UBIACTION 2021

Edited by Florian Lang, Robin Welsch, Luke Haliburton, Fiona Draxler, Sebastian Feger, Steeven Villa, Matthias Hoppe, Pascal Knierim, Ville Mäkelä, Albrecht Schmidt

UBIACTION 2021

6th Seminar on Ubiquitous Interaction
February 11, 2021, Munich, Germany

Edited by

Florian Lang
Robin Welsch
Albrecht Schmidt

UBIACTION 2021 – Vol. 6

Editors

Florian Lang, Robin Welsch, Luke Haliburton, Fiona Draxler,
Sebastian Feger, Steeven Villa, Matthias Hoppe,
Pascal Knierim, Ville Mäkelä, Albrecht Schmidt

Human-Centered Ubiquitous Media Group
Institut für Informatik
Ludwig-Maxmilians-Universität München
hsuc@um.ifi.lmu.de

ACM Classification 1998 - H.5 INFORMATION INTERFACES AND PRESENTATION

ISBN-13: 979-8511284576

Publication date
11. February 2021

Vorwort

Over the course of the year 2020, the global community has witnessed how a virus caused an unprecedented shift in our habitual reality. Every aspect of life has - at least partially - been moved online: Working, studying, workouts and doctor visits, staying in touch with friends and family. All of these activities have primarily happened in front of screens. As a result, we have spent much more time at home and surrounded by and engaging with technology.

Even before this pandemic, researchers in the field of Ubiquitous Computing have worked towards what Mark Weiser called "the age of calm technology" -a future in which technology fades into the background of our everyday lives while always being available to support and augment humans.

As part of the 6th Advanced Seminar on Human-Centered Interaction in Ubiquitous Computing led by Prof. Dr. Albrecht Schmidt, chair of Human-Centered Ubiquitous Media at the Ludwig-Maximilians-University Munich, 13 students have reviewed recent research projects in this field and compiled their insights in this book.

In the following chapters, readers will learn about a wide variety of topics such as novel interaction paradigms, intelligent feedback in the context of physical exercise and medicine, recent developments in assistive systems for users with visual or hearing impairments, the role technology can play in regards to stress and well-being, the use of intelligent systems for education and learning as well as design considerations for companions in video games.

The results show that we are well on our way towards an even more pervasive integration of technology in numerous parts of our lives. They raise the hope that with the progression of ubiquitous and intelligent systems, we might be better equipped for future crises and future living.

Inhaltsverzeichnis

Editors: F. Lang, R. Welsch, L. Haliburton, F. Draxler, S. Feger, S. Villa, M. Hoppe, P. Knierim, V. Mäkelä, A. Schmidt. February 11, 2021. Munich, Germany.

Participants

Annalena Streichert

Annalena.Streichert@campus.lmu.de
Mensch-Computer-Interaktion

Julia Huber

Huber.Ju@campus.lmu.de
Mensch-Computer-Interaktion

Benedikt Pirker

Benedikt.Pirker@campus.lmu.de
Medieninformatik

Yu Sun

Yu.Sun@campus.lmu.de
Mensch-Computer-Interaktion

Andrea Baumann

Andrea.Baumann@campus.lmu.de
Informatik

Svenja Yvonne Schött

S.Schoett@campus.lmu.de
Medieninformatik

Miriam Leona Metz

Metz.Miriam@campus.lmu.de
Mensch-Computer-Interaktion

Mariam Ali Hussain

Mariam.Ali@campus.lmu.de
Mensch-Computer-Interaktion

Elizabeth Maria Bouquet

E.Bouquet@campus.lmu.de
Mensch-Computer-Interaktion

Carmen Schanderl

Carmen.Schanderl@campus.lmu.de
Mensch-Computer-Interaktion

Laurence Hannes Niklas Lerch

L.Lerch@campus.lmu.de
Informatik

Viktoriia Rakytianska

V.Rakytianska@campus.lmu.de
Medieninformatik mit Mediengestaltung

Tamia Eliza Josephine Bosch

T.Bosch@campus.lmu.de
Informatik

Proxemische intelligente Benutzerschnittstellen

Annalena Streichert

Ludwig-Maximilians-Universität München, München, Deutschland
annalena.streichert@campus.lmu.de

Zusammenfassung

Die Anzahl und Komplexität digitaler Geräte in unserem Alltag steigt immer weiter an. Daher werden die Interaktionen mit der Vielzahl an Geräten immer schwieriger. Um trotzdem nahtlose Interaktionen zu gewährleisten, werden neue Interaktionsmöglichkeiten benötigt. Während wir für die Interaktionen mit anderen Personen zwischenmenschliche Beziehungen nutzen, bleibt diese Möglichkeit bei ubiquitären Computersystem meistens unbeachtet. Proxemische Benutzerschnittstellen nutzen diese räumlichen Beziehungen zwischen Menschen und Geräten, um natürlichere und nahtlose Interaktionen mit ubiquitären Technologien zu ermöglichen. In dieser Arbeit wird ein Überblick über vorhandene Literatur zu proxemischen intelligenten Schnittstellen gegeben, um aktuelle Forschungslücken in der Literatur zu identifizieren.

2012 ACM Computing Classification Human-centered computing → Ubiquitous and mobile computing

Keywords and phrases Proxemik; Proxemische Interaktionen; Ubiquitäre Computersysteme.

1 Einführung

> „When you walk up to your computer, does the screen saver stop and the working windows reveal themselves? Does it even know if you are there? How hard would it be to change this? Is it not ironic that, in this regard, a motion-sensing light switch is „smarter“ than any of the switches in the computer...?“ - Buxton [4]

Buxton adressiert in diesem Statement den Bedarf an neuen Interaktionsmöglichkeiten für Computersysteme. Mit der zunehmenden Anzahl an Geräten im Alltag wird dieses Thema immer wichtiger. Bereits vor 30 Jahren schilderte Weiser [24] die Vision, dass die nächste Ära der Interaktion mit Computern durch die steigende Anzahl von Geräten im täglichen Leben dominiert wird.

Cite as: Annalena Streichert. Proxemische intelligente Benutzerschnittstellen. In *6th Seminar on Ubiquitous Interaction (UBIACTION 2021)*. Editors: F. Lang, R. Welsch, L. Haliburton, F. Draxler, S. Feger, S. Villa, M. Hoppe, P. Knierim, V. Mäkelä, A. Schmidt. February 11, 2021. Munich, Germany. pp. 1:1–1:16.

Heute ziehen neben traditionellen Computern und Smartphones immer mehr neuartige digitale Geräte in unser Leben ein, wie zum Beispiel Smartwatches oder Smart Home Geräte. Die Technologie ist allgegenwärtig. Nach Weisers Vision soll sich die Technologie nahtlos in das Leben der Menschen einfügen, sodass sie sich im Hintergrund befindet und nur bei Bedarf in den Vordergrund rückt. Um diesen Übergang zu ermöglichen, müssen ubiquitäre Computersysteme (Ubicomp) ihre Umgebung wahrnehmen und interpretieren. Heute ist die Mehrheit der Geräte jedoch blind für die Anwesenheit anderer Geräte und für nicht-computergestützte Aspekte der Umgebung, wie zum Beispiel Menschen und räumliche Anordnungen. Für intelligente Interaktionen, vorallem zwischen mehreren Geraten, ist es notwendig: (1) zu erkennen, welche Geräte steuerbar sind; (2) ein bestimmtes Gerät aus der großen Anzahl auszuwählen; (3) Informationen über seinen Status abzurufen; und (4) das Gerät in angemessener Weise zu steuern [15].

Dort kommt die Proxemik zum Einsatz. Im Alltag nutzen Menschen proxemische Distanzen für Interaktionen. Je nach Vertrautheit halten wir Abstand zu Personen, oder orientieren uns zu ihnen, wenn wir sie ansprechen. Wenn wir zum Beispiel eine Person nach dem Weg fragen möchten, orientieren wir uns zu ihr und bewegen uns auf sie zu bis wir, je nach Vertrautheit, einen bestimmmten Abstand erreicht haben. Dabei wird das auf die Person zubewegen bereits als implizite Interaktion wahrgenommen. Dieses Wissen über proxemische Beziehungen kann genutzt werden, um intelligente Schnittstellen für ubiquitäre Computersysteme zu implementieren. Genauso wie Menschen zunehmendes Engagement und Intimität erwarten, wenn sie sich anderen nähern, sollten sie auch zunehmende Konnektivität und Interaktionsmöglichkeiten erwarten können, wenn sie sich oder ihre Geräte in unmittelbare Nähe zueinander und zu anderen Dingen in der Umgebung bringen [6]. Das Ziel dieser Arbeit ist, einen Überblick über vorhandene Literatur zum Thema proxemische intelligente Benutzerschnittstellen zu geben und Forschungslücken zu identifizieren.

2 Ubiquitäre Computersysteme

Mark Weiser führte bereits vor 30 Jahren den Begriff der ubiquitären Computersysteme ein, der das allgegenwärtige Vorhandensein von Computertechnlogien in allen möglichen Arten von Geräten beschreibt [24]. In seiner Vorstellung besitzt jede Person mehrere digitale Geräte in unterschiedlichen Größen, je nach Aufgabe, die durch Netzwerke verbunden sind. Weiser hat außerdem vorhergesagt, dass die Computer in die alltägliche Umgebung der Menschen einziehen werden, eingebettet in alle Arten von Alltagsgegenständen und Räumen. Trotz der Allgegenwärtigkeit beschrieb er, dass die Technologie zu einem unsichtbaren Werkzeug wird, so dass sich Nutzer auf die Aufgabe konzentrieren

können [24].

Damit Nutzer sich auf die Aufgabe konzentrieren können, sollte Technologie an dem Ort zugänglich sein, an dem sie benötigt wird und unsichtbar im Hintergrund bleiben, bis zu dem Moment, in dem sie benötigt wird [25]. Das Konzept wurde von Weiser und Brown als „ruhige“ Technologie eingeführt, um eine Möglichkeit aufzuzeigen, sich auf mehr Informationen einzustellen, sich aber gleichzeitig weniger mit ihnen zu beschäftigen.

Heute haben wir eine Vielzahl an digitalen Geräten in unserem Leben, die auch schon teilweise in Alltagsgegenstände eingebettet sind. Ein wichtiger Aspekt, der noch zu Weisers Vision fehlt, sind nahtlose Interaktionen zwischen Nutzern und ihren Geräten und zwischen den verschiedenen Geräten. Nur wenn Nutzer zwischen Technologien nahtlos wechseln können, können diese in den Hintergrund rücken und Kontinuität mit den bestehenden Arbeitsabläufen gewährleistet werden. Dafür müssen fließende Übergänge zwischen den Funktionsräumen ermöglicht werden [11] . Ishii und Ulmer haben das Konzept als nahtlose Kopplung von Bits und Atomen beschrieben [12].

Um diese Konzepte zu realisieren werden Ansätze benötigt, die nahtlose Interaktionen mit der sich im Hintergrund befindenden ruhigen Technologie und fließenden Übergänge zwischen der in den Vordergrund tretenden Aktivität und der peripheren Wahrnehmung im Hintergrund ermöglichen [17].

2.1 Herausforderungen

Die allgegenwärtige Nutzung von Technologie erfordert leicht zu bedienende Schnittstellen und eine positive Benutzererfahrung. Die Menschen erwarten, dass sie Geräte nutzen können, ohne sich mit den zugrunde liegenden Konzepten oder technischen Details auseinandersetzen zu müssen [19]. Um diese positive Benutzererfahrung zu schaffen, müssen die folgenden von Marquardt et al. [17] ausgearbeiteten Design Herausforderungen mit dem Fokus auf proxemische Interaktionen beachtet werden (auf der Basis von früheren Arbeiten von Bellotti et al. [2]).

Interaktionsmöglichkeiten aufzeigen

Traditionelle User Interfaces schlagen mögliche Aktionen vor. Da Ubicomp Technologien jedoch in den Hintergrund rücken und nicht immer sichtbar sind, werden Alternativen benötigt, Interaktionsmöglichkeiten aufzuzeigen. Daraus ergibt sich die Frage, wie Systeme so gestaltet werden können, dass sie die geeigneten Interaktionsmöglichkeiten offenbaren, auch wenn sie sich im Hintergrund oder im Übergang zwischen Hintergrund und Vordergrund befinden.

Lenken von Aktionen

Die Interaktion mit einem Gerät ist eindeutig, da sie über eine bestimmte Eingabemethode erfolgt. In einer Ubicomp Umgebung ist die Eingabe durch verschiedene Eingemethoden möglich. Daher muss das System erkennen, ob die Interaktion an das System gerichtet ist oder ignoriert werden soll.

Verbindungsaufbau zwischen Geräten

In Ubicomp Umgebungen müssen die verschiedenen Geräte miteinander verknüpft werden, um nahtlose Interaktionen zu ermöglichen. Dafür ist es wichtig, zu kontrollieren, welche Geräte sich zu welchem Zeitpunkt verbinden, sodass die Interaktionsbedürfnisse, Privatsphäre und Sicherheit berücksichtigt werden.

Feedback bereitstellen

Bei Ubicomp Systemen ist es noch wichtiger als bei traditionellen Computern Feedback bereitzustellen, um den Nutzern zu signalisieren, ob ihre Interaktion erfolreich war. Dabei muss berücksichtigt werden, dass die Aufmerksamkeit der Nutzer zwischen Vordergrund und Hintergrund wechselt.

Vermeiden und Korrigieren von Fehlern

Das System sollte Möglichkeiten bereitzustellen, Fehler zu korrigieren. Da Ubicomp Systeme Sensortechnologien für die Eingabe nutzen, kommt es öfter zu falschen Interpretationen der Eingabe.

Verwaltung von Privatsphäre und Sicherheit

Mit der Anzahl der möglichen Interaktionen mit der Technologie steigen auch die Risiken für die Privatsphäre und der Bedarf an größerer Sicherheit. Daher stellt sich die Frage, wie die Risiken minimiert werden können, ohne gleichzeitig all die positiven Aspekte von Ubicomp zu verlieren.

3 Proxemik

Im alltäglichen Leben nutzen Menschen räumliche Beziehungen für Interaktionen. Änderungen der räumlichen Beziehungen sind eine implizite Form der Kommunikation. Je nach Vertrautheit orientieren wir uns an Personen, wenn wir sie ansprechen, wir rücken näher an Objekte heran, für die wir uns interessieren, und wir stehen oder sitzen relativ zu anderen, je nachdem, was gerade ansteht [17]. Proxemik beschreibt die Theorie über die Wahrnehmung und Verwendung von zwischenmenschlichen Distanzen durch Menschen, um

ihre Interaktionen mit anderen Menschen zu vermitteln. In verschiedenen Studien wurden Muster untersucht, wie bestimmte physische Distanzen mit sozialer Distanz korrelieren, wenn Menschen interagieren.

3.1 Persönlicher Raum

Sommer definiert den persönlichen Raum als Distanz, die eine Person zwischen sich und anderen Personen um sie herum einnimmt [20]. Änderungen in dem Bereich des persönlichen Raums beeinflussen, wie Personen interagieren und kommunizieren. Nach einer Studie von Hecht et al. [10] ist der persönliche Raum kreisförmig. Es wurden jedoch geschlechtsspezifische Unterschiede bei der zwischenmenschlichen Distanz festgestellt.

3.2 Halls Proxemik

Edward Hall hat 1966 das Konzept der Proxemik eingeführt, um zusammenhängende Beobachtungen und Theorien über die Nutzung des Raumes durch den Menschen zu beschreiben [9]. Dabei liegt der Fokus auf messbaren Distanzen zwischen Menschen bei Interaktionen. Hall beschreibt, wie Menschen Proxemik wahrnehmen, interpretieren und nutzen, um Beziehungen zu anderen Menschen aufzubauen. Dabei korreliert nach Hall physische Distanz mit sozialer Distanz. Diese Distanzen können in diskrete Zonen eingeteilt werden:

- Intim (0–45cm): z.B. die Distanz von Personen in enger Beziehung oder in einer angeregten Auseinandersetzung. Dort werden die meisten Sinneseindrücke angesprochen. Der Bereich wird normalerweise nur mit Erlaubnis betreten (mit Ausnahmen aufgrund von Umgebungseinschränkungen).
- Persönlich (0.45-1.2m): z.B. bei Interaktionen mit Freunden oder Familie. Die Distanz entspricht ca. der Armlänge, sodass die andere Person berührt werden kann und Verständigung in geringer Lautstärke möglich ist.
- Sozial (1.2-3.65m): z.B. Interaktionen in einer formelleren Umgebung. In der Distanz ist es schwieriger, die andere Person zu berühren und die Kommunikation erfolgt mit höherer Lautstärke.
- Öffentlich (>3,65 m): z.B. Abstand eines Redners zum Publikum. Personen müssen lauter sprechen und der andere primäre sensorische Input ist das Sehen.

Die Theorie der Proxemik kann jedoch nicht nur auf Interkationen angewendet werden, sondern auch auf „die Organisation des Raumes in Häusern und Gebäuden und schließlich die Gestaltung der Städte“ [8]. Objekte eines Raumes können in fixed und semi-fixed Features unterteilt werden. Fixed Features beinhalten alle unbeweglichen Objekte, wie zum Beispiel die Anordnung von Räumen, Wänden und Türen [17]. Semi-fixed Features beinhalten die räumliche

Anordnung von Objekten im Raum, die bewegt werden können, wie zum Beispiel Möbel, Stühle oder Tische.

4 Proxemik für ubiquitäre Computersysteme

Informationen über proxemische Beziehungen zwischen Personen und ihren Geräten können genutzt werden, um neue Interaktionen für ubiquitäre Computersysteme zu entwickeln. Bisher beachten nur wenige Forschungsprojekte Wissen über räumliche Beziehungen. Die Projekte, die Proxemik miteinbeziehen beachten jedoch noch nicht die feineren Nuancen der Distanz, Orientierung, Bewegung, Position und Identität [17]. Marquardt et al. [17] haben Entitäten festgelegt, zwischen denen proxemische Beziehungen möglich sind:

- Personen (einzelne Personen oder kleine Gruppen)
- Große interaktive digitale Oberflächen (z.B. Whiteboards)
- Informationsgeräte (z.B. digitale Bilderrahmen)
- Persönliche tragbare Geräte (z.B. Smartphone)
- Fixed Features (z.B. Wände) und semi-fixed Features (z.B. Möbel)

4.1 Dimensionen der Proxemik

Marquardt et al. haben aus den vorgestellten Theorien fünf proxemische Dimensionen abgeleitet [17].

Distanz

Distanz ist die grundlegende Dimension zur Beschreibung räumlicher Beziehungen. Sie beschreibt die messbare Länge zwischen Entitäten, wie zum Beispiel Personen, Geräte, Objekte, fixed/semi-fixed Features in der Umgebung. Distanz kann als präzise Messung oder grobe Kategorisierung dargestellt werden. Gemessen werden kann die absolute Positionen oder die relative Entfernung. Aktualisierungen können in diskreten Stufen oder kontinuierlich durchgeführt werden (z. B. Abstandsaktualisierung bei sehr großer Nähe, Eintritt in eine größere Entfernung oder Übergang von einer Zone in eine andere).

Orientierung

Orientierung beschreibt die Richtung, in die eine Entität blickt. Diese kann bei Menschen durch die Körperorientierung [21], Körperhaltung oder die Blickrichtung bestimmt werden. Bei Geräten oder Objekten kann die Orientierung eine klar definierte Vorderseite erfordern, zum Beispiel bei Displays. Die Orientierung kann entweder relativ zwischen zwei Entitäten oder absolut zu einem festen Punkt in der Umgebung gemessen werden. Sie kann qualitativen Begriffen (z.B. „zugewandt“ oder „abgewandt“) oder quantitative Werten (z.B. Winkel) dargestellt werden. Dafür können Vektoren im Raum genutzt werden, um zu bestimmen, worauf eine Person zeigt oder schaut.

Bewegung

Bewegung wird durch absolute oder relative Änderungen von Position, Orientierung, Geschwindigkeit und Beschleunigung einer Einheit über die Zeit definiert. Dadurch kann aufgezeigt werden, wie sich eine Person einem Gerät nähert, ob die Person schneller oder langsamer wird und ob die Person die Richtung ändert.

Identität

Die Entitäten im Raum und zugehörige Daten werden durch eindeutige Bezeichnungen, wie zum Beispiel Namen oder Kategorien, einer Identität zugeordnet. Informationen über die Identität sind wichtig, um die Entitäten von einander zu unterscheiden. Die benötigte Granularität hängt dabei von der Systemfunktionalität ab. Zum Beispiel sind grob gefasste Kategorien (z.B. Person, keine Person) für ein öffentliches Display ausreichend, das die gleichen Informationen anzeigt, sobald sich eine Person nähert. Für einen privaten Computer, der personalisierte Inhalte anzeigen soll, muss jedoch zwischen feiner gefassten Kategorien (z.B. Namen der verschiedenen Nutzer) unterschieden werden.

Standort

Für die Beschreibung des Standorts, an dem die Interaktion stattfindet, können qualitative und quantitative Aspekte miteinbezogen werden. Zu den qualitativen Aspekten gehört der Kontext, der die Umgebung als Ganzes beschreibt (z.B. Haus oder Büro) und Metainformationen über die Nutzung des Raumes enthält (z.B. Nutzung als Besprechungsraum für ein bestimmtes Projekt). Zu den quantitativen Aspekten gehört das Layout, das die Konfiguration des Raumes beschreibt, wie zum Beispiel die Position der fixed und semi-fixed Features im Raum.

4.2 Design von proxemischen Interaktionen

Die fünf vorgestellten Dimensionen können genutzt werden, um neue Interaktionsmöglichkeiten in Systeme zu integrieren. Marquardt et al. haben dafür das „Proxemic interactions framework“ vorgestellt [17].

- Eingabe: Sensortechnologie stellt Daten über aufgezeichnete Entitäten bereit.
- Standort: Kontext der Umgebung und Details über semi-fixed und fixed Features

- Identität: Informationen über Entitäten ermöglichen dem System entweder die Unterscheidung zwischen Personen und Geräten oder - bei feinerer Granularität - eindeutige Namen von Personen und Geräten
- Relativer Abstand zwischen diesen Entitäten.
- Orientierung zwischen Entitäten
- Bewegung und Orientierung durch die Bestimmung von Änderungen der beiden vorherigen über die Zeit
- Verhaltensregeln: legen fest, wie das System proximale Beziehungen zwischen Entitäten interpretiert und wie sie in Systemaktionen umgesetzt werden
- Ausgabe: abhängig von den Aktionen löst das System entsprechende Ausgabeaktionen aus

4.3 Bewältigung der UbiComp Herausforderungen mithilfe von Proxemik

Im folgenden werden Möglichkeiten aufgezeigt, wie die Herausforderungen aus 2.1 mithilfe von Proxemik bewältigt werden können (nach Marquardt et al. [17]). Danach können Techniken, die mit Blick auf die sozialen Erwartungen an die Proxemik entwickelt werden, als universeller Weg zur Bewältigung vieler Herausforderungen in Ubicomp eingesetzt werden.

4.3.1 Herausforderung 1: Interaktionsmöglichkeiten aufzeigen

In der realen Umgebung wissen wir was verfügbar ist, indem wir uns umschauen. In der Ubicomp Umgebung müssen explizit sonst verborgene Angebote auf dem Bildschirm eines Geräts visualisiert werden. Ähnlich wie Menschen einander beim Start der Kommunikation begrüßen [13], sollten Ubicomp Systeme andere Entitäten „begrüßen “, indem sie mögliche Interaktionen aufzeigen [13, 17]. Um das zu ermöglichen ist es wichtig, dass das System erkennt, welche Personen anwesend sind und was ihre Absichten sind. Ein Beispiel dafür ist Buxtons smartes Lichtsystem, das mithilfe von Bewegungserkennung Lichter an- und ausschaltet [3]. Im realen Leben nutzen Menschen Proxemik bei der Begrüßung und sozialen Interaktion: man nimmt andere Personen aus der Ferne wahr, wird zunehmend aufmerksam und engagiert, wenn sich die Person zuwendet und nähert, und beginnt dann zu interagieren, wenn sie sich in einem geeigneten proxemischen Bereich befindet. Dieses Verhalten kann auch für digitale Systeme genutzt werden, um den Übergang von Aufmerksamkeit zur Interaktion zu implementieren. Beispiele hierfür sind öffentliche Displays, wie zum Beispiel von Vogel und Balakrishnan [23]. Ihr Prototyp unterscheidet zwischen vier Interaktionsphasen mit einem fließenden Übergang

von der impliziten öffentlichen Interaktion, bis zur nahen expliziten Interaktion. Dadurch kann das Display zwischen öffentlichen und privaten, auf die Person abgestimmten, Informationen wechseln. Der Phasenwechsel wird erst durch implizite Interaktion wie Körperbewegung, Körperposition und Kopfausrichtung signalisiert. Mit zunehmender Nähe werden explizite Gesten und Berührungen genutzt.

4.3.2 Herausforderung 2: Lenken von Aktionen

Um Interaktionen an ein bestimmtes Gerät zu lenken, gibt es verschiedene Möglichkeiten. Eine Möglichkeit sind diskrete Zonen für Interaktionen, in denen sich die Person befinden muss, damit die Eingabe akzeptiert wird. Eine weitere Möglichkeit sind attentive User Interfaces (AUIs). AUIs nutzen Augenfixierungen um zu bestimmen, ob die Aumerksamkeit (der Blick) einer Person auf das System gerichtet ist. Diese Information wird dann genutzt, um die Reaktion eines Systems anzupassen [22]. Zum Beispiel kann dadurch nur auf dem Gerät eine Benachrichtigung ausgegeben werden, dem Nutzer in dem Moment ihre Aufmerksamkeit schenken. Es können außerdem fixed und semi-fixed Features miteinbezogen werden, um zu erkennen ob die Interaktionen an das System gerichtet sind. Ein Beispiel hierfür ist der Media Player von Ballendat et al., der den Film startet, wenn eine Person sich setzt [1]. Ein weiterer Ansatz ist, Bewegungsverläufe zu nutzen. Im alltäglichen Leben kann unterschieden werden, ob wir zum Beispiel direkt auf eine Person zugehen oder stattdessen schnell an ihr vorbei. In ähnlicher Weise können Ubicomp-Systeme die Bewegungen von Menschen und Geräten interpretieren, um Aktionen zu steuern. Wenn sich mehrere Geräte in der Nähe einer Person befinden, können diese anhand von Distanz oder Identität gefiltert werden, um die Anzahl möglicher Ziele einzugrenzen.

4.3.3 Herausforderung 3: Verbindungsaufbau zwischen Geräten

Um nahtlose Interaktion zu ermöglichen, muss kontrolliert werden, wie sich ein Gerät mit anderen Geräten verbindet. Gleichzeitig muss jedoch Privatsphäre und Sicherheit gewährleistet werden. Als Bedingung für eine Verbindung kann die Proxemik genutzt werden. Mithilfe von Distanz, Identität und Orientierung ist es möglich zu bestimmen, welche Geräte sich in naher Umgebung befinden. Nur zu diesen wird dann eine Verbindung aufgebaut. Die Implementierung kann als binäre Funktion (nah = verbunden), oder als progressiver Verbindungsprozess erfolgen, wie von Kray et al. vorgeschlagen [14]. Sie nutzen räumliche Regionen um Mobiltelefone herum, um Geräteverbindungen auf- und abzubauen oder eine Datenübertragung zu initiieren. Wenn sich ein Gerät über drei diskrete Regionen bewegt, wird zunächst eine Vorschau auf eine Medienübertragung angezeigt, wobei die Übertragung erst nach dem Eintritt in eine nähere Region beginnt.

4.3.4 Herausforderung 4: Feedback bereitstellen

Im Gegensatz zu traditionellen Computern haben viele Ubicomp Systeme kein graphisches Display. Daher werden neue alternative Ausgabemethoden benötigt, wie zum Beispiel visuelle Lichter, Töne, Sprache oder physisch bewegte Objekte, wie in tangible UIs. Informationen über physische Orientierung und Entfernung können genutzt werden, um eine angepasste Ausgabe zu ermöglichen. So können die Details der Ausgabe, oder zum Beispiel die Lautstärke an diese Informationen angepasst werden. Je nach proxemischer Beziehung kann außerdem die Auswahl der geeigneten Ausgabemethode variieren. Zum Beispiel könnte das System ein akustisches Signal einer visuellen Ausgabe vorziehen, wenn der Nutzer von dem Bildschirm abgewendet ist.

4.3.5 Herausforderung 5: Vermeiden und Korrigieren von Fehlern

Eine Möglichkeit, eine falsch interpretierte oder unbeabsichtigte Aktion zu korrigieren ist, die umgekehrte Aktion auszuführen. Das System kehrt dann zum vorherigen Zustand zurück. Anstatt die umgekehrte Aktion auszuführen, ist es auch möglich, eine explizite Aktion zu definieren, die eine frühere Aktion rückgängig macht. Besonders für Aktionen mit hoher Auswirkung, wie zum Beispiel das Löschen von Informationen, kann Proxemik genutzt werden, um die Aktion abzusichern. Ein Ansatz dafür ist, dass eine Aktion nur dann ausgeführt werden kann, wenn sich eine Person in unmittelbarer Nähe eines Geräts befindet. Alternativ könnten Aktionen mit hoher Auswirkung eine proxemische Beziehung in mehreren Dimensionen erfordern, zum Beispiel nahe am Gerät sein und auf den Bildschirm schauen. Außerdem könnten Aktionen provisorisch sein und rückgängig gemacht werden, wenn sich der Nutzer entfernt, bevor er die Aktion bestätigt hat.

4.3.6 Herausforderung 6: Verwaltung von Privatsphäre und Sicherheit

Der Zugriff auf Ubicomp-Systeme kann in Abhängigkeit von der wahrgenommenen Nähe von Personen, Geräten oder anderen Objekten reguliert werden. Die Entfernung zwischen Entitäten kann genutzt werden, um die Menge an Informationen zu bestimmen, die zwischen ihnen ausgetauscht werden. Nach Fiskin et al. gilt dabei das Prinzip „Distanz impliziert Misstrauen“ [5].

Zusätzlich zu der Distanz können Systeme auch Orientierung, Identität oder Standort verwenden, um die Menge der geteilten Informationen beeinflussen. Datenschutzrelevante Informationen könnten zum Beispiel nur sichtbar sein, solange die Person auf den Bildschirm schaut, aber verborgen, sobald sie wegschaut oder eine andere Person mit auf den Bildschirm schaut. Ubicomp-Systeme könnten außerdem die Identität des Nutzers erkennen, und je nach Identität zu restriktiveren Datenschutz- und Sicherheitseinstellungen wechseln, wenn andere Personen oder Geräte in der Umgebung erkannt werden. Ein weiterer Faktor der berücksichtigt werden kann ist der Standort. Je nach Art der Umgebung können die Einstellung angepasst werden. In einem Großraumbüro zum Beispiel wird eine höhere Sicherheitsstufe benötigt als zu Hause.

Ein anderer Ansatz ist, Systeme zu entwerfen, die die Erwartungen der Menschen an den persönlichen Raum respektieren. Das Ubicomp-System kann die gleichzeitige Interaktion mehrerer Personen so beeinflussen, dass ein solches Maß an Privatsphäre für alle Beteiligten erhalten bleibt

5 Ansätze für proxemische Benutzerschnittstellen

5.1 Toolkits und Frameworks

Proximity Toolkit. Das Toolkit von Marquardt et al. [16] erleichtert Programmierern den Zugang zu proxemischen Informationen zwischen Personen, Objekten und Geräten in einer Ubicomp-Umgebung auf kleinem Raum. Es besteht aus vier Komponenten:

- Proximity Toolkit Server: ermöglicht es mehreren Clients auf die Proximitätsinformationen zuzugreifen
- Tracking-Plug-in-Module: verbinden verschiedene Trackingsysteme mit dem Toolkit und streamen die Daten an den Server
- Visuelles Monitoring Tool: visualisiert aufgezeichnete Entitäten mit deren proxemischen Beziehungen
- Application programming interface (API): ereignisgesteuerte Programmierbibliothek, ermöglicht Zugriff auf die proxemischen Interaktionen

Proxemic Mobile App Framework. Perez et al. [18] haben ein Framework implementiert, das aktuelle Trends in der Mobiltechnologie in Bezug auf die Möglichkeiten der Erfassung und Verarbeitung verschiedener Arten von Daten nutzt, um proxemischen Anwendungen zu erstellen. Es soll Entwickler bei der Implementierung und bei der Verwaltung von proxemischen Informationen unterstützen, um die passende Kombination von Dimensionen festzulegen. Das Framework besteht aus drei Komponenten:

- Proxemische Zonen Modul: ermöglicht die Definition der Größe der proxemischen Zonen.

- DILMO Modul: definieren der passenden DILMO (= Distanz, Identität, Standort, Bewegung, Orientierung) Kombination.
- API: stellt die Verbingung zwischen den proxemischen Zonen und dem DILMO Modul dar. Es ermöglicht die Verarbeitung von proxemischen Informationen und Werten und stellt Klassen und Methoden zur Verfügung, um die proxemischen Zonen zu definieren.

5.2 Person-zu-Gerät Interaktionen

Das im Folgenden beschriebene Beispiel veranschaulicht die Gestaltungsmöglichkeiten für proxemische Interaktionen zwischen Personen und den Geräten in ihrer Umgebung.

Proxemischer Media Player. Ballendat et al. [1] haben einen Media Player entwickelt, der proxemische Beziehungen für Interaktionen nutzt. Die Anwendung unterstützt das Durchsuchen, Auswählen und Abspielen von Videos auf einem großen, an der Wand montierten Bildschirm. Mithilfe des Proximity Toolkits von Marquardt et al. [16] wird die Position und Orientierung von Personen, Objekten und anderen digitalen Geräten in der Nähe aufgezeichnet.

Das System wird aktiviert, wenn eine Person den Raum betritt und zeigt kontinuierlich mehr Inhalte an, wenn sich der Abstand der Person zum Display verringert. Wenn sich die Person in geringer Nähe zum Display befindet, ermöglicht es explizite Interaktion durch direkte Berührung. Wenn die Person dann auf dem vor dem Display stehenden Sofa Platz nimmt, wird implizit auf Vollbildansicht umgeschaltet. Für den Media Player wird keine Fernbedienung benötigt. Explizite Interaktionen sind durch physische Artefakte (z.B. Handy, Stift), ausgelöst durch Abstand und Orientierung, möglich.

Ballendat et al. [1] haben auch das Prinzip der Attentiven User Interfaces (AUIs) [22] eingebaut. Die Videowiedergabe wird pausiert, wenn die Person einen Anruf entgegennimmt, liest, oder sich mit einer anderen Person unterhält.

Die Anwendung erkennt, wenn eine weitere Person den Raum betritt und zeigt daraufhin grundlegende Informationen wie den Videotitel an. Wenn die Person sich nähert, wird eine geteilte Ansicht angezeigt, die eine detailliertere Videobeschreibung beinhaltet. Sobald die Person in Reichweite des Displays ist, erhält sie die volle Kontrolle. Wenn alle Personen den Raum verlassen wird das Display ausgeschaltet.

5.3 Gerät-zu-Gerät Interaktionen

Die zunehmende Anzahl digitaler Geräte in unserer Umgebung, von persönlichen Smartphones bis hin zu großen stationären Bildschirmen, erhöht die Möglichkeiten der Interaktion mit digitalen Inhalten. Trotzdem ist der Transfer von Informationen zwischen den Geräten noch schwierig. Es muss

beachtet werden, welche Geräte miteinander kommunizieren sollen und welche Informationen dabei ausgetauscht werden [17].

Cross-Device Interaction. Cross-Device Interaktionen ermöglichen die Freigabe von Inhalten und die Kontrolle bei der Zusammenarbeit an einem Ort. Gronbaek et al. [7] schlagen einen auf Proxemik basierenden Ansatz zur Gestaltung flexibler geräteübergreifender Interaktionen vor. Ihr Fokus liegt auf der Ermöglichung zwischenmenschlicher Beziehungen durch ein flexibles Zusammenspiel zwischen Menschen, interaktiven Geräten und Merkmalen in der Umgebung (z.B. Möbeln). Sie stellen drei Prototypen für Cross-Device Interaktion vor: *Slam-to-Share*, *Stick-to-Share* und *Show-to-Share*. Slam-to-Share nutzt physische Gesten in Form von zwei aufeinanderfolgenden Schlägen auf eine Tischoberfläche, um eine Ad-hoc-Zuordnung zwischen den Geräten mehrerer Benutzer zu ermöglichen. Mithilfe von Stick-to-Share können Assoziation zwischen Inhalten und Artefakten in der Umgebung realisiert werden. Aufkleber ermöglichen den digitalen Zugriff. Durch das Anheften von Papierartefakten in einen gemeinsam genutzten Raum ist der digitale Zugriff dauerhaft für alle verfügbar. Show-to-Share ist eine Präsentationsanwendung, bei der der Zugriff auf die Steuerung der Präsentation durch Manipulation der Größe und Ausrichtung von QR-Tags verändert werden kann. Dadurch können nur Personen, die sich in einer bestimmten proxemischen Zone befinden, die Präsentation steuern.

6 Fazit und Ausblick

In dieser Arbeit wurden vorangegangene Arbeiten zu proxemischen intelligenten Benutzerschnittstellen vorgestellt. Die Nutzung von proxemischen Distanzen für intelligente Benutzerschnittstellen hat großes Potential, die Herausforderungen, die ubiquitäre Computersysteme mit sich bringen, zu bewältigen. Die fünf vorgestellten Dimensionen der Proxemik liefern eine Vielzahl an neuen Interaktionsmöglichkeiten, indem sie einzeln oder kombiniert implementiert werden. Trotzdem sind sie momentan in fast keinen Ubicomp-Systemen zu finden. Der vorgestellte Media Player ist ein gutes Beispiel für die Möglichkeiten, die die Proxemik bietet [1]. Er kann als Startpunkt für weitere Systeme gesehen werden, da er eine Reihe von neuartigen Möglichkeiten vorstellt, Proxemik im Interaktionsdesign zu berücksichtigen.

Ein großes Thema bei proxemischen Schnittstellen sind Verhaltensregeln. Je nach Nutzer kann das Verhalten gegenüber Geräten variieren, weshalb sich die Systeme an ihre Nutzer anpassen und eine Balance zwischen der Beurteilung von proxemischen Informationen und impliziter oder expliziter Interaktion erreichen sollten. Zukünftige Arbeit könnte sich damit befassen, wie diese Personalisierung erfolgen kann.

Bisherige Arbeiten beschäftigen sich größtenteils mit kleinen Ubicomp-Umgebungen, wie zum Beispiel einzelnen Räumen. Diese Forschung sollte auf größere Bereiche und mehr Geräte ausgeweitet werden, um alle Vorteile der Proxemik auszunutzen.

Literatur

1 T. Ballendat, N. Marquardt, and S. Greenberg. Proxemic interaction: designing for a proximity and orientation-aware environment. In *ACM International Conference on Interactive Tabletops and Surfaces*, pages 121–130, 2010.

2 V. Bellotti, M. Back, W. K. Edwards, R. E. Grinter, A. Henderson, and C. Lopes. Making sense of sensing systems: Five questions for designers and researchers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '02, page 415–422, New York, NY, USA, 2002. Association for Computing Machinery.

3 B. Buxton. Integrating the periphery & context: A new model of telematics. In *Graphics Interface*, pages 239–239, 1995.

4 W. Buxton. Living in augmented reality: Ubiquitous media and reactive environments. *Video Mediated Communication*, pages 363–384, 1997.

5 K. P. Fishkin, S. Roy, and B. Jiang. Some methods for privacy in rfid communication. In *European Workshop on Security in Ad-hoc and Sensor Networks*, pages 42–53. Springer, 2004.

6 S. Greenberg, N. Marquardt, T. Ballendat, R. Diaz-Marino, and M. Wang. Proxemic interactions: The new ubicomp? *Interactions*, 18(1):42–50, 2011.

7 J. E. Grønbæk, M. S. Knudsen, K. O'Hara, P. G. Krogh, J. Vermeulen, and M. G. Petersen. Proxemics beyond proximity: Designing for flexible social interaction through cross-device interaction. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, page 1–14, New York, NY, USA, 2020. Association for Computing Machinery.

8 E. T. Hall. A system for the notation of proxemic behavior. *American anthropologist*, 65(5):1003–1026, 1963.

9 E. T. Hall. *The hidden dimension*. Garden City, NY: Doubleday, 1966.

10 H. Hecht, R. Welsch, J. Viehoff, and M. R. Longo. The shape of personal space. *Acta Psychologica*, 193:113 – 122, 2019.

11 H. Ishii, M. Kobayashi, and K. Arita. Iterative design of seamless collaboration media. *Commun. ACM*, 37(8):83–97, 1994.

12 H. Ishii and B. Ullmer. Tangible bits: Towards seamless interfaces between people, bits and atoms. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, CHI '97, page 234–241, New York, NY, USA, 1997. Association for Computing Machinery.

13 A. Kendon. *Conducting interaction: Patterns of behavior in focused encounters*, volume 7. CUP Archive, 1990.

14 C. Kray, M. Rohs, J. Hook, and S. Kratz. Group coordination and negotiation through spatial proximity regions around mobile devices on augmented tabletops. In *2008 3rd IEEE International Workshop on Horizontal Interactive Human Computer Systems*, pages 1–8, 2008.

15 D. Ledo, S. Greenberg, N. Marquardt, and S. Boring. Proxemic-aware controls: Designing remote controls for ubiquitous computing ecologies. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '15, pages 187–198, New York, NY, USA, 2015. Association for Computing Machinery.

16 N. Marquardt, R. Diaz-Marino, S. Boring, and S. Greenberg. The proximity toolkit: Prototyping proxemic interactions in ubiquitous computing ecologies. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, UIST '11, pages 315–326, New York, NY, USA, 2011. Association for Computing Machinery.

17 N. Marquardt and S. Greenberg. *Proxemic Interactions: From Theory to Practice*. Morgan & Claypool Publishers, 2015.

18 P. Pérez, P. Roose, Y. Cardinale, M. Dalmau, D. Masson, and N. Couture. Mobile proxemic application development for smart environments. In *18th International Conference on Advances in Mobile Computing and Multimedia (MoMM '20)*, 2020.

19 A. Schmidt, B. Pfleging, F. Alt, A. Sahami, and G. Fitzpatrick. Interacting with 21st-century computers. *IEEE Pervasive Computing*, 11(1):22–31, 2012.

20 R. Sommer. Studies in personal space. *Sociometry*, 22(3):247–260, 1959.

21 R. Sommer. Personal space. the behavioral basis of design. 1969.

22 R. Vertegaal and J. S. Shell. Attentive user interfaces: the surveillance and sousveillance of gaze-aware objects. *Social Science Information*, 47(3):275–298, 2008.

23 D. Vogel and R. Balakrishnan. Interactive public ambient displays: Transitioning from implicit to explicit, public to personal, interaction with multiple users. In *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology*, UIST '04, page 137–146, New York, NY, USA, 2004. Association for Computing Machinery.

24 M. Weiser. The computer for the 21 st century. *Scientific American*, 265(3):94–105, 1991.

25 M. Weiser and J. S. Brown. Designing calm technology. *PowerGrid Journal*, 1(1):75–85, 1996.

Simulation von haptischem Feedback in Virtueller Realität als Unterstützung medizinischer Verfahren

Julia Huber

Ludwig-Maximilians-Universität München, München, Deutschland
huber.ju@campus.lmu.de

Zusammenfassung

Um mit dem Einsatz von technischen Geräten eine detailgetreue Wiedergabe der Realität darzustellen, forschen Wissenschaftler nach den besten Wegen, nicht nur visuelle und auditive, sondern auch haptische Informationen in der virtuellen Realität mitzuteilen. Durch diverse Arten von Rückmeldung soll schließlich eine Illusion von Haptik entstehen. Auch im Feld der virtuellen Chirurgie ist die Erlangung von haptischen Auskünften von großer Bedeutung, da so neue diagnostische und interventionelle Verfahren gefunden werden können. Diese Arbeit trägt verschiedene Methoden der haptischen Simulation vor und zeigt an Beispielen auf, wie das in der Medizin Anwendung findet. Die Illusion einer haptischen Rückmeldung ist ein großes Forschungsfeld und es gibt vielfältige Herangehensweisen an dieses Thema, ob durch die visuelle, auditive oder auch haptische Umsetzung. Wegen bestehender Limitationen wird weiter nach Optimierung gesucht, dennoch werden die Methoden bereits angewandt, die auch in Zukunft die virtuelle Realität revolutionieren und auch in der Medizin eine immer wichtigere Rolle spielen werden.

2012 ACM Computing Classification Human-centered computing → Human computer interaction (HCI) → Interaction devices → Haptic devices

Keywords and phrases Virtuelle Realität; Haptisches Feedback; Medizin.

1 Einleitung

1.1 Motivation des Forschungsgebiets in der Medizin

Neue Technologien in der virtuellen Realität und der Robotik sind im Aufstreben und haben Potential, gängige Methoden zu revolutionieren [15]. Die Bedeutung von Simulatoren haptischen Feedbacks nimmt dabei zu [4]. Die Rückmeldung fördert das Gefühl für Schwerkraft, Trägheit und Berührungen und soll eine intuitive Bestätigung für die Interaktion mit Objekten geben. Dabei erlangt man Kenntnisse über Gewicht und Bewegung und ob die Objekte

Cite as: Julia Huber. Simulation von haptischem Feedback in Virtueller Realität als Unterstützung medizinischer Verfahren. In *6th Seminar on Ubiquitous Interaction (UBIACTION 2021)*. Editors: F. Lang, R. Welsch, L. Haliburton, F. Draxler, S. Feger, S. Villa, M. Hoppe, P. Knierim, V. Mäkelä, A. Schmidt. February 11, 2021. Munich, Germany. pp. 2:1–2:14.

erfolgreich berührt oder gegriffen wurden [2]. Das Vertrauen des Benutzers in seine virtuellen Taten soll dabei gefestigt werden [3]. Anwendung finden solche Simulatoren längst auch im Feld der Medizin und der Gesundheitsversorgung. Bereits jetzt sind klinisch validierte, leistungsstarke Simulatoren weltweit im Einsatz [15].

1.2 Aufbau der Arbeit

Das Ziel dieser Arbeit liegt darin, die Funktionsweise der Simulatoren von haptischem Feedback in der virtuellen Realität aufzuzeigen und die Rolle dieser im medizinischen Umfeld zu eruieren. Dazu sollen erst diverse Möglichkeiten der Simulation dargelegt werden und anschließend auf den Stand der Technik in der Medizin übertragen werden. Dabei soll auch auf die Grenzen des Forschungsgebietes eingegangen werden, was mit einem derzeit zukünftigem Ausblick beantwortet werden soll. Am Ende wird die Bedeutung der erforschten Möglichkeiten für die Medizin kurz diskutiert.

2 Überblick über die Umsetzung der Illusion von haptischem Feedback

Die visuellen und akustischen Sinnesmodi profitieren von einer großen Verfügbarkeit hochwertiger Anzeigegeräte. In Kombination mit einer genauen und latenzarmen Umsetzung ermöglichen diese die Entwicklung überzeugender Simulationen, die ausreichen, um bei den Benutzern ein starkes Gefühl der Präsenz zu erzeugen. In vielen Anwendungsbereichen ist es wahrscheinlich, dass auch die Berührung ein überzeugender Faktor für die Präsenz sein kann [16]. Im Folgenden wird auf die unterschiedlichen Weisen der Sinnesansprachen eingegangen, um die Funktionsweise der haptischen Illusion zu erläutern.

2.1 Visuelle Rückmeldung

Visuelles Feedback in einer virtuellen Umgebung beeinflusst die haptische Wahrnehmung. L écuyer et al. [10] haben das Phänomen der „visuellen Dominanz“ in Bezug auf die Einschätzung der Steifigkeit von virtuellen Objekten untersucht. Demnach kann eine verzerrte visuelle Darstellung von Objekten die wahrgenommene Steifigkeit bei der Manipulation eines haptischen Feedbackgeräts stark verfälschen. In einer Studie von L écuyer et al. wurde gezeigt, dass es ein perzeptuelles Offset gibt zwischen der Wahrnehmung einer realen Feder und der Wahrnehmung einer pseudo-haptischen Feder, die mit visuellem Feedback simuliert wurde. Die pseudo-haptische Feder wurde im Vergleich zur realen Feder unterschätzt. Mit Hilfe einer psychophysikalischen Methode fanden die Forscher heraus, dass der Wahrnehmungs-Offset (oder der Punkt

der subjektiven Gleichheit) im Durchschnitt 9% betrug. Desweiteren wurde ein Tiefen- und Perspektiveneffekt beobachtet: Objekte, die sich in größerer Entfernung befinden, werden als steifer wahrgenommen [5].

Die unmittelbare visuelle Rückmeldung ist wesentlich für das Ergebnis der virtuellen Interaktion. Eine Studie hat gezeigt, dass eine Verzögerung des visuellen Feedbacks größere negative Auswirkungen haben kann auf die Erfüllung von Aufgaben. Erstens bestand eine lineare Beziehung zwischen der Latenz und der Zeit, die die Teilnehmer brauchten, um sich im virtuellen Umfeld zum Ziel zu bewegen. Außerdem hingen die Auswirkungen der Verzögerung von der Schwierigkeit der Aufgabe ab: Je schwieriger die Aufgabe, desto größer der durch Latenz verursachte Nachteil. Eine Latenz im visuellen Feedback erhöht sowohl den Zeitaufwand für die Ausführung einer Aufgabe als auch die Anzahl der gemachten Fehler. Die Beziehung zwischen dem Grad der Latenz und der Verschlechterung der Leistung wurde gemessen und als linear beschrieben: Die Geschwindigkeit nahm proportional mit der Verzögerung ab [8].

Dabei kommt es aber nicht nur auf die vordergründige Visualisierung von Tatbeständen an, sondern auch auf das, was für den Nutzer unterbewusst im Hintergrund der gezeigten Szene abläuft. Bei einem Verkehrssimulationsmodell von Herpers et al. [6] bedeutet dies beispielsweise, dass die Verkehrsteilnehmer als Software-Agenten implementiert wurden, deren Position und Bewegung ein realistisches Verhalten zeigen sollen. Auch wenn die Aktoren gerade nicht aktiv zu sehen sind, soll ihr Standort jederzeit neu berechnet werden können, sobald diese benötigt werden, um dem Zuschauer ein Gefühl von realen Umständen zu geben [6].

2.2 Auditive Unterstützung

Auch auditives Feedback kann dabei helfen, eine verstärkte Illusion von Kräften zu suggerieren.

In einer Studie wurde untersucht, wie anhand der Sonifikationsmethode die Kontrolle und das Bewusstsein der Bediener beim virtuellen Greifen verbessert werden können. Diese Methode ist dazu da, die Kraftverstärkung auf das Klangtempo abzubilden. Um die Auswirkungen zu beobachten, wurden der Index des Greifrisikos, die Zeit der Aufgabenerfüllung und die subjektive Arbeitsbelastung unter verschiedenen Fällen von Feedback-Schnittstellen und verschiedenen Arten von Materialien gemessen und analysiert. Das Klangtempo wurde durch die Veränderung der Dauer der ausgewählten Audiosignale erzeugt. Die Signale wurden dann den Bedienern in Bezug auf die Änderungen der Antriebskraft präsentiert. Das Ergebnis zeigt, dass Audiosignale Auswirkung auf eine schnelle Alarmierung haben und die Aufmerksamkeit und Kontrolle der Bediener steigern. Im Experiment werden die Audiosignale durch Klangtempo, Frequenz und Lautstärke dargestellt. Das Tempo ist die Geschwindigkeit der Audiosignale. Es wird durch zwei Schlagwellen mit einer bestimmten Dauer dargestellt. Dabei nehmen die Bediener im Experiment wahr, dass die treibende Kraft während des Greifvorgangs umso höher ist, je schneller das Tempo ist [18].

2.3 Haptisches Feedback

Die Berechnung von zwei Wahrnehmungsvariablen dient der Untersuchung der haptischen Wahrnehmung:

- Just Noticeable Difference (JND): der Punkt, an dem es keinen Wahrnehmungsunterschied zwischen zwei Reizen gibt, d. h. jeder Reiz mit einer relativen Differenz zu einem Referenzreiz, die kleiner als die JND ist, wird als gleich dem Referenzreiz wahrgenommen.
- Point of Subjective Equality (PSE): wird verwendet, um zwei Reize unterschiedlicher Beschaffenheit zu vergleichen, und entspricht dem Punkt, an dem sie als genau gleich wahrgenommen werden [5].

Vielzählige haptische Möglichkeiten wurden schon entwickelt, um physikalische Kräfte zu simulieren und tatsächliche haptische Rückmeldung zu geben. Die am weitesten verbreiteten hapischen Lösungen sind Wearables, Handhelds, Mid-Airs und Encounter-Types. Während es bei Handhelds geerdete und nicht geerdete Modelle gibt, sind Encounter-Types konventionellerweise geerdet. Handhelds und Mid-Airs erfordern, dass der Benutzer ein Gerät trägt. Der Interaktionsraum ist oft auf die Hände des Benutzers beschränkt, da virtuelle Objekte nicht mit anderen Körperteilen gefühlt werden können. Haptische Geräte vom Encountered-Typ sind meist gruppierte Roboterarme, die sich so bewegen, dass der Benutzer auf den End-Effektor des Roboterarms trifft, wenn er mit einem virtuellen Objekt in Kontakt kommt. Diese Geräte erfordern nicht, dass der Benutzer ein Gerät tragen oder ein Werkzeug halten muss, was den Kontakt mit allen Körperteilen ermöglicht. Außerdem bietet die Haptik im Gegensatz zu anderen Formen haptischer Lösungen, eine physische Oberfläche, die es dem Benutzer ermöglicht, Objekte direkt zu berühren und zu manipulieren. Trotz dieser Vorteile hat die geerdete Encounter-Haptik Mängel, die ihre Verwendung einschränken, wie zum Beispiel die hohen Kosten und der begrenzte Arbeitsbereich. Das Arbeitsvolumen ist durch den Raum, den der Roboterarm überspannt, begrenzt, und selbst ein mobiler Roboter auf Rädern setzt der Höhe des Arbeitsraums eine Grenze. Darüber hinaus erfordern Roboterarme eine komplexe Bewegungsplanung, um um Hindernisse in der Umgebung herum zu navigieren. Um diese Einschränkungen zu beheben, haben Forscher begonnen, schwebende haptische Geräte zu erforschen, die Quadcopter in der virtuellen Realität verwenden [1]. Bei der Koppelung von Interaktionen mit haptischen Feedback, wird eine Verbesserung der Tiefenwahrnehmung von den Benutzern wahrgenommen [16].

Am Beispiel eines Nadeleinstichs kann erläutert werden, wie viele Faktoren für eine lebensechte Simulation der Haptik berücksichtigt und umgesetzt werden müssen: Steifigkeit, Schneid- und Reibungskräfte an der Nadelspitze und am Schaft, Nadelauslenkung, also die Biegung, Gewebe-Deformation und nicht-axiale Kräfte [13].

Desweiteren spielt die Beschaffenheit der Instrumente und des Gewebes eine Rolle. Interaktionen von starren chirurgischen Instrumenten mit festem Gewebe müssen einen hohen Widerstand („virtuelle Wand“) simulieren. Das flexible chirurgische Instrument würde sich nach der Krafteinleitung verbiegen und wird hauptsächlich durch das Einklemmen der Finger bedient. Der Finger spürt den Kraftabprall, wenn sich das flexible chirurgische Instrument verbiegt. Die Simulation des Kraftaufpralls muss eine genaue „Kraft-Weg-Beziehung“ ergeben, die der Steifigkeit des Instruments entspricht [7].

Überdies fallen Vibrationsmuster bei der Interpretation von Haptik ins Gewicht. In einer weiteren Studie wurde gezeigt, dass die wahrgenommene Nachgiebigkeit einer Platte, die ein Studienteilnehmer mit den Füßen betreten sollte, mit der Intensität des Vibrationsfeedbacks monoton anstieg und in geringerem Maße von der zeitlichen oder Frequenzverteilung der Rückkopplung abhing. Wenn sowohl die Plattensteifigkeit, also die inversive Nachgiebigkeit, als auch die Vibrationsamplitude manipuliert wurden, blieb der Effekt bestehen, wobei beide Faktoren zur Wahrnehmung der Nachgiebigkeit beitrugen. Ein signifikanter Einfluss der Vibration wurde sogar für Amplituden beobachtet, die nahe an den psychophysischen Nachweisschwellen lagen. Diese Ergebnisse zeigen, dass vibrotaktile sensorische Kanäle für die Wahrnehmung der Oberflächen-Nachgiebigkeit von großer Bedeutung sind, und legen nahe, dass Korrelationen zwischen vibrotaktilen sensorischen Informationen und motorischer Aktivität für die Kontrolle der menschlichen Fortbewegung von größerer Bedeutung sein können [17].

Laut einer gründlichen Evaluierung der haptischen Kraftausgabe mit einem bereits vorgeschlagenen Framework ist visuo-haptisches Rendering allein auf Basis von Bilddaten und Segmentierung ohne Einbeziehung von Oberflächen- und Tetraedermeshes erfolgreich [3].

2.4 Kombination aus Sinnesansprachen

Es besteht die Theorie, dass neuronale Systeme möglicherweise nicht einzelne sensorische Ausgaben unabhängig voneinander verarbeiten, sondern dass stattdessen multisensorische Interaktionen untereinander bestehen. Auch bei der taktilen Wahrnehmung läuft die visuelle Verarbeitung mit und ist notwendig für eine optimale taktile Wahrnehmung bestimmter Objekteigenschaften [11]. Das ausgeprägte haptische Feedback ist beispielsweise sehr wichtig bei der virtuellen Darstellung des Greifens von flexiblen Materialien, wie z.B. Schaumstoff. Eine hohe Reaktionskraft muss gegeben sein, da die Kraft nur dann erkannt werden kann, wenn solche Materialien ausgiebig und spürbar gegriffen werden [19].

Im Prinzip sollten alle simulierten Sinneseindrücke miteinander verbunden sein - zum Beispiel sollte ein schallproduzierendes Objekt lauter werden, wenn es visuell näher erscheint, und der Schall sollte aus der gleichen Richtung kommen, aus der wir das Objekt sehen. Ebenso sollte man bei der Haptik die Kante eines Objekts genau an der Stelle fühlen können, an der man diese Kante sieht. Die genaue Lokalisierung von Haptik ist jedoch technisch schwieriger zu erreichen als die Lokalisierung von auditiven und visuellen Hinweisen. Dies liegt vor allem an dem größeren Wahrnehmungsspielraum bei der Lokalisierung der Reizquelle über akustische und visuelle Hinweise. Das als „Bauchredner-Effekt" bekannte Phänomen, bei dem ein Audiostimulus, der räumlich näher an einem visuellen Stimulus liegt, als vom Ort des visuellen Stimulus ausgehend wahrgenommen wird, wurde auch für die räumliche Dominanz anderer sensorischer Reize gegenüber akustischen Reizen nachgewiesen. Man ist also zu einem gewissen Grad tolerant gegenüber Ungenauigkeiten bei der visuell-akustischen Ko-Lokation als bei der visuell-haptischen Ko-Lokation [16].

Bei einer Greif-Aufgabe waren die Bediener in der Lage, die für die Ausführung erforderliche Kraft zu differenzieren, indem sie die Kraftrückmeldung zum einen grafisch überwachten und die Antriebskraft mit Hilfe von Audiosignalen und Reaktionsmomenten von Joysticks steuerten, was zeigt, dass Fähigkeit der Erkennung der Reaktionskraft in verschiedenen Modalitäten die Leistung der Bediener bei der Manipulationsaufgabe erhöht hat.

Die Wahrnehmung der Kraft in gemischten Modi hilft, Kraftvariationen zu unterscheiden und das Sicherheitsbewusstsein beim Greifen von Materialien zu verbessern. Insgesamt lässt sich sagen, dass die Kombination der verschiedenen Rückmeldungen zur Maximierung der Kontrolle und Kraftunterscheidung führt.

