

## **Physischer oder virtueller Prototyp? Ein vergleichender Usability-Test am Beispiel eines Blutdruckmessgerätes**

Simon KÖNIG, Niels HINRICHER, Simon SIEBERS, Claus BACKHAUS

*Zentrum für Ergonomie und Medizintechnik, FH Münster  
Bürgerkamp 3, D-48565 Steinfurt*

**Kurzfassung:** In dieser Arbeit wurde ein vergleichender Usability-Test mit einem Blutdruckmessgerät und einer Nachbildung in der Virtuellen Realität (VR) durchgeführt. In dem Test mussten 24 Probanden jeweils drei für das Gerät typische Use Cases abarbeiten. Jeder dieser Use Case wurde in mehrere Handlungsschritte gegliedert. Ein Versuchsleiter bewertete diese hinsichtlich der Handlungskompetenz der Probanden mit Hilfe eines dreistufigen Bewertungsverfahrens. Aus diesen Bewertungen wurden Erfolgsquoten berechnet, die mittels Mann-Whitney-U-Test ( $\alpha = .05$ ) auf Unterschiede untersucht wurden. Die Unterschiede in den Erfolgsquoten erwiesen sich bei allen Use Cases sowie den einzelnen Handlungsschritten als nicht signifikant. Die Evaluation der Usability von Prototypen in der Virtuellen Realität ist somit möglich. Es gilt zu prüfen, ob die Ergebnisse dieser Arbeit auf komplexere Medizinprodukte übertragbar sind.

**Schlüsselwörter:** virtueller Prototyp, Usability-Test, virtuelle Realität, VR, Mensch-Maschine-Schnittstelle

### **1. Situation**

Schätzungsweise sterben allein in amerikanischen Krankenhäusern zwischen 44.000 und 98.000 Menschen aufgrund von Behandlungsfehlern (Kohn et. al. 2000). Menschliches Versagen gilt dabei als Hauptursache für Behandlungsfehler im Zusammenhang mit Medizinprodukten (Backhaus 2010). Um das Auftreten von Behandlungsfehlern zu vermeiden werden bereits vor der Markteinführung Nutzertests durchgeführt. Diese Nutzertests werden von allen relevanten Regularien gefordert (FDA & CDRH 2016; IEC 62366-1 2015). Ein Ansatz, um schwerwiegende Bedienfehler schon in der Entwicklung von Medizinprodukten zu vermeiden, ist die nutzerzentrierte Gestaltung. Kern dieser Methode ist die Konzentration auf die Anforderungen und Kenntnisse der Anwender (DIN EN ISO 9241-210 2020). Auf Basis dieser Anforderungen und Kenntnisse werden mehrere Prototypen entwickelt und von den Anwendern in Nutzertests evaluiert. Dies ermöglicht eine iterative Verbesserung der Gebrauchstauglichkeit eines Produktes. Die Herstellung und Evaluation dieser Prototypen ist jedoch sehr ressourcen- und kostenaufwendig (Zhou & Rau 2019). Virtuelle Prototypen könnten diesem Problem entgegenwirken, da sie schneller entwickelt und leichter abgewandelt werden können als herkömmliche Prototypen (Choi et. al. 2015).

Ein virtueller Prototyp ist eine Computersimulation eines physischen Produktes, das wie ein physischer Prototyp präsentiert, analysiert und getestet werden kann (Wang 2002). Zur Darstellung und Interaktion kann die Virtual Reality-Technologie genutzt werden. Für den Einsatz von virtuellen Prototypen im nutzerzentrierten Gestaltungsprozess müssen die Nutzer jedoch auch möglichst natürlich mit dem Produkt interagieren können. Der Einsatz von VR-Technologie und der Nutzen von Prototypen in der

Virtuellen Realität in Verbindung mit Nutzertests in der Medizintechnik sind noch nicht ausreichend erforscht.

In dieser Arbeit wurde untersucht, ob es hinsichtlich der Erfolgsquoten von Nutzern signifikante Unterschiede zwischen einem physischen und einem virtuellen Prototyp gibt.

