

Ausarbeitungsleitfaden für Nutzerstudien zur Evaluation von XR-Interfaces in der Produktentwicklung

Jakob Harlan, Dr.-Ing. Benjamin Schleich, Prof. Dr.-Ing. Sandro Wartzack

Abstract

Technologien der erweiterten Realität (extended reality, XR) finden im gesamten Produktentwicklungsprozess Anwendung. Insbesondere Systeme zur aktiven Erzeugung und Veränderung digitaler Produktdaten bieten noch viel Potential. Die Erforschung und Entwicklung dieser immersiven Interfaces beruht maßgeblich auf der Evaluation durch Nutzerstudien, denn nur so kann die Einsatztauglichkeit der Mensch-Maschine-Schnittstellen seriös beurteilt und verglichen werden. Bei der Konzeptionierung und Durchführung dieser Nutzerstudien gibt es viel zu beachten. In dieser Arbeit wird ein Leitfaden für das Design von Evaluationen von XR Interfaces für die Produktentwicklung präsentiert. Zu Beginn müssen die Fragestellungen festgelegt werden, welche die Studie beantworten soll. Ausgehend von diesen müssen die zu testenden Versuchsbedingungen, die gestellten Aufgaben, die aufgenommenen Metriken, die gewählten Probanden und der geplante Ablauf festgelegt werden. Zusätzlich zu der allgemeinen Darlegung wird das Vorgehen anhand eines Fallbeispiels angewandt. Die Gestaltung einer Nutzerstudie zur Evaluation der Usability und Eignung eines neuartigen Virtual Reality Interfaces zur natürlichen Gestaltung von Vorentwürfen wird vorgestellt.

Keywords: Virtuelle Realität, Erweiterte Realität, Nutzerstudie, Produktentwicklung, Mensch-Maschine-Schnittstelle.

1 Einleitung

Technologien der erweiterten Realität (extended reality, XR) bieten für den gesamten Produktentwicklungsprozess viel Potential, vom VR Konzeptdesign bis zur AR Wartungsunterstützung (Coburn et al., 2017). Es befinden sich bereits eine Vielzahl solcher Technologien vornehmlich zur Betrachtung und Analyse von 3D Produktmodellen, insbesondere für Design Reviews, im Einsatz. Schnittstellen (Interfaces) zum aktiven Erzeugen und Modifizieren von Produkt- und Bauteilgeometrie haben allerdings noch einen geringeren Reife- und Verbreitungsgrad. Die Entwicklung und Erforschung solcher Systeme ist sehr aufwändig, was nicht zuletzt an noch jungen Entwicklungsplattformen liegt. So wird relativ viel Zeit benötigt, um die angedachten Funktionen der Interfaces zu implementieren, was dazu führt, dass wenig Zeit zur eigentlichen Evaluation des Interfaces bleibt (Ashtari et al., 2020). Doch nur durch entsprechende Nutzerstudien kann seriös die Einsatztauglichkeit der entwickelten Schnittstellen beurteilt werden.

Aufgrund der Vielzahl der zu berücksichtigenden Aspekte stellt insbesondere für Forscher mit Produktentwicklungshintergrund die Konzeptionierung und Durchführung geeigneter Nutzerstudien oftmals eine Herausforderung dar. Daher sollen in dieser Arbeit wichtige Fragestellungen bei der Gestaltung solcher Nutzerstudien offengelegt, relevante Literatur präsentiert und ein Leitfaden zur Ausarbeitung gegeben werden. Zusätzlich zur allgemeinen Darlegung wird ein konkretes Fallbeispiel eines VR-Interfaces zur Erstellung von Vorentwürfen präsentiert.