In einem psychophysikalischen Test, der innerhalb dieser Studie durchgeführt wurde, wurde die Analyse durch Berechnung des Verhältnisses zwischen der Stimulusintensität ΔI und fünf Referenzstimuli durchgeführt. Ein kleinerer Wert des Verhältnisses entspricht einer höheren Empfindlichkeit der Wahrnehmung der Benutzer gegenüber Rückmeldungen. Der Test umfasst Stimuli, die auf Kraftrückmeldung, auditivem und visuellen Feedback basieren. Die Wahrnehmung des auditiven Feedbacks war fast dreimal besser als reines Kraftfeedback. Die Kraftwahrnehmung der Bediener im visuellen Modus war dreimal besser als bei der Verwendung von Joysticks und einmal besser als bei der Verwendung von Audiosignalen. Die Ergebnisse zeigen, dass die Bediener für jede Modalität die Kraft mit unterschiedlicher Empfindlichkeit wahrnehmen. Die Einschränkung bei der Wahrnehmung des Stimulus zeigt, dass die Integration aller Stimuli in ein effektives multimodales Feedback notwendig ist, um die menschliche Leistung zu verbessern [19].

3 Stand der Technik: Aktuelle Anwendungsgebiete in der Medizin

Durch die Anwendung von den genannten Simulationsmöglichkeiten soll Virtuelle Realität der Medizin unter Anderem dabei helfen, die Ausbildung durch neue Einblicke in einen virtuellen Patienten zu verbessern und insofern das Risiko des Patienten zu minimieren. Gleichzeitig bietet eine Vielzahl von Visualisierungstechniken neue Einblicke in die menschliche Anatomie [4]. Das macht eine präoperative Eingriffsplanung zur Steigerung des Patientenwohls möglich und spart dabei Behandlungskosten, wie zum Beispiel durch die Reduzierung von Nadelrepositionierungen bei Leberpunktionen [12]. Ein wesentlicher Punkt ist das virtuelle Training von invasiven chirurgischen Eingriffen, das anhand von Simulatoren, die eine visuo-haptische Benutzeroberfläche bieten, unterstützt wird [14]. Die virtuelle Chirurgie bietet Chirurgen die Möglichkeit, verschiedene Arten von chirurgischen Eingriffen zu erleben und zu erlernen, indem sie visuelle und physische Informationen liefert. Das notwendige Operationstraining kann nach Bedarf wiederholt werden und somit werden lange Trainingszeiten verringert [7]. Im Folgenden werden ein paar Beispiele aus der Medizin dargelegt.

3.1 Entwicklung von virtueller Spritzengebung

Im Bereich der virtuellen Spritzengebung gibt es bereits einige Ansätze, die in der Medizin angewandt werden. In Bezug auf die Haptik müssen die Simulationen des Nadeleinstichs die Steifigkeit, die Schnitt- und Reibungskräfte an der Nadelspitze und am Schaft nachahmen. Wie bereits erwähnt, sind die Nadelauslenkung, die Gewebedeformation und nicht-axiale Kräfte ebenfalls relevant für realistische, visuo-haptische Simulationen in Echtzeit [13].

Ein Ansatz beschreibt eine haptische Vorrichtung mit sechs Grad Bewegungsfreiraum zur Rückmeldung von Kräften, die dem Einstechen und Drehen der Nadel widerstehen. Ein verbessertes haptisches Volumen-Render-Verfahren wird zur Berechnung der Kräfte für das haptische Feedback verwendet, sodass bei der Anwendung Echtzeit-3D-Visualisierung mit optionaler Stereoansicht einen Überblick über die punktierte Region geben können. Dabei ermöglichen 2D-Visualisierungen von orthogonalen Schnitten einen detaillierten Eindruck des anatomischen Kontextes und erfahrene Ärzte können durch das unterschiedliche haptische Verhalten dieser Strukturen Informationen über die

Nadelposition erhalten. Die virtuellen haptischen Modelle werden aus segmentierten Bilddaten generiert und die Berechnung der Kräfte werden in Echtzeit in Abhängigkeit von Nadelposition, Nadeldrehung, Einstichwinkel und lokalen Gewebeeigenschaften berechnet. Das haptische Gerät wird dann verwendet, um die gerenderten Kräfte an den Benutzer zurückzugeben. Widerstandskraft, Oberflächenreibung und Viskosität von medizinischen Strukturen werden basierend auf einem haptischen Volumen-Render-Ansatz simuliert [4].

In einer weiteren Studie wurde ein neuartiges Konzept im Kontext der Leberpunktionssimulation unter Verwendung eines Force-Feedback-Evaluations-Frameworks untersucht, bei dem direkt das Nadelkraft-Volumen gerendert werden soll. Die haptischen Algorithmen dieses Simulatorsystems basieren zum einen auf teilweise segmentierten Patientenbilddaten und zum anderen auf einem nichtlinearen, an Organgrenzen wirksamen Federmodell. Der neue Simulator AcusVR-4D mit US Imaging simulierter Nadelführung, der in Abbildung 1 zu sehen ist, wird als Trainings- und Planungsinstrument eingesetzt. Das vorherige System bestand aus einem Haptikgerät mit einer Kombination aus Shutterbrille und einem Röhren-Monitor für die Immersion in der Virtuellen Realität. Dieser 3D-Simulator war hauptsächlich für das Training von Lumbalpunktionen gedacht, bei denen eine kleine Menge Nervenwasser aus dem Bereich der Lendenwirbel entnommen wird [13].

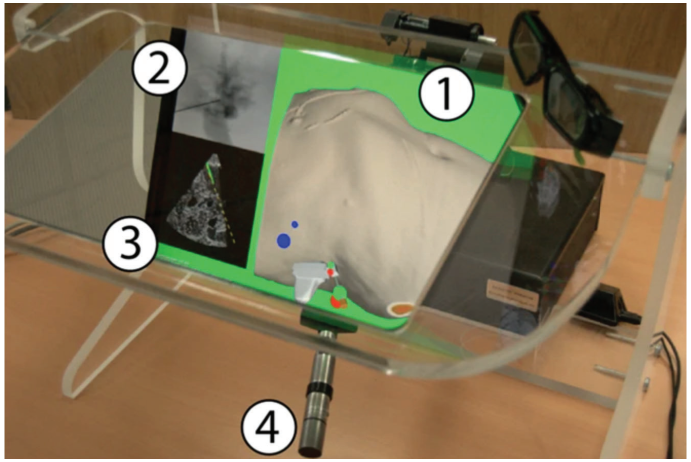

Abbildung 1 Haptische Workbench und Simulation: (1) Haupt-Render-Fenster, (2) Röntgen-Render-Fenster, (3) Ultraschall-Render-Fenster, (4) haptischer Gerätegriff [13] .

3.2 Kardiovaskuläre Chirurgie

Ein weiterer Anwendungsfall der virtuellen Haptik-Simulation findet sich in kardiovaskulären Systemen. Diese betreffen das Herz und das Gefäßsystem. Hierfür wurde ein dynamisches Modell eines Führungsdrahtes entwickelt, also eines dünnen Drahtes, der zur Einführung, Stabilisierung und Positionierung von Kathetern dient. Mit diesem soll die Kollision eines Führungsdrahtes im Blutgefäß oder in der Rohrleitung des Kräfte-Feedback-Geräts simuliert werden können. Der Führungsdraht ist dynamisch und dabei für das Kräfte-Feedback zuständig. Eingesetzt wird er in der virtuellen kardiovaskulären Chirurgie und der Mensch-Roboter-Mikrointeraktions-Chirurgie.

Die ausgeübte Kraft ist sehr gering und kann schnell durch die Bewegung und Reibung des Aktuators und das Zittern der Hand beeinflusst werden, weshalb eine Anforderung an die Robustheit des Systems besteht. Dann wird ein Terminal-Gleitmodusregler entworfen, um die Robustheit und Konvergenz der Kraftfeedbacksteuerung zu verbessern. Da die Kraftinformationen bei virtuellen kardiovaskulären Eingriffen wichtig sind, muss die virtuelle Operation dem Bediener ein Gefühl für die Kraft vermitteln, das gleich einer realen Operation ist.

Es gibt hier zwei Kategorien des Kraftfeedbackgeräts:
Zum einen gibt es die starren Geräte, wie es die meisten virtuellen chirurgischen Instrumente sind, einschließlich Skalpell-Simulatoren, Simulatoren der virtuellen Gefäßnaht-Chirurgie und Simulatoren der laparoskopischer Chirurgie, bei der der Bauchraum von innen mit einer Kamera untersucht wird. Die Reibung, Trägheit und andere Faktoren würden die Genauigkeit des starren Kraftinteraktionsgeräts beeinflussen.
Es gibt aber auch flexible Geräte, einschließlich des Führungsdrahtes und des Katheters für die virtuelle kardiovaskuläre Chirurgie. Die flexiblen Instrumente würden sich unter Krafteinwirkung verbiegen, und ihr dynamisches Modell ist nichtlinear.

Die aktuelle Modellierungsmethode besteht darin, das Impedanzmodell durch die Beziehung zwischen Verschiebung und Kraftrückführung zu erstellen. Es handelt sich bei dem chirurgischen Instrument um einen flexiblen Körper, sodass das Steuerungssystem die Auswirkung der Biegeverformung auf die dynamische Leistung des Systems berücksichtigen muss. Da der Zustand der Biegeverformung variiert, ändert sich die dynamische Leistung mit der Veränderung der virtuellen Umgebung [7].

3.3 Katheterlegung

Kathetergeräte bieten einen einfachen und schmerzlosen Zugang zum Inneren des menschlichen Körpers durch natürliche Öffnungen und Gefäße anhand von dünnen, flexiblen Drähten und Schläuchen. Bei der Einsetzung ist allerdings keine Beurteilung physikalischer Eigenschaften des Gewebes im Körperinneren möglich wegen der Bewegung des Gewebes und der Übertragungsbeschränkungen der Kathetergeräte bezüglich Nachgiebigkeit, Reibung und Rückschlag. Deshalb ist die Bereitstellung eines haptischen Feedbacks während der manuell abtastenden Palpationsverfahren hilfreich, um die taktilen Informationen, die dem Arzt bei Kathetereingriffen zur Verfügung stehen, abzubilden. Durch ein bewegungskompensiertes aktuiertes Kathetersystem für eine haptische Wahrnehmung von sich schnell bewegenden Gewebestrukturen soll dies ermöglicht werden. Die Informationen, die den Ärzten zur Verfügung stehen, werden dabei über das hinaus erweitert, was derzeit durch Röntgen- oder Ultraschallbilder bereitgestellt ist, um beispielsweise das Gewebe um eine Herzklappe herum zu ertasten und zu untersuchen, ob sie verkalkt oder verengt und stenotisch ist.

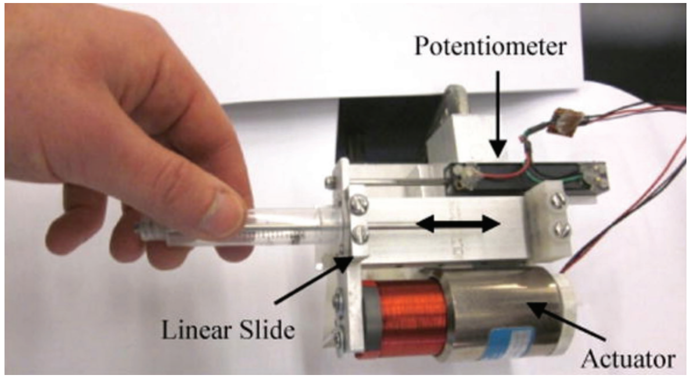

Abbildung 2 Die betätigte haptische Schnittstelle des Katheters [9] .

Der betätigte Katheter, zu sehen in Abbildung 2, ist mit einem Kraftsensor an der distalen Spitze, die vom Herzen weiter entfernt liegt, und einer Kraft-Feedback-Schnittstelle ausgestattet, die es dem Benutzer ermöglicht, die Position des Katheters einzustellen, während er die Kräfte an der Katheterspitze wahrnimmt. Die Wirksamkeit dieses Geräts und der Schnittstelle wurde durch eine psychophysikalische Studie evaluiert, in der verglichen wurde, wie genau Benutzer verschiedene Materialien, die an einem Herzbewegungssimulator angebracht sind, unter Verwendung des haptischen Geräts und eines herkömmlichen manuellen Katheters unterscheiden können. Die Haptik verbessert die Fähigkeit des Benutzers zur Unterscheidung der Materialeigenschaften und die Gesamtzahl der Fehler wird im Vergleich zum manuellen Kathetersystem um die Hälfte verringert.

Das Problem liegt darin, dass die lange, flexible Beschaffenheit von Herzkathetern, durch die sie sich leicht in den Körper einführen lassen, sie auch schlecht macht bei der Übertragung von Kraftfeedback-Informationen an den Bediener. Wenn ein Katheterwerkzeug mit dem Gewebe in Kontakt kommt, wird die Kontaktkraft durch die Katheternachgiebigkeit und Reibungsverluste durch Dichtungen und viskose Flüssigkeiten ausgeglichen [9].

4 Grenzen des Forschungsbereichs

Die Wahrnehmung der haptischen Illusion ist vor allem dann optimiert, wenn der Zeitpunkt der Rückmeldung stimmt. Beispielsweise kann jedoch eine schlechte Netzwerkverbindung die sinnvolle Zusammenarbeit in virtuellen Umgebungen erheblich einschränken. Eine Studie zeigt, dass eine Latenz im visuellen Feedback sowohl den Zeitaufwand für die Ausführung einer Aufgabe als auch die Anzahl der gemachten Fehler erhöht. Für eine Analyse der Handschrift beispielsweise mussten die Teilnehmer auf einem Telescriber schreiben, einem Vorläufer des Faxes, das zur Übertragung der Handschrift über ein Netzwerk verwendet wird. Dabei wurde beobachtet, dass eine Verzögerung zwischen der Bewegung der Stifte durch die Teilnehmer und der Sichtbarkeit der resultierenden Markierungen dazu führte, dass Buchstaben nicht nur mehr langsamer geschrieben wurden, sondern auch mehr in der Form variierten. Die Beziehung zwischen dem Grad der Latenz und der Verschlechterung der Leistung scheint linear zu sein: Die Schreibgeschwindigkeit nahm proportional mit der Verzögerung ab [8].

Speziell in der Medizin liegen die Grenzen darin, dass es allgemein sehr schwierig ist, die Simulatoren-Systeme herzustellen, da der Eingriff in neue Patientendaten sehr viel Zeit und Aufwand für die patientenspezifische Vorbereitung der manuellen Segmentierung erfordert. Die Simulatoren sind aber auf einen vollständig segmentierten 3D-Patienten angewiesen. Der hohe Segmentierungsaufwand verhindert oft den Einsatz von Simulatoren für Virtuelle Realität im klinischen Umfeld.

Können diese Probleme umgangen werden, muss man allerdings immer noch die Bewegungsdynamik des Patienten beachten. Die Geräte müssen an die Anatomie des Patienten angepasst werden, da sie sonst nicht zum Einsatz kommen können. Auch das kann Hindernisse für die Entwicklung der Feedback-Geräte darstellen [13].

5 Evaluierung der Ergebnisse und Ausblick in den weiteren technischen Fortschritt in der Medizin

Zusammenfassend lässt sich sagen, dass bereits sehr viele Studien und Forschungen zum Thema Simulation von haptischer Rückmeldung und teils sogar physikalischer Kräfte bestehen und sehr viele verschiedene Herangehensweisen dafür gefunden wurden. Visuelle, auditive und haptische Rückmeldungen haben bereits einzeln schon Effekt, doch je mehr Sinne pro Zeit angesprochen werden, desto eingängiger ist die Illusion von Haptik. Einige Methoden wurden schon auf medizinische Verfahren angepasst oder weiterentwickelt, wie in den vorgestellten Fällen der virtuellen Spritzengebung und Katheterlegung oder auch in der kardiovaskulären Chirurgie. Neben der Entwicklung aktueller Anwendungen gehen weitere Entwicklungen in der Mikrochirurgie vorangehen [15].

Neben logistischen Problematiken wie dem Aufwand bei der patientenspezifischen Informationsbeschaffung und der virtuellen Umsetzung davon, bestehen auch noch technische Umsetzungsschwierigkeiten und Möglichkeiten zur Verfeinerung. Am Beispiel des virtuellen Katheters steht noch Arbeit aus, um das Konzept zu opitimieren. Es ist geplant, dass das betätigte Kathetersystem miniaturisiert wird für eine bessere Arbeit mit klinischen Kathetertechniken. Eine Verbesserung der Bewegungskompensation für die Reduzierung des haptischen Rauschens, das durch Zielverfolgungsfehler verursacht wird, ist notwendig. Mit der Optimierung ist es aber möglich, dass dieses System es Klinikern eines Tages ermöglichen wird, die Organe und Gewebestrukturen im Körper zu ertasten und genauer mit ihnen zu interagieren, wodurch sich möglicherweise eine neue Welt von diagnostischen und interventionellen Verfahren eröffnet, die mit einem Katheter durchgeführt werden können [9].

Ebenso wird der virtuelle Ansatz in der Medizin auch in vielen anderen Bereichen moderne Techniken der Diagnosen und Behandlungen begünstigen, da durch das haptische Feedback neue Informationen gesammelt werden können.

Literatur

1 Parastoo Abtahi, Benoit Landry, Jackie Junrui Yang, Marco Pavone, Sean Follmer, and James A. Landay. Beyond the force: Using quadcopters to appropriate objects and the environment for haptics in virtual reality. Association for Computing Machinery, 5 2019. doi:10.1145/3290605.3300589.

2 Inrak Choi, Heather Culbertson, Mark R. Miller, Alex Olwal, and Sean Follmer. Grabity: A wearable haptic interface for simulating weight and grasping in virtual reality. pages 119–130. Association for Computing Machinery, Inc, 10 2017. doi:10.1145/3126594.3126599.

3 Dirk Fortmeier, Andre Mastmeyer, Julian Schröder, and Heinz Handels. A virtual reality system for ptcd simulation using direct visuo-haptic rendering of partially segmented image data. *IEEE Journal of Biomedical and Health Informatics*, 20:355–366, 1 2016. doi:10.1109/JBHI.2014.2381772.

4 M Färber, J Heller, F Hummel, C Gerloff, and H Handels. Virtual reality based training of lumbar punctures using a 6dof haptic device.

5 Yoren Gaffary, Benoit Le Gouis, Maud Marchal, Ferran Argelaguet, Bruno Arnaldi, and Anatole Lecuyer. Ar feels 'softer' than vr: Haptic perception of stiffness in augmented versus virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, 23:2372–2377, 11 2017. doi:10.1109/TVCG.2017.2735078.

6 Rainer Herpers, Wolfgang Heiden, Michael Kutz, David Scherfgen, Ulrich Hartmann, Jens Bongartz, and Oliver Schulzyk. Fivis–bicycle simulator in the immersive visualisation environment "immersion square". In *Forschungsspitzen und Spitzenforschung*, pages 61–68. Springer, 2009.

7 Zhi Hu, Yueying Wang, Su Yu, and Guohua Cui. A method to enhance fidelity of force feedback control in virtual and human-robot micro interaction cardiovascular intervention surgery. *International Journal of Advanced Manufacturing Technology*, 105:4883–4897, 12 2019. doi:10.1007/s00170-019-03936-3.

8 Caroline Jay, Mashhuda Glencross, and Roger Hubbold. Modeling the effects of delayed haptic and visual feedback in a collaborative virtual environment. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 14(2):8–es, 2007.

9 Samuel B. Kesner and Robert D. Howe. Discriminating tissue stiffness with a haptic catheter: Feeling the inside of the beating heart. pages 13–18, 2011. doi:10.1109/WHC.2011.5945454.

10 Anatole Lécuyer. Simulating haptic feedback using vision: A survey of research and applications of pseudo-haptic feedback. *Presence: Teleoperators and Virtual Environments*, 18(1):39–53, 2009.

11 Nadia Magnenat-Thalmann and Ugo Bonanni. Haptics in virtual reality and multimedia. *IEEE MultiMedia*, 13(3):6–11, 2006.

12 André Mastmeyer. Patientenspezifische virtual-reality-simulationen: Ultraschallgestützte nadelpunktionen in atmungsbewegten virtuellen patienten. 2017.

13 Andre Mastmeyer, Dirk Fortmeier, and Heinz Handels. Evaluation of direct haptic 4d volume rendering of partially segmented data for liver puncture simulation. *Scientific Reports*, 7, 12 2017. doi:10.1038/s41598-017-00746-z.

14 Andre Mastmeyer, Matthias Wilms, and Heinz Handels. Interpatient respiratory motion model transfer for virtual reality simulations of liver punctures. 7 2017. URL: http://arxiv.org/abs/1707.08554.

15 Rory McCloy and Robert Stone. Virtual reality in surgery. *BMJ*, 323(7318):912–915, 2001. URL: `https://www.bmj.com/content/323/7318/912`, `arXiv:https://www.bmj.com/content/323/7318/912.full.pdf`, `doi:10.1136/bmj.323.7318.912`.

16 David Swapp, Vijay Pawar, and Céline Loscos. Interaction with co-located haptic feedback in virtual reality. *Virtual Reality*, 10(1):24–30, 2006.

17 Yon Visell, Bruno L Giordano, Guillaume Millet, and Jeremy R Cooperstock. Vibration influences haptic perception of surface compliance during walking. *PLoS one*, 6(3):e17697, 2011.

18 Ahmad Anas YUSOF, Takuya KAWAMURA, and Hironao YAMADA. Operational evaluation of a construction robot tele-operation with force feedback. 43(1):8–15, 2011.

19 Ahmad Anas Yusof, Takuya Kawamura, and Hironao Yamada. Evaluation of construction robot telegrasping force perception using visual, auditory and force feedback integration. *Journal of Robotics and Mechatronics*, 24:949–957, 2012. `doi:10.20965/jrm.2012.p0949`.

Intelligentere Lernerinnerungen im Mobilen Kontext

Benedikt Pirker

Ludwig-Maximilians-Universität München, München, Deutschland
`Benedikt.Pirker@campus.lmu.de`

Zusammenfassung

Kontinuierliche Wiederholung und Übung sind essentiell um signifikante Lernfortschritte zu erzielen. Mobile Lernapplikationen bieten ihren Nutzern die Möglichkeit durch Benachrichtigungen an Lernaufgaben erinnert zu werden. Jedoch werden diese Benachrichtigungen aktuell zu starren Zeitpunkten ausgelöst, was für Nutzer und Nutzerinnen häufig unpraktisch ist. Unvorhersehbare Tagesabläufe und eine höhere Priorisierung anderer Tätigkeiten können dazu führen, dass der Nutzer bzw. die Nutzerin zur vorgesehenen Zeit eine andere Aktivität ausführt. Eine Lernerinnerung würde dann häufig als störend und sogar unerwünscht empfunden werden. Um dieses Problem zu lösen, können intelligente Erinnerungsstrategien implementiert werden, welche die Erinnerungen auf die Tagesabläufe und Bedürfnisse der Nutzer und Nutzerinnen abstimmen. Zusätzlich können Methoden implementiert werden, welche eine Erinnerung der Lern-Applikation überflüssig machen. Dieser Artikel gibt einen Überblick über verschiedene Strategien und diskutiert ihre Implementierung in einer Lernapplikation. Die Strategien werden allgemein und mit Lernbezug betrachtet. Bei der Implementierung wird auf verschiedene Gestaltungsmöglichkeiten, technische Herausforderungen und diverse weitere Herausforderungen genauer eingegangen.

2012 ACM Computing Classification Applied computing → Education → E-learning; Human-centered computing → Ubiquitous and mobile computing → Ubiquitous and mobile computing theory, concepts and paradigms → Mobile computing

Keywords and phrases M-Learning; Zeitplanung; Intelligente Erinnerungen; Benachrichtigungen; Lernen.

1 Einführung

Mit dem Siegeszug des persönlichen Computers ('PC') haben sich elektronische Medien in fast allen Bereichen unseres Lebens etabliert [25, 14]. Das Potential des sogenannten elektronischen Lernens ('E-Learning'), also die Nutzung

Cite as: Benedikt Pirker. Intelligentere Lernerinnerungen im Mobilen Kontext. In *6th Seminar on Ubiquitous Interaction (UBIACTION 2021)*. Editors: F. Lang, R. Welsch, L. Haliburton, F. Draxler, S. Feger, S. Villa, M. Hoppe, P. Knierim, V. Mäkelä, A. Schmidt. February 11, 2021. Munich, Germany. pp. 3:1–3:16.

elektronischer Medien für Lernaufgaben, wurde dabei schon sehr früh erkannt und genutzt [39, 20, 26]. Vorteile wie Individualisierung, Zentrale Lieferung der Inhalte, Zeit- und Ortsunabhängigkeit, Interaktivität, fortgeschrittene Visualisierungsmöglichkeiten, unterschiedliche Aufnahmekanäle wie Audio, Video haben E-Learning auch für Aufgabenbereiche interessant gemacht, welche an sich nichts mit elektronischen Medien zu tun haben, wie z.B. Mathematik oder Sprachen [8, 5]. Durch das Internet können Lerninhalte unabhängig vom eigenen Standort und der aktuellen Tageszeit abgerufen werden. Dies hat Fernuniversitäten zu höherer Popularität verholfen und Menschen die Möglichkeit gegeben abseits einer klassischen Lerninstitution neues Wissen und Fähigkeiten zu erwerben [29]. Aber auch Schulen und herkömmliche Universitäten haben elektronische Medien längst in ihre Lernangebote integriert [30, 24]. Zum Zeitpunkt des Verfassens dieses Artikels zeigt sich durch die globale Pandemie und die daraus resultierenden Ausgangsbeschränkungen das Liefern von Lerninhalten über das Internet als so wichtig wie nie zuvor [35]. Jedoch sieht man auch gerade in dieser Zeit, dass E-Learning noch nicht in der Lage ist, herkömmliche Lernmethoden gänzlich zu ersetzten. Nachteile wie (die Schulung der Pädagogen auf E-Learning, die nötige Selbstdisziplin und Selbstlernkompetenz, technische Herausforderungen, die reduzierte soziale Nähe zum Lehrer/Pädagogen machen es für Instituionen mit jüngeren Kursteilnehmern (z.B. Unterstufe, Grundschule) wünschenswert wenn möglich wieder zum klassischen Unterrichtsstil zurück zu kehren [17, 41].

2 Mobiles Lernen

Mobiles Lernen ('M-Learning') stellt eine Weiterentwicklung des E-Learnings dar. Dabei handelt es sich um Lernen mit ubiquitären und mobilen Medien. Heutzutage spricht man dabei meist vom Lernen auf Mobiltelefonen und Tablet-Computern.

2.1 Chancen und Herausforderungen bei mobilem Lernen

Verglichen mit E-Learning verändern sich bei M-Learning durch den unterschiedlichen Geräteformfaktor und die teils stark wechselnde Lernumgebung einige Aspekte. Ein großer Vorteil ist die hohe Flexibilität von Zeit und Ort. Der Nutzer / die Nutzerin hat das mobile Gerät meistens bei sich, kann also 'immer und überall' lernen. Durch die Nähe zum Mobiltelefon können Nutzer und Nutzerinnen ebenfalls viel besser an ihre Lerneinheiten erinnert werden, zum Beispiel durch Vibrations- bzw. Ton-Alarme und Benachrichtigungen. Die Fülle an Sensoren eines Mobiltelefons kann zusätzlich Aufschlüsse über den Umgebungskontext und sogar den emotionalen Zustand des Nutzers / der Nutzerin geben und dabei helfen, die richtige Lernmethode auszuwählen. Befindet

sich die Person zu Hause, können aufgrund der ruhigeren Umgebung andere Lektionen vorgeschlagen werden, wie etwa in der U-Bahn (z.B. Spracheingabe wird möglich). Hat die Person wiederholt negative Emotionen, gefallen ihr die Lern-Lektionen eventuell nicht. Hier könnte auf eine Lernmethode gewechselt werden, die in der Vergangenheit positive Emotionen ausgelöst hat. Mobiles lernen bring jedoch auch einige Nachteile mit sich. Die kleinere Größe eines mobilen Gerätes führt dazu, dass der Lerninhalt dafür optimiert werden muss. Eine hohe Flexibilität der Lernumgebung kann auch zu einer raschen Änderung des Ortes führen, welche das Lernen erschwert oder wegen der man das Lernen sogar unterbrechen muss. Umgebungsgeräusche können zusätzlich manche Eingabe- und Ausgabemodalitäten (wie Sprach Ein- und Ausgabe) behindern. Sowohl die kleine Bildschirmgröße, als auch die unberechenbare Lernumgebung haben dazu geführt, dass aktuell meist kleinere Lern-Lektionen implementiert werden, welche ständig unterbrochen und fortgeführt werden können. Bei Sprachlernen eignet sich zum Beispiel Vokabeltraining sehr gut um die Nachteile von M-Learning zu kompensieren und Videokurs-Applikationen bieten ihre Lektionen meist in vielen kleinen Teilvideos an. Diese Art des Lernens wird als Mikrolernen (englisch: 'Microlearning') bezeichnet.

2.2 Aktuelle Lernerinnerungsmethoden und deren Effekt auf den Lernerfolg

Das Erlernen komplexer Fähigkeiten erfordert kontinuierliche Wiederholung und Übung, damit signifikanter Fortschritt erzielt wird. Um die Lernmotivation der Nutzer und Nutzerinnen aufrecht zu erhalten, haben sich sogenannte Echtzeitbenachrichtigungen bzw. **'Push-Notifications'** unter bestimmten Voraussetzungen als valide Methode bewiesen. Hier ist es wichtig dass ihr Inhalt wichtig und relevant genug ist und sie nicht zu häufig verwendet werden, da sie sonst als störend empfunden werden können [31]. Als wichtig und relevant werden im allgemeinen Benachrichtigungen über eingehende Nachrichten von anderen Menschen, über Veranstaltungen in der echten Welt oder mit Bezug auf eine Person in den Kontakten der Nutzer eingestuft [34].

Ein weiterer beliebter Ansatz um die intrinsische Motivation der Nutzer und Nutzerinnen zu erhöhen ist die Integration von Spielelementen in die Lernumgebung, auch bekannt als **'Gamification'**. Wenn dies korrekt implementiert wird, kann dadurch die Motivation häufiger mit der Lern-Applikation zu interagieren gesteigert werden [38, 13]. Anstatt einer einfachen Lernerinnerung, würde ein Nutzer bzw. eine Nutzerin daran erinnert werden, ein bestimmtes Ziel zu erreichen. Dieses Ziel würde dem Nutzer / der Nutzerin eventuell wichtiger erscheinen, als eine einfache Lernerinnerung und er / sie wäre dadurch vielleicht eher gewillt die vorgeschlagene Lektion zu absolvieren. Falls die Lern-Lektion nicht wie geplant absolviert wird, könnte es dafür spielerische Konsequenzen geben. Jedoch existiert, nach aktuellem Wissensstand des Autors, keine Forschung welche die Effektivität von Lernerinnerungen mit Gamification-Charakter untersucht, oder diese mit einfachen Push-Notification-basierten Lernerinnerungen vergleicht. Forschung, welche die allgemeine Wirkung von Gamification Ansätzen in Lernumgebungen untersucht, ist im Gegensatz dazu bereits verfügbar.

Fan und Wang (2020) haben in ihrer Studie beobachtet, dass die kompetitive und kooperative Interaktion mit Mitlernern durch Gamification einen positiven Effekt auf die Lernleistung der Nutzer und Nutzerinnen haben kann [21]. Bei der Studie wurden Gamification Techniken in Englisch Lernapps verwenden. In der Studie von Su & Cheng korrelierte eine höhere Motivation der Teilnehmer und Teilnehmerinnen ebenfalls mit höherem Erfolg [38]. Das Gamification-Konzept von Bicen & Kocakoyun führte dazu, dass das Lernen als lustiger und unterhaltender empfunden wurde, was wiederum zu einer höheren Beteiligung der Probanden und Probandinnen führte [13].

Um eine Überblick über den Status Quo der Erinnerungs- und Motivationssteigerungsfunktionen in den aktuell populärsten Lernapplikationen geben zu können, werden in der folgenden Sektion die Top Vier der jeweils populärsten Lern-Apps im Google Play Store analysiert. Dabei werden die Apps nach ihren Kategorien gruppiert. Diese sind Sprachen-Lernapps und kursbasierte Lernapps. Die Downloadzahlen entsprechen dem Zeitpunkt Januar 2021.

2.2.1 Sprachen-Lernapps

- Duolingo (100 Millionen+ Downloads) [3]
- Rosetta Stone (10 Millionen+ Downl.) [9]
- Mondly (10 Millionen+ Downl.) [7]
- Babbel (10 Millionen+ Downl.) [1]

Duolingo, Rosetta Stone und Mondly bieten lediglich die Möglichkeit einen fixen Zeitpunkt für Push-Notifications auszuwählen. Babbel bietet auf den ersten Blick mehr Alltagsintegration für die Erinnerung an, wie z.B. 'Morgens', 'Mittags', 'Unterwegs' oder 'In meiner Pause'. Jedoch verbirgt sich hinter diesen Bezeichnungen auch die Auswahl einer starren Uhrzeit und von Wochentagen. In Abbildung 1 ist eine solche Erinnerung zu sehen.

Duolingo und Mondly bieten ähnliche Gamification-Features innerhalb der mobilen Applikation. Lektionen bringen Erfahrungspunkte, welche genutzt werden können um höhere Level zu erreichen. Zusätzlich kann man sich mit seinen Freunden und mit anderen Spielern in globalen Ranglisten vergleichen. Eine unterscheidende Eigenschaft sind die sogenannten 'Streaks' innerhalb von Duolingo. Hier werden Nutzer und Nutzerinnen belohnt, wenn sie ihre Lektionen jeden Tag absolvieren. Duolingo sendet im Zusammenhang mit Gamification auch Push-Notifications. Diese weisen unter anderem auf laufende

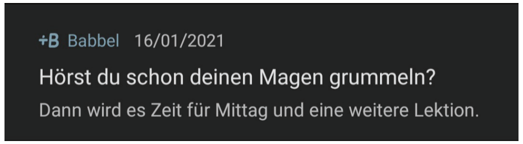

Abbildung 1 Beispiel einer vermeintlichen Integration eines Alltagserignisses in einer Echtzeitbenachrichtigung [1].

Streaks und erreichbare Ziele hin. Falls man die App über eine längere Zeit vernachlässigt und der Streak damit unterbrochen wurde, wird man ebenso mit einer Push-Notification informiert. Sowohl Babbel, als auch Rosetta Stone bieten aktuell keine Gamification Funktionen. Alle genannten Sprachen Lernapps bieten keine Integration von Standort Daten.

2.2.2 Kursbasierte Lernapps:

- Coursera (10 Millionen+ Downl.) [2]
- Udemy (10 Millionen+ Downl.)[10]
- LinkedIn Learning (5 Millionen+ Downloads) [6]
- edX (5 Millionen+ Downl.) [4]

In Coursera können Benachrichtigungen für laufende Kurse aktiviert werden, welche z.B. auf kommende Deadlines hinweisen. Diese können zusätzlich mit dem eigenen Kalender synchronisiert werden. Lernerinnerungen sind für jeden Kurs individuell möglich, wobei hier der Wochentag und ein fixer Zeitpunkt gewählt werden kann. Nutzer und Nutzerinnen können auch für jeden Kurs individuelle Ziele setzen, wie wie viel man zu welcher Zeit geschafft haben will und welche Fähigkeiten man am Ende des Kurses erlangt haben möchte. Udemy bietet lediglich die Möglichkeit für Erinnerungen zu fixen Zeitpunkten (Wochentag plus Zeit) und weder Gamification noch andere Erinnerungsfunktionen. LinkedIn Learning hat sogenannte 'Weekly Goals Reminder'. Man setzt sich ein Ziel (z.B. 15Min, 30Min Lernen) für die Woche und LinkedIn erinnert den Nutzer / die Nutzerin gelegentlich daran. Fixe Termine, an denen man Lernen möchte, kann man sich innerhalb der Plattform nicht einrichten. Die edX App bietet keine Erinnerungsfunktionen oder Motivationshelfer. Die Plattform sendet lediglich Erinnerungen über E-Mail.

Alle genannten, kursbasierten Lern-Applikationen bieten zum Zeitpunkt des Verfassens keine Gamification-Funktionen.

3 Strategien für intelligente Lernerinnerungen im mobilen Kontext

Die in Kapitel 2 genannten aktuell verbreiteten Methoden zur Steigerung der kontinuierlichen Lern-Aktivititäten von Nutzern und Nutzerinnen stellen bereits eine Verbesserung gegenüber M-Learning Modellen ohne sie dar [31, 38]. Jedoch gibt es noch Szenarien, in denen diese nicht funktionieren bzw. welche nicht abgedeckt werden. Echtzeitbenachrichtigungen leiden aktuell unter ihrer geringen Flexibilität. Viele Menschen haben einen Tagesablauf, welcher Ihnen nicht erlaubt am selben Wochentag immer zu selben Zeit verfügbar zu sein. Benachrichtigungen welche zu einem unpassenden Zeitpunkt gesendet werden, werden möglicherweise als störend empfunden und können zu einer Unterbrechung der aktuell fortlaufenden Aufgabe einer Person führen [28]. Dies kann wiederum negative Konsequenzen auf die erbrachte Leistung innerhalb einer Aufgabe [12], auf den emotionalen Zustand der Person [11] und die Gesamtdauer der primären Aufgabe haben [16]. Echtzeitbenachrichtigungen, welche zu einer Zeit geschickt werden, in denen Nutzer und Nutzerinnen beschäftigt sind, werden zwar oft wahrgenommen, jedoch häufig auch bewusst ignoriert, falls ihr Inhalt für den Empfänger bzw. die Empfängerin nicht relevant oder interessant genug ist [28].

In den folgenden Sektionen werden Ansätze vorgestellt, welche Nutzer und Nutzerinnen intelligenter als mit aktuell verbreiteten Methoden erinnern sollen, ihre Lern-Lektionen zu absolvieren. In Sektion 3.1 und 3.2 werden Push-Notification-basierte Methoden behandelt und in Sektion 3.3 werden Konzepte vorgestellt, um Erinnerungen schrittweise überflüssig zu machen. Wie in Sektion 2.2 bereits erwähnt, scheint die Forschung auf dem Bereich Lernerinnerungen mit Gamification-Charakter noch nicht sehr fortgeschritten zu sein, weshalb auf Gamification Ansätze in den folgenden Sektionen nicht genauer eingegangen wird.

3.1 Manuelle Platzierung einer Erinnerung in opportunen Zeitfenstern

Im menschlichen Alltag finden sich oft Zeitpunkte, die man weder mit produktiven noch unterhaltenden Tätigkeiten füllt. Mobile und ubiquitäre Geräte ermöglichen durch ihre hohe Verfügbarkeit diese Zeitpunkte gezielt mit Aktivitäten zu belegen. Durch technische Mechanismen und / oder Wissen über den Alltag von Personen, können opportune Erinnerungen manuell von einem Menschen platziert werden. Ein Beispiel wäre die Verwendung von Bluetooth-Beacons, welche beim mobilen Gerät eine Benachrichtigung auslösen, wenn es in der Nähe ist. Die Beacons könnten an Orten Platziert werden, an denen die Person oft eine Pause macht und daher verfügbarer sein könnte. Diese Art der Erinnerungs-Platzierung wird in der folgenden Sektion als manuelle Platzierung bezeichnet. Die Fehleranfälligkeit dieser Systeme ist relativ gering, jedoch ist ihre Flexibilität dafür beschränkt.

Cai et. al. (2017) haben mit WaitSuite ein System entwickelt, in dem vordefinierte Szenarien, in denen bekannte Wartezeiten auftreten, von den Forschern mit entsprechenden Lern-Lektionen gefüllt wurden [15]. Eingebettete Auslösemechanismen stellen dabei sicher, dass die Lern-Applikation in der jeweiligen Situation jedes mal aufgerufen wird. Das System ist in fünf verschiedene Untersysteme gegliedert, welche alle auf eine unterschiedliche Art des Wartens spezialisiert sind. Diese sind:

- ElevatorLearner: Der Nutzer / die Nutzerin lernt während er / sie auf den Aufzug wartet.
- PullLearner: Der Nutzer / die Nutzerin lernt nachdem er / sie den oberen Rand des Smartphone Bildschirms nach unten gezogen hat um ein Neuladen der Applikation zu bewirken.
- WifiLearner: Der Nutzer / die Nutzerin lernt während sich sein / ihr Computer mit dem Wi-Fi verbindet.
- EmailLearner: Der Nutzer / die Nutzerin lernt während seine / ihre E-Mail versendet wird.
- WaitChatter: Der Nutzer / die Nutzerin lernt während er / sie auf eine Antwort von einem Konversationspartner in einer Textnachrichten Applikation wartet.

ElevatorLearner und PullLearner hatten dabei die höchste Rate an Interaktionen. EmailLearner hatte etwas niedrigere Interaktion und WifiLearner und WaitChatter eine besonders niedrige Interaktionsrate. Die Autoren beschreiben dabei das Verhältnis zwischen Wartezeit und Reaktionszeit (also Zeit die benötigt wird, um zu einer Lernaufgabe und zurück zu wechseln) als ausschlaggebend für die Interaktionsrate. Je größer die Wartezeit gegenüber der Reaktionszeit ist, desto höher wird die Interaktion der Nutzerinnen und Nutzer mit der Lernaufgabe.

Ein naiver Ansatz, wie eine Push-Notification zu einem starren Zeitpunkt kann verbessert werden, wenn der Zeitpunkt auf den Tagesablauf des Nutzers / der Nutzerin abgestimmt wird. Localytics, eine Marketing Firma fand in ihrer Analyse, dass die beste Zeit für eine Push-Notification im Allgemeinen um 10:00 Lokalzeit sei. Die Klick-Rate liegt hier bei 15% [40]. Die Marketing Firma MailChimp lieferte eine noch genauere Analyse ihrer Klickraten für E-Mail-basierter Werbung [23]. Im Allgemeinen ähnelt die Verteilung der Klicks über die Tageszeit der von Localytics, 10:00 scheint auch bei MailChimp die populärste Zeit zu sein [23]. Zusätzlich zeigt die Analyse, dass sich die Verteilung der Klickrate bei Studierenden im Gegensatz zu Mittvierzigern oder Rentner und Rentnerinnen leicht nach hinten verschiebt (ca. zwei Stunden).

Besonders interessant ist der Einfluss des ausgeübten Berufes. Bleibt die Spitze der Klickzahlkurve bei Anwälten noch bei 9:00-10:00, so verschiebt sie sich bei Kellnern und Kellnerinnen auf 13:00 und bei Neugeborenenpfleger und -pflegerinnen scheint es keine klaren Spitzen zu geben, vermutlich aufgrund des Schichtbetriebes.

Fischer et al. (2011) beobachteten in ihrer Studie, dass Nutzer und Nutzerinnen eine mobile Push-Notification eher akzeptieren, wenn diese direkt nach dem Beenden einer mobilen Interaktion präsentiert wird [22]. Die getesteten Interaktionen waren dabei der Empfang einer SMS oder das Beenden eines Sprach-Anrufs.

3.2 Automatische Erkennung des richtigen Zeitpunktes einer Erinnerung

Da es viele Zeitpunkte im Alltag einer Person geben kann, welche sich für Lernen gut eignen würden, jedoch diese durch ihre Anzahl und Diversität schwer manuell mit Benachrichtigungen verknüpfbar sind, bietet sich die automatische Voraussage des richtigen Zeitpunktes an, z.B. mittels Machine-Learning-Methoden. Mit automatischer Erkennung werden in diesem Absatz immer Einschätzungen eines Algorithmus bezeichnet, welche eventuell anfällig für falsche Vorhersagen sind.

Dingler et al. (2017) entwickelten das System QuickLearn, welches sehr kurze Lern-Lektionen ('Microlearning') in einer M-Learning Umgebung testet [19]. Verschiedene Aspekte wie die Präsentationsform der Lektionen, automatische Erinnerungen, Interaktionsmodalitäten und der Lernkontext wurden dabei genauer untersucht. Der richtige Zeitpunkt für die automatischen Benachrichtigungen wurde mit Hilfe eines Machine Learning Classifiers erkannt. Der Classifier versuchte dabei zu bestimmen, wann die Nutzer und Nutzerinnen gelangweilt sind (mit Hilfe einer Fülle von Sensordaten als Eingabewerte). Dies passiert nur wenn der Bildschirm gerade eingeschaltet wurde und wenn die letzte Benachrichtigung länger als 20Min her ist. Ebenso erhalten Nutzer und Nutzerinnen keine Benachrichtigungen zwischen 23:00 Uhr und 07:00 Uhr. Falls die Nutzer nicht gelangweilt scheinen, wird nur in einem von neun Fällen eine Benachrichtigung gesendet. Die Ergebnisse der Studie haben gezeigt, dass Langeweile nicht mit der Bereitschaft zu Lernen korrelieren zu scheint. Wie die Autoren anmerken, könnten kurze Lerneinheiten nicht die richtige Stimulation sein, um die Langeweile Nutzer und Nutzerinnen zu verringern. Was jedoch beobachtet werden konnte ist, dass Probanden stärker bereit waren Microlearning Einheiten zu absolvieren, wenn sie unterwegs waren. Microlearning Applikationen und automatische Benachrichtigungen für mobile bzw. transitäre Umgebungen zu optimieren, könnte daher äußert erfolgreich sein.

Steil et al. (2018) entwickelten ein System, welches voraussagen konnte, ob die Aufmerksamkeit eines Nutzers bzw. einer Nutzerin bald auf sein / ihr Mobiltelefon gerichtet sein wird, oder nicht [37]. Dabei werden der menschliche Blick ('Gaze Detection'), der visuelle Umgebungskontext ('Object Detection', Face Detection', 'Semantic Scene Segmentation', 'Depth Reconstruction', 'Inertial Sensor Data Head') und die Sensoren und Applikationsdaten des Mobiltelefons als Eingabe für den Vorhersagealgorithmus verwendet. Für die Sammlung der Daten des visuellen Umgebungskontextes wurde ein Blickerfassungsgerät (Eye-Tracker) auf dem Kopf getragen und mit einem Laptop verbunden, welcher sich im Rucksack der Testperson befand. Auf dem Blickerfassungsgerät befand sich zusätzlich eine Stereokamera, welche in die Blickrichtung der Probanden ausgerichtet war und über ein Gyroskop Informationen über die Kopfbewegung sammelte. Die Autoren merken an, dass diese eher sperrige Zusammensetzung der Messgeräte eventuell das Aufmerksamkeitsverhalten der Probanden beeinflusst haben könnte.

Mehrotra et al. (2015) entwickelten ein Machine-Learning-System, welches die Akzeptanz einer Echtzeitbenachrichtigung aufgrund ihres Inhaltes und Informationen über den Kontext, in dem sie geliefert wird, vorhersagen konnte [27]. Eine Benachrichtigung wurde als 'akzeptiert' gewertet, wenn der Nutzer / die Nutzerin auf diese geklickt hat, was wiederum zu einem Start der zugehörigen Applikation führte. Die Vorhersage der Akzeptanz einer Benachrichtigung, kann dabei helfen zu Entscheiden, ob der aktuelle Moment passend ist um diese zu schicken. Falls die vorhergesagte Akzeptanz niedrig ist, könnte davon ausgegangen werde, dass der Zeitpunkt unpassend ist und der Nutzer / die Nutzerin durch die Benachrichtigung gestört werden würde.

3.3 Erzeugung von Gewohnheiten durch Erinnerungen

Gewohnheiten können uns repetitive Aufgaben als selbstverständlich erscheinen und automatisiert ausführen lassen. Dadurch besteht gar kein Bedürfnis mehr, an eine Aufgabe erinnert werden zu müssen. Lernapplikationen können diese Eigenschaft der menschlichen Psyche verwenden, um die Motivation ihrer Nutzer und Nutzerinnen langfristig hoch zu halten.

WaitSuite hatte laut Interviews mit den Probanden einen weiteren Effekt: Die Nutzer und Nutzerinnen haben nach einiger Zeit Gewohnheiten geformt, bei denen das Lernen automatisch mit der Wartezeit assoziiert wurde. Konnte ein Nutzer / eine Nutzerin ihre Lernaktivität nicht wahrnehmen, bestand die Option diese Später nachzuholen, was diese Gewohnheitsbildung zusätzlich unterstützt hat [15].

Stawarz, Cox und Blandford (2014) entwickelten Richtlinien für ein System, welches die Gewohnheitsbildung bei der Einnahme von Arznei unterstützen soll [36]. Sie weisen darauf hin, dass genaue Einhaltung der vorgeschriebenen Einnahme-Intervalle von Medizin essentiell für eine korrekte Behandlung ist. Jedoch führen Vergesslichkeit und unvorhersehbare Tagesabläufe immer noch häufig zur Nichteinnahme des Medikamentes. Herkömmliche mobile Erinnerungsapplikationen für Medikamente funktionieren dabei ähnlich wie Lernerinnerungen: zu fixen Wochentagen und Zeitpunkten. Um dabei zu helfen die Einnahmekontinuität der Patienten zu verbessern, empfehlen die Autoren einen gewohnheitsbildenden Ansatz, bei dem die Medikamenteneinahme unbewusst mit einer wiederkehrenden Gewohnheit assoziert wird, was eine zusätzliche Erinnerung irgendwann überflüssig macht. Für die Implementierung einer mobilen Applikation, welche eine neue Gewohnheit zu erzeugen und diese mit der Medikamenteneinnahme zu verbinden soll, empfehlen die Autoren folgende Kernfunktionen:

- Nutzer und Nutzerinnen muss zuerst dabei geholfen eine neue Gewohnheit anzulegen (z.B. um 8:00 nach dem Zähneputzen)
- Die Nutzer und Nutzerinnen sollen anschließend Sicherheitsbenachrichtigungen anlegen und individualisieren. Diese werden aktiviert, falls die gewohnte Tätigkeit nicht wie üblich stattfindet.
- Um besser den Überblick zu behalten, muss Nutzern und Nutzerinnen die Möglichkeit geboten werden zu überprüfen, ob ein Medikament zuvor, bzw. an einem bestimmten Tag eingenommen wurde.

Rodríguez et al. (2020) entwickelten ein System welches über Verarbeitung der Umgebungsgeräusche die Aktivität zweier Testprobandinnen erkennen konnte und dadurch eine Erinnerung zur Medizineinnahme passend zur vorher festgelegten Aktivität sendete [33]. Die Erinnerung erfolgt in Form eines Sprachassistenten, welcher sein Konversationsmodell auf den Kontext anpasst. Der Sprachassistent fragt ebenso nach, ob das Medikament eingenommen wurde und trägt damit zur Einhaltung der Arzneieinnahme und zur Gewohnheitsbildung bei.

4 Integration intelligenter Erinnerungsmethoden in mobile Lernapplikationen

In diesem Abschnitt werden diverse intelligente Erinnerungsmethoden aus Kapitel 3 in eine mögliche mobile Lernumgebung integriert. Dabei wird zuerst ein Konzept vorgestellt, wie eine solche mobile Applikation funktionieren könnte. Danach wird auf technische Implementierung, deren Herausforderungen und weitere Herausforderungen eingegangen. Die vorgestellten Konzepte und Implementierungsmöglichkeiten entsprechen den Überlegungen des Autors und sollen als Inspiration für zukünftige Forschungsprojekte dienen.

4.1 Mögliche Richtlinen für die Implementierung einer intelligenteren mobilen Lernumgebung

Eine relativ einfache Methode um voreingestellte Zeitpunkte der Push-Nachrichten besser auf die Nutzer und Nutzerinnen abstimmen zu können, wäre eine kurze Frage: 'Was ist dein Beruf?', da hier der durchschnittlich beste Zeitpunkt besser geschätzt werden kann. Um die Menge an Benachrichtigungen schrittweise reduzieren zu können, sollten Gewohnheiten mehr in die Erinnerungsfunktionen integriert werden. Falls eine Erinnerung ignoriert wird, sollte bei einer zweiten Push-Notification auf die Gewohnheit eingegangen werden (z.B. 'Hast du nach dem Frühstück deine Lern-Lektion absolviert?'). Automatische Kontext Erkennung kann eine große Hilfe sein um zu erkennen, ob Gewohnheiten zum angegebenen Zeitpunkt erledigt wurden, oder nicht. Ein Beispiel wäre das Erkennen von Zähneputzen durch Audio-Aufnahme, oder das Erkennen des Aufenthaltes im öffentlichen Verkehr durch GPS-Daten. Unterstützung von Bluetooth-Beacons könnte ebenfalls sehr praktisch sein, um Erinnerungen zu Platzieren (z.B. Person befindet sich neben dem Bett -> automatisches Auslösen von Mini Lern-Lektion). Erinnerungen sollten nur dann sofort ausgelöst werden, wenn sie für den Nutzer auch sehr relevant sind (z.B. wenn sie einen sozialen Aspekt haben), oder eine Gewohnheit erkannt wird. In allen anderen Fällen sollte die Applikation vorher überprüfen, ob der aktuelle Zeitpunkt günstig ist, um eine Nachricht zu senden und diese ggf. verschieben, bis sich ein opportuner Moment ergibt. Wurde die geplante Gewohnheit über längere Zeit nicht erkannt, können allgemeine opportune Momente genutzt werden, um den Nutzer zu erinnern, wie z.B. das Ende einer mobilen Tätigkeit und die damit zusammenhängende kognitive Verfügbarkeit. Zusätzliche Verbesserung um zu erkennen, ob eine Push-Notification tatsächlich wahrgenommen wurde, wäre Aufmerksamkeitserkennung. All die Annahmen, die getroffen werden, können mit längerer Nutzung der Applikation und mit vielen Nutzern immer weiter verbessert werden. Die Hypothese: 'Der beste Zeitpunkt für Anwälte ist um 10:00', kann durch Nutzer-Daten überprüft werden. Durch diese kann man auch erkennen, wie hoch die Klickrate von Nutzern und Nutzerinnen bei der automatischen Erkennung einer Gewohnheit ist. Falls diese zu niedrig ist könnte man überlegen, ob die Gewohnheits-Erkennung noch zu schlecht, oder die Annahme 'Kombination von Gewohnheit und Erinnerung führt zu besserer Klickrate' falsch ist.

4.2 Technische Implementierung und Herausforderungen

Die Integration von Gewohnheiten in Push-Notification Texte ist rein technisch gesehen nichts besonderes. Hier ist eine hohe Abdeckung von Gewohnheiten und eine Individualisierung der Texte auf die verschiedenen Gewohnheiten eine

Herausforderung. Auch könnte es schwierig sein zu wissen wann ein Beruf am besten die voreingestellte Benachrichtigung bekommt, ohne davor viel Daten gesammelt zu haben. Eine naive Annahme am Anfang (wann beginnst du zu arbeiten, wann hörst du auf -> schicke die Nachricht am Ende des ersten Drittel der Arbeitszeit) plus anschließendes Data-Mining nach mehreren Monaten wäre eine Möglichkeit um hier bessere Ergebnisse zu bekommen. Bluetooth-Beacons lassen sich vergleichsweise leicht in eine Applikation Integrieren, da es vorgefertigte Bibliotheken gibt [32]. Was jedoch herausfordernd sein könnte, ist die automatische Erkennung von Gewohnheiten. Machine-Learning-Classifier lieferten in den Studien teilweise noch falsche Ergebnisse und müssten in einer realen Anwendung kontinuierlich verbessert werden [33]. Ähnlich ist es bei der Erkennung von opportunen Momenten für Echtzeitbenachrichtigungen [27]. Für die Erkennung der Umgebung und der Aktivitäten der Nutzer und Nutzerinnen gibt es bereits APIs, welche den Entwicklungsaufwand verringern können (z.B. Google Activity Recognition API [18]).

4.3 Weitere Herausforderungen

Eine Applikation welche alle oben genannten Fähigkeiten integriert, wäre aufwändig in der Entwicklung und der darauf folgenden Analyse der zahlreichen neuartigen Komponenten. Hier wäre es eventuell sinnvoller die Komponenten schrittweise zu entwickeln und auf ihre Effektivität zu testen.

Die Eingabe zahlreicher persönlicher Informationen beim Start der Lern-Applikation kann auf die Nutzer und Nutzerinnen abschreckend wirken. Hinzu kommt, das manche Menschen ihre persönlichen Informationen nicht teilen, oder einer Lern-Applikation die Berechtigung zur Aufnahme von Audio oder dem Sammeln von Standort-Daten nicht geben möchten. Die Verarbeitung dieser Daten ist außerdem aus Datenschutz-Aspekten als sensibel zu betrachten. Nutzer und Nutzerinnen müssen immer völlig aufgeklärt werden, wozu ihre Daten gespeichert werden und was damit passiert. Die Verarbeitung sollte wenn möglich ausschließlich am Gerät erfolgen und nicht an einen Server gesendet werden. Um die nötige Transparenz zu bieten könnte es auch hilfreich sein, den Quellcode der Applikation öffentlich zugänglich zu machen.

5 Fazit und Ausblick

Intelligentere Lernerinnerungen können dabei helfen die oft als störend empfundenen, herkömmlichen starren Echtzeitbenachrichtigungen zu verbessern. Relativ einfach zu implementierende Ansätze wie die Individualisierung der vorgeschlagenen Benachrichtigungs-Zeit auf die Berufe der Nutzer, könnten bereits dabei helfen die Nutzer und Nutzerinnen zum Lernen zu erinnern, wenn diese auch tatsächlich verfügbar sind. Darüber hinaus bieten Ansätze der

Gewohnheitsbildung eine Möglichkeit um die Notwendigkeit von Lernerinnerungen zu minimieren, was wiederum in weniger Unterbrechung für den Nutzer / die Nutzerin resultieren würde. Automatische Erkennung von Gewohnheiten könnte ebenfalls genutzt werden, um den Nutzer bzw. die Nutzerin nur in einem passenden Zeitpunkt an das Lernen zu erinnern. Andere Möglichkeiten dafür sind die eigenständige manuelle Platzierung von Erinnerungshelfern (z.B. mittels Bluetooth-Beacons), oder automatische Erkennung von opportunen Momenten für Push-Notifications. Ein Aspekt welcher in der Literatur leider (nach aktuellem Wissensstand des Autors) gar nicht erforscht ist, ist die Effektivität von Lernerinnerungen mit Gamification-Charakter. Ebenso wurden einige der vorgestellten Konzepte erst mit kleinen Probanden- und Propandinnen-Gruppen getestet, was fortführende Studien interessant machen könnte. Wenn man den aktuellen Stand der betrachteten Lern-Applikationen mit den zahlreichen vorgestellten, intelligenteren Konzepten vergleicht, so erkennt man dass das Potential von Lern-Applikationen definitiv noch nicht ausgeschöpft ist. Dies gilt sowohl für der Forschung, als auch für den kommerziellen Bereich der mobilen Lern-Applikationen.

Literatur

1 Babbel. Accessed January 14, 2021 from `https://play.google.com/store/apps/details?id=com.babbel.mobile.android.en`.
2 Coursera. Accessed January 14, 2021 from `https://play.google.com/store/apps/details?id=org.coursera.android`.
3 Duolingo. Accessed January 14, 2021 from `https://play.google.com/store/apps/details?id=com.duolingo`.
4 edX. Accessed January 14, 2021 from `https://play.google.com/store/apps/details?id=org.edx.mobile`.
5 Lengalia. Accessed January 12, 2021 from `https://www.lengalia.com/en/spanish-for-universities-schools.html`.
6 LinkedIn Learning. Accessed January 14, 2021 from `https://play.google.com/store/apps/details?id=com.linkedin.android.learning`.
7 Mondly. Accessed January 14, 2021 from `https://play.google.com/store/apps/details?id=com.atistudios.mondly.languages`.
8 Mumie. Accessed January 12, 2021 from `https://www.mumie.net/`.
9 Rosetta Stone. Accessed January 14, 2021 from `https://play.google.com/store/apps/details?id=air.com.rosettastone.mobile.CoursePlayer`.
10 Udemy. Accessed January 14, 2021 from `https://play.google.com/store/apps/details?id=com.udemy.android`.
11 Piotr D. Adamczyk and Brian P. Bailey. If not now, when?: The effects of interruption at different moments within task execution. *Conference on Human Factors in Computing Systems - Proceedings*, pages 271–278, 2004.

12 Brian P. Bailey, Joseph A. Konstan, and John V. Carlis. Measuring the effects of interruptions on task performance in the user interface. *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, 2:757–762, 2000. doi:10.1109/ICSMC.2000.885940.

