

Papier oder Computer? Eine Studie zum Einsatz unterschiedlicher Gestaltungsmedien bei Usability Tests

Maren Mäuselein
User Interface Design GmbH
Friedrichsring 46
68161 Mannheim
maren.maeuselein@uidesign.de
www.uidesign.de

Joachim Machate
User Interface Design GmbH
Teichnacher Str. 38
71634 Ludwigsburg
joachim.machate@uidesign.de
www.uidesign.de

Johannes Schäfer
User Interface Design GmbH
Teichnacher Str. 38
71634 Ludwigsburg
johannes.schaefer@uidesign.de
www.uidesign.de

Abstract

Bei der Erstellung von Prototypen für Usability Tests stellt sich oft die Frage, ob die Wahl des Gestaltungsmediums Auswirkungen auf die Testergebnisse hat. Diese Studie vergleicht die Medien

Papier und Computer miteinander und kommt zu dem Schluss, dass Papierprototypen eine gleichwertige Alternative zu computerbasierten Prototypen sind.

Keywords

Usability Test, Vergleichsstudie, Prototyping, Papierprototyp, Gestaltungsform

1.0 Einleitung

Die Umsetzung von Konzeptideen in User Interface Prototypen ist eine zentrale Methode im benutzerzentrierten Gestaltungsprozess. Durch Prototypen können Bedienkonzepte visualisiert und erfahrbar gemacht werden; die Kommunikation über ein Konzept wird erleichtert und Missverständnisse können vermieden werden (Burmester, Görner 2003, S.51).

In Bezug auf Usability Tests bedeutet dies, dass diese mit Hilfe von Prototypen schon während der Entwicklungsphase eines Produkts möglich sind. Frühzeitiges Testen führt dazu, dass Schwachstellen des Konzeptes früh aufgedeckt und Kosten für Verbesserungen und Anpassungen niedrig gehalten werden können (Nielsen 1993).

Die Palette möglicher Arten von Prototypen für den Einsatz bei Usability Tests reicht von einfach, schnell und kostengünstig herzustellenden Papierprototypen bis hin zu interaktiven computerbasierten Prototypen, die der Nutzer unter Umständen nicht mehr von dem fertigen Produkt unterscheiden kann. Die große Anzahl und Unterschiedlichkeit verfügbarer Prototypformen macht eine Ent-

scheidung darüber, welche Art Prototyp die passende ist, oft schwierig.

In dieser Studie gingen wir den Fragen nach, ob es unterschiedliche Verhaltensweisen seitens der Studienteilnehmer beim Umgang mit papierbasierten gegenüber computerbasierten Prototypen gibt und ob sich je nach Medium unterschiedliche Bedienprobleme entdecken lassen.

2.0 Die Studie

Ziel der durchgeführten Studie war es, die Ergebnisse von Usability Tests mit Papierprototypen und computerbasierten Prototypen miteinander zu vergleichen. Von dem Testobjekt – einem Online-Mediencenter zur Verwaltung und Bearbeitung von Bildern, Musik und Videos – wurden zwei korrespondierende Prototypen erstellt, mit denen anschließend Tests mit insgesamt zehn Teilnehmern durchgeführt wurden. Fünf Testpersonen bearbeiteten Aufgaben mit dem Papierprototyp, fünf weitere mit dem computerbasierten Prototyp. Alle Teilnehmer hatten außerdem die Gelegenheit, den jeweils anderen Prototypen kennenzulernen. Die Ergebnisse der Tests wurden hinsichtlich verschiedener Fragestellungen untersucht, wie z. B.:

- Anzahl und Schwere der gefundenen Probleme
- Problemkategorien
- Bereitschaft der Teilnehmer, das Produkt zu kommentieren und eigene Ideen zu äußern
- Präferenz der Teilnehmer für einen Prototyp
- Anzahl der Fehler und Probleme bei der Aufgabenbewältigung

2.1 Rahmenbedingungen

Die Studie wurde an der Universität Paderborn im Rahmen einer Abschlussarbeit durchgeführt. Insgesamt wurden 10 Teilnehmer, ausschließlich Studenten, rekrutiert. Die Teilnehmer waren zwischen 20 und 30 Jahren alt und verfügen alle über eine hohe Erfahrung im Umgang mit Computern. Auf Grund der finanziellen und zeitlichen Rahmenbedingungen konnten die Tests nur mit dieser relativ homogenen Gruppe durchgeführt werden. Bei der Interpretation der Ergebnisse müssen deshalb die Einschränkungen der Teilnehmergruppe – Alter und Erfahrung – mit in Betracht gezogen werden.

