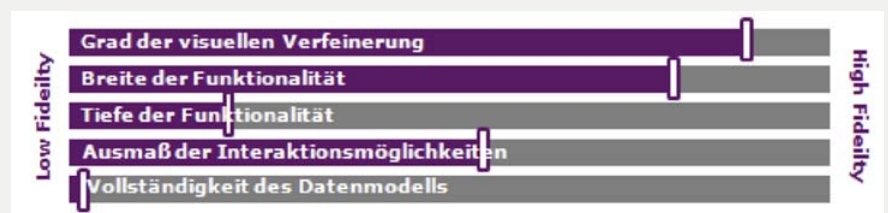


Abb. 1.
Unabhängige Manipulation einzelner Fidelity-Dimensionen (Quelle: eigene Darstellung).



Ausarbeitung von Bereichen investiert werden, die für eine Evaluation gar nicht notwendig wären.

1.4. Befunde zum Einfluss der visuellen Verfeinerung

Wenngleich das Konzept der mixed Fidelity eindeutig Vorteile gegenüber der ursprünglichen dichotomen Unterscheidung bietet, so gibt es dennoch bislang kaum Studien, die sich mit dem Einfluss einzelner Dimensionen auf z. B. die Identifikation von Usability-Problemen oder das Nutzerverhalten im Usability-Test befassen. Aus der bisherigen Forschung zur Fidelity und auch zur Ästhetik von Prototypen lässt sich jedoch die Vermutung ableiten, dass vor allem die Dimension der visuellen Verfeinerung ein beeinflussender und somit entscheidender Faktor sein könnte.

Häufig wird empfohlen, einen Prototyp erst dann mit einem kompletten Design auszustatten, wenn dieses auch evaluiert werden soll. Dies wird zum einen damit begründet, dass die Nutzer ansonsten zu viele Kommentare zu kosmetischen Aspekten des Prototyps abgeben und sich weniger auf funktionale Aspekte konzentrieren (Snyder, 2003; McCurdy, 2006). Auch legen manche Studien nahe, dass Nutzer größere Hemmungen haben, einen Prototyp zu kritisieren, der über ein komplett ausgearbeitetes Layout verfügt und dementsprechend „vollständig“ wirkt (z. B. Snyder, 2003).

Darüber hinaus gibt es aber auch Studien, die belegen, dass die Bewertung der Ästhetik eines Systems auch die Bewertung der subjektiv empfundenen Usability beeinflusst – je „schöner“ das System eingestuft wird, desto höher wird auch die Usability bewertet. Eine Studie zu diesem Thema stammt von Tractinsky et al. (2000).

Die Teilnehmer dieser Studie wurden in drei Gruppen aufgeteilt und bewerteten zunächst neun Screens in Bezug auf die Kriterien „Ästhetik“, „Usability“ und „Menge der Informationen“. Im Anschluss daran führte jede Versuchsperson mit

einem der Screens einige Aufgaben durch, wobei die tatsächliche Usability der Screens experimentell manipuliert war (niedrig vs. hoch, operationalisiert durch die Reaktionszeit des Systems oder der Buttons). Die Teilnehmer wurden den Kategorien hohe bzw. geringe Usability zufällig zugeordnet.

Im Anschluss an die Aufgabenbearbeitung bewerteten die Teilnehmer das System, mit dem sie gearbeitet hatten, erneut in Bezug auf die drei bereits

genannten Kriterien. Es zeigte sich dabei, dass die Bewertung der subjektiven Usability nur wenig von der tatsächlichen Usability des Systems beeinflusst wurde, sondern eher davon, wie schön die Nutzer das System fanden.

Dies legt die Vermutung nahe, dass Nutzer einen Prototyp hoher visueller Verfeinerung hinsichtlich der Usability besser bewerten könnten als einen funktionsgleichen Prototyp geringer Verfeinerung, wenn sie diesen aufgrund der mangelnden Ausgestaltung als weniger schön empfinden.