2 Stand der Technik

Nutzerstudien bilden ein zentrales Werkzeug der nutzerorientierten Gestaltung und sind seit der Entwicklung von graphischen Benutzeroberflächen essenziell zur Verbesserung deren Nutzbarkeit. Viel zitierte Standardwerke geben einen umfangreichen Einblick in die Möglichkeiten zur Evaluation von Mensch-Maschine-Schnittstellen (Albert & Tullis, 2013; Sauro & Lewis, 2016). Wie ein Großteil der Literatur, behandeln diese klassische zweidimensionale Interfaces, vieles ist aber auf immersive Applikationen übertragbar und auch Vorgehen speziell für diese wurden bereits zusammengestellt (Gabbard et al., 1999). Zur Bewertung der User Experience wird grundsätzlich zwischen der formativen und summativen Evaluation unterschieden. Formative Erhebungen werden kontinuierlich während der Interfaceentwicklung durchgeführt. Hierbei werden durch einen Entwickler oder in minimalen Nutzertests die größten Nutzbarkeitsprobleme identifiziert und ausge bessert. Eine summative Analyse beschreibt hingegen die Evaluation eines finalisierten Interfaces durch Aufnahme von Kenngrößen und dem Vergleich dieser mit Zielsetzungen und ggf. auch alternativen Systemen (Albert & Tullis, 2013).

Nutzertests stellen nur eine Evaluationsmethode von Vielen dar. Weniger aufwendige Alternativen basieren darauf das Nutzerverhalten mit Hilfe von Heuristiken vorherzusagen. Dabei können entweder analytische Methoden verwendet werden (Blandford et al., 2008) oder das untersuchte System von Experten bewertet werden (Sutcliffe & Gault, 2004). Die aufwändigste Methode der Evaluation ist die Feldstudie. Hierbei wird ein neuartiges XR-Interface einer Produktivabteilung in der Industrie zu Verfügung gestellt, der Eindruck der Nutzer erfragt und der Einfluss auf die entstandenen Produkte analysiert. Für die Evaluierung wissenschaftlicher Prototypen hat sich allerdings der Vergleich durch Nutzerstudien in kontrollierter Umgebung etabliert. Beispielsweise wurde in (Toma et al., 2012) sowohl Teil- als auch Baugruppenmodellierung in einem immersive CAVE Umgebung vorgestellt und gegen klassisches Desktop-CAD evaluiert. Feeman et al. (2018) haben einen VR-CAD Modellierer für moderne VR-Brillen implementiert. In ihrer Studie wurde mit Hilfe von Juroren untersucht ob Nutzer kreativer in

VR als in klassischer Umgebung arbeiten. In einer aktuellen Studie haben Seybold und Mantwill (2020) in virtueller Realität erzeugbare 3D-Skizzen getestet.

Wie man sieht, gibt es schon einige Arbeiten mit Nutzerstudien zur Evaluation von immersiven Mensch-Maschine-Schnittstellen im Kontext der Produktentwicklung. Was zurzeit noch fehlt ist eine generelle Vorgehensweise, anhand derer die Nutzerstudien einheitlich erarbeitet werden können. Motiviert durch den Stand der Technik widmet sich der Beitrag der Forschungsfrage, wie Nutzerstudien zur Evaluation von XR-Interfaces für die Produktentwicklung systematisch anhand eines Leitfadens konzipiert werden können. Die erarbeiteten Erkenntnisse erlauben es damit Produktentwicklungswissenschaftler/-innen, ihre entwickelten XR-Interfaces schneller und zielgerichteter evaluieren zu können.

3 Ausarbeitungsleitfaden für Nutzerstudien

Im Folgenden werden sechs Aspekte zur strukturierten Gestaltung einer Nutzerstudie detailliert vorgestellt, zu welchen Abbildung 1 eine Übersicht liefert.

	1. Fragestellung - Was soll untersucht werden
	2. Versuchsbedingungen - Welche Konfigurationen werden verglichen?
	3. Aufgabenstellung - Was müssen die Probanden machen?
	4. Metriken - Welche Ergebnis- und Leistungsorientierten Größen werden erhoben?
	5. Probanden - Wer soll teilnehmen? Wie viele Teilnehmer werden benötigt?
	6. Ablauf - Wie läuft die Studie pro Probanden ab?

Abbildung 1: Ausarbeitungsleitfaden für Nutzerstudien von immersiven Interfaces für die Produktentwicklung in 6 Schritten.