Es wurden vier Aufgaben mit bis zu drei Teilaufgaben formuliert, die zum Einen verschiedene Teile und die wichtigsten Funktionen der Anwendung abdeckten und die zum Anderen teilweise bereits Problempotential enthielten. Die Aufga-

ben für beide Prototypen waren identisch. Um einen einheitlichen Testablauf zu gewähren wurde die Moderation in einem Testleitfaden vorgegeben.

Der computerbasierte Prototyp wurde von den Testpersonen eigenhändig bedient, den Papierprototypen setzte die Moderatorin während der Bearbeitung in einer zusätzlichen Rolle als „Computer“ entsprechend zusammen.

2.2 Die Prototypen

Bei der Erstellung der Prototypen wurde großer Wert darauf gelegt, eine möglichst hohe Vergleichbarkeit zu gewährleisten. Beide Prototypen zeigen eine Anwendung ohne grafischen Designanteil, da dieser auf Papier nur schwer zu imitieren ist. Trotzdem wurden die medien-spezifischen Vorteile beider Techniken beachtet.

2.2.1 Der computerbasierte Prototyp

Der computerbasierte Prototyp wurde als HTML Seiten dargestellt (Abb. 1). Insgesamt umfasst der Prototyp 20 HTML Seiten, zwei Stylesheets und 30 Bilddateien.



Abb 1: Startscreen des HTML Prototyps

Eine schwarz-weiße Gestaltung des Prototypen und der angezeigten Bilder wurde bewusst vermieden. Zwar wurde kein Designaufwand investiert, dem Vorteil eines HTML Prototypen, einfach und schnell Farbe zu benutzen und Fotos darzustellen wurde jedoch Rechnung getragen. Die Funktionalität des Proto-

typs beschränkt sich auf die vorgesehenen Aufgaben.

2.2.2 Der Papierprototyp

Der Papierprototyp wurde mit einfachen Materialien hergestellt:

- Papier in DIN A 4 und DIN A 3
- Bleistift und verschiedene schwarze Stifte
- ablösbares Klebeband

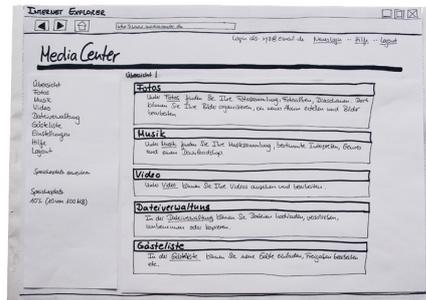


Abb 2: Startseite des Papierprototyps

Der Hintergrund mit unveränderlichen Elementen – einem schematisch dargestellten Browserfenster und der Kopfzeile des Testobjekts – wurden auf einem DIN A 3 Blatt gezeichnet, das als Hintergrund für den restlichen Prototyp dient. Die Navigation und der Content wurden austauschbar je nach Inhalt mit ablösbarem Klebeband auf dem Hintergrund fixiert (Abb. 2). Alle Beschriftungen wurden handschriftlich erstellt, Linien frei Hand gezogen und statt fotografischer Bilder wurden einfache Strichzeichnungen verwendet.

2.3 Durchführung

Für fünf Testpersonen begann die Studie mit dem Papierprototyp, für die fünf anderen mit dem computerbasierten Prototypen. Die Teilnehmer wurden zwar auf die Existenz eines zweiten Prototyps hingewiesen, waren sich der Ziele der Studie jedoch während der Aufgabenbearbeitung nicht bewusst.

Nach einer Begrüßung und Vorstellung des Testobjekts wurden die Testper-

sonen zunächst gebeten, die Startseite und ihre Elemente frei zu kommentieren. Nach dieser einführenden Aufgabe bearbeiteten die Teilnehmer die festgelegten Aufgaben der Reihe nach. Im Anschluss an die Aufgaben wurden sie hinsichtlich der Bedienbarkeit des Testobjekts befragt. Anschließend bekamen die Testpersonen die Möglichkeit, Aufgaben mit dem jeweils anderen Prototyp zu bearbeiten und anschließend ihre Präferenz zu äußern.