13 Huseyin Bicen and Senay Kocakoyun. Perceptions of students for gamification approach: Kahoot as a case study. *International Journal of Emerging Technologies in Learning (iJET)*, 13(02):72–93, 2018. URL: https://online-journals.org/index.php/i-jet/article/view/7467.

14 BVDW. Digitale Nutzung in Deutschland 2018. Technical report, 2018. URL: https://www.bvdw.org/themen/publikationen/detail/artikel/digitale-nutzung-in-deutschland-2018/.

15 Carrie J. Cai, Anji Ren, and Robert C. Miller. Waitsuite: Productive use of diverse waiting moments. *ACM Trans. Comput.-Hum. Interact.*, 24(1), March 2017. doi:10.1145/3044534.

16 Edward Cutrell, Mary Czerwinski, and Eric Horvitz. Notification, Disruption, and Memory: Effects of Messaging Interruptions on Memory and Performance. *Conference on Human-Computer Interaction Interact 2001*, (1999):263–269, 2001.

17 Deutsche Welle. Corona-Pandemie: Herausforderung digitales Lernen, 2020. Accessed January 12, 2021 from https://www.dw.com/de/corona-pandemie-herausforderung-digitales-lernen/a-56172231.

18 Google Developers. Google Activity Recognition API. Accessed January 21, 2021 from https://developers.google.com/location-context/activity-recognition.

19 Tilman Dingler, Dominik Weber, Martin Pielot, Jennifer Cooper, Chung-Cheng Chang, and Niels Henze. Language learning on-the-go: Opportune moments and design of mobile microlearning sessions. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*, MobileHCI '17, New York, NY, USA, 2017. Association for Computing Machinery. doi:10.1145/3098279.3098565.

20 Edtechhistory.org.uk. THE 1970S – A DECADE OF DEVELOPMENT. URL: http://www.edtechhistory.org.uk/2020/06/02/the-1970s-a-decade-of-development/.

21 Jun Fan and Zhen Wang. The impact of gamified interaction on mobile learning app users' learning performance: the moderating effect of users' learning style. *Behaviour & Information Technology*, 0(0):1–14, 2020. arXiv:https://doi.org/10.1080/0144929X.2020.1787516, doi:10.1080/0144929X.2020.1787516.

22 Joel E. Fischer, Chris Greenhalgh, and Steve Benford. Investigating episodes of mobile phone activity as indicators of opportune moments to deliver notifications. *Mobile HCI 2011 - 13th International Conference on Human-Computer Interaction with Mobile Devices and Services*, pages 181–190, 2011. doi:10.1145/2037373.2037402.

23 John Foreman. mailchimp, 2014. Accessed January 05, 2021 from https://mailchimp.com/resources/insights-from-mailchimps-send-time-optimization-system/.

24 Michael Gaebel, Veronika Kupriyanova, Rita Morais, and Elizabeth Colucci. E-Learning in European Higher Education Institutions. (December):1–92, 2014. URL: http://www.eua.be/Libraries/publication/e-learning_survey.pdf?sfvrsn=2.

25 Thomas Hess, Clemens Koob, Florian Mann, and Stephanie Sendobry. Studie LIFE - Unser Leben vernetzt und digital. *Munich School of Management Magazine - Das Magazin der Fakultät für Betriebswirtschaft der Ludwig-Maximilians-Universität München (10:2009)*, pages 6–9, 2009. URL: https://www.wim.bwl.uni-muenchen.de/download/epub/studie_life.pdf.

26 Eckhart Kauntz. Bei der dritten falschen Antwort sagt der Computer vor, 1987. URL: http://www.rissberger.de/pages/faz_87-06-25.html.

27 Abhinav Mehrotra, Mirco Musolesi, Robert Hendley, and Veljko Pejovic. Designing content-driven intelligent notification mechanisms for mobile applications. *UbiComp 2015 - Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 813–824, 2015. doi:10.1145/2750858.2807544.

28 Abhinav Mehrotra, Veljko Pejovic, Jo Vermeulen, Robert Hendley, and Mirco Musolesi. My phone and me: Understanding people's receptivity to mobile notifications. *Conference on Human Factors in Computing Systems - Proceedings*, pages 1021–1032, 2016. doi:10.1145/2858036.2858566.

29 Bastian Pauly and Nina Brandau. Fast jeder Zweite bildet sich online weiter, 2018. Accessed January 12, 2021 from https://www.bitkom.org/Presse/Presseinformation/Fast-jeder-Zweite-bildet-sich-online-weiter.html.

30 Dominik Petko. *Lernplattformen in Schulen*. VS Verlag, 1 edition, 2010. URL: http://library1.nida.ac.th/termpaper6/sd/2554/19755.pdf.

31 X. Pham, T. Nguyen, W. Hwang, and G. Chen. Effects of push notifications on learner engagement in a mobile learning app. In *2016 IEEE 16th International Conference on Advanced Learning Technologies (ICALT)*, pages 90–94, 2016. doi:10.1109/ICALT.2016.50.

32 Radius Network. Android Beacon Library, 2016. Accessed January 21, 2021 from http://altbeacon.github.io/android-beacon-library/configure.html.

33 Marcela D. Rodríguez, Jessica Beltrán, Maribel Valenzuela-Beltrán, Dagoberto Cruz-Sandoval, and Jesus Favela. Assisting older adults with medication reminders through an audio-based activity recognition system. *Personal and Ubiquitous Computing*, 2020. doi:10.1007/s00779-020-01420-4.

34 Alireza Sahami Shirazi, Niels Henze, Tilman Dingler, Martin Pielot, Dominik Weber, and Albrecht Schmidt. Large-scale assessment of mobile notifications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing*

Systems, CHI '14, page 3055–3064, New York, NY, USA, 2014. Association for Computing Machinery. doi:10.1145/2556288.2557189.

35 Statistisches Bundesamt. Digitales Lernen nimmt stark zu - Statistisches Bundesamt, 2020. Accessed January 12, 2021 from https://www.destatis.de/DE/Presse/Pressemitteilungen/2020/12/PD20_N081_63.html.

36 Katarzyna Stawarz, Anna L. Cox, and Ann Blandford. Don't forget your pill! designing effective medication reminder apps that support users' daily routines. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, page 2269–2278, New York, NY, USA, 2014. Association for Computing Machinery. doi:10.1145/2556288.2557079.

37 Julian Steil, Yusuke Sugano, Philipp Müller, and Andreas Bulling. Forecasting user attention during everyday mobile interactions using device-integrated and wearable sensors. *arXiv*, pages 1–13, 2018.

38 C-H. Su and C-H. Cheng. A mobile gamification learning system for improving the learning motivation and achievements. *Journal of Computer Assisted Learning*, 31(3):268–286, 2015. doi:10.1111/jcal.12088.

39 Patrick Suppes, Max Jerman, and Guy Groen. Arithmetic drills and review on a computer-based teletype. In *The Arithmetic Teacher*, volume 13, pages 303–309. 2020. URL: https://web.archive.org/web/20160305160358/http://suppes-corpus.stanford.edu/articles/comped/54-5.pdf, doi:10.5951/at.13.4.0303.

40 Josh Todd. localytics, 2016. Accessed January 05, 2021 from http://marketingland.com/5-must-push-notification-best-practices-mobile-marketers-172981.

41 Eliza Trapp. Digitaler Unterricht – was hat sich für Lehrende durch Corona verändert?, 2020. Accessed January 12, 2021 from https://www.bpb.de/lernen/digitale-bildung/werkstatt/317021/digitaler-unterricht-was-hat-sich-fuer-lehrende-durch-corona-veraendert.

Assistenztechnologie in persönlichen Bereichen für Menschen mit Sehbehinderung

Yu Sun

Ludwig-Maximilians-Universität München, München, Deutschland
Yu.Sun@campus.lmu.de

Zusammenfassung

Technologie hat seit jeher den Alltag der Menschen erleichtert. Heutzutage gibt es eine gewisse Menge von Technologien, welche das Potential haben, mittels gutem Design und spezifischer Anpassungen, das Leben beeinträchtigter Menschen zu verbessern. In dieser Arbeit wird untersucht, wie Assistenztechnologie Menschen mit Sehbehinderung im Alltagsleben helfen kann. Dafür wird zunächst erklärt, wie die Informationsübermittlung und Interaktion zwischen Menschen mit Sehbehinderung und Maschinen durch übliche Hilfsmittel und neuartige Technologien möglich ist, und wie sie verschiedene Arten von Feedback wie z.B. visuelles Feedback, akustisches Feedback und haptisches Feedback anbieten könnten. Danach wird erläutert, welche Technologien für Menschen mit Behinderung, vor allem mit Sehbehinderung, bereits in der Anwendung sind oder Anwendungsmöglichkeiten haben, und in welchen Szenarien sie angewandt werden sollten. Anhand ausgewählter Studien werden daraufhin konkrete Beispiele für das Design und die Realisierung der assistiven Systeme für Menschen mit Sehbehinderung beleuchtet.

2012 ACM Computing Classification CCS → Human-centered computing → Accessibility → Accessibility systems and tools

Keywords and phrases Assistive Technologie; Sehbehinderung; Barrierefreiheit; Haptisches Feedback; Sprachsteuerung.

1 Einführung

Laut des im Oktober 2019 veröffentlichten World Report on Vision der Weltgesundheitsorganisation (WHO) leiden weltweit mindestens 2,2 Milliarden Menschen an Sehbehinderung oder Blindheit [19]. Die Technologien, die wir gebaut haben, sind jedoch nicht zugänglich für alle Benutzergruppen, z.B. die Popularität von touchscreen-basierten Geräten ist ein riesiges Problem für sehbeeinträchtigte Nutzer. „Technologie ist am besten, wenn sie allen die

Cite as: Yu Sun. Assistenztechnologie in persönlichen Bereichen für Menschen mit Sehbehinderung. In *6th Seminar on Ubiquitous Interaction (UBIACTION 2021)*. Editors: F. Lang, R. Welsch, L. Haliburton, F. Draxler, S. Feger, S. Villa, M. Hoppe, P. Knierim, V. Mäkelä, A. Schmidt. February 11, 2021. Munich, Germany. pp. 4:1–4:14.

gleichen Möglichkeiten gibt [11].“ Assistive Technologie, welche nicht nur die Technologie selbst zugänglich macht, sondern auch in mehr Situationen ein unabhängigeres Leben von beeinträchtigten Menschen ermöglicht, ist deswegen wichtig im Bereich Menschen-Computer-Interaktion. Aus dem angloamerikanischen Sprachraum übernommen bzw. übersetzt (von assistive technology) werden hierfür auch die Begriffe assistierende Technologie oder Unterstützungstechnologie verwendet.

Assistenztechnologie befasst sich primär mit Hilfsmitteln für Menschen mit Behinderungen [18]. Hiermit werden nicht nur speziell angefertigte oder angepasste Hilfsmittel bezeichnet wie zum Beispiel Rollstühle, Lifte oder andere Alltagshilfen, sondern auch Möglichkeiten der Anpassung (engl.: adaptability) und Zugänglichkeit/Barrierefreiheit (engl.: accessibility) von herkömmlichen Produkten mit universellem Design speziell für Menschen mit Behinderungen. Das Ziel der Assistenztechnologie ist es, die Kluft zwischen dem, was ein beeinträchtigter Mensch tun möchte und dem, was die vorhandene Infrastruktur ihm bzw. ihr ermöglicht, zu überbrücken. Sie besteht aus Ausstattungen, Geräten und Systemen, die Hindernisse von beeinträchtigten Menschen überwinden und ihre uneingeschränkte und gleichberechtigte Beteiligung an allen Aspekten der Gesellschaft unterstützen können.

In dieser Arbeit liegt der Fokus auf persönlichen Bereichen von Menschen mit Sehbehinderung. Das Alltagsleben in persönlichen Bereichen umfasst eine Vielzahl verschiedener Aktivitäten. Die meisten von ihnen finden in einer häuslichen Umgebung statt, aber einige Aktivitäten wie z.B. die Benutzung von Geldautomaten werden in dieser Arbeit auch erwähnt.

2 Herausforderungen im Alltagsleben von Menschen mit Sehbehinderung

Die Herausforderungen, die Menschen mit Sehbehinderung haben, sind nicht gut verstanden. Brady et al. [2] haben eine Studie durchgeführt, bei der über 5.000 Nutzer mit Sehbehinderung ungefähr 40.000 Fragen über eine Applikation gestellt haben. Die Nutzer konnten zuerst die alltäglichen Gegenstände, über die sie Fragen hatten, photographieren. Anschließend konnten sie die Fotos und die Fragen auf dieser Plattform posten. Übliche Fragen waren z.B. welches Medikament ist zu sehen? Welche Farbe hat der Gegenstand usw. Das Ergebnis dieser Studie kann dabei helfen, ein besseres Verständnis von den Problemen, mit denen die sehbehinderten Menschen konfrontiert sind, zu erlangen. Brady et al. haben eine Taxonomie der Fragentypen erstellt und es gibt folgende Kategorien: Identifikation, Beschreibung und Lesen. Folgendes ist die Auflistung der Taxonomie, wobei für jede Kategorie repräsentative Beispiele zur Veranschaulichung angegeben sind:

- Identifikation
  - Kein Kontext: Was ist das?
  - Kontextuell: Könnten Sie mir bitte sagen, ob dies ein Diät-Pepsi oder ein normaler Pepsi ist?
  - Medizin: Was sind diese Pillen?
- Beschreibung
  - Zustand (ein/aus): Ist ein Licht an?
  - Computer/TV Bildschirm: Was steht auf diesem Bildschirm?
- Lesen
  - Digitales Display: Könnten Sie mir bitte sagen, welche Texte/Nummer auf diesem Ofen/Wecker steht?
  - Kochen: Könnten Sie mir bitte die Kochanweisungen für eine Mahlzeit vorlesen?
  - Badezimmer: Könnten Sie mir die Einzelheiten für dieses Hygieneprodukt geben?

Durch diese Studie wird ein besseres Verständnis für die Herausforderungen gewonnen, mit denen sehbeeinträchtigte Menschen im täglichen Leben konfrontiert sind. Es ist bemerkenswert dass die Fragen zu Essen und persönliche Pflege relativ häufig gestellt wurden und die Leute Schwierigkeiten hatten, in diesen Bereichen unabhängig zu sein.

3 Arten von Feedback für Menschen mit Sehbehinderung und assistive Technologie

Bei vielen alltäglichen Aufgaben, wie z.B. die genaue Position eines Stuhls finden und darauf sitzen, den auf dem Boden verlorenen Gegenstand aufheben, sich die Zahnbürste im Badzimmer holen etc., müssen sich Menschen mit schwerem Sehverlust auf andere Sinne verlassen, hauptsächlich auf Haptik und Hören. Bei manchen Situationen können sie sich auch auf ihren visuellen Sinn verlassen, wenn das Restsehvermögen mittels visueller Hilfsmittel bei dem Erledigen der Aufgaben ausreichend ist.

3.1 Visuelles Feedback

Gängiges Hilfsmittel: Bildschirmlupe
Weltweit gibt es 1,2 Milliarden Menschen mit einer Sehbehinderung, die mit Brillen oder Kontaktlinsen nicht behoben werden kann [19]. Jedoch haben viele dieser Menschen eine verbleibende Sehfähigkeit und könnten visuelle Hilfsmittel verwenden. Ein gängiges Werkzeug ist die Bildschirmlupe, die eine Vergrößerung der Bildschirminhalte bewirkt und dadurch ein verbessertes visuelles Feedback anbieten kann.

Technologien für neuartiges visuelles Feedback: Projektion und AR
In gefährlichen Szenarien wie z.B. Navigation in Räumen oder auf Treppen können Menschen mit schlechter Sicht zwar sehen, aber stehen häufiger in Gefahr oder können sich nicht schnell fortbewegen. Mithilfe von neuartigem visuellen Feedback z.B. Projektion und AR(Augmented Reality) können sich die Menschen schneller fortbewegen und ein sicheres Gefühl bekommen. Zhao et al. [21] haben Visualisierungen für eine projektionsbasierte AR Plattform und eine intelligente Brille entworfen. Visuelle Highlights können direkt auf die Treppe projiziert werden. Die Visualisierung zeigt die Position des Benutzers, ohne dafür die Treppe mit zusätzlichen Geräten erweitern zu müssen. Zur Evaluation wurden 12 Teilnehmer getestet, mit dem Ergebnis dass die projektionsbasierte AR die Schrittgeschwindigkeit und auch die von den Teilnehmern selbst gemeldete psychologische Sicherheit erhöht hat.

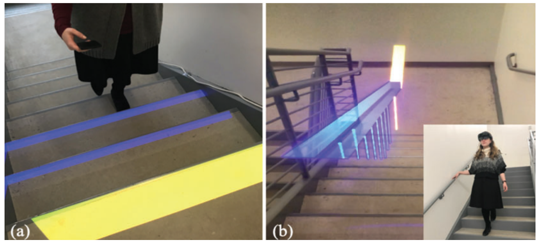

Abbildung 1 (a) zeigt die Visualisierungen für eine projektionsbasierte AR Plattform, die bei der Navigation auf Treppen hilft. (b) zeigt eine intelligente Brille zur Erleichterung der Treppennavigation [21].

3.2 Haptisches Feedback

Auf Haptik basierte Hilfsmittel sind wichtig für Menschen mit Sehbehinderung. Herkömmliche Hilfsmittel wie z.B. der Blindenstock, der Blindenhund und die Blindenschrift sind effizient. Sie können ansonsten fehlende Informationen zur Verfügung stellen und die natürlichen Fähigkeiten der menschlichen Sinne erweitern.

Brailleschriftdarstellung
In einem Projekt des Massachusetts Institute of Technology wurde eine touchscreen-basierte Steuerung eines Ofens mithilfe von Blindenschrift für sehbeeinträchtigte Menschen zugänglich gemacht [11]. Die Idee ist es, dass die Blindenschrift auf eine Metallplatte geschmolzen wird, welche mit Ausschnitten an der Stelle der Touchelemente als Rahmen für das Touchdisplay

verwendet wird. Die sehbeeinträchtigten Menschen können anhand der Braille-Beschriftungen neben den Ausschnitten identifizieren, welcher Touchbutton darunter ist. Es bietet eine Gelegenheit, dass Menschen mit Sehbehinderung den Ofen mit touchbasierter Darstellung verwenden und zum Haushalt beitragen können.

Abbildung 2 Ein Ofen Display, welches mittels Blindenschrift die visuellen Informationen zu haptischen Informationen konvertiert hat [11].

Die auf dem gleichen Konzept (Beschriftungs- und Overlayssystem) basierten Designs sind vielfältig. Ein Beispiel wäre ein taktiler Mikrowellenaufkleber, der durch verschiedene Formen die Funktionen der Touchbuttons darstellt. Ein Aufkleber in der Form von einem seitlichen Dreieck repräsentiert den Startknopf und ein X-Aufkleber dient zum Abbrechen usw. Nach dem Aufkleben an der Stelle der Touchelemente ist diese Haushaltsmaschine für sehbehinderte Menschen wieder zugänglich.

Aktualisierbares Brailledisplay
Außderdem existieren aktualisierbare Brailledisplays, welche ein Computer-Ausgabegerät für blinde Menschen sind und Zeichen in Brailleschrift darstellen. Üblicherweise werden sie durch Screenreader angesteuert, die Zeichen in ausgewählten Bildschirmbereichen auslesen und in Computerbraille darstellen. An der Braillezeile sind Steuertasten angebracht, mit denen der dargestellte Bildschirmausschnitt verschoben werden kann. Alternativ können Screenreader auch eine Sprachausgabe bieten. Gegenüber dem Vorlesen sind Braillezeilen genauer und geben Wort für Wort wieder. Somit kann auch die Rechtschreibung direkt überprüft werden, ohne dass die Sprachausgabe die Wörter buchstabieren muss. Zahlreiche elektronische Geräte können Brailledisplay unterstützen, z.B. das iPhone von Apple kann durch Bluetooth mit einem kabellosen Brailedisplay angeschlossen werden. Wenn man Text bearbeitet, zeigt das Brailledisplay den Text im Kontext an und die Änderungen werden automatisch von Brailletext

in gedruckten Text konvertiert. Ein Brailledisplay mit Eingabetasten kann man auch verwenden, um ein iPhone zu steuern, wenn die iPhone-Funktion VoiceOver aktiviert ist.

Haptisches Feedback durch Vibration

Vibrotaktiles Feedback (wird auch als Vibrationsfeedback bezeichnet) ist eine Art der Kommunikation. Haptik bietet Vorteile für die Zugänglichkeit von vielen Geräten für Menschen mit Sehbehinderung, z.B. der Geldautomat. Viele Geldautomaten wurden zwar mit einem Audio-System ausgestattet, bei denen ein Benutzer einen Kopfhörer an den Geldautomat anschließen kann, aber dies kann die Wahrnehmung der Umgebung reduzieren, weshalb die Nutzer für eine Attacke bei der Abhebung des Geldes anfällig sind. Ein vibrierendes Keyboard wäre eine gute Lösung dafür. Cassidy et al. [3] haben eine haptische Geldautomat-Schnittstelle für Menschen mit Sehbehinderung entworfen. Kleine Vibrationsmotoren wurden an der Unterseite jeder Tastenkappe von Nummer 1 bis 6 angebracht. Die Vibrationsmotoren konnten dann in einer geeigneten Reihenfolge vibrieren und die Nutzer wurden geführt, ihre Hände an einer bestimmten Stelle am Geldautomat zu bewegen.

3.3 Akustisches Feedback

Laut Fritz et al. [6] ist die Informationsbandbreite von dem visuellen Kanal, dem Hörkanal und dem taktilen Kanal jeweils $10^6, 10^4$ und $10^1 - 10^2$ Bits pro Sekunde. Man kann deswegen davon ausgehen, dass die Informationen, die vom akustischen Kanal wahrgenommen werden, wesentlich mehr im Vergleich zu denen von dem taktilen Kanal sind. Eine gängige Technologie, die von 82.5% der Menschen mit Sehbehinderung bei der Nutzung des Smartphones verwendet wurde, ist der Bildschirmleser [20]. Daneben gibt es immer mehr neuartige Technologien aufgrund von Fortschritten in der Verarbeitung natürlicher Sprache (NLP) und der Bildverarbeitung unter Verwendung neuronaler Netze. Wenn sie zusammengebracht werden und in einem assistiven System integriert werden könnten, wäre es effizient und hilfreich für Menschen mit Sehbehinderung.

Forscher in Stanford [14] haben ein sprechendes auf künstliche Intelligenz (KI) basiertes System für Kaffeemaschinen eingeführt, bei dem Sprachsynthese und Bildverarbeitungs-Technologie kombiniert werden. Ein Nutzer setzt zuerst ein Headset mit Kamera auf, dann verarbeitet dieses Gerät die durch die Aufnahme erhaltenen Live Informationen und gibt dem Nutzer bei der Bedienung des Touchscreens Anweisungen. Der Nutzer kann anhand der akustischen Kommandos den Finger nach oben oder nach unten bewegen und dann den richtigen Button drücken.

Abbildung 3 Nutzer bestellt Kaffee auf Touchscreen mithilfe von Sprachanleitung und Bilderkennung [14].

Ein solches System ist nicht nur hilfreich bei der Bedienung des Touchscreens. Blindtool [4] ist eine auf tiefem Lernen (deep learning) basierte mobile Applikation, die Blinden einen „Sinn für Vision“ gibt. Die App sagt dem Nutzer welcher Gegenstand erkannt wurde und wie gut die Erkennung funktioniert hat. Das Faltungs-Neuronale Netzwerk in der App kann 1000 Klassen verstehen.

4 Interaktionsmöglichkeiten

Im vorherigen Kapitel wurden verschiedene Arten von Feedback vorgestellt, und zwar visuelles Feedback, akustisches Feedback und haptisches Feedback. Durch Feedback auf verschiedene Sinne oder Verstärkung des vorhandenen Vermögens wird es möglich, dass Menschen mit Sehbehinderung gleichwertige Informationen gewinnen könnten. Aber Feedback ist unidirektional und um Kommunikation zwischen Menschen und Maschinen zu ermöglichen, muss man den Maschinen Input geben und mit ihnen interagieren können. Eine Benutzerschnittstelle (englisch User Interface) ist die Stelle oder Handlung, mit der ein Mensch mit einer Maschine oder einem Arbeitsgerät in Interaktion tritt [16]. Im einfachsten Fall ist das ein Lichtschalter: Er gehört weder zum Menschen, noch zur „Maschine“ (Lampe), sondern ist die Schnittstelle zwischen beiden. In diesem Kapitel liegt der Fokus auf Touch-Benutzerschnittstelle (Touch user interface), Gestenbasierte Interaktion, und Sprachbenutzerschnittstelle (voice user interface).

4.1 Touch-Benutzerschnittstelle

Eine Touch-Benutzerschnittstelle (TUI) ist eine Technologie, die auf dem Tastsinn (Haptik) basiert [17]. Während eine grafische Benutzerschnittstelle auf dem Sehsinn beruht, ermöglicht eine TUI dem Benutzer, besonders solchen mit Sehbehinderung, eine auf taktiler oder Braille-Eingabe basierte Interaktion.

Die Fortschritte in der Touchscreen-Technologie haben die Verbreitung von Touchscreens erhöht, aber Touchscreens sind für Benutzer mit Sehbehinderung weitgehend unzugänglich. Ein Multi-Touch-Screen (Bildschirm mit Mehrfingergestenerkennung) ist eine besondere berührungsempfindliche Oberfläche für die Eingabe von Daten. Kane et al. [8] haben Slide Rule eingeführt, eine Multi-Touch-Interaktionstechnik, mit der sehbeeinträchtigte Benutzer mit Touchscreen-Anwendungen interagieren können. Anstatt beispielsweise in der Mail-Anwendung auf eine Schaltfläche zum Weiterleiten zu tippen, leiten Benutzer eine Nachricht weiter, indem sie mit dem Finger eine Rechtsbewegung ausführen.

4.2 Gestenbasierte Interaktion

Menschen mit Sehbehinderung können im Alltagsleben in einschränkende Situationen geraten. Wenn z.B. beide Hände bereits für eine Tätigkeit benötigt werden, ist es schwierig gleichzeitig auf Geräte wie Smartphones zuzugreifen. Patil et al. [13] haben einen neuartigen Interaktionsmodus an einem Blindenstock erforscht und GesturePod, ein einfach zu integrierendes Gerät, entwickelt, welches an jedem Stock befestigt werden kann und mit ihm ausgeführte Gesten erkennen kann.

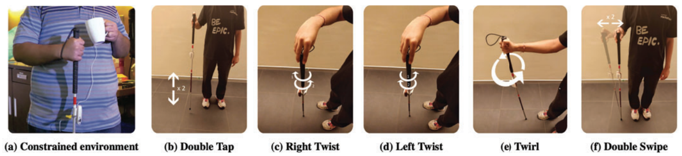

Abbildung 4 (a) zeigt dass Menschen mit Sehbehinderung in eingeschränkten Situationen sind (beide Hände werden bereits für eine Tätigkeit benötigt). (b)-(f) stellen verschiedener Gesten dar, die in dieser Studie verwendet wurden [13].

Mit GesturePod kann ein Benutzer allgemeine Aufgaben auf seinem Smartphone ausführen, ohne sie zu berühren oder das Telefon aus der Tasche zu ziehen. Eine neuartige, effiziente Pipeline für maschinelles Lernen wird verwendet, um das Gestenerkennungsmodell zu trainieren und einzusetzen. Die zur Verfügung stehenden Gesten sind Doppeltippen, Rechtsdrehung, Linksdrehung, doppeltes Kippen und drehen.

4.3 Sprachinteraktion

Ein Voice User Interface (VUI) ist eine Technologie, mit der Benutzer ihre Stimmen für die Interaktion mit Computern und Geräten verwenden können [1]. Es ist integrierbar in vielen verschiedenen Systemen und Anwendungsszenarien und funktioniert gut in Kombination mit anderen Technologien. Es ist erweiterbar auf mobile Geräte, Wearables, AR und VR(Virtual Reality), und ebnet den Weg für konversationelle und multimodale Interaktionen.

Bei vielen technischen Geräten wird Sprachsteuerung für Menschen mit beeinträchtigtem Sehvermögen eingesetzt. Ein Beispiel wären mobile Geräte von Apple. Bei ihnen gibt es alle möglichen Funktionen mit Sprachinteraktion:

- Siri, ein Sprachassistent, der dem Nutzer bei vielen alltäglichen Dingen helfen kann.
- Diktierfunktion, mit der man nicht mehr tippen muss.
- Audiobeschreibungen, mit der man Filme mit detaillierter Audiobeschreibung von jeder Szene erleben kann.
- Gesprochene Inhalte, mit denen man sich fast jeden Text vorlesen lassen kann.
- VoiceOver, mit dem man hören kann, was auf dem Bildschirm passiert.

Sprachinteraktion kann auch in Wearables integriert werden. Kim et al. [10] haben ein System entwickelt, mit dem Nutzer durch ihre Stimme interagieren können. In diesem System wird eine Kamera auf Brusthöhe positioniert, um Bilder aufzunehmen. Ein Raspberry Pi 3 Model B+ wird auf eine 3D gedruckte Halterung montiert und wird an der Taille des Nutzers angebracht. Dieses System fragt den Nutzer, nach welchem Objekt er sucht und erkennt die Antwort des Nutzers, um den Namen des Objekts zu wissen. Befindet sich das Objekt in Reichweite, fährt das System fort und beschreibt die genaue Position des Objekts, z.B. „Es befindet sich rechts vor Ihnen.“ Andernfalls teilt das System dem Nutzer mit, dass kein Objekt gefunden wurde und fordert ihn auf, sich umzudrehen oder den Namen eines anderen Objekts zu sagen.

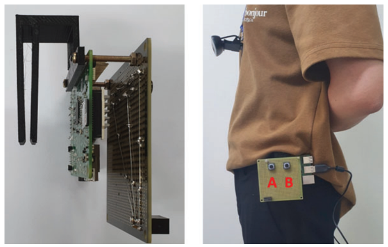

Abbildung 5 (a) zeigt eine gedruckte Halterung. (b) zeigt freihändige Geräte [10].

5 Anwendungsszenarien und Beispiele

In diesem Kapitel werden Anwendungsszenarien und konkrete Beispiele im Kontext des Alltagslebens beleuchtet. Das tägliche Leben umfasst alle verschiedenen Aktivitäten, die möglicherweise jeden Tag oder sogar mehrmals am Tag durchgeführt werden. Diese Aktivitäten können wie folgt weiter klassifiziert werden:

- Persönliche Pflege
- Essen Zubereitung und Verzehr
- Steuerung der Haushaltsgeräte
- Zeitmessung und Alarmierung
- Geld, Finanzen und Einkaufen
- Sexuelle und reproduktive Aktivitäten

5.1 Essen Zubereitung und Benutzung der Küche

In der Küche wird assistive Technologie häufig verwendet, z.B. sprechender Timer, vibrierender Timer, sprechende Küchenwaage, sprechender Thermometer, sprechender Ofen, elektrische Bratpfanne mit taktilen Bedienelementen, Flüssigkeitsstandsanzeige mit akustischem Feedback usw. Sie sind praktisch und sind schon am Markt verfügbar. Eine Flüssigkeitsstandsanzeige kann verhindern, dass ein Nutzer eine Tasse überfüllt, weil die Wasserlinie nicht mehr sichtbar ist.

Ein ferngesteuerte Fleischthermometer mit Grillalarm verfügt über eine Sonde, die in die Mitte des Fleisches eingesetzt ist, und einen drahtlosen Sender, der Temperaturmesswerte an den Empfänger sendet, der sich in der Entfernung befindet. Eine Spracherinnerung warnt den Benutzer, wenn das Fleisch "fast fertig" und "fertig" ist. Das Gerät kann im Ofen, auf einer Kochplatte oder auf einem Außengrill verwendet werden.

5.2 Steuerung der Haushaltsgeräte

Die Steuerung der Haushaltsgeräte umfasst die Verwendung von Waschmaschinen und Staubsaugern, das Ein- und Ausschalten von Lichtern, das Öffnen und Schließen von Vorhängen, Jalousien und Türen sowie die Steuerung des Heizungssystems.

Menschen mit Sehbehinderung haben Schwierigkeiten beim Ein- und Ausschalten von Elektrogeräten, da es für sie ziemlich mühsam ist die elektrischen Schalter zu erreichen, welche z.B. an der Wand montiert sind. Tan et al. [15] haben ein assistives Gerät für Sehbehindete Menschen entwickelt, um Haushaltsgeräte praktischer zu steuern. Auf ein speziell entwickeltes Fernbedingungspanel werden die Braille-Beschriftungen zusammen mit den mechanischen

Tasten hinzugefügt, damit die Menschen besser wissen können, welche Elektrogeräte sie ein- oder ausschalten möchten. Auf dem Fernbedingungspanel ist eine Miniatur-Sprachanzeige eingebettet, um zu bestätigen, welche Elektrogeräte sie ein- oder ausgeschaltet haben. WiFi-Technologie wird als übertragendes Medium verwendet. Die drahtlose Natur des Netzwerkes ermöglicht den Benutzern, von jedem beliebigen Ort die Elektrogeräte zu steuern, solange sie innerhalb der Internet-Abdeckung bleiben. Nach der Evaluation haben mehr als 75% der Teilnehmer angegeben, dass sie gerne oder sehr gerne dieses Gerät zuhause verwenden.

Abbildung 6 WiFi basiertes Fernbedingungspanel mit Braille-Beschriftungen für elektrische Haushaltsgeräte [15].

5.3 Persönliche Pflege und Versorgung der Gesundheit

Medikamentmanegement ist ein komplexer Prozess und Teil des Alltagslebens, jedoch ist es viel schwieriger für Menschen mit Sehbehinderung. Technologie kann eine wichtige Rolle spielen, um die Mendikamenteneinhaltung zu verbessern [12]. Bei der Verwaltung der Medikamente haben Forscher bereits RFID (Radio frequency identification) verwendet [7]. Außerdem wurden Barcodes und zwei-dimensionale QR Codes eingesetzt, um medizinische Fehler zu vermeiden [9]. Solche Technologien können auch in assistiven Systemen eingebettet werden, damit die Einhaltung von Medikamenten bei sehbehinderten Menschen verbessert werden kann. Farhadyar et al. [5] haben ein assistives Medikamentenmanagement System entwickelt, welches auf IoT (Internet of things) basiert und RFID Technologie und mobile Computing anwendet. Es besteht aus einer Android Applikation, einer Medikamentenbox und einem RFID Gerät. Ein Bluetooth Modul wurde ins System integriert und Vibrationsmotoren der Medikamentenbox wurden mit Microkontroller verbunden. Bei der Registrierung werden alle Informationen zu den Medikamenten eingegeben und die aufklebbaren RFID-Tags werden in ein bestimmtes Fach der Medikamentenbox gelegt. Anschließend bietet das System zum festgelegten Zeitpunkt eine Erinnerung, indem es eine akustische Warnung aktiviert. Dann werden die Vibrationsmotoren angeschaltet und helfen beim Suchen der Medikamente. Die Kommunikation zwischen dem Handy und dem RFID Gerät wird durch Bluetooth hergestellt.

Abbildung 7 Nutzer können beim Registrierungsprozess die Informationen zu jeweiligen Medikamenten in einer Android Applikation eingeben [5].

Dieses System bietet Feedback welches vom Tastsinn des Benutzers empfangen werden kann, damit Menschen mit Sehbehinderung im Alltagsleben ihre Medikamente besser managen können.

6 Diskussion und Ausblick

Diese Arbeit bietet einen Überblick über einige der Technologien, mit denen Menschen mit Sehbehinderung bei alltäglichen Aktivitäten unterstützt werden. Assistenztechnologien vom Low-End bis zum High-End wie z.B. Bildschirmlupe, Brailleschriftdarstellung, Sprachsteuerung, Bilderkennung, RFID-Beschriftung usw. können gut alleine funktionieren oder mit anderen integriert werden. Sie können verschiedene Arten von Feedback anbieten, und zwar visuelles Feedback, akustisches Feedback und haptisches Feedback. Betrachtet man den aktuellen Stand der assistiven Technologie, sieht man sowohl Vorteile als auch Nachteile. Einerseits stehen verschiedene, sehr nützliche und oft geniale Geräte zur Verfügung oder könnten bald in der Zukunft angewendet werden. Auf der anderen Seite sind viele dieser Geräte ziemlich teuer, oft ein Mehrfaches der Kosten der entsprechenden Geräte für Sehende, und es

gibt nicht so viele Auswahlmöglichkteiten wie für Sehende. Menschen mit Sehbehinderung könnten auch schwer von neuartigen Technologien erfahren, oder die Produkte haben keine gute Dokumentation in lokalen Sprachen. Die Lösungen, die in außereuropäischen Sprachen verfügbar sind, sind relativ selten und einige Geräte sind nur in Englisch verfügbar. Es kann sein, dass mangelnde Informationen und hohe Kosten es verhindern könnten, dass Menschen mit Sehbehinderung die Geräte im Alltagsleben ausprobieren und anwenden, von denen sie profitieren könnten. Aber wenn all diese Probleme in der Zukunft gelöst werden, könnten Menschen sehr davon profitieren.

Literatur

1 Amazon. What Is a Voice User Interface? Retrieved Febuary 2, 2021 from `https://developer.amazon.com/en-GB/alexa/alexa-skills-kit/vui`.

2 Erin Brady, Meredith Morris, Yu Zhong, Samuel White, and Jeffrey Bigham. Visual challenges in the everyday lives of blind people. pages 2117–2126, 04 2013. doi:10.1145/2470654.2481291.

3 Brendan Cassidy, Gilbert Cockton, and Lynne Coventry. A haptic atm interface to assist visually impaired users. 10 2013. doi:10.1145/2513383.2513433.

4 Joseph Paul Cohen. Blindtool – a mobile app that gives a 'sense of vision' to the blind with deep learning, 2015. URL: `https://github.com/ieee8023/blindtool`.

5 Kiana Farhadyar, Reza Safdari, Ahmad Beh-Pajooh, and Iman Nematollahi. Assistive medication management system for users with visual impairment. 06 2018. doi:10.3233/978-1-61499-868-6-53.

6 Jason Fritz and Kenneth Barner. Design of a haptic data visualization system for people with visual impairments. *Rehabilitation Engineering, IEEE Transactions on*, 7:372 – 384, 10 1999. doi:10.1109/86.788473.

7 Bryan Houliston, David Parry, and Robert Ticehurst. Procedural error identification in ward-based drug dispensing via rfid. volume 16, 01 2011.

8 Shaun Kane, Jeffrey Bigham, and Jacob Wobbrock. Slide rule: Making mobile touch screens accessible to blind people using multi-touch interaction techniques. pages 73–80, 01 2008. doi:10.1145/1414471.1414487.

9 Carol Keohane, Catherine Yoon, Matthew Ditmore, Anne Bane, Osnat Levtzion-Korach, Thomas Moniz, Jeffrey Rothschild, Allen Kachalia, Judy Hayes, William Churchill, Stuart Lipsitz, Anthony Whittemore, David Bates, and Tejal Gandhi. Effect of bar-code technology on the safety of medication administration. *The New England journal of medicine*, 362:1698–707, 05 2010. doi:10.1056/NEJMsa0907115.

10 Taeyong Kim, Sanghong Kim, Joonhee Choi, Youngsun Lee, and Bowon Lee. Say and find it: A multimodal wearable interface for people with visual impairment. pages 27–29, 10 2019. doi:10.1145/3332167.3357104.

11 MIT. The technology we build, 2013. Retrieved Febuary 2, 2021 from https://youtu.be/_t2FSzdOFak.

12 Eyal Oren, Ellen Shaffer, and B Guglielmo. Impact of emerging technologies on medication errors and adverse drug events. *American journal of health-system pharmacy : AJHP : official journal of the American Society of Health-System Pharmacists*, 60:1447–58, 08 2003. doi:10.1093/ajhp/60.14.1447.

13 Shishir Patil, Don Dennis, Chirag Pabbaraju, Nadeem Shaheer, Harsha Simhadri, Vivek Seshadri, Manik Varma, and Prateek Jain. Gesturepod: Enabling on-device gesture-based interaction for white cane users. pages 403–415, 10 2019. doi:10.1145/3332165.3347881.

14 Stanfordonline. Stanford Seminar - Accessibility and the AI Autumn, 2020. Retrieved Febuary 2, 2021 from https://www.youtube.com/watch?v=PJE_gnTreBo.

15 Chi Tan, Wai Wong, Chun Loh, and Min Soe. Wi-fi based smart home electrical appliances remote control for visually impaired person. pages 153–158, 06 2020. doi:10.1109/I2CACIS49202.2020.9140090.

16 Wikipedia. Benutzerschnittstelle. Retrieved Febuary 2, 2021 from https://de.wikipedia.org/wiki/Benutzerschnittstelle.

17 Wikipedia. Touch user interface. Retrieved Febuary 2, 2021 from https://en.wikipedia.org/wiki/Touch_user_interface.

18 Wikipedia. Unterstützungstechnologie. Retrieved Febuary 2, 2021 from https://de.wikipedia.org/wiki/Unterst%C3%BCtzungstechnologie.

19 World Health Organization. *World report on vision*, volume 214. 2019.

20 Hanlu Ye, Meethu Malu, Uran Oh, and Leah Findlater. Current and future mobile and wearable device use by people with visual impairments. *Conference on Human Factors in Computing Systems - Proceedings*, 04 2014. doi:10.1145/2556288.2557085.

21 Yuhang Zhao, Elizabeth Kupferstein, Brenda Castro, Steven Feiner, and Shiri Azenkot. Designing ar visualizations to facilitate stair navigation for people with low vision. 08 2019. doi:10.1145/3332165.3347906.

Electronic Travel Aids - A Survey of Recent Developments

Andrea Baumann

Ludwig-Maximilians-Universität München, München, Deutschland
Andrea.Baumann@campus.lmu.de

Abstract

Approximately one billion people worldwide suffer from visual impairments, which drastically affect their independent mobility. Since the 1950s researchers have therefore worked on Electronic Travel Aids (ETAs) to mitigate these mobility problems and allow Visually Impaired Persons (VIPs) to navigate safely and effectively. This paper presents a comparative survey of nine recently developed ETAs, to inform the research community of the progress made in assistive technology for VIPs. The projects themselves differ in their developmental stages with some being only very early prototypes while others have already conducted first user studies. The comparison indicates a trend towards low-cost designs, which are also affordable for VIPs with a lower income. Nonetheless, the development of human-computer interfaces suitable for VIPs remains a challenge, just as it has been in the past, and more research addressing this issue has to be conducted.

2012 ACM Computing Classification Human-centered computing → Accessibility → Accessibility technologies

Keywords and phrases Electronic-Travel-Aids; Visual Impairment; Haptic; Auditory; Navigation; Obstacle Detection.

1 Introduction

According to recent estimations made by the WHO, in 2020 there were at least 1 billion visually impaired persons (VIPs) worldwide [30]. Approximately 49 million of these being blind [2]. Additionally, there are estimations that 87 % of the world's blind live in developing countries [27]. However, even in industrialised countries such as the United States of America, the majority of those who are visually impaired have a lower income. Out of 19 million visually impaired Americans over 70 % are not under full-time employment and over a quarter lives below the poverty line [22]. As described by Marston et al. [19, 17] this high unemployment rate may at least in part be due to the difficulties VIPs face while travelling independently. While sighted people can

Cite as: Andrea Baumann. Electronic Travel Aids - A Survey of Recent Developments. In *6th Seminar on Ubiquitous Interaction (UBIACTION 2021)*. Editors: F. Lang, R. Welsch, L. Haliburton, F. Draxler, S. Feger, S. Villa, M. Hoppe, P. Knierim, V. Mäkelä, A. Schmidt. February 11, 2021. Munich, Germany. pp. 5:1–5:18.

use visual cues, maps and signs to gain knowledge about spatial information, VIPs often face the problem of losing track of their location in open spaces or even in what direction they are facing. Walking to new locations where no learned path can be followed is challenging. Some of the most difficult tasks VIPs face when travelling independently are finding locations such as train platforms or bus stops, finding the correct transport vehicle as well as changing between transportation methods in a safe and efficient way. These difficulties often result in VIPs refraining from travelling in the first place with only 15 % of VIPs leaving their homes daily [18, 15].

The navigation tools traditionally used by VIPs trying to overcome these problems, are white canes and guide dogs. A white cane is the simplest, cheapest and most reliable navigation aid and there are many variants such as *Long Canes, Symbol Canes* and *Guide Canes* [3, 7]. Despite that, there are still various reasons as to why VIPs avoid using it. The most important factors being the fear of striking either people or fragile objects, the fear of collisions with obstacles above the height detectable by the cane, and the social stigma associated with using a white cane [15].

Guide Dogs on the other hand generally allow for faster travel and often act as a deterrent to personal attacks such as thievery. There are, however, also numerous reasons why a guide dog might not be a VIPs desired travel aid. Dogs need daily care which requires time and is a huge responsibility not everyone might be ready for. VIPs also need living accommodations large enough for the animal and unlike a white cane, guide dogs lead to ongoing costs such as food and veterinarian bills. Last but not least a VIP might not be fond of dogs or might even be allergic to them [28].

Ever since the 1950s, researchers have therefore worked on Electronic Travel Aids (ETAs) in an attempt to resolve the issues VIPs face while travelling independently [23]. This research focuses primarily on two basic problems: 1) Obstacle avoidance in the VIPs immediate vicinity of up to 5 meters and 2) Orientation and Navigation [24].

ETAs are often either hand-held [12, 31], cane-mounted [14, 26] or head-mounted [29, 10] and their input methods can be classified into two major categories: *sensor input* and *camera input* [20]. However, the prices for commercially available ETAs vary from 90 up to 800 USD and are therefore oftentimes unaffordable [22]. Since only recently developed ETAs were to be surveyed for this paper, nine ETAs published during the years 2019 and 2020 were selected for review and will be described briefly. Six of these prototypes use auditory feedback and the remaining three projects focus on haptic feedback. Although using auditory feedback tends to be cheaper, haptic feedback, unlike auditory feedback, does not block a VIPs ability to hear ambient noise. The different systems were picked because of their novel

approaches to mitigating the problems VIPs face when travelling independently. However, the projects differ in their developmental stage with some being only very early prototypes while others have already conducted first user studies. For every prototype, the pros and cons will be stated clearly. However, these differences in development, as well as the information provided about the ETAs, do not allow for a standardisation of advantages and disadvantages. Following the presentation of these prototypes, a discussion comparing the projects with one another concludes the paper. The discussion is based on four important ETA characteristics *Free-hands, Free-ears, Wearable* and *Simple*. Focusing on not requiring the user to hold the device, to leave the VIPs hearing unimpaired, to allow for multitasking while using the ETA and last but not least, to be easy to use without extensive training [3].

2 Auditory Feedback

For a VIP the ability to hear ambient noise is essential. Despite this being a widely known fact [3, 1] a majority of ETAs developed still use auditory instructions as a guiding method for VIPs. However, the simultaneous presence of ambient noise and the auditory instruction produced by the ETA increases the likelihood of both sounds interfering with one another [1]. Nonetheless, the decision to use auditory feedback is no recent trend and has also been present in previous, older projects [3]. One explanation for this might be the fact that auditory feedback methods tend to be cheaper than haptic ones [25]. This section presents and analyses six recently developed prototypes all of whom use some form of auditory instructions for navigation and/or obstacle detection.

2.1 Lodz Project

Skulimowski et al. developed a prototype (see Figure 1) for interactive sonification, i.e. the conveyance of information via non speech audio, of U-depth images [25]. The system combines a depth camera with an Android OS smartphone and open in-ear headphones. The depth camera has a depth reconstruction range of 40 centimetre - 5 meter and is mounted on the user's head via a dedicated headgear made of elastic straps. The VIP can adjust the covered distance using the accompanying android app to a distance smaller than 5 meters. Afterwards, the selected distance is verbally announced using Text-To-Speech (TTS). The pre-processing procedures and the U-depth map calculations happen on the Android platform in real-time, with depth being converted into sound and mapped to the frequency scale to represent obstacle distances. The lower the frequency, the further away from the object. If an

(a)

(b)

Figure 1 a) The left image shows the hardware components of the Lodz Project. The scanner is mounted on a plastic head and the Android application containing the depth map is displayed. b) On the right a VIP using the system is shown [25].

obstacle is close, an alarm sounds over the headphones. The android app offers three different operation modes.

While in the **Interactive Sonification Mode** the user can select a certain scene area for sonification by touching the desired area on the U-depth map presented in the application. The second operation mode **Promixity Mode** is the default mode of the system and does not require interaction with the screen. If this mode is active, warnings will only be given for the closest obstacles and out of ten depth ranges only the first three will be sonified. A vertical fling over an object on the depth map activates the **Verbal Mode** and information about the distance to the object in centimetre is conveyed verbally.

Pros: The used depth scanner has a weight of 95 grams and is therefore both lightweight and compact. A user study on indoor mobility tasks was conducted with three VIP, all of whom were familiar with smartphones. It was discovered that the ETA was easy to use and a ten-minute introduction to the system was sufficient for successful usage.

Cons: The offered Field of View (FoV) is rather narrow when compared to human sight (58° horizontally, 45° vertically). However, according to the VIP who tested the device, this was a welcome restriction as it helps to limit the amount of sonified information. However, the **Interactive Sonification Mode** and the **Verbal Mode** require active interaction with the smartphone.

2.2 CREATION

CREATION was developed by Manjari et al. at the Bennett University in Greater Noida, India [16]. Their prototype consists of a Raspberry Pi 4, a Pi camera, an ultrasonic sensor, a MI power bank, a vibration motor as well as Bluetooth earphones. The developers decided upon using Raspberry Pi 4 as it performed twice as fast as the Raspberry Pi 3, which they also tested. CREATION is mounted on a cane. Pi camera captures frames in real-time and resizes them to 224px x 224px. Afterwards, the object detection models *You Only Look Once (YOLO)* and *Single Shot Multibox Detector (SSD)* classify the captured objects as either car or person. If a car or a person is present a voice-based audio output communicates the nature and range of the object to the VIP through the Bluetooth earphones. Additionally, the integrated vibration motor activates if the sonar sensor detects an obstacle within its range. The 10.000 mAH capacity of the chosen power bank allows for 24 hours of usage. However intensive use is not recommended.

Pro: CREATION allows for real-time processing and is a low cost, light-weight enriched aid.

Cons: The main contribution of the project itself lies with the developed object recognition algorithm. However, at the time of publication, the algorithm only recognised two different classes: *Car* and *Person*. Furthermore, the prototype seems to be untested for usability in everyday tasks, as neither a lab- nor a field-study with either blind-folded or visually impaired participants was reported.

2.3 Smart Cap

Smart Cap [20] is an ultrasonic sensor-based system. For the prototype, three ultrasonic sensors, an ATmega microcontroller, an Arduino board, and a buzzer were built into a commercially available cap (see Figure 2). The sensors are facing three different directions. If an object enters a sensor's set vicinity (less than 1 meter) the microcontroller activates the buzzer. However, the distance which triggers buzzer activation can be modified as long as it does not exceed the range of the sensor. As all three sensors activate the same buzzer, unique buzzing frequencies for all three directions were used so that the user can interpret the location of the object correctly.

Pros: Smart Cap is a low-cost device allowing for real-time processing. As the device is worn on the head the hands of the VIP are free and the user can move around without restrictions. Since a commercially available cap is used, Smart cap is also inconspicuous and can therefore be worn without attracting attention. Some initial testing of the prototype gave promising results, achieving a 95% object detection accuracy.

Cons: In comparison to other projects presented in this paper Smart Cap is a simplistic prototype (only one buzzer producing one type of sound feedback was used). The reason behind this is the developers aim to keep Smart Cap as low in costs as possible. Additionally, no user study has been conducted. The authors tested the developed model by blindfolding themselves as well as simulating the model in Proteus 8.

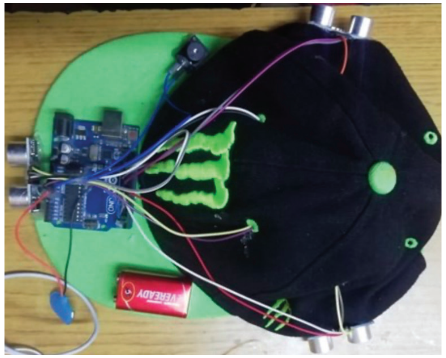

Figure 2 This image depicts the Smart Cap prototype. The components of the device are attached to the top of a commercially available cap [20].

2.4 Blind People Playing Sport

This project by Ferrand et al. [6] has been designed for sports practice and is intended to be used in a controlled environment. For this, a spatialised virtual stereo sound source is created, via which a VIP can be guided along a fixed path. This ETA is a low-power ARM platform and utilises a low-latency real-time location system (RTLS) and an assembly of three low-cost sensors for head tracking. Audio output is provided via headphones. The RTLS consists of a D-GNSS module used for outdoor localisation. The module has a small portable patch antenna which allows for an average location accuracy of 10-20 centimetre. For indoor localisation, a UWB chip has been used. UWB uses a radio signal's Time of Flight, i.e. the time it takes for a signal to travel from transmitter to receiver, for distance measurement. This results in an accuracy of 10 centimetres within a 30 x 30 meter square. The Head-Tracker can be connected to the main system via either Bluetooth or USB. When using the ETA there are two steps to follow: First, a sighted person records a

track trajectory using the device. Then a VIP starts the guidance mode at a position close to the recorded track and follows the sound source along the recorded trajectory.

Pros: The usage of a spatialised sound source with real-time localisation allows for the VIP to follow along with a recorded track naturally with minimal training involved and the user can follow the sound almost immediately. Performance and speed can however be improved through practice. The ETA itself is compact and weighs less than 1kg with an autonomy of about 2 hours. Furthermore, the prototype was tested with a sample containing both sighted as well as blind people.

Cons: As the system currently does not support obstacle detection the ETA is only suitable for usage in controlled environments. Moreover, even though a user study has been conducted the sample size was small with only four participants of which only two were visually impaired.

2.5 Clunj-Napoca Project

Clunj-Napoca Project was developed by Sebestyen et al. at the Technical University of Clunj-Napoca [24]. The project consists of a two-part solution: An Arduino based device for obstacle avoidance and a mobile application for navigation. Both modules can be used standalone or in combination with each other. The developers put a lot of thought into the component selection to achieve a low priced end product. For feedback, an auditory output was chosen. However, the developers explicitly advise placing the earphones behind the ears where the sound is then transported to the inner ear via the scalp bone.

Obstacle Avoidance:

For the obstacle avoidance module, they used an Arduino Uno based on its scalability, compactness and low price. An ultrasound distance sensor, Gyroscope sensor and dual audio amplifier were connected to it and placed into a 3D printed case. The ETA is supposed to be placed on a walking stick for maximum performance. If the module is used in combination with the navigation system then a Wi-Fi module needs to be attached as well. During system use, a distinction is made between two working modes.

In the **Beginner Mode** there are no scanning movements with the stick. The users instead hold the walking stick in a fixed position in front of themselves. If an object is detected intermittent sounds are generated in both ears. Depending on the measured distance to the obstacle one of three sound levels is chosen. In the **Advanced Mode** VIPs use the stick for horizontal scanning motions. Depending on the position of the detected obstacle, sounds are generated in the left ear, both ears or the right ear of the VIPs.

The system automatically switches between these two working modes, depending on a scanning motion being present or not.

Navigation:

The navigation module is run on a mobile device and uses the capabilities of a typical, commercially available mobile phone such as its GPS receiver, orientation sensor, acceleration sensor and its internet connection. Via the internet connection, the module accesses maps, paths and points of interest from a server which represents the project's third and final component. While much of the data for the maps and paths can be gathered from already available sources, sighted people can also manually enhance a path with additional information such as restaurants in proximity to a chosen path. The navigation system then relays relevant information to the user such as where to turn at cross-points and offers verbal indications of relevant points of interest.

Pros: The Clunj-Napoca Project uses a 4.8 volts rechargeable power supply which ensures 42 hours of usage. Even though they have decided on using an auditory feedback method their placement instructions for the earphones showed their consideration for the possible impacts of this decision as the blocking of ambient sound can be avoided this way. The ETA is capable of real-time processing and both cheap and portable. By including obstacle detection in addition to navigation capabilities the ETA is furthermore not restricted to controlled environments. The beginner mode is easy to use and the given feedback can be interpreted quickly without necessitating extensive training.

Cons: Advanced mode on the other hand seems to be less intuitive and some training is required for successful usage. Additionally, neither a study with VIPs nor with blindfolded, sighted people has been conducted.

2.6 Thessaly Project

The ETA shown in was developed by Dimas et al. at the University of Thessaly as a four components system [4]. It consists of a stereoscopic depth-aware RGB camera which is connected via a USB cable to the BCU, a Raspberry Pi BCU orchestrating communication between user and cloud services, a wearable Bluetooth headset and a cloud infrastructure. Additionally, a low-end mobile phone connected to the wearable system is required to act as a hotspot device so communication with the cloud is possible. The design of the system followed the user requirements listed by Iakovidis et al. [13] which led to a sunglasses shaped ETA. The wearable system itself performs risk assessment, while computationally intensive tasks such as object recognition and detection are performed on the cloud infrastructure. For the obstacle detection, a risk map is generated based on depth-values and combined with

a saliency map generated using a Generative Adversarial Network (GAN). The deep learning object classification model (LB-FCN light) then recognises an object based on the obstacle regions detected. The detection distance for an object to be a risk object is set to two meters. Object position and its labels are then communicated to the Bluetooth headset via the built-in TTS synthesizer of the BCU. As a VIP approaches an obstacle the termini *low, medium* and *high* are used to communicate the risk of collision in addition to the spatial location of the object and its labels. Objects detected repeatedly are only reported once until they become relevant again. Moreover, if a person is getting closer to an object the notification frequency increases until a STOP message is produced if the VIP continues to approach a high-risk obstacle.

Pros: The obstacle detection method developed for this project is in real-time and does not require the training of domain-specific data. Therefore obstacle detection capabilities are not limited to only certain objects. With the ETA coming as sunglasses the user's hands are free and the VIP can move around freely without attracting attention. The cloud infrastructure is furthermore suitable for horizontal scaling therefore extendable based on future needs. The glasses themselves were created with a human-centred design process.

Cons: The system currently is suitable for outdoor navigation only. Moreover, because of the aluminium case of the camera, the prototype is rather heavy which may cause discomfort for the user. There has also not been a user study. Only the algorithms themselves were tested. Field tests with VIPs and/or blind-folded participants are planned for the future.

3 Haptic Feedback

Haptic Feedback methods are based on the sense of touch. ETAs which communicate information necessary for navigation and/or obstacle detection via touch usually do so through vibration patterns [5]. They utilise the vast surface area skin offers (for example a 70 kilogram man has a skin surface area of 1.7 square-metres [9]) and the thus conveyed spatial information is then interpreted by the nervous system as directional information [1]. This way the conflict of information sources that is present with auditory feedback can be avoided, allowing for a safer guidance method [1].

3.1 IGuide

Even though IGuide has been categorised as utilising a haptic feedback method, the ETA developed by Harshitha et al. [11] uses both auditory as well as tactile feedback. The prototype which comes in the form of a boot has three ultrasonic sensors, a Wi-Fi module, two vibration motors, an APR module, Arduino, a moisture sensor, an emergency button and a battery integrated. Based on the data collected by the sensors, the microcontroller circuit constructs a logical map of the environment. The ultrasonic sensors are placed to detect floor and knee level obstacles and obstacles on the boot's outside respectively. Their measuring range spans from three centimetres up to three meters. Once an obstacle has been detected and is 25cm away from the user, the APR module produces an auditory output with the corresponding vibration motor generating haptic feedback simultaneously. A speaker is used for auditory feedback. The speaker can however be substituted with wireless headphones. With the help of the integrated moisture sensor, wet flooring can be detected which can prevent slipping accidents. The push of the emergency button sends the VIP‘s live location data to a predefined person.

Pros: IGuide is a cost-effective and lightweight prototype that does not require extensive training

Cons: As no information on conducted prototype testing has been communicated conclusions on the effectiveness of the prototype can currently not be drawn.