3.1 Fragestellungen

Die Planung einer Nutzerstudie muss damit begonnen werden, ihre Ziele festzustecken. Dafür können Fragestellungen formuliert werden, die die Studie beantworten soll. Die *Design Research Methodology* hilft bei der Findung und Argumentation der Fragestellungen (Blessing & Chakrabarti, 2009). Alle nachfolgenden Entscheidungen in Bezug auf das Studiendesign sollten der Beantwortung der jeweils gestellten Fragestellung dienen. Eine oft gestellte Frage ist, ob die Nutzbarkeit eines Interface ausgereift genug ist. Weitere beispielhafte Fragestellungen sind: Werden mit einem neuartiges System Produktmodelle in einer höheren Qualität erzeugt, Zusammenhänge schneller erkannt oder Arbeitsschritte effizienter erledigt?

3.2 Versuchsbedingungen

Um die Studienergebnisse einordnen zu können, hilft es unterschiedliche Bedingungen miteinander zu vergleichen. Immersive Interfaces können in vielen Fällen gegen ihre etablierten Gegenstücke getestet werden, beispielsweise VR Bauteil- und Baugruppenmodellierung im Vergleich zu klassischen CAD-Tools (Feeman et al., 2018; Toma et al., 2012). Dabei gibt sowohl die Möglichkeit, dass alle Testpersonen alle Bedingungen testen (*Within-Subjects*) oder dass sie jeweils nur einen Teil der Versuchsbedingung testen (*Between-Subjects*) wenn der Lerneffekt zu groß wäre. Wird *within-subjects* getestet müssen, um ein Bias durch die Reihenfolge zu verhindern, Versuchs-Konfigurationen durchmischt werden.

3.3 Aufgabenstellung

Um die Ergebnisse der Teilnehmer vergleichen zu können, sollten alle Probanden die gleiche Aufgabe ausüben. Bei der Auswahl der Aufgabenstellung gilt es zwischen Realitätsnähe, Studiendauer und Auswertbarkeit abzuwägen. So sollte die Aufgabe nicht zu umfangreich und komplex sein, sodass es im zeitlich begrenzten Rahmen einer Studie verstanden und bearbeitet werden kann. Allerdings darf die Aufgabe nicht zu stark vereinfacht werden, um die Übertragbarkeit auf reale Anwendungsfälle zu gewährleisten. Das verwendete Beispiel sollte sich für die Bearbeitung mit dem neuartigen Interface anbieten, allerdings nicht dafür maßgeschneidert sein, um einen fairen Vergleich zu bieten. Auch ist es wichtig die Zielsetzung für den Nutzer klar zu kommunizieren. Ob ein Nutzer seine Aufgabe möglichst schnell, möglichst genau oder möglichst kreativ umsetzen soll hat erheblichen Einfluss auf die dargelegte Leistung. Es sollte die geeignetste Zielformulierung für den getesteten Anwendungsfall gewählt und besonderes Augenmerk daraufgelegt werden, dass alle Teilnehmer die gleiche Formulierung präsentiert bekommen. Werden mehrere Aufgaben gestellt, sollten auch diese zur Vermeidung von Lerneffekten gemischt werden.

3.4 Metriken

Da grob formulierte Fragestellungen oft nicht direkt beantwortet werden können, müssen bestimmbare Größen gefunden werden, welche in der Studie erhoben und im Anschluss ausgewertet werden. Dabei kann zwischen ergebnis- bzw. leistungsorientierten Metriken (*User Performance*) und nutzerorientierten Metriken unterschieden werden (*User Satisfaction*) (Albert & Tullis, 2013).

