

# UX Messung mithilfe des UEQ+ Frameworks

Ein Erfahrungsbericht am Beispiel eines digitalen Zeitungsangebots

Heike Sandkühler  
NOZ Digital GmbH  
Osnabrück, Germany  
heike.sandkuehler@noz-digital.de

Martin Schrepp  
TI SAP CP Experience  
SAP SE  
Walldorf, Germany  
martin.schrepp@sap.com

Jörg Thomaschewski  
Hochschule Emden/Leer  
Emden, Germany  
joerg.thomaschewski@  
hs-emden-leer.de

## ABSTRACT

Eine gute User Experience ist für den Erfolg digitaler Produkte sehr wichtig. Daher ist die kontinuierliche Messung dieser Qualität von hohem Interesse. Hierzu bieten sich Online-Fragebögen an, da Kunden leicht über einen Link im Produkt angesprochen werden können. Wir beschreiben einen solchen Ansatz am Beispiel der Messung der User Experience eines digitalen Zeitungsangebots mit dem UEQ+ (N=1770 Teilnehmer). Der UEQ+ ist eine Sammlung von UX-Skalen, die beliebig kombiniert werden können. Damit kann man einen exakt auf die Erfordernisse des zu evaluierenden Produkts zugeschnittenen Fragebogen erzeugen. Wir beschreiben in diesem Beitrag, wie eine kontinuierliche Messung des digitalen Zeitungsangebots der *NOZ Medien & mh:n Medien* erstellt wurde. Im Detail wird gezeigt, anhand welcher Kriterien die relevanten UEQ+ Skalen ausgewählt wurden und wie der entsprechende Online-Fragebogen ins Produkt eingebettet wurde. Erste Ergebnisse der Produktevaluationen werden vorgestellt. Hierbei ergaben sich auch interessante Erkenntnisse zur Stabilität der UEQ+ Messungen und zur Beziehung der UEQ+ Ergebnisse mit dem Net Promotor Score. Weiterhin gehen wir darauf ein, wie die Ergebnisse im Unternehmen präsentiert werden und zur Verbesserung des Angebots eingesetzt werden.

## CCS CONCEPTS

• Human-centered computing • Human computer interaction • HCI theory, concepts and models

## KEYWORDS

User Experience, Messung von UX, Fragebögen, UEQ+

## 1 Einleitung

Eine gute User Experience (UX) ist für den langfristigen Markterfolg von digitalen Dienstleistungen oder Produkten sehr

wichtig. Kunden haben hier in der Regel die Auswahl zwischen verschiedenen Anbietern und können diese auch leicht wechseln, wenn sie mit der Qualität des Angebots nicht zufrieden sind. Daher ist die kontinuierliche Messung der UX Qualität für den langfristigen geschäftlichen Erfolg digitaler Produkte und Dienstleistungen ein zentrales Erfordernis.

Gerade für die Evaluation von digitalen Angeboten, bieten sich Online-Fragebögen zur Evaluation der subjektiv wahrgenommenen UX an. Ein Link auf den Fragebogen kann leicht im digitalen Angebot platziert werden. Damit ist es möglich, mit geringem Aufwand die Daten einer großen Anzahl von Nutzern zu erheben.

Allerdings ist es nicht immer ganz einfach, einen geeigneten UX-Fragebogen zu finden. Einerseits ist das Erstellen eines validen Fragebogens komplex und zeitaufwändig, andererseits decken die vorhandenen UX-Fragebögen nicht immer alle UX-Aspekte ab, die für das zu evaluierende Produkt wichtig sind [1].

Der UEQ+ [2, 3] ist eine Erweiterung des UEQ [4], der versucht dieses Problem durch einen modularen Ansatz zu lösen. Der UEQ+ ist kein fertiger, direkt anwendbarer Fragebogen, sondern eine Sammlung von UX-Skalen, die frei zu einem konkreten Fragebogen kombiniert werden können. Wir beschreiben in diesem Beitrag den Aufbau und erste Erfahrungen eines Prozesses zur kontinuierlichen Messung der UX-Qualität des digitalen Angebots der *NOZ Medien & mh:n Medien* mit Hilfe des UEQ+ Frameworks.

## 2 Warum ist die Messung der UX wichtig?

Zu den digitalen Nachrichtenprodukten der *NOZ Medien & mh:n Medien* zählen die ePaper-App, die News-App und die News-Portale. In diesem Beitrag wird die Messung der UX-Qualität anhand der ePaper-App vorgestellt. Dieses Produkt beinhaltet die digitale Version der Tageszeitung (ePaper) sowie alle Beilagen, die sonst mit der Tageszeitung mitgeliefert werden. Zudem kann auf ältere Ausgaben zugegriffen werden. Die App enthält weiterhin ein zusätzliches Magazin, sowie aktuelle Meldungen aus dem News-Portal. Abbildung 1 zeigt die Gestaltung der ePaper-App in der Ansicht auf einem Tablet.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s). MuC'20 Workshops, Magdeburg, Deutschland © Proceedings of the Mensch und Computer 2020 Workshop on «Quantitative Methoden zur Messung von User Experience». Copyright held by the owner/author(s).

