

Alterseffekte bei der Nutzung von Online-Mediatheken - eine Usability-Evaluation

Franzisca Maas, Robert Tscharn, Jörn Hurtienne

Lehrstuhl für Psychologische Ergonomie, Julius-Maximilians-Universität Würzburg

Zusammenfassung

Die Online-Mediatheken der ARD und des ZDF bieten Nutzerinnen und Nutzern aller Altersgruppen die Möglichkeit TV- und Radioinhalte orts- und zeitunabhängig zu konsumieren. Trotz deren häufiger Nutzung durch verschiedene Altersgruppen finden sich hier in der Literatur jedoch keine Untersuchungen zu altersspezifischen Aspekten der Gebrauchstauglichkeit. In der vorliegenden Studie wurden daher diese beiden am meisten verbreiteten öffentlich-rechtlichen Online-Mediatheken sowohl anhand analytischer Usability-Methoden als auch einer empirischen Usability-Testung mit jungen und älteren Probanden verglichen. Die Ergebnisse zeigen, dass ältere Versuchsteilnehmer durchgängig größere Probleme bei der Benutzung der Mediatheken haben und dass das zwar modernere, aber teilweise zu simple Design der ZDF-Mediathek vor allem für die älteren Probanden problematisch war.

1 Einleitung

Laut der aktuellen ARD/ZDF-Onlinestudie von 2016 sind Onlineaktivitäten mit „Medien“ (TV, Video, Audio, Artikel) die zweithäufigste Internet-Aktivität der Deutschen. Nutzten 2013 nur 21% „zumindest selten“ die Mediatheken der Sender, so waren es 2016 schon 37% und nicht nur Jugendliche versorgen sich über das Internet mit Bewegtbild, sondern auch vermehrt ältere Nutzer (ARD/ZDF-Onlinestudie 2016). Wollen die Sender jedoch langfristig, dass ihre Online-Angebote auch genutzt werden, müssen sie nicht nur dafür sorgen, dass ihre Inhalte im Netz zur Verfügung stehen, sondern müssen diese auch nutzerfreundlich gestalten. Dies gilt insbesondere dann, wenn die Sender nicht nur Jugendliche, für die der Umgang mit Netzinhalten eine Selbstverständlichkeit ist, sondern auch die durchschnittlich 61-jährigen Tagesschau-Zuschauer (Emcke, 2013) mit ihrem Online-Angebot überzeugen wollen. Die vorliegende Arbeit vergleicht die aktuellen Versionen der beiden Mediatheken daher hinsichtlich der in der ISO-Norm 9241-11 (ISO,

1999) definierten Usability-Dimensionen 1) Effektivität, 2) Effizienz und 3) Nutzerzufriedenheit. Um mögliche Methodeneinflüsse zu untersuchen, wurden dabei die heuristische Evaluation nach Nielsen (Nielsen, 1994) sowie ein Usability-Test mit 27 Nutzern verschiedener Altersgruppen anhand vordefinierter Usecases durchgeführt. Zudem werden die verwendeten Methoden und mögliche Einschränkungen diskutiert.

2 Testobjekte

Die Mediatheken der ARD und des ZDF wurden ausgewählt, da sie als Produkte der öffentlich-rechtlichen Sender von der breiten Bevölkerung finanziert werden und somit auch für alle nutzbar sein sollten. Beide ermöglichen den Online-Zugriff auf die Beiträge der Sender, wobei der Online-Auftritt des ZDF zuletzt Ende 2016 komplett überarbeitet wurde. Die ARD-Mediathek ist unter www.ardmediathek.de, die ZDF-Mediathek unter www.zdf.de zu finden. Beide Seiten bieten neben der Option im linearen Fernsehen verpasste Sendungen nachzuschauen auch die Möglichkeit bestimmte Beiträge zu einer digitalen Merkliste hinzuzufügen, das TV-Programm nachzuschlagen, gezielt nach Sendungen oder Beiträgen zu suchen sowie das Live-Programm zu verfolgen. Die vorliegende Arbeit beschränkt sich auf die Desktopversionen der Mediatheken.