3.4.1 User Performance

Ergebnis- bzw. leistungsorientierte Metriken messen auf unterschiedliche Arten die Erfüllung einer Aufgabe. Am direktesten lässt sich dies messen, wenn für eine Aufgabe

gezählt werden kann, wie oft sie erfolgreich bzw. fehlerhaft ausgeführt wurde. Zum Beispiel aus wie vielen Versuchen eine Zielschaltfläche wie oft korrekt ausgewählt wird. Ob ein neuartiges Interface Effizienzvorteile bietet, kann meist ebenfalls gut gemessen werden. Hier bietet es sich an, die Zeit, die Nutzer für eine spezielle Aufgabe brauchen, aufzunehmen. Andersherum kann auch gezählt werden wie viele (erfolgreiche) Arbeitsschritte in einer festgelegten Zeit geschafft werden (Feeman et al., 2018; Toma et al., 2012). Eine weitere Option ist die Bestimmung der erreichten Präzision durch Messen der Abweichung von einer Zielstellung, beispielsweise beim Platzieren eines Objektes.

Demgegenüber können subjektive Leistungsmerkmale, wie beispielsweise die Kreativität der durch das XR-Interface erzeugten Ergebnisse, oft nicht direkt gemessen werden. Eine etablierte Technik zur Erhebung dieser Daten ist die *Consensual Assessment Technique* (CAT) (Amabile, 1982). Hierbei werden nach der Durchführung der Studie alle Resultate einem Gremium vorgelegt. Diese bewerten unabhängig voneinander die anonymen Ergebnisse hinsichtlich gewünschter Kriterien. Das können Eigenschaften wie Kreativität, Tauglichkeit und auch Ästhetik sein. Die Bewertenden sollten sich in der Domäne auskennen, der Umfang ihrer Expertise darf aber unterschiedlich sein. Je nach getesteten Anwendungsfall kann diese Jury aus dem Forscherumfeld oder der Industrie gewonnen werden.

3.4.2 User Satisfaction

Nutzerorientierte Messgrößen bestimmen unterschiedlichste Gefühle, Meinungen und Erfahrungen der Anwender. Diese haben großen Einfluss auf den Erfolg eines Interfaces, denn unabhängig von der Qualität der Ergebnisse wird sich eine Anwendung, mit welcher Anwender unzufrieden sind, nicht gegen die Konkurrenz durchsetzen (Albert & Tullis, 2013). Die üblichste Form Empfindungen der Nutzer aufzunehmen sind Fragebögen, die die Anwender nach der Erfüllung einer Aufgabe nach ihrem eigenen Ermessen ausfüllen. Für viele unterschiedliche Eigenschaften gibt es bereits etablierte Fragebögen:

- Zur Bestimmung von Nutzbarkeit (Usability) sind *System Usability Scale* (SUS) (Brooke, 1996) und *USE Questionnaire* (Lund, 2001) weit verbreitet. Während ersterer mit 10 Fragen nur konkret die Nutzbarkeit aufnimmt gibt der zweite durch 30 Fragen getrennt Aufschluss über Nützlichkeit, Benutzerfreundlichkeit, Lernbarkeit und Zufriedenheit des Interfaces. Da XR-Schnittstellen oft durch Begeisterung und Faszination punkten können, kann sich der ausführlichere Fragebogen lohnen.

- Kognitive Belastung kann beispielsweise mit Hilfe des NASA-Task Load Index (NASA-TLX) erhoben werden (Hart & Staveland, 1988). Dies ist besonders relevant für intuitive immersive Anwendungen, die es versuchen dem Nutzer zu ermöglichen der Bedienung eines Systems weniger mentale Aufmerksamkeit schenken zu müssen.
- Ein hohes Präsenzepfinden der Nutzer in der virtuellen Welt sorgt für Zufriedenheit und verbesserte Leistung. Eine ganze Reihe konkurrierender Fragebögen existieren über welche Lombard et al. im Zuge der Vorstellungen ihres eigenen Fragenkatalogs *The Temple Presence Inventory* einen guten Überblick geben (Lombard et al., 2009).
- Einen guten Überblick in Bezug auf weitere relevante Fragestellungen hinsichtlich der Empfindungen der Nutzer bietet der von Tcha-Tokey et al. vorgestellte Ansatz. Bei diesem auf immersive Systeme spezialisierten Fragebogen werden 82 Fragen, in 10 Subskalen, gestellt (Tcha-Tokey et al., 2016).