3.0 Ergebnisse

Als Ergebnisse wurden außer den bemerkten Usabilityproblemen auch Äußerungen der Teilnehmer, Fehler bei der Aufgabenbearbeitung sowie Antworten aus verschiedenen Befragungen vor und nach den Aufgaben als Daten gewonnen. Diese wurden gemäß dem Prototyp, mit dem die Teilnehmer gearbeitet hatten, gruppiert. Gruppe A hatte mit dem Papierprototyp, Gruppe B mit dem computerbasierten Prototypen Aufgaben bearbeitet.

3.1 Kommentare und Ideen

Die Äußerungen der Teilnehmer während der Tests wurden nachträglich gruppiert. Unter „Ideen“ wurden alle konstruktiven Äußerungen sowie Verbesserungsvorschläge zusammengefasst, „Kommentare“ bezeichnen alle Kommentare zu Elementen oder Strukturen der Anwendung. Rein beschreibende Kommentare wurden nicht beachtet, alle anderen wurden unterteilt in Äußerungen über Erwartungen (Expectations) und bewertende Kommentare (Judgements).

Die Ergebnisse zeigen, dass der Papierprototyp die Teilnehmer mehr dazu angeregt hat, die Startseite frei zu kommentieren wie es in der einleitenden Aufgabe gefordert war (15 Äußerungen in Gruppe A), während die Teilnehmer der Gruppe B, die mit dem computer-

basierten Prototyp gearbeitet hat, deutlich weniger Äußerungen [8] machten).

Während der Aufgabenbearbeitung lieferten beide Prototypen die gleiche Anzahl konstruktiver Ideen (24). Interessanter sind hier die Kommentare. Gruppe B zeigte eine deutlich größere Bereitschaft, die Anwendung zu bewerten (28 zu 15 wertenden Kommentaren wie „das finde ich gut / blöd / unpraktisch“) (Abb. 3).

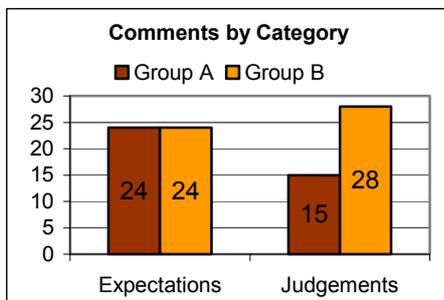


Abb 3: Gruppe B (Computer) neigte deutlich eher dazu, die Anwendung zu beurteilen

3.2 Usabilityprobleme: Anzahl und Schwere

Die gefundenen Usabilityprobleme wurden ebenfalls nach Gruppen aufgeteilt. Beide Gruppen haben jeweils genau 21 Probleme identifiziert, wobei fünf Probleme in jeder Gruppe einzigartig waren.

Den Problemen wurde außerdem ein Schwereindikator zugewiesen. Die Skala reicht von eins (das Problem verhindert die Fertigstellung der Aufgabe) bis vier (eher subtile Probleme, die eine Verbesserung in zukünftigen Produkten bringen könnten) (Dumas und Redish 1999). Die Auswertung nach Gruppen hat gezeigt, dass Gruppe A, die mit dem Papierprototypen gearbeitet hat, zwei Probleme der geringsten Schwere mehr gefunden hat als Gruppe B (drei statt eins), während Gruppe B jeweils ein Problem mehr der Schweren eins und zwei identifiziert hat (sechs statt fünf, bzw. acht statt sieben). Von den fünf Problemen, die jede

Gruppe exklusiv identifiziert hat, wurden jeweils vier nur von einer Testperson bemerkt.

Der Trend, dass die Gruppe mit dem computerbasierten Prototyp im Schnitt etwas schwerwiegendere Probleme gefunden hat kann nach dieser Studie noch nicht zu Schlussfolgerungen führen und müsste zunächst weiter verifiziert werden.

3.3 Usabilityprobleme: Problemkategorien

Aus den gefundenen Problemen wurden nach einer qualitativen Analyse sieben Kategorien herausgebildet:

- Konsistenz (Consistency)
- Prägnante Grafik (Distinct Graphics)
- Lokalität (Locality)
- Systemrückmeldung (System Response)
- Systemstruktur (System Structuring)
- Orientierung (User Orientation)
- Ausdruck (Wording)

Eine quantitative Analyse anhand der Anzahl der gefundenen Probleme liefert hier keine tragbaren Schlussfolgerungen (Abb. 4). Deshalb ist es nötig, die Probleme genauer zu betrachten und den Kontext, in dem sie aufgetreten sind, zu beleuchten.