<https://doi.org/10.18420/muc2020-ws105-244>

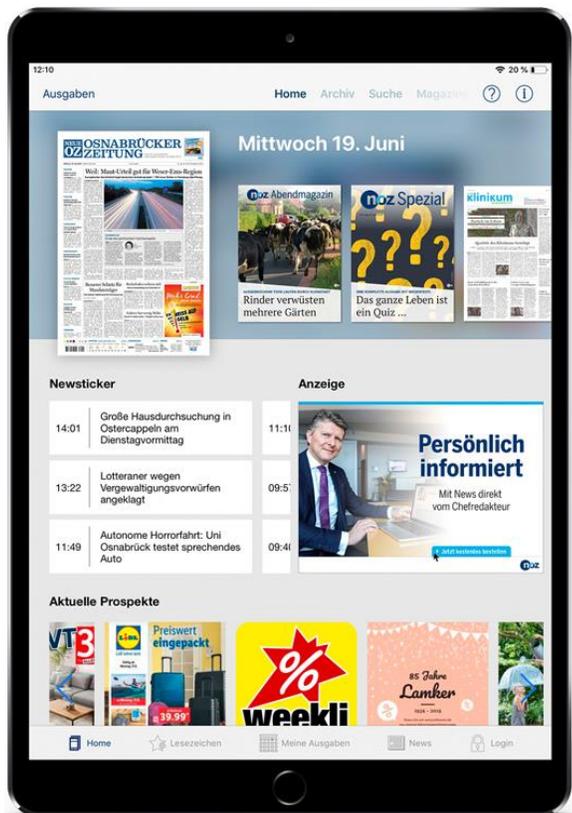


Abbildung 1: Die ePaper-App beinhaltet die digitale Zeitung und wird hauptsächlich auf dem Tablet genutzt.

Das Messen der UX-Qualität ist ein wichtiger Schritt in Richtung einer nutzerzentrierten Produktentwicklung. Es stellt ein wichtiges Tool dar, um die UX der Produkte kontinuierlich zu verbessern, was sich wiederum positiv auf die Nutzerzufriedenheit und somit auch auf die Kundenerhaltung auswirkt.

Konkret erfüllt das kontinuierliche Messen der UX Qualität zum einen den Zweck der Qualitätssicherung. Mithilfe eines UX-Monitorings können Veränderungen in der wahrgenommenen UX über Zeiträume hinweg sichtbar gemacht werden. Zum anderen dient die Erhebung als eine Art Kompass in der Produktentwicklung. Die Skalen des UEQ+ decken unterschiedliche UX-Qualitäten ab, welche zueinander ins Verhältnis gesetzt werden können. Außerdem kann die Bewertung einer UX-Skala mit dem Wichtigkeitsurteil der Skala verglichen werden, wobei deutlich wird, welche Aspekte der UX aus Sicht der Nutzer verbesserungsbedürftig sind. So dienen die Ergebnisse als Tool zur Priorisierung in der Weiterentwicklung der Produkte.

Ein Fragebogen bietet sich zur Erhebung der UX für die Produkte der NOZ Medien & mh:n Medien an, da die Nutzer direkt in dem Produkt angesprochen werden können.

Die Bewerbung des Online-Fragebogens in der App ist in Abbildung 2 dargestellt. Die Gestaltung des Fragebogens aus Teilnehmersicht zeigt Abbildung 3.

### 3 Erstellung des Fragebogens

Die Erstellung eines eigenen Fragebogens zur Messung der relevanten UX-Aspekte kam aus Zeitgründen nicht in Frage, d.h. es sollte ein Standardfragebogen verwendet werden.

Ein typisches Praxis-Problem bei der Verwendung von Standardfragebögen ist, dass es für eine gegebene Kombination von zu messenden UX-Aspekten oft keinen passenden Fragebogen gibt. Hier setzt die Idee des UEQ+ Frameworks [2, 3] an. Der UEQ+ ist eine Sammlung von UX-Skalen (diese erweitern die im UEQ [4] bereits vorhandenen Skalen, daher der Name UEQ+), die beliebig kombiniert werden können. Alle Skalen haben dabei das gleiche Format.

Eine Skala des UEQ+ besteht (ähnlich zum UEQ) aus 4 Items in Form eines semantischen Differentials. Da die Skalen beliebig miteinander kombiniert werden sollen, sind die Items einer Skala gruppiert. Ein kurzer einleitender Satz setzt einen Kontext für die gemeinsame Interpretation aller Items einer Skala. Zusätzlich wird für die Berechnung einer KPI (Key Performance Indikator) noch die Wichtigkeit der Skala für das betreffende Produkt abgefragt.