Sollen die Nutzererfahrungen offen und flexibel aufgenommen werden, kann ein *Think Aloud* verwendet werden. Hierbei wird der Nutzer von der Studienleitung zum kontinuierlichen Verbalisieren seiner Gedanken motiviert. Audio- oder Videoaufnahme des Vorganges erleichtert das Auswerten. Diese Technik liefert Rückmeldung zu allen Aspekten des Interfaces und hilft Stärken und Schwächen aufzudecken. Die Erkenntnisse sind allerdings aufgrund ihrer fehlenden Struktur nur schwer vergleichbar. Ganz im Gegensatz dazu stehen objektiv messbare nutzerorientierte Metriken, wie beispielsweise *Eye-* und *Motiontracking*. Hierbei werden physische Vorgänge aufgenommen und bieten eine objektive Grundlage zur Analyse und zum Vergleich. *Motiontracking* Daten können etwa verwendet werden, um die Ergonomik eines XR-Interfaces zu bewerten.

3.5 Probanden

In der Literatur gibt es unterschiedliche Meinungen dazu, wie viele Probanden für eine Nutzerstudie nötig sind. Es wird berichtet, dass schon mit 5 Probanden bis zu 80 Prozent aller Usability Probleme erkannt werden können (Virzi, 1992), andere Autoren fordern 10 Probanden oder mehr (Faulkner, 2003). Auch für die Anzahl der Teilnehmer sollte sich an der Zielstellung der Studie orientiert werden: Sollen nur die größten Usability Probleme aufgedeckt werden, um das Interface zu verbessern, genügen bereits kleine Studien mit ca. 5 Probanden. Mehrere solcher kleinen, formativen Test sind oft effektiver als eine größere Studie. Zum einen werden so kleinere Änderungen am Interface direkt getestet und zum anderen können in einem einzelnen Test kleinere Probleme unentdeckt bleiben, wenn größere Defizite des Interfaces die Nutzbarkeit

sehr einschränken (Albert & Tullis, 2013). Soll allerdings die Effektivität eines neuartigen Systems für einen konkreten Anwendungsfall mit Hilfe von messbaren Metriken stochastisch untersucht werden, sind deutlich mehr Probanden nötig. Beispielsweise testeten die im Stand der Technik vorgestellten Studien von immersiven XR-Interfaces in der Produktentwicklung zwischen 8 (Toma et al., 2012) und 27 (Seybold & Mantwill, 2020) Probanden.

Neben der reinen Anzahl der Probanden gilt es auch zu beachten, welcher Personenkreis getestet werden soll. Ob sich ein Tool an beispielsweise bestimmte Branchenspezialisten, allgemeine Produktentwickler oder beliebige Privatanwender richtet, sollte sich auch in den Testpersonen widerspiegeln. Um Vorerfahrungen in die Auswertung miteinfließen zu lassen können, sollten diese vor der Studie per Fragebogen aufgenommen werden. Gerade die Erfahrungen mit XR-Technologie spielt eine wichtige Rolle. Insbesondere Probanden, für die diese Technologie neu ist, sind oft fasziniert abgelenkt und es fällt ihnen schwer, sich auf die Aufgabe zu konzentrieren (Ashtari et al., 2020).

3.6 Ablauf

Ein Ablaufplan hilft dem durchführenden Wissenschaftler, bei jedem Probanden die gleichen Bedingungen zu schaffen und nichts Wichtiges zu vergessen. Dieser kann pro Probanden erzeugt werden, um auch direkt die Durchmischung der Versuchsbedingungen und Aufgabenstellungen abzudecken. Üblicherweise wird nach der Begrüßung als erstes eine Einverständniserklärung zur Teilnahme und eine Belehrung unterschrieben. Für Studien mit XR-Interfaces muss in der Belehrung auf potenzielle VR-Krankheitserscheinungen hingewiesen werden und dem Probanden klar sein, dass die Studie beim Aufkommen von Unwohlsein jederzeit abgebrochen werden kann. Nach diesen Formalia können relevante personenbezogene Daten erhoben werden (Alter, XR-Erfahrung etc.), die Interfaces ausprobiert, die Aufgaben durchgeführt und abschließende Fragebögen ausgefüllt werden.