Die bisher aufgezeigten Befunde machen deutlich, dass hinsichtlich des Einflusses der Dimension „visuelle Verfeinerung“ noch Klärungsbedarf besteht. Fördert ein nur grob ausgearbeiteter Prototyp mehr und qualitativ hochwertigere Probleme zutage als ein Prototyp, der bereits über ein vollständiges Layout verfügt? Führt das Vorhandensein visueller Details in einem Prototypen-Test dazu, dass sich Nutzer eher auf eben diese „kosmetischen“ Aspekte beziehen und somit weniger Feedback zu funktionalen Aspekten abgeben? Und beeinflusst die ästhetische Wirkung die Beurteilung der Usability?

Eine eindeutige Antwort auf diese Fragen kann aus den bisher skizzierten Befunden nur schwer abgeleitet werden, da all diese Studien neben der visuellen Verfeinerung noch weitere Faktoren variierten (wie z. B. das Medium, die Informationsarchitektur). Somit konnte nicht zweifelsfrei geklärt werden, ob Gruppenunterschiede jeweils tatsächlich auf die Optik des Prototyps

zurückzuführen waren oder durch Wechselwirkung mit weiteren Faktoren entstanden.

In der im Folgenden beschriebenen Studie wurde daher zum ersten Mal gezielt der Einfluss der visuellen Verfeinerung auf das Verhalten von Nutzern und die Problemerkennung im Usability-Test untersucht.

Dabei wurde von folgenden Annahmen ausgegangen:

- Prototypen von hoher und geringer visueller Verfeinerung fördern eine ähnliche Anzahl und Qualität von Usability-Problemen zu Tage (Sefelin et al., 2003; Virzi, Sokolov, & Karis, 1996).
- Der Prototyp von hoher visueller Verfeinerung wird im Test besser bezüglich der Attraktivität besser bewertet als der Prototyp von geringer visueller Verfeinerung. Dementsprechend wird auch die Usability des Prototyps von hoher visueller Verfeinerung besser bewertet als die des Prototyps von geringer visueller Verfeinerung (Tractinsky et al., 2000).

2. Grundlagenstudie zur visuellen Verfeinerung

2.5.

Methodik und Ablauf der Studie

Bei der durchgeführten Untersuchung handelt es sich um ein einfaktorielles Design mit zwei Stufen, bei der die visuelle Verfeinerung als unabhängige Variable auf den Stufen „hoch“ und „niedrig“ variiert wurde. Hierzu wurde ein Zwei-Gruppen-Vergleich durchgeführt. Jeweils eine Gruppe (à 12 Testpersonen) durchlief einen Usability-Test mit einem Webseiten-Prototyp, der über geringe visuelle Verfeinerung verfügte (Abb. 2) und hauptsächlich in Graustufen gehalten war. Weitere 12 Testpersonen durchliefen den gleichen Test mit einem Prototyp von hoher visueller Verfeinerung (Abb. 3). Im Hinblick auf die übrigen vier Dimensionen der Fidelity (nach McCurdy, 2006) unterschieden sich die beiden Prototypen nicht voneinander. Auf diese Weise konnte gewährleistet werden, dass ein eventueller Unterschied

zwischen den Gruppen eindeutig auf die Manipulation der unabhängigen Variablen (visuelle Verfeinerung) zurückzuführen ist. [Abb. 2], [Abb. 3]

Im Rahmen der durchgeführten Untersuchung war hierbei vor allem der Einfluss auf folgende abhängigen Variablen von Interesse:

- Anzahl und Qualität erkannter Usability-Probleme
- Bewertung der Attraktivität und der subjektiv empfundenen Usability

Beim dem verwendeten Testgegenstand handelte es sich um einen frühen Konzeptentwurf für ein online-basiertes Nutzertagebuch – da keiner der Testpersonen bisher mit einem solchen Tool gearbeitet hatte, konnte so auch ein Einfluss eventueller Erfahrungswerte der Nutzer ausgeschlossen werden.