Das folgende Beispiel der Skala *Effizienz* zeigt das Format einer UEQ+ Skala:

Für das Erreichen meiner Ziele empfinde ich das Produkt als									
langsam	<input type="radio"/>	schnell							
ineffizient	<input type="radio"/>	effizient							
unpragmatisch	<input type="radio"/>	pragmatisch							
überladen	<input type="radio"/>	aufgeräumt							
Die durch diese Begriffe beschriebene Produkteigenschaft ist für mich									
Völlig	<input type="radio"/>	Sehr							
unwichtig	<input type="radio"/>	wichtig							

Neben den schon im UEQ vorhandenen Skalen *Attraktivität*, *Effizienz*, *Durchschaubarkeit*, *Steuerbarkeit*, *Stimulation* und *Originalität* enthält der UEQ+ im Moment noch die Skalen *visuelle Ästhetik*, *Anpassbarkeit*, *Nützlichkeit*, *intuitive Bedienung*, *Wertigkeit*, *Vertrauen*, *Inhalts-Seriosität*, *Inhalts-Qualität*, *Akustik* und *Haptik*. Diese Skalen wurden anhand empirischer Ergebnisse zur Wichtigkeit bestimmter UX Aspekte für verschiedene Produktkategorien erstellt [1,7]. Einige weitere Skalen werden in Kürze verfügbar sein.

Die drei Nachrichtenprodukte erfüllen unterschiedliche Nutzerbedürfnisse und werden in unterschiedlichen Nutzungskontexten genutzt. Der Vorteil des UEQ+ liegt darin, dass die zum Produkt passenden UX-Aspekte ausgewählt werden können. Im Folgenden wird erläutert, wie sich die Wahl der verwendeten UX-Skalen für die ePaper-App zusammensetzt. Für den zu erstellenden Fragebogen wurden folgende Skalen des UEQ+ ausgewählt:

- *Inhaltsqualität*: Subjektiver Eindruck, dass die Informationen, die das Produkt dem Nutzer bereitstellt, aktuell, gut aufbereitet und interessant sind. Diese Skala wurde gewählt,

weil die Inhaltsqualität bei einem Nachrichtenprodukt eine zentrale Rolle spielt.

- **Effizienz:** Nutzer haben den Eindruck, dass sie ihre Aufgaben ohne unnötigen Aufwand lösen können. Für die ePaper-App wurde diese Skala gewählt, weil vermutet wurde, dass die Schnelligkeit, sowie die Übersichtlichkeit der App einen wichtigen Faktor in der Nutzung darstellt. Da die ePaper-App von den Nutzern meist täglich genutzt wird, ist es wichtig, dass das Produkt schnell und ohne große Aufwände genutzt werden kann.
- **Wertigkeit:** Subjektiver Eindruck, dass das Produkt von hoher Qualität und professionell gestaltet ist. Diese Skala wurde gewählt, weil die ePaper-App in der digitalen Produktpalette ein hochpreisiges Premium-Produkt darstellt.
- **Durchschaubarkeit:** Subjektiver Eindruck des Benutzers, dass es leicht ist, sich mit dem Produkt vertraut zu machen. Diese Skala wurde gewählt, da viele Nutzer der ePaper-App von der gedruckten Zeitung auf die digitale Zeitung umsteigen und anfangs noch wenig Erfahrungen mit der Technik haben. Ein verständliches und leicht zu durchschauendes Produkt ist hier besonders wichtig, da auch der Altersdurchschnitt tendenziell höher ist als in anderen digitalen Produkten.
- **Nützlichkeit:** Subjektiver Eindruck, dass die Verwendung des Produkts Vorteile bringt, Zeit spart oder die persönliche Produktivität verbessert. Diese Skala wurde gewählt, da bereits bekannt ist, dass für viele Nutzer der ePaper-App die Nützlichkeit besonders relevant ist. Vorteile, wie das Einsparen des Papiers oder das ortsunabhängige Zeitungslesen sind wichtige Faktoren für viele Nutzer.
- **Attraktivität:** Gesamteindruck des Produkts. Mögen die Benutzer das Produkt oder nicht? Diese Skala wurde gewählt, um die allgemeine Zufriedenheit zu ermitteln.

Zusätzlich zur UX-Qualität ist für *NOZ Medien & mh:n Medien* noch die Weiterempfehlungsbereitschaft sehr relevant. Kunden, die vom Angebot so angetan sind, dass sie dieses Freunden und Bekannten weiterempfehlen, sind wichtig zur Gewinnung neuer Abonnenten. Solche Kunden haben auch eine geringe Kündigungswahrscheinlichkeit für ihr Abonnement. Für die Messung dieser Eigenschaft wurde der Net Promotor Score (kurz NPS) [5] in den Fragebogen mit aufgenommen.

Abhängig von der angegebenen Wahrscheinlichkeit zur Weiterempfehlung werden alle Teilnehmer der Umfrage in drei Gruppen eingeteilt: *Detractors*, *Neutrals* und *Promoters*.

Anhand eines Freitextkommentars werden *Promoters* gebeten zu erläutern, warum sie das Produkt weiterempfehlen würden, *Neutrals* werden gebeten zu beschreiben, was an der App verbessert werden sollte und *Detractors* werden nach einer Begründung ihrer Bewertung gefragt.