4 Fallbeispiel

Laut VDI 2223 verwenden Konstrukteure zur Entwicklung von Vorentwürfen Handskizzen. CAD Programme seien für diesen Schritt ungeeignet da sie eine zu hohe Gestaltfestlegung fordern (VDI, 2018). Wir haben ein VR-Interface entwickelt, das eine intuitive und bewusst unscharfe Gestaltung von Vorentwürfen mit natürlichen Handgesten ermöglicht. Dieses soll die Vorteile der Handskizze mit denen eines CAD verknüpfen, so dass digitale dreidimensionale Modelle schnell und einfach gestaltet werden können.



Abbildung 2: Interaktionen im VR-System zu Erstellung von Vorentwürfen

In Abbildung 2 sind dessen Funktionalitäten dargestellt. Das System ermöglicht die Erzeugung und direkte Dimensionierung von geometrischen Primitiven. Die so erzeugten Bausteine können durch natürliche Hand- und Fingerinteraktionen bewegt, gefärbt, gelöscht und mit booleschen Operationen kombiniert werden (Harlan et al., 2020). Durch die intuitive Bedienung und spielerischen Interaktionen soll es Anwendern unabhängig ihrer Erfahrungen in dreidimensionaler Modellierung möglich sein, Bauteil- oder Produktideen als räumlichen Vorentwurf zu erzeugen. Mit Hilfe des in Abschnitt 3 vorgestellten Ausarbeitungsleitfadens ist die Evaluation dieses VR-Interfaces geplant worden und wird im Folgenden vorgestellt.

4.1 Fragestellungen

Die Nutzerstudie soll auf drei Fragestellungen Antworten liefern:

- Ist das Interface ausgereift genug, sodass Nutzer es ohne Probleme verwenden können?
- Unterscheiden sich handskizzierte Vorentwürfe von Modellen, die mit dem Interface erstellt wurden hinsichtlich ihrer Eignung und Kreativität?
- Sind Produkterwickler daran interessiert das neuartige System einzusetzen?

4.2 Versuchsbedingungen und Aufgabenstellung

Entsprechend der Motivation durch die VDI 2223 wird die traditionelle Handskizze mit dem immersiven System verglichen. Dabei werden drei unterschiedliche Aufgaben in jeder Versuchsbedingungen ausgeführt.

Tabelle 1: Versuchsbedingungen und Aufgabentypen der gestalteten Studie.

Vorgehensweisen	Aufgabentypen
Perspektivische Handskizze auf Papier mit Bleistift und Radiergummi.	Nachskizzieren eines vorgegebenen Produktes
VR-Vorentwurf in dem entwickelten Interface.	Änderungsskizze an einem gegebenen Modell / einer gegebenen Zeichnung.
	Skizzieren eines neuen Modells dessen Anforderungen vorgegeben wird.

Die unterschiedlichen Aufgabentypen sollen dabei helfen die Vorgehensweisen in verschiedenen Szenarien zu testen und so Einblicke in ihre jeweiligen Stärken und Schwächen zu bekommen. Um den Einfluss von Lerneffekten auf die Auswertung zu vermeiden, wird sowohl die Reihenfolge der Vorgehensweisen als auch die Reihenfolge der Aufgabentypen für jeden Probanden durchmischt.

4.3 Metriken

Zur Beantwortung der gestellten Fragestellungen werden jeweils zwei Metriken zur *User Performance* und zur *User Satisfaction* aufgenommen.

Als erste ergebnisorientierte Metrik kommt die benötigte Zeit für die Aufgaben zum Einsatz. Die Zeit startet erst wenn die Aufgabe verstanden und tatsächlich gestartet wurde und läuft bis der Proband mit dem Vorentwurf zufrieden ist. Weiterhin werden die erzeugten Vorentwürfe durch ein *CAT* bewertet. Zehn fachkundige Jurymitglieder schätzen jeden Entwurf hinsichtlich seiner Qualität, der gezeigten Kreativität und Einsatztauglichkeit ein.