Jede der Testpersonen absolvierte im Rahmen des Tests fünf Aufgaben mit dem Prototyp. Im Anschluss an die Aufgabebearbeitung erfolgte eine sogenannte Videokonfrontation (Hamborg & Greif, 1999). Zudem füllten die Testpersonen abschließend einen Fragebogen aus, mit dem die Attraktivität und die subjektive empfundene Usability der Seite erfasst wurden. Der Fragebogen bestand aus 14 der 21 Items aus dem AttrakDiff (Hassenzahl et al., 2003) – je sieben davon erfassen die wahrgenommene Attraktivität bzw. die wahrgenommene Usability (operationalisiert über die Items zur pragmatischen Qualität aus dem Fragebogen).

2.6.

Zentrale Ergebnisse

2.6.1.

Anzahl und Qualität gefundener Usability-Probleme

Die Auswertung zeigte, dass sich die Anzahl der gefundenen Usability-Probleme in beiden Bedingungen nicht signifikant voneinander unterscheidet. So konnten bei hoher visueller Verfeinerung insgesamt 30 Probleme identifiziert werden; bei geringer Verfeinerung waren es 28 (jeweils ohne Redundanzen). Darüber hinaus wurden

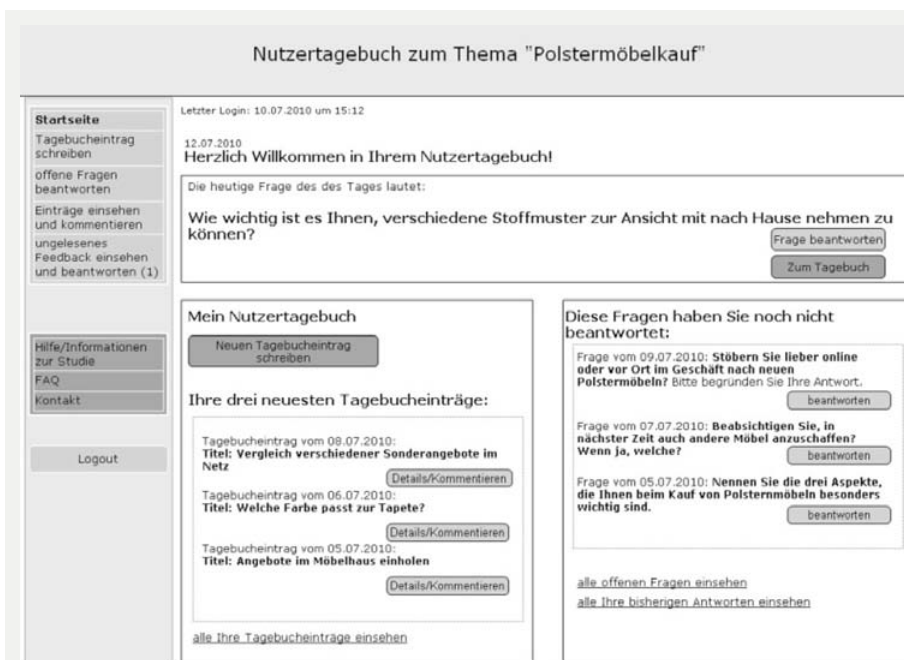


Abb. 2.

Prototyp geringer visueller Verfeinerung



Abb. 3.

Prototyp hoher visueller Verfeinerung



	Geringe Verfeinerung	Hohe Verfeinerung
Anzahl der Probleme pro Gruppe (inkl. Redundanzen)	41	36
Anzahl der Probleme pro Gruppe (ohne Redundanzen)	30	28
Anzahl alleinig gefundener Probleme	7	5
Anzahl aller gefundenen Probleme insgesamt über beide Gruppen (ohne Redundanzen)*	49	

Tab. 1.
Anzahl identifizierter Usability-Probleme bei geringer bzw. hoher visueller Verfeinerung