Diese Freitextantworten werden zur Auswertung in Kategorien eingeteilt. So wird erkennbar, welche konkreten Produkteigenschaften gelobt oder kritisiert werden. Die Daten haben eine wichtige Funktion in der anschließenden Interpretation der Ergebnisse und für die Ableitung von konkreten Handlungsempfehlungen.



Abbildung 2: Bewerbung des Online-Fragebogens in der ePaper-App.

Die Inhalte in der ePaper-App empfinde ich als...

1	2	3	4	5	6	7
---	---	---	---	---	---	---

veraltet aktuell

1	2	3	4	5	6	7
---	---	---	---	---	---	---

uninteressant interessant

1	2	3	4	5	6	7
---	---	---	---	---	---	---

schlecht aufbereitet gut aufbereitet

1	2	3	4	5	6	7
---	---	---	---	---	---	---

unverständlich verständlich

Dass die Inhalte in der ePaper-App aktuell, interessant, gut aufbereitet und verständlich sind, ist für mich...

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Völlig unwichtig Sehr wichtig

Abbildung 3: Online-Fragebogen aus Sicht des Teilnehmers.

## 4 Bisherige Befragungen

Die Umfrage wurde bisher einmalig in jedem Produkt durchgeführt. Voraussichtlich wird dies zwei- bis dreimal im Jahr wiederholt. Die Nutzer nehmen an der Umfrage freiwillig teil. Es wird kein Incentive für die Teilnahme vergeben.

Bei den ePaper-Apps wurden zur Bewerbung der Umfrage zwei Werbemittel eingesetzt: Eine dauerhafte Grafik wurde auf der Startseite der App platziert und ein Layer, der sich beim Start der App öffnet wurde eingerichtet. Abbildung 2 zeigt diese Grafik auf der Startseite der ePaper-App. Abbildung 3 zeigt einen Ausschnitt

aus dem Fragebogen (Skala Inhaltsqualität) aus Sicht des Teilnehmers.

Die Dauer des dauerhaften Werbemittels auf der Startseite betrug 13 Tage und hatte eine durchschnittliche Klickrate von 0,45%. Der Layer wurde an zwei Tagen ausgespielt und hatte eine durchschnittliche Klickrate von ca. 6,1%. Teilgenommen haben insgesamt 3327 Nutzer. Täglich lesen im Durchschnitt ca. 71.000 Nutzer das ePaper in der ePaper-App.

Neben dem UEQ+ wurden die Teilnehmer in der Umfrage nach ihrem Alter, Geschlecht, ihrer Nutzungsintensität, Gerätenutzung und der Nutzung bestimmter App-Bereiche gefragt. Außerdem wurde nach der Wiederempfehlungswahrscheinlichkeit (NPS) gefragt und die Teilnehmer hatten die Möglichkeit als Freitextantwort einen Kommentar abzugeben.

Das Ausfüllen des Fragebogens dauert im Durchschnitt ca. 6 Minuten. Bei der Berechnung der durchschnittlichen Bearbeitungsdauer wurden Teilnehmer, die länger als 30 Minuten brauchten nicht berücksichtigt, da man hier davon ausgehen muss, dass die Bearbeitung der Umfrage zwischendurch unterbrochen wurde.

### 5 Erfahrungen mit dem UEQ+

Wir zeigen am Beispiel einer Befragung (N=1770 Teilnehmer), welche Daten für die Auswertung anfallen. Tabelle 1 zeigt die Skalenmittelwerte und Standardabweichungen der verwendeten Skalen.

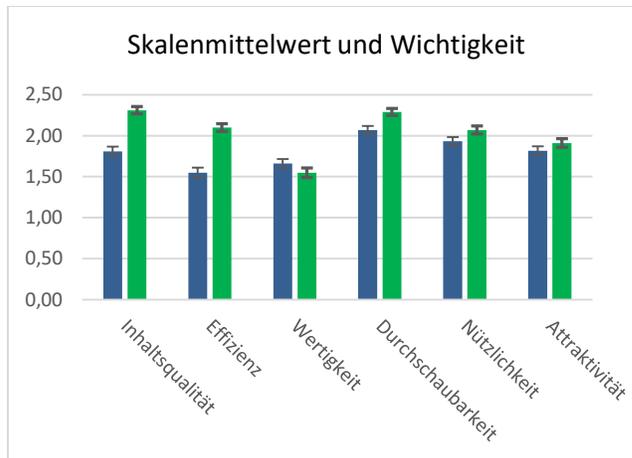
**Tabelle 1:** Skalenmittelwerte (MW), Varianz und Standardabweichungen (STD) für die Studie mit N=1770 Teilnehmern.

Skala	MW	Varianz	STD
Inhaltsqualität	1,81	1,38	1,18
Effizienz	1,55	1,68	1,29
Wertigkeit	1,66	1,42	1,19
Durchschaubarkeit	2,07	1,33	1,15
Nützlichkeit	1,93	1,14	1,07
Attraktivität	1,82	1,28	1,13

Die Mittelwerte (die Skalen sind von -3 bis +3 kodiert) zeigen schon recht deutlich, dass das Angebot bzgl. seiner UX sehr positiv bewertet wird.