3 Heuristische Evaluation

Anhand von Niensens zehn Heuristiken (Nielsen, 1994) wurden beide Systeme von sechs Evaluatoren (Studierende der Mensch-Computer-Systeme, Grundkenntnisse in Usability-Evaluation durch den Kurs „Usability und Softwareergonomie“) frei exploriert und aufgetretene Probleme klassifiziert (zum Beispiel Fehlervermeidung oder Sichtbarkeit des Systemstatus). Anschließend wurden gefundene Fehler im Team je nach Schwere in fünf Kategorien eingeteilt, die von 0 (kein Problem für die Usability) bis 4 (Usability-Katastrophe) reichen. Dabei wurden bei der ARD-Mediathek insgesamt 35 und für die ZDF-Mediathek 25 Probleme gefunden. Tabelle 1 zeigt die gefundenen Schweregrade pro Mediathek. Der Modus beider Mediatheken war jeweils bei Schweregrad 2.

	Schweregrad 0	Schweregr. 1	Schweregr. 2	Schweregr. 3	Schweregr. 4
ARD	0	13	16	4	0
ZDF	0	9	14	2	0

Tabelle 1: Schweregrade der Heuristischen Evaluation

4 Usability-Test

Basierend auf der heuristischen Evaluation wurde ein empirischer Usability-Test durchgeführt, um die identifizierten Schwachstellen mit realen Nutzern zu quantifizieren. Aufgrund der vergleichbaren Ergebnisse aus der heuristischen Evaluation wurde erwartet, dass sich die beiden Mediatheken in Bezug auf ihre Effektivität und Effizienz nicht signifikant unterscheiden würden. Die ZDF-Mediathek hatte jedoch kurz vor Beginn der heuristischen Evaluation ein neues, moderneres Design erhalten. Die Hypothese lautete daher, dass die ZDF-Mediathek insgesamt auf der Dimension „Zufriedenheit“ besser als die ARD-Mediathek abschneiden sollte, jedoch nicht auf den Dimensionen „Effektivität“ und „Effizienz“. Zudem wurden insgesamt mehr Schwierigkeiten (höhere Bearbeitungsdauer, mehr Fehler, höherer mentaler Aufwand) für die ältere Benutzergruppe über beide Mediatheken hinweg erwartet. Schließlich wurde vermutet, dass das moderne Design der neuen ZDF-Mediathek vor allem jüngeren Nutzern zugutekommen würde, wohingegen ältere Nutzer möglicherweise sogar besser mit dem konservativeren Design der ARD-Mediathek zurechtkommen sollten.

4.1 Stichprobe

Der Test wurde mit 27 Probanden (13 weiblich) zwischen 17 und 76 Jahren durchgeführt. Das Durchschnittsalter der Stichprobe lag bei 41 Jahren ($SD = 19.12$). Es wurden zwei Altersgruppen gebildet: 12 Probanden, die jünger als 40 Jahre waren ($M = 21.17$, $SD = 3.30$) sowie 15 Probanden mit 40 Jahren oder älter ($M = 56.87$, $SD = 7.95$).

4.2 Erhobene Daten

Die gemessenen Variablen wurden an die drei Dimensionen der ISO-9241-11 angelehnt.

Effektivität	Effizienz	Zufriedenheit
Anzahl gelöster Usecases	Zeit in Sekunden pro Usecase	Intuitivität: QUESI (Naumann und Hurtienne, 2010)
Fehler (Zurück-Button im Browser)	Anstrengung: SEA-Skala (Eilers, Nachreiner, & Hänecke, 1986)	Attraktivität: AttrakDiff (Hassenzahl, 2003)