Kategorie	Geringe Verfeinerung	Hohe Verfeinerung
Wording	14	11
Navigation allgemein	3	1
unklare Unterteilung Tagebuch vs. Fragen	2	3
Navigation zu persönlichen Nachrichten	5	0
Seitenaufbau (Informationsarchitektur)	5	6
Kommentarfeld (unklare Funktion)	1	2
Gestaltung (Ablenkung durch graphische Elemente)	0	4
Logout	0	1
Gesamt	30	28

Tab. 2.
Anzahl identifizierter Usability-Probleme bei geringer bzw. hoher visueller Verfeinerung differenziert nach Problemkategorie.

unter jeder Bedingung einzelne Probleme identifiziert, die jeweils nur hier aufgedeckt wurden. [Tab. 1]

Für diejenigen Bereiche, in der nur von je einer Gruppe ein Usability-Problem festgestellt wurde, zeigte sich zumindest teilweise ein Unterschied zwischen den Gruppen: So wurden unter geringer visueller Verfeinerung mehr Probleme aufgedeckt, die sich auf die Navigation und das Wording bezogen (Fehlermeldung, Eingabeaufforderung in einem Formular). [Tab. 2]

Vor allem Probleme hinsichtlich der Navigation und hinsichtlich des Wording wurden im Vergleich zur hohen Verfeinerung häufiger gefunden. So fiel nur den Testpersonen, die mit dem schwarz-weißen Prototypen arbeiten auf, dass einer der Navigationspfade sehr viele Klicks erforderte, um ans Ziel zu gelangen. Dies wurde als sehr umständlich bewertet – vor allem im Hinblick auf die häufige Nutzung dieser Funktion. Bei hoher Verfeinerung fiel dieser Sachverhalt den Testpersonen nicht auf, da sie aufgrund der farblichen Hervorhebung schnell ans Ziel gelangten – sie erkannten jedoch nicht, dass dieser Prozess bei häufiger Nutzer sehr

zeitaufwendig ist. Hohe visuelle Verfeinerung förderte hingegen eher die Identifikation von Problemen zu Tage, die auf die Gestaltung einzelner Seitenelemente zurückzuführen waren. So wurden einzelne Testpersonen laut eigener Aussage z. B. von graphischen Elementen „abgelenkt“ oder sie bemängelten die farbliche Gestaltung einzelner Seitenbereiche.

2.6.2. Bewertung der Attraktivität und der subjektiv empfundenen Usability

Die Auswertung des Fragebogens zur Bewertung der Attraktivität und der subjektiv empfundenen Usability des Tools zeigt, dass sich im Hinblick auf die Bewertung ein signifikanter Unterschied in Bezug auf die Bewertung der Attraktivität ergab ($df=13$; $p=0,01$), nicht aber hinsichtlich der Usability ($df=13$; $p=0,118$), wobei auch hier die Bewertung des Prototyps geringer visueller Verfeinerung tendenziell schlechter ausfällt wie Abb. 4 erkennen lässt. [Abb. 4]

Weiterhin ergab die Analyse, dass die Korrelation zwischen der Bewertung Attraktivität und Usability bei hoher Verfeinerung generell höher ausfiel ($r=.863$) als bei geringer Verfeinerung ($r=.606$).

3. Implikationen für die Praxis

Die Ergebnisse zeigen, dass sich die visuelle Verfeinerung nicht auf die Anzahl identifizierter Usability-Probleme im Test auswirkt. Auch hinsichtlich der Qualität der Probleme wird deutlich, dass sich hier kein wesentlicher Unterschied ergibt. So gab es hohe Übereinstimmungen in der Problemerkennung in Bezug auf die Funktionen und Seitenbereiche, die die Nutzung der Anwendung beeinträchtigen.

Ogleich der Unterschied in Problemquantität und -qualität nicht signifikant war, lässt sich dennoch die Tendenz erkennen, dass geringe visuelle Verfeinerung dazu führt, dass sich die Testpersonen eher auf funktionale Aspekte konzentrieren wie z. B. das Wording und die Navigationsstruktur.