Der UEQ+ erhebt zu jeder Skala auch ein Wichtigkeitsurteil. Die folgende Abbildung 4 zeigt die Mittelwerte und die erhobenen Wichtigkeitsurteile.

Wenn die Skalen des UEQ+ gut gewählt sind, sollten diese auch von den Teilnehmern der Befragung als sehr relevant für ihre Gesamtbeurteilung bewertet werden (s. Abb.4, grüne Werte). Das ist für diese Untersuchung offensichtlich der Fall. Insofern sind die Wichtigkeitsurteile auch eine gute Möglichkeit anhand der Daten zu prüfen, ob die gewählten Skalen auch eine hohe Relevanz aus Sicht der Teilnehmer besitzen.



**Abbildung 4:** Skalenmittelwerte (blau) und Wichtigkeitsurteile (grün).

Die Wichtigkeitsurteile werden verwendet, um aus den einzelnen Skalenwerten eine Gesamtbewertung zu berechnen. Hierfür werden die Skalenwerte pro Person mit der relativen Wichtigkeit der Skala gewichtet [2, 3]. Der Mittelwert über alle diese Bewertungen resultiert in der KPI des Produkts, die in dieser Studie 1,82 beträgt. Der ebenfalls ermittelte Wert für den *Net Promoter Score* betrug 0,36. Beide Werte deuten auf eine sehr hohe Zufriedenheit mit dem Angebot hin.

Die Reliabilität der Skalen wurde durch Cronbachs Alpha [6] bestimmt. Die entsprechenden Werte waren durchgängig sehr hoch: *Inhaltsqualität* 0,89, *Effizienz* 0,88, *Wertigkeit* 0,93, *Durchschaubarkeit* 0,95, *Nützlichkeit* 0,95, und *Attraktivität* 0,96.

Das erhobene Wichtigkeitsurteil wird, wie schon beschrieben, zur Berechnung der UEQ+ KPI (Kennzahl, die einen Wert für die Gesamtbeurteilung über alle Skalen ausdrückt) verwendet, indem die Skalenwerte mit der relativen Wichtigkeit pro Skala gewichtet werden (für Details, siehe [2, 3]).

Eine für den UEQ+ methodisch relevante Frage ist daher, wie stabil die Wichtigkeitsurteile sind. Theoretisch sollte die Wichtigkeit eines UX-Aspekts für die Gesamtbeurteilung eines Produkts vor allem von der Art des Produkts abhängen [1, 7].

Tabelle 2 zeigt die mittlere Wichtigkeit und die zugehörige Standardabweichung (in Klammer) der verwendeten UEQ+ Skalen für 4 Untersuchungen zu digitalen Angeboten von *NOZ Medien & mh:n Medien*.

Für die Messung wurden die ePaper-Apps der vier zur Mediengruppe gehörenden Verlage untersucht. Die vier Produkte sind in ihrer technischen Umsetzung identisch und unterscheiden sich nur aufgrund ihrer redaktionellen Inhalte. Zudem ist die Preisgestaltung den unterschiedlichen Märkten angepasst. Ein Digital-Abo kostet für Produkt A, B und C 28,90 € und für Produkt D 22,90 € pro Monat.

**Tabelle 2:** Wichtigkeitsurteile aus 4 Studien. Der Wert in Klammer im Spaltenkopf beschreibt die Anzahl der erhobenen Daten pro Produkt.

Skalen	Studien (Teilnehmer)			
	A (242)	B (1125)	C (1770)	D (190)
Inhaltsqualität	2,24 (1,09)	2,25 (0,99)	2,31 (0,93)	2,27 (0,89)
Effizienz	2,06 (1,01)	2,07 (1,00)	2,10 (1,00)	2,08 (0,94)
Wertigkeit	1,40 (1,74)	1,46 (1,33)	1,55 (1,27)	1,71 (1,17)
Durchschaubarkeit	2,34 (0,85)	2,20 (0,96)	2,29 (0,92)	2,27 (0,97)
Nützlichkeit	2,02 (1,12)	2,02 (1,03)	2,07 (1,03)	2,08 (1,03)
Attraktivität	1,85 (1,17)	1,76 (1,24)	1,91 (1,13)	1,89 (1,22)

Wie zu erwarten, unterscheiden sich die Wichtigkeitsurteile zu den Skalen nicht wesentlich.