3.2 VVIT Project

This project from researchers at the Vedavyasa Institute of Technology (VVIT) uses ultrasonic waves for obstacle detection, measuring the elapsed time between sound wave generation and the sound wave coming back, for distance to obstacle estimation [21] . The prototype consists of multiple sensors such as an ultrasonic sensor, a Light Dependent Resistance (LDR) sensor and a Passive Infrared (PIR) sensor, a panic alert button, a microcontroller, a buzzer, an LED indicator, a vibration motor, record and playback IC for audio output as well as a GSM-GPS module. Once an object has been detected by the ultrasonic sensor, the vibration motor produces tactile feedback. To avoid accidental collisions with seeing people during night time, the LDR sensor activates the LED light connected to the front of the device, once a certain level of darkness has been reached. The LED then illuminates brightly and effectively acts as a flashlight. Furthermore, if the PIR sensor detects the nearby presence of another person, the playback IC produces an audio output. In case of emergency, a press of the panic alert button produces a buzzing sound to alert bystanders. Additionally, an SMS message containing the live location data of the user is being sent to all numbers saved in the microcontroller.

Pros: The developers deliberately decided to use haptic feedback for their low-cost device, instead of the more commonly used audio output. This way the device is not only accessible for visually impaired users but also VIPs with auditory impairment.

Cons: Although the described prototype is quite complex, certain things such as how the device is intended to be used (whether it is intended to be placed on a cane, handheld, head-mounted, etc.) or what kind of output the playback IC produces for nearby people are left uncovered. Conclusions on the effectiveness of the prototype can also not be drawn as no information regarding the testing of the device has been presented.

3.3 Open Source Project

The prototype (see Figure 3) developed by Petsiuk et al. [22] utilises an ultrasonic sensor in a 3-D printed bracelet case. There were two different bracelet models developed having one or two vibration motors respectively. An especially developed Arduino software performs distance measurements covering a range from 2 cm up to 4 m which corresponds to the specifications of the sensor. The implemented algorithm then utilises two vibration motors to convey point-distance information with variations in frequency and amplitude, depending on the distance to the obstacle. However, the prototype is intended as an independent addition to traditional tools and not as a replacement for such.

Pros: Petsiuk et al.'s prototype does not require complex calibrations or extensive training. It additionally consists of only a small number of readily available and inexpensive components resulting in a material bill of less than 24 USD. This makes the device extremely affordable. The ETA has furthermore been tested with five sighted blindfolded people performing nine primary everyday navigation and guidance tasks. The chosen tasks included both indoor and outdoor navigation as well as the avoidance of collisions with pedestrians. Since the device is in the form of a bracelet, the VIP‘s hands remain free for other tasks.

Cons: Because of its narrow scanning angle, the device only scans in front of the user. To avoid head-level obstacles, users have to sweep the head area along the direction in which they are moving. Moreover, small and fast-moving obstacles can not be detected. Although a user study has been conducted the indoor and outdoor environments were familiar to the users. This might have affected the participants' performance. The ETA has also not yet been tested with visually impaired people.

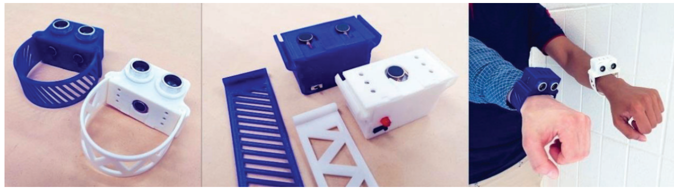

Figure 3 This image shows both prototype versions of the Open Source Project. The version using two vibration motors for feedback is blue and the version using only one vibration motor is white. Additionally two people wearing the bracelets are depicted [22].

4 Discussion

While the points speaking for or against the individual projects have already been discussed in their corresponding sections, we now aim to put the presented ETAs into comparison with one another. We are aware that a direct comparison might seem disadvantageous for some of the presented projects, given their different development stages. However, since we are focusing on the concept rather than the current state of the device itself, this is not a problem. As previously mentioned we are using the ETA characteristics *Free-hands, Free-ears, Wearable* and *Simple* as evaluation criteria. These are in accordance with the four characteristics important for ETAs, as stated by Dakopoulos and Bourbakis [3] in their 2010 survey of ETA devices. Table 1 presents a summary of the ETA characteristics met by each of the covered projects. They will additionally be covered extensively in the following.

Free-hands: Even though most ETAs aim to completely substitute conventional travel aids such as white canes, VIPs might wish to use their previous navigation device in addition to an ETA. Therefore it is important that VIPs do not have to hold the ETA in their hands [3]. Out of the presented projects, all but the *Lodz Project (2.1)* meet these criteria. Even though *Lodz Project* uses a head-mounted camera system VIPs need their hands to use the accompanying smartphone application. As the developers of the *VVIT Project (3.2)* have not stated whether the device is intended to be hand-held, head-mounted or used otherwise it is ambiguous whether the project does meet the criteria or not. *CREATION (2.2)* and the *Clunj-Napoca Project (2.5)* strictly speaking require a cane which is handheld, however as a visually impaired person most likely uses one for navigation anyways, they can be considered as fulfilling the criteria.

Free-ears: Free-ears refers to a VIP‘s ability to hear ambient noise while using the ETA. A requirement which several of the devices opting for an auditory feedback method do therefore not fulfil. Nonetheless designing an audio output based ETA does not automatically eliminate the Free-ears criteria. Out of the six presented auditory prototypes 1/3 (*Smart Cap (2.3)* and *Clunj-Napoca Project (2.5)*) managed to both use auditory feedback and be Free-ears. *Clunj-Napoca Project (2.5)* even gave explicit instructions to place the earphones necessary for the ETA behind instead of inside the ear, to avoid said restrictions to a VIPs ability to hear. The three haptic feedback based prototypes all fulfil the Free-ears criteria. IGuide which utilises both haptic as well as auditory feedback methods decided on using a speaker instead of earphones and the *VVIT Project (3.2)* expressibly decided upon haptic feedback for the prototype to be accessible for VIPs with an auditory impairment as well.

Wearable: The wearable characteristic requires the ETA to be augmented into all actions. Therefore offering flexibility and the possibility to multitask during usage [3]. All systems but *Lodz Project (2.1)* meet these requirements since they need to operate the smartphone application renders VIPs unable to something else simultaneously. Although *CREATION (2.2)* and the *Clunj-Napoca Project (2.5)* are not worn on the VIP's body but placed on a cane, this fact does not impair the users flexibility. Despite nothing being known on the intended use of the *VVIT Project (3.2)* the prototype's compactness leads to the assumption of it being not hindering the users' flexibility either.

Simple: An ETA is considered simple if the system is easy to use and does not require extensive training [3]. For more than half of the presented ETAs, no user study has been conducted yet. Therefore conclusions on the devices ease of use can not be drawn. However *The Lodz Project (2.1), Blind People Playing Sport (2.4), Clunj-Napoca Project (2.5)* and the *Open Source Project (3.3)* do not require extensive training before the prototypes can be used successfully.

As just noted, more than half of the projects had not conducted a user study. Out of the projects where a study had been conducted, *Lodz Project (2.1)* had a sample size of N=3, *Blind People Playing Sport (2.4)* had a sample of N=4 (two of whom were visually impaired) and the *Open Source Project (3.3)* had a sample size of N=5 (all of whom were sighted individuals). This tendency is applicable for ETA research in general as it usually has extremely small sample sizes or has even been carried out by only the device's developers [8].

Looking back at the introduction we see that numerous VIPs suffer from lower-income or live below the poverty line. A detail which several of the recently developed ETAs seem to consider as they actively tried to keep their

ETA Characteristics				
	Free-hands	Free-ears	Wearable	Simple
Lodz Project	X	X	X	✓
CREATION	(✓)	X	(✓)	X
Smart Cap	✓	✓	✓	X
Blind People Playing Sport	✓	X	✓	✓
Clunj-Napoca Project	(✓)	✓	(✓)	✓
Thessaly Project	✓	X	✓	X
IGuide	✓	✓	✓	X
VVIT Project	(✓)	✓	✓	X
Open Source Project	✓	✓	✓	✓

Table 1 Table summarising the characteristics met by the individual projects

device as cheap and therefore affordable, as possible. The *Clunj-Napoca Project (2.5)* especially carefully selected their components to achieve high performance with as low an expense as possible. It is furthermore distinctive that devices which achieved to be of low cost used ultrasonic sensors. This goes along with the fact that although devices utilising stereo cameras might have a more precise input, they are also more expensive [24].

5 Conclusion

After carefully analysing the above-presented systems it becomes apparent that just as it has been in the past [3], the main challenge of developing a successful ETA is not the development of the technology itself but rather the development of a suitable interface between system and user. Table 1 reinforces this, as only two out of the nine presented projects (one for each feedback modality) managed to fulfil all four characteristics which are, according to Dakopoulos and Bourbakis [3], important for a prevailing ETA. The *Clunj-Napoca Project (2.5)* has proven that the decision to utilise auditory feedback does not render a device unable to meet these criteria if the ramifications are considered and counteracted properly. The *Open Source Project (3.3)* on the other hand demonstrated that haptic feedback methods are not necessarily more expensive than auditory ones. Rendering haptic feedback ETAs as financially accessible as auditory ones and therefore worth deeper exploration. With this in mind, a whole new generation of both auditory and haptic ETAs is possible. It furthermore becomes apparent that there might be a possible trade-off to consider between price and effectiveness. Recent research focused on developing prototypes that are affordable for everyone. Often choosing

low-cost ultrasonic sensors rather than stereo cameras which might have a more precise input but are also more expensive [24]. Whether this trade-off is worthwhile remains yet to be seen.

6 Acknowledgements

We wish to thank the authors of the Lodz Project, Smart Cap and the Open Source Project for permitting their images to be used in this publication.

References

1 Arnav Bharadwaj, Saurabh Bhaskar Shaw, and Daniel Goldreich. Comparing tactile to auditory guidance for blind individuals. *Frontiers in Human Neuroscience*, 13:443, 2019.

2 Rupert RA Bourne, Jaimie Adelson, Seth Flaxman, Paul Briant, Michele Bottone, Theo Vos, Kovin Naidoo, Tasanee Braithwaite, Maria Cicinelli, Jost Jonas, et al. Global prevalence of blindness and distance and near vision impairment in 2020: progress towards the vision 2020 targets and what the future holds. *Investigative Ophthalmology & Visual Science*, 61(7):2317–2317, 2020.

3 Dimitrios Dakopoulos and Nikolaos G Bourbakis. Wearable obstacle avoidance electronic travel aids for blind: a survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40(1):25–35, 2009.

4 George Dimas, Dimitris E Diamantis, Panagiotis Kalozoumis, and Dimitris K Iakovidis. Uncertainty-aware visual perception system for outdoor navigation of the visually challenged. *Sensors*, 20(8):2385, 2020.

5 Jan BF van Erp, Katja I Paul, and Tina Mioch. Tactile working memory capacity of users who are blind in an electronic travel aid application with a vibration belt. *ACM Transactions on Accessible Computing (TACCESS)*, 13(2):1–14, 2020.

6 Sylvain Ferrand, François Alouges, and Matthieu Aussal. An electronic travel aid device to help blind people playing sport. *IEEE Instrumentation & Measurement Magazine*, 23(4):14–21, 2020.

7 CABVI Central Association for the Blind and Visually Impaired. What are the different types of white canes? accessed: 19.01.2021 at. URL: `https://www.cabvi.org/articles/what-are-the-different-types-of-white-canes/`.

8 Nicholas A Giudice and Gordon E Legge. Blind navigation and the role of technology. *The engineering handbook of smart technology for aging, disability, and independence*, 8:479–500, 2008.

9 Lowell A Goldsmith. My organ is bigger than your organ. *Archives of dermatology*, 126(3):301–302, 1990.

10 José L González-Mora, A Rodriguez-Hernandez, LF Rodriguez-Ramos, L Díaz-Saco, and N Sosa. Development of a new space perception system for blind people, based on the creation of a virtual acoustic space. In *International Work-Conference on Artificial Neural Networks*, pages 321–330. Springer, 1999.

11 R Harshitha, Sabha Aafrin Razvi, G Sahana, A Shiva Shankaran, and D Mahesh Kumar. Iguide an electronic aid for visually impaired. *Journal of VLSI Design and Signal Processing*, 5:6–11, 2019.

12 Andreas Hub, Joachim Diepstraten, and Thomas Ertl. Design and development of an indoor navigation and object identification system for the blind. *ACM Sigaccess Accessibility and Computing*, (77-78):147–152, 2003.

13 Dimitris K Iakovidis, Dimitrios Diamantis, George Dimas, Charis Ntakolia, and Evaggelos Spyrou. Digital enhancement of cultural experience and accessibility for the visually impaired. In *Technological Trends in Improved Mobility of the Visually Impaired*, pages 237–271. Springer, 2020.

14 Sound Foresight Technology Ltd. Ultracane - putting the world at your fingertips. accessed: 01.01.2021 at. URL: `https://www.ultracane.com/soundforesigntechnologyltd/ultracane`.

15 Shachar Maidenbaum, Shlomi Hanassy, Sami Abboud, Galit Buchs, Daniel-Robert Chebat, Shelly Levy-Tzedek, and Amir Amedi. The "eyecane", a new electronic travel aid for the blind: Technology, behavior & swift learning. *Restorative neurology and neuroscience*, 32(6):813–824, 2014.

16 Kanak Manjari, Madhushi Verma, and Gaurav Singal. Creation: Computational constrained travel aid for object detection in outdoor environment. In *2019 15th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, pages 247–254. IEEE, 2019.

17 James R Marston and Reginald Golledge. Removing functional barriers: Public transit and the blind and vision impaired. 2001.

18 James R Marston and Reginald G Golledge. The hidden demand for participation in activities and travel by persons who are visually impaired. *Journal of Visual Impairment & Blindness*, 97(8):475–488, 2003.

19 James R Marston, Reginald G Golledge, and C Michael Costanzo. Investigating travel behavior of nondriving blind and vision impaired people: The role of public transit. *The Professional Geographer*, 49(2):235–245, 1997.

20 Vipul Parihar, Yogesh Rohilla, and Kusum Kumari. Ultrasonic sensor based smart cap as electronic travel aid for blind people. In *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*, pages 873–877. IEEE, 2020.

21 C Periasamy, V Krishnapriya, et al. Electronic travel aid for visually impaired people along with a panic alert system. *International Journal of Computer Communication and Informatics*, 1(1):17–21, 2019.

22 Aliaksei L Petsiuk and Joshua M Pearce. Low-cost open source ultrasound-sensing based navigational support for the visually impaired. *Sensors*, 19(17):3783, 2019.

23 Santiago Real and Alvaro Araujo. Navigation systems for the blind and visually impaired: Past work, challenges, and open problems. *Sensors*, 19(15):3404, 2019.

24 Gheorghe Sebestyen, Anca Hangan, and Anca Misin. Travel assistant for persons with visual disabilities. In *2020 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR)*, pages 1–6. IEEE, 2020.

25 Piotr Skulimowski, Mateusz Owczarek, Andrzej Radecki, Michal Bujacz, Dariusz Rzeszotarski, and Pawel Strumillo. Interactive sonification of u-depth images in a navigation aid for the visually impaired. *Journal on Multimodal User Interfaces*, 13(3):219–230, 2019.

26 Iwan Ulrich and Johann Borenstein. The guidecane-applying mobile robot technologies to assist the visually impaired. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 31(2):131–136, 2001.

27 Ramiro Velázquez. Wearable assistive devices for the blind. In *Wearable and autonomous biomedical devices and systems for smart environment*, pages 331–349. Springer, 2010.

28 Second Sense Beyond vision loss Kathy Austin. White cane vs. guide dog: Why or why not? accessed: 19.01.2021 at. URL: `http://www.second-sense.org/2016/09/white-cane-vs-guide-dog-why-or-why-not/`.

29 Jamie Ward and Peter Meijer. Visual experiences in the blind induced by an auditory sensory substitution device. *Consciousness and cognition*, 19(1):492–500, 2010.

30 WHO. World health organisation. blindness and vision impairment. accessed: 30.12.2020 at. URL: `https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment`.

31 Dan Yuan and Roberto Manduchi. A tool for range sensing and environment discovery for the blind. In *2004 Conference on Computer Vision and Pattern Recognition Workshop*, pages 39–39. IEEE, 2004.

The Social Side of Digital Stress

Svenja Yvonne Schött

Ludwig-Maximilians-Universität München, München, Deutschland
s.schoett@campus.lmu.de

Abstract

Digital factors play a role in the experience of stress that is a common phenomenon in today's society. As stress is detrimental to health and quality of living, research tries to assess what causes this stress and find coping mechanisms. In this chapter, we provide an overview of social causes and social solutions to digital stress. Digital interactions amplify three kinds of social pressures, namely constant awareness of other people, magnified social pressure and directly negative interactions. We present strategies for coping with these social-digital stressors and uncover which of these stressors have yet to receive adequate coping recommendations. Technology also has social features that can improve human wellbeing. Increased awareness of this dichotomy of digital interactions empowers designers to avoid amplifying social stressors through technology, as well as utilise the social benefits of technology to increase wellbeing.

2012 ACM Computing Classification • Human-centered computing→Human computer interaction (HCI) • Human-centered computing→Ubiquitous and mobile computing • Applied computing→Law, social and behavioral sciences

Keywords and phrases Digital Stress; Digital Wellbeing; Social Pressure.

1 Introduction

Stress is a common issue in today's society. In a 2018 survey, 74% of participants reported to have experienced general stress levels high enough to render them unable to cope during the previous year [30]. On a global scale, 35% of people reported feeling a lot of stress the day before a survey [14].

While stress can be the result of various tension inducing factors like illness or psychological factors [39], using information technology can also cause stress. This stress individuals experience has been labelled technostress [42] or digital stress. Technostress can result from the constant confrontation with an abundance of information available through the internet. This phenomenon is called information overload [28, 31, 35] and is detrimental to wellbeing, as the human brain is not equipped to process the amount of information adequately [35]. Further, technical devices like smartphones themselves can

Cite as: Svenja Yvonne Schött. The Social Side of Digital Stress. In *6th Seminar on Ubiquitous Interaction (UBIACTION 2021)*. Editors: F. Lang, R. Welsch, L. Haliburton, F. Draxler, S. Feger, S. Villa, M. Hoppe, P. Knierim, V. Mäkelä, A. Schmidt. February 11, 2021. Munich, Germany. pp. 6:1–6:18.

have negative effects on mental wellbeing [8, 38]. If the device is too complex or lacks usability, using it can induce stress [32]. Furthermore, social factors can contribute to digital stress. For example, social networking sites [28] and the societal expectation to be reachable at any moment [35] are associated with stress.

However, technology can also enhance human wellbeing through strengthened social connectedness, access to social support, mental health support and enjoyment [1, 6, 31, 38, 44, 45].

In this paper, we focus on this social side of digital stress. After providing an overview over useful background knowledge on stress (see Section 2.1), we categorise social-digital stressors into three major groups: Constant Awareness, Social Pressure and Directly Negative Interactions (see Section 2.2). The objective is to showcase how social pressure can contribute to digital stress. Further, we highlight how people cope with social-digital stressors (see Section 3). These discussed coping strategies range from technical approaches like the development of smarter notifications to social strategies for individuals like reaching out for help. Finally, we present how social support, companionship and positive interactions can be utilised in the digital space to improve wellbeing (see Section 4).

2 Causes of Digital Stress

2.1 General Stress

Before diving deeper into the digital side of stress we want to give a general introduction to stress. This introduction will cover what stress is and why it is so detrimental and should thus be researched.

From a medical perspective, stress can be defined as a cumulative [32] ”physical, mental, or emotional [agent] that causes [...] tension” [39]. The cause which elicits stress is called a stressor [35]. Illness and injury are internal stressors. External stressors range from stressful social situations to the environment and also include psychological factors [39].

While it may feel like everyone is constantly stressed nowadays and that being stressed is normal, stress can have serious consequences. Aside from the detrimental effect on the quality of life, stress has effects on health.

Not only is stress a contributing factor to anxiety and depression [39], it is also a factor in an abundance of physical conditions like irritable bowel syndrome or high blood pressure [39]. The immune system [17] and fertility [27] can also suffer from stress if stress isn’t being coped with effectively.

While stress responses and triggers are different on an individual basis [20], different devices can be used to estimate mental stress. Physiological reactions to stress like increased heart and breathing rate and skin conductance can

be measured with ECGs (electrocardiograms), PPGs (photoplethysmograms), respiration sensors and EDAs (electro-dermal activity sensors). Aside from this, video, audio and motion sensors can be used to capture behavioural responses to stress [23]. Further, psychological tests can be used to measure how stressful situations are and classify them e.g. on the Perceived Stress Scale (PSS) [10]. These measurements can thereafter be used to judge the effectivity of stress-reducing interventions.

When trying to understand stress from a psychological perspective the transactional model of stress by Lazarus and Folkman is a widely accepted model. According to them, stress is caused when a person’s environment is overwhelming, surpassing the limit of their resources and endangering their wellbeing [45]. Person and environment are interdependent [25] and the stress levels depend on the individual experiencing stress, the situation, time frame (as stress may change over time [25]) and the outcome [24]. The way the individual perceives the stressful situation [25] is defining for whether the stress is detrimental to their wellbeing (distress) or beneficial (eustress) [35].

It should be noted that stress can be seen as a subset of emotions called the stress emotions consisting of ”anger, anxiety, guilt, shame, sadness, envy, jealousy, and disgust” [24]. These emotions are responses to issues with the relationship between humans and their environment [24].

People experiencing stress need to deal with the stress, which is called coping. Coping is defined as the cognitive and behavioural efforts which people need to take to manage psychological stress caused by person–environment transactions [25, 45]. More than one coping strategy can be employed at once [15]. At the same time, coping strategies change over time depending on the situational context [24].

There are two main strands of coping strategies: problem-focused and emotion-focused approaches. Problem-focused approaches aim at changing the stressful situation itself [24]. Emotion-focused approaches try to decrease the emotionally perceived stress e.g. through distraction [45]. These coping approaches focus on changing the thoughts and feelings towards the stressor using cognitive processes like reappraisals and disengagement [42], while the stressful conditions themselves remain unchanged [24]. Appraisal is the ongoing process of assessing the own situation and what effects it has on ones own wellbeing and is defining for the coping process. Primary appraisal of stress depends on harm experienced in the past, anticipated harm from the stress and the chances of potentially growing from the challenge. Secondary appraisal regards if the stressful relationship with the environment can be changed [25].

A further way to distinguish different coping approaches is to classify them by style and proces. In this context, process refers to the efforts of managing stress, which are influenced by context and time. Style refers to the personality traits that influence coping, as the experience of stress differs from person to person [24].

2.2 Social-Digital Stress

Digital stress is a specific type of stress caused by technology. While digital media usage is not directly related to increased stress levels in most cases, the rise of digital media has caused changes to the environment we live in and navigate on a daily basis [16]. In many countries, access to communication technology is essential for participating in society, not only for socialisation but also for obtaining access to information and entertainment. Societal standards arise based on the opportunity of access to the internet, requiring constant communication and availability [4].

There are many contributing factors to stressful experiences with technology. However, social interactions online, as well as social structures in the real world can play an important role in enabling or contributing to digital stress.

Constant awareness of other people One of the main contributing factors to social stress in the digital landscape is the constant awareness of other people. Seeing other peoples posts -be it people one knows in real life, acquaintances made online or celebrities- can make people feel connected (see Section 4). However, knowing about the experiences other people make can cause psychological distress.

First of all, people tend to post highlight-reels of their lives online [41]. Showing their lives in the best light and only sharing the positive experiences: the amazing vacation they took, the perfect looking bread they baked, the selfie where they look stunning. In many cases, one does not get to see the 100 other selfies they took that looked just a bit worse, one does not get to see them fresh out of bed or right after a mental breakdown - but internet users do get to see those moments when they experience them themselves. The picture people paint of their lives online rarely portray those low notes (as they may make them vulnerable), but realistically everybody has problems. If a person is constantly confronted with the positive versions of others online, this can make the own life pale in comparison. This negative perspective on oneself can cause people to be depressed [41], overcompensate and stress [21], striving for an ideal they can never reach. Aside from putting energy into the way they portray themselves online, this includes making adaptations to the own life. For example, there have been instances where people who regularly use photo filters that distort the face become so discontent with the way their real face looks that they have plastic surgery in order to look more like filters [36].

As mentioned above, being online can raise awareness for the positive experiences other people make. At the same time, interactions on social

media increase the awareness of negative instances in other peoples lives. This awareness of other people causes emotional distress, as people are empathetic to other people. Constantly being confronted with negative experiences and expecting oneself to care about them is detrimental to ones own wellbeing [16].

Further, being aware of other peoples experiences can cause fear of missing out (FoMO). FoMO is one of the main ways of unhealthy attachments to social media. People suffering from FoMO have trouble with not being online constantly for fear of lost opportunities, experiences and not seeing and reacting to content soon enough [2], especially if the information and interactions are only temporarily available. FoMO is further associated with the desire to be perceived as popular by others. Excessive smartphone use and especially notifications [35] can increase the chances of developing FoMO [13]. FoMo is associated with negative wellbeing, as it leads to anxiety, problematic sleep patterns, lack of concentration, decreased control over ones own emotions and dependence on social media for gratification [2].

Magnified social pressure Another social factor that can contribute to digital stress is social pressure. Interactions with technology can magnify already existing social pressures, influencing behaviour, as well as the portrayal of oneself. Going hand in hand with the awareness of how other people portray themselves online, people become more conscious of how they portray themselves. If the way others perceive an individual does not match with the desired image they want to give off, social stress ensues [32]. The pressure to portray oneself in a certain way can thus evolve into a stressor. This type of digital stress is especially prevalent in people who have a high need for social acceptance and elicit people-pleasing behaviours [8]. While societal expectations exist independently of the digital world, the digital landscape provides opportunities for constant confrontation with social pressures like expectations and norms of how one should behave and the way one should deal with daily life [4]. Social pressure can lead to the pattern where users try to conform their usage of social networks to the usage patterns of their friends, which is a condition under which social network use leads to stress [28]. Further, social pressure leads to digital overuse and the development of smartphone addiction [4, 8], both of which are detrimental to wellbeing.

Maintaining relationships online can cause stress, as individuals are expected to effectively and constantly use digital media to communicate and keep in touch [4, 16]. Dealing with social demands e.g. providing social support to friends [28] is part of modern daily life [4]. Constantly being expected to be in contact through the means of mobile devices can lead to fatigue and chronic stress [19]. This societal pressure is further amplified through notifications. Instead of only checking messages when the individual has the desire or need to do so, habitual checking behaviour ensues. This unhealthy notification-checking behaviour driven by the reminders to communicate with others is a source of heightened anxiety levels [8].

Further, Weinstein et al. defined socio-digital challenges, called the Type 2 stressors, that affect teenagers trying to navigate and maintain close, real-life relationships in the digital world [43, 44]. One of these stressors is the feeling of being smothered by the constant availability to other people, who the individual may want to be connected to. Aside from this, other people may pressure individuals to give them access to their accounts or even intimate pictures for the sake of proving their trust. Dealing with requests like this can cause stress. Lastly, teenagers can experience stress when their privacy is violated through their contacts breaking into their online accounts and/or devices without permission [43].

Another main contributing factor to digital stress is the societal expectation of immediacy. While analogue means like letters can take days to arrive at the recipient, mobile instant messages and emails arrive right after being sent. The speed of message transmission can evolve into an issue when combined with the societal expectation that people are online all the time and thus reachable [35]. The expectation of constant availability can lead the sender to expect a quick response from the recipient of the message [7, 34]. The prospect of immediate responses is detrimental to the relationship for both sender and receiver. The sender, who initiated the conversation when it was convenient for them, may not receive an answer in a timely fashion, thus violating their expectations. The receiver of the message feels the pressure to comply with the senders' expectation of an immediate response [34] under the prospect of possibly hurting the relationship and dampening their own popularity in case of a delayed response [2]. In addition to already existing social pressure through notifications, instant messaging services often display when someone was last online as a predictor for users attentiveness. Aside from being useless in predicting whether the message will actually be read soon, this feature creates further social pressure [34].

Aside from instant messages, emails cause stress. While they are relevant to obtain critical information in the workplace, social norms to respond as fast as possible [29] can lead to overuse, stress and at the end of the day even burnout [31]. Further, internet users reportedly find it stressful and exhausting when they have a lot of messages left to reply to [5]. Not only direct messages have the effect of channelling social pressure, but expecting oneself to react to social media posts in a timely fashion can also result in stress [42].

Negative interactions Using technology, directly or indirectly negative interactions with other people can occur. Socialisation itself can be a stressor [5, 43], which is why it can contribute to social digital stress. Social media can give us even more opportunities for socialization. This can have both positive and negative consequences. The negative effects are especially prominent for interactions that are in itself destructive like cyberbullying or anonymous hate.

Being the victim of hostile interactions online can cause stress, which we can thus classify as social-digital stress. Cyberbullying, as the name suggests, is bullying that takes place in the digital space. It is characterised as intentional, aggressive and harassing actions against an individual intended to harm and socially exclude them and display a power imbalance. These are executed repeatedly over time using digital mediums [22, 26]. The digital nature of the interactions enables cyberbullying through anonymity and possible large audiences [12]. Cyberbullying has been proven to negatively affect mental health for a long time [40]

There are three kinds of directly hostile interactions that act as digital stressors, called Type 1 stressors: personal attacks, public shaming and humiliation, and impersonation. While personal attacks are directed at the victim through private messages, public shaming refers to the action of posting mean comments about someone publicly or even sharing private information like nude photos to embarrass the targeted individual. Impersonation is another way to harm someone. Here, the perpetrator pretends to be the victim online through hacking the attacked persons account or creating fake accounts and then mocking, slandering or embarrassing them [43].

Negative interactions online are not reserved for social media and personal messaging. Multiplayer online games, especially Multiplayer Online Battle Arena (MOBA) like League of Legends and Dota, are a breeding ground for toxic behaviour. In contrast to cyberbullying, toxic behaviour takes place over shorter periods and as a direct reaction to negative events like the own failure in the game [22]. For example, losing a game of League of Legends may result in teammates insulting each other. The "high competitiveness, [...], lack of immediate repercussions, and [...] sense of anonymity" [22] in the MOBA genre leads to a toxic social environment, that causes frustration and generally bad social relations amongst player, which in turn leads to stress and a decrease in wellbeing [22].

As discussed, there are various social factors that contribute to or cause digital stress. These social factors, which we group into the categories constant awareness of other people, social pressure and directly negative interactions, are detrimental to wellbeing and therefore need to be coped with.

3 Coping with Social-Digital Stress

In order to understand how people can cope with social-digital stressors, we now provide an overview of various applicable coping strategies.

Generally speaking, raising awareness for the way people use their smartphone [8] through self-reflection is a strategy for coping with digital stress. This self-reflection can be achieved through conscious confrontation with and manipulation of the stressor e.g. through visualisation [35]. Visualisations of stressors can lead to more open conversations and consideration of other users [35]. The increased awareness for smartphone usage ties in with the coping mechanisms distraction and avoidance of stressors [42, 45]. The smartphone itself can be used as a distraction from socially stressful interactions [45]. Consumed media can act as an escape from the stressful reality [16]. In an experiment by Yau et al. [47] the passive use of phones as a distraction had the same positive effects on mood as texting friends and acquiring social support that way. Even the mere prospect of being able to use a phone in case of stressful interactions without actually using it can act as a buffer from harmful experiences, as well as social exclusion [18]. Internet use is an effective coping strategy in this way but has to be carefully administered as not to replace other coping strategies, as this may lead to internet overuse [45]. While avoiding or ignoring a stressful situation is a common recommendation teens give each other concerning socio-digital stressors [44], it is an emotion-focused strategy and does therefore not actually help improve the stressed individuals situation [44].

Coping with the constant awareness of other people Certain skills, like the development of high self-compassion, can help in the fight against social stressors [21]. However, the list of active coping strategies for detrimental effects on mental health of being constantly aware of positive and negative events in other people's lives is quite short. The area with the most research focus is how to cope with the Fear of Missing Out on digital events. In order to deal with FoMO, Alutaybi et al. recommend socio-technical approaches that include social media design, self-talk, as well as checklists and strive to improve online resilience [2]. In particular, they talk about how social media design can be enriched with technical elements like auto-replies, filtering and setting statuses to act as FoMO countermeasures. In addition to this, they recommend that users build digital resilience to be empowered to utilise problem-solving coping strategies. Learning about how FoMO occurs, managing ones own expectations and self-regulation can play a powerful role in combating FoMO. A tool combining the elements of the socio-technical approach is FoMO-R (Fear of Missing Out Reduction) [2].

Coping with social pressure As discussed in Section 2.2, the expectation of immediacy created through digital means of communication can be detrimental to wellbeing.

The proposed coping strategies for dealing with the pressure to respond to messages immediately involve technical approaches. There are calls for improved, smarter notifications, and even for the use of AI integrated into smartphones that may protect users [8]. Independent of technological advances, users can cut down their email traffic or resort to batching if there is a lot of email traffic [29]. Different apps have been developed that tackle the social expectations coming from immediacy in different ways. Delaying smartphone notifications is one approach that caused participants of a study to feel less stressed, as it removed potential stressors [35]. Going one step further, the sender of the message could be given the option of whether to send a notification at all, after being briefed on the recipients current status. While recipients would need to privately share and update their status, this approach could lead to a decrease in unimportant and untimely notifications, resulting in less social pressure to reply immediately. From a privacy perspective, it should be noted that in this approach the sender is only informed about the current status reactively if they wrote the recipient a message [7].

The last-seen or last-online status displayed in many messaging apps leads to a heightened feeling of pressure to respond. However, it has little predictive information on actual attentiveness. Depending on the application, these status displays can be manually turned off. A remedy that does not remove the feature completely would be the employment of machine-learning for calculating a more accurate prediction of users current attentive states. If the sender is notified through their smartphone that the recipient is busy and probably won't respond soon, the recipient of the message feels less pressured to respond to the message immediately. Criteria used in identifying the attentiveness could be general screen activity, interactions with the notification centre or even the proximity sensor. These predictions not always being accurate is another positive factor, as it creates room for Butler Lies like 'Sorry! I just saw your message', which can additionally alleviate social pressure [34].

Unrelated to behaviour in response to messages, digital media can mirror social pressure to portray oneself in a certain way. While researchers have uncovered this stressor, specific coping strategies have yet to be proposed. Further, social pressure can be the direct result of interactions with people. Toxic behaviour from real-life acquaintances that transcends to involve the digital realm can be hard to deal with, especially as they are oftentimes related to peoples desire to maintain the relationship even after it evokes stress. Peers instruct teens dealing with such stressors to communicate directly e.g. draw boundaries, express feelings and get a clear understanding of the perpetrators motivations. Further, the problem-focused and active approach of eliminating future stressful encounters by cutting ties altogether e.g. through breaking up and/or stopping communication with anyone involved in the digital stress completely are common coping recommendations [44].

Coping with negative interactions Active problem-solving strategies like de-escalation are the most effective for reducing offline bullying [33]. This reflects the way teenagers cope with cyberbullying and other directly hostile interactions online. The most common recommendation for coping with hostility online is to get help, which is an active and problem-focused approach. Further, confronting the person involved in the digital stressor, as well as utilising digital solutions are recommendations for coping with digital stress related to negative interaction. Digital solutions involve the changing of privacy settings and passwords or blocking and unfollowing the hostile person [44].

4 Using Social Strategies to Improve Wellbeing

Digital interactions are not purely stressful and detrimental to wellbeing; they can actually be used to actively improve mental wellbeing. While the previous section focused on coping with stressful media use, the upcoming section focuses on coping (with life) through media use. More specifically, we will present research projects that devote themselves to utilising the social aspects of digital media to improve wellbeing.

Technology has become a vital tool for coping with stress and other negative emotions in everyday life [38, 45]. Mobile devices, internet services like social media, TV and games provide a vast range of different options for coping with stress [45], ranging from a plenitude of mindfulness apps to social support on social media [38]. Media can be seen as a facilitation, so to say additional tool or dimension for coping strategies that are possible in an analogue setting. Alternatively, media use can be viewed as a distinct coping strategy [45].

Social support online One of the most prevalent positive effects of being active online is the opportunity for getting social support. Social support describes both verbal and nonverbal communication that alleviates stress and uncertainties. Generally, social support can be divided into five categories: information support, tangible assistance, esteem support, network support, and emotional support [1]. Social interactions can help with solving specific issues actively. Aside from this, the feeling of being socially connected alone can have positive effects on mental health [10].

Social media does not contain the word social without reason. Online communities on social media and forums provide a space for acquiring social

support [1, 31, 38, 44, 45]. Being a part of online communities creates the opportunity to share feelings, express emotions and receive support and specific advice on coping in return [31]. The feeling of belonging to a specific digital community is related to social wellbeing [3]. Online activity allows people to connect to others who have the same experiences as them, which is ideal for coping with specific stressors like mourning. Aside from specific issues, social support contributes to general mental health support [38]. Reaching out for social support online is considered a problem-focused coping style [46].

Social media can further be used for self-disclosure. Self-disclosure refers to the practice of posting about oneself and has a mediating effect on stress and specifically the stressor loneliness. While self-disclosure is considered an active coping strategy offline, it is a passive strategy when utilised online [15]. Even when nobody actually listens, using social media in this way is beneficial to wellbeing.

Aside from utilising online communities for the provision of social support, technology can be used to personally connect to people who can alleviate stress by using mobile phones [45] to call or text [47]. In mobile environments, staccato social support is especially prevalent. Staccato social support is defined through especially short and fast supporting interactions [1]. Mobile communication can buffer the effects of stressors and thus improve well-being [47], as people have the power to reach out directly after stressful situations and get support. It acts as an extension of human communication, surpassing the need for being in the same place at the same time [5, 47]. This is especially useful right now, as meeting people face-to-face has become problematic. Communication with digital tools can even strengthen relationships that originated outside of the digital space [11]. People who text friends report better moods and portray lower stress levels than those who use their phones to watch videos or who do nothing at all [47].

Using technology as a means for connecting with other people can have positive effects on wellbeing. Ultimately, the positive effects of social relations may even outweigh the negative effects of stress [16].

The Smartphone as a Companion As previously discussed, smartphones enable us to connect to other people, thus being a beneficial tool for wellbeing. However, the use of the smartphone itself, independently of who we connect with through it, can act as a buffer between users and negative experiences.

Nowadays, people rarely leave the house without their smartphones. The device provides constant opportunities for digital experiences. While the access to smartphones enables people to be in incessant contact with other people to e.g. obtain social support as previously discussed, the smartphone itself can act as a form of companion for alleviating stress. Carolus et al. propose that we constantly interact with smartphones in a way that resembles relationships to other humans and calls this digital companionship. The human-human like tie between users and their phones is the result of the way the device communicates. Humans subconsciously react to the communication in a way shaped by social rules and norms that apply to human communication. The model of Carolus et al. evaluates relationships to smartphones through the characteristics psychological closeness, trust and preoccupation. In turn, the characteristics trust and preoccupation have effects on stress and coping with stress. Regarding closeness, smartphones are the most relevant device and even more relevant than some humans. While closeness itself has little effect on stress, social relationships are inherently important for coping with stress. Just like human-human relationships, human-smartphone relationships can be both a source of stress and a resource for coping with stress. Higher levels of trust, defined through reliability and dependability, cause the relationship to be less stressful and have more associations with coping with stress, not only for companionship with humans but also with devices like smartphones. Thus, if a trust relationship can be established, the device can be useful in reducing stress [5].

Going one step further, the companion-like traits of smartphone devices can be purposefully used to create stress management applications. Technology is a great way for collecting information like emotions to create intelligent digital companions that can support people dealing with stress-related exhaustion syndrome in a personalised way. As previously discussed, stress and coping differ on an individual basis. Thus, using socially intelligent and adaptive companions that personalise their responses based on collected data opens possibilities that can be integrated in interface design of devices which we carry around with us every day anyways. The digital companion can collect emotion data communicated through tangible and non-verbal means for using them in the context of stress management. This information can thereafter be used by researchers and designers to effectively counteract stressors [20].

However, while companionship with devices can have positive effects on wellbeing, it does not replace real human contact. Further, unhealthy attachments should be avoided as these are detrimental to wellbeing [8].

Creating positive experiences In order to improve wellbeing, the previously discussed methodologies lie in the way we use our smartphones or the way we utilise technology for acquiring social support. Another approach to improve wellbeing is to specifically design technology to foster positive experiences. The following projects demonstrate how this is being done with the social nature of technology in mind.

Rodgers et al. developed a device for communicating wellbeing called MoodCloud. MoodCloud is composed of an app and an ambient display utilising colours. Rodgers et al. argue that wellbeing is socially constructed to some extend, as our expectations of what wellbeing is are the result of interactions and societies norms. On this premise, MoodCloud is used to express, discuss and determine wellbeing collaboratively, inciting reflection. At the core, wellbeing is seen as an interaction [37].

Technology can further be used to incite acts of kindness through gamified means. While being on the receiving end of kindness is nice, intentionally executing acts of kindness -positive activities like generosity or gratitude- improves happiness and wellbeing. To incite positive activities, Ciocarlan et al. suggest the use of persuasive games. Not only do the encouraged acts of kindness have positive effects on wellbeing, but the digital persuasive game also plays a role in tending to mental illness, as individuals can benefit from early detection of symptoms and access support from afar [9].

5 Discussion & Conclusion

Through the means of a literature review, this paper aims to shed light on the social component of digital stress and highlight what role social factors play in making the digital experience both more and less stressful.

We identify three major groups of social-digital stressors: constant awareness of other people, social pressure and directly negative interactions. Through means of technology, people communicate both positive and negative experiences of their lives. We hypothesise that this constant awareness of others is detrimental to wellbeing due to the sheer amount of other people individuals are regularly in contact with. The primary strategies for dealing with the detrimental effects of the constant awareness of other people are emotion-focused, striving for changes in the troubled individual. Researchers suggest and try to facilitate the development of self-compassion and online-resilience, as these are associated with better management of stressful situations. However, technology should work for people - otherwise, it doesn't work. Ubicomp can play an important part in reducing digital stress by focusing research efforts on the adaption of technology to decrease stressors and helping users deal with them in a problem-focused way.

In contrast, researchers suggest technical and problem-focused approaches for coping with social pressure. The social pressure that exists in the real world can be magnified in the digital landscape. Without completely changing social structures, technical design changes in the way attentiveness is being communicated can alleviate people of the stress to behave a certain way to e.g. respond to messages immediately. The social pressure from others to portray oneself a certain way has yet to receive adequate attention and specific coping suggestions.

Researchers have accumulated a catalogue of problem-focused strategies to cope with directly negative social interactions online. Adapting the digital landscape through removing features like anonymity, which is known to incite hostility, could further prove to be a reasonable approach.

Aside from social stressors prevalent in the digital landscape, we identify social features that improve wellbeing. The provision of social support online is especially well researched and is actively used for coping with real-life stressors. Further, even the possibility of access to digital devices can buffer stressful experiences. Communicating more and more positively can additionally improve wellbeing. Thus, technology should support these types of behaviour.

The findings suggest that social factors play an important role in empowering digital stressors. However, only few of the papers on digital stress covered both issue and solution (e.g. [34]). We call for future work to identify coping strategies for concrete stressors to enable suitable coping strategies for different stressful situations. Aside from this, we highlight how digital media is already being used to reduce stress and propose that future work makes use of this by incorporating it in research on coping with digital stress.

References

1 Phil Adams, Eric P.S. Baumer, and Geri Gay. Staccato social support in mobile health applications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 653–662, New York, NY, USA, Apr-May 2014. Association for Computing Machinery. doi:10.1145/2556288.2557297.

2 Aarif Alutaybi, Dena Al-Thani, John McAlaney, and Raian Ali. Combating fear of missing out (fomo) on social media: The fomo-r method. *International Journal of Environmental Research and Public Health*, 17(17), August 2020. doi:10.3390/ijerph17176128.

3 Moritz Büchi, Noemi Festic, and Michael Latzer. How social well-being is affected by digital inequalities. *International Journal of Communication*, 12:3686–3706, 2018. URL: http://ijoc.org/index.php/ijoc/article/view/8780.

4 Moritz Büchi, Noemi Festic, and Michael Latzer. Digital overuse and subjective well-being in a digitized society. *Social Media + Society*, 5, 2019. doi:10.1177/2056305119886031.

5 Astrid Carolus, Jens F. Binder, Ricardo Muench, Catharina Schmidt, Florian Schneider, and Sarah L. Buglass. Smartphones as digital companions: Characterizing the relationship between users and their phones. *New Media & Society*, 21:914–938, 2019. doi:10.1177/1461444818817074.

6 Marta E. Cecchinato, John Rooksby, Alexis Hiniker, Sean Munson, Kai Lukoff, Luigina Ciolfi, Anja Thieme, and Daniel Harrison. Designing for digital wellbeing: A research & practice agenda. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI EA '19, pages 1–8, New York, NY, USA, 2019. Association for Computing Machinery. doi:10.1145/3290607.3298998.

7 Hyunsung Cho, Jinyoung Oh, Juho Kim, and Sung-Ju Lee. I share, you care: Private status sharing and sender-controlled notifications in mobile instant messaging. In *Conference Companion Publication of the 2020 on Computer Supported Cooperative Work and Social Computing*, CSCW '20 Companion, pages 13–17, New York, NY, USA, 2020. Association for Computing Machinery. doi:10.1145/3406865.3418571.

8 Elizabeth F. Churchill. Designing for digital well-being. *Interactions*, 27(1):26–28, December 2019. doi:10.1145/3373362.

9 Ana Ciocarlan, Judith Masthoff, and Nir Oren. Kindness is contagious: Study into exploring engagement and adapting persuasive games for wellbeing. In *Proceedings of the 26th Conference on User Modeling, Adaptation and Personalization*, UMAP '18, pages 311–319, New York, NY, USA, 2018. Association for Computing Machinery. doi:10.1145/3209219.3209233.

10 Sheldon Cohen, Tom Kamarck, and Robin Mermelstein. A global measure of perceived stress. *Journal of Health and Social Behavior*, 24(4):385–396, 1983. URL: https://www.jstor.org/stable/2136404.

11 Katie Davis. Friendship 2.0: Adolescents' experiences of belonging and self-disclosure online. *Journal of Adolescence*, 35:1527–1536, December 2012. doi:10.1016/j.adolescence.2012.02.013.

12 Katie Davis, David P. Randall, Anthony Ambrose, and Mania Orand. 'i was bullied too': stories of bullying and coping in an online community. *Information, Communication & Society*, 18(4):357–375, 2015. doi:10.1080/1369118X.2014.952657.

13 Jon D. Elhai, Jason C. Levine, Robert D. Dvorak, and Brian J. Hall. Fear of missing out, need for touch, anxiety and depression are related to problematic smartphone use. *Computers in Human Behavior*, 63:509–516, 2016. doi:10.1016/j.chb.2016.05.079.

14 Gallup 2019 global emotions report. Gallup, 2019. Accessed January 3, 2021 from https://www.gallup.com/analytics/248909/gallup-2019-global-emotions-report-pdf.aspx.

15 Elodie Gentina and Rui Chen. Digital natives' coping with loneliness: Facebook or face-to-face? *Information & Management*, 56, 2019. doi:10.1016/j.im.2018.12.006.

16 Keith N. Hampton, Weixu Lu, and Inyoung Shin. Digital media and stress: the cost of caring 2.0. *Information, Communication & Society*, 19(9):1267–1286, 2016. doi:10.1080/1369118X.2016.1186714.

17 Tracy Bennett Herbert and Sheldon Cohen. Stress and immunity in humans: a meta-analytic review. *Psychosom Med*, 55(4):364–379, Jul-Aug 1993. doi:10.1097/00006842-199307000-00004.

18 John F. Hunter, Emily D. Hooker, Nicolas Rohleder, and Sarah D. Pressman. The use of smartphones as a digital security blanket: The influence of phone use and availability on psychological and physiological responses to social exclusion. *Psychosomatic Medicine*, 80:345–352, May 2018. doi:10.1097/PSY.0000000000000568.

19 Kayoko Ikeda and Kazutoshi Nakamura. Association between mobile phone use and depressed mood in japanese adolescents: a cross-sectional study. *Environ Health Prev Med*, 19(3):187–193, 2014. doi:10.1007/s12199-013-0373-3.

20 Monika Jingar and Helena Lindgren. Tangible communication of emotions with a digital companion for managing stress: An exploratory co-design study. In *Proceedings of the 7th International Conference on Human-Agent Interaction*, HAI '19, pages 28–36, New York, NY, USA, 2019. Association for Computing Machinery. doi:10.1145/3349537.3351907.

21 Rebecca Keyte, Lauren Mullis, Helen Egan, Misba Hussain, Amy Cook, and Michail Mantzios. Self-compassion and instagram use is explained by the relation to anxiety, depression, and stress. *Journal of Technology in Behavioral Science*, 2020. doi:10.1007/s41347-020-00186-z.

22 Bastian Kordyaka. Digital poison - approaching a theory of toxic behavior in moba games. In *Thirty Ninth International Conference on Information Systems*, ICIS 2018, Atlanta, Georgia, USA, dec 2018. Association for Information Systems. URL: https://aisel.aisnet.org/icis2018/behavior/Presentations/18/.

23 Saewon Kye, Junhyung Moon, Juneil Lee, Inho Choi, Dongmi Cheon, and Kyoungwoo Lee. Multimodal data collection framework for mental stress monitoring. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, UbiComp '17, page 822–829, New York, NY, USA, 2017. Association for Computing Machinery. doi:10.1145/3123024.3125616.

24 R. S. Lazarus. Coping theory and research: Past, present, and future. *Psychosomatic Medicine*, 55:234–247, 1993. doi:10.1097/00006842-199305000-00002.

25 Richard S. Lazarus and Susan Folkman. Transactional theory and research on emotions and coping. *European Journal of Personality*, 1:141–169, 1987. doi:10.1002/per.2410010304.

26 Nathaniel Levy, Sandra Cortesi, Urs Gasser, Edward Crowley, Meredith Beaton, June Casey, and Caroline Nolan. Bullying in a networked era: a literature review. *Berkman Center Research Publication*, 2012-17:1–61, 2012. doi:10.2139/ssrn.2146877.

27 Germaine M Buck Louis, Kirsten J Lum, Rajeshwari Sundaram, Zhen Chen, Sungduk Kim, Courtney D Lynch, Enrique F Schisterman, and Cecilia Pyper. Stress reduces conception probabilities across the fertile window: evidence in support of relaxation. *Fertil Steril*, 95(7):2184–2189, Jun 2011. doi:10.1016/j.fertnstert.2010.06.078.

28 Christian Maier, Sven Laumer, Christoph Weinert, and Tim Weitzel. The effects of technostress and switching stress on discontinued use of social networking services: a study of facebook use. *Information Systems Journal*, 25, March 2015. doi:10.1111/isj.12068.

29 Gloria Mark, Shamsi T. Iqbal, Mary Czerwinski, Paul Johns, Akane Sano, and Yuliya Lutchyn. Email duration, batching and self-interruption: Patterns of email use on productivity and stress. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 1717–1728, New York, NY, USA, 2016. Association for Computing Machinery. doi:10.1145/2858036.2858262.

30 Mental health statistics: stress. Mental Health Foundation, 2018. Accessed January 4, 2021 from https://www.mentalhealth.org.uk/statistics/mental-health-statistics-stress.

31 Nour Mheidly, Mohamad Y. Fares, and Jawad Fares. Coping with stress and burnout associated with telecommunication and online learning. *Frontiers in Public Health*, 8:672–679, 2020. doi:10.3389/fpubh.2020.574969.

32 Neema Moraveji and Charlton Soesanto. Towards stress-less user interfaces: 10 design heuristics based on the psychophysiology of stress. In *CHI'12*, CHI' 12, pages 1643–1648, New York, NY, USA, May 2012. Association for Computing Machinery. doi:10.1145/2212776.2223686.

33 Melissa M. Mahady Wilton Wendy M. Craig Debra J. Pepler. Emotional regulation and display in classroom victims of bullying: Characteristic expressions of affect, coping styles and relevant contextual factors. *Social Development*, 9:226–245, May 2000. doi:10.1111/1467-9507.00121.

34 Martin Pielot, Rodrigo de Oliveira, Haewoon Kwak, and Nuria Oliver. Didn't you see my message? predicting attentiveness to mobile instant messages. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 3319–3328, New York, NY, USA, 2014. Association for Computing Machinery. doi:10.1145/2556288.2556973.

35 Romina Poguntke. *Understanding Stress Responses Related to Digital Technologies*. PhD thesis, Institut für Visualisierung und Interaktive Systeme der Universität Stuttgart, 2020.

36 Kamleshun Ramphul and Stephanie G Mejias. Is "snapchat dysmorphia" a real issue? *Cureus*, 10, March 2018. doi:10.7759/cureus.2263.

37 Shannon Rodgers, Bernd Ploderer, Brittany Maloney, and Jason Hang. Designing for wellbeing-as-interaction. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI EA '19, pages 1–6, New York, NY, USA, 2019. Association for Computing Machinery. doi:10.1145/3290607.3312901.

38 Mattias Rost, John Rooksby, Alexandra Weilenmann, Thomas Hillman, Pål Dobrin, and Juan Ye. Mobile wellbeing. In *Proceedings of the 9th Nordic*

Conference on Human-Computer Interaction, NordiCHI '16, New York, NY, USA, 2016. Association for Computing Machinery. doi:10.1145/2971485.2987676.

39 MD, FACP, FACR William C. Jr. Shiel. Medical definition of stress. MedicineNet. accessed December 19, 2020 from https://www.medicinenet.com/stress/definition.htm.

40 Peter K. Smith, Jess Mahdavi, Manuel Carvalho, Sonja Fisher, Shanette Russell, and Neil Tippett. Cyberbullying: its nature and impact in secondary school pupils. *The Journal of Child Psychology and Psychiatry*, 49:376–385, 2008. doi:10.1111/j.1469-7610.2007.01846.x.

41 Mai-Ly N. Steers, Robert E. Wickham, and Linda K. Acitelli. Seeing everyone else's highlight reels: How facebook usage is linked to depressive symptoms. *Journal of Social and Clinical Psychology*, 33(8), October 2014. doi:10.1521/jscp.2014.33.8.701.

42 Monideepa Tarafdar, Christian Maier, Sven Laumer, and Tim Weitzel. Explaining the link between technostress and technology addiction for social networking sites: A study of distraction as a coping behavior. *Information Systems Journal*, 30:96–124, January 2020. doi:10.1111/isj.12253.

43 Emily C. Weinstein and Robert L. Selman. Digital stress: Adolescents' personal accounts. *New Media & Society*, 18:391–409, 2016. doi:10.1177/1461444814543989.

44 Emily C. Weinstein, Robert L. Selman, Sara Thomas, Jung-Eun Kim, Allison E. White, and Karthik Dinakar. How to cope with digital stress: The recommendations adolescents offer their peers online. *Journal of Adolescent Research*, 31:415–441, 2016. doi:10.1177/0743558415587326.

45 Lara Wolfers and Frank Schneider. Using media for coping: A scoping review. *Communication Research*, July 2020. doi:10.1177/0093650220939778.

46 Kevin B. Wright and Stephen A. Rains. Weak tie support preference and preferred coping styles as predictors of perceived credibility within health-related computer-mediated support groups. *Health Communication*, 29:281–287, 2014. doi:10.1080/10410236.2012.751084.

47 Joanna C. Yau and Tao Yi Reich, Stephanie M.and Lee. Coping with stress through texting: An experimental study. *Journal of Adolescent Health*, pages 1–7, August 2020. doi:10.1016/j.jadohealth.2020.07.004.

Unimodal and Multimodal Augmented Feedback for Physical Exercise

Miriam Metz

Ludwig-Maximilians-Universität München, München, Deutschland
metz.miriam@campus.lmu.de

Abstract

Physical exercise is an important tool to gain and maintain health and fitness. Feedback can help trainees to learn about the correct execution technique of exercises, increase the effectiveness of their training, and keep engagement and motivation levels high. Access to such feedback from a human coach is often limited but technology could step in and provide individual feedback at scale. Yet the questions remain how feedback technologies need to be designed to prove beneficial and which feedback modalities are most effective for the application in physical exercise. This work presents an overview of recent research projects and aims at comparing visual, auditory, haptic, and multimodal feedback. It was found that every modality comes with a specific set of implications, strengths, and weaknesses which is why the design of feedback systems has to take the user's context, skill level and needs into account. While it is challenging to compare modalities in regards to effectiveness, differences in cognitive load and perceptual precision are elaborated. Further research can help to develop more explicit design guidelines for effective feedback technologies and thereby facilitate a fuller realization of their potential.

2012 ACM Computing Classification Human-centered computing → Human computer interaction (HCI)→ Interaction techniques;

Keywords and phrases Visual feedback; Auditory Feedback; Haptic Feedback; Multimodal Feedback; Physical Exercise

1 Introduction

With the COVID-19 pandemic, daily life for many people around the world has fundamentally changed. In many regions across the globe, public life and social habits have been widely shut down by various government-imposed restrictions. In particular, fitness and sports routines were abruptly interrupted as gyms, yoga studios, or public pools were closed for several months. With limited access to group fitness classes, physiotherapy, or personal trainer sessions,

Cite as: Miriam Metz. Unimodal and Multimodal Augmented Feedback for Physical Exercise. In *6th Seminar on Ubiquitous Interaction (UBIACTION 2021)*. Editors: F. Lang, R. Welsch, L. Haliburton, F. Draxler, S. Feger, S. Villa, M. Hoppe, P. Knierim, V. Mäkelä, A. Schmidt. February 11, 2021. Munich, Germany. pp. 7:1–7:24.

sedentary times in the day-to-day life during quarantine increased, ultimately multiplying health risks.

To maintain health and fitness, many have turned to remote and digital training alternatives. Common examples range from fitness apps (e.g. Freeletics [2], Nike Training Club [5]) and physiotherapy apps (e.g. Kaia Health [4], Vivira [8]) to Youtube videos (e.g. Blogilates [1]). Local gyms and personal trainers have started to offer remote consultation or fitness classes via video conferencing technologies [3]. Additionally, the home workout gear market is booming (e.g. Peloton [6]). Overall, a shift towards exercising at home instead of at dedicated facilities has happened.

Even beyond restrictions during a pandemic, exercising at home offers several benefits. It is *time-saving* as trainees don't have to commute to a facility or don't have to wait for potentially occupied equipment. It *saves money* as no membership fees for gyms or studios have to be paid. It allows for *flexible* integration into people's lives as they don't have to adhere to limited opening hours or fixed class schedules. With adequate data security measures in place, there is an opportunity for *more privacy* than when working out in the gym or a group. Therefore, trainees can exercise without feeling *self-conscious*. Lastly, a home workout can be completed at the trainee's individual pace [7].

Yet, working out at home also comes with certain problems and risks. The lack of guidance usually provided by an instructor can lead to *incorrect execution* of physical exercises. This, in turn, can lead to *ineffectiveness* of the exercises or even *injury or re-injury* [34, 33]. Uncontrolled exercise can further cause a state of *overtraining* and exhaustion in the trainees and the missing social context, as well as the lack of indirect accountability that normally comes from that, can *reduce motivation* and commitment to training.

In this context, advances in technology (e.g. through wearable technologies, computer vision, and machine learning) and design (e.g. novel forms of visualization, feedback, or interaction) could help to lower these risks by providing trainees with automated, augmented feedback, guiding them through workouts, correcting movements and posture, counting repetitions and holding times and encouraging trainees to keep exercising — scalably without the need to have a human instructor present. Thanks to technology- and AI-powered feedback systems, thousands of people can receive custom individual feedback simultaneously.

The goal of this work is to review recent HCI literature that focuses on supporting individual, physical exercising with technologically enabled feedback. A special focus is set on analyzing different feedback modalities (visual, auditory, haptic, and multimodal) and comparing their advantages and disadvantages in the context of individual, physical exercise.

2 Background

2.1 Physical Exercise

Physical exercise is generally defined as "a subset of physical activity that is planned, structured, and repetitive and has as a final or an intermediate objective the improvement or maintenance of physical fitness". Exercise is thereby differentiated from mere physical activity in regards to two aspects: its organized and iterative nature, as well as the goal to "improve or maintain physical fitness" [10].

Physical exercise can take place in a diverse range of settings. It can be performed individually (e.g. going for a run alone) or in a group (e.g. joining an indoor cycling fitness class at the gym). It can be self-guided (e.g. performing a set of pre-defined exercises in the gym) or instructor-guided (e.g. training with a personal trainer or physiotherapist). Finally, physical exercise can be done remotely from home (e.g. doing yoga at home by following a video), outdoors, or on-site at a dedicated facility (e.g. working out at the gym).