## 2. Methode

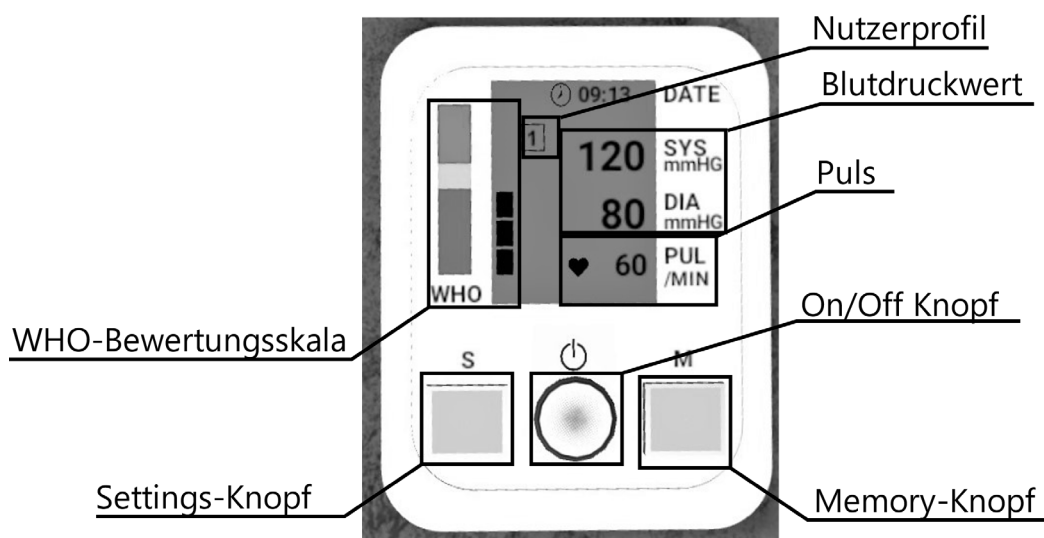
### 2.1 Versuchsaufbau

Für den Vergleich des physischen mit dem virtuellen Prototyp wurden insgesamt 24 Studierende und Angestellte der FH Münster in zwei Gruppen zu je 12 Probanden eingeteilt. Die Einteilung erfolgte randomisiert unter Beachtung der Gleichverteilung des Alters, des Geschlechtes und der Nutzung einer Sehhilfe. In der Gruppe des physischen Nutzertests besaßen nach eigener Auskunft zwei Probanden keine, sieben wenig und drei viel Vorerfahrung mit Blutdruckmessgeräten. In der Gruppe des virtuellen Nutzertests hatte ein Proband keine Vorerfahrung, neun Probanden wenig und zwei viel Erfahrung. Des Weiteren besaßen nach eigener Auskunft sechs Probanden keine, vier wenig und zwei viel Erfahrung im Umgang mit VR-Technologie.

Für die Durchführung des physischen Nutzertests kam ein Blutdruckmessgerät der Fa. Mondpalast zum Einsatz. Die Interaktionen der Probanden mit dem Blutdruckmessgerät wurden mit einer Kamera (GoPro Hero 5, GoPro Inc., USA) aufgezeichnet.

Für den virtuellen Nutzertest wurde eine HTC Vive (HTC Corporation, Taiwan) verwendet. Der Kopfumfang und der Linsenabstand des Head-Mounted-Displays (HMD) wurde für jeden Probanden individuell eingestellt. Der virtuelle Nutzertest wurde mittels Bildschirmaufnahme auf einem PC dokumentiert.

Der virtuelle Prototyp wurde in der Entwicklungsumgebung Unity (Vers. 2019.3.0a7, Unity Technologies, USA) erstellt und mit C# (Microsoft Corporation, USA) programmiert. Für die Interaktion in der VR wurde ein Infrarotsensor (Leap Motion Controller, Leap Motion Inc., USA) verwendet, der die Hände der Probanden erfasst und in die VR-Umgebung projiziert. Der virtuelle Prototyp ist in Abbildung 1 dargestellt und entspricht dem physischen Prototypen in Gestalt und Funktion.



**Abbildung 1:** Bildschirmaufnahme des virtuellen Prototyps mit Funktionsbeschreibung.

Bei Tastendruck erhielten die Probanden technisch bedingt kein haptisches Feedback, sondern nur ein Klick-Geräusch als akustisches Feedback. Da mit dem virtuellen Prototyp keine tatsächliche Blutdruckmessung möglich war, wurde bei Auslösung ein zufälliger systolischer Wert zwischen 110 und 180 mmHg generiert. Abhängig von der Systole wurde ein diastolischer Wert in einem physiologisch sinnvollen Rahmen generiert und ein Puls zwischen 50 und 80 festgelegt.