Das Nutzerempfinden wird jeweils direkt nach dem Absolvieren der drei Aufgaben mit einer Vorgehensweise über einen *USE* Fragebogen aufgenommen. Zusätzlich wird, nachdem beide Systeme genutzt wurden, in einem letzten Fragebogen ermittelt, welche Variante der Proband für die einzelnen Aufgabentypen bevorzugt. Hier werden auch weitere Aspekte, wie beispielsweise in welchem Vorgehen die Nutzer das Gefühl hatten potenzielle Probleme ihres Entwurfes besser vorhersehen zu können, ermittelt.

4.4 Teilnehmer

Durch die Durchmischung der Reihenfolge von Versuchsbedingungen und der Aufgaben gibt es insgesamt 12 unterschiedliche Abläufe ($2! * 3! = 12$). Für eine erhöhte Aussagekraft soll jeder Ablauf zwei Mal durchgeführt werden, sodass 24 Probanden für die Studie geplant werden. Für eine hohe Relevanz der Ergebnisse sollten die Teilnehmer Produktmodellierungserfahrungen haben und wenn möglich auch schon VR-Technologie genutzt haben. So wird vermieden, dass durch eine mögliche Faszination

von der Technologie das Ergebnis verfälscht wird. Vorerfahrungen in beiden Bereichen werden durch einen Selbsteinschätzungsfragebogen aufgenommen.

4.5 Ablauf

Für einen reibungslosen Ablauf der Studie wird für jeden Probanden ein Ablaufplan vorbereitet anhand dessen die Studienleitung durch die Tätigkeiten führt. Ein Beispiel für einen solchen ist in Tabelle 2 zu sehen.

Tabelle 2: Studienablaufplan für einen Probanden

#	Tätigkeit
1	Begrüßung und kurze Einführung in die Thematik
2	Belehrung, Einverständniserklärung und ggf. Beantwortung von Fragen
3	Aufnahme personenbezogener Daten wie Vorerfahrungen und Alter
4	Bekanntmachen mit allen Funktionen des VR-Interfaces
5	Start der Video- und Bildschirmaufnahme
6	Versuchsbedingung I: VR-Interface für Vorentwürfe <ul style="list-style-type: none">- Erstellung eines neuen Entwurfes- Nachahmung eines bestehenden Entwurfes- Erweiterung eines bestehenden Entwurfes- USE Fragebogen zum VR-Interface
7	Versuchsbedingung II: Handskizze <ul style="list-style-type: none">- Erstellung eines neuen Entwurfes- Nachahmung eines bestehenden Entwurfes- Erweiterung eines bestehenden Entwurfes- USE Fragebogen zur Handskizze
8	Ende der Video- und Bildschirmaufnahme
9	Fragebogen zur persönlichen Präferenz
10	Verabschiedung

5 Fazit

In diesem Beitrag haben wir auf Basis etablierter Literatur ein Ausarbeitungsleitfaden für Nutzerstudien für immersive Mensch-Maschine-Schnittstellen in der Produktentwicklung vorgestellt. Diesen haben wir zur Planung der Evaluation eines VR-Interface zu dem natürlichen Erzeugen von Vorentwürfen angewandt. Sobald es die Pandemie-situation erlaubt, gilt es den erstellten Plan in einer Vorstudie zu prüfen und auszubessern. Im Anschluss kann die vollständige Studie durchgeführt und ihre Ergebnisse ausgewertet werden. Der Leitfaden soll wie uns, auch anderen Wissenschaftler in der Produktentwicklung eine Hilfestellung zur strukturierten Auswertung von entwickelten XR-Unterstützungen bieten.

Danksagung

Die Autoren danken der Deutschen Forschungsgemeinschaft für die Förderung des Forschungsvorhabens WA 2913/34-1.