Aufgrund der Stabilität der Wichtigkeitsurteile ist es eine Option, diese in zukünftigen Befragungen nicht mehr mit zu erheben, sondern die Gewichtung aus bestehenden Untersuchungen zu verwenden. Damit kann der notwendige Aufwand zum Ausfüllen des Fragebogens durch die Teilnehmer weiter reduziert werden. Eine spannende Frage ist natürlich auch, wie der Zusammenhang zwischen dem Net Promoter Score [5] und den gemessenen UX-Eigenschaften ist. Natürlich kann man nicht annehmen, dass die Bereitschaft ein Produkt weiterzuempfehlen ausschließlich von der UX des Produkts abhängt. Hier können andere Faktoren, z.B. der Preis des Produkts, die generelle Wahrnehmung der Marke, das Verhalten des Personals beim Verkaufsgespräch oder bei Service-Anfragen, eine Rolle spielen. Dieser Zusammenhang geht damit über die reinen Produktqualitäten hinaus.

Für die untersuchten digitalen Zeitungsangebote ergibt sich aber ein sehr starker Zusammenhang zwischen der UX KPI und dem NPS. Dieser zeigt sich sowohl in Bezug auf die Gesamtskala NPS10 (Werte von 0 bis 10) des NPS als auch für die eigentliche Skala NPS3 mit den drei Kategorien *Detractor* (Skalenwert 0 bis 6), *Neutral* (Skalenwert 7 bis 8) und *Promoter* (Skalenwert 9 bis 10). Die folgende Tabelle 3 zeigt die beobachteten Korrelationen für die vier Studien aus Tabelle 2.

**Tabelle 3:** Korrelationen zwischen NPS und der UX KPI für 4 Studien.

NPS Skala	Studien (Teilnehmer)			
	A (242)	B (1125)	C (1770)	D (190)
NPS10	0,77	0,81	0,73	0,56
NPS3	0,70	0,74	0,69	0,61

Die Korrelationen sind alle sehr hoch, d.h. die UX KPI ist offenbar ein sehr guter Prädiktor für die Bereitschaft das Produkt weiterzuempfehlen.

Offenbar ist also für digitale Zeitungsangebote vor allem die UX für die Weiterempfehlungsbereitschaft relevant. Das deutet darauf hin, dass die gewählten Skalen einerseits den Gesamteindruck der Abonnenten gut abdecken und andererseits weitere Faktoren eine untergeordnete Rolle für die Bereitschaft spielen, das Angebot weiterzuempfehlen.

Der UEQ+ ist erst seit 2019 verfügbar [2, 3]. Deshalb existiert noch kein zum UEQ vergleichbarer Benchmark für die Skalenmittelwerte. Ein solcher ist wegen der Modularität des Ansatzes auch schwer bereitzustellen. Nicht alle Skalen werden

gleich häufig verwendet werden und damit wird es für einige seltener verwendete Skalen auch sehr lange dauern, bis ausreichend Daten vorhanden sind.

Für die Anwendung bei *NOZ Medien & mh:n Medien* spielt das aber eine untergeordnete Rolle. Der Fragebogen wird hier kontinuierlich zur Evaluation der Produkte eingesetzt, d.h. hier entsteht durch wiederholte Messung ganz automatisch eine Wissensbasis von Untersuchungen, mit denen man bestimmen kann, wie gut ein neu evaluiertes Produkt abschneidet.

Zusätzlich wird neben dem UEQ+ auch der NPS verwendet, den man als Außenkriterium zur Berechnung von Cut-Points [8, 9] heranziehen kann. Solche Cut-Points können helfen die gemessenen Skalenmittelwerte des UEQ+ besser zu interpretieren.

Die meisten von uns werden Cut-Points aus der Medizin kennen. Typische Beispiele sind Grenzwerte für den Blutdruck oder die Konzentration von Cholesterin im Blut. Ein Cut-Point teilt gemessene Werte immer in zwei Kategorien. Zum Beispiel gelten Cholesterin-Werte von >200 (das ist der Cut-Point) als bedenklich, Werte <200 als unbedenklich. D.h. ein Cut-Point ist eine einfache Entscheidungshilfe, mit der gemessene Werte leichter interpretiert werden können. Für den UX Bereich wurden solche Cut-Points von [8, 9] als Entscheidungshilfen vorgeschlagen. Die Berechnung von Cut-Points benötigt ein Außenkriterium, was in unserem Fall der NPS ist.

Wir beschreiben im Folgenden die Berechnung von Cut-Points anhand der Daten aus der oben im Detail beschriebenen Studie. Das Verfahren ist denkbar einfach, d.h. kann einfach selbst ohne spezielle Software durchgeführt werden. Es eignet sich daher gut, um schnell zu einer Entscheidungshilfe zu kommen. Zur Beschreibung brauchen wir aber einige Definitionen.