## 2.2 Versuchsablauf

Der Versuchsablauf war für beide Nutzertests identisch. Im Anschluss an eine standardisierte Einführung, in der die Probandendaten erhoben und der Versuchsablauf erklärt wurde, erfolgte die Bearbeitung der Use Cases „Blutdruck messen“, „Blutdruck aufrufen“ und „Benutzer wechseln“.

Die drei Use Cases waren in einzelne Handlungsschritte gegliedert, welche nacheinander von den Probanden bearbeitet wurden. Die zu dem jeweiligen Use Case gehörigen Handlungsschritte sind der Tabelle 2 im Ergebnisteil zu entnehmen.

## 2.3 Bewertung der Handlungskompetenz

Jeder Handlungsschritt wurde durch den Versuchsleiter während der Tests nach einem 3-Punkte-Schema bewertet. Das Bewertungsschema beinhaltete die Bewertungsstufen „Gut“, „Mittel“ und „Schlecht“. Tabelle 1 nennt die Kriterien, auf deren Grundlage die Bewertung der Handlungsschritte erfolgte.

**Tabelle 1:** *Verwendete Kriterien zur Bewertung der Handlungsschritte bezüglich der Handlungskompetenzen der Probanden.*

Bewertung	Beschreibung
Gut	<ul style="list-style-type: none"><li>- Zügige Bedienung ohne Hilfestellung</li><li>- Fehlerfreie Durchführung</li></ul>
Mittel	<ul style="list-style-type: none"><li>- Längeres zögern vor der Bedienung</li><li>- Fehler werden ohne Hilfe durch den Versuchsleiter korrigiert</li></ul>
Schlecht	<ul style="list-style-type: none"><li>- Ausführung nach Hilfestellung des Versuchsleiters</li><li>- Fehler wird nicht selbstständig korrigiert</li></ul>

Aus den relativen Häufigkeiten der in Tabelle 1 genannten Bewertungsstufen wurden nach Nielsen (2001) Erfolgsquoten berechnet:

$$\text{Erfolgsquote} = \frac{\sum \text{Gut} + \sum \text{Mittel} * 0,5}{\text{Probanden} * \text{Handlungsschritte}} * 100$$

Die Erfolgsquoten wurden sowohl für jeden Use Case als auch für jeden Handlungsschritt separat berechnet. Mittels Mann-Whitney-U-Test ( $\alpha = .05$ ) wurden die Erfolgsquoten des virtuellen und des physischen Nutzertests auf Unterschied getestet. Es wurden folgende Hypothesen überprüft.



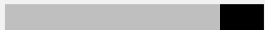

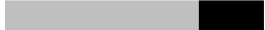


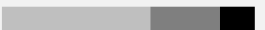


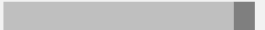
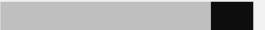




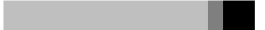






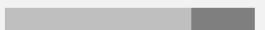


**H<sub>1</sub>:** Es gibt einen Unterschied in den Erfolgsquoten der einzelnen Use Cases zwischen dem physischen und dem virtuellen Nutzertest. **Mit H<sub>01</sub>:** Es gibt keinen Unterschied in den Erfolgsquoten der einzelnen Use Cases zwischen dem physischen und dem virtuellen Nutzertest.

**H<sub>2</sub>:** Es gibt einen Unterschied in den Erfolgsquoten der einzelnen Handlungsschritte zwischen dem physischen und dem virtuellen Nutzertest. **Mit H<sub>02</sub>:** Es gibt keinen Unterschied in den Erfolgsquoten der einzelnen Handlungsschritte zwischen dem physischen und dem virtuellen Nutzertest.

### 3. Ergebnisse

Die Ergebnisse des Vergleichs der Erfolgsquoten im physischen und virtuellen Nutzertest sind Tabelle 2 zu entnehmen. Die relativen Häufigkeiten der Bewertungsstufen (vgl. Tabelle 1) sind als Balkendiagramme dargestellt.