Tabelle 2: Gemessene Variablen für die Dimensionen der ISO-9241-11

4.3 Versuchsablauf

Die Usability-Tests wurden bei den Versuchspersonen zuhause durchgeführt um einen möglichst natürlichen Nutzungskontext zu gewährleisten. Nach der Begrüßung folgte zunächst eine Datenschutz- und Einverständniserklärung und -belehrung. Anschließend bearbeiteten die Probanden die Usecases für die erste Mediathek und füllten nach jeder Aufgabe die SEA-Skala aus. Es wurde ein within-subject-Versuchsdesign angewendet, da die Aufgaben aufeinander aufbauten. Die Hälfte der Probanden startete mit der ARD-Mediathek und die andere Hälfte mit der ZDF-Mediathek, um Positionseffekte konstant zu halten. Mit jeder Mediathek wurden sechs Usecases durchgeführt. Als erstes (1) sollten die Probanden den Abendnachrichtenbeitrag (ARD: Tagesschau 20Uhr, ZDF: heute 19Uhr) des Vorabends finden. Anschließend (2) sollte das gefundene Video im Vollbildmodus gestartet, der Ton abgestellt und der Vollbildmodus geschlossen werden. Daraufhin (3) sollte der Nachrichtenbeitrag der jeweiligen digitalen Merkliste hinzugefügt werden. Als nächstes sollten die Probanden (4) herausfinden, welcher Beitrag nach dem aktuell im Fernsehen Laufenden kommt. Zum Schluss (5) sollte der zuvor in der Merkliste gespeicherte Beitrag gestartet und (6) nur dieser aus der Liste gelöscht werden. Nach Abschluss aller Aufgaben der ersten Mediathek wurde den Probanden der QUESI und AttrakDiff vorgelegt. Währenddessen wurde die jeweils andere Mediathek gestartet. Für diese wurde der gleiche Ablauf verwendet. Abschließend wurden die Probanden nach einer persönlichen Präferenz für eine der beiden Seiten sowie möglichen Gründen für die Entscheidung gefragt.

4.4 Ergebnisse

Bezüglich der *Effizienz* der beiden Seiten zeigte eine 2 (Mediathek: ARD vs. ZDF; within) x 2 (Alter: junge vs ältere Benutzer; between) ANOVA, entgegen unserer Erwartung, dass die Usecases in der ARD-Mediathek signifikant schneller bearbeitet wurden $F(1,25) = 12.96, p = .001, \eta_p^2 = .34$ und jüngere Nutzer schneller in der Bearbeitung waren als ältere $F(1,25) = 127.17, p < .001, \eta_p^2 = .60$. Insgesamt ist jedoch zu beachten, dass die Nutzergruppe der über 40-Jährigen in beiden Mediatheken im Mittel über 50 Sekunden pro Usecase benötigte, was einen hohen Wert für diese Art von Aufgabe darstellt. Da bei genauerer Betrachtung die schnellere Bearbeitung aller Usecases in der ARD hauptsächlich aufgrund eines der komplexeren Usecases (Nr.5: „gespeicherten Beitrag aus Merkliste starten“) zustande kam, wurde dieser noch einmal getrennt betrachtet. Dabei zeigte sich im Gegensatz zu allen Usecases eine signifikante Interaktion zwischen der Mediathek und dem Alter der Nutzer $F(1,25) = 8.55, p = .007, \eta_p^2 = .26$. Vor allem ältere Nutzer hatten große Probleme bei der Bewältigung des Usecases in der ZDF-Mediathek, während die Jüngeren im ZDF nur wenig schlechter waren als in der ARD. Bei der SEA-Skala wurde die ARD-Mediathek signifikant weniger anstrengend eingeschätzt $F(1,25) = 9.232, p = .006, \eta_p^2 = .270$, und ältere Nutzer berichteten eine höhere Anstrengung in beiden Mediatheken $F(1,25) = 19.45, p < .001, \eta_p^2 = .44$. Eine Interaktion konnte nicht gefunden werden $F(1,25) = 3.148, p = .088$.

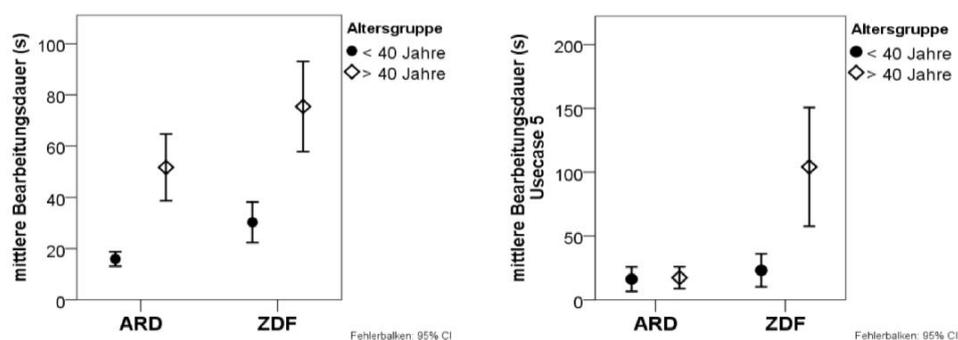


Abbildung 1: mittlere Bearbeitungsdauer (s) über alle Usecases (links) und für Usecase 5 (rechts)