Wir haben pro Teilnehmer  $i$  und UEQ+ Skala die Werte der Items und die Wichtigkeitsbeurteilung der Skala. Wir berechnen jetzt pro Teilnehmer einen Wert  $KPI(i)$  [2, 3], indem wir den Skalenwerte pro Person mit den relativen Wichtigkeiten multiplizieren und über alle Skalen summieren. Dieser Wert liegt zwischen -3 und +3 (7-stufige Skala). Zusätzlich haben wir für jede Person den NPS-Wert  $NPS3(i)$  in  $\{-1, 0, +1\}$ , d.h.  $NPS3(i) = +1$  steht für *Promoter*,  $NPS3(i) = 0$  für *Neutral* und  $NPS3(i) = -1$  für *Detractor*. Wir berechnen jetzt für jedes  $c$  in  $[-3, +3]$  folgende Kennwerte:

$$Sensitivity(c) = \frac{| \{i | NPS3(i) = 1 \wedge KPI(i) > c \} |}{| \{i | NPS3(i) = 1 \} |}$$

$$Specificity(c) = \frac{| \{i | NPS3(i) < 1 \wedge KPI(i) < c \} |}{| \{i | NPS3(i) < 1 \} |}$$

D.h.  $Sensitivity(c)$  ist die Anzahl der *Promoter* im Datensatz, deren KPI größer  $c$  ist, d.h. die bei Verwendung des Cut-Points  $c$  zur Prognose richtig als *Promoter* erkannt würden.  $Specificity(c)$  ist die Anzahl der *Detractors* und *Neutrals* im Datensatz, die bei Verwendung des Cut-Points  $c$  zur Prognose richtig als *Nicht-Promoter* erkannt wurden. Der Youden-Index

$$(Sensitivity(c) + Specificity(c)) - 1$$

beschreibt eine Möglichkeit beide Kennwerte gemeinsam zu optimieren. Der optimale Cut-Point ist derjenige Wert für  $c$ , für

den der Youden-Index maximal ist, d.h. dieser optimale Wert  $c$  diskriminiert am besten zwischen *Promotern* und *Nicht-Promotern*. Analog kann man einen Wert  $c'$  berechnen, der zwischen *Detractors* und *Nicht-Detractors* optimal unterscheidet. Diese beiden Werte kann man nun als einfache Entscheidungshilfe zur Bewertung der UEQ+ KPI verwenden.

D.h. ein Wert für die KPI  $> c$  deutet darauf hin, dass der betreffende Kunde sehr zufrieden ist, ein Wert  $< c'$  deutet auf einen eher unzufriedenen Kunden hin.

Die Berechnung dieser beiden Cut-Points aus zwei Studien ergab recht ähnliche Werte:

- Studie 1:  $c = 5,95$  und  $c' = 5,15$
- Studie 2:  $c = 5,95$  und  $c' = 4,95$

Diese Werte segmentieren die befragten Teilnehmer recht gut in Bezug auf die NPS Werte:

**Tabelle 4:** Segmentierung der Teilnehmer zweier Studien anhand der berechneten Cut-Points.

Studie 1	
KPI(i) > 5,95	Promoters 743, Neutrals 147, Detractors 31
5,95 > KPI(i) > 5,15	Promoters 157, Neutrals 296, Detractors 52
KPI(i) < 5,15	Promoters 24, Neutrals 112, Detractors 208
Studie 2	
KPI(i) > 5,95	Promoters 387, Neutrals 78, Detractors 12
5,95 > KPI(i) > 4,95	Promoters 101, Neutrals 242, Detractors 53
KPI(i) < 4,95	Promoters 4, Neutrals 57, Detractors 191

D.h. auch wenn hier noch kein Benchmark vorhanden ist, kann man mit dieser Methode doch zu einer einfachen Entscheidungsheuristik kommen, die hilft, UEQ+ Messergebnisse einfach zu interpretieren. Allerdings sind diese Heuristiken produktspezifisch und abhängig vom verwendeten Außenkriterium (in unserem Fall dem NPS), d.h. nicht gut auf andere Produkte übertragbar.

## 6 Kommunikation der Ergebnisse im Unternehmen

Die Ergebnisse werden in einer Präsentation für Produktmanager und Geschäftsführung aufbereitet. Die Ergebnisse der Skalen sowie die Wichtigkeitsurteile und der davon abgeleitete Handlungsdruck werden aufgezeigt. Je nach Produkt, wird verglichen, ob sich der Abo-Status oder das Alter der Teilnehmer auf die Bewertung auswirken. Es wird aufgezeigt, ob es eine Veränderung der Werte zur Vorwelle gibt.

Zusätzlich werden der erhobene NPS (Weiterempfehlungswahrscheinlichkeit) des Produkts und die erhobenen Freitextantworten präsentiert. Die Freitextantworten werden kategorisiert und aufbereitet. Gemeinsam mit den Daten der UEQ+ Erhebung ergibt sich so ein Gesamtbild. Es wird gezeigt, wie sich die Bewertung der UX und des NPS verändert hat und an welchen Stellen Handlungsdruck vorliegt. Die qualitative Auswertung der Freitextantworten gibt zudem Rückschlüsse auf konkrete Handlungsempfehlungen.

Zusammen mit dem Produktmanagement wird im Anschluss eine UX-Roadmap erstellt, anhand derer konkrete Maßnahmen festgehalten werden, die auf die aufgezeigten Probleme und Schwachstellen einzahlen.