Dementsprechende Probleme wurden z. T. bei geringer Verfeinerung erkannt, bei hoher Verfeinerung jedoch nicht. Das deutlichste Beispiel hierfür in der vorliegenden Studie war die Navigation von der Startseite zu einer persönlichen Nachricht für den Nutzer (siehe Tabelle 2, „Navigation zu persönlichen Nachrichten“). Bei hoher Verfeinerung war der Weg zu dieser Nachricht für die Nutzer aufgrund einer farblichen Hervorhebung in der Navigation recht leicht zu erkennen. Dass jedoch die Navigation dorthin sehr viele Klicks erfordert und dadurch sehr zeitaufwändig war, wurde nur unter geringer Verfeinerung erkannt. Der Navigationspfad wurde daraufhin in der anschließenden Optimierungsphase deutlich verkürzt, wodurch die Nutzerfreundlichkeit und die Effizienz der Nutzung erhöht werden konnten. Wäre nur mit einem Prototyp von hoher Verfeinerung getestet worden, wäre das Problem im Test nicht zu Tage gefördert worden.

Hieraus lässt sich schließen, dass funktionale Aspekte wie Navigation, Wording oder Informationsarchitektur einer Anwendung mit einem Prototyp geringer Verfeinerung ggf. sogar besser evaluiert werden können als mit einem Prototyp hoher Verfeinerung, da sie aufgrund der mangelnden visuellen Ausgestaltung eine Konzentration der Nutzer auf die grundlegenden funktionalen Aspekte fördern. Einschränkend muss jedoch gesagt werden, dass dies nicht für jede Art von Anwendung gilt – hierauf soll weiter unten noch einmal eingegangen werden.

Die Analyse der Bewertung der subjektiv empfunden Usability legt zudem die Vermutung nahe, dass die empfundene Attraktivität bei hoher Verfeinerung stärker auf die Bewertung der Usability einwirkt als dies bei geringer Verfeinerung der Fall ist. Die Bewertung der Usability und der Einfluss der empfundenen Attraktivität auf die wahrgenommene Usability wichen in beiden Bedingungen jedoch nicht signifikant voneinander ab. Dennoch ließ sich die Tendenz dahingehend erkennen, dass der Prototyp hoher Verfeinerung hinsichtlich seiner Usability besser bewertet wurde – und das, obwohl sich Qualität

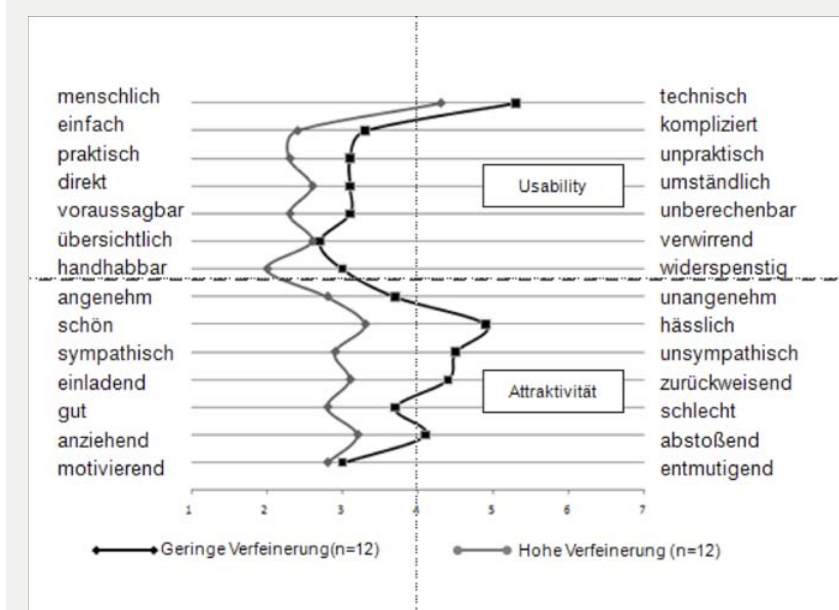


Abb. 4.
Bewertung der subjektiv empfundenen
Usability und der Attraktivität

und Quantität der identifizierten Usability-Probleme in beiden Bedingungen nicht signifikant voneinander unterschieden. Auf Grundlage der Daten kann allerdings keine eindeutige Aussage darüber getroffen werden, inwiefern die visuelle Verfeinerung eines Prototyps die Bewertung der Usability tatsächlich beeinflusst – diese Frage sollte daher mit einer größeren Stichprobe noch einmal erörtert werden.