Literaturverzeichnis

- Albert, W., & Tullis, T. (2013): *Measuring the user experience: Collecting, analyzing, and presenting usability metrics*. Newnes.
- Amabile, T. M. (1982): *Social psychology of creativity: A consensual assessment technique*. *Journal of Personality and Social Psychology*, 43(5), 997–1013.
- Ashtari, N., Bunt, A., McGrenere, J., Nebeling, M., & Chilana, P. K. (2020): *Creating Augmented and Virtual Reality Applications: Current Practices, Challenges, and Opportunities*. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–13). Association for Computing Machinery.
- Blandford, A. E., Hyde, J. K., Green, T. R. G., & Connell, I. (2008): *Scoping Analytical Usability Evaluation Methods: A Case Study*. *Human-Computer Interaction*, 23(3), 278–327.
- Blessing, L. T. M., & Chakrabarti, A. (2009): *DRM, a Design Research Methodology*. Springer-Verlag.
- Brooke, J. (1996): *Sus: A "quick and dirty" usability*. *Usability Evaluation in Industry*, 189.
- Coburn, J. Q., Freeman, I., & Salmon, J. L. (2017): *A Review of the Capabilities of Current Low-Cost Virtual Reality Technology and Its Potential to Enhance the Design Process*. *Journal of Computing and Information Science in Engineering*, 17(3).
- Faulkner, L. (2003): *Beyond the five-user assumption: Benefits of increased sample sizes in usability testing*. *Behavior Research Methods, Instruments, & Computers*, 35(3), 379–383.
- Feeman, S. M., Wright, L. B., & Salmon, J. L. (2018): *Exploration and evaluation of CAD modeling in virtual reality*. *Computer-Aided Design and Applications*, 15(6), 892–904.
- Gabbard, J. L., Hix, D., & Swan, J. E. (1999): *User-centered design and evaluation of virtual environments*. *IEEE Computer Graphics and Applications*, 19(6), 51–59.
- Harlan, J., Schleich, B., & Wartzack, S. (2020): *Linking a game-engine with CAD-software to create a flexible platform for researching extended reality interfaces for the industrial design process*. *Proceedings of the 31st Symposium Design for X (DFX2020)*, 169–178.
- Hart, S. G., & Staveland, L. E. (1988): *Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research*. In P. A. Hancock & N. Meshkati (Eds.), *Advances in Psychology* (Vol. 52, pp. 139–183). North-Holland.
- Lombard, M., Ditton, T. B., & Weinstein, L. (2009): *Measuring presence: The temple presence inventory*. 1–15.
- Lund, A. M. (2001): *Measuring usability with the use questionnaire*. *Usability and User Experience Newsletter of the STC Usability SIG*, 8(2), 3–6.

Sauro, J., & Lewis, J. R. (2016): *Quantifying the User Experience: Practical Statistics for User Research*. Morgan Kaufmann.

Seybold, C., & Mantwill, F. (2020): *3D Sketching in VR Changing PDM Processes*. In F. Nyffenegger, J. Ríos, L. Rivest, & A. Bouras (Eds.), *Product Lifecycle Management Enabling Smart X* (Vol. 594, pp. 297–310). Springer International Publishing.

Sutcliffe, A., & Gault, B. (2004): *Heuristic evaluation of virtual reality applications*. *Interacting with Computers*, 16(4), 831–849.

Tcha-Tokey, K., Christmann, O., Loup-Escande, E., & Richir, S. (2016): *Proposition and Validation of a Questionnaire to Measure the User Experience in Immersive Virtual Environments*. *International Journal of Virtual Reality*, 16(1), 33–48.

Toma, M. I., Gîrbacia, F., & Antonya, C. (2012): *A comparative evaluation of human interaction for design and assembly of 3D CAD models in desktop and immersive environments*. *International Journal on Interactive Design and Manufacturing (IJDeM)*, 6(3), 179–193.

VDI. (2018): *VDI 2223—Methodisches Entwerfen technischer Produkte*.

Virzi, R. A. (1992): *Refining the test phase of usability evaluation: How many subjects is enough?* *Human Factors*, 34(4), 457–468.

Kontakt

M. Sc. Jakob Harlan

Dr.-Ing. Benjamin Schleich

Prof. Dr.-Ing. Sandro Wartzack

Lehrstuhl für Konstruktionstechnik KTmfk

Martenstraße 9

91058 Erlangen

harlan@mfk.fau.de

<https://www.mfk.tf.fau.de/>