Depending on the context and trainee, the focus, goals, and intentions behind physical exercise can differ. In physiotherapy, physical exercise is used for rehabilitative purposes after injury or surgery, to relieve or prevent pain and restore strength and range of movement [33]. In a general fitness context, physical exercise is used to establish and maintain physical fitness, usually with health in mind. And in (competitive) sports, physical exercising aims beyond the maintenance of fitness with a stronger focus on performance improvement of both amateur and professional athletes.

All of these areas are unified by one key aspect: the correct technique and execution of the physical exercises. For trainees to learn about this correct technique, a specific set of skills needs to be learned for which trainees engage in practice over time. Training is often moderated by an instructor who guides, evaluates the performance, and provides feedback that is frequently discussed and adapted to the situation [20].

2.2 Physical Exercise in HCI

Over the last ten years, interest in sports and physical exercise has also increased in the HCI community. One important milestone was the creation of a special interest group for sports [25]. According to Vidal et al., this interest is based on different trends such as a performance-focused quantified-self movement, developments in home training systems, and advances in sensing technologies and wearables [37]. The research focus lies on aspects such as creating and maintaining engagement [38], supporting social and group experiences [26], and real-time performance evaluation and technique correction. Specifically for engagement, further research areas have emerged that investigate physical activity in play and games, exertion games (short "exergames"), and the use of gamification [24].

2.3 Feedback Theory

In the context of physical exercise and sports, feedback can be defined as information conveyed during or after the execution of a movement. In contrast to feed*back*, the term feed*forward* refers to information on how to complete future movements [33].

Feedback can be classified according to different dimensions: The terms *intrinsic or internal feedback* and *augmented or extrinsic feedback* differentiate the source of the feedback. Intrinsic feedback describes sensory afferences - the signals that the brain receives from the body. This type of feedback is always present during physical activity. Augmented feedback is defined as information that cannot be perceived without an external source. It is provided by a trainer or a display of any modality [31].

Concurrent feedback refers to information that is provided in real-time during an exercise, and *terminal feedback* refers to information provided after an exercise [31].

Positive reinforcing feedback is appreciative and motivational, inspires better performance and continued engagement, whereas *negative critical feedback* is mainly informative and aims to support the trainee in correcting errors and getting closer to the target performance. Both feedback types can have "informational content which can either be evaluative (e.g. good, bad, right, wrong) or quantitative (e.g. too fast, too slow, too less)" [28].

Another categoric separation is based on whether or not the information provided via feedback includes an interpretation. Different terminology is used to specify this difference, sometimes to highlight a particular aspect. *Uninterpretative feedback* is often used synonymously with *open-ended feedback* and provides the trainee with quantitative information [34]. Another term for this type is *non-self-referential feedback* which is used in contrast with *self-referential feedback* where the latter refers to an interpretation that compares the current performance of a trainee to a previous performance of that same trainee [30]. *Interpretative feedback* is also referred to as *evaluative feedback* or as mentioned before *self-referential feedback*. This type puts quantitative measurements into context through comparison with benchmarks, the target performance, or previous performances [34].

Two further important design aspects when applying feedback in a system are the *frequency* and *modality* in which it is provided. For frequency, an option is to preplan exactly how often feedback is given, for example after fixed

time intervals. It can also be changed randomly or systematically reduced reflecting the trainee's improvements. This way, the frequency can send an implicit message about training progression [30]. Feedback can be transported using different modalities such as visual, auditory, haptic, or - by combining several modalities - multimodal displays. Depending on the context of use and the trainee's condition, some feedback modalities might be more effective or accessible than others. Feedback modalities will be further discussed in the upcoming chapters.

3 Technology-enabled Augmented Feedback for Physical Exercise

This section starts with a brief overview of the prerequisites of providing technology-enabled feedback to trainees, namely sensing relevant bodily signals, defining the target state and modeling the user state. Subsequently, the application of feedback in physical exercise is discussed by looking at the field through the lens of feedback modalities.

3.1 Prerequisites

The process of providing a trainee with feedback during physical exercise follows several recurring phases. Initially and often continuously, a set of signals relevant to the performance or technique is captured using suitable sensing technology. Based on the data, a user model can be generated, kept up to date, and optionally compared to a defined target model. The derived information can be interpreted by analyzing the target-actual difference and directed to the user as feedback. These phases can be repeated in order to continuously provide new feedback. The following sections elaborate on the three phases: sensing, defining the target state, and user modeling.

3.1.1 Sensing

The foundation for enabling technology-provided feedback in the context of physical exercising is the sensing of relevant signals. Most often, relevant signals measure either bodily processes or technique parameters. Signals based on bodily processes are among others electrodermal activity (also: skin conductance response) or heart rate [34]. Relevant technique parameters vary depending on the exercise. Examples include quantitative measures such as the number of repetitions, holding time, or the overall time spent doing an exercise, as well as qualitative technique descriptors such as body balance [34, 21], the Center of Pressure (CoP) [14], the range, extent, and trajectory of movement, the posture [33], the activation of specific muscles or the pace [21].

Suitable sensing technology has to comply with a set of prerequisites. It needs to be capable of sensing a signal relevant to support physical exercise, it should not constrain body movements and extensive calibration should not be required [14]. Sensors can be divided into stationary and wearable sensors.

Stationary sensors are not attached to the trainee's body but instead placed at a suitable distance. These sensors are mostly used for motion tracking. Advantages are that they are readily available, easy to set up, and relatively low cost. Drawbacks include strict area and positioning requirements, referring to the necessary available space and distance to the user, and low accessibility as they often can't track users with wheelchairs or walkers which might be particularly relevant for application in the context of physiotherapy [33]. Examples of optical tracking systems are Microsoft Kinect, Vicon, and Qualysis. Microsoft Kinect is reported to be less accurate than Vicon and Qualysis whereas the latter are more accurate than IMUs (see paragraph on wearable sensors) [33, 19]. Other projects rely on cameras of mobile devices such as smartphone cameras which users often already own.

Wearable sensors are sensing units that can be worn at the wrist or other limbs as a band, at the chest as a strap, or in the shoes. Apart from optical heart rate sensors, inertial measurement units (IMUs) are commonly used. They usually consist of a multi-axis accelerometer, a gyroscope, and a magnetometer, and can be combined with Bluetooth radios for the transmission of data [14, 23]. Additionally, they are low cost, portable, and therefore often used in research prototypes for activity recognition, gait analysis, and rehabilitation. Single IMUs can be used for simple activity recognition whereas multiple IMUs allows for more accurate posture recognition. Today, they are usually included in smartphones and smartwatches [14]. IMUs can be used in combination with other sensors. Smart insoles, for example, link IMUs with pressure sensors achieving a sensing unit that is unobtrusive while wearing it in the shoes [14, 15].

3.1.2 Defining the Target State

Generally, there are different methods of how to define the target state of an exercise or a set of exercises. Specifying a pre-determined model for what is considered desired and correct allows for the automatization of instruction and feedback [34]. For that, it is necessary to simplify and isolate parameters that are representative of the target state and that can be captured by the sensing technology [21]. This leads to a close relationship between the specification and the used sensors: usually either a desired specification of the target state informs the decision of what sensing technology to use or vice versa [20].

Target models can vary largely in complexity. A simple example would be defining the desired number of repetitions of an exercise, in which case the target state is described merely by a number. A more complex example would be to capture the motion data of an expert performing a plank on video, deriving the relative joint positions and angles between them, and defining the acceptable deviation range for all angles so that the execution can still be considered correct.

3.1.3 Modeling the User State

As user models are usually benchmarked against the target models, their designs are closely related. Therefore, similar to models of the target state, user models diverge greatly in their complexity. The most simple form is storing and regularly updating measured sensor values. Beyond that, user modeling can get more complex, for example when powered by artificial intelligence. A research area of interest is *Human Pose Estimation* which aims at creating a model of the human body by analyzing - often in real-time - input images or videos using computer vision technology. Models can be created from a monocular camera-input resulting in 2D models or with additional cameras or depth sensors resulting in 3D models.

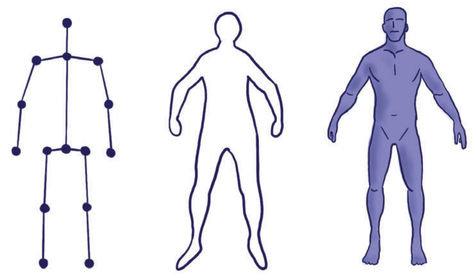

Figure 1 Conceptual sketch of skeleton-based modeling, contour-based modeling and volume-based modeling (from left to right).

There are different approaches to modeling the human body (see Figure 1). Skeleton-based models represent a set of joint locations and the corresponding limb orientation in the form of a graph with vertices and edges. Contour-based models are based on rough contour and shape information of the body. Volume-based models are 3D models based on geometric shapes and meshes. Research making use of the latest developments in deep learning technologies aims at maximizing the accuracy of these models [11].

Project	Exercise(s)	Visual	Auditory	Haptic
Physio@Home [33]	Various	X		
GymSoles [14]	Deadlift, Squat	X		
GymSoles++ [13]	Deadlift, Squat	X		
BalBoa [36]	Handstand	X		
BodyLights [34]	Various	X		
Guerra et al. [19]	Various		X	
Jymmin' [18]	Various		X	
Eyes-free Yoga [29]	Yoga		X	
GymSoles [14]	Deadlift, Squat			X
GymSoles++ [13]	Deadlift, Squat			X
Peeters et al. [27]	Indoor Cycling			X
Bial et al. [9]	Cycling			X
Schneider et al. [30]	Indoor Cycling	X	X	
Verrusio et al. [35]	Various	X		X

Table 1 Reviewed research projects (by project name or authors if no name was given) and the exercises and feedback modalities that were studied.

3.2 Application in Physical Exercise

The following section gives an overview of recent research projects (see Table 1) and important insights related to technology-enabled feedback for physical exercise. The chapter is structured by feedback modality, thus reviewing visual, auditory, haptic, and multimodal feedback.

Overall, visual feedback seems to be the most while haptic feedback the least explored [31]. Possible reasons are that haptic feedback requires additional hardware and that the range of information that can be communicated via haptic feedback is strongly limited. Additionally, the perceptive precision of haptic feedback is lower than with visual or auditory feedback.

3.2.1 Visual Feedback

In the context of physical exercise, visual feedback is particularly relevant as vision dominates other senses regarding the perception of spatial information. This means learning new movements can be facilitated by allowing trainees to observe and imitate based on visual demonstration such as a live demonstration by an instructor or a video recording of such. As opposed to an analog demonstration, digital visualizations can be simplified and abstracted which allows for the reduction of the presented information to the essential and most relevant aspects. Trainees can be presented with visual feedback throughout and after physical exercise. Means for visual displays are currently usually

screens - both static as when mounted to a wall and mobile as with devices such as tablet computers, smartphones, or smartwatches. Alternatively, head-mounted displays, smart glasses (e.g. Google Glass), and virtual reality headsets, as well as projections onto surfaces such as the wall, can be used [31].

With *Physio@Home*, Tang et al. developed a prototype that guides users through physiotherapy exercises. The process is supported by a “Wedge visualization system” providing real-time visual guides, as well as single- or multi-camera views. By visually encoding a set of complex technique parameters using the “Wedge”, the researchers were able to provide the user with a lot of information at once: The completed and incompleted portions of the desired movement (through an arc), the current and target angle of the relevant limb (by showing the numbers), the direction in which the movement has to be completed (using an arrow), the target movement speed (by animating the arc and arrow), as well as some correctional hints regarding the correct limb position (see Figure 2).

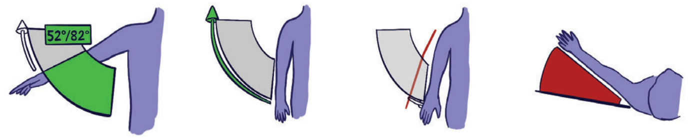

Figure 2 Conceptual sketch of the visual elements of the Wedge visualization: Movement arc and degree numbers, directional arrow, nearest correct arm position, topdown angle deviation (from left to right).

Additionally, the user's movements were tracked, recorded, and displayed on a large TV screen in front of them. In the single-view version, the user saw a frontal perspective of themselves, similar to standing in front of and looking in a mirror. In the multi-view version, a bird's eye perspective of the user was displayed in addition.

The goal of a conducted user study was to understand if different conditions had an impact on the users' accuracy in performing the physiotherapy exercises as well as to learn about the users' preferences. Four combinations were tested in a within-subject study, varying the visualization type (Wedge vs. simple video recording) and the number of camera views (single perspective vs. multiple perspectives).

Overall, users were able to achieve the highest accuracy of correct execution with the combination of multi-camera views and the Wedge visualization. Interestingly, however, there was no clear preference for any of the visualizations among all participants. Despite the higher accuracy that could be achieved with the Wedge visualization, the corrective feedback was often perceived as too strict. Further, participants reported that the Wedge was visually too complex and required a higher degree of attention [33].

What can be learned from this project is that while it is possible to design sophisticated visualizations rich in information that even achieve to improve the accuracy of movement execution, a successful design needs to take the cognitive load, context, and needs of the user into account. Correcting the user to achieve perfect form might lead to higher execution accuracy but it can equally reduce engagement and motivation. Additionally, the results of this project stress the importance of isolating and visualizing the most relevant technique parameters of an exercise. Less seems to have to potential to be more when it comes to visual feedback.

GymSoles by Elvitigala et al. is another project investigating visual feedback that aimed at visualizing the Center of Pressure (CoP) while doing two popular strength exercises: squats and deadlifts. For these exercises, the bodyweight distribution on the feet is an important indicator of correct execution. The GymSoles prototype consisted of pressure-sensitive, smart insoles that could be worn in the shoes as well as an on-screen visualization depicting the CoP as a dot moving between the heel and toes (see Figure 3).[1]

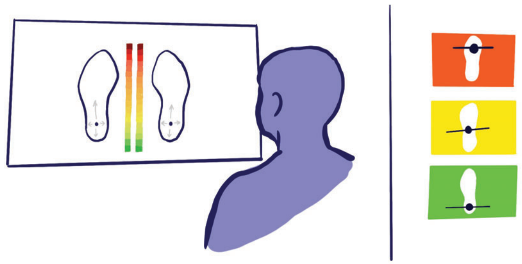

Figure 3 Conceptual sketch of the on-screen visualization of the Center of Pressure (CoP) used for GymSoles (left) as well as the revised and color-coded visualization used for GymSoles++ (right).

As a result of a user study, it was found that for beginner trainees visualizing the CoP significantly improved body posture during the performed exercises. For advanced trainees who were familiar with the correct technique, the feedback turned out to be rather disturbing and detracting attention away from the body and the correct execution. While the visual feedback allowed for a high degree of precision, some participants criticized how the on-screen visualizations required constant monitoring [14]. These results show that the skill level of trainees can have an effect on what kind of feedback is valuable for them. Further, a static installation of a visual display doesn't seem to be ideal for squats and deadlifts as the movements involve different orientations of the head. Lastly, while visual feedback can be effective, it also draws a high level of attention to itself and thereby away from the body and the intrinsic feedback the trainee could perceive instead. In a second iteration, *GymSoles++*, the researchers improved their prototype based on the gathered insights. In particular, the visual feedback was moved from a static, large display in the room to a mobile smart glasses display to reduce the negative impact on the head posture. Additionally, they simplified the visualization by color-coding the difference between the current CoP and the ideal CoP (see Figure 3). This way, the trainee could deliberately decide on when to focus on the feedback and when not: the color codes were still perceptible without explicit visual focus and the exact CoP location could be inspected when focussing on the display [13].

BalBoa is a research project by Vidal et al. that, similar to GymSoles, investigated visual feedback regarding bodyweight distribution. In this case, the physical exercise of interest was the handstand and the goal was to explore how giving feedback on balance could support training for it. The trainee would place their hands on a pressure-reactive surface on the ground and was provided with immediate visual feedback, which would allow them to correct their performance in real-time. The feedback consisted of a visual representation of the hands on the ground with RGB LEDs placed at the corresponding positions of points where the pressure values were measured using the surface. The pressure intensity on the respective pressure point was represented through a nuanced color scale (see Figure 4).

In a preliminary experiment with the prototype, the researchers found that while the immediacy of the feedback was praised, information about the absolute weight distribution on the hands was not perceived as helpful as in the target state of a handstand different areas of the hand need to hold different amounts of weight. That means that with the tried version of the prototype trainees would see a multicolored visualization that needed to be correctly interpreted first, causing a high cognitive load. As a result, the researchers plan to develop a revised version of BalBoa where the pressure ranges will need to be refined to allow for more helpful corrections during the handstand. Additionally, they reported that the chosen color scale (see Figure 4) needed to be explained to participants during the first use [36]. This might be traced

[1] The project additionally explored vibrotactile feedback on the feet. The relevant results will be presented in the chapter on haptic feedback.

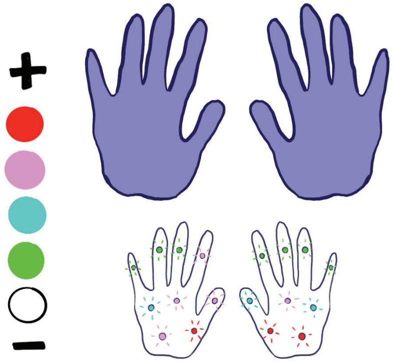

Figure 4 Conceptual sketch of BalBoa's LED feedback during a hand stand. The color scale reflects the incremental pressure, as follows: white (no pressure is applied); green (small pressure); blue (medium pressure); purple (high pressure); red (maximum pressure).

back to a violation of principles on color usage in information visualization which state that for ratio data the order and distance of values should be represented in the chosen color scale which is not achieved with the selected colors. Additionally, colors come with attached meanings and values based on common usage patterns (e.g. green = correct, red = wrong) which should be taken into consideration in the design to avoid confusing users [32].

BodyLights by Vidal et al. uses visualization in a different way than the previously described projects. Here, wearable laser lights ("BodyLights") were used to visualize body movement during strength training. The lights can be attached to the body (e.g. at joints that are relevant to a specific exercise) or to equipment (e.g. at both ends of a barbell). During movement, the laser light is projected onto a surface (e.g. training mat, wall) and offers an external, visual representation of the movement (see Figure 5). A personal trainer defines relevant target marks (e.g. post-its, tape lines on the wall) to which the trainee can align their movements.

Participants reported on numerous positive effects and benefits from training with the prototype. BodyLights helped to augment a set of training parameters such as pace, starting and ending points of movements, body position in relation to the used equipment and vice versa, muscle activation (when trembling or shaking), or a combination of these. It increased the trainee's autonomy by enabling them to identify errors on their own. Furthermore,

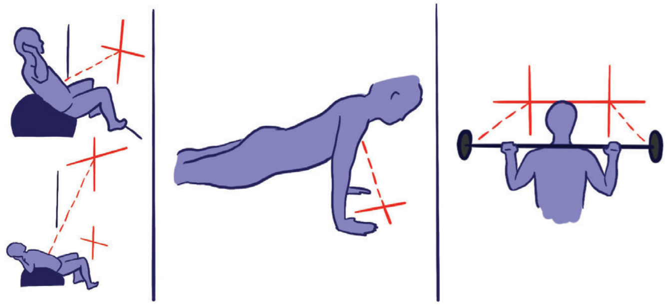

Figure 5 Conceptual sketches of BodyLights in use for a Swiss ball abs exercise marking start and end, for a plank exercise and for a barbell suqat (from left to right).

trainees reported that the laser projections offered an external focal point of attention that helped to decide what aspects of the technique and execution to focus on. Finally, the versatile options to use BodyLights with different exercises, as well as the potential to adapt the feedback to individual trainees, were praised. Besides the various strengths of BodyLights, two weaknesses were identified: As in other previously presented projects, there was an effect that visual attention was strongly demanded. In addition, some participants mentioned that it was stressful for them to see the limitations of their own body and execution technique reflected so precisely in the augmented visual feedback [34]. This shows that augmented feedback comes with a flipside: On the one hand, it has the potential to support the trainee, on the other hand, the increased perception of one's own capabilities can also induce body-image issues. Yet, it is noteworthy that even a relatively simple visualization such as a laser light projection has the potential to provide meaningful feedback.

3.2.2 Auditory Feedback

Auditory feedback can be provided in various different forms, for example through verbal cues, sounds, music, or rhythm, and by different means, such as speakers or headphones. Due to the various dimensions of sound such as volume, pitch, or tone, as well as different display attributes such as timing and localization, auditory feedback allows for the encoding of multidimensional data which might serve as the basis for augmented feedback [31].

An interesting approach to using sounds during physical movements is *sonification*. In their *review of sonification for use in physiotherapy*, Guerra et al. define sonification as "a technique to transform data into sound so that the properties and relationships between data variables are communicated using audible signals" [19]. The variability of how this concept can be applied is high. A simple example could be to map the volume of a continuously played sound to movement speed and thereby supporting a trainee to monitor the uniformity and consistency of movement. In the review, the researchers were able to conclude that sonification generally facilitates the conveyance of data towards a physiotherapy trainee. Positive reported effects are improvements in movement quality and control, increased movement and body awareness, improved performance (in particular postural control and range of movement) as well as encouragement to move. Further, more frequent feedback lead to increased retention of the target movement pattern, and melodic sonification was found to be superior to rhythmic sonification. While the review stresses that sonification seems promising to be used as a feedback technique for physical exercise, it also concludes that the design details of when and how to apply sonification to achieve these benefits need to be researched further [19].

With *Jymmin*, Fritz et al. studied a specific example of sonification using musical feedback to support movements. They compared passive music listening during physical exercise on fitness machines (e.g. stepper, stomach trainer) with Jymmin. "Jymmin", coming from the combination of "gym" and "jammin'", is an approach to music-making through the rhythmical operation of fitness machines. The resulting sounds are pre-defined, have a clear underlying metric pulse, and are mapped to the movements of the machines. In several studies, the researchers were able to show that Jymmin had a positive impact on mood and engagement during exercise [17], reduced the perceived exertion [18] and the perceived pain [16].

With *Eyes-Free Yoga*, Rector et al. brought up an interesting aspect of designing feedback of different modalities: accessibility. Auditory feedback, in particular, has the potential to serve visually impaired users in the context of physical exercise. The prototype provided audio-only feedback during six different yoga poses. Users were tracked using a Microsoft Kinect, their movements were analyzed and verbal corrections and auditory feedback were given. In a study focusing on qualitative feedback from visually impaired and blind participants, two conditions were compared: training with instructions only and training with instructions and feedback. They found that users with visual impairment preferred the prototype *with* verbal feedback and corrections yet the corrections were sometimes perceived as too strict. Some verbal instructions were easier to understand and follow than others which highlights the importance of experimenting with different design alternatives. Overall the verbal cues helped to understand the target posture which encouraged participants to participate in on-site group yoga classes due to a better understanding of the exercises. One further point that was mentioned was a desire for variety in positive feedback to keep up the encouraging effect it had. This shows that beyond designing and studying feedback in regards to its effectiveness, it is important to create an enjoyable and engaging experience and to explore the effects of feedback over time [29].

3.2.3 Haptic Feedback

Haptic feedback refers to both, tactile as well as kinesthetic feedback. Tactile feedback is conveyed and perceived on the skin, for example through vibrations or pressure. Kinesthetic feedback refers to intrinsic feedback mechanisms such as receptors in the muscles that allow us to feel the pose of the body. This section will mainly focus on the first as tactile feedback can be actuated externally, for example with vibration motors or robots [31]. Similar to visual and auditory feedback, haptic feedback can be designed by varying specific characteristics such as frequency, modulation, and waveform. Compared to the first two modalities, haptic feedback offers less of these characteristics, thereby reducing the possible design variations in this modality. The advantage is that it leads to a comparatively lower cognitive load [22]. Therefore, it can be particularly useful to communicate warnings or to give hints [9].

As part of *GymSoles*, a project previously discussed in the chapter on visual feedback, vibrotactile feedback was evaluated. The design goal was the same: Supporting squats and deadlifts by providing the trainee with feedback regarding their Center of Pressure (CoP) during training. This was achieved by attaching eight vibration motors to the sidewalls of a pair of sports shoes, thereby placing the feedback output spatially close to where the input data is measured. While the CoP was shifting on the sole, the vibration frequency of the motors closer to the CoP was increased. When comparing the designs of the visual and vibrotactile feedback, no quantitative difference regarding performance accuracy and no qualitative preference was found. Further, the vibrotactile feedback was perceived as more subtle thereby helping to focus more on the technique and body. Yet, some participants reported that vibrations were loud and still felt very alarming [14]. Based on this feedback, the updated prototype, *GymSoles++*, placed the vibration motors directly within the smart insole. This way, the haptic feedback would be provided on the foot sole rather than the sides of the feet and the volume during vibrations wasn't as loud [13].

Haptic feedback can be particularly valuable in usage contexts where the visual and auditory channels can't be blocked with feedback as they are required for other tasks. An example is *cycling* on the streets where it is important to monitor the street traffic. Peeters et al. aimed at understanding the impact of exercise intensity on the perception of vibrotactile cues in the case of cycling. Therefore, they tested the perceptibility of vibration signals on the thighs and the spine in a stationary indoor cycling setting. In their study, the researchers found that vibration signals are well perceived on both thighs and spine during cycling. The perceptual accuracy on the spine was significantly higher. On the thighs, the knee and middle of the thigh allow for the best perceptual accuracy whereas the upper thigh towards the hip joint resulted in the lowest accuracy. This was explained based on higher muscle and fat mass in these regions [27]. This work lays a promising foundation for further exploration of vibrotactile feedback during cycling. Bial et al. investigated a different approach by applying haptic feedback on the feet during cycling. The goal was to communicate the target pedaling frequency in order to attain a certain heart rate by alternating vibration signals on the left and right foot. Advantages of placing the actuators inside of the shoes as opposed to other body areas are that it isn't necessary to wear wristlets or gloves and that there wouldn't be any contact interruptions as there might be with the handles. One interesting finding of a user study was that the feedback was perceived as motivating to keep up with the vibration pattern [9].

3.2.4 Multimodal Feedback

Multimodal, as opposed to unimodal feedback, refers to combining several feedback modalities into one unified experience. Effectively, this could mean audiovisual, audiohaptic, visuohaptic, and audiovisuohaptic feedback. Not many recent research projects focus on specific configurations of multimodal feedback during physical exercise, yet, one important category to mention is feedback provided by robots.

Schneider et al. investigated and compared the use of *socially assistive robots* with display instructions in *indoor cycling training* in a long-term study over 18 days. The concept was focused on people in isolated environments such as space missions or arctic stations. In the development phase, the researchers modeled different phases and different movements of the training plan and prepared a detailed design of the feedback mechanism included a large variety of feedback types. As quantitative feedback, both non-self-referential feedback (e.g. reporting of the average heart rate, cadence, and power values), as well as self-referential feedback (by comparing the performance to a past performance while additionally monitoring aspects such as the maximal heart rate) were provided. The latter could result in explicitly encouraging feedback if the current performance was better than previous ones. Further, qualitative performance feedback was provided, e.g. "You were very good during speed intervals!", and the robot occasionally gave advice backed by sport scientific knowledge. As soon as enough data was available per session type, a trend analysis was provided.

The study focused on testing the user's compliance with the provided feedback by comparing the robot giving auditory instructions and feedback to a non-interactive text display showing instructions and feedback for the user to read. The robot condition was perceived as both more challenging and more motivating. Additionally, the user compliance was significantly higher with the robot because the feedback was automatically provided and didn't have to be read. The researchers concluded that it wasn't clear if the robot was the deciding factor making the difference as the comparison condition merely consisted of text instead of, for example, audio output [30]. An interesting aspect of this project is that while the robot is providing primarily auditory feedback, its physical, embodied presence might have an impact on how users deal with and react to the provided feedback. A further study comparing the robot to just auditory design could shed light on this.

Besides socially-assistive robots, another type of inherently multimodal technologies are *exoskeletons* providing both haptic and visual feedback to their wearers. Recent research evaluating their application in physical exercise focuses on the context of rehabilitation of upper or lower limbs, gait training or older adults as a target user group. One example of such an exoskeleton is the Human Body Posturizer (HBP), a full-body orthesis consisting of four elements: a back element, a cranio-cervical helmet, a lombo-sacral sacrum element and a lower limb element. After Di Russo et al. were able to demonstrate that training with the HBP can improve mobility and motor control [12], Verrusio et al. conducted an interesting study comparing the effects of physical training with and without the HBP on depression levels of eldery patients. After training for walking, balance and posture strength with a physical therapist for 45 minutes three times a week for six months only the group that trained with the exoskeleton showed a significant reduction in depression levels. The researcher hypothesize that this effect might originate from the circumstance that training with the HBP allows for a faster achievement of noticeable improvements and results leading to a feeling of success and achievement [35]. To the best of our knowledge, dedicated research on the effects of exoskeleton-enabled feedback and its modalities hasn't been performed.

3.3 Comparing Feedback Modalities

Based on the insights gathered in this review, it is clear that feedback modalities as a whole can't be ranked or measured against each other in regards to effectiveness. The effectiveness of feedback depends on many more factors than modalities, such as task complexity, the trainee's skill level, and design decisions [31].

Visual feedback, both concurrent and terminal, was found to be effective depending on the context. Concurrent visual feedback is more effective for complex tasks, whereas terminal visual feedback for simple tasks [31]. It allows for more precision to convey information than other modalities, especially compared to haptic feedback [14]. Several examples have shown that if visual feedback is presented to the user throughout physical exercise, it requires and strongly draws the trainee's focus, thereby drawing attention away from the body, technique, and environment [15]. Yet, it was found that this attention-drawing effect can have positive sides as well when trying to learn complex movements. In that context, the feedback offers a focal point of attention that helps to guide the trainee's attention through the complex movements and thus helps to avoid cognitive overload by simplifying the movement momentarily and making it more accessible [31].

Auditory feedback has the advantage that it is perceptible without loading the full attention capacity of the trainee [15]. For the most part, it offers lower communication precision and depth for conveying information than visual feedback but more than haptic feedback [19]. Sonification of movements was found to be a powerful feedback tool, especially for simple tasks, with a still underexplored potential.

Haptic feedback comes with seemingly strong drawbacks as it initially takes longer to interpret, the perceptual precision and range of easily communicated information is limited and it requires dedicated hardware. Nonetheless, it shares one advantage with auditory feedback as it is perceptible without full focus and thereby perceived as more subtle and less distracting as opposed to visual feedback. This promises to be beneficial for feedback to advanced trainees that don't require as much feedback on execution as well as in scenarios where the visual or auditory channels can't be used for feedback [15].

Multimodal feedback might have the potential to combine the strengths of all modalities into a seamless, individual feedback experience that could prove especially effective for complex physical exercises. Just as human instructors or physiotherapists rarely rely on one single modality for instruction, multimodal feedback might combine and switch between them [31].

4 Discussion

The goal of this section is to critically reflect on the application of technologically enabled feedback for trainees during physical exercise.

Replacing instructors vs. supporting instructors. An important aspect to reflect on when discussing feedback that is automatically provided by technology is the implicit change of the role of the instructor. With technology, the potential for personalized, more accessible, and scalable feedback rises. This could explain why the current overarching design approach is to automatize the processes of instruction and correction. This is problematic as the limitations

of technology to assess execution lead to a focus on isolated parameters. The underlying understanding that correct technique can be specified is risky as in reality the trainee's skills and context might influence what can be considered correct and designs that don't acknowledge that could fail to meet the actual user needs [34]. Overall, technology to support exercise is still considered impractical, requiring technical expertise, and expensive [14]. An emerging alternative design approach is to aim at supporting an instructor rather than replacing them. Both design directions have advantages and disadvantages and should be able to co-exist to serve different users, needs and scenarios.

Impact on autonomy, agency, and learning. The *guidance* hypothesis describes the risk of developing a dependency on feedback when being provided with it permanently during the technique acquisition phase. When the feedback is later removed for trainees with a feedback dependency, their performance can be worse than the one of trainee's that never received feedback [31]. Compared to the feedback provided by a human instructor, technology has the potential to provide continuous and more precise feedback to trainees. This could lead to an even more severe reliance on this feedback, increasing dependency and reducing autonomy, agency, and learning in trainees. To avoid this, a successful design should aim at reducing these effects by facilitating the development of intrinsic error detection and correction capabilities in trainees. The minimum requirement for an acceptable approach should be to be transparent. Even if a system doesn't specifically address the risk, trainees should be informed about them in order to be empowered to make a decision about usage.

Impact on body awareness. One goal of physical exercise can be to better get to know one's body and capabilities, to increase body awareness, and to mindfully focus on oneself. Introducing technology to the practice can distract and draw attention towards invasive devices or feedback mechanisms. Visual feedback seems to draw attention away from the body most strongly (e.g. [34, 14, 33]). Additionally, by making the trainee's abilities measurable and more visible, feedback can even increase the stress related to negative aspects such as the limitations of the trainee's fitness level. This could further lead to newly developing or reinforced body image and mental health issues [34]. Therefore, the degree of distraction should be added to the studied parameters when investigating feedback modalities and designers should be mindful of possible negative consequences of their products and projects.

5 Limitations

Research projects focusing on feedback in group training settings, exergames, and mixed, virtual, or augmented reality have been excluded from this work to limit the scope in favor of a more focused review.

6 Conclusion and Future Work

To conclude, visual, auditory, haptic, and multimodal feedback all come with their own set of relevant design parameters, implications, advantages, and limitations that need to be considered when aiming at designing effective and engaging experiences. Even though the context of physical exercise requires adaptation, it is important to conform to established principles and design guidelines from other application areas. Further, the effectiveness of feedback shouldn't only be measured based on whether it allows users to increase the execution accuracy of exercises or whether it manages to provide meaningful corrections. Instead, engagement, fun, and a certain level of humaneness are important in order to ensure a pleasant user experience and create the potential for long-term value. Therefore, it is important to take aspects such as the cognitive load of a feedback design and the user's context, skill level and needs into account when designing feedback for physical exercise. A certain feedback design might only work in certain circumstances and, for an individual user, feedback most likely has to change over time to stay effective.

An observation that can be made when looking at studies on feedback for physical exercise is that a variety of feedback designs in different modalities are studied in regards to a variety of usage scenarios and evaluation parameters. Thereby, the opportunity for future work in this space seems vast as it is lacking extensive, systematic research. One direction could be to compare different designs of the same modality to elaborate on advantages and disadvantages and design recommendations more clearly. In regards to auditory feedback, the emerging field of movement sonification could be more explicitly investigated. Evidence suggests that sonification has beneficial effects but the details of the sound design need to be studied further. Apart from a small subset of projects, little work focusing on physical exercising and feedback throughout considers the accessibility aspects of different modalities. Yet, auditory and haptic feedback could have great potential for visually impaired or blind users. Further, it would be interesting to explore how visual and/or haptic feedback could be optimized for users with hearing loss.

References

1 blogilates YouTube Statistics (Website, accessed Jan 21, 2021). URL: `https://socialblade.com/youtube/user/blogilates`.

2 Freeletics (Website, accessed Jan 4, 2021). URL: `https://www.freeletics.com/en/training/`.

3 Hosting Workouts Over Zoom (Website, accessed Jan 5, 2021). URL: `https://blog.zoom.us/best-practices-for-hosting-zoom-workouts/`.

4 Kaia Health, AI Lab (Website, accessed Jan 4, 2021). URL: `https://www.kaiahealth.com/ai-platforms/motion-coach/`.

5 Nike Training Club App (Website, accessed Jan 21, 2021). URL: `https://www.nike.com/ntc-app`.

6 Peloton Interactive, Inc. | Class A Common Stock Historical Data | Nasdaq (Website, accessed Mar 28, 2021). URL: `https://www.nasdaq.com/market-activity/stocks/pton/historical`.

7 Strength Training at Home: Workouts With and Without Equipment (Website). URL: `https://www.healthline.com/health/exercise-fitness/strength-training-at-home`.

8 Vivira Health Lab (Website, accessed Jan 5, 2021). URL: `https://www.vivira.com/`.

9 Dominik Bial, Thorsten Appelmann, Enrico Rukzio, and Albrecht Schmidt. Improving cyclists training with tactile feedback on feet, 2012. `doi:10.1007/978-3-642-32796-4_5`.

10 Carl J. Caspersen, Kenneth E. Powell, and Gregory M. Christenson. Physical Activity, Exercise, and Physical Fitness: Definitions and Distinctions for Health-Related Research Synopsis. Technical report.

11 Yucheng Chen, Yingli Tian, and Mingyi He. Monocular Human Pose Estimation: A Survey of Deep Learning-based Methods. `doi:10.1016/j.cviu.2019.102897`.

12 Francesco Di Russo, Marika Berchicci, Rinaldo Livio Perri, Francesca Romana Ripani, and Maurizio Ripani. A Passive Exoskeleton Can Push Your Life Up: Application on Multiple Sclerosis Patients. *PLoS ONE*, 8(10), oct 2013. `doi:10.1371/journal.pone.0077348`.

13 Don Samitha Elvitigala, Denys Matthies, Chamod Weerasinghe, Yilei Shi, and Suranga Nanayakkara. GymSoles++: Using Smart Wearbales to Improve Body Posture when Performing Squats and Dead-Lifts. In *ACM International Conference Proceeding Series*, pages 1–3, New York, NY, USA, mar 2020. Association for Computing Machinery. `doi:10.1145/3384657.3385331`.

14 Don Samitha Elvitigala, Denys J.C. Matthies, Löic David, Chamod Weerasinghe, and Suranga Nanayakkara. Gymsoles: Improving squats and dead-lifs by visualizing the user's center of pressure. In *Conference on Human Factors in Computing Systems - Proceedings*, pages 1–12, New York, NY, USA, may 2019. Association for Computing Machinery. `doi:10.1145/3290605.3300404`.

15 Corinna Feeken, Dag Ennenga, Merlin Wasmann, Heiko Möller, Wilko Heuten, and Susanne Boll. Climbingassist - Direct vibro-Tactile feedback on climbing technique. In *UbiComp 2016 Adjunct - Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 57–60. Association for Computing Machinery, Inc, sep 2016. `doi:10.1145/2968219.2971417`.

16 Thomas H. Fritz, Daniel L. Bowling, Oliver Contier, Joshua Grant, Lydia Schneider, Annette Lederer, Felicia Höer, Eric Busch, and Arno Villringer. Musical agency during physical exercise decreases pain. *Frontiers in Psychology*, 8(JAN), jan 2018. doi:10.3389/fpsyg.2017.02312.

17 Thomas H. Fritz, Johanna Halfpaap, Sophia Grahl, Ambika Kirkland, and Arno Villringer. Musical feedback during exercise machine workout enhances mood. *Frontiers in Psychology*, 4:1–7, 2013. doi:10.3389/fpsyg.2013.00921.

18 Thomas H. Fritz, Samyogita Hardikar, Matthias Demoucron, Margot Niessen, Michiel Demey, Olivier Giot, Yongming Li, John Dylan Haynes, Arno Villringer, and Marc Leman. Musical agency reduces perceived exertion during strenuous physical performance. *Proceedings of the National Academy of Sciences of the United States of America*, 110(44):17784–17789, oct 2013. doi:10.1073/pnas.1217252110.

19 Joao Guerra, Lee Smith, Domenico Vicinanza, Brendon Stubbs, Nicola Veronese, and Genevieve Williams. The use of sonification for physiotherapy in human movement tasks: A scoping review, 2020. doi:10.1016/j.scispo.2019.12.004.

20 Honorata Jakubowska. *Skill Transmission, Sport and Tacit Knowledge: A Sociological Perspective*. 2017.

21 Mark L. Latash and Vladimir Zatsiorsky. *Biomechanics and motor control: defining central concepts*. 2015.

22 Jani Lylykangas, Veikko Surakka, Jussi Rantala, Jukka Raisamo, Roope Raisamo, and Esa Tuulari. Vibrotactile information for intuitive speed regulation. In *People and Computers XXIII Celebrating People and Technology - Proceedings of HCI 2009*, pages 112–119. British Computer Society, 2009. doi:10.14236/ewic/hci2009.13.

23 Dan Morris, T. Scott Saponas, Andrew Guillory, and Ilya Kelner. RecoFit: Using a wearable sensor to find, recognize, and count repetitive exercises. In *Conference on Human Factors in Computing Systems - Proceedings*, pages 3225–3234. Association for Computing Machinery, 2014. doi:10.1145/2556288.2557116.

24 Florian Mueller, Martin R. Gibbs, and Frank Vetere. Taxonomy of exertion games. In *Proceedings of the 20th Australasian Conference on Computer-Human Interaction: Designing for Habitus and Habitat, OZCHI'08*, pages 263–266. ACM Press, 2008. doi:10.1145/1517744.1517772.

25 Florian Mueller, Joe Marshall, Rohit A. Khot, Stina Nylander, and Jakob Tholander. Understanding sports-HCI by going jogging at CHI. In *Conference on Human Factors in Computing Systems - Proceedings*, volume 18, pages 869–872, New York, NY, USA, 2015. Association for Computing Machinery. doi:10.1145/2702613.2727688.

26 Florian Mueller, Shannon O'Brien, and Alex Thorogood. Jogging over a distance: Supporting a "jogging together" experience although being apart. In *Conference on Human Factors in Computing Systems - Proceedings*, pages 2579–2584, New York, New York, USA, 2007. Association for Computing Machinery. doi:10.1145/1240866.1241045.

27 Thomas Peeters, Eric Van Breda, Wim Saeys, Evi Schaerlaken, Jochen Vleugels, Steven Truijen, and Stijn Verwulgen. Vibrotactile Feedback during Physical Exercise: Perception of Vibrotactile Cues in Cycling. *International Journal of Sports Medicine*, 40(6):390–396, 2019. doi:10.1055/a-0854-2963.

28 Arkalgud Ramaprasad. On the definition of feedback. *Behavioral Science*, 28(1):4–13, 1983. doi:10.1002/bs.3830280103.

29 Kyle Rector, Cynthia L. Bennett, and Julie A. Kientz. Eyes-Free Yoga: An Exergame Using Depth Cameras for Blind and Low Vision Exercise. Technical report, 2013. URL: http://yourshapegame.ubi.com/fitness-evolved-2012/.

30 Sebastian Schneider, Luise Süssenbach, Ingmar Berger, and Franz Kummert. Long-term feedback mechanisms for robotic assisted indoor cycling training. In *HAI 2015 - Proceedings of the 3rd International Conference on Human-Agent Interaction*, pages 157–164, New York, NY, USA, 2015. Association for Computing Machinery, Inc. doi:10.1145/2814940.2814962.

31 Roland Sigrist, Georg Rauter, Robert Riener, and Peter Wolf. Augmented visual, auditory, haptic, and multimodal feedback in motor learning: A review. *Psychonomic Bulletin and Review*, 20(1):21–53, 2013. doi:10.3758/s13423-012-0333-8.

32 Samuel Silva, Beatriz Sousa Santos, and Joaquim Madeira. Using color in visualization: A survey. *Computers and Graphics (Pergamon)*, 35(2):320–333, 2011. doi:10.1016/j.cag.2010.11.015.

33 Richard Tang, Anthony Tang, Xing-Dong Yang, Scott Bateman, and Joaquim Jorge. Physio@Home: Exploring visual guidance and feedback techniques for physiotherapy exercises. 2015. doi:10.1145/2702123.2702401.

34 Laia Turmo Vidal, Hui Zhu, and Abraham Riego-Delgado. BodyLights: Open-Ended Augmented Feedback to Support Training Towards a Correct Exercise Execution. In *Conference on Human Factors in Computing Systems - Proceedings*. Association for Computing Machinery, apr 2020. doi:10.1145/3313831.3376268.

35 Walter Verrusio, A. Renzi, F. Cecchetti, F. Gaj, M. Coi, M. Ripani, and M. Cacciafesta. The Effect of a Physical Training with the Use of an Exoskeleton on Depression Levels in Institutionalized Elderly Patients: A Pilot Study. *Journal of Nutrition, Health and Aging*, 22(8):934–937, oct 2018. URL: https://link.springer.com/article/10.1007/s12603-018-1044-2, doi:10.1007/s12603-018-1044-2.

36 Laia Turmo Vidal, Luis Parrilla Bel, Elena Márquez Segura, and Annika Waern. Balboa: A balancing board for handstand training. In *Conference* on *Human Factors in Computing Systems - Proceedings*. Association for Computing Machinery, 2019. doi:10.1145/3290607.3312909.

37 Laia Turmo Vidal, Elena Márquez Segura, and Annika Waern. Movement correction in instructed fitness training: Design recommendations and opportunities. *DIS 2018 - Proceedings of the 2018 Designing Interactive Systems Conference*, pages 1041–1054, 2018. doi:10.1145/3196709.3196789.

38 Paweł W. Woźniak, Pascal Knierim, and Matthias Hoppe. Health, fun, and engagement: Computing technologies that support physical activity. *XRDS: Crossroads, The ACM Magazine for Students*, 25(4):16–20, 2019. doi:10.1145/3331014.

Designing for Disengagement

Mariam Ali Hussain

Ludwig-Maximilians-Universität München, München, Deutschland
Mariam.Ali@campus.lmu.de

Abstract

While commercial design seeks to capture and hold the attention of the user and increase the users' engagement with the product, this work questions the so-called attention economy and calls for designing for disengagement. This can be understood as designers creating less dependence on the technology and making it easier for users to step away from it. A literature review was conducted to understand the problems of dependence on and addiction to technology and to find approaches for design implications in this area. We present three application areas in which designing for disengagement can be applied: Digital navigation tools, mindfulness apps and spelling and grammar checkers. The implication for designers is to think about their product's purposes outside the moments of use so that users are able to transfer the knowledge or practice in their daily life. We highlight opportunities and challenges for future research and we notice the need to establish terminology within this sub-field.

2012 ACM Computing Classification Human-centered computing → Interactive systems and tools

Keywords and phrases Reflective HCI; Design; Attention Economy; Disengagement.

1 Introduction

For companies in the commercial area, it is crucial how much attention the consumer pays to their products and how to capture and hold their attention more efficiently. This is the so called "attention-economy" [13]. This term explains that attention is not only a resource for companies but also a currency: Users pay for a service with their attention. YouTube for example, recommends videos that the individual user is more likely to watch on the first page or while watching a video to keep the user as long attentive as possible. The autoplay feature of YouTube amplifies this effect. The advertisement played before and during each video is a reason to keep audience's attention fixed on the successive videos. The more time a user spends on the app the more ad revenue. This means that companies are economically motivated to engage their users as long as possible.

Cite as: Mariam Ali Hussain. Designing for Disengagement. In *6th Seminar on Ubiquitous Interaction (UBIACTION 2021)*. Editors: F. Lang, R. Welsch, L. Haliburton, F. Draxler, S. Feger, S. Villa, M. Hoppe, P. Knierim, V. Mäkelä, A. Schmidt. February 11, 2021. Munich, Germany. pp. 8:1–8:16.

A major factor that influences the user's attention is the design of the product. The interface design that intentionally encourages users for a specific behavior like spending more time than wanted or purchasing for features that are not needed is known as malicious interface design [7] or the dark patterns of user experience design [9], see Figure 1. The endless scrolling through Instagram, YouTube or Twitter or the recommended content that keeps the user engaged are only a few more examples [25]. These come to the expense of the user: People describe that they experience excessive and habitual use of technology that they later regret or find meaningless [39, 20].

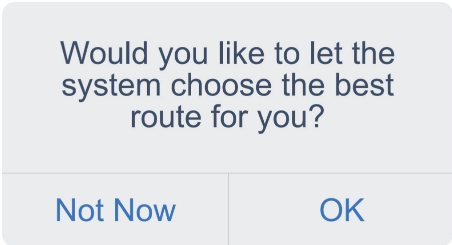

Figure 1 Exemplary interface for a dark pattern. Users are not given a "No" option. Image: [4]

Research in software design mainly focuses on promoting user's engagement [29, 32] and using it as a measurement for the evaluation of a system or application [39, 1]. This work reflects on the question when designers of technology should support disengagement or create less dependence so that it becomes easy for users to step away from the technology. Therefore, a literature review was conducted and structured as follows:

First, we present related works and design agendas like Slow Technology and Mindfulness in Human-Computer-Interaction which address the similar problem and show attempts on how to solve it (see Chapter 2).

The next section is about technology addiction and dependence on technology and their differences. The triggers that cause obsessive phone usage are explained and why people sometimes rely on technology too much (see Chapter 3).

Moreover, three appliances that are prime for designing for disengagement are presented: Digital navigation tools, mindfulness apps and spelling and grammar checkers (see Chapter 4).

In addition, this work also identifies gaps for future research opportunities and shows that there is a need to establish terminology within this area (see Chapter 5).

In conclusion, we summarize the findings of the literature review and the implication of designing for disengagement (see Chapter 6).

2 Related Work

2.1 Slow Technology

Slow Technology is a design agenda which promotes more reflection and slower pace in technology usage. It is created as a countermovement against the more efficient and faster becoming technologies and lifestyles. The idea behind slow technology is to develop technologies that are slow so that users are able to regain control of the their attention and become more mindful [10]. This work has a similar goal but with a different approach. While slow technology emphasizes on the speed aspect, this work focuses on changing the user's interactions with technology to become more independent and to make it easy to move away from it.

Odom et al. described the design and making of an artifact called the Slow Game in order to show how slowness can be implemented in a game [30]. The design artifact consists of a cube with a low-resolution display made out of LED lights. They implemented a different version of the classic "snake" mobile game in which the player has to move the snake that gets continually longer without moving into itself. Instead of pressing on buttons, in the Slow Game the movements of the snake are controlled by the physical rotations of the cube, see Figure 2. The special aspect is that a user can only make one movement a day which means approximately after 18 hours.

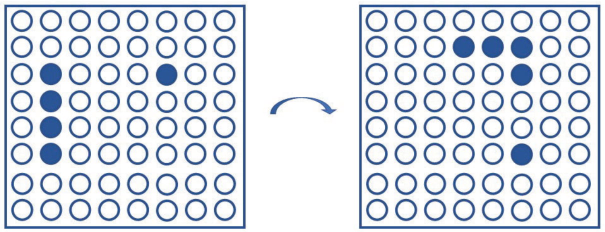

Figure 2 Possible interaction of the Slow Game cube. Image: [30]

2.2 Mindfulness in Human-Computer-Interaction

Mindfulness in human-computer-interaction (HCI) does not have a clear definition. Projects in this field intend to support moments of mindfulness, stress reduction and better relaxation [38]. What they all have in common is the enhancement of the mental well-being. Lukoff et al. emphasizes the importance of transferring mindfulness from apps to everyday life so that users become more independent when practicing mindfulness. They recommend to design for less dependence in the first place by leaving out redundant design elements. They also propose to gradually reduce the scaffolding so that users disengage from the tool. Last but not least, they point out to design for long-term effects to practicing mindfulness in everyday life and not only in formal practices [18]. These aspects are just as important in other fields outside of mindfulness applications.

The app "Calm", which became the iPhone App of the year by Apple in 2017 [37], offers a range of mindfulness guided and unguided meditation practice modules. The app allows users to choose specific goals like better sleep, stress reduction or improving focus to provide suitable content. There is a free version but in order to get access to the full modules, one have to pay for a premium version. A study found out that Calm helps to reduce stress and improve mindfulness in short-term contexts for college students [12]. There is a need to explore the long-term effectiveness and acceptability of such meditation apps. In our opinion, since Calm is a rather consumer-based app because of the tracking and reminders, there is room to improve the concept through disengagement and through transfer practices in everyday life without using the app. Another research argues that the app developers should not claim that daily practice is optimal because they found out that some study participants perceived a feeling of failure and demotivation when they could not keep up with the practice [5].

2.3 Considering Not to Design

Even if we address the issue of designing for disengagement, we think it is just as important that in some cases it is better to not design at all. The main purpose of disengagement is to support people being independent of technology and confident in their own abilities. Baumer et al. [2] reflect on the question when technology should be not implied. They present three questions that need to be considered during the design process in order to think about when the invention might not be appropriate.

The first question is to think about whether no or less technology can be used for the same purpose.

The second question tackles the problem that in some cases technology can cause trouble or harm even if at first it seems like the technological inventions is improving a specific situation.

The third question deals with the difference between a transformation of a problem and the problem itself.

For future research, specific cases can be explored in which it is better to not design at all.

3 Addiction and Dependence

Being dependent on technology or being addictive to technology have different meanings. In this work, we consider both meanings, but not interchangeably, because on the one hand we ask the question on how to support disengagement and therefore less dependence on technology and on the other hand we deal with the question on how to make it easy for users to step away from technology and therefore being less addictive to them.

3.1 Technology Addiction

When talking about technology addiction in this work, it means the obsessive use of technology. Tran et al. investigated in triggers that start obsessive phone use [39]: Moments of downtime, tedious tasks, social awkwardness, anticipation and sometimes people check their phones for no specific reason, see Figure 3. After understanding the reasons, they provide three design agenda to create a meaningful use of mobile phones [39].

1. They recommend to create meaningful experiences for users instead of focusing on lockout mechanisms.
2. They uncover the need for designers that promote tasks that go beyond the current moment of use and see it also as an investment in the future.
3. They propose to give users the control for when they are likely to spend time on the phone and when they are likely to spend no time. This can be done by avoiding dark patterns that capture and hold users' attention.

The most important findings are that in order to avoid technology addiction, it is important to create a meaningful experience for the user and allow them to stop using the technology whenever they want.

According to Lanette and Mazmanian, it is not always clear how to differentiate between beneficial and detrimental phone use. This can result to the issue of incorrectly labeling individuals as addicts. Because the usage of smartphones changed and increased in the digital age, there is a need to redefine what can be considered as normal and what can be considered as a clinical addiction [16].

Trigger	Description	Example
Unoccupied moment	Any moment of downtime with no obvious alternative stimuli	*"I feel like whenever I'm bored, I just check if I have any notifications or texts or I go on Instagram and I just scroll through even though there's no real purpose…If I'm not around friends or if I'm laying down in bed and I can't fall asleep or I'm walking to class or if I'm in class and I don't want to pay attention anymore, things like that."*
Tedious task	Any effortful activity	*"You get into the video, and then you realize you don't want to go back to homework. You just keep watching the videos even if they're not good, because it's more fun than homework."*
Social awkwardness	Situations that deviate from social norms or leave the user feeling uncomfortable	*"You probably think more about what other people must be thinking about you, instead of thinking about what you're seeing on your phone screen. Instead you're like, 'Oh people probably think I'm a loser sitting over here all alone."'*
Anticipation	An expectation of social or informational rewards	*"If I'm talking to someone on Snapchat or if I'm texting them and they haven't responded, I'll continuously check up on that."*

Figure 3 An overview of the triggers for compulsive phone usage according to Tran et al. with statements from the study participants. Image: [39]

Minagawa and Fujinami propose a modified version of the Smartphone Addiction Proneness scale (SAPS) introduced from Kim et al. [14] to measure the level of addiction to smartphone. The SAPS consists of 15 questions related to the excessive use of smartphone in the four categories: Disturbance of adaptive functions, virtual life, withdrawal, and tolerance, see Figure 4. The SAPS score is extracted from the information of how long and how frequent the smartphone device is used. With the help of a regression model, Minagawa and Fujinami proposed a technique to automatically recognize the SAPS score based on the daily smartphone use [24].

Subdomain	Items	No.
Disturbance of Adaptive Functions	My school grades dropped due to excessive smartphone use.	1
	I have a hard time doing what I have planned (study, do homework, or go to afterschool classes) due to using smartphone.	5
	People frequently comment on my excessive smartphone use.	9
	Family or friends complain that I use my smartphone too much.	12
	My smartphone does not distract me from my studies.	13*
Virtual Life Orientation	Using a smartphone is more enjoyable than spending time with family or friends.	2
	When I cannot use a smartphone, I feel like I have lost the entire world.	6
Withdrawal	It would be painful if I am not allowed to use a smartphone.	3
	I get restless and nervous when I am without a smartphone.	7
	I am not anxious even when I am without a smartphone.	10*
	I panic when I cannot use my smartphone.	14
Tolerance	I try cutting my smartphone usage time, but I fail.	4
	I can control my smartphone usage time.	8*
	Even when I think I should stop, I continue to use my smartphone too much.	11
	Spending a lot of time on my smartphone has become a habit.	15

* Reverse-coded items.
doi:10.1371/journal.pone.0097920.t002

Figure 4 The SAPS questionnaire which contains 15 questions divided into the categories disturbance of adaptive functions, virtual life, withdrawal, and tolerance. Image: [14]

Regardless of the definition of addiction, some people experience a feeling of stress or disgust when they lack control or self-regulation in their mobile phone usage [19]. Additionally, social media apps cause a fear of missing out which causes anxiety [20].

Overuse of smartphones can also lead to unproductivity. Moran and Salazar introduce the term "the Vortex" which they describe as follows:

> The Vortex is a user-behavior pattern that begins with a single intentional interaction followed by a series of unplanned interactions. This unplanned chain of interactions creates a sense of being "pulled" deeper into the digital space, making the user feel out of control [26].

Excessive smartphone use can create feelings of dissatisfaction and unproductivity and therefore solutions to reduce them need to be further investigated.

3.2 Dependence on Technology

When talking about dependence on technology in this work, we define it as relying more on technology than on oneself. The term "automation bias" explains this issue: It refers to the phenomenon that people trust excessively in automated systems which can sometimes lead to trusting in false automated information [27]. This is investigated in human-factors research [40, 31] and we could not find any research covering automation bias in the field of HCI. Baxter and Kabi collected three main reasons for automation bias [3].

- People tend to believe that technology's analytical power is superior to that theirs and therefore overestimate the performance and precision of technologies [40].
- People prefer to put the least cognitive effort into human decision-making and therefore hand over the decision to the technology [40].
- When a system executes the same task, people tend to put less effort into it or hand over the responsibility to the system [6].

There is a need to explore automation bias in the field of HCI and investigate times where it occurs in everyday life when using technical devices. This work provides some examples which are found during the literature review.

4 Application Areas

In the following, we will introduce three application areas where designing for disengagement could be effectively applied. The purpose is to highlight opportunities for future research in this sub-field. We chose these three areas because there is supportive research that amplifies the problems and needs.

4.1 Digital Navigation Tools

The desired effect of a navigation tool is that the user eventually reaches the intended destination. Some research claim that mobile navigation systems support wayfinding but neglect to support people in learning to orient themselves in an environment or in learning the route [28, 15, 17]. Michon and Denis suggest landmarks for reorientation on the route. Landmarks are local features which should support users in focusing on the environmental aspects [23]. What if the navigation tool would not provide the exact path but rather easy recognizable features along a route? Navigation tools that enhance the user's spatial knowledge create less dependence on the technology and therefore can be considered as a use case when designing for disengagement.

The example of the navigation system shows that the actual problem which is a user who can not orient themselves is not tackled but instead a short-term solution - to get to the route by instruction - is provided. It would be interesting to know how a navigation tool would look like if it enhances also spatial knowledge. Not only the effect on the spatial knowledge can be evaluated but also how applicable the idea is in the real world because, as stated in the section about dependence on technology, people tend to choose the option with the least cognitive effort.

Digital Navigation tools like Google Maps help to navigate to a certain place. Google Maps would show the exact path on the mobile device, that one has to go in order to reach their destination. Figure 5 shows an example of how landmarks in a navigation app might show for example only a picture and maybe a name of the landmark and a very short instruction about the direction instead of the exact path. The user will pay more attention on the environment and may learn the route and the user becomes eventually at some point independent from the navigation tool.

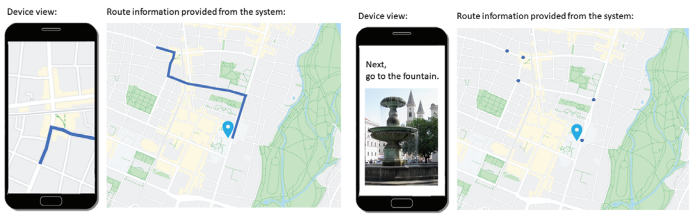

Figure 5 Comparison of a traditional navigation tool (left) and a navigation with landmarks tool (right). Map: [21], Image: [35]

Another idea in this area could be considering augmented reality (AR) as a tool in order to blend landmarks with the real world environment. McMahon

et al. showed that AR was the most effective way for navigating in an unknown environment for people with intellectual disabilities because of this aspect [22]. It would be interesting to know if this applies to other groups and if this method supports reorientation.