Für die Konzeption von Prototypen zum Zwecke der nutzerbasierte Evaluation von Anwendungen lassen sich aus der Studie abschließend einige Hinweise ableiten:

1. Prototypen von geringer visueller Verfeinerung sind für die Evaluation funktionaler Aspekte ebenso gut geeignet wie Prototypen von hoher visueller Verfeinerung. Sie fördern eine ähnlich hohe Anzahl von Usability-Problemen zu Tage wie Prototypen hoher Verfeinerung und auch die Qualität der gefundenen Probleme unterscheidet sich nicht wesentlich voneinander.
2. Die Ergebnisse legen nahe, dass

- sich Versuchspersonen bei geringer visueller Verfeinerung tendenziell eher auf funktionale Aspekte konzentrieren (wie z. B. Eindeutigkeit der Navigation, Wording) und somit ggf. Probleme identifizieren, die bei hoher Verfeinerung unentdeckt bleiben.
3. Zudem legen die Ergebnisse die Vermutung nahe, dass hohe visuelle Verfeinerung trotz ähnlicher identifizierter Usability-Probleme eher dazu führt, dass der Prototyp in Bezug auf seine Usability besser bewertet wird als dies bei geringer Verfeinerung der Fall ist. Inwiefern dies tatsächlich der Fall ist, bleibt abzuklären – dennoch sollte dieser Aspekt berücksichtigt werden, wenn die Usability eines Prototyps zusätzlich mit Hilfe eines Fragebogens erhoben wird – insbesondere wenn mit hoher visueller Verfeinerung getestet wird. Nicht zuletzt, da es ebenfalls denkbar ist, dass ein solcher Prototyp auch deshalb besser bewertet wird, da die Anwendung bereits recht „vollständig“ wirkt und die Versuchspersonen ihre



Bewertung daher aus Rücksicht auf die Entwickler weniger kritisch ansetzen (Snyder, 2003).

Dennoch ist ein Test mit geringer visueller Verfeinerung nicht für alle Anwendungen empfehlenswert. Vielmehr muss vor der Entscheidung über den Grad der visuellen Verfeinerung vorab die genaue Fragestellung bzw. das Ziel der Evaluation analysiert werden. Generell lässt sich sagen, dass Prototypen geringer visueller Verfeinerung vor allem eignen für Fragestellungen, die sich auf funktionale Aspekte einer Anwendung beziehen wie z. B.:

- die Evaluation früher Konzepte und Ideen
- Wording oder Navigationsstruktur
- Verständlichkeit von Funktionalitäten und Features

In anderen Fällen ist ein Test mit geringer visueller Verfeinerung eher nicht empfehlenswert. Dies gilt z. B. wenn:

- emotionale Qualitäten erhoben werden sollen (Joy of Use, Wirkung des Designs)
- das Design dem Nutzer eine wesentliche Orientierungshilfe auf der Seite bietet (z. B. beim Test von Anwendungen, die sehr viel Text oder anderes Datenmaterial enthalten, dass z. B. durch farbliche Hinterlegung oder Hervorhebung voneinander abgegrenzt werden soll).

In diesen beiden Fällen ist eine visuelle Verfeinerung durchaus notwendig, wenn auch in unterschiedlicher Ausprägung: Da die Optik im ersteren Fall ein zentraler Bestandteil der Fragestellung ist, sollte das Aussehen des Prototyps der späteren Anwendung so ähnlich wie möglich sein, um auch z. B. den Joy of Use des späteren Systems korrekt vorherzusagen. Dies macht auch deutlich, dass die visuelle Verfeinerung oftmals erst gegen Ende des User-Centered-Designprozesses über hohe Fidelity verfügen kann, da ein genau ausgearbeitetes Layout zumeist erst zu einem späteren Zeitpunkt festgelegt wird.