Für das Produktmanagement ist die Erhebung der Daten ein hilfreiches Mittel, um begründete Entscheidungen im Rahmen der Produktentwicklung und Produktoptimierung treffen zu können. Aufgrund der quantitativen Methodik stellen die Ergebnisse eine gute Argumentationsgrundlage dar. Aus Sicht des UX-Managements bietet die Erhebung ein wichtiges Tool, um den Nutzern eine Stimme in der Produktentwicklung zu geben.

## 7 Zusammenfassung und Ausblick

Die Erhebung der UX mithilfe des UEQ+ dient in erster Linie dem UX-Monitoring. Es zeigt eine grobe Einordnung der wahrgenommenen UX der Produkte. Zudem kann eine Verschlechterung oder Verbesserung der UX von Produkten über einen längeren Zeitraum sichtbar gemacht werden.

Zusätzlich dienen die Ergebnisse zur Priorisierung in der Produktentwicklung. Zum einen können die ermittelten Wichtigkeitsurteile der UX-Merkmale in Relation zueinander gesetzt werden. So zeigt sich, welche UX-Merkmale für die Nutzer besonders relevant sind. Und zum anderen zeigt der Grad, zu dem die Erwartungen der Nutzer hinsichtlich einer UX-Skala erfüllt werden, an, bei welchen UX-Merkmalen Handlungsdruck besteht und wie groß dieser ist. Aufwände in der Produktentwicklung können so aus UX-Sicht priorisiert werden. Die Maßnahmen, die voraussichtlich auf UX-Merkmale mit hoher Relevanz oder starkem Handlungsdruck einzahlen, werden höher priorisiert.

Um diese Priorisierung vorzunehmen, ist es jedoch notwendig zu verstehen, welche Maßnahmen in der Produktentwicklung auf ein UX-Merkmal einzahlen. Der UEQ+ hilft, die Komplexität der Fragestellung "Was ist gute UX" zu verdeutlichen, es hat sich in der Praxis allerdings herausgestellt, dass das Ableiten von konkreten Handlungsempfehlungen alleine anhand der UEQ+ Daten schwierig ist. Es können lediglich Vermutungen darüber angestellt werden, was für die Nutzer ausschlaggebend für ihre Bewertung war. Daher wird zusätzlich mithilfe der Abfrage des NPS und der darauffolgenden Freitextantwort eine qualitative Methodik eingesetzt. Die Freitextantworten werden kategorisiert, wodurch Folgendes deutlich wird:

1. Gründe für die positive Bewertung der Promotoren.
2. Hinweise dazu, was am Produkt verändert werden müsste, damit Neutrals zu Promotoren werden.
3. Gründe für die schlechte Bewertung der Detractors.

Anhand dieser Erkenntnisse ist die Interpretation der UEQ+ Daten einfacher und es können konkrete Verbesserungsmaßnahmen abgeleitet werden. Je nachdem, wie hoch der Einfluss einer Maßnahme auf eine der UX-Skalen ist, desto höher wird die Maßnahme priorisiert.

Die Erhebung der UX, wie sie in diesem Beitrag beschrieben ist, steht bei der *NOZ Medien & mh:n Medien* jedoch noch immer am Anfang. Die Methodik und der Umgang mit den Ergebnissen muss daher weiter etabliert und sein Mehrwert geprüft, sowie der Aufwand gegenüber dem Nutzen geprüft werden. Ein Ziel ist es

daher, das Vorgehen zukünftig effizienter zu gestalten. Zudem soll die Repräsentativität der Daten verbessert werden, indem alle Kunden auf regelmäßiger Basis zu den Produkten befragt werden.

## Literatur

- [1] Winter, D., Hinderks, A., Schrepp, M., & Thomaschewski, J. (2017). Welche UX Faktoren sind für mein Produkt wichtig?. Mensch und Computer 2017-Usability Professionals.
- [2] Schrepp, M., & Thomaschewski, J. (2019). Design and Validation of a Framework for the Creation of User Experience Questionnaires. *International Journal of Interactive Multimedia & Artificial Intelligence*, 5(7).
- [3] Schrepp, M., & Thomaschewski, J. (2019). Eine modulare Erweiterung des User Experience Questionnaire. Mensch und Computer 2019-Usability Professionals.
- [4] Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and evaluation of a user experience questionnaire. In *Symposium of the Austrian HCI and Usability Engineering Group (63-76)*. Springer, Berlin, Heidelberg.
- [5] Reichheld, F. F. (2003). The one number you need to grow. *Harvard business review*, 81(12), 46-55.
- [6] Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, pp. 297-334.
- [7] Santoso, H. B., & Schrepp, M. (2019). The impact of culture and product on the subjective importance of user experience aspects. *Heliyon*, 5(9), e02434.
- [8] Thielsch, M. T., Thielsch, C., & Hirschfeld, G. (2019). How informative is informative? Benchmarks and optimal cut points for E-Health Websites. Mensch und Computer 2019-Workshopband.
- [9] Hirschfeld, G., & Thielsch, M. T. (2015). Establishing meaningful cut points for online user ratings. *Ergonomics*, 58(2), 310-320.