4.2 Mindfulness Apps

Lukoff et al. offer three suggestions for designers on how to apply mindfulness practices on everyday life so that people do not become dependant on the scaffolding in relaxation-oriented apps [18].

First, they propose to minimize design elements as much as possible in order to make it easy for people to not rely on the app in the first place. For example this might be done by omitting redundant design elements like sounds and visuals and instead focusing on simple voice guidance.

Second, by decreasing scaffolding little by little, users will reach a certain state in which they are able to meditate without the app or in which the app only has the functionality of a timer for example.

Third, it is suggested to focus on long-term changes in the mindfulness capacity in order that users are able to practice mindfulness in everyday life and not only during meditation. They recommend gentle notifications to remind them to practices mindfulness at certain moments during the day.

Apps that support mindfulness and relaxation should not only help users only in the time when they are using the app but also for times outside the use and in their daily life.

These insights can provide a good basis for implementing a project in this area because we could not find a successful implementation in our literature review.

4.3 Spelling and Grammar Checkers

Hardré mentions spelling and grammar checkers as an example of how people overtrust in technology. According to him, users overtrust in word processing systems because they believe that only because a word is not marked as incorrect then there must be no error. But they overlook the possibility that the word is used in the false context or is spelled wrong because the system interpreted it as another word [11], see Figure 6. This is an example for how people overestimate the performance and accuracy of technologies.

This scenario made us think about another use case: When people use spelling and grammar checkers and choose the suggested word from the auto-correction, they might hand over the responsibility to the technology and not pay attention to the mistake. A study showed that spell-checker helped students to revise their spelling during the dictation exercise but after the use

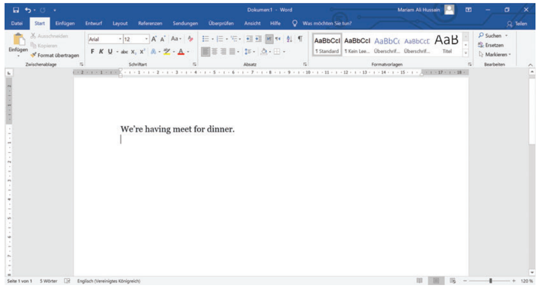

Figure 6 A spelling error which is not labeled as one from the word processing system because it is written correctly. The word "meat" and "meet" are pronounced the same but "meat" makes more sense in this context.

of spell-checkers they repeated the errors [34]. Hence, people gain a short-term benefit from word processing systems but it does not improve their literacy skills.

Last but not least, for the spelling and grammar checkers we could not find any research on how word processing systems can enhance the user's literacy skills. Therefore, it is an open opportunity to think about how such a tool would look and work like. Feedback hints could help to think about one's mistakes. For example, when there is a spelling mistake in a sentence, the spelling checker could give a hint like "there is one letter too much", see Figure 7. Users might remember the hint the next time when they are writing the same word again.

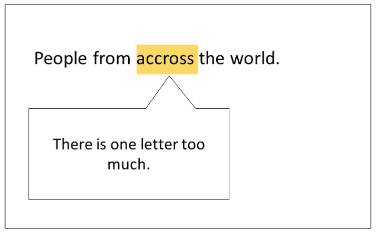

Figure 7 A spelling checker that gives a hint for the user because one c is redundant.

There is another idea which is inspired by the digital well-being tools that say how much screen time one spends on an app. How about visualizing the

current spelling mistakes to inform the user about them? Figure 8 shows a sketch to better represent the idea. Users might be more aware about their spelling and grammar errors.

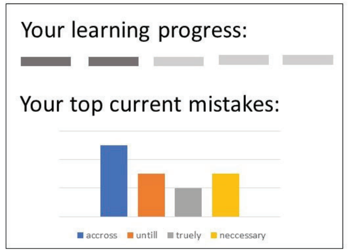

Figure 8 A sketch of a spelling checker system that shows an overview about the last spelling mistakes of a user.

We provided use cases and approaches to design for disengagement and encourage designers and researchers to extract concrete ideas and implement a design artifact. The design process of an application in this area could reveal hidden challenges and offer insights that illustrate how such challenges can be productively handled through design practice.

Moreover, an evaluation of the created design artifact could help to understand how efficient these products are in order to fight technology addiction and dependence.

5 Gaps and Challenges for Future Research

Our analysis revealed different research and design approaches that represent interesting opportunities and directions. Throughout the paper we have discussed application areas, problems and approaches when it comes to designing for disengagement. In this section we want to highlight gaps and challenges for future research. We invite researchers, designers and also users to contemplate about the appropriateness as well as the need for designing for disengagement.

5.1 Conflicting Interests with the Attention Economy

In the attention economy, commercial companies make profit when people use their products as much as possible and as long as possible. This makes it difficult to persuade the economic view on the purposes of the design for disengagement. People might use the products less or even stop using them.

Apple and Google take the approach to include apps which should help users from getting distracted by having track of when and for how long they spend time on apps and devices [8]. Future research can investigate if other companies provide similar approaches and how effective the current systems are in order to prevent technology addiction. It is just as important to ask the question how could they be made more effective e. g. if there are better ways to visualize the screen time information. How often should the user be notified about their results and how should they be informed? They can be notified regularly or only when they are excessively using their phones.

Even if it sounds promising that big companies like Apple and Google recognized this problem, it is still questionable how far they will go when it comes to change the design or concept of their products directly in order to support disengagement. Because the mentioned approaches from these companies only provided tracking and notification systems and therefore neglect the relationship between how much and why people spend time on their devices. In our opinion, this is a challenge that might be further discussed.

5.2 Establishing Terminology

This sub-field has no standardized terminology yet. Pierce called the intentional negation of technology the “undesign of technology” [33]. He uses the term to refer to a spectrum ranging from the inhibition, to displacement, to erasure of technology. His purpose is to question how the intentional negation of design plays a role in HCI. Satchell and Dourish investigated in the “non-use of technology” [36]. They focused on the varieties of why and how people not use technology. Baumer and Silberman called for rethinking situations when technologies might be not appropriate, e. g. “the implication not to design” [2].

The terminologies from the literature describe the sub-field closely but they do not cover the topic as a whole. We propose the term “designing for disengagement” because the purpose of this research is to make people less dependent on technology and to make it easy for them to step away from it which is reflected in the word “disengagement”.

6 Concluding remarks

This work provided a literature review and established the term designing for disengagement. Technology addiction and technology dependence are the reasons for a need to create and design artifacts that makes users independent and/or help them to improve themselves. This topic provides opportunities and gaps for future research. There are open questions and implications that are presented in the following. This requires to think for long-term benefits people can have from a product even in moments when they are not using it. There is no standardized terminology for this sub-field yet. This highlights the complexity of designing for disengagement: On the one hand it should help users to be less dependent on technology and on the other hand it should help them in their self-regulation. However, there are barriers that need to be considered: The attention and engagement of the user with a digital product is a valuable resource to businesses. Companies make money on the engagement of the user, e. g. every time when a user watches an ad or pays for a service. This leads to the further question on how to make it economically attractive to design for disengagement.

We hope to encourage researchers and practitioners in HCI communities to work towards specific guidelines and standards for designing for disengagement.

7 Acknowledgments

We thank all the volunteers, and all publications support and staff, who wrote and provided helpful comments on previous versions of this document.

References

1 Firdaus Banhawi and Nazlena Mohamad Ali. Measuring user engagement attributes in social networking application. In *2011 International Conference on Semantic Technology and Information Retrieval*, pages 297–301. IEEE, 2011.

2 Eric P.S. Baumer and M. Six Silberman. When the implication is not to design (technology). In *Conference on Human Factors in Computing Systems - Proceedings*, 2011. doi:10.1145/1978942.1979275.

3 Novartis Baxter and Fresenius Kabi. Understanding human over-reliance on technology. *Long-Term Care*, 5(5):2, 2017.

4 Michael Chromik, Malin Eiband, Sarah Theres Völkel, and Daniel Buschek. Dark patterns of explainability, transparency, and user control for intelligent systems. In *IUI Workshops*, volume 2327, 2019.

5 Joseph Clarke and Steve Draper. Intermittent mindfulness practice can be beneficial, and daily practice can be harmful. an in depth, mixed methods

study of the "calm" app's (mostly positive) effects. *Internet Interventions*, 19:100293, 2020.

6 Enrico Coiera. Technology, cognition and error. *BMJ quality & safety*, 24(7):417–422, 2015.

7 Gregory Conti and Edward Sobiesk. Malicious interface design: Exploiting the user. *Proceedings of the 19th International Conference on World Wide Web, WWW '10*, (April 2010):271–280, 2010. doi:10.1145/1772690.1772719.

8 Larry Dignan. Apple, google have similar phone addiction approaches with ios, android, Jun 2018. URL: https://www.zdnet.com/article/apple-google-have-similar-phone-addiction-approaches-with-ios-android/.

9 Colin M. Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L. Toombs. The dark (patterns) side of UX design. *Conference on Human Factors in Computing Systems - Proceedings*, 2018-April:1–14, 2018. doi:10.1145/3173574.3174108.

10 Lars Hallnäs and Johan Redström. Slow technology–designing for reflection. *Personal and ubiquitous computing*, 5(3):201–212, 2001.

11 Patricia L Hardré. When, how, and why do we trust technology too much? In *Emotions, Technology, and Behaviors*, pages 85–106. Elsevier, 2016.

12 Jennifer Huberty, Jeni Green, Christine Glissmann, Linda Larkey, Megan Puzia, and Chong Lee. Efficacy of the mindfulness meditation mobile app "calm" to reduce stress among college students: Randomized controlled trial. *JMIR mHealth and uHealth*, 7(6):e14273, 2019.

13 Emmanuel Kessous. The attention economy between market capturing and commitment in the polity. *Œconomia. History, Methodology, Philosophy*, (5-1):77–101, 2015.

14 Dongil Kim, Yunhee Lee, Juyoung Lee, JeeEun Karin Nam, and Yeoju Chung. Development of korean smartphone addiction proneness scale for youth. *PloS one*, 9(5):e97920, 2014.

15 Antonio Krüger, Ilhan Aslan, and Hubert Zimmer. The effects of mobile pedestrian navigation systems on the concurrent acquisition of route and survey knowledge. In *International Conference on Mobile Human-Computer Interaction*, pages 446–450. Springer, 2004.

16 Simone Lanette and Melissa Mazmanian. The smartphone" addiction" narrative is compelling, but largely unfounded. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–6, 2018.

17 Heinrich Löwen, Jakub Krukar, and Angela Schwering. Spatial learning with orientation maps: The influence of different environmental features on spatial knowledge acquisition. *ISPRS International Journal of Geo-Information*, 8(3):149, 2019.

18 Kai Lukoff, Ulrik Lyngs, Stefania Gueorguieva, Erika S. Dillman, Alexis Hiniker, and Sean A. Munson. From ancient contemplative practice to the app store: Designing a digital container for mindfulness. In *DIS 2020 - Proceedings of the 2020 ACM Designing Interactive Systems Conference*, 2020. doi:10.1145/3357236.3395444.

19 Kai Lukoff, Cissy Yu, Julie Kientz, and Alexis Hiniker. What makes smartphone use meaningful or meaningless? *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(1):1–26, 2018.

20 Ludmila Lupinacci. 'absentmindedly scrolling through nothing': liveness and compulsory continuous connectedness in social media. *Media, Culture & Society*, page 0163443720939454, 2020.

21 Google Maps. Kartendaten © 2021 geobasis–de/bkg (© 2009). URL: https://www.google.de/maps/.

22 Don D McMahon, Cate C Smith, David F Cihak, Rachel Wright, and Melinda M Gibbons. Effects of digital navigation aids on adults with intellectual disabilities: Comparison of paper map, google maps, and augmented reality. *Journal of Special Education Technology*, 30(3):157–165, 2015.

23 Pierre-Emmanuel Michon and Michel Denis. When and why are visual landmarks used in giving directions? In *International Conference on Spatial Information Theory*, pages 292–305. Springer, 2001.

24 Satoru Minagawa and Kaori Fujinami. Estimating smartphone addiction proneness scale through the state of use of terminal and applications. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*, pages 726–729, 2018.

25 Christian Montag, Bernd Lachmann, Marc Herrlich, and Katharina Zweig. Addictive features of social media/messenger platforms and freemium games against the background of psychological and economic theories. *International journal of environmental research and public health*, 16(14):2612, 2019.

26 Kate Moran and Kim Salazar. The vortex: Why users feel trapped in their devices, Oct 2018. URL: https://www.nngroup.com/articles/device-vortex/.

27 Kathleen L Mosier and Linda J Skitka. Automation use and automation bias. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 43, pages 344–348. SAGE Publications Sage CA: Los Angeles, CA, 1999.

28 Stefan Münzer, Hubert D Zimmer, and Jörg Baus. Navigation assistance: A trade-off between wayfinding support and configural learning support. *Journal of experimental psychology: applied*, 18(1):18, 2012.

29 Heather O'Brien and Paul Cairns. *Why engagement matters: Cross-disciplinary perspectives of user engagement in digital media*. Springer, 2016.

30 William Odom, Ishac Bertran, Garnet Hertz, Henry Lin, Amy Yo Sue Chen, Matt Harkness, and Ron Wakkary. Unpacking the thinking and making behind a slow technology research product with slow game. In *Proceedings of the 2019 on Creativity and Cognition*, pages 15–28. 2019.

31 Raja Parasuraman and Dietrich H Manzey. Complacency and bias in human use of automation: An attentional integration. *Human factors*, 52(3):381–410, 2010.

32 Christopher Peters, Ginevra Castellano, and Sara De Freitas. An exploration of user engagement in hci. In *Proceedings of the International Workshop on Affective-Aware Virtual Agents and Social Robots*, pages 1–3, 2009.

33 James Pierce. Undesigning technology: Considering the negation of design by design. *Conference on Human Factors in Computing Systems - Proceedings*, pages 957–966, 2012. doi:10.1145/2207676.2208540.

34 Hazelynn Rimbar. The influence of spell-checkers on students'ability to generate repairs of spelling errors. *Journal of Nusantara Studies (JONUS)*, 2(1):1–12, 2017.

35 Rufus46. Wikimedia, 2013. Retrieved Febuary 2, 2021 from https://commons.wikimedia.org/wiki/File:Brunnen_am_Geschwister-Scholl-Platz-4.jpg. License: https://creativecommons.org/licenses/by-sa/3.0/deed.en.

36 Christine Satchell and Paul Dourish. Beyond the user: use and non-use in hci. In *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7*, pages 9–16, 2009.

37 Layna Smith. What is calm? Calm Help Center. Retrieved Febuary 2, 2021 from https://support.calm.com/hc/en-us/articles/115002474527-What-is-Calm-.

38 Nađa Terzimehić, Renate Häuslschmid, Heinrich Hussmann, and MC Schraefel. A review & analysis of mindfulness research in hci: Framing current lines of research and future opportunities. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2019.

39 Jonathan A. Tran, Katherine S. Yang, Katie Davis, and Alexis Hiniker. Modeling the engagement-disengagement cycle of compulsive phone use. In *Conference on Human Factors in Computing Systems - Proceedings*, 2019. doi:10.1145/3290605.3300542.

40 Christopher D Wickens, Simon Hollands, Justin Gand Banbury, and Raja Parasuraman. *Engineering psychology and human performance*. Psychology Press, 2015.

Exploring the Design of Companions in Video Games

Elizabeth Maria Bouquet

Ludwig-Maximilians-Universität München, München, Deutschland
e.bouquet@campus.lmu.de

Abstract

Companions are game characters that accompany the player throughout a longer part of the gameplay, complementing his character or skill set and serving as part of the narrative. Because of their enduring presence in the game, the companion character has a great impact on the player experience. There are several different aspects, that should be considered when designing such a character, including its overall behavior, awareness of surroundings and emotional intelligence. This work illuminates how a companion character can influence the player's game experience. It provides an overview about the design characteristics and existing examples of companions in popular games.

2012 ACM Computing Classification CCS → Human-centered computing → Human computer interaction (HCI)

Keywords and phrases Companions; Video Games; Non-Player Characters; Artificial Intelligence in Games.

1 Introduction

The word "*companion*" derives from the old french term "*compaignon*", which literally means "*one who breaks bread with another*" [7]. Companions have already appeared in the shape of sidekicks or allies in popculture and literacy: Sherlock Holmes and Dr. Watson [1], Frodo and Samwise [2], Batman and Robin [3], Rick and Morty [4]. While sidekicks assist the central character and complement the hero(ine)'s abilities, allies can have equal abilities to the main character - companions can take on both roles [4, 31, 8].

[1] Fictional characters created by Sir Arthur Conan Doyle, 1886
[2] Fictional characters created by J. R. R. Tolkien, 1954
[3] Fictional characters from the comic series Batman published by DC Entertainment
[4] Fictional characters from the cartoon Rick and Morty by J. Roiland and D. Harmon

Cite as: Elizabeth Maria Bouquet. Exploring the Design of Companions in Video Games. In *6th Seminar on Ubiquitous Interaction (UBIACTION 2021)*. Editors: F. Lang, R. Welsch, L. Haliburton, F. Draxler, S. Feger, S. Villa, M. Hoppe, P. Knierim, V. Mäkelä, A. Schmidt. February 11, 2021. Munich, Germany. pp. 9:1–9:16.

In the history of video games, there have been plenty of different companions: *Yoshi* as one of the oldest popular companions in video games, the fairy *Navi* accompanying the hero Link through the land of Hyrule, Gerald's horse *Roach*, *Ellie* from *The Last of Us*, and Kratos's son *Atreus* [24, 25, 6, 22, 29]. All of them appear in different shapes, sizes and behaviors and what they have in common is their enduring presence in the whole game experience.

While companions are an omnipresent design approach in the game industry, there has not been much research in this field yet. That's why this work is focusing on the design of companions and tries to explore the design of companions in popular games in order to comprehend how certain design characteristics are currently applied in the game industry. Due to the limited time frame, this work can only provide one particular example of a companion per characteristic. This gives an opportunity for future work to complement the outcomes of these explorations with more various examples and maybe find common highlights and downsides of current companion designs. These insights could then lead to possible ideas for improvements or best practices.

After defining what distinguishes a companion from other video game characters, the following gives a brief insight of the perception of virtual characters from the player's perspective and its impacts on the game experience (see section 2). Section 3 illuminates the design aspects from different perspectives in current literature and complement them with examples of existing companion characters of popular game titles.

2 Virtual Characters in Video Games

Emmerich et al. categorize roughly three different game character types: Avatars as the representation of the player, Non-Player Characters (NPCs) and Game Agents, both controlled by the computer. [8]

According to Warpefelt et al., an NPC can be defined as every character that is part of the virtual game world, not controlled by the player, and that acts as an active personality [39]. There exist different NPC types, when categorizing them by their functionality. Warpefelt et al. describe four metatypes for NPCs depending on their functionality they provide for the player: functions, adversaries, friends, and providers. They can also be a mix of these types [27, 39]. In addition, they define eleven types of NPC roles: *"Buy, sell and make stuff"*, *"Provide services"*, *"Provide combat challenges"*, *"Provide mechanical challenges"*, *"Provide loot"*, *"Give or advance quests"*, *"Provide narrative exposition"*, *"Assist the player"*, *"Act as an ally in combat"*, *"Accompany the player"*, and *"Make the place look busy"*. [38]

Pinchbeck introduces the term of a "*persistent NPC*" which "*appears repeatedly or has a defineable role in the world and plot*". According to Pinchbeck, persistent NPCs are mainly introduced as allies. Companion characters can serve as goal-givers, tutors, add emotional layer, or they can deliver a certain atmosphere [27]. To give some examples: *Aku Aku* from *Crash Bandicoot* [20] serves as a sage guide and tutor, *Elizabeth* from *Bioshock Infinite* [12] as an ally, guide, while also providing an emotional layer. *Ellie* from *Last of Us* [22] delivers atmosphere and an emotional layer by being the young protégé of the player character.

In the following, "*companions*" will be defined as non-player characters with a defined role in the plot and with an appearance over a longer period of the gameplay.

2.1 Believability of a Game Character

Like fantasy characters are vitalizing the story of a movie or book, virtual game characters can make a digital world become alive, as characters are important to create a believable world and game story. "*Well-designed, socially capable characters can significantly increase game enjoyment and are explicitly requested by players*" [15, 39].

> "Achieving characterhood is also what separates NPCs from the 'dead' things in the game world. For example, a vendor NPC is often functionally not much different from a vending machine in that it can provide the player with goods in exchange for some kind of currency, but the 'theatrics' associated with the transactions, such as speech or visual presentation, makes the game experience different. Although the difference is small in purely technical terms (playing sounds and animations) it does provide something for the player to use when they anthropomorphize an NPC as a character" [39].

Because of their (nearly) continuous presence throughout the gameplay, companions are the NPCs with the highest influence on the player's experiences and expectations [8]. They can function as the major antagonist or as a pure atmospheric element [10].

In the field of design, the most important aspect is not only the piece of art itself, but also how the observer perceives it. In the case of companions, the design should fit together with the player's expectations in order to fulfill a positive game experience.

There are four typical patterns that describe how players perceive NPCs according to Harth [10]:

- *primordial expectation of triviality*: The player sees the NPC not as a social person but rather as part of a program based on predictable mathematical operations

- *rigorous attestation of triviality*: This state is based on the previous one, but the player expects the NPC's behavior to be based on its own code
- *hybridization of players*: The NPC is perceived as an agent with own intentions and more intelligent behavior
- *temporary attestation of personhood*: The player perceives an NPC as being a personhood

Believability is an ambiguous term, which has different interpretations. The overall purpose of a believable character is to have a design that matches the player's expectations [8, 15].

According to Lankowski and Björk, believability is dependent on the following aspects: sense of self, awareness of surroundings, visual body damage, initiative, emotional attachment, contextual conversational responses, goal-driven personal development, own agenda. [14]

On the other hand, Lee and Heeter addressed believability of NPCs in general as being determined by the visual appearance, emotions, the personality, the motivation, and social relations. [15]

Furthermore, the emotional bond between the player and the companion character affects the companion's believability. Emotional Attachment to digital characters can be as real as to living beings. In the study conducted by Bopp et al. players described that they felt "*kind of connected [...]*" to the game character, emphasized with them or simply "*took up character goals as their own*". Bopp et al. also investigated different types of emotional attachments to game characters identified with key emotions and mapped to the degree of perceived emotional versus functional value of the character. [3, 2]

3 Different Aspects of Companion Design

The following chapter describes seven characteristics of designing a companion character for video games. The chapter includes the following design terms:

1. Visual Representation
2. Behavior
3. Personality and Social Relations
4. Internal and External Awareness
5. Emotional Intelligence
6. Communication
7. Story and Gameplay Relevance

This list of characteristics is an aggregation of three different perspectives on character design by Emmerich et al., Lankowski and Björk, as well as Lee and Heeter [8, 14, 15] (see Table 1 for detailed mapping). Each aspect is expanded with several aspects from previous literature and examples from

Emmerich et al.		Lankowski and Björk	Lee and Heeter	Mapped Characteristics
General Characteristics	Appearance	-	Visual Appearance	Visual Representation
	Personality	-	Personality	Personality and Social Relations
	Agenda	Own Agenda	Motivation	Awareness
General Capabilities	Awareness (Internal + External)	Sense of Self Awareness of Surroundings	-	Awareness
	Emotional Intelligence	Emotional Attachment	Emotions	Emotional Intelligence
	Social Relations	-	Social Relations	*Personality and Social Relations*
Behavior	Context Sensitivity	Contextual Conversational Responses	-	*Awareness + Communication + Emotional Intelligence*
	Autonomy	-	-	Behavior
	Initiative + Activity	Initiative	Motivation	*Behavior + Awareness*
Communication Capabilities	Communication with Player Communication with NPCs	Contextual Conversational Responses	- -	Communication
Relation to the Player Character	Interdependence	-	-	*Personality and Social Relations + Story Relevance*
	Power Dynamics	-	-	
	Obligations	-	-	
Significance	Story Relevance Gameplay Relevance	Goal-Driven Personal Development	- -	Story and Gameplay Relevance

Table 1 Mapping of the three different list of characteristics that describe a believable character

popular games. A summarized version of the content in this chapter is provided in table 2.

3.1 Visual Representation

The visual representation is one of the most crucial features of a game character. It is the first to be perceived by a player and based on this, the player infers from their first impression what a character is capable of and how the player can interact with it. This design aspect includes all visual hints regarding gender, age, ethnicity, status, observable behavior, and animations [15].

Moreover, companions can appear in different shapes: besides humans, companions can be animals, e.g. the horse *Roach* from *The Witcher Series*, robots, e.g. *Clank* from *Ratchet and Clank*, magical creatures, e.g. the fairy *Navi* from *The Legend of Zelda Series*, and even inanimate objects like the *Companion Cube* in the *Portal Series* [6, 11, 25, 34].

Furthermore, Pinchbeck differentiates between characters who are visually present in the game world and those who are integrated by means of communication and thus only have an auditory presence [27]. Not being visually present does not implicate that the characters don't have an impact on the player's game experience. Examples for characters that are rarely visible in the gameplay, but do influence the narrative and the experience, are *Cortana* from the game series *Halo* and *Wheatley* from the game *Portal 2* [5, 35].

3.2 Behavior

The behavior of a character is based on its implemented artificial intelligence. The implementation of a believable behavior pattern becomes more complex with the incrementing complexity of the mechanics, narrative and environment of a game [13]. There are several methods to implement this [1]. Controllers of NPCs can be carefully crafted behavior trees, hierarchical finite state machine or automatic planning techniques like hierarchical task networks and goal oriented action planning. Macindoe et al. introduced their own Monte-Carlo cooperative planning: sidekicks that can reason about how their actions affect their knowledge about the intentions of a human collaborator [16].

Moreover, Warpefelt et al. defines a Game Agent Matrix. It describes on the one hand how flexible a game character reacts to changes in its environment, and on the other hand, how it understands complex social structures and goals. They categorize companion characters as the most complex characters according to this matrix system [37].

In current games, there are fully and semi-autonomous companions [31]. While full autonomous characters are able to act completely independent from the player, the behavior of semi-autonomous characters can be influenced by the player in some cases [18]. According to van den Herik et al., virtual allies should be able to adapt to the player's playing style [36]. In order to provide a good solution for adaptive behavior, inferring human goals is a premise for the companion implementation. It can be gained from observing game actions from the player's side, which has already been topic in research [16]. Besides, the adaptive behavior of game characters regarding fight tactics and game mechanics, the behavior of current NPCs, especially the social behavior is still not matured enough to convince the player to perceive the character as personhood throughout the whole gameplay [39].

Ellie from the *Last of Us* is a companion that acts autonomously while accompanying the player. She also adapts her behavior to the player's actions in order to follow the player's strategy as well as possible. However, her behavior seems to be mostly scripted, which seems to be the case in the majority of games that feature companions with autonomous behavior [22]. The designers of her character decided to not make her a burden for the player, but to seem more self-dependent and helping. Her artificial intelligence includes a behavior that is supporting the player in combat by e.g. stunning enemies by throwing a brick. When in danger, she hurries to the main character Joel for cover, and for the overall gameplay she is programmed to stay always by his side. During non-cutscenes, her character is also perfectly shipped by small interactions and animations, that show her fidgeting, whistling or making spontaneous comments. All these design decisions increase her believability as a personality and can create a strong relation between the player and her [19, 9].

3.3 Personality and Social Relations

Lee and Heeter state that "*Personality defines the uniqueness and peculiar qualities of computer characters that distinguish them from other computer characters.*" A personality can be defined as a set of unique qualities of a character. These qualities can be seen in relation to the model of psychological traits [15]. Ochs et al. recites the (OCEAN) personality model which features the five traits openmindedness, conscientiousness, extraversion, agreeableness and neuroticism. Depending on the level of each trait, the personality colludes differently with the player character (as cited by [26]).

Extraversion and Neuroticism will also be important in the context of *Emotional Intelligence* (see Chapter 3.5).

Niess and Wozniak elaborated an experiential framework for companion technologies, which references smart home companions, e.g. Amazon's Alexa. The framework is based on Hassenzahl's hedonic and pragmatic concepts and on Mekler and Hornbaek's eudaimonic concept [17]. Niess and Wozniak added another layer on top by introducing empathetic experience based on four concepts from philosophy of empathy [23]. This concept could be transferred to the companions in video games in order to give them a more significant social role in the partnership with the player and to shift their relation from a simple stimulus-reaction-interaction to a more human-like relation (as mentioned by [23] in the context of HCI in general).

Lee and Heeter also note that a character that has social relations with other non-player characters supports their believability towards the player. Furthermore, when the player perceives a social connection with the game character, the character will be perceived as more believable [15].

Besides, Emmerich et al. also states that the relationship to the player is an even more important factor for the companion's believability. The companion should at least fight and travel on the side of the player and share some of his goals [8]. But on the other side, permanent NPCs can have ambiguous goals and even turn into the opponent character at the end of the game, as mentioned by Pinchbeck. [27]

The relation between the player and the companion can be defined within the following three dimensions [8]:

- *Interdependence* defines how the player and the companion depend on each other, how their abilities complement each other and how they help each other in different situations based on their different skill repertoire.
- *Power Dynamics* determine how the power is divided between the player and the companion and if one of them is more dominant than the other.
- *Obligations* between player character and companion are often described in the back stories of the character or they are introduced step by step when making progress in the game play. The player can then receive

information about the relationship to the companion and "*increase the sense of obligation*" on both sides. This has also an impact on the emotional attachment and caring about the companion from the player's perspective.

The companion *Daxter* from Jak and Daxter is a great example to show how a complementary personality can accentuate the personality of the player avatar. Jak, the player character, is a more openminded and brave character but he is also rather introvert and doesn't show any emotions or humor. Jak as a character alone could make the whole game experience rather monotonous. However, Daxter as a nearly pure extrovert and slightly nervous character, who doesn't bother to show his emotions (in most of the cases his fear or annoyance), he introduces a good portion of humor to the game play and attaches well to Jak's personality [21].

3.4 Internal and External Awareness

The internal awareness of a companion character describes its self-perception. This includes it personal agenda and goals, its overall motivations and its initiative to take actions. This self-awareness gives the companion a more believable appearance. Self-initiated acting is however not only depending on the internal goals but also on the context of the situation. Besides, according to Lankowski and Björk, believable NPCs should also be aware of the events that happen in their environment. The characters should react to the changes, even when the player is not controlling them. In the context of external awareness, Lankowksi and Björk introduce the term "*Visual Body Damage*", which describes that the character's body should visually react to external interactions and that damage with long-lasting effects in real life should also affect the virtual body in a similar way i.e. by leaving scars [14].

Princess Yorda from the game *ICO* is an example for how external awareness can be implemented. She shows fear when she perceives an enemy by covering her eyes with her hands. When she is pulled away or pushed to the ground, her body is reacting realistically to those physical impacts. Throughout the game, she is always watching the enemies that are pursuing her and the player character.

Moreover, there is an exceptional feature which also accentuates the relation between Yorda and Ico, the main character, in a more realistic way: The player has to continuously press the "R1" key on their game controller in order to hold Yordas hand (in other games, the shoulder keys of the gamepad are mainly used for triggering weapons or accelerating vehicles, i.a. in Halo [5]). When the player releases the key, Yorda will stop to follow him and will probably get in trouble. This leads to a more haptic feedback for the characters's relationship and gives Yorda a realistic awareness of her environment, when she reacts

realistically to the player's interaction. When Ico is running while holding her hand, she is also adapting her running speed to his, looking like she's actually pulled by Ico's hand. [30]

3.5 Emotional Intelligence

Showing and reacting to emotions according to the player's expectations can underpin the perception of the companion as a personhood and strengthen the bondage between player and virtual character.

> "When creating a character and wishing the player to perceive this as a person [...], it follows that not only the thematic elements (like visual appearance) must be met but also the interactive qualities of self-awareness, self-impelled actions, expression of emotions and ability to use languages." [14]

According to Ochs et al. implementing emotional intelligence to a character is a complex approach and depends on several factors, i.a. the personality traits of the character: Especially the *neuroticism* level, which describes the personality's tendency to rapidly react emotionally to events or change the emotional state, and the *extraversion* level, which describes how much emotions a character is willing to show (both described previously in chapter 3.3). Moreover, the implementation of emotional intelligence also depends on the fact, that it is still unclear in the field of psychology, how many types of emotions exist and how many of them an individual really experiences in his life [26]

Having a look on existing games, Ravenet et al. states, that "*very few games endow the NPCs with enough autonomous capabilities to trigger adaptive emotional behaviors*" [28]. Many games show deep emotional situations in cutscenes or predefined sequences, but in the game play itself, there is rather a lack of emotional variety and flexibility. Emmerich et al. refers to the game *The Witcher 3: The Wild Hunt*, which has well acted emotions and believable reactions in the cutscenes, but doesn't match the same level of emotional intelligence in the rest of the game [28, 8].

However, according to Ravenet et al.'s research, there exist several different models to implement emotional intelligence, that provide some kind of state machine for internal emotional states and transitions for different kinds of emotions. In addition, the companion should also be able to react to any events or situations in the game with an appropriate emotion state. In order to relate more to the player's character, the companion should also have the ability to adapt its emotional state to the current emotions of the player's avatar [28].

The companion character *Elizabeth* from *Bioshock Infinite* is able to show and transition between several different emotional states. She presents her emotions with facial expressions, different tones in her voice and gestures which let her appear more human-like. She reacts appropriately and understandable in certain situations. However, it should be mentioned that her emotional behavior is apparently following a predefined script [12].

Emotional Intelligence seems to be a valuable feature of a companion, that has a powerful impact on the believability of the character but also on the emotional attachment between the player and the character. It seems to be profitable to invest more future research to investigate this subject in the context of video game companions.

3.6 Communication

The ability to use natural language is a quality that should not be missed when designing a more believable character according to. Natural language is composed of non-verbal communication on the one hand, including movement, posture, and facial expressions, as well as verbal communication on the other hand. Moreover, spoken language should be emphasized with facial expressions to give a more human-like nuance when talking to the player or other NPCs. Besides, the term of "*contextualized conversational responses*" determines a communicative intelligence for the game character that is always aware of the conversational context. This would possibly avoid situations in which the game character repeats always the same sentence when asked the same question over and over again. And it could be possible for the character to react to an interruption by the player. [14]

The game studio *Valve Corporation* developed a face engine called the *Facial Expressions Primer* which is based on about 40 facial muscles – it is capable of having characters convey a wide field of emotions by manipulating those facial muscles leading to more human like facial expressions [32]. In the game *Half Life 2* produced by valve, the companion character *Alyx Vance* shows how well this engine works. She is not only able to verbally communicate with the player, but she is also capable of emphasizing her statements with facial expressions and gestures, as well as showing different emotions. The engine has also a feature that controls the gaze direction of a character. That's why Alyx is always aware of the player's position and can speak directly to them. [33]

However, this term seems not to be technically matured in several different video games as well as in the research field about video game companions and could need some more investigations in future research.

3.7 Story and Gameplay Relevance

Lastly, when designing a companion character, its relevance for the game narrative and also for the gameplay as well as the mechanics should be considered. Harth advises against the risk that a game character can only be perceived as lifeless, functional game element and not as a personhood [10], which could appear for companions with no significant relevance to the game.

Pinchbeck et al. describe that a companion does not have to be significantly relevant in form of a gameplay function but can anyway be relevant for the game story. Its diegetical properties are not only decorative but can support the player's actions and the game experience [27].

On the one hand, a character can take a major role for the game story by being present in cutscenes and dialogues. On the other hand, it can have a leading in game part and significantly influence the gameplay without carrying the story of the game [8]. For a companion design, there are several ways to make it relevant for a game and it can also be relevant for both, the gameplay and the game's story. They assume that "*well-perceived companions feature high significance with regard to at least one of the two aspects*". Virtual characters should also somehow adapt to the progressing gameplay.

The companion *Atreus* from *God of War* is a great example to demonstrate how the believability of a companion character grows when his personality and behavior keep developing while the game story's continuing. Atreus is the main character's son (Kratos) and accompanies him on his journey. Kratos teaches him how to survive in this world and Atreus supports him in solving puzzles and later also in combat. Atreus even has an own skill tree, which shows how dynamically the companion can evolve its character. Besides that, he repeatedly reminds the player of exploring more of the environment [29].

4 Conclusion and Future Work

There are many aspects that have to be considered when designing a companion character for a video game: its visual appearance, personality, behavior patterns, awareness of surroundings, emotions, social relations and communicative skills. The listings of characteristics for video game companions by Emmerich et al. [8] as well as Lankowski and Björk's and Lee and Heeter's definitions of a believable non-player character [14, 15] show some congruent thoughts and can be a guideline when conceptualizing a companion.

Adaptive and flexible behavior, emotional intelligence, communication and overall internal and external awareness still seem to be wide fields that have to be explored, when looking at companions and other game characters in current

Characteristic	Short Description
Visual Representation	All visual clues - the first feature to be perceived by the player
Behavior	The overall behavior pattern of the companion depending on the game's mechanics and the level of autonomy the character has
Personality and Social Relations	The personality traits of the companion and his relation to the player character and other NPCs
Internal and External Awareness	Awareness of the companions own personal goals as well as of its reactivity to external events
Emotional Intelligence	Companion's ability to show and react to emotions as well as believably transition between emotional states
Communication	The ability to communicate with the player character and other NPCs using natural language, including non-verbal communication
Story and Gameplay Relevance	The relevance of the companion towards the gameplay and/or the narrative

Table 2 A summary of the design characteristics of a companion game character presented in Chapter 3

AAA game titles [5], while however, there has been significant improvements over the time and more hardware and software resources.

Moreover looking at some popular companion characters shows that it is not always needed to completely fulfill the whole bucket list of characteristics and that there are several different ways to create a character that the players care about and that they enjoy being accompanied by. Due to a limited scope, the analysis in this work is limited to one companion per characteristic. In further research, this analysis could be extended with more companions in order to get a wider impression of what has already been done in game design. Furthermore, all collected insights could lead to a rather flexible list of characteristics that is dependent on the type of companion character a game designer wants to create.

Besides the field of video games, research in the field of companion characters could also be useful in the research and development of other types of companions, such as robotic companions or smart home companions like Amazon's Alexa, in order to improve the interaction between human and virtual character and make it more animated.

[5] "AAA" or "Triple-A" is an informal categorization of games that are produced by major game publishers and therefore have significantly high budget for development

References

1 Sander Cornelus Johannus Bakkes. *Rapid adaptation of video game AI*. PhD thesis, Tilburg University, 2010.

2 Julia Ayumi Bopp, Elisa D Mekler, and Klaus Opwis. Negative emotion, positive experience? emotionally moving moments in digital games. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 2996–3006, 2016.

3 Julia Ayumi Bopp, Livia J Müller, Lena Fanya Aeschbach, Klaus Opwis, and Elisa D Mekler. Exploring emotional attachment to game characters. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, pages 313–324, 2019.

4 Ron Buchanan. "side by side":the role of the sidekick. *Studies in Popular Culture*, 26(1):15–26, 2003. URL: `http://www.jstor.org/stable/23414984`.

5 Microsoft Studios Bungie. Halo - game series. Game [Xbox, Xbox 360, Xbox One, PC], 2001. [29th January 2021]. URL: `https://www.xbox.com/en-us/games/halo?source=lp`.

6 Bandai Namco Games CD Project Red. The witcher 3: Wild hunt. Game [Playstation 4, Xbox One, PC], 2015. [29th January 2021]. URL: `https://store.steampowered.com/app/292030/The_Witcher_3_Wild_Hunt/`.

7 Oxford Dictionaries. *Paperback Oxford English Dictionary*. Oxford University Press, 7 edition, 2001.

8 Katharina Emmerich, Patrizia Ring, and Maic Masuch. I'm glad you are on my side: How to design compelling game companions. In *The Annual Symposium*, pages 141–152, 10 2018. URL: `http://dx.doi.org/10.1145/3242671.3242709`, `doi:10.1145/3242671.3242709`.

9 Megan Farokhmanesh. How naughty dog created a partner, not a burden, with ellie in the last of us. Polygon, March 2014. [4th January 2021]. URL: `https://www.polygon.com/2014/3/22/5531146/the-last-of-us-ellie-was-designed-to`.

10 Jonathan Harth. Empathy with non-player characters? an empirical approach to the foundations of human/non-human relationships. *Journal For Virtual Worlds Research*, 10(2), 2017. `doi:10.4101/jvwr.v10i2.7272`.

11 Sony Computer Entertainment Insomniac Games. Ratchet & clank (reboot). Game [Playstation 4], 2016. [29th January 2021]. URL: `https://store.playstation.com/de-de/product/EP9000-CUSA01073_00-RCPS400000000000`.

12 2K Games Irrational Games. Bioshock infinite. Game [Playstation 3, PC, Xbox 360], 2013. [29th January 2021]. URL: `https://2k.com/en-US/game/bioshock-infinite/`.

13 John Laird and Michael VanLent. Human-level ai's killer application: Interactive computer games. *AI magazine*, 22(2):15–15, 2001. URL:

https://ojs.aaai.org/index.php/aimagazine/article/view/1558, doi:10.1609/aimag.v22i2.1558.

14 Petri Lankoski and Staffan Björk. Gameplay design patterns for believable non-player characters. In *DiGRA Conference*, pages 416–423, 2007.

15 Michael Sangyeob Lee and Carrie Heeter. What do you mean by believable characters?: The effect of character rating and hostility on the perception of character believability. *Journal of Gaming & Virtual Worlds*, 4(1):81–97, 2012. doi:10.1386/jgvw.4.1.81_1.

16 Owen Macindoe, Leslie Kaelbling, and Tomás Lozano-Pérez. POMCoP: Belief space planning for sidekicks in cooperative games. In *Proceedings of the Eighth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 10 2012. URL: https://dl.acm.org/doi/10.5555/3014629.3014637, doi:10.5555/3014629.3014637.

17 Elisa D Mekler and Kasper Hornbæk. Momentary pleasure or lasting meaning? distinguishing eudaimonic and hedonic user experiences. In *Proceedings of the 2016 chi conference on human factors in computing systems*, pages 4509–4520, 2016.

18 Ville Mäkelä and Albrecht Schmidt. I don't care as long as it's good: Player preferences for real-time and turn-based combat systems in computer rpgs. In *Conference: Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '20)*, 11 2020. URL: http://dx.doi.org/10.1145/3410404.3414248, doi:10.1145/3410404.3414248.

19 Colin Moriarty. Naughty dog talks combat in the last of us. IGN, February 2012. [4th January 2021]. URL: https://www.ign.com/articles/2012/02/13/naughty-dog-talks-combat-in-the-last-of-us.

20 Activision Naughty Dog. Crash bandicoot: N. sane trilogy. Game [Playstation 4, Xbox One], 2017. URL: https://store.playstation.com/de-de/product/EP0002-CUSA07399_00-CRASHNSANETRLOGY.

21 Sony Computer Entertainment Naughty Dog. Jak and daxter. Game [Playstation 2, Playstation 4], 2001. URL: https://store.playstation.com/de-de/product/EP9000-CUSA07934_00-JAKDAXTER0BUNDLE.

22 Sony Computer Entertainment Naughty Dog. The last of us. Game [Playstation 4], 2014. [29th January 2021]. URL: https://store.playstation.com/de-de/product/EP9000-CUSA00556_00-THELASTOFUS00000.

23 Jasmin Niess and Paweł W. Woźniak. Embracing companion technologies. *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*, Oct 2020. URL: http://dx.doi.org/10.1145/3419249.3420134, doi:10.1145/3419249.3420134.

24 Nintendo. Super mario world. Game [SNES], 1992. [29th January 2021]. URL: https://www.nintendo.de/Spiele/Super-Nintendo/Super-Mario-World-752133.html.

25 Nintendo. The legend of zelda: Ocarina of time, 1998. [3th December 2020]. URL: https://www.nintendo.de/Spiele/Nintendo-64/The-Legend-of-Zelda-Ocarina-of-Time-269536.html.

26 Magalie Ochs, Nicolas Sabouret, and Vincent Corruble. Simulation of the dynamics of nonplayer characters' emotions and social relations in games. *IEEE Transactions on Computational Intelligence and AI in Games*, 1(4):281–297, 2009. URL: http://dx.doi.org/10.1109/TCIAIG.2009.2036247, doi:10.1109/TCIAIG.2009.2036247.

27 Dan Pinchbeck. An analysis of persistent non-player characters in the first-person gaming genre 1998-2007: a case for the fusion of mechanics and diegetics. *Eludamos: Journal of Computer Game Cultures*, 3(2):261–279, 2009. URL: http://www.eludamos.org/index.php/eludamos/article/view/vol3no2-9.

28 Brian Ravenet, Florian Pecune, Mathieu Chollet, and Catherine Pelachaud. Emotion and attitude modeling for non-player characters. In Kostas Karpouzis Georgios N. Yannakakis, editor, *Emotion in Games*, pages 139–154. Springer, 2016.

29 Sony Interactive Entertainment SIE Santa Monica Studio. God of war. Game [Playstation 4], 2018. [29th January 2021]. URL: https://store.playstation.com/de-de/product/EP9000-CUSA07410_00-0000000GODOFWARN.

30 Sony Interactive Entertainment Team ICO. Ico. Game [Playstation 2, Playstation 3], 2001. [29th January 2021]. URL: https://store.playstation.com/en-cy/product/EP9000-BCES00141_00-LBPDLCSONYCK0004.

31 Jonathan Tremblay and Clark Verbrugge. Adaptive companions in FPS games. In *Proceedings of the Eighth International Conference on the Foundations of Digital Games*, pages 229–236. Society for the Advancement of the Science of Digital Games, 2013.

32 Sodabus (Username). Facial expressions primer. Valve Software Documentation, July 2005. [4th January 2021]. URL: https://developer.valvesoftware.com/wiki/Facial_Expressions_Primer.

33 Electronic Arts Valve. Half life 2. Game [PC], 2004. [29th January 2021]. URL: https://store.steampowered.com/app/220/HalfLife_2/.

34 Electronic Arts Valve. Portal. Game [PC, Playstation 3], 2007. [29th January 2021]. URL: https://store.steampowered.com/app/400/Portal/.

35 Electronic Arts Valve. Portal 2. Game [PC, Playstation 3, Xbox 360, Xbox One], 2011. [29th January 2021]. URL: https://store.steampowered.com/app/620/Portal_2/.

36 HJ van den Herik, HHLM Donkers, and PHM Spronck. Opponent modelling and commercial games. In G. Kendall and S. Lucas, editors, *Proceedings of the IEEE 2005 symposium on computational intelligence and games*, pages 15–25. Essex University, 2005. Pagination: 11.

37 Henrik Warpefelt, Magnus Johansson, and Harko Verhagen. Analyzing the believability of game character behavior using the game agent matrix. In *DiGRA Conference*, 2013.
38 Henrik Warpefelt and Harko Verhagen. Towards an updated typology of non-player character roles. In *Proceedings of the International Conference on Game and Entertainment Technologies*, 2015.
39 Henrik Warpefelt and Harko Verhagen. A model of non-player character believability. *Journal of Gaming & Virtual Worlds*, 9(1):39–53, 2017. doi:10.1386/jgvw.9.1.39_1.

Approaches to Enhance the Musical Experience of People with Hearing Impairment

Carmen Schanderl

Ludwig-Maximilians-Universität München, München, Deutschland
Carmen.Schanderl@campus.lmu.de

Abstract

This report highlights various ways in which the musical experience of people with hearing impairments can be improved. It identifies reasons why hearing problems exist and how they can be improved with conventional methods such as air conduction hearing aids and bone conduction hearing aids. It also addresses alternative ways in which people in the deaf community can hear and feel the music better, to give it deeper meaning. These include, in particular, visual representations of the music and haptic feedback through vibrations over the whole body or targeted areas of the body. To make the transition from studies to real life, examples are given of well-known musicians, with hearing problems, and how they overcame their conditions and were able to perform musical activities. Integrated into social life were alternative methods such as vibrotactile feedback and visual representations through a concert for the deaf.

2012 ACM Computing Classification Human-centered computing → Accessibility → Accessibility technologies

Keywords and phrases Haptic Feedback; Vibrotactil; Visual; Deaf; Hearing Impaired.

1 Introduction

Music accompanies us all our lives. Whether it's the alarm clock that wakes someone up in the morning with a selected song, or just the radio in the car. Babies are also often sung to sleep or have music played to them in the womb even before they are born. Music can also increase the feeling of happiness. Salimpoor et al. found out, that listening to music can release dopamine [25]. Dopamine is a neuronal messenger that gives feelings of happiness, it also has a motivating effect and helps to concentrate [9, 25]. In everyday life, it is not only music that is part of it, sounds can also serve as warning signals. For example, the smoke alarm, a fire brigade siren or the horn of a car. So music

Cite as: Carmen Schanderl. Approaches to Enhance the Musical Experience of People with Hearing Impairment. In *6th Seminar on Ubiquitous Interaction (UBIACTION 2021)*. Editors: F. Lang, R. Welsch, L. Haliburton, F. Draxler, S. Feger, S. Villa, M. Hoppe, P. Knierim, V. Mäkelä, A. Schmidt. February 11, 2021. Munich, Germany. pp. 10:1–10:16.

and sounds are not only part of the entertainment, but can also be lifesavers. But what if someone is not able to hear? The World Health Organization estimates that around 1.5 billion people worldwide are affected by some degree of hearing loss [29]. 430 million of them need hearing rehabilitation to address hearing impairment. Forecasts indicate that by 2050, up to 2.5 billion people are estimated to have some degree of hearing impairment, of which 700 million will need rehabilitation. [28] The following pages describe causes of hearing loss, which may be due to nature, self-inflicted, or illness. Furthermore, conventional supportive aids for hearing improvement are presented and compared with alternative, new approaches which can enhance the musical experience of hearing impaired people. For this purpose, various studies have presented that work with haptic and visual feedback. Participants in these studies were very enthusiastic and discovered that not only hearing is necessary for a satisfying musical experience [19]. "Factors that contribute to the enjoyment of music are feeling vibrations or seeing visual displays." [19]

2 How the Ear works

The sensory organ ear is on the one hand responsible for balance and the other hand for hearing. Anatomically, the ear is divided into three components. These include the outer ear, middle ear, and inner ear. The outer ear shows the visible pinna, which captures the sound and transmits it through the ear canal to the eardrum. After the vibrations which are caused by the sound waves, passes through the eardrum, they receive the middle ear. The eardrum transmits the vibrations to three small bones. These bones move when the eardrum transmits the vibrations and then reaches the inner ear. The inner ear includes the Cochlea which looks like a snail shell and makes it possible to hear and included the sense of balance. The cochlea is filled with fluid and contains many small hair cells. These receive sound waves and trigger electrical signals in the hearing nerve, which transmits the sound waves to the brain. [22]

2.1 Reasons for Hearing Impairment and different Types

People who are not able to hear, mostly have problems with sensorineural hearing, conductive hearing, or both a so-called mixed hearing loss. In sensorineural hearing loss, the concerned has a poorer quality of sound and difficulty hearing soft tones in the sounds that are still audible. The cause is in the inner ear, where the hair cells, the auditory nerve, or both may be damaged. In the case of impaired conductivity, on the other hand, one possible cause could be that sound is disturbed from reaching the inner ear. For example, due to ear problems such as inflammation of the outer or middle

ear. In conductive hearing loss, the cause can often be corrected, whereas, in sensorineural hearing loss, the cause is more often permanent hearing loss. Fortunately, there are treatment options in both cases, which are discussed in the next subchapter. [23] Further causes for hearing impairment can be aging, genetic causes, chronic ear infection, complications at birth, use of particular drugs, certain infectious diseases, or exposure to excessive noise [29].

2.2 Conventional Modalities to support Hearing and Listening Comprehension

There exist different possibilities to support hearing for people with hearing impairment. For example, air conduction and bone conduction hearing aids. The bone conduction hearing aids for example can be used, when the air cannot transmit the sound, like with conductive hearing loss. The system behind this is that the vibration of the sound goes directly through the cranial bone and not through the ear canal via the air through the eardrum into the inner ear, as is the case with air conduction hearing aids. [13] Nanayakkara et al. mentioned in their paper, that the best-known aids for a better music experience are sign language and subtitles because they are easily available [19].

There exist much different sign languages where people can communicate. For the different languages exists different signs for words and grammar. For communication hand movements are used and even the whole body to underline expressions and thoughts. Examples of the different languages are the American Sign Language (ASL) or the British Sign Language (BSL) and people with normal hearing often use the Sign Supported English or SSE to communicate with people with hearing impairment. [31]

Subtitles are provided to enable viewers to understand film content more easily. There are various types of subtitles, including closed captioning (CC) or subtitles for the deaf or hard of hearing, also known as SDH. Subtitles are used when another language is spoken in the video that you do not speak. closed captures, on the other hand, are in the same language and also add events that are happening, such as “door slamming”, “car noise” or the name of the person who is speaking. CCs are therefore well suited for deaf or hard of hearing viewers. The last category the SDH combines the other two categories. It is for people designed who do not hear anything from the video and do not understand the spoken language. This means that the translated conversations plus sound effects can be read along. [1]

3 People with Hearing Impairment and the Music

People with hearing problems also participate in musical activities. Among other things, they move or dance to music, sign or sing a song, or simply listen to music. [2] Nanayakkara et al. also found in their study that there is a difference in the degree of hearing loss whether people do participate or not. While almost 80 per cent of participants with partial hearing loss reported participating in musical activities, only up to 30 per cent of participants with profound hearing loss did so. [19]

The following sections will discuss how the brain of people with hearing impairments transforms and also uses the part of the brain that is actually responsible for hearing. It continues with musicians who are deaf or partially hearing impaired and the techniques they have used to overcome their conditions.

3.1 Transformations in the Brain of People with Hearing Impairments

Besides mood, listening to music or practising music can improve learning for example the mathematical performance, and activates different parts of the brain [26]. The University of Washington also shows that “brains of deaf people rewire to ‘hear’ music” [27]. An explanation of why people can enjoy musical activities could be, that they sense vibrations in the auditory cortex. The auditory cortex is normally used to hear. [27] Dr Dean Shibata, at the time an assistant professor of radiology at the University of Washington, who made the findings says “vibrational information has essentially the same features as sound information — so it makes sense that in the deaf, one modality may replace the other modality in the same processing area of the brain. It’s the nature of the information, not the modality of the information, that seems to be important to the developing brain.” [27] Furthermore, Good et al. described the areas in the brain that change in people with hearing impairment and are relevant for music perception. In the part of the secondary auditory cortex, an increased activation to visual and vibrotactile stimuli can be observed [10]. In one experiment, Dr Shibata scanned the brains of participants with normal hearing and participants who are deaf while their hands were exposed to vibrations. In both groups, activity was detected in the area of the brain where vibrations are processed. In the deaf participants, the area in which sounds are normally perceived was also activated, which was not the case with the normal hearing participants. [27] Visual stimuli also activate the auditory cortex. In one study, hearing and deaf people were shown a moving dot pattern and the deaf participants again showed activity in the auditory cortex. [7]

3.2 Examples of Musicians with Deafness or Hearing Impairment

Another interesting way of interacting with music is shown by some famous musicians. The stories of two musicians are described below. They include Ludwig van Beethoven, a German composer and pianist, and Evelyn Glennie, a British percussionist. Both have in common that they were not hearing impaired from birth, but were challenged with hearing problems at a young age.

Ludwig van Beethoven Ludwig van Beethoven started losing his hearing at the age of 26 and was totally deaf in his last years of life [20]. Beethoven continued to compose after his hearing deteriorated. This worked through the vibrations of his piano [10] and with some of the latest developments in acoustic hearing aids [20]. The gadget with which he tries to overcome his hearing problems was an Ear Trumpet, a Resonating Plate, and the famous Wooden Drumstick which was developed by Johan Nepomuk Maelzel. The principle behind the hearing aid was the reverse function of a megaphone. Megaphones intensify the outgoing sound and the hearing trumpet intensifies the incoming sound and directs it directly into the auditory canal. [20] Beethoven also uses a Resonating Plate. This has been constructed in such a way that when placed on the piano, it operated as a sound conductor, making the sound seem more pronounced. [20] The constructor was Conrad Graf, which also builds a special piano for Beethoven. The mechanism behind this was that it amplified the sound of the piano so that he could hear the music better. [20] The most interesting method for this article is the Wooden Drumstick. Beethoven attached a wooden drum stick to the resonating body of his piano and held it with his teeth so that the vibrations were transmitted [21]. And further, over bone conduction, the vibrations were transmitted to his inner ear [20]. This method, which was described in chapter 2.3, is another way of hearing sounds other than through the air, through the eardrums. There, the vibrations are transmitted directly to the inner ear via the jawbone [15, 30]. Further “a far greater range of frequencies is transmitted via bone conduction of sound compared with purely tactile stimulation.” [19]

Evelyn Glennie Evelyn Glennie is a British profoundly deaf percussionist. At the age of 8, the hearing problems started and at the age of 11 they became more severe and she had to wear hearing aids. When sounds were amplified by the hearing aids, it was uncomfortable, even painful for her, and everything was distorted and confusing to decipher. It even went so far as to affect her balance and sense of touch. So she decided to take off the hearing aids and rely more on her body and less on her ears. So her body acted like a resonating chamber with which she could "hear". When playing percussion instruments, she realized that she could feel the vibrations in different parts of her body. Her teacher trained these senses with her. She put her hands on the wall outside the music room and her teacher played two notes on the drums and asked her which was the higher and which was the lower. She felt the higher notes in the upper part of her hand while she felt the lower notes down to her wrist. She was also aware of a door slamming or the phone ringing and learned to use her ability more and more as she became more dependent on it. [5] She feels "the low sounds [...] mainly in (her) legs and feet and high sounds might be particular places on (her) face, neck, and chest" [5].

4 Alternative Modalities to support Hearing and Listening Comprehension

Young adults, in particular, want to have a satisfying musical experience when they attend cultural programs such as a theatre or a concert. Through vibrotactile and visual technologies, people with hearing impairments can enjoy music and normal hearing is not a prerequisite to participate in such activities. [10]

In the following sections, alternative methods and modalities are presented that can be used to achieve this. It also analyses how different body parts are used for music perception.

4.1 Methods and Modalities

In addition to traditional methods such as subtitles or sign language, alternative methods to enhance music enjoyment are becoming increasingly popular. Of particular note are vibrotactile and visual techniques that can enhance the musical experience. [10] Strong basses are particularly preferred, where the vibrations are felt more strongly [19]. A method that also requires vibrations is "Speaker Listening", where listeners feel the vibrations of a loudspeaker and can thus perceive the music. [10] Linking vibrotactile and visual technologies, for example, a disco for the deaf was also integrated into social life in a project by 7UP, an American soft drink brand, and Martin Garrix, a world-renowned DJ, in collaboration with Fake Love, a media agency. They organized a concert for deaf people. Different devices were installed to the beat of Martin Garrix music. For example, special vibrating platforms on which the guests could stand and feel the vibrations caused by the sound through their bodies. Furthermore, loudspeaker walls with water experiments were created. Water and coloured water that was in plexiglass cubes were vibrated by different frequencies from speaker cones that were under the cubes. And this created different patterns and splashes in time with the music. Further, cameras were installed that observed the water movement and recorded this data and converted it into visual images. The images were projected on large projection screens and side-mounted screens, and everything went back to what the DJ was playing and what the water experiments were doing. And they also gave the participants vibrating backpacks that they could wear and feel the vibrations directly on their bodies. [14] The history of visualizing Music goes back to the early 20th century. Oskar Fischinger, for example, depicted music visually by working with shapes and geometric patterns. To visualise classical music, Walt Disney produced a film without spoken passages, using only animation. Another film director named Norman McLaren expressed his visual idea of music through hand drawings. [19] Mitroo et al. presented an approach that combines musical notes with colours. They incorporated pitches, notes, chords, etc. to create moving images in synchronisation with the music. [18] Well-known music players such as iTunes, Winamp or the Windows Media Player can visualise songs. Likewise, visual effects are seen on screens at concerts and live performances by musicians. A Tac-Tile Sounds System (TTSS) is a kind of portable music floor. The TTSS can convert sound into vibrations, which can then be felt all over the body while sitting, lying or standing. This product was developed by Russ Palmer, an international music therapist who himself has a hearing impairment. The aim was to create both a therapeutic aid and a recreational product that can be used in clinics and at home. For example, to perceive speech, music or rhythm. [24]

4.2 Where People feel the Music

Touch can be perceived through the surface receptors of the skin, the so-called mechanoreceptors. The Vater-Pacini corpuscles, Merkel cell receptors, Krause corpuscles, Meissner corpuscles and Ruffini corpuscles belong to the touch sensors of the skin. [17] The Vater-Pacini corpuscles perceive vibrations and convert them into an electrical signal, which reaches the brain through nerve pathways via the spinal cord [16]. Humans can perceive frequencies of music differently in several regions of the body. Evelyn Glennie and Russ Palmer reported on their experiences where they specifically perceive low and high tones. Evelyn Glennie also mentioned that "with very low frequency vibration the ear starts becoming inefficient and the rest of the body's sense of touch starts to take over." [6] She even can manage to distinguish the approximate pitch of tones, this is probably due to the fact that she was able to hear the tones once, and so can compare the feeling through her memory. [5] Nanayakkara et al. mentioned another interesting observation, that "preliminary testing suggested that the Haptic Chair was capable of providing, not only haptic sensory input (via the sense of touch) but also bone conduction of sound via ear or directly to the CNS." [19]

Higher Tones Russ Palmer describes that in Music the higher tones can be felt in the fingers, head and hairs [24]. Just as Evelyn Glennie also describes that she feels the higher tones in the upper part of her body, such as the face, neck and chest. In the study of Nanayakkara et al., the vibrations of notes, rhythms and instrument combinations are sensed by the fingertips. Small domes made of solid, thin plastic are attached to a chair with armrests. These transmit the vibration to the hands, and the armrests to the forearms. [19]

Middle Tones According to Russ Palmer, the middle tones are felt in the abdomen, chest and arms [24].