**Tabelle 2:** Vergleich der Erfolgsquoten im physischen und im virtuellen Nutzertest.

Use-Case	Handlungsschritte	Erfolgsquote (physisch)	Erfolgsquote (virtuell)	Sig.
Blutdruck messen	1. Messwert aufnehmen	 92%	 88%	.775
	2. Messwert aufnehmen	 92%	 63%	.128
	3. Messwert aufnehmen	 88%	 67%	.242
	<b>Gesamt</b>	 90%	 72%	.060
Blutdruck aufrufen	Durchschnittswert finden und ablesen	 42%	 79%	.068
	vorletzten Messwert finden und ablesen	 96%	 83%	.713
	höchsten Messwert finden und ablesen	 100%	 92%	.514
	niedrigsten Messwert finden und ablesen	 100%	 96%	.755
	<b>Gesamt</b>	 84%	 88%	.378
Benutzer wechseln	Einstellungsmenü aufrufen	 29%	 54%	.198
	Benutzerwechsel vornehmen	 83%	 75%	.671
	Auswahl bestätigen	 79%	 88%	.671
	<b>Gesamt</b>	 64%	 72%	.410

Die Ergebnisse des Mann-Whitney-U-Tests ( $\alpha = .05$ ) belegen, dass für alle untersuchten Use Cases und Handlungsschritte die Nullhypothesen **H<sub>01</sub>** und **H<sub>02</sub>** beibehalten werden müssen. Es konnten keine signifikanten Unterschiede bezüglich der Erfolgsquoten festgestellt werden.

## 4. Diskussion

Eine deutliche Limitation des virtuellen Nutzertests verglichen mit dem realen Test war das fehlende haptische Feedback. Dies könnte insbesondere einen Einfluss auf den Use Case „Benutzer wechseln“ gehabt haben. Beim Handlungsschritt „Einstellungsmenü aufrufen“ mussten die Probanden die mittlere Taste für zwei Sekunden gedrückt halten. Der Tastendruck war dabei aufgrund der Verdeckung durch die Hand nur unzureichend ersichtlich und zudem nicht ertastbar. Einige Probanden interagierten durch das eingeschränkte Feedback so langsam mit dem Gerät, dass das Menü möglicherweise unbewusst geöffnet wurde. Durch die Verwendung von VR-Handschuhen mit haptischem Feedback, z.B. mittels Resonanzaktuatoren, könnte die Bedienung der Geräte, insbesondere bei Stellaufgaben wie dem Gedrückt halten eines Tasters, realitätsgetreuer umgesetzt werden.

Die Darstellung von Textelementen geringer Größe, wie sie auf dem Prototyp zu finden sind, ist mit der HTC Vive limitiert. Dies führt zu Schwierigkeiten bei der Lesbarkeit. Um die Lesbarkeit von Textelementen in der VR zu erhöhen ist eine höhere Bildschirmauflösung notwendig, die neuere HMD-Modelle wie beispielsweise die Valve Index (Valve Corporation, USA) bieten.

Trotz der Limitationen durch das fehlende haptische Feedback und der begrenzten Auflösung konnte kein signifikanter Einfluss der VR auf die Erfolgsquoten und somit auf die Handlungskompetenz der Nutzer nachgewiesen werden. Prototypen in der VR könnten somit die klassischen Usability-Tests mit physischen Prototypen sinnvoll ergänzen oder bei idealer digitaler Umsetzung sogar ersetzen. Die Arbeiten von Kuutti et al. (2001) sowie von Bruno und Muzzupappa (2010) stützen diese Annahme. Kuutti et al. (2001) kommen in ihrer Studie zu dem Schluss, dass ihr virtueller 2D-Prototyp eines Mobiltelefons grundsätzlich zur Identifizierung von Usability-Problemen geeignet ist. Die Qualität der Simulation sei jedoch maßgeblich für den Testerfolg verantwortlich. Bruno und Muzzupappa (2010) führten virtuelle Usability-Tests mit den Benutzeroberflächen eines Mikrowellenherdes und einer Waschmaschine durch und bezeichnen die VR als eine valide Alternative zu herkömmlichen Usability-Tests.

Es gilt zu prüfen, ob diese Ergebnisse auch auf komplexere Medizinprodukte übertragbar sind. Auch der Einfluss der VR auf weitere Usability-Attribute wie beispielsweise Effektivität, Erlernbarkeit, Anwenderakzeptanz, User Experience oder Anwenderbelastung müssen weiter untersucht werden.