In Bezug auf die *Effektivität* zeigte die 2 x 2 ANOVA mit der abhängigen Variable „mittlere Fehleranzahl“, dass in der ARD-Mediathek im Mittel signifikant weniger Fehler gemacht wurden $F(1,25) = 9.075, p = .041, \eta^2 = .16$, und ältere Nutzer nicht signifikant mehr Fehler machen als jüngere Nutzer $F(1,25) = 4.123, p > .05$. Eine signifikante Interaktion gibt es nicht $F(1,25) = .057, p > .05$. Die totale Anzahl erfolgreich bearbeiteter Usecases unterschied sich kaum (ARD: 145; ZDF: 135; total: 162), in beiden Mediatheken konnten jedoch die meisten Usecases gelöst werden. Entgegen unserer Hypothese zur *Zufriedenheit*, zeigte die 2 x 2 ANOVA in Bezug auf die Attraktivität, dass die ARD-Mediathek als attraktiver empfunden wurde $F(1,25) = 4.670, p = .040, \eta^2 = .16$, und ältere Nutzer beide Mediatheken als weniger attraktiv einschätzen $F(1,25) = 8.246, p = .008, \eta^2 = .25$. Das gleiche Muster zeigt sich auch bei der Einschätzung der intuitiven Benutzung (QUESI) (Haupteffekt Mediathek: $F(1,25) = 6.875, p = .015, \eta^2 = .22$; Haupteffekt Alter: $F(1,25) = 26.18, p < .001, \eta^2 = .50$).

Die abschließende qualitative Befragung ergab, dass 19 der 26 Versuchspersonen lieber die ARD-Mediathek benutzen würden (ZDF: 6 Personen, unentschieden: 2 Personen). Zudem ist der Anteil derjenigen, die die ARD-Mediathek bevorzugen, etwas größer bei den Älteren (73%), als bei den Jüngeren (67%). Gründe für die Bevorzugung der ARD-Mediathek waren beispielsweise, dass diese „vertrauter“ bzw. „übersichtlicher“ sei und „alle Funktionen ‚ausgeschrieben‘ statt spezielle Buttons“ seien. Problematisch sei bei der ARD vor allem ihr „überfrachtetes“ Design. Beim ZDF wurden die „schöne Gestaltung“ und das „moderne Design“ gelobt. Negativ fiel vielen Nutzern jedoch das „zu simple und dadurch verwirrende Design“ und „hinter Symbolen versteckte Funktionen“ auf. Manche waren sogar der Meinung die schönere Optik sei nur „zum Showing-Off“ ohne die Funktionalität zu verbessern. Alle Daten sind unter www.osf.io/6m7cw verfügbar.

5 Diskussion

Die Usability der Mediatheken der ARD und des ZDF wurde anhand von zwei klassischen Methoden verglichen: Heuristische Evaluation und Usability-Test. Die heuristische Evaluation ließ vermuten, dass die beiden Produkte grundsätzlich bedienbar sind, jedoch teilweise Probleme in der Benutzung auftreten. Der empirische Usability-Test konnte diesen Eindruck bestätigen. Die Erwartung, dass sich die beiden Mediatheken in Bezug auf Effektivität und Effizienz nicht unterscheiden, konnte nicht bestätigt werden, die Mediathek der ARD schnitt in fast allen gemessenen Variablen besser ab als die ZDF-Mediathek. Auch die Hypothese, dass die ZDF-Mediathek aufgrund ihrer moderneren Optik in Bezug auf die Zufriedenheit der Nutzer besser abschneiden würde, konnte nicht bestätigt werden, die ARD wurde sogar als attraktiver und intuitiver bewertet.