Lower Tones The low tones can be felt down to the lower body. By legs and feet, as Palmer describes [24]. Evelyn Glennie also feels the lower tones in her legs and feet [5].

Back of the body In various studies such as the Haptic Chair by Nanayakkara et al. or EmotiChair by Karam et al. the back was used to feel the music, having vibrotactile outputs attached to chairs stimulating the back. Karam et al. justified their approach in such a way that the back offers plenty of space to attach vibrotactile simulators. Furthermore, arms and legs can still be moved freely when using vibrotactile stimulators on the back. People are also already familiar with vibrations in this area, such as the vibrations caused by driving a car. [3]

5 Studies that evaluate the alternative Modalities

In comparison with the conventional modalities to support hearing and listening comprehension, the alternative methods not just works with vibrations through the jawbone or skull bone, but through the whole body. Some researchers try to improve the musical experience of people with hearing problems and conducted various studies. The following is a study by Nanayakkara et al. who developed a music display and Haptic Chair. Then the “Emoti Chair” by Karam et al. is described and the third study is the “Tactile Composition” by Gunther et al. which is about a full-body vibrotactile stimulator.

5.1 Music Display and Haptic Chair

The study of Nanayakkara et al. designs a music display and a haptic chair. Special about the chair is, that there is no electronic pre-process to amplify or change the vibrotactile feedback. The researchers justify this with the cause, that any modification could disrupt the experience. The basic function of the

Haptic Chair is to transform music and sound inputs into a tactile output that can be perceived through the chair.

Background The study consisted of two parts, a background study with 41 participants and a user evaluation study with 43 participants. In both studies, the participants had varying degrees of hearing impairment. The background study was conducted on one side to find out whether people with hearing loss participate in musical activities or not, and which music genres they prefer. These results if people with hearing loss participate in musical activities have already been discussed in chapter 3. The genres they prefer are music with stronger bass, this is due to the stronger vibrations that music with bass causes. The results of the background study provided the basis for the choice of music for the user evaluation study. The user evaluation study aimed to find out whether the design of the visual display and the haptic chair can improve the musical experience of people with hearing impairment. [19]

Visual Display The visual display uses colours and shapes to show note beginnings, note duration, the pitch of a note, volume, instrument type and key changes. High and low notes vary with the shapes, and loud and soft notes vary with the brightness. The colours represent different instruments. The background colour also plays a role and adapts to the musical key. Synaesthetes were used to create a link between the musical key and the background. Synaesthetes are people who see colours when they hear sounds. [19]

Haptic Chair The chair is designed to enhance the musical experience by sensing the vibrations generated by music in many different parts of the body. They had to choose a flexible material and at the same time, stable, for which the well-known Ikea chair “Poäng” with its accompanying footrest is ideally suited. The selected loudspeakers also had to transmit the vibrations to different materials, so they chose special contact loudspeakers suitable for this purpose. The vibrations were transmitted from the arm bones, finger bones, hand bones, wrist bones, lumbar spine and legs. The speakers were attached to the armrest, footrest and backrest. Special attention was also paid to the high-frequency sounds and extra plastic domes were placed on the armrest at the level of the hands to further amplify the vibration. Through the amplification and the transmission to the arm bones, the vibrations were conducted through the whole body. [19]

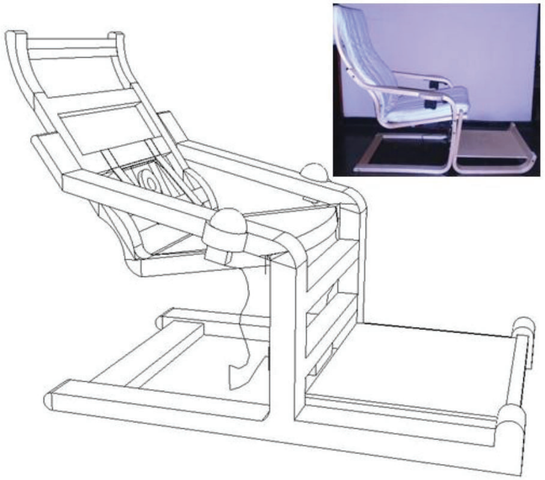

Figure 1 "Draft of Haptic Chair and actual Haptic Chair" [19]

Procedure The selected test songs based on the background study ran through four different trials (Tabelle 1). The four various test versions could be composed of three different components: Music, visual display and the Haptic Chair. The music was always played on normal membrane speakers. The other variations were with or without the Haptic Chair, or with or without the visual display. [19]

Trial	Visual Display	Haptic Chair	Task
A	off	off	Follow the music
B	on	off	Follow the music while paying attention to the visual display
C	off	on	Follow the music while paying attention to the vibrations provided via the Haptic Chair
D	on	on	Follow the music while paying attention to the visual display and vibrations provided via the Haptic Chair

Table 1 "Four trials for a piece of music" [19]

Results That the attempt to improve the musical experience of the hearing impaired was successful can be seen in the results of the study. Most of the participants preferred run C and D, in both trails the Haptic Chair was included. None of the participants preferred run A or run B. What these two have in common is that both runs were without haptic feedback from the Haptic Chair. 46 per cent of the participants chose run D as their favourite, which is music, visuals and Haptic Chair. The participants justified this by saying that in addition to the haptic feedback, a clearer meaning can be assigned to the visualisation, which was not the case in run B. And as many as 54 per cent chose run C music only along with the Haptic Chair. This shows that the haptic feedback through vibrations has a very great impact on the musical perception of people with hearing impairments. [19]

Related works Similar works included the EmotiChair by Karam et al. and a Tactile Composition or aesthetic composition for the sense of touch by Gunther et al. [12, 11].

The EmotiChair is an audio-tactile display that converts sound into tactile vibrations using a Model Human Cochlea (MHC). "The MHC adopts a scaled-down version of the hair cells, and uses audio speakers to stimulate the mechanoreceptors of the skin." [3] The MHC prototype was a chair with output channels attached in a logical order (Abblidung 2) [3, 12]. "Preliminary results from this work suggest that the MHC can potentially offer a more effective means of expressing basic emotional information from music in a vibrotactile display when the signals more closely match the perceptual elements of the music such as melody and instrument parts." [3]

The development by Gunter et al. introduces a system that allows to perceive and feel vibrations on the skin surface. It is a vibrotactile whole-body simulator with 13 sound transducers that are worn on the body and produces music in tactile representation. The sound transducers were distributed over the entire surface of the body because this is the best way to recreate the posture of a dancer. The high-frequency transducers are distributed along both arms and legs. The transducers on the legs are placed on the upper side so that there is still the possibility to sit down. The low-frequency transducer has been intuitively placed on the lower back because it has been observed that low-frequency vibrations can be felt in the lower back during a concert, for example. In Figure 3 (Abblidung 3) you can see the transducers. "The small black circles represents the high-frequency transducers and the large circle on the lower back represent the Interactor." [11]

Figure 2 MHC-Prototype [12]

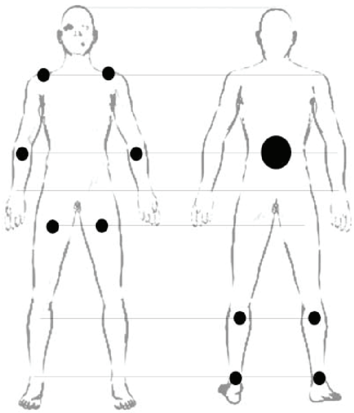

Figure 3 Body Simulator [11]

6 Conclusion

This paper aims to find existing ways to bring music closer to people with hearing impairments. For this purpose, different musicians with certain degrees of hearing loss were presented and how they were able to overcome their deficits. It also compared conventional hearing aids and alternative methods of "hearing" and presented studies that have explored these alternative methods.

A quote from Evelyn Glennie: "Hearing is basically a specialized form of touch. Sound is simply vibrating air which the ear picks up and converts to electrical signals, which are then interpreted by the brain. The sense of hearing is not the only sense that can do this, touch can do this too." [6]. This can be proved by making a self-experiment, for example when a big truck drives by, the low frequencies of its hum can also be perceived as vibrations, or when the radio is on, the vibrations can be felt through a piece of paper or a balloon between the hands. The National Technical Institute of the Deaf in Rochester also implements this, where everyone who attends the musical is given a balloon to feel the musical vibrations with their fingertips. [27] Russ Palmers TacTile Sound System can help people with hearing impairment in this way not just to perceive music or rhythm, but also speech or draw attention to everyday sounds [24]. Also, a sign language interpreter who assisted in the Nanayakkara et al. study tested the Haptic Chair in speech training for deaf people and noticed a positive difference when the Haptic Chair was used in speech training [19].

Conventional hearing aids compared to the alternative hearing aids, amplify sounds, or direct them to the ear via bone conduction to hear sounds. Alternative aids, such as vibrotactile devices or visual displays, bring music closer to people in a different way. However, this is not inferior to listening to music normally. This has been proven by research at Washington University, which has shown that the auditory cortex in the brain is activated when vibrations are felt. Hearing is no longer necessary to have a satisfying musical experience or to participate in musical activities, as demonstrated by Evelyn Glennie and Ludwig van Beethoven. Through their fine sense and use of the other senses, such as the sense of touch, they manage to make music and perceive music. This approach should be integrated more into social and everyday life, as it was done by Martin Garrix and 7UP. That it is normal to go to a concert and have vibrating platforms to match the frequencies of the music or vibrating backpacks to transfer the vibrations directly to the body.

Returning to the alarm signals mentioned in the introduction, there are different ways of alerting people with hearing impairments. For example, light signalling systems, vibration alarm devices, light and vibration alarm clocks [4]. In Switzerland, for example, there is a nationwide alarm system that is suitable for people with hearing impairments. The alarm is sent via an app in which the hearing-impaired person receives a push message on their mobile phone. [8]

In summary, there are vibrotactile methods and visual displays already in development that can enhance the musical experience. The different frequencies can be perceived in different parts of the body and then, through practice, assigned to different pitches. These approaches can be used not only to bring music closer to people with hearing impairments, but also to enrich everyday life, improve education such as language training, and provide increased safety with adapted alerting systems.

References

1 Access Innovation Media Pty Limited 2017. Subtitles for the deaf or hard-of-hearing (sdh) - subtitles, closed captions, and sdh., 2017. Retrieved February 2, 2021 from `https://blog.ai-media.tv/blog/what-is-sdh`.

2 Darrow Alice-Ann. The role of music in deaf culture: Implications for music educators, 1993.

3 Carmen Branje, Michael Maksimowski, Maria Karam, Deborah I. Fels, and Frank Russoy. Vibrotactile display of music on the human back. pages 154–159, 2010. `doi:10.1109/ACHI.2010.40`.

4 Deutscher Gehörlosen-Bund e.V. Technische hilfsmittel. Retrieved February 2, 2021 from `http://www.gehoerlosen-bund.de/faq/medien,%20tv,%20kommunikation#Technische_Hilfsmittel`.

5 Glennie Evelyn. Evelyn-glennie-deaf-and-music-questions. Retrieved February 2, 2021 from https://www.evelyn.co.uk/about/student-section/.

6 Glennie Evelyn. Hearing essay, 1 2015. URL: https://www.evelyn.co.uk/hearing-essay/.

7 E. M. Finney, I. Fine, and K. R. Dobkins. Visual stimuli activate auditory cortex in the deaf. *Nature Neuroscience*, 4:1171–1173, 2001. doi:10.1038/nn763.

8 FGB-FSS Schweizerische Gehörlosenbund. Alertswiss, 10 2018. Retrieved February 2, 2021 from https://www.sgb-fss.ch/endlich-ein-alarm-fuer-gehoerlose/.

9 Neurolab GmbH. Dopamin, 9 2017. Retrieved February 2, 2021 from https://neurolab.eu/infos-wissen/wissen/neurotransmitter/dopamin/.

10 Arla Good, Maureen J. Reed, and Frank A. Russo. Compensatory plasticity in the deaf brain: Effects on perception of music. *Brain Sciences*, (4):560–574, 2014. URL: https://www.researchgate.net/publication/267449378, doi:10.3390/brainsci404056.

11 Eric Gunther, Glorianna Davenport, and Sile O'Modhrain. Cutaneous grooves: Composing for the sense of touch. *NIME-02*, 2002. URL: https://mf.media.mit.edu/pubs/conference/CutaneousGrooves.pdf.

12 Maria Karam, Gabe Nespoli, Frank Russo, and Deborah I. Fels. Modelling perceptual elements of music in a vibrotactile display for deaf users: A field study. pages 249–254, 2009. doi:10.1109/ACHI.2009.64.

13 David W. Maidment, Alex B. Barker, Jun Xia, and Melanie A. Ferguson. Effectiveness of alternative listening devices to conventional hearing aids for adults with hearing loss: a systematic review protocol. *BMJ Open*, 6:11683, 2016. URL: http://dx.doi.org/10.1136/bmjopen-2016-011683, doi:10.1136/bmjopen-2016.

14 Meadow Matthew. Martin garrix & 7up team up for "concert for the deaf", 3 2016. Retrieved February 2, 2021 from https://www.youredm.com/2016/03/18/martin-garrix-7up-team-concert-deaf/.

15 MED-EL. Air conduction vs. bone conduction: Candidacy guide for bone conduction system, March 2019. Retrieved February 2, 2021 from https://blog.medel.pro/bone-conduction-candidacy-audiogram/.

16 Dr. med. Nonnenmacher. Mechanorezeptoren, 7 2019. Retrieved February 2, 2021 from https://medlexi.de/Mechanorezeptoren.

17 Dr. med. Nonnenmacher. Vater-pacini-körperchen, 7 2019. Retrieved February 2, 2021 from https://medlexi.de/Vater-Pacini-K%C3%B6rperchen.

18 J B Mitroo, Nancy Herman, and Norman I Badler. Movies from music: Visualizing musical compositions.

19 Suranga Nanayakkara, Elizabeth Taylor, Lonce Wyse, and S. H. Ong. An enhanced musical experience for the deaf: Design and evaluation of a music display and a haptic chair. *CHI*, 2009.

20 Antonio Perciaccante, Alessia Coralli, and Neil G. Bauman. Beethoven: His hearing loss and his hearing aids. *Otology & neurotology : official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 41:1305–1308, 10 2020. doi:10.1097/MAO.0000000000002755.

21 Ferreira Bento Ricardo. Beethoven's deafness, the defiance of a genius. *Intl. Arch. Otorhinolaryngol.*, 2009.

22 RNID. How the ear works, April 2020. Retrieved February 2, 2021 from https://rnid.org.uk/information-and-support/ear-health/how-the-ear-works/.

23 RNID. Types and causes of hearing loss and deafness, April 2020. Retrieved February 2, 2021 from https://rnid.org.uk/information-and-support/hearing-loss/types-of-hearing-loss-and-deafness/.

24 Palmer Russ. "tac-tile sound system" (ttss). URL: http://www.kolumbus.fi/riitta.lahtinen/tactile.html.

25 Valorie Salimpoor, Mitchel Benovoy, Kevin Larcher, Alain Dagher, and Robert Zatorre. Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nature neuroscience*, 14:257–62, 02 2011. doi:10.1038/nn.2726.

26 Vincent J. Schmithorst and Scott K. Holland. The effect of musical training on the neural correlates of math processing: A functional magnetic resonance imaging study in humans. *Neuroscience Letters*, 354:193–196, 1 2004. doi:10.1016/j.neulet.2003.10.037.

27 Neary Walter. Brains of deaf people rewire to 'hear' music, November 2001. Retrieved February 2, 2021 from https://www.washington.edu/news/2001/11/27/brains-of-deaf-people-rewire-to-hear-music/.

28 2021 WHO. Who deafness and hearing loss, 3 2021. Retrieved February 2, 2021 from https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss.

29 2021 WHO. Whodeafness, 3 2021. Retrieved February 2, 2021 from https://www.who.int/news-room/facts-in-pictures/detail/deafness.

30 Styleman Wim. Was ist knochenleitung (bone conduction)?, December 2017. Retrieved February 2, 2021 from https://www.bone-conduction.com/de/was-ist-knochenleitung/#:~:text=Wie%20funktioniert%20Knochenleitung%20(bone%20conduction)%3F&text=Wenn%20Schall%20in%20das%20Ohr,einer%20bestimmten%20Weise%20zu%20vibrieren.

31 Disabled World. Sign language and assistive hearing devices. *Disabled World*, 9 2020. URL: https://www.disabled-world.com/disability/types/hearing/communication/.

How Safe are Your IoT Devices? Strategies and Ethical Perspectives on Security and Privacy Awareness through Interactive Tools

Laurence Lerch

Ludwig-Maximilians-Universität München, München, Deutschland
l.lerch@campus.lmu.de

Abstract

In recent years, IoT devices have flooded the markets and users' homes. With increased complexity and connectivity, keeping track of the privacy and security of numerous devices has become a challenge. Simultaneously, a lack of user interfaces limits the ability to make informed decisions about the usage. This paper explores named problems from an interdisciplinary perspective that covers computer science and humanities, privacy and security, as well as usability and ethics. A comprehensive literature review reveals that historical and legal foundations regarding privacy and consent pave the way for the overall concept of *awareness* in the context of IoT devices. In this sense, awareness means that users are conscious of why and to which extent data collectors use, store, and secure their data. As practical opportunities to increase such awareness, recent research has proposed numerous approaches – namely labels and certification schemes but also interactive tools. Whereas the former may be easy to use, the examination shows that static methods face serious obstacles in a rapidly changing IoT world. Interactive tools can respond to such challenges by being dynamic and adaptable. In both approaches, the paper presents user interfaces as crucial and helpful for usability, focusing on icons as an example. Still, designers and developers have to consider the investigated ethical caveats of misuse and manipulation. The theoretically elaborated foundations of this work provide the basis for ethical solutions combining both the simplicity of labels and the advantages of dynamic tools.

2012 ACM Computing Classification Security and privacy → Human and societal aspects of security and privacy → Usability in security and privacy

Keywords and phrases Internet of Things; Usable Security; Ethics; Privacy; Awareness.

Cite as: Laurence Lerch. How Safe are Your IoT Devices? Strategies and Ethical Perspectives on Security and Privacy Awareness through Interactive Tools. In *6th Seminar on Ubiquitous Interaction (UBIACTION 2021)*. Editors: F. Lang, R. Welsch, L. Haliburton, F. Draxler, S. Feger, S. Villa, M. Hoppe, P. Knierim, V. Mäkelä, A. Schmidt. February 11, 2021. Munich, Germany. pp. 11:1–11:22.

1 Introduction

Being away from technology and being forgotten by technology appears to be an unreachable and abstract aim – today more than ever. Already twenty years ago, Marc Langheinrich saw at the intersection of privacy protection and ubiquitous computing "the frightening vision of an Orwellian nightmare-come-true, where countless 'smart' devices with detailed sensing and far-reaching communication capabilities will observe every single moment of our lives, so unobtrusive and invisible that we won't even notice!" [26, p. 289] With the rapid development of the Internet-of-Things (IoT), this vision seems to have come true. But: How did we get to this point? And how could and should we react?

Indeed, many people are concerned about their security and privacy in the digital revolution. Nevertheless, most of them do not act like it and do not try to reduce the influence of companies, states, and others on their data; an observation called the *Privacy Paradox* [36, 57]. Nowhere is the differentiating situation between increased data collection and decreased opportunities to inform users more evident than in the IoT. Omnipresent and ubiquitous systems, be they in healthcare, autonomous cars, or smart homes [41], are interconnected and collect the data that users generate and leave behind. At the same time, these devices link to the surrounding reality in such a way that it may be challenging to distinguish virtuality.

This work proposes *awareness* as a key challenge to address these problems with a necessarily holistic and interdisciplinary theme [8]. With this, the first section deals with ethical considerations and the legal framework. Furthermore, it clarifies with current research terms about *trust*, *transparency*, and *consent*. Based on this, a crucial concept of *awareness* is to be developed, which serves as a basis for ethically justifiable handling of privacy and security. The second part of this analysis specifies the research object to the situation of IoT devices and focuses on state-of-the-art technologies to inform the user. The downsides of these approaches are examined in comparison as well. Afterward, a short digression into the field of user interfaces questions the possibilities and limitations of icons as a prominent example for understandable and usable communication. The relationship between subjective decisions and objective dangers, as well as the possibility of suitable interfaces as an answer to the Privacy Paradox, have a decisive role to play in every part of the elaboration. Hence, this work contributes to research:

- A theoretically elaborated conceptualization of awareness in the context of privacy and security as well as its relationship to trust and transparency with a view to previous research and literature.
- An overview of current and practical approaches and tools for raising awareness of privacy and security in IoT.
- An assessment and investigation of the ethical hazards posed by user interfaces and icons.
- A distinction between privacy and security in terms of subjective perception and objective risk.
- Conceptual proposals and holistic solutions that can provide a perspective for future research.

Before elaborating mentioned aspects, a first clarification is necessary: The whole topic is not just a matter of *security*. It is also important to explicitly include the idea of *privacy* with a humanistic and liberal background, which is prominent and well-argued in western culture and countries. In this analysis, lack of privacy is understood as a significant threat against the idea that every human being should be able to decide whether and when and to what extent he or she wants to give data away. Therefore, privacy and security belong to each other as they are interdependent. Privacy is not entirely feasible when security is disregarded. Consequently, the importance of security is here understood as deduced from personal decisions about privacy in a judicial process.

2 Foundations of Privacy, Law, and Awareness

As already covered with the relationship of security and privacy, such an elaboration is not only a practical question of computer science and human-computer-interaction but also a genuinely critical penetration of applied ethics by looking at the reasons and principles for appropriate actions. Thus, it is about the importance of a humanistic-scientific and normative argumentation, which asks: Why is there a need to talk about this specific field at all? How should the developers, distributors, and users of such IoT devices act to allow appropriate privacy and security? Finally: How can practical solutions and paths for this claim be implemented? This aspect goes hand-in-hand with discussed problems of moral philosophy by Elettra Bietti: She sees that "some work in moral philosophy, particularly in its connections with technology, is criticized for not going far enough prescriptively" [7, p. 213]. It is, therefore, the aim of this section to illuminate and evaluate theoretical considerations on consent and awareness but only with the underlying outlook of the following practical and technological consequences. To this end, it is worthwhile first to examine the foundations of privacy in history.

2.1 Privacy in History and Ethics

One particularly famous and often cited aspect of privacy was set by Samuel D. Warren and Louis D. Brandeis in their paper *The Right to Privacy* in 1890 as they argue for "the more general right of the individual to be let alone" [55, p. 205]. Although privacy in history was not always as prominent as it might appear today [25], the concept of privacy itself is hence by no means something that only the elapsed twenty years were capable of creating. On the contrary, the roots of our perception of privacy, connected with social, cultural, and legal concepts, trace back to the 19th century [43] and even beyond, when privacy is fundamentally associated with an anthropological understanding of humanity as individuals with individual rights.

Already in antiquity, distinctions were drawn between household and state, for example, by Aristotle [50]. As Capurro [9] argues, privacy is also linked to its counterpart, namely *publicness*. Nonetheless, in a disruptive and rapidly changing world, such divisions are becoming more difficult. Social media, semi-public discourses, and not least the interconnection of virtuality and reality make it increasingly difficult to draw a sharp line. The consequence, of course, is that absolute privacy and concrete, practical definitions are hardly achievable. Despite this problem and trying to summarize privacy's meaning to two terms, many researchers argue for either *access* or *control* or both [47]. The further sections of this paper show that especially *control* is implicitly stressed for informing users – not only as a means that users can decide about privacy settings, but to give them the opportunity to be aware of the risks and concerns on which they can indirectly control their data flow by using or not-using IoT devices.

From a socio-political perspective, as Rutherford and Rutherford [46] recognize, most philosophers have pleaded for privacy as essential in and for societies. In this context, an argumentation from both a deontological and a consequentialist perspective appears possible. Whereas the former seems reflected in current jurisprudence and legal frameworks as they often posit human dignity as the foremost norm, assesses the latter the ethical action against the background of harm. One prominent example of consequentialism is John Stuart Mill [33]. As Taylor notices, for Mill, there "are areas of life where the intrusion of law or public censure can only reduce human happiness—areas where the individual's autonomy and individuality should not just be recognised but encouraged" [50, p. 64]. Keulen and Kroeze further state that Mill "stressed that in a liberal democracy, the freedom of private individuals should not be limited by a bureaucratic state or other unnecessary forms of state control; interference in one's private life should be only allowed when an individual harms someone else" [25, p. 31], which is naturally coherent with his famous harm principle. Thanks to the technological revolution and ubiquitous computers, it is questionable whether the liberal state does not have to actively intervene in the spheres of individuals when they are at the mercy of mass surveillance – something not far from what Keulen and Meier

call the "paradox of the liberal state" [25, p. 31].

In the more recent research, privacy and security are linked to *trust*, which is a nebulous umbrella term for many different concepts, as McKnight and Chervany [32] show. However, from an ethical perspective, attention must be drawn here to a possible misunderstanding by the everyday usage of this term: Trust is not and should by no means be the ultimate goal of all privacy considerations since it represents a consequence-focused reduction of the original goal, namely to increase privacy and security. In many cases, trust is a consequence of proper security and privacy and not vice versa. If trust becomes the ultimate goal in a highly complex, opaque world with IoT devices, severe conclusions arise to merely enable users to *feel* and *believe* familiarity and trust. Conversely, as Ishmaev points out, "even outright distrust can be morally justified in specific contexts, and this observation invites us to consider the real value of trust in the IoT" [23, p. 205].

Another striking distinction related to trust draw McKnight et al. with *trust in people* and *trust in technology* [31]. They note that "trust in technology will often reflect positive/negative emotions people develop toward a technology" [31, p. 12:5]. Hence, also trust in technology is exposed in its meaning to the limitation of becoming an emotionalized and subjectivistic norm with its ultimate goal *trust by users* and is easily susceptible to manipulation. Overall, Ishmaev aptly summarizes:

> "However, it is crucial for policies and technological solutions to focus on the trustworthiness of systems and not just the psychological trust attitudes of users. If the latter becomes a goal framed in the model of 'trust deficit', instead of the promised benefits, consumer IoT systems can bring about the dystopian vision of a ubiquitous surveillance economy." [23, p. 229]

If trust is to be included in the discussion, it must be considered against the background of transparency: Transparency creates trust, if and only if transparency shows the *actual* reality of privacy and security. However, also transparency is not without its disadvantages in practical implementations. As Acquisti et al. [3] note, the limitations in transparency and control are severe as it alone cannot necessarily positively influence users' decision-making process. As one interesting and, in their view, promising solution to this issue, they propose interventions aimed at anticipating and counteracting known barriers and constraints faced by individuals when making privacy decisions, which could be feasible with nudges. Thus, the revelation reality of privacy by transparency has to be done in a usable and understandable way.

In a nutshell, one observation summarizes the dilemma of privacy in particular: the so-called *Privacy Paradox*. It describes the ambivalence between people's opinion of privacy as critical and their behavior of acting little or not at all following their opinion [57, 36]. Reasons are user interfaces, education levels, and social norms that lead to the Privacy Paradox and associated fatigue [60, 56]. Furthermore, it appears that the Privacy Paradox also plays a significant role in consent. In their study, Bashir et al. [6] show insufficient comprehension and voluntariness in the consent process due to an asymmetry between users and service providers. However, before more in-depth insights into consent and related awareness, it is worth taking a brief look at the current legal situation, especially in Europe, since, as van der Sloot [52] recognizes, there is an intense exchange about how privacy should be thought of: Is it a personal right? A right that guarantees control over data? Or even a property right?

2.2 Legal Context and the GDPR

As it is with the foundations, legal statements are not just emerged since the rise in the popularity of IoT devices. Many different regulatory frameworks were pronounced since the beginning of the information age in the mid-20th century. These are, for example in the United States, the *Freedom of Information Act* of 1968, the *Privacy Act* of 1974, or the *Computer Security Act* of 1987 [46]. From the European human rights perspective, incorporation into the *European Convention on Human Rights* and the *EU Charter of Fundamental Rights* should be mentioned [52].

A notable milestone in the EU was reached with the General Data Protection Regulation (GDPR), applicable from 2018 [2]. Despite all the positive directions [58], the GDPR is by no means uncontroversial, especially in the area of technology and social media as well as from the perspective of businesses [49, 17, 29]. Mangini et al. [30] note that mainly the "right to be forgotten" (GDPR Art. 17) is expensive and complicated to implement. They cite as reasons lack of knowledge, short deadlines, too much effort, ambiguity, and non-clarity of the legislation. A recent study by Zaeem and Barber [58] shows that the GDPR has made progress in protecting user data, but when it comes to editing and deleting, more effort appears necessary.

It is not possible within the scope of this paper to fully encompass the meaning and content of the GDPR. Nevertheless, it is worth pointing out a concept essential for dealing with data and data protection: *consent*. From a legal perspective, consent is not always necessary for the GDPR since it is only one possibility out of a total of six to base data processing on the framework. Simultaneously, the requirements – namely informed, specific, unambiguous, and freely given (GDPR Art. 4) [2] – are probably most difficult to fulfill in practice [52]. However, the concepts of consent serve as a foundation for increased user awareness, as the following sections show.

Before getting into more detail about awareness, one aspect that was not discussed so far: According to Porambage et al. [41], the legal framework must be thought of globally. In a decentralized world of the Internet and data-flows, privacy is indeed not only a question of objective claims but also of metaethically normative possibility, culture-relativistic scopes, and intergovernmental practicability. Seemingly objective approaches face intense problems as, especially in privacy and security, different cultures may find different answers. Historically too, sociocultural dependencies and changing norms require, according to Keulen and Kroeze [25], the treatment of privacy as a context-dependent phenomenon. From a descriptive perspective, regulations that are thought of globally appear to be culturally relatable despite a normative claim. Although this issue requires more in-depth research, the subsequent context focuses on Western and European circumstances due to a limited scope.

2.3 Consent and Awareness for Privacy

Consent may be famous for its complex practical implementation and fulfillment, especially regarding particularly prominent check-boxes, cookies, and consent management platforms (CMP) [37]. Thus, it can be said already at this point that this work can and will by no means fulfill the *legal* framework. Instead, concerning consent and its problems, the aim is to show from an *ethical* perspective opportunities and ways how users can practicably gain *awareness* over their data themselves with modern technologies presented in the research. Three possible problems in direct or indirect terms of consent are to theoretically identify first:

- Users de facto agree to consent, although they would not do so if they would know the consequences. This matter leads to an asymmetry between users and providers or companies as, for example, Bashir et al. find in the area of cloud services [6]. It might also be possible since many users do not read privacy policies at all [38] or since they are nudged to decisions [37].
- Due to a lack of sanctions in practice, consent is not queried or is understood to be implicit [37, 51]. Hence, the user may not express his or her decisions.
- Unconsciously or consciously, IoT devices are insecure and pose a security risk *despite* an informed consent which results in reduced privacy. Thus, this represents a crucial issue between the related concepts of privacy and security.

Derived from an understanding of liberties and people as individuals, it is not an ethical problem if users are truly fully aware of the consequences and decide to accept all these consequences voluntarily, even though the consequences are disproportional to the benefit of giving the consent. However, this full knowledge seems like an abstract idea, which is hardly achievable in an opaque and heterogeneous world. The proposed way of getting closer to such an ideal is the idea of achieving *awareness* through *informing* and *knowledge*. As the following thoughts show, every single consideration directs to these terms.

Another way to approach the subject is through the *Software Engineering Code of Ethics and Professional Practice* [1], for which Gotterbarn et al. [16] laid the foundation in 1997. Interestingly, differences in the versions arise not only in detail but also in fundamental shifts. For Gotterbarn et al., the product was still the first principle, whereas now the public has moved to the top place. Version 5.2 argues that in "all these judgments concern for the health, safety and welfare of the public is primary; that is, the 'public interest' is central to this code" [1]. Admittedly, with the normative principle of the *public interest*, some questions and problems arise since it may not be evident in all circumstances what is meant by it concretely in practice. Moreover, it is to ask to what extent the relationship of a rather consequentialist principle can harmonize with duty-ethical, legislative, and cultural-relativist ones.

To make these abstract principles more concrete, Harris [20] proposes guidelines for Web design, namely that (1) information should not be hidden, (2) information should not be used or transferred without proper consent, (3) consent should be properly obtained, and that (4) privacy should be maintained. All of them are also sensible for IoT devices since inquiry of consent is understood as a prerequisite and maintaining privacy as a necessity. Overall, not *consent* alone is necessary but *informed consent*, as also the GDPR requires. The question, however, remains how to concretely avoid information hiding and implement *properly* obtaining of consent.

In summary, the problems surrounding informed consent are more significant than they might appear at first glance. Although there are undoubtedly desirable approaches in research and practice, they are either normatively too weak, laborious to regulate, without incentives for companies, or hardly practicable. Nevertheless, the challenging situation becomes promising when the previous discussion of consent in ethics is expanded to include the notion of broader *awareness*. *Awareness* means that the user is conscious of the extent to which and how data are protected and secured. The immediately following questions are: How can such awareness for privacy and security be achieved? Which aspects play a decisive role in implementing in practice? To sum up: How can IoT and awareness be connected when limited user interfaces and the highest complexity and connectivity makes it nearly impossible for citizens to recognize the consequences of such devices?

The idea of relating awareness to the Privacy Paradox itself can already be found in an examination of Pötzsch [42]. From her analysis, two insights are of particular importance: (1) First, she describes that "Awareness is based

on an individual's attention, perception and cognition of physical as well as non-physical objects. The state of being aware of something fades away as soon as there is no longer any stimulus present" [42, p. 228]. Awareness is thus only possible when a continuous stimulus is present. (2) Furthermore, she recognizes that privacy-awareness information is sometimes of a general nature and user-independent, e.g. privacy disclaimers on websites, and sometimes personally connected to an individual and user-specific, e.g. feedback from policy evaluations. Especially on the Internet, as a place of different ideas, socialization, and cultures, both aspects enrich the discussion. A recent study, which investigated the importance of impulsivity as a specific personal characteristic for the Privacy Paradox, supports the relevance of user-specific centering [4].

Indeed, even such a focus on awareness has to contend with criticism of practicability, adjustability, and responsible parties. That some companies might have little interest in such an implementation and self-disclosure about their services and devices is certainly not to be dismissed. Simultaneously, there is a great sense for open-source information, above all in technology and computer science. The ability and knowledge of experts can be used to participate quite concretely in creating such information systems. Given the relevance and concern for data protection, it may be reasonable that society will welcome such an opportunity with warm hands, but with a view on the Privacy Paradox, only if the information is prepared appropriately and with knowledge of the need for *awareness*, as the following chapters examine.

3 Interactive Tools as a Means of Raising Awareness

Now that the previous sections have mentioned practicability as a decisive aspect of privacy, the question arises of how these considerations can be implemented and used. In this context, the following section takes up former research [39, 12]. The aim is to grasp this research in its relevance. This also covers the ethical risks and opportunities of icons. Furthermore, possibilities and ways are proposed to improve these technologies and tools in practice. Hence, this part offers concrete suggestions for future implementations.

3.1 Background and Related Work in Security and Privacy

The noteworthy sub-areas in privacy and security of IoT devices range from low-level tactics to high-level principles, from the implementation of hardware to middleware, via software engineering to the already introduced issues of consent and user empowerment [15, 40, 41]. According to Wanbil et al. [54], privacy can be achieved through technical and social solutions. Technical approaches describe the attempt to prevent unauthorized access and loss of data. On the other hand, social approaches mean, for example, creating acceptance and awareness for the use of data. In such a classification, this work has to combine both [39]. Technical considerations as criteria are crucial in determining specific objective recommendations for a user; social approaches include behavior and usability.

Specifically, increasing the security of IoT devices has been a stated goal for some time resulting in comprehensive research [13, 19, 28, 44, 48]. Even though there are unquestionable approaches to implement security in the field of IoT [5], at least in computer science, the aspect of ethical and legal principles of privacy appears to have received less attention in previous research. As security can be understood as a component to enable privacy and achieve it in the first place, it is by no means to underestimate its importance. On the other hand, it is also worth examining the prerequisites and possibilities of privacy for IoT. For example, machine learning methods appear to be a promising way to minimize risks and increase privacy [59].

A more well-known approach is *Privacy by Design*, which has also found appeal in the area of IoT devices [40]. However, as Porambage et al. [41] recognize, some open questions and limitations arise there as well. First, they call for a definition of a general model for IoT privacy; from a perspective of awareness, the principles developed by Langheinrich [26, 27] appear to be useful and adaptable. Second, the development of privacy-enhancing technologies with the possibility of scalability and heterogeneity of IoT. And third, the implementation and integration with a perfect balance of privacy policies, localization and tracking requirements, and sensitive data access control mechanisms. However, the extent to which one can speak of *perfect balance* remains worthy of discussion.

These are welcome efforts to enable privacy to a greater extent and from the very beginning. However, as Williams et al. [57] recognize in their study, there is still a significant lack of awareness in the Internet-of-Things, which leads to an even greater prevalence of the Privacy Paradox described above. To take a closer look at precisely this *awareness*, the following section describes ways of informing users and empowering them to make decisions proactively based on the knowledge they have gained.

3.2 Labels and Interactive Tools in Recent Research

One approach to creating awareness among users can be introducing a label system when purchasing IoT devices, as has already been the case for some time on numerous food packages and in the area of nutrition [18, 21, 34] and even privacy in general [24]. In principle, research from this area makes a valuable contribution to the purchase and use of IoT devices, especially for users' perceptual and decision-making abilities. Emami et al. [12] determine

with their study that the participants perceived the label they designed for IoT devices as accessible and useful. An interesting aspect of their label is that it addresses both privacy and security risks separately. They also note that many participants did not pay attention to privacy and security before purchasing IoT devices but became more concerned after some time.

Nevertheless, the rapid technological development with numerous patches, updates, and upgrades limits rigid and static labeling in practice. Thus, it may be both that, unlike the original label, an update has resolved significant security issues. Moreover, there is an inherent risk that initially positively labeled products will experience a reduction in security and privacy over time. It is not advisable to use labels that may also be less present when purchasing online. As another related term, also *certification schemes*, as Ishmaev [23] discusses in the context of trust, require specification, adaptation, and flexibility. Indeed, certifications could be recalled or reissued, and they might be trustworthy. As a severe downside, these proposals require a specific organization to advocate for these certifications for entire trustworthiness and standardization. Furthermore, binding certifications may evoke liability issues and legal concerns after data accidents and breaches. Without clarification of such matters, certification systems do not appear to be feasible at this time.

A more dynamic approach describe Oser et al. [39] with their tool SAFER. Their study investigates how an IoT device risk assessment framework can help increase users' awareness in a corporate environment where the context of their study is companies with a BYOD policy ("Bring Your Own Device"), which may create serious security risks by IoT devices. In their study, they refer to the aforementioned research by Emami-Naeini et al. [12], which shows how security and privacy could influence IoT purchasing behavior. Oser et al. conducted their study with 20 participants in mixed-method evaluations. The tool SAFER itself is based on four different components: Device Identification, Vulnerability Enrichment, Scoring, and Frontend. As user interface (UI), they offer a *guided* and a *rich* version. The former is based on a traffic light to show the overall risk; the latter allows more detailed information. Overall, SAFER's UI proves to be comprehensive and useful. Especially color-based device risk assessment conveyed the most important information and was perceived helpful.

An exciting and vital finding of their study is the significant tension between subjective perception and objective risk of devices. For example, users have greater concerns about a CCTV camera labeled as low-risk than about an e-book reader, also labeled as low-risk. As they rightfully conclude, a focus on the difference between risk assessment of the firmware of a device and the importance of privacy is necessary. They also state that a divergence between the assessment of the tool and the perceived complexity of a device reduces trust in the tool. It seems evident that this also affects the user's awareness and can become an extension of the Privacy Paradox as, despite stated opinion, the action is not taken according to objective standards. Furthermore, this issue links to the previously shown distinction between user-independent and user-specific awareness [42]. Whereas independent approaches allow a higher degree of objectivity, specific approaches offer room for individualization. A means to enable awareness in such situations is implementing suitable user interfaces.

3.3 User Interfaces and Icons for Awareness

At first glance, user interfaces of interactive tools may be seen as a promising means for privacy and security awareness in interactive tools – just think of the numerous possibilities such as color design, traffic light systems, and conspicuous warnings. Despite these advantages and general research in positive nudging and privacy [10], it is essential to point out risk from an ethical perspective: User interfaces can help create awareness among users, but at the same time, these methods can also be used to push users in unwanted and unethical directions. Namely, when preset manipulation becomes the maxim instead of individual decision support. Such a focus also reveals dangers that can fall under "dark patterns" [35]. At the same time, such a perspective opens up the critical question: Who is the decision-maker who may set the limits of *acceptable* risks? Is this always the user by conducting user-centered and individual assessments? Or not rather the developer and designer of the tool? In summary, UI is similar to the *value* of trust described earlier: Trust as an ultimate goal is normatively as weak as awareness through user interfaces alone. The question remains – a question that scientists and developers should always ask – what assumptions and boundaries should be made.

Nevertheless, when this happens, user interfaces open up numerous opportunities. As a guideline, the mentioned aspects of consent can now be abstracted to awareness. In this sense, the principle of awareness should:

- Familiarize users with the consequences of the purchase, use, and consent.
- In doing so, provide guidance for creating awareness when operators and manufacturers of IoT devices unlawfully or lawfully do not ask for consent.
- Include risks and lessons learned from security breaches in decision-making that may exist independent of consent and privacy awareness.

Due to the limited scope, the following paragraphs discuss icons as part of user interfaces, their possibilities, and their problems as a means to achieve named goals. A central advantage of icons is that the GDPR (Art. 12) explicitly speaks of providing information "in combination with standardized icons in

order to give in an easily visible, intelligible and clearly legible manner a meaningful overview of the intended processing" [2]. It is noteworthy already at this point that this has to be done *in combination*, whereby Rossi and Palmiran [45] recognize that the GDPR has not specified this circumstance further. The legal situation of icons through the GDPR is further and comprehensively discussed by Efroni et al. [11].

Regardless of the lack of legal standardization, icons can be helpful in enabling usable information about the consequences of use. In their research, Rossi and Palmiran [45] explore which iconographies could be used as motifs to implement privacy concepts understandably. The tool by Oser et al. [39] discussed earlier also uses icons. Rossi and Palmiran, however, recognize a crucial and understandable point: In the development of such icons, a "constant opposition between simplicity and preciseness of representation arises. Whereas the former is a fundamental feature to ensure usability and scalability of the visual elements to any dimension, the latter is important to convey the exact meaning of the corresponding concept" [45, p. 83]. Since an icon may, due to simplicity and plasticity, communicate less or even wrongly interpreted information, this addresses limitations that continue to give the aforementioned *combination* a raison d'être. Besides, the question arises about the extent to which individual and cultural differences reveal differences in interpretation [22].

A further critical thought stands in connection to the already mentioned decision-makers. It is crucial and of decisive relevance who designs them and with what purposes [53]. Manufacturers of IoT devices will certainly pursue different interests in design and use than government actors or NGOs may do, as the latter want to make users aware of privacy and security risks. Although perhaps perceived helpful, binary recommendations for use or non-use (like green and red icons) may also increase this risk from an ethical perspective since users could be pushed to a decision based on predefined criteria by external parties. What makes it clear once again: Trust can be pursued in different ways by different actors with different goals. Icons that convey a sense of trust while hiding privacy and security risks are an inherent danger of such an implementation and the lack of standardization through the GDPR or government regulations regarding icons.

Rossi and Palmiran [45] even go so far as to demand that data subjects have to be trained to understand the concept behind the icons. Whether such responsibility from a political-practical perspective is to be located with state actors remains controversial. Nevertheless, such an approach can also find its way into interactive tools: By using icons in principle, it is possible to gain more in-depth background experience through short tutorials, playful exercises, and gamification strategies. One step in this direction is already implemented with SAFER by providing explanatory tooltips [39]. Unquestionably, this approach requires further empirical research in order to be able to investigate the theoretical possibility of even deeper guidance in such an interactive tool. In summary, however, it can be said that user interfaces and especially icons can be embedded well in such tools after ethical risks and challenges are considered in the implementation and design. As long as the legislation does not require standardized icons, such a design process should still be accompanied by independent parties to assess the named ethical hazards.

3.4 Discussion and Further Research

A significant issue raised earlier is the tension between static labels and dynamic requirements in a rapidly changing world of IoT devices. Emami-Naeini et al. [12] point out in their study that there is certainly the possibility of an *interactive* online label or a QR code on the printed static label. It would be worth further investigation whether information on the static label about privacy and security would still be necessary at all during a purchase process or whether a prominent notice with a QR code pointing to a constantly updating website or tool would suffice. With adaptations and extensions for private use, this could build on SAFER by Oser et al. [39]. To investigate this option empirically, a study is needed to ask whether many users are not discouraged from this increased effort of a QR code. Should it turn out that such a possibility is widely used, new ways and opportunities will open up in practice to significantly increase awareness among users.

Such a holistic approach would also address a second problem already discussed: The different levels of knowledge and subjective experiences of users, making it challenging to objectively address all subjective privacy concerns to fit each individual. As previous research points out, it is necessary to discuss users' perception of risk, especially when it comes to higher complexity and sensitive data [12, 39]. Whereas security by its nature appears to be more objective, privacy with discussed consent and control is rather subjective. Using an interactive tool that starts with a QR code at the point of purchase may allow a wide and personalized range of accuracy and individuality. Users could decide whether they want to receive more in-depth information about the device in addition to a simple, initial overview, as in SAFER's *guided* and *rich* view. At the same time, the individualization of such a tool could pave the way for customized personalization to each user. Setting these preferences, for example, with more innovative guides such as educational quizzes or interactive images could also be helpful for people who have little to no knowledge and experience so far. From an ethical and humanistic point of view, further research is needed with a specific focus on individual perception and cultural relativistic criticism.

Finally, as indicated earlier [42], stimuli can be a practical way to create

ongoing awareness related to the user experience. This idea is connectable with information and content since a rapidly changing technology may receive updates and face security incidents with the following changed risk recommendations. As this idea may be beneficial during the ownership and the use, it would go beyond the sole purchase process and open up chances for permanent assisting. This consequence is not a disadvantage since users are more concerned in terms of their IoT devices over time [12]. In practice, many ways are imaginable for implementation, such as real-time notifications in a mobile environment like smartphones or tablets.

To sum up, the long-term hope of upcoming research would be an empirically examined tool that, similar to diet or nutrition apps, is not seen as a burden but as a companion, that can be individualized for every single person regardless of the level of knowledge and awareness, that is used daily through habit or stimuli, that encompasses the private environment with the business environment, and that combines privacy with security.

4 Conclusion

As manifestos about the current situation of IoT "move out from descriptive to predictive mode" [14, p. 10] and reveal how important it is for designers, developers, and individuals to think about societal questions of a digital age, it is also needed to question the current topics of privacy and security through an ethical lens. In this work, after presenting a perspective of history and the meaning of privacy, concepts and approaches of the recent research were examined. Especially *awareness* and *consent* were shown as principles on which an increase of privacy and security is feasible. Further, it became apparent that theoretical considerations must include the legal situation – on the one hand, to do justice to the legal framework, and on the other hand, in order to emphasize possibilities of design.

The second part of the work combined the discussed principles and the technological implementation in the Internet-of-Things (IoT). Against the background of recent privacy and security research, current approaches for informing users were examined. The analysis was able to show that existing technologies can already make a notable contribution to achieving the stated goal of awareness. Above all, the user interface is able to support such an aim, although a sole focus on trust has been cited as an attendant hazard. Opportunities for upcoming research arise in this conceptual interface area as the previous approaches may also be applied for holistic solutions in daily life. This work provides the theoretical foundation with the hope that eventually one day, this will lead to greater privacy and security awareness of the individual and society.

References

1 Software Engineering Code of Ethics and Professional Practice (Version 5.2) as recommended by the ACM/IEEE-CS Joint Task Force on Software and jointly approved by the ACM and the IEEE-CS as the standard for teaching and practicing software engineering. Retrieved March 1, 2020 from `https://ethics.acm.org/code-of-ethics/software-engineering-code/`.

2 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Official Journal of the European Union (OJ)*, 59(L 119):1–88, 2016.

3 Alessandro Acquisti, Idris Adjerid, and Laura Brandimarte. Gone in 15 Seconds: The Limits of Privacy Transparency and Control. *Security & Privacy, IEEE*, 11(4):72–74, 2013. `doi:10.1109/MSP.2013.86`.

4 Zahra Aivazpour and V. Srinivasan (Chino) Rao. Information Disclosure and Privacy Paradox: The Role of Impulsivity. *SIGMIS Database*, 51(1):14–36, 2020. `doi:10.1145/3380799.3380803`.

5 Masoud Barati, Ioan Petri, and Omer F. Rana. Developing GDPR Compliant User Data Policies for Internet of Things. In *Proceedings of the 12th IEEE/ACM International Conference on Utility and Cloud Computing*, UCC'19, pages 133–141, New York, NY, USA, 2019. Association for Computing Machinery. `doi:10.1145/3344341.3368812`.

6 Masooda Bashir, Carol Hayes, April D. Lambert, and Jay P. Kesan. Online Privacy and Informed Consent: The Dilemma of Information Asymmetry. In *Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community*, ASIST '15, pages 1–10, USA, 2015. American Society for Information Science.

7 Elettra Bietti. From Ethics Washing to Ethics Bashing: A View on Tech Ethics from within Moral Philosophy. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, FAT* '20, pages 210–219, New York, NY, USA, 2020. Association for Computing Machinery.

8 Johannes Buchmann, editor. *Internet Privacy: Eine multidisziplinäre Bestandsaufnahme/A multidisciplinary analysis*. Springer, Berlin, Heidelberg, 2012. `doi:10.1007/978-3-642-31943-3`.

9 Rafael Capurro. IT and Privacy from an Ethical Perspective Digital Whoness: Identity, Privacy and Freedom in the Cyberworld. Introduction. In Johannes Buchmann, editor, *Internet Privacy Eine multidisziplinäre Bestandsaufnahme/A multidisciplinary analysis*, chapter 2.1, pages 64–66. Springer, Berlin, Heidelberg, 2012. `doi:10.1007/978-3-642-31943-3_2`.

10 Eun Kyoung Choe, Jaeyeon Jung, Bongshin Lee, and Kristie Fisher. Nudging People Away from Privacy-Invasive Mobile Apps through Visual Framing. In Paula Kotzé, Gary Marsden, Gitte Lindgaard, Janet Wesson, and Marco

Winckler, editors, *Human-Computer Interaction – INTERACT 2013. Lecture Notes in Computer Science*, volume 8119, pages 74–91, Berlin, Heidelberg, 2013. Springer. `doi:10.1007/978-3-642-40477-1_5`.

11 Zohar Efroni, Jakob Metzger, Lena Mischau, and Marie Schirmbeck. Privacy Icons: A Risk-Based Approach to Visualisation of Data Processing. *European Data Protection Law Review*, 5(3):352–366, 2019. `doi:10.21552/edpl/2019/3/9`.

12 Pardis Emami-Naeini, Henry Dixon, Yuvraj Agarwal, and Lorrie Faith Cranor. Exploring How Privacy and Security Factor into IoT Device Purchase Behavior. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pages 1–12, New York, NY, USA, 2019. Association for Computing Machinery. `doi:10.1145/3290605.3300764`.

13 Alaba Fadele, Mazliza Othmana, Ibrahim Hashem, and Faiz Alotaibi. Internet of Things Security: A Survey. *Journal of Network and Computer Applications*, 88:10–28, 2017. `doi:10.1016/j.jnca.2017.04.002`.

14 Ester Fritsch, Irina Shklovski, and Rachel Douglas-Jones. Calling for a Revolution: An Analysis of IoT Manifestos. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 1–13, New York, NY, USA, 2018. Association for Computing Machinery. `doi:10.1145/3173574.3173876`.

15 Sameera Ghayyur, Primal Pappachan, Guoxi Wang, Sharad Mehrotra, and Nalini Venkatasubramanian. Designing Privacy Preserving Data Sharing Middleware for Internet of Things. In *Proceedings of the Third Workshop on Data: Acquisition To Analysis*, DATA '20, pages 1–6, New York, NY, USA, 2020. Association for Computing Machinery. `doi:10.1145/3419016.3431484`.

16 Don Gotterbarn, Keith Miller, and Simon Rogerson. Software Engineering Code of Ethics. *Commun. ACM*, 40(11):110–118, 1997. `doi:10.1145/265684.265699`.

17 Samuel Greengard. Weighing the Impact of GDPR. *Communications of the ACM*, 61(11):16–18, 2018. `doi:10.1145/3276744`.

18 Klaus G. Grunert and Josephine M. Wills. A review of European research on consumer response to nutrition information on food labels. *Journal of Public Health*, 15(5):385–399, 2007. `doi:10.1007/s10389-007-0101-9`.

19 Amina Harit, Abdellah Ezzati, and Rachid Elharti. Internet of Things Security: Challenges and Perspectives. In *Proceedings of the Second International Conference on Internet of Things, Data and Cloud Computing*, ICC '17, New York, NY, USA, 2017. Association for Computing Machinery. `doi:10.1145/3018896.3056784`.

20 James K. Harris. Ethical Issues in Web Design. *Journal of Computing Sciences in Colleges*, 25(2):214–220, 2009.

21 Kristy L. Hawley, Christina A. Roberto, Marie A. Bragg, Peggy J. Liu, Marlene B. Schwartz, and Kelly D. Brownell. The science on front-of-

package food labels. *Public Health Nutrition*, 16(3):430–439, 2013. doi:10.1017/S1368980012000754.

22 Leif-Erik Holtz, Katharina Nocun, and Marit Hansen. Towards Displaying Privacy Information with Icons. In Simone Fischer-Hübner, Penny Duquenoy, Marit Hansen, Ronald Leenes, and Ge Zhang, editors, *Privacy and Identity Management for Life. Privacy and Identity 2010. IFIP Advances in Information and Communication Technology*, volume 352, pages 338–348, Berlin, Heidelberg, 2011. Springer. doi:10.1007/978-3-642-20769-3_27.

23 Georgy Ishmaev. Rethinking Trust in the Internet of Things. In Ronald Leenes, Rosamunde van Brakel, Serge Gutwirth, and Paul De Hert, editors, *Data Protection and Privacy: The Internet of Bodies*, Computers, Privacy and Data Protection, pages 203–230. Hart Publishing, Oxford, 2019. doi:10.5040/9781509926237.ch-009.

24 Patrick Gage Kelley, Joanna Bresee, Lorrie Faith Cranor, and Robert W. Reeder. A "Nutrition Label" for Privacy. In *Proceedings of the 5th Symposium on Usable Privacy and Security*, SOUPS '09, New York, NY, USA, 2009. Association for Computing Machinery. doi:10.1145/1572532.1572538.

25 Sjoerd Keulen and Ronald Kroeze. 1. Privacy from a Historical Perspective. In Bart van der Sloot and Aviva de Groot, editors, *The Handbook of Privacy Studies*, pages 21–56. Amsterdam University Press, Amsterdam. doi:10.1515/9789048540136-002.

26 Marc Langheinrich. Privacy by design – Principles of privacy-aware ubiquitous systems. In Gregory D. Abowd, Barry Brumitt, and Steven Shafer, editors, *Ubicomp 2001: Ubiquitous Computing. Lecture Notes in Computer Science*, volume 2201, pages 273–291, Berlin, Heidelberg, 2001. Springer. doi:10.1007/3-540-45427-6_23.

27 Marc Langheinrich. A Privacy Awareness System for Ubiquitous Computing Environments. In Gaetano Borriello and Lars Erik Holmquist, editors, *UbiComp 2002: Ubiquitous Computing. Lecture Notes in Computer Science*, volume 2498, pages 237–245, Berlin, Heidelberg, 2002. Springer. doi:10.1007/3-540-45809-3_19.

28 Huichen Lin and Neil W. Bergmann. IoT Privacy and Security Challenges for Smart Home Environments. *Information*, 7(3):44, 2016. doi:10.3390/info7030044.

29 Thomas Linden, Rishabh Khandelwal, Hamza Harkous, and Kassem Fawaz. The Privacy Policy Landscape After the GDPR. *Proceedings on Privacy Enhancing Technologies*, 1:47–64, 2020. doi:10.2478/popets-2020-0004.

30 Vincenzo Mangini, Irina Tal, and Arghir-Nicolae Moldovan. An Empirical Study on the Impact of GDPR and Right to Be Forgotten - Organisations and Users Perspective. In *Proceedings of the 15th International Conference on Availability, Reliability and Security*, ARES '20, New York, NY, USA, 2020. Association for Computing Machinery. doi:10.1145/3407023.3407080.

31 D. Harrison McKnight, Michelle Carter, Jason Bennett Thatcher, and Paul F. Clay. Trust in a Specific Technology: An Investigation of Its Components and Measures. *ACM Trans. Manage. Inf. Syst.*, 2(2), 2011. doi:10.1145/1985347.1985353.

32 D. Harrison McKnight and Norman L. Chervany. Trust and Distrust Definitions: One Bite at a Time BT. In Rino Falcone, Munindar Singh, and Yao-Hua Tan, editors, *Trust in Cyber-societies*, pages 27–54. Springer, Berlin, Heidelberg, 2001.

33 John Stuart Mill. *On Liberty*. Cambridge University Press, Cambridge, 2011. doi:DOI:10.1017/CBO9781139149785.

34 Lisa M. Soederberg Miller and Diana L. Cassady. The effects of nutrition knowledge on food label use. A review of the literature. *Appetite*, 92:207–216, 2015. doi:10.1016/j.appet.2015.05.029.

35 Arvind Narayanan, Arunesh Mathur, Marshini Chetty, and Mihir Kshirsagar. Dark Patterns: Past, Present, and Future: The Evolution of Tricky User Interfaces. *Queue*, 18(2):67–92, 2020. doi:10.1145/3400899.3400901.

36 Helen Nissenbaum. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press, Stanford, 2009. doi:10.1515/9780804772891.

37 Midas Nouwens, Ilaria Liccardi, Michael Veale, David Karger, and Lalana Kagal. Dark Patterns after the GDPR: Scraping Consent Pop-Ups and Demonstrating Their Influence. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, pages 1–13, New York, NY, USA, 2020. Association for Computing Machinery. doi:10.1145/3313831.3376321.

38 Jonathan A. Obar and Anne Oeldorf-Hirsch. The biggest lie on the Internet: ignoring the privacy policies and terms of service policies of social networking services. *Information, Communication & Society*, 23(1):128–147, 2020. doi:10.1080/1369118X.2018.1486870.

39 Pascal Oser, Sebastian Feger, Paweł W. Woźniak, Jakob Karolus, Dayana Spagnuelo, Akash Gupta, Stefan Lüders, Albrecht Schmidt, and Frank Kargl. SAFER: Development and Evaluation of an IoT Device Risk Assessment Framework in a Multinational Organization. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(3), 2020. doi:10.1145/3414173.