## 5. Literatur

- Backhaus C (2010) Usability-Engineering in der Medizintechnik: Grundlagen – Methoden – Beispiele. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg.
- Bruno F, Muzzupappa M (2010) Product interface design: A participatory approach based on virtual reality. In: International Journal of Human-Computer Studies 68 (5), S. 254–269. DOI: 10.1016/j.ijhcs.2009.12.004.
- Choi S, Jung K, Noh SD (2015) Virtual reality applications in manufacturing industries: Past research, present findings, and future directions. In: Concurrent Engineering 23 (1), S. 40–63. DOI: 10.1177/1063293X14568814.
- DIN EN ISO 9241-210 (2020) Ergonomie der Mensch-System-Interaktion – Menschzentrierte Gestaltung interaktiver Systeme.
- FDA, CDRH (2016) Applying Human Factors and Usability Engineering to Medical Devices – Guidance for Industry and Food and Drug Administration Staff. <https://www.fda.gov/media/80481/download>.
- IEC 62366-1 (2015) Medical Devices. Part 1: Application of Usability Engineering to Medical Device
- Kohn LT, Corrigan JM, Donaldson MS (2000) To Err is Human: Building a Safer Health System. DOI: 10.17226/9728.

- Kuutti K, Battarbee K, Säde S, Mattelmäki T, Keinonen T, Teirikko T, Tornberg A-M (2001) Virtual prototypes in usability testing. In: Proceedings of the 34<sup>th</sup> Annual Hawaii International Conference on System Sciences. DOI: 10.1109/HICSS.2001.926546.
- Nielsen J (2001) Success rate: the simplest usability metric. In: Jakob Nielsen's Alertbox 18. <https://www.nngroup.com/articles/success-rate-the-simplest-usability-metric/>.
- Wang GG (2002) Definition and Review of Virtual Prototyping. In: Journal of Computing and Information Science in Engineering 2 (3), S. 232–236. DOI: 10.1115/1.1526508.
- Zhou X, Rau P-LP (2019) Determining fidelity of mixed prototypes: Effect of media and physical interaction. In: Applied ergonomics 80, S. 111–118. DOI: 10.1016/j.apergo.2019.05.007.



Gesellschaft für  
Arbeitswissenschaft e.V.

## Arbeit HUMAINE gestalten

67. Kongress der  
Gesellschaft für Arbeitswissenschaft

Lehrstuhl Wirtschaftspsychologie (WiPs)  
Ruhr-Universität Bochum

Institut für Arbeitswissenschaft (IAW)  
Ruhr-Universität Bochum

3. - 5. März 2021

---

## GfA-Press

---

**Bericht zum 67. Arbeitswissenschaftlichen Kongress vom 3. - 5. März 2021**

**Lehrstuhl Wirtschaftspsychologie, Ruhr-Universität Bochum  
Institut für Arbeitswissenschaft, Ruhr-Universität Bochum**

Herausgegeben von der Gesellschaft für Arbeitswissenschaft e.V.  
Dortmund: GfA-Press, 2021  
ISBN 978-3-936804-29-4

NE: Gesellschaft für Arbeitswissenschaft: Jahresdokumentation

Als Manuskript zusammengestellt. Diese Jahresdokumentation ist nur in der Geschäftsstelle erhältlich.

Alle Rechte vorbehalten.

© **GfA-Press, Dortmund**

**Schriftleitung: Matthias Jäger**

im Auftrag der Gesellschaft für Arbeitswissenschaft e.V.

Ohne ausdrückliche Genehmigung der Gesellschaft für Arbeitswissenschaft e.V. ist es nicht gestattet:

- den Kongressband oder Teile daraus in irgendeiner Form (durch Fotokopie, Mikrofilm oder ein anderes Verfahren) zu vervielfältigen,
- den Kongressband oder Teile daraus in Print- und/oder Nonprint-Medien (Webseiten, Blog, Social Media) zu verbreiten.

Die Verantwortung für die Inhalte der Beiträge tragen alleine die jeweiligen Verfasser; die GfA haftet nicht für die weitere Verwendung der darin enthaltenen Angaben.

**Screen design und Umsetzung**

© 2021 fröse multimedia, Frank Fröse

[office@internetkundenservice.de](mailto:office@internetkundenservice.de) · [www.internetkundenservice.de](http://www.internetkundenservice.de)