Bezüglich der beiden Altersgruppen zeigte sich wie erwartet, dass die ältere Nutzergruppe langsamer war und die Benutzung als anstrengender bewertet. Ältere machten aber im Mittel nicht mehr Fehler als die jüngere Gruppe. Dies könnte jedoch auch an einer geringeren Risikobereitschaft bei älteren Nutzern liegen (Barnard, Bradley, Hodgson, & Lloyd, 2013). Die Älteren empfanden beide Mediatheken als weniger attraktiv und intuitiv als die Jüngeren, was möglicherweise mit einer größeren Erfahrung der unter 40-Jährigen mit Online-Inhalten zu erklären ist. Die qualitative Befragung am Ende macht deutlich, dass das modernere Design der ZDF-Mediathek jedoch auch für viele jüngere Nutzer kein uneingeschränkter Vorteil ist, wenn mit dem neuen Design konventionelle Interaktionsmöglichkeiten versteckt werden. Die deutlich höhere Bearbeitungsdauer der älteren Probanden stellte in diesem Test ein eher kleineres Problem dar, da den Nutzern ausreichend Zeit zur Lösung der Aufgaben bereitstand. Es ist jedoch zweifelhaft, ob diese Nutzer bei einer realen Nutzung der Mediatheken die Persistenz aufbringen würden ebenso lange nach einer Lösung zu suchen, oder ob sie nicht eher nach einiger Zeit mit einem hohen Grad an Frustration aufgeben würden. So könnte die momentane Nutzerfreundlichkeit der beiden Systeme schlussendlich zu einer Abkehr (vor allem älterer Nutzer) führen. Wollen die beiden öffentlich-rechtlichen Sender ihr Hauptpublikum, das schon heute die ältere Altersgruppe ist, auch künftig mit ihren Inhalten digital erreichen, sollte diese Nutzergruppe gezielter bei der Gestaltung ihrer Mediatheken beachtet werden.

Kritisch zu bewerten, ist die Tatsache, dass in dieser Evaluation nur mit sechs Usecases gearbeitet wurde, die möglichst vergleichbare Bereiche der Mediatheken abdecken sollten. Diese Auswahl übersieht möglicherweise spezielle Probleme der einzelnen Seiten. Die Autoren unterstreichen explizit, dass die berichteten Unterschiede beider Mediatheken nur für die hier untersuchten Usecases interpretiert werden sollten.

Abschließend lässt sich festhalten, dass beide Mediatheken Usability-Probleme aufweisen, die gerade bei älteren Menschen die Nutzung erschweren. Nicht zuletzt zeigt unsere Studie, dass man gerade bei älteren Nutzern nicht ohne Weiteres von den Ergebnissen einer Heuristischen Evaluation auf das Verhalten und die Einschätzung realer Probanden schließen kann.

Literaturverzeichnis

ARD/ZDF-Onlinestudie (2016). *Videonutzung im Internet 2011 bis 2016*, <http://www.ard-zdf-onlinestudie.de/index.php?id=550>, zuletzt aufgerufen am 05.06.2017.

Barnard, Y., Bradley, M., Hodgson, F., & Lloyd, A.D. (2013). Learning to use new technologies by older adults: Perceived difficulties, experimentation behaviour and usability. *Computer Human Behaviour* 29(4), 1715-1724.

DIN EN ISO 9241 (1999). *Ergonomie der Mensch-System-Interaktion. Teil 11. Anforderungen an die Gebrauchstauglichkeit; Leitsätze*. Berlin: Beuth.

Eilers, K.; Nachreiner, F.; Hanecke, K. (1986): Entwicklung und Überprüfung einer Skala zur Erfassung subjektiv erlebter Anstrengung. *Zeitschrift f. Arbeitswissenschaft* 40(4), S. 215-224.

Emcke, C. (2013). Aller Tage Abend. In: *Zeitmagazin*. Nr. 26, 20. Juni 2013, S. 22-32 (<http://www.zeit.de/2013/26/berichterstattung-tagesthemen-emckes-expeditionen/komplettansicht>, zuletzt geprüft am 05.06.2017).

Hassenzahl, M., Burmester, M., & Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In: Szwillus, G., und Ziegler, J., (Hg.): *Mensch & Computer 2003*, Bd. 57. Wiesbaden: Vieweg+Teubner Verlag, 187-196.

Naumann, A., & Hurtienne, J. (2010). Benchmarks for intuitive interaction with mobile devices. In: *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services*, 401-402. Lisbon, Portugal: ACM.

Nielsen, J. (1994). *Usability engineering*. Oxford: Elsevier.