Im zweiten Fall kann eine geringe visuelle Verfeinerung sogar dazu führen, dass die

Nutzer im Test negativ gegen den Prototyp beeinflusst werden, da besonders bei einer hohen Informationsmenge auf der Seite die Orientierung sehr schwer fällt, wenn nicht auf graphische Elemente zurückgegriffen wird. Hier muss die visuelle Verfeinerung sich zwar nicht detailgetreu am späteren Layout orientieren; sie sollte aber zumindest soweit ausgearbeitet sein, dass sie dem Nutzer hinreichende Orientierung bietet. Dies sollte auch ohne Weiteres schon in früheren Phasen des Konzeptionsprozesses möglich sein.

Abschließend lässt sich demnach festhalten, dass – bei entsprechender Fragestellung – auf eine detaillierte visuelle Ausarbeitung von Prototypen verzichtet werden kann. Besonders wenn funktionale Aspekte im Vordergrund stehen, sollte man explizit davon Abstand nehmen, einen Prototyp mit einem kompletten Layout auszustatten. Denn geringe visuelle Verfeinerung lenkt die Aufmerksamkeit der Nutzer eher auf eben diese funktionalen Aspekte und die tatsächliche Usability einer Anwendung. Eine Beeinflussung des Urteils durch die Optik der Anwendung ist hier weniger wahrscheinlicher als bei hoher Verfeinerung. So kann es teilweise sogar empfehlenswert sein, einen Prototyp auch dann mit nur geringer visueller Verfeinerung zu testen, wenn bereits ein ausgearbeitetes Layout vorliegt, um die grundsätzliche Bedienbarkeit einzelner Funktionen oder des gesamten Konzeptes zu testen.

Literatur

4. Bias, R. G. & Mayhew, D. J., (Hrsg.) (1994). Cost-Justifying Usability. San Francisco: Morgan Kaufmann Publishers.
5. Hamborg, K.-C. & Greif, S. (1999). Hierarchische Aufgabenanalyse. In: H. Dunckel (Hrsg.). Handbuch psychologischer Arbeitsanalyseverfahren. Zürich: vdf.
6. Hassenzahl, M., Burmeister, K., & Holler, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In: J. Ziegler & G. Szwillus (Hrsg.): Mensch und Computer 2003. Interaktion in Bewegung. Stuttgart: Teubner, S. 187-196.
7. McCurdy, M. (2006). Breaking the fidelity barrier: An examination of current characterization of prototypes and an example of a mixed fidelity success. CHI 2006 Proceedings, 1233-1243.
8. Nielsen, J. (1990). Paper versus Computer Implementations as Mockup scenarios for heuristic evaluation. In D. Diaper (Hrsg.), Human Computer-Interaction: Interact 90, S. 315-320. Amsterdam: Elsevier.
9. Rudd, J., Stern, K. & Isensee, S. (1996). Low vs. high-fidelity prototyping debate. Interactions, 3 (1), 76-85.
10. Sefelin, R., Tscheligi, M., & Giller, V. (2003). Paper prototyping - what is it good for?: a comparison of paper- and computer-based low-fidelity prototyping. CHI '03 extended abstracts on Human factors in computing systems, 778-779.
11. Snyder, C. (2003). Paper prototyping: The fast and easy way to design and refine user interfaces. San Francisco: Morgan Kaufman.
12. Tractinsky, N., Katz, A.S., & Ikar, D. (2000). What is beautiful is usable. Interacting with Computers, 13, S. 127-145.
13. Virzi, R. A., Sokolov, J. L. & Karis, D. (1996). Usability Problem Identification using both low- and high-fidelity prototypes. Proceedings of CHI 96, S. 236-243.