Zwei der drei Probleme, die exklusiv von Gruppe A (Papier) in der Kategorie *Distinct Graphics* identifiziert wurde,

scheinen mediumsspezifisch zu sein, d.h. sie würden in dieser Form in der fertigen computerbasierten Anwendung nicht auftauchen (z.B. Unterscheidung klickbare / nicht klickbare Items). Ein interessantes Ergebnis ist, dass alle Probleme, die von Gruppe B in dieser Kategorie gefunden wurden, ebenfalls von den Nutzern des Papierprototyps identifiziert wurden. In einem frühen Entwicklungsstadium, in dem noch kein Grafikdesign stattgefunden hat und sich Probleme in dieser Kategorie hauptsächlich auf Strukturen und Anordnungen von Elementen beziehen, beeinflusst das Medium Papier nicht die Offenlegung von Problemen.

Die Probleme, die die meisten Nutzer betrafen wurden von beiden Prototypen aufgedeckt. Die meisten Probleme, die nur von einer Gruppe aufgedeckt wurden, wurden auch nur von jeweils einer Testperson bemerkt. Das exklusiv von Gruppe A bemerkte Problem in der Kategorie *System Response* wurde von zwei Testpersonen bemerkt.

Das Problem in der Kategorie *Wording* wurde von drei Teilnehmern der Gruppe B bemerkt und mit der größten Schwere bewertet. Besonders interessant ist hierbei, dass das Problem ebenfalls mediumsspezifisch ist. Hierbei handelt es sich um Missverständnisse, die mit dem Begriff „Seite“ einhergingen. Während die Nutzer des Papierprototypen mit „Seite“ automatisch eine neue Al-

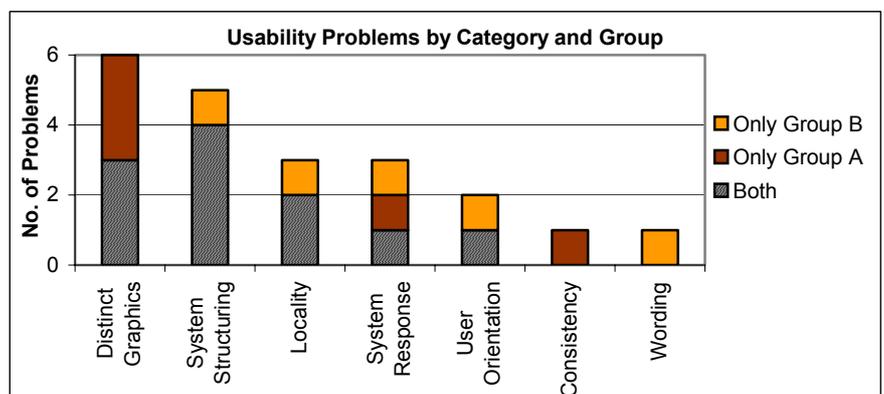


Abb 4: Verteilung der Usabilityprobleme nach Kategorien

bumseite assoziierten, da diese auch jeweils auf einem neuen Blatt Papier gezeigt wurde, hatten die Teilnehmer der Gruppe B große Probleme mit Aufgaben, die sich auf das Album bezogen (blättern, Seite löschen etc.). Da sich am PC auf einer Albumseite nur das angezeigte Bild mit beschreibendem Text ändert, kam es hier zu Missverständnissen. Der Papierprototyp hat diese Problematik nicht aufdecken können. Da das Endprodukt eine computerbasierte Anwendung ist, wurde an diesem Beispiel deutlich, dass Tests auf dem tatsächlichen Medium während der Entwicklung unumgänglich sind.

3.4 Präferenz

Nach der Bearbeitung der Aufgaben mit dem Prototyp wurden die Testpersonen gebeten, einige der Aufgaben mit dem jeweils anderen Prototypen zu wiederholen. Anschließend wurde ihre Präferenz für einen der Prototypen abgefragt: fünf Teilnehmer bevorzugten den computerbasierten Prototypen, vier den Papierprototypen und ein Teilnehmer enthielt sich (Abb. 5).

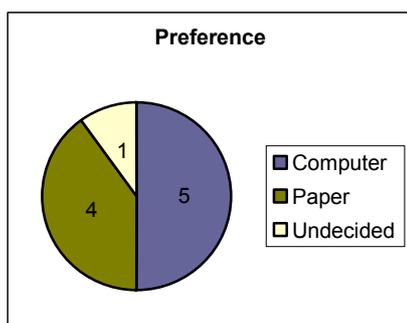


Abb 5: Die Frage, mit welchem Medium die Teilnehmer lieber arbeiten würden, ergab kein eindeutiges Ergebnis

Die Stimmen waren innerhalb der Gruppen gleichmäßig verteilt (Gruppe A [Papier]: je zwei Stimmen, Gruppe B: drei für den HTML Prototypen, zwei für den Papierprototypen). Die Testpersonen, die den Papierprototypen vorzogen gaben als Begründung an, er ließe ihnen

mehr Freiraum und sei flexibler, einer sagte es mache mehr Spaß mit ihm zu arbeiten. Der computerbasierte Prototyp wurde bevorzugt, weil den Teilnehmern die Arbeit mit Computer und Maus vertraut ist und der Prototyp dem späteren Produkt ähnlicher ist.

3.5 Bewältigung der Aufgaben

Gruppe A hatte etwas mehr Schwierigkeiten bei der Bearbeitung der Aufgaben. Die Teilnehmer machten hier bei sechzehn Aufgaben Fehler, während dies bei Gruppe B nur bei vierzehn Aufgaben der Fall war. Außerdem benötigten die Teilnehmer der Gruppe A bei fünf Aufgaben explizite Hilfestellung, während Gruppe B diese nur bei vier Aufgaben brauchte. Trotzdem bewertete Gruppe A die Handhabung des Produktes anschließend etwas besser als Gruppe B. Wie in 3.1 beschrieben waren die Teilnehmer in Gruppe B wertender als Gruppe A und schienen der Anwendung kritischer gegenüberzustehen.

4.0 Fazit

Die Auswertung der Tests hat gezeigt, dass beide Gestaltungsmedien, Papier und Computer, zu vergleichbaren Ergebnissen führen: sowohl die Anzahl konstruktiver Ideen als auch die Menge an gefundenen Usabilityproblemen war in beiden Gruppen gleich. Es gab keine signifikanten Unterschiede in der Art der Probleme, die erkannt wurden (Problemkategorien) und es wurde kein Prototyp eindeutig von den Teilnehmern bevorzugt.

Unterschiede in den Ergebnissen zeigten sich beispielsweise darin, dass die Nutzer des Papierprototypen mehr freie Äußerungen während der einleitenden Aufgabe, den Startscreen zu beschreiben machten, während der computerbasierte Prototyp bei der Bearbeitung der Aufgaben zu deutlich

mehr wertenden Kommentaren anregte.

Unter den Voraussetzungen, wie sie für diese Studie gegeben waren, hat der Papierprototyp sich als gleichwertig effektiv wie ein computerbasierter Prototyp für das Aufdecken von Problemen in Usability Tests erwiesen. Die Empfehlung von Rudd u.a. (Rudd et al. 1996), die davon abraten low-fidelity Prototypen (und damit in ihrem Verständnis auch Papierprototypen) für Nutzertests zu verwenden, kann nach den Erfahrungen in dieser Studie nicht bestätigt werden.

Allerdings haben die Ergebnisse auch gezeigt, dass ein Papierprototyp nie vollständig Tests auf dem Zielmedium ersetzen kann. Während des Entwicklungsprozesses werden Tests am PC für Software irgendwann notwendig sein, um medienspezifische Probleme aufzudecken.

Die komplette Studie wurde als Diplomarbeit an der Universität Paderborn im Studiengang Medienwissenschaft unter dem Titel „Paper Prototypes vs. Computer-based Prototypes in a User-centered Design Process“ durchgeführt und beschrieben und ist auf Anfrage bei den Autoren verfügbar.

5.0 Literatur

- Burmester, M.; Görner, C. (2003): Das Wesen benutzerzentrierten Gestaltens. In: Machate, J.; Burmester, M. (Ed.): User Interface Tuning. Benutzungsschnittstellen menschlich gestalten. Frankfurt: Software & Support Verlag, S. 47-66.
- Dumas, J. S.; Redish, J. C. (1999): A Practical Guide to Usability Testing. Exeter: Intellect Books.
- Nielsen, Jakob (1993): Usability Engineering. Boston: AP Professional.
- Rudd, Jim; Stern, Ken; Isensee, Scott (1996): Low vs. High-Fidelity Prototyping Debate. Interactions, Volume 3 Issue 1, 76-85.