40 Charith Perera, Ciaran McCormick, Arosha K. Bandara, Blaine A. Price, and Bashar Nuseibeh. Privacy-by-Design Framework for Assessing Internet of Things Applications and Platforms. In *Proceedings of the 6th International Conference on the Internet of Things*, IoT'16, pages 83–92, New York, NY, USA, 2016. Association for Computing Machinery. doi:10.1145/2991561.2991566.

41 Pawani Porambage, Corinna Schmitt, Pardeep Kumar, Andrei Gurtov, Mika Ylianttila, and Athanasios Vasilakos. The Quest for Privacy in the Internet

of Things. *IEEE Cloud Computing*, 3(2):36–45, 2016. doi:10.1109/MCC.2016.28.

42 Stefanie Pötzsch. Privacy Awareness: A Means to Solve the Privacy Paradox? In Vashek Matyáš, Simone Fischer-Hübner, Daniel Cvrček, and Petr Švenda, editors, *The Future of Identity in the Information Society*, pages 226–236, Berlin, Heidelberg, 2009. Springer. doi:10.1007/978-3-642-03315-5_17.

43 Megan Richardson. *The Right to Privacy: Origins and Influence of a Nineteenth-Century Idea*. Cambridge University Press, Cambridge, 2017. doi:10.1017/9781108303972.

44 Rodrigo Roman, Jianying Zhou, and Javier Lopez. On the features and challenges of security and privacy in distributed internet of things. *Computer Networks*, 57(10):2266–2279, 2013. doi:10.1016/j.comnet.2012.12.018.

45 Arianna Rossi and Monica Palmiran. What's in an Icon?: Promises and Pitfalls of Data Protection Iconography. In Dara Hallinan, Ronald Leenes, Serge Gutwirth, and Paul De Hert, editors, *Data Protection and Privacy: Data Protection and Democracy*, Computers, Privacy and Data Protection, pages 59–92. Hart Publishing, Oxford, 2020. doi:10.5040/9781509932771.ch-003.

46 Rebecca H. Rutherfoord and James K. Rutherfoord. Privacy and Ethical Concerns in Internet Security. In *Proceedings of the 2010 ACM Conference on Information Technology Education*, SIGITE '10, pages 131–134, New York, NY, USA, 2010. Association for Computing Machinery. doi:10.1145/1867651.1867686.

47 Marijn Sax. Privacy from an Ethical Perspective. In Bart van der Sloot and Aviva de Groot, editors, *The Handbook of Privacy Studies*, chapter 3, pages 143–172. Amsterdam University Press, Amsterdam, 2018. doi:https://doi.org/10.1515/9789048540136-006.

48 Eman Shaikh, Iman Mohiuddin Mohiuddin, and Ayisha Manzoor. Internet of Things (IoT): Security and Privacy Threats. In *2019 2nd International Conference on Computer Applications Information Security (ICCAIS)*, pages 1–6, 2019. doi:10.1109/CAIS.2019.8769539.

49 Sophie Stalla-Bourdillon and Alison Knight. Data Analytics and the GDPR: Friends or Foes? A Call for a Dynamic Approach to Data Protection Law. In Ronald Leenes, Rosamunde van Brakel, Serge Gutwirth, and Paul De Hert, editors, *Data Protection and Privacy: The Internet of Bodies*, Computers, Privacy and Data Protection, pages 249–276. Hart Publishing, Oxford, 2019. doi:10.5040/9781509926237.ch-011.

50 Roger Taylor. No Privacy without Transparency. In Ronald Leenes, Rosamunde van Brakel, Serge Gutwirth, and Paul De Hert, editors, *Data Protection and Privacy: The Age of Intelligent Machines*, Computers, Privacy and Data Protection, pages 63–88. Hart Publishing, Oxford, 2017. doi:10.5040/9781509919376.ch-003.

51 Christine Utz, Martin Degeling, Sascha Fahl, Florian Schaub, and Thorsten Holz. (Un)Informed Consent: Studying GDPR Consent Notices in the Field. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, CCS '19, pages 973–990, New York, NY, USA, 2019. Association for Computing Machinery. doi:10.1145/3319535.3354212.

52 Bart van der Sloot. Privacy from a Legal Perspective. In Bart van der Sloot and Aviva de Groot, editors, *The Handbook of Privacy Studies*, chapter 2, pages 63–136. Amsterdam University Press, Amsterdam, 2018. doi:10.1515/9789048540136-004.

53 Iris van Ooijen and Helena U. Vrabec. Does the GDPR Enhance Consumers' Control over Personal Data? An Analysis from a Behavioural Perspective. *Journal of Consumer Policy*, 42(1):91–107, 2019. doi:10.1007/s10603-018-9399-7.

54 Lee Wanbil, Wolfgang Zankl, and Henry Chang. An Ethical Approach to Data Privacy Protection. *ISACA Journal*, 6:1–9, 2016.

55 Samuel D. Warren and Louis D. Brandeis. The Right to Privacy. *Harvard Law Review*, IV(5):193–220, 1890.

56 Meredydd Williams, Jason R. C. Nurse, and Sadie Creese. The Perfect Storm: The Privacy Paradox and the Internet-of-Things. In *11th International Conference on Availability, Reliability and Security (ARES)*, pages 644–652, 2016. doi:10.1109/ARES.2016.25.

57 Meredydd Williams, Jason R. C. Nurse, and Sadie Creese. Privacy is the Boring Bit: User Perceptions and Behaviour in the Internet-of-Things. In *2017 15th Annual Conference on Privacy, Security and Trust (PST)*, pages 181–190, 2017. doi:10.1109/PST.2017.00029.

58 Razieh Nokhbeh Zaeem and K. Suzanne Barber. The Effect of the GDPR on Privacy Policies: Recent Progress and Future Promise. *ACM Transactions on Management Information Systems*, 12(1), 2020. doi:10.1145/3389685.

59 Mengyao Zheng, Dixing Xu, Linshan Jiang, Chaojie Gu, Rui Tan, and Peng Cheng. Challenges of Privacy-Preserving Machine Learning in IoT. In *Proceedings of the First International Workshop on Challenges in Artificial Intelligence and Machine Learning for Internet of Things*, AIChallengeIoT'19, pages 1–7, New York, NY, USA, 2019. Association for Computing Machinery. doi:10.1145/3363347.3363357.

60 Riaz Ahmad Ziar, Rafiulllah Omar, Irfan Ahmad, and Saber Niazy. Information Privacy Paradox and Fatigue in IoT. *Kardan Journal of Engineering and Technology*, 1(1):37–46, 2019.

Automatic creation of content for mobile language learning applications

Viktoriia Rakytianska

Ludwig-Maximilians-Universität München, München, Deutschland
V.RakytianskaL@campus.lmu.de

Abstract

Learning a language is a complex and time-intensive process. Nowadays with high access to mobile phones, it became possible for everyone to learn a language. However, it is difficult and time-consuming to create personalized content for all users. Therefore there are already some approaches to create such content automatically, but most of them deal with text, focusing on vocabulary learning. Automatic generation of other types of content, such as image, audio, or video still needs to be developed. Phonetics, grammar, and lexicology are critical while creating learning content. Despite the type of content it should be linguistically and pedagogically approved. There are a lot of technical challenges when creating different types of content for learning a language on mobile devices. In addition, the definition of the target group of learners and their needs is key.

2012 ACM Computing Classification Applied computing → Education → E-learning

Keywords and phrases Automatically generated content; mobile learning; language learning; personalization.

1 Introduction

The demand for mobile learning increased over the last years, recently further driven by COVID-19 [2, 37, 40]. Therefore, creating automatically personalized content for learning systems is nowadays of high demand. Technological development and global access to the Internet and mobile devices proved that the learning process can be accessed from anywhere anytime [30]. UNESCO in its report from 2013 writes that the flexibility of mobile learning "allows people to study during a long break or while taking a short bus ride [30]." Larry D. Rosen in his book "Rewired: Understanding the iGeneration and the Way They Learn" talks about "anytime, anywhere education, which provides opportunities outside the classroom." He points out that we "can gather content anytime and from anywhere in the world, [...] can communicate with

Cite as: Viktoriia Rakytianska. Automatic creation of content for mobile language learning applications. In *6th Seminar on Ubiquitous Interaction (UBIACTION 2021)*. Editors: F. Lang, R. Welsch, L. Haliburton, F. Draxler, S. Feger, S. Villa, M. Hoppe, P. Knierim, V. Mäkelä, A. Schmidt. February 11, 2021. Munich, Germany. pp. 12:1–12:16.

people anywhere and anytime. It is precisely this social web that provides untold [...] opportunities for mobile learning" [39]. The number of smartphone users worldwide today surpasses three billion and is forecast to further grow by several hundred million in the next few years [10]. This leads to a growing demand of user expectations. "The users of the new mobile learning applications expect usability (efficiency, effectiveness, and satisfaction) as a 'de-facto characteristic' for any mobile learning application [4]." Amir Dirin et al. [11] concluded that "the user expectations [...] were much more than just the mere functionalities of the application or just extending the uses of the mobile devices". They talk about expectations that are "beyond instrumental".

The problem is that it is very time-consuming to create comprehensive learning content for mobile learning applications. Therefore, there are already a few approaches for automatically generating content. This paper gives an overview of some case studies and existing leading mobile language learning applications. First, the content itself is examined in more detail (2 Theoretic Background of Automatic Content Creation with Importance of Linguistic (Chapter 2.1) and Automatic generation of Content (Chapter 2.2.). After analyzing language learning applications (3.1 Leading Mobile Language Learning Applications), selected research paper case studies are illustrated (3.2 Selected Case Studies). Afterwards, challenges of language learning applications and content creation are summarized (3.3 Challenges). The focus lays on technical challenges and analysis of automatically generated content of personal relevance for learners. Finally, it is examined whether and how such automatically generated and personalized content could become easily established in the future (4 Discussion and 5 Conclusion).

2 Theoretic Background of Automatic Content Creation

The field of mobile technology is very fast-moving. Therefore any theoretical approach has to be periodically revised [41], making it difficult to define, categorize and generate content for mobile devices.

Generally, *mobile content* refers to any media content, which is texts, images, videos, music, or podcasts, that can be accessed and retrieved via a mobile device [45]. Scolari et al. [41] classify mobile content as "genre-based" and "purpose-based". Genre-based content is not only related to the commercial environment and everyday events, but is also influenced by mainstream media content (e.g. news, advertisement, etc.). Purpose-based content inform, promote, educate or entertain. They also mention that definition of content depends on the application. As this work deals with content for mobile language learning applications, its genre is defined as "foreign languages content" and purpose as "educational content". The focus lays on four types of content: text, audio, video, or image.

When learning a foreign language, learning content must be correct from the linguistic perspective to ensure native learning experience (2.1 Importance of linguistics). In addition, linguistics play a significant part in understanding the difference between conversational speech, formal speech, and rules about using words in different cultures [6]. There are two ways to generate content from developer perspective: manually by creating own content or the usage of automatic content creation technology. In this work, the focus lies on automatically generated content (2.2 Automatically generated content).

2.1 Importance of Linguistic

"Linguistics can be broadly defined as the scientific study of language [6]." Traditionally there are five core branches of linguistics: *phonetics* (the study of speech sounds in general), *phonology* (the study of the sound systems of individual languages), *morphology/grammar* (the study of the creation, structure, and form of words), syntax/grammar (the study of structural units larger than one word, i.e. phrases and sentences), and *semantics* (the study of word and sentence meaning) [6]. *Pragmatics* (the study of meaning in context) and *sociolinguistics* (the study of the relationship between language and society) are two other branches of linguistics [6, 18, 20]. Learning a language with a mobile language learning application can focus on a specific branch of linguistics or all branches together. Each branch of linguistic requires a specific type of content, which later on will contribute into success in each part of the language learning process. When learning a foreign language, be it academic or social level, the focus usually lays on these three branches of a language:*phonetics*, *grammar*, and *lexicology* [6, 23].

Phonetics studies human sounds, and **Phonology** classifies these sounds within a particular language. Audio, video, and text[1] as content types can help learn the pronunciation of words, phrases and sentences and perfect listening comprehension [17].

Grammar represents a language system, explaining the forms and structure of words and their arrangement in sentences. It is a core of written and spoken communication [6]. Therefore, audio, video, and text as content types

[1] Here: transcription

can be used to learn the language's grammar.

Lexicology studies the meaning of words and their semantic relations [6]. All four types of content (audio, video, text and image) can be used to learn lexicology.

2.2 Automatic Generation of Content

Learning content should be presented to the learners instead of them looking for it themselves [38]. There are two types of automatically generated content: local/context-aware content and personalized content. *Local content* is based on the users' location information and surrounds them with relevant information. *Context-aware content* surpasses geographical location and mixes features from the environment with the mobile personal profile or person's activities. Such content can help users make better decisions and find much suitable content [19]. Personalizing content for users is still a very young discipline. The goal of *personalized content* is to ensure that it corresponds to the users' needs and interests. The systems that generate personalized content are called *recommender systems*. There are two types of such systems: *collaborative* and *content-based methods*. *Collaborative filtering methods* are based on user interactions with items to generate new recommendations and then identify the same taste users. *Content-based filtering methods* base their predictions on user's information. Recommendations here don't depend on the contributions of other users. Mostly, the services use both of these systems to make a better prediction [7]. Such and similar algorithms are used today in services such as Twitter (text), Spotify(audio), Netflix (video), and Instagram (image).

3 Language Learning Applications

3.1 Leading Mobile Language Learning Applications

As to SimilarWeb here is the leading language learning applications for web, iOS and Android as of February 2021 [42]. According to visits per month[2], which includes both application and web traffic, ten language learning applications stand out the most. With a large gap Duolingo is a leading application for language learning with 71.20 Millions visits per month. Seven applications such as HiNative, Babbel, italki, Fluentu, Memrise, Roseta Stone and Busuu have millions of visits per month [42].

There is no significant difference in rating of the presented language learning applications for both iOS [3] and Android [32] systems. However, as for

[2] One user can have several visits.

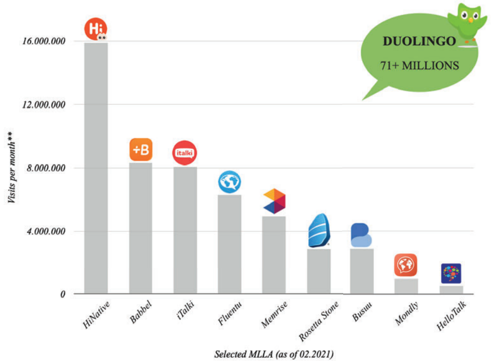

Figure 1 Monthly visits of selected Mobile language learning applications

February 2021 Memrise and Mondly are the most well-rated applications for learning a language.

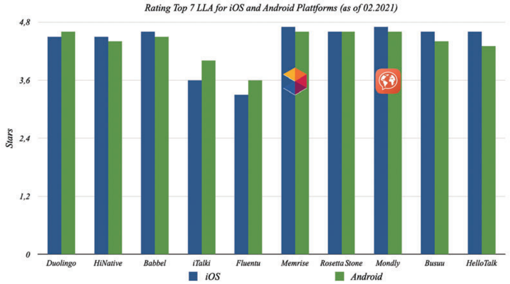

Figure 2 Rating of selected mobile language learning applications

The goal is to look at the following leading applicaitons on the market and find out if they personalize their content and generate it automatically.

Duolingo uses AI to deliver personalised language learning [46]. Before starting the language course the learner needs to complete a proficiency assessment, where each answered question is determined by the previous answer. This builds a language profile for each learner and helps to start the language from the right level. Using a concept called spaced repetition, language lessons are designed so users practice personalized tasks over longer and longer intervals. Using AI and machine learning Duolingo can make intelligent predictions and personalised the learning content. Duolingo has online events, where one can participate in conversations with native speakers and other learners [12]. Moreover, the learner can also make his or her own courses and present them to other learners. Besides, there is a Duolingo podcast, where the learner can listen to audio and see the original text, which makes it easy to follow the conversations. [13]. As of march 2021 there are three supported languages: Spanish for English speakers, French for English speakers and English for Spanish speakers. Everyone learns differently, therefore there are variety of features to motivate every learner [14]. For example, Leaderboards. Every week the learner will be paired with 30 other learners to find out the best performer. This feature is optional and can be disabled by user any time [15]. Duolingo has integrated chatbot, equipped with AI algorithms to understand the user context and generate personalized responses [31]. The bots have different personalities. There's Chef Robert, Renée the Driver and Officer Ada, for example. They react differently to the learner's answers and try to simulate a real conversation. In case the learner stucked, there is a "help my reply" button. For now the bots support chats and spoken conversations are in plans for the future.

HiNative [24] is a Q&A app for language learners that want to have questions answered by native speakers. It is a collaborative platform, where everyone helps each other learning a language. There is no automatically generative content because of the concept of the application.

Babbel [4] has a speech recognition feature, so one can practice. There is also a review manager, that asks if you remember learned words. This makes the words stay in your long term memory. Babel proposes to learn grammar though dialog. This is a dynamic content that depends on the answers of the learner. Learning content is created manually by the team of linguists and native speakers.

italki [26] has customized learning features, where one can choose a teacher for 1-0-1 lessons based on the goals and interests of the learner. There is no automatically generated content, as lessons are done by thousands of teachers and therefore depends on each teacher.

FluentU [21] presents video content for language learners, such as movie trailers, music videos, news or talks. Interests and skill levels are taken into consideration. All captions are subtitled and translated. The learner can click on any word to see an in-context definition, along with example sentences.

Memrise [33] focuses on vocabulary learning with personalization in mind. The learners can take a photo of any object and feed to the app to know the name of the object in the user's desired language. Also, one can set daily goals

and learning reminders, as well as join leader-boards to compete against other learners.

Roseta Stone [43] has focus on phonetics and uses TruAccent speech engine to ensure users get the right articulation. It uses augmented reality and embedded translation features to teach users in real time.

Mondly [36] has incorporated chat-bots in its online learning as well as via a mobile app where the learner can choose to either type or speak the responses.

Busuu [8] creates the content manually by expert linguists, enhanced with machine learning technology. Busuu offers smart features like personalized study plans and intelligent vocabulary trainer, allowing learners to study at their own pace. There is a possibility to get feedback from native speakers and even choose one-to-one online tutoring. The learner has control of the learning plan and progress, because there is a monitor that shows the overall and individual performance with the Management Platform's dashboards.

HelloTalk [22] focuses on learning a language from native speakers by chatting. It is a tandem logic, where the native speaker teaches his or her language and the learner does the same in return. The learner chat with language partners via text, voice recordings, voice calls, video calls, and even doodles. HelloTalk interface has a built-in aids for translation, pronunciation, transliteration, and corrections. Moreover, the learner can chat with not only with one native speaker, but even join group chats for a collaborative learning experience.

3.2 Selected Case Studies

This section gives an overview of 19 selected case studies for learning applications. The case studies use either text, audio, video or image as a medium to transfer learning content to the learner. Most of the selected case studies are related to text as the main medium, while audio, video and image are less implemented. In addition, the case studies were segmented along two dimensions: Technology and Content. At the dimension Technology, some of the case studies relate to web implementations, while others are focused on mobile. The Content dimension differentiates between Personalisation focused technologies as well as Automatisation focused technologies.
The following text related case studies are illustrated below: *Twasebook* (Case study 1), *FeedLearn* (Case study 4), *MicroMandarin* (Case study 8), *Wait-Learning* (Case study 9), *Enhancing Authentic Web Pages* (Case study 11)). In addition, *Mobile Adaptive Call* (Case study 3) and *Context-Sensitive Microlearning* (Case study 18) for audio content as well as *Smart Subtitles* (Case study 19) for video content are described in the following.

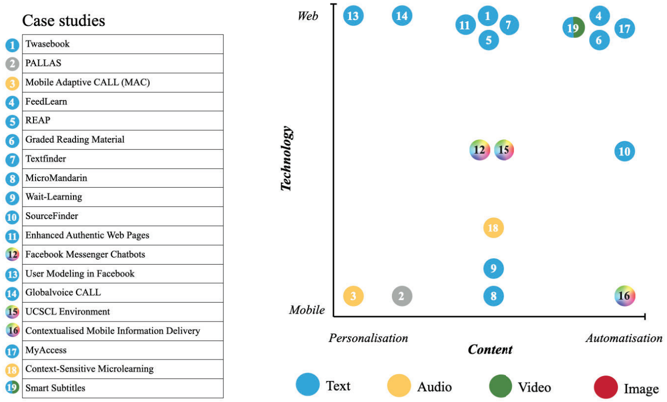

Figure 3 Overview of segmented case studies

Text related case studies (selected). *Twasebook* (Case study 1) [10]: The web application identifies status updates from Twitter in the learners target language. It pulls up the latest tweets that mention the searched term or word in combination with an automatic translation. The proof of concepts utilizes large available open content from social networks to provide and train the learner with everyday vocabulary. Overall, it helps the learner to improve its vocabulary, especially for less common words.

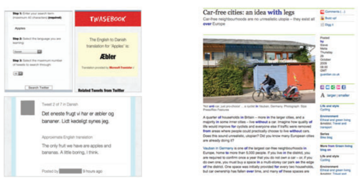

Figure 4 Left: Twasebook; right: FeedLearn

FeedLearn (Case study 4) [28]: This application uses Facebook feeds for microlearning. Implemented as a Google Chrome extension, it shows users interactive quizzes to answer without leaving their Facebook feeds. *Micro-Mandarin*(Case study 8) [16]: Asian contextualized language microlearning via mobile, especially throughout the day with many breaks and waiting times at different places. Based on the location-based service Foursquare, the application provides context relevant learning content to learners in large cities. *Wait-Learning* (Case study 9) [9]: While the user awaits a response during instant messaging conversations, WaitChatter translates contextual relevant foreign language vocabulary just in time and provides it to the user. In addition, the application provides micro-quizzes. Enhancing Authentic Web Pages (Case study 11) [34]: The web based application automatically provides visual input enhancement for web pages. This supports the language awareness of learners, especially for difficult patterns such as lexical classes, gerunds and infinitives or, condition and phrasal verbs.

Audio related case studies (selected). *Mobile Adaptive Call* (Case study 3) [44]: The mobile application helps Japanese-English learners to distinguish the native phonemic contrast between native and non-native language in speech via audio. It calculates the learner's proficiency and checks against a database to systematically track learning improvement. The idea behind is mainly to reduce discrimination based on pronunciation. *Context-Sensitive Microlearning* (Case study 18) [5]: Based on a sensor-triggered mobile phone application, the learner receives audio feedback with words or sentences based on the object being in touch with. For instance, receiving messages in the household around a sofa, when sitting on the furniture.

Figure 5 Left: Mobile adaptive call; right: context-sensitive microlearning

Video related case studies (selected). *Smart Subtitles* (Case study 19) [29]: Providing smart and interactive subtitles for videos while watching to improve vocabulary learning. Common video subtitles are often more comprehensive, not tailored to learning and result in less learning success. Repeated onomatopoeic reproduction of passages during the film help the learner to better remember the language.

3.3 Challenges

The following challenges focus exclusively on technical aspects. Strategic challenges are not taken into account (e.g. deciding which technology to use with regard to the other app providers). The well-known learning app providers are already facing most of the following challenges today, although isolated challenges still represent an open technical problem that will have to be solved in the future.

Mobile Device Issues:

- Adaptation to different mobile systems (iOS, Android);
- Handling patents so that no patent rights are infringed;
- Increased user expectations and demands;
- Screen size and the methods of inputting [1, 25].

Additionally, other issues such as storage capacity, processor speed, battery life, and compatibility of devices (in terms of both operating systems and transfer of large amounts of data) have also been raised as points needing consideration in implementing mobile devices in learning contexts [27].

Automation/personalization Challenges:

- Efficiency increase in data processing (e.g. processing grammar of texts; integration of existing algorithms);
- Necessity of a content model.

There are examples where the content such as text and picture is generated automatically. However, there are no applications, that generate video or audio content automatically.

Recommender Systems Challenges [35]:

- integration of content based recommendation systems into existing language apps and make it work properly;
- large amount of data needed, so recommendations can work effectively;
- cold start problem;
- scarcity problem (users must have a possibility of rating the system);
- privacy.

Development of new technologies:

- Difficulty of automation (examples: Machine Learning needs time and many data sets to work well;
- Different and changing data formats for audio, video, text, image, etc.);
- Data protection (necessary: user consent) Selection of methods for personalization that make the majority of all users "happy": (1) user selects topics himself - e.g. "Christmas", (2) app automatically suggests content to user based on various dimensions: Time, location, demographics (age, gender, etc.) and/or links to social media (Facebook, Instagram, Twitter, LinkedIn, newspaper, Spotify, YouTube).

4 Discussion

Creating automatically generated and personalized content, not just for language learning applications, be it mobile or web, is still a difficult and challenging task. The work presented here offers some overview of existing mobile language learning applications as well as selected approaches and some solutions on how to integrate these into a language learning process. However, the existing approaches (case studies) have not yet been investigated in long-term studies and with significant amount of participants.

The leading language learning mobile applications on the market rarely generate learning content automatically. Nevertheless, almost all of them present themselves as an application that personalizes the content. However, it is far away from the truth. These applications rarely provide possibilities to personalize any type of content, be it text, image, audio or video. However, some applications like Duolingo or Rosetta Stone give an opportunity for some personalized features, like personal learning plan or personal feedback from a native speaker. HelloTakl makes it also possible to give and to receive a feedback from own messages (texts), but it is rather the whole concept of the application.

It is well known, that a language has a dynamic evolutionary nature. It lives with people through the people. No wonder there are already some approaches to learn the language with a help of social media, like Twitter, Facebook or Instagram. Some propose a web browser gadgets to learn the language with the help of Netflix. Also, many people learn language through the lyrics of their favorite band or singer. It is understandable, because in films, song, newspapers and social media posts one can find a truly live language like it is nowadays. So, it brings a thought of possibility to integrate for example existing multilingual translation services like Google Translator or DeepL in an existing language learning applications.

Using recommender systems became very popular in the past years. There is a huge potential to use them in language learning applications as well. However, there are some challenges of doing that. A lot of data must have been processed and evaluated beforehand, so the system can individually adapt to the user. Nowadays, it is difficult to keep collected data anonym and secure it, not mention the legal perspective of this topic. In addition, a large community is necessary in order to match content well with other data. Only then better suggestions can be offered for the target user.

Another possibility will be an integration of language learning feature to the social media (Facebook, Instagram, Youtube, Twitter, Netflix, Newspaper). Facebook developed a chatbot for such purpose. One can learn from the posts with a help of chatbot or translate the posts directly into English. Youtube and

Netflix have an option to use subtitles in their videos. Vice versa social media can be integrated in existing language learning app or a separate application can be created like Twasebook.

Even if it will be possible for a user to create a personal automatically generated content like it is already partly possible in the learning app Duolingo,, there are a lot of technical and pedagogical challenges. A key challenge is to select, sort and proof the right content from a large data base for the right user. A teenager might learn a language with Snapchat, a grandfather with a classical language learning application or still using a book. In addition, how can learners be motivated to use these technologies at all? The question of whether automatic content generation and personalization can sustainably increase progress and finally success in learning a foreign language has not yet been clearly answered. Too few applications have been tested on a large scientific scale and examined for their long-term effect. The relevance of the topic is given and its potential should be further explored.

5 Conclusion

Development of technologies in many spheres of our lives has produced a lot of data about us. Automatically generated and personalized content can be based on large available information. While we can receive a recommendation for an item to buy a film, a song, an article in a newspaper or even a post in social media, it is still challenging to receive such content when learning a foreign language. Technically it should be possible to cover linguistic, pedagogical and user perspectives. None of the leading mobile learning language applications presents such a possibility in an advanced way. There are only some personalized features, including an individual learning plan or live learning via chat. Nevertheless, some approaches automatically generate the content, but only the text. Approaches for another types of content for mobile language learning applications such as image, audio and video are not yet developed. The question remains: If automatically personalized content for language learning applications is developed, how will it influence the learning progress?

References

1 Glenn Stockwell . Using mobile phones for vocabulary activities: Examining the effect of platform. *Language Learning and Technology*, 14, 06 2010.

2 Antonius Setyawan Nur Agung, Monika Widyastuti Surtikanti, and Charito A. Quinones. Students' perception of online learning during COVID-19 pandemic: A case study on the english students of STKIP pamane talino. *SOSHUM : Jurnal Sosial dan Humaniora*, 10(2):225–235, jul 2020. doi:10.31940/soshum.v10i2.1316.

3 AppStore. https://www.apple.com/app-store/. Accessed: 2021-03-23.

4 Babbel. Accessed: 2021-01-15. URL: https://de.babbel.com.

5 Jennifer S. Beaudin, Stephen S. Intille, Emmanuel Munguia Tapia, Randy Rockinson, and Margaret E. Morris. Context-sensitive microlearning of foreign language vocabulary on a mobile device. In Bernt Schiele, Anind K. Dey, Hans Gellersen, Boris de Ruyter, Manfred Tscheligi, Reiner Wichert, Emile Aarts, and Alejandro Buchmann, editors, *Ambient Intelligence*, pages 55–72, Berlin, Heidelberg, 2007. Springer Berlin Heidelberg.

6 Annette Becker and Markus Bieswanger. *Introduction to English Linguistics*. UTB GmbH, February 2017. URL: https://www.ebook.de/de/product/28205495/annette_becker_markus_bieswanger_introduction_to_english_linguistics.html.

7 J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez. Recommender systems survey. *Knowledge-Based Systems*, 46:109–132, jul 2013. doi:10.1016/j.knosys.2013.03.012.

8 Busuu. Accessed: 2021-01-13. URL: https://www.busuu.com.

9 Carrie J. Cai, Philip J. Guo, James Glass, and Robert C. Miller. Wait-learning: Leveraging conversational dead time for second language education. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems*. ACM, apr 2014. doi:10.1145/2559206.2581183.

10 Graeme W. Coleman and Nick A. Hine. Twasebook: A "crowdsourced phrasebook" for language learners using twitter. In *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design*, NordiCHI '12, page 805–806, New York, NY, USA, 2012. Association for Computing Machinery. doi:10.1145/2399016.2399157.

11 Amir Dirin and Marko Nieminen. User experience evolution of m-learning applications. 04 2017. doi:10.5220/0006370301540161.

12 Duolingo. Duolingo online events. Accessed: 2020-12-10. URL: https://events.duolingo.com.

13 Duolingo. Duolingo podcasts. Accessed: 2020-12-10. URL: https://podcast.duolingo.com.

14 Duolingo. Guide for leaders in education. Accessed: 2020-12-15. URL: https://duolingo-data.s3.amazonaws.com/s3/schools/active/Duolingo_for_Schools_Guide.pdf.

15 Duolingo. What are leaderboards? Accessed: 2021-01-06. URL: https://support.duolingo.com/hc/en-us/articles/360035931732-What-are-Leaderboards-.

16 Darren Edge, Elly Searle, Kevin Chiu, Jing Zhao, and James A. Landay. Micromandarin: Mobile language learning in context. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, page 3169–3178, New York, NY, USA, 2011. Association for Computing Machinery. doi:10.1145/1978942.1979413.

17 Universität Duisburg Essen. Phonetics and phonology, January 2021. Accessed: 2020-12-09. URL: `https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/`.

18 Ralph Fasold. *An introduction to language and linguistics*. Cambridge University Press, Cambridge, UK New York, 2006.

19 Claudio Feijóo, Ioannis Maghiros, Fabienne Abadie, and José-Luis Gómez-Barroso. Exploring a heterogeneous and fragmented digital ecosystem: Mobile content. *Telematics and Informatics*, 26(3):282–292, aug 2009. `doi:10.1016/j.tele.2008.11.009`.

20 Edward Finegan. *Language : its structure and use*. Thomson Wadsworth, Boston, MA, 2008.

21 FluentU. Accessed: 2021-03-17. URL: `https://www.fluentu.com`.

22 HelloTalk. Accessed: 2021-01-13. URL: `https://www.hellotalk.com`.

23 Raymond Hickey. Standards of english: Codified varieties around the world (studies in english language series). cambridge, u.k.: Cambridge university press. 2012. xxii + 421 pp. hb (9780521763899) £70.00. *Journal of Sociolinguistics*, 18(5):708–713, nov 2014. `doi:10.1111/josl.12097`.

24 HiNative. Accessed: 2021-01-15. URL: `https://hinative.com/en-US`.

25 Patricia ThorntonChris Houser. Using mobile phones in education. In *The 2nd IEEE International Workshop on Wireless and Mobile Technologies in Education, 2004. Proceedings*. IEEE. `doi:10.1109/wmte.2004.1281326`.

26 italki. italki. Accessed: 2021-01-15. URL: `https://www.italki.com`.

27 Marguerite Koole. *A Model for Framing Mobile Learning*. 03 2009.

28 Geza Kovacs. FeedLearn: Using Facebook Feeds for Microlearning. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, apr 2015. `doi:10.1145/2702613.2732775`.

29 Geza Kovacs and Robert C. Miller. Smart subtitles for vocabulary learning. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, apr 2014. `doi:10.1145/2556288.2557256`.

30 Rebecca Kraut. *Policy guidelines for mobile learning*. United Nations Educational, Scientific and Cultural Organization, Paris, 2013.

31 Frederic Lardinois. Duolingo's chatbots help you learn a new language, October 2016. Accessed: 2021-01-10. URL: `https://techcrunch.com/2016/10/06/duolingos-chatbots-help-you-learn-a-new-language/`.

32 Google LLC. Google play, 2021. Accessed: 2021-03-22. URL: `https://play.google.com/`.

33 Memrise. Accessed: 2021-01-17. URL: `https://www.memrise.com`.

34 Detmar Meurers, Ramon Ziai, Luiz Amaral, Adriane Boyd, Aleksandar Dimitrov, Vanessa Metcalf, and Niels Ott. Enhancing authentic web pages for language learners. In *Proceedings of the NAACL HLT 2010 Fifth Workshop on Innovative Use of NLP for Building Educational Applications*, pages 10–18, Los Angeles, California, June 2010. Association for Computational Linguistics. URL: `https://www.aclweb.org/anthology/W10-1002`.

35 Marwa Hussien Mohamed, Mohamed Helmy Khafagy, and Mohamed Hasan Ibrahim. Recommender systems challenges and solutions survey. In *2019 International Conference on Innovative Trends in Computer Engineering (ITCE)*. IEEE, feb 2019. `doi:10.1109/itce.2019.8646645`.

36 Mondly. Accessed: 2021-01-13. URL: `https://www.mondly.com`.

37 Aziz Naciri, Mohamed Amine Baba, Abderrahmane Achbani, and Ahmed Kharbach. Mobile learning in higher education: Unavoidable alternative during COVID-19. *Aquademia*, 4(1):ep20016, apr 2020. `doi:10.29333/aquademia/8227`.

38 Sobah Abbas Petersen and Jan-Kristian Markiewicz. PALLAS: Personalised language learning on mobile devices. In *Fifth IEEE International Conference on Wireless, Mobile, and Ubiquitous Technology in Education (wmute 2008)*. IEEE, mar 2008. `doi:10.1109/wmute.2008.17`.

39 Larry D. Rosen. *Rewired*. ST MARTINS PR 3PL, March 2010. URL: `https://www.ebook.de/de/product/9147333/larry_d_rosen_rewired.html`.

40 V. Sathishkumar, R. Radha, Ar Saravanakumar, and K. Mahalakshmi. E-learning during lockdown of covid-19 pandemic: A global perspective. *International Journal of Control and Automation*, 13:1088–1099, 06 2020.

41 Carlos Alberto Scolari, Juan Miguel Aguado, and Claudio Feij. Mobile media: Towards a definition and taxonomy of contents and applications. *International Journal of Interactive Mobile Technologies (iJIM)*, 6(2):29, apr 2012. `doi:10.3991/ijim.v6i2.1880`.

42 SimilarWeb. February 2021 overview, 2021. Accessed: 2021-03-21. URL: `https://www.similarweb.com`.

43 Rosetta Stone. Rosetta stone. Accessed: 2021-01-13. URL: `https://www.rosettastone.com`.

44 Maria Uther, Iraide Zipitria, James Uther, and Pushpendra Singh. Mobile adaptive CALL (MAC): a case-study in developing a mobile learning application for speech/audio language training. In *IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE'05)*. IEEE, 2005. `doi:10.1109/wmte.2005.46`.

45 Ryte Wiki. Mobile content. Accessed: 2021-01-21. URL: `https://de.ryte.com/wiki/Mobile_Content`.

46 Kevin Yancey and Burr Settles. "hi, it's duo": The ai behind the meme, September 2020. Accessed: 2021-03-09. URL: `https://blog.duolingo.com/hi-its-duo-the-ai-behind-the-meme/`.

New Interaction Paradigms with Multi-device Interaction in Mixed Realities

Tamia Eliza Josephine Bosch

Ludwig-Maximilians-Universität München, München, Deutschland
t.bosch@campus.lmu.de

Abstract

In this paper, we examine the usage of multiple devices to interact in Mixed Realities. We present different input techniques, including controller-based input, motion tracking as well as eye-gaze-based tracking. Each of these options is demonstrated with examples in current research papers. For a new interaction paradigm with Mixed Realities, combining other smart devices appears to be inevitable. Creating a seamless interaction between the digital and physical world is a crucial part of forming a new interaction paradigm. Therefore further research is presented which combines smartphones or tablets with Augmented Reality. These projects offer interaction and direct manipulation of the environment. Through the usage of cross-device interaction in these researches, the usual display space is extended into the augmented realm. This new space offers an additional 3D presentation of content. There are still difficulties concerning accurate tracking and uniform way of interaction to be suited for usage in everyday life.

2012 ACM Computing Classification Human-centered computing → Mixed / augmented reality

Keywords and phrases Mixed Reality; Augmented Reality, Cross-Device Computing; Interaction paradigms.

1 Introduction

The borders of a smartphone should not be the edges of the display. Using the smartphone as a remote control or connecting it to other smart devices is already commonly used. Researchers are thinking further and want to develop new cross-device interaction between smartphones or other smart devices and Mixed Reality (MR). This broad field of interaction in MR includes multi-device interaction with Virtual Reality (VR) and Augmented Reality (AR), which is approached in different research papers. This paper aims to present current research and find new interaction paradigms that emerge out of these different interaction models.

Cite as: Tamia Eliza Josephine Bosch. New Interaction Paradigms with Multi-device Interaction in Mixed Realities. In *6th Seminar on Ubiquitous Interaction (UBIACTION 2021)*. Editors: F. Lang, R. Welsch, L. Haliburton, F. Draxler, S. Feger, S. Villa, M. Hoppe, P. Knierim, V. Mäkelä, A. Schmidt. February 11, 2021. Munich, Germany. pp. 13:1–13:16.

Researchers are examining new interaction methods aiming for an immersive experience to diminish the borders between the physical world and smart devices. Natural and intuitive handling for users with multiple smart devices in MR is vital for a new interaction paradigm. The ability to use the immediate surrounding space as a display surface and connecting it to seamless interaction with our current smart devices opens up a new dimension for interaction [27].

The standard smartphone is a prevalent device, which is therefore commonly used and known among current end users. The idea is to widen the use of this ubiquitous tool. Researches present different ideas to use the device for interaction in VR and AR. Different input techniques and interaction possibilities with smartphones as controllers or input devices will be presented in section 3.1. Since the device already offers well-tested features like precise touch input or haptic feedback via vibration, the camera can also be used as a tracking device [19, 29]. Therefore the device is suitable for multipurpose usage. Furthermore, emerging problems and limitations that come with the usage of these ideas and interaction methods are presented in section 3.2.

In section 4 new interaction paradigms in MR that derive from the current research will be discussed. This is followed by pointing out research gaps in multi-device interaction and the overall conclusion and outlook obtained from the discussion.

Before presenting more research papers in detail, we need to define important contextual terminology and mark out the area of research.

2 Definitions and Current State of the Art

This section will give an introduction to key terminology and its definitions. The terms AR and VR are explained in more detail and differentiated from each other, as well as the definition of cross-device interaction. Furthermore, relevant research papers will be discussed in detail. These will form the basis of further investigations for new interaction paradigms in MR.

2.1 Reality–Virtuality Continuum

The Reality-Virtuality Continuum was formulated by Milgram and Kishimo [17] to represent the broad scope of MR. As shown in Figure 1 MR is located between the two extremes. On the left side of the spectrum, there is the real physical environment on the right side is the completely virtual environment, where everything is computer-simulated. Thus, MR is defined as an environment somewhere in between these two extrema. It needs to represent objects from both the real world and virtual reality within one display.

Milgram describes the illustrated term of Augmented Virtuality (AV), as a technology that is primarily in a virtual environment and includes a few

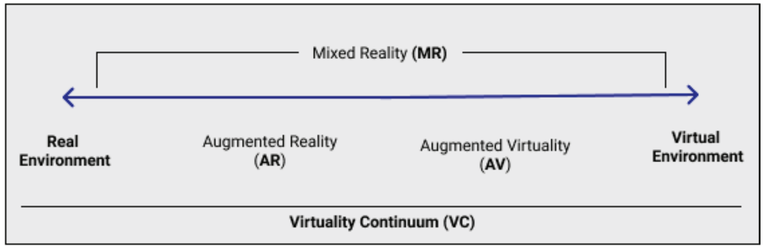

Figure 1 Reality-Virtuality Continuum simplified [17].

real-world objects [18]. This term is currently not often used but describes a technology on the edge to complete Virtual Environments.

Azuma [1] defines the term Augmented Reality (AR) as a variation of Virtual Reality (VR). In VR the user moves in a completely artificial area. This means that no part of the surrounding real world is visible. AR complements the surrounding reality with virtual objects, which ideally fit into the environment. It is further defined as a system with three main characteristics. It firstly combines real and virtual secondly, it is interactive in real-time, and lastly, it is registered in 3D. This clear definition avoids limiting technology in the way that some researchers associate AR with the need for Head-Mounted-Displays (HMDs). This paves the way for AR systems without the need for HMDs and their limitations.

There are no clear boundaries in the Virtuality Continuum (Figure 1), and therefore, the investigated topic of multi-device interaction in MRs is rather to be classified as an area of this broad-spectrum, which contains the three previously mentioned characteristics.

2.2 Cross-device Interaction

Cross-device interaction spans a wide research topic that has been growing rapidly. Brudy [5] conducted many papers covering this large field and presented a taxonomy, which is now widely used as a stable reference. In Brudys taxonomy, the definition of cross-device and multi-device are equivalent, this also applies on this paper. The terms define the usage of any smart devices spanning from displays, laptops, smartphones, tablets, large surfaces, tangibles, headsets, and IoT devices.

The great advantage of cross-device interaction is to overcome the limitations of individual devices. Therefore the advantages of multi-device interaction are a part of finding new ways to interact with digital content. Above all, it is important to find a uniform terminology to explore this topic further.

Important for an effective cross-device interaction is a suitable process for data exchange. This often requires a tracking system that reliably tracks the individual input devices on the one hand and the user of that device on the other. There are different systems with different strengths and weaknesses. Usually those tracking systems can be divided into the following two categories: *outside-in* and *inside-out*. Outside-in requires a static sensor in the room, which tracks the position mostly through optical tracking. Therefore outside-in systems mainly use depth cameras, which are currently the dominant technology. This means the user needs to be in the field of view of the device, otherwise the tracking is lost. Brudy anticipates this system to be more user context-aware, which is key for integrating this system in cross-device interaction. A different approach is **inside-out** tracking, it uses sensors that are already integrated into the device. This means these devices can define their own positions in the room without the need for cameras in the surrounding. Therefore inside-out tracking is practical for mobile cross-device applications since these systems are independent of external cameras. For orientation, the tracking system mainly uses acoustic, radio, and recently also optical signals while using device cameras. This kind of tracking provides a 2D position of the devices. A reliable 3D position with non-spatial sensors is still a major difficulty. [5]

The majority of cross-device interaction research with the use of mobile devices is developed in conjunction with other 2D devices like wearables, tablets, and also large display [29]. Even though hand and gesture tracking is an important part of interaction in MR it is not a necessity in combination with multi-device interaction and according to the previously defined characteristics concerning AR systems. An idea of cross-device interaction within MRs is presented by Heun [10] and the project of the Reality Editor. The combination of AR techniques and cross-device interaction allows the Reality Editor to operate different smart devices within one interface. The goal is to avoid the installation of multiple apps for every single device. With the use of the direct environment in connection with AR, smart devices can be operated easily and flexibly. The Reality Editor maps a graphical interface on top of the tangible physical surfaces with the help of augmented reality. With this interface, the devices can now be manipulated, new functionalities can be created and these can also be logically connected to other devices in the room. It is possible to use this interface to operate and manipulate smart lights. Further, the user can link the switch to devices such as a radio or other objects.

3 Possibilities and Limitations in Current Research

The following section outlines the necessary tools and interaction techniques, which will build the foundation of new interaction paradigms between smart devices and MRs. First, different input methods are shown. Furthermore, the opportunities of smartphones as input devices and the advantages of using the device as a controller are pointed out. Also, the current problems and limitations concerning certain interaction methods, especially in shared social spaces, are demonstrated. New researches where smartphone cameras are used as tracking devices with Six Degrees of Freedom (6DoF) are also presented.

3.1 Input Techniques and Interaction

The way a user naturally interacts with certain smart devices is the basis for researchers to study different input techniques and expand them further. Firstly, we need to clarify which devices we investigate for input techniques. Since the focus is on interaction in MRs, we will mainly look at methods that are used in AR but also include some research in the field of VR. Consequently, this also includes HMDs, which are frequently used in this research area and play an important role in the interaction model. For entry-level users, devices like Oculus Go or Google Daydream are commonly used. But the lack of support of tracking with 6DoF is still a missing feature. [19]

Interaction in VR is aiming to detach the user from the real world entirely. With an HMD an illusion is created where the user is immersed in a completely virtual environment [28]. This detachment in VR is supported by an interaction that is mostly controller-based and, therefore, renounces the immediate environment. The user can operate in the displayed environment with natural hand movements and active selections of objects but needs to hold external controllers in both hands [9, 24].

There are efforts to integrate daily used devices like smartphones into the virtual environment. This means that interaction with real-world devices is possible within the simulation. Through that, users are not distracted from their surroundings and can stay in their immersive experience. To make this possible, Zhang and Billinghurst [28] researched the display and usability of smartphones within a virtual environment. This means that the user does not have to leave the current experience if he wants to interact with his smartphone.

Eye-tracking, gesture input, and also a combination of these two input methods are currently researched as an interaction technique. The difficulty in the development of eye-tracking input is to avoid the so-called Midas Touch problem [25]. The idea is that the users can select an object simply by looking at it for a longer period of time than average. The problem with this is an unintentional selection of objects with the gaze of the user. This kind of accidental selection can also occur in the area of gesture control.

The combination of target selection by gaze and the further manipulation of the object on a touch display was researched by Pfeuffer [21]. This gaze-touch combines the advantages of both interaction techniques. The gaze is used for selection, and the multi-touch input is precise enough for further manipulations, this widens the interaction possibilities for the user. This gaze-touch combination tries to avoid the Midas Touch problem since only the targeting is gaze-based, and the manipulation of objects prevails with common multi-touch gestures. This combination of input techniques enhances a seamless interaction with smart devices.

Other papers further explored this combination by adding the distinction of eye vs. head gaze input. Pathmanathan [20] therefore compared these two different gaze methods for interaction in AR. Further, object manipulation was carried out with external clicker devices. The participants of the user study preferred targeting via head gaze, though the gaze tracking with a Microsoft HoloLens 2 was often prone to error. Another downside of the preferred head tracking is the quick fatigue because users are unaccustomed to move the head so frequently. Other studies try to circumvent the Midas Touch problem by making the targeting of an object depend on the dwell time. A threshold is set that is longer than the normal dwell time to activate a selection [25]. Another approach is continuous gaze tracking. The study by Khamis [14] primarily examines the implementation of smooth pursuit eye movements in VR. The pursuits in VR works for selection. However, it is error-prone once a participant uses it while walking. There are still selection delays due to threshold settings. To accelerate this, the combination with a controller can be tested in future work.

Manipulation and rotation of objects in AR and the use of 6DoF represent another challenge of eye-tracking input. Liu's study [16] investigated the rotation of objects using eye-gaze for HMDs. The focus was rather on the usability of three different input methods. No preferred method could be discovered in the user study, but there were different advantages for different task types. This research aims to provide another input method of 3D modeling and editing for people who may have limitations with direct hand interaction.

Another important aspect of a good user experience in MRs is haptic feedback. This is hardly given in the examples mentioned. The use of handheld controllers could support this, but the use of these is rather unnatural. Haptic feedback has a significant impact on the immersive experience and is an important factor in making a physical interaction as real as possible [4].

The most intuitive way to interact with virtual objects is with bare hands or the use of instrumented gloves [24]. With the usage of gloves, the tracking of finger motions is possible paired with a wearable, like a wrist tracker it can also provide orientation and position [11]. Novel hand tracking methods do not need gloves and enable free-hand interaction, which supports an immersive experience. Technologies such as Leap Motion and Oculus Quest support this. Theoretically, there are no spatial restrictions, but the user's hands must be within the tracking area of the camera. Otherwise, tracking is lost. [6]

3.1.1 Smartphone as Input Device

The usual operation of a smartphone is mostly done with touch gestures. In order to make the input methods more diverse, new input paradigms were researched in the Paper by Zhu [29]. Interactions with a smartphone, such as orientation, grasping, tilting, and also shaking, were investigated. These interaction methods, which are currently not common, can be used to overcome the physical limitations of the smartphone and expand input and output ranges. To ensure an immersive experience, the need for haptic feedback was also mentioned. The use of a smartphone as an input device or controller automatically includes this kind of feedback.

The combination of smartphones with AR devices is also being explored. The limits of the smartphone display can be extended with AR and thus offer additional visualization and manipulation surfaces. In most researches, the external space was mainly used for the visualization of objects in 3D space. The input techniques continued to be mainly via the precise touch gestures using the display of the smartphone. [15, 23, 29]

Ren [23] tested different input methods for window management on a smartphone in combination with an AR headset. The participants should perform the window management task under four different aspects. Semantic touch input, the immediate posture of the smartphone, position and movement awareness, and mid-air gesture were tested. The participants mainly focused on the smartphone as a tool for window management, with the recommendation to include the position of the smartphone for discrete management operations. But also, the mid-air gestures were suggested in combination with pinching gestures.

A similar project called Bishare [29] also combined the advantages of the precise input with a smartphone and AR interactions. The idea was to use the advantages of cross-device interaction with a true bidirectional usage of both smartphone and AR HMDs. This offered a seamless interaction between the devices, which was also paired with freehand gestures. The idea was to display 3D Objects either directly on the smartphone display, augmented beside or above the smartphone, or displayed in the direct environment. This interaction with the content was possible either through hand gestures, like pinching or swiping, or direct manipulation on the touch display.

3.1.2 Smartphone as a Controller

Another use of the smartphone in MRs is to turn it into a controller. This was realized with the project Pocket6 [2]. This controller offers an inexpensive and simple alternative to expensive and outside-in tracked controllers. Through a simple app, the smartphone can be turned into a precise controller. The smartphone can track the user's input in real-time as a Six Degrees of Freedom (6DoF) controller. It delivers accurate results, and through a suitably selected algorithm, tracking is also possible while walking. This means that movements, whether sitting or standing, do not restrict the input.

Besides, the smartphone can also be used as an inside-out tracking controller in the VR/AR HMD area. Inside-out tracking is used to avoid the problems like occluded areas from the sensors. In the Paper of Hattori [9] an HMD was combined with a smartphone that can estimate its position and angle via image recognition. It acts as a controller, which calculates its own position. This information is then transmitted to the HMD. This can then perform ray-casting based on the given information. This combination of HMD with the smartphone is another cheap and simple alternative for otherwise more expensive controllers.

3.2 Limitations in Interaction

The previous papers primarily examined interactions in which a user is at the center. Immersive VR and AR experiences are mostly limited to the actual user. These previously discussed interaction methods and cross-device interactions were mostly tried with one user at a time and in a restricted setup and space. Therefore, there is further research that wants to include the environment of everyday life. Furthermore, there are ideas to decrease the immersive experience and make the virtual or augmented world accessible to outsiders.

Considering the current field of multiplayer games in VR, every party requires suitable equipment to join the experience. The Share VR project [8] wants to give people in the room the opportunity to join in a VR multiplayer game. Using mobile displays together with floor projections enabled users without an HMD to experience collaborative gaming. To get a true visualization of the virtual world, this was combined with positional tracking of the outside player to enable the VR experience. The ShaRe [13] project takes a similar approach to AR with HMDs. They were aiming to include the people in the vicinity, without the normally required devices to interact with VR. In ShaRe the augmented content displayed to the person wearing the HMD was displayed onto a flat surface with a projector. The projection includes the surrounding participants. During the project, inaccuracies occurred due to unstable marker tracking. Another problem was conflicting visualizations of the projected content

The discussed interaction models are mostly restricted to be used at home or for games. There is not much research on the usability of cross-device interaction in MRs beyond the discussed environments. However, there is some research that investigates the everyday usability of AR HMD in social and shared places.

Gugenheimer's paper [7] is about the usability of HMD in public and about the social acceptance of people. The focus of current research is mainly in the area of technological barriers. This mainly looks at tracking, input, locomotion, and the field of view. Therefore, it is essential to explore the social components and identify barriers at an early stage to shield users from the outside.

4 Discussion about new Interaction Paradigms in MRs

This section reflects on already mentioned papers and, moreover, presents additional research that fits into the seamless interaction model of this paper with MRs. For new interaction paradigms, it is important that used technologies are intuitively usable and not prone to error, as well as a seamless interaction between multiple devices. Therefore, it is crucial to find out how future interaction with cross-device interaction can look like.

To enable real seamless interaction with and among multiple devices, implementations must be found that work towards minimizing the borders between the digital and the real world. Projects like the previously mentioned reality editor [10] display interfaces directly on the physical object. These interfaces can be captured with a tablet or smartphone and enable direct manipulation with one or more devices. These types of overlay interfaces can also be placed on simple non-digital objects. Benko [3] used this idea with the combination of AR and physical objects of the environment. During this research, AR was used to overlay computer-generated content paired with information on registered physical objects. This enabled users to access and view detailed information about individual items in the environment.

The previously discussed models are passive and only enable the users to view content rather than actively interact with it. For a new innovative interaction model, truly bidirectionally input methods are needed. The inclusion of a smartphone or other smart devices offers more interaction possibilities. Other user studies investigate it in combination with augmented reality [29, 19, 27]. The biggest advantage of including a smartphone is the easy usability, especially the precise touch input [29]. Mb widening the interaction surface, the combination with AR and gesture input was researched. This approach attempts to reduce the boundaries between the digital and the physical world and improves the seamless interaction between devices[27]. Quian [22] investigated the touch-less execution of everyday smartphone tasks. The interaction was attempted with two different depths. The user could interact with the phone without even physically touching it. The participants predominantly preferred the interaction from a distance.

Another active interaction model was researched in the project as Ubii (Ubiquitous interface and interaction) [12], which aims to map the digital world in the real world. In this way, the user's perception should be expanded and not just limited to the screen. Users could manipulate and move files by pinch gestures. Thus, with the help of this interaction, it is possible to copy and drag documents to other devices located in the same room. Files can be dragged to printers, other computers, or projector screens. The documents are visually visible through AR, outside the actual screens.

A similar but more complex approach of interaction between the physical environment and surrounding devices is called Mergereality [27]. The goal is to create a cross-device interaction that is executed by gestures in augmented reality. The concept aims to further minimize the boundaries between the digital and physical environments. The prototype used Oculus Rift VR headset in combination with gesture sensing through Leap Motion. Since the project aims to interact with the real world, the headset was modified to turn it into a see-through AR device. The interaction with the prototype was tested inside a room with surrounding smart devices.

Three specific interaction possibilities were tested. Notes written on a tablet can be dragged by hand from the display to an AR environment. Thus, the advantages of precise input on a tablet were optimally exploited. The availability of the notes in the immediate environment thus makes the information visible for others. Another demo is based on the same basic idea of interaction with gestures. This time, however, the advantage of the 3-dimensional environment was used. So you can seamlessly drag a 3D model from a computer into the AR environment. The third demo made it possible to select a color within the room and use it to tint a smart light. For the implementation, pinch gestures were used, as this gesture is similar to an eyedropper. This setup enabled the user to pick a color inside the room and apply it to the smart light via the described hand gesture. The project aimed to unite the augmented and the real world. In further researches studies, the researchers wanted to add more IoT devices for interaction.

A different approach to including smartphones in interaction with MRs was re-purposing them as a controller. The extended use of smartphones as controllers were already mentioned in 3.1.2. The resulting multi-device interaction offers new interaction opportunities in different areas. This can be used in gaming, or multi-device interaction [2, 19, 15]. This offers a cheap,

efficient, intuitive, and easily accessible 6DoF controller for everyone since this is made possible through an app [2].

Deriving a new interaction paradigm from the discussed researches is difficult since the studies are mostly placed directly in the AR or VR area, including headsets and other devices. Nevertheless, it is possible to spot similar goals and problems throughout the studies. The overall aim was to create a seamless interaction with intuitive input methods, which expand the usual display surface and allow manipulating the direct environment. This consequently aims for a uniform input method, which according to the research, lies in between touch input, gestures, and eye-gaze input. Since the Midas Touch problem [25, 16] is recurring, the input is less error-prone when executed through a smart device like a smartphone or tablet [29, 27]. Smart devices offer more interaction techniques such as orientation, grip, tilting, whacking offer additional, which need to be further explored in practice [29]. The displaying of augmented content is either realized through HMD or projections. AR glasses were more prominent in the research, but this often contradicted other users' inclusion without the required equipment.

5 Research Gaps and Limitations in Multi-device Interaction

Input techniques like gesture control still lack precise control and uniform interaction gestures. Since the goal is natural interaction, hand gestures appear to be the most intuitive way of interaction with MRs. However, emerging problems like unintentional selection due to the lack of gesture knowledge or precision of input are still occurring. Either uniform gestures improve the form of input, or the research needs to lean more towards handheld devices with precise input options. [25, 16, 19] The need for additional input options is either found through external controllers, which are still very expensive, or need to be further researched in the field of re-purposing smartphones as controllers [2]. Including smartphones as controllers were researched, but the device offers more capabilities using the touchscreen, and the haptic feedback can therefore be further explored [29, 19]. Currently, the input method is dependent on the intended usage, this means, more research is needed in the area of independent gesture control, as currently, different input methods offer different advantages for the respective tasks [12, 23, 22]. A uniform model for gesture interaction should therefore be found.

In the current research, precise scans of the room and the interactive objects had to be made in advance [27, 12]. This limits spontaneous or quick interactions with the immediate environment. An idea of displaying information through constant personal projection is given by ambient mobile pervasive display prototype [26]. But a practical solution is still needed and therefore, further research in the are of instant interaction with MRs without a cumbersome previous setup.

MRs and the multi-device interaction investigated must be brought closer to the everyday user and focus on the advantages of cross-device interaction. Many of the VR and AR researches require high quality and, therefore, expensive devices. Not many people possess the needed devices. Currently, VR or AR glasses are only dominantly used for gaming [24]. Therefore the demand to include ubiquitous devices is important to emerge the describe interaction ideas to a broader spectrum of people. With this interaction with MRs can be brought the everyday user. Thus, ideas like Ubii [12] or Bishare [29] should be further explored. As well as expanding the usage of smart devices in the research of Mergereality [27]. To facilitate sharing information and its accessibility in the direct environment and further enhance direct manipulation of the environment.

6 Conclusion and Outlook

The challenge of MRs combined with multi-device interaction embraces a large research field [1, 17]. Several different input and interaction techniques were presented, with their advantages and their disadvantages. Deciding which input and interaction technique to use mainly depends on the area of operation. The combination of different input methods makes it possible to circumvent some of the displayed weaknesses. There are still limitations concerning these interactions, which need further research. There are emerging challenges of using MR in shared social spaces and the overall acceptance of these methods. More research is needed to define a paradigm for seamless multi-device interaction successfully. Especially with the focus of the involvement of smart devices, MRs become more accessible for a new diverse group of users.

References

1 Ronald T Azuma. A survey of augmented reality. *Presence: Teleoperators & Virtual Environments*, 6(4):355–385, 1997.

2 Teo Babic, Harald Reiterer, and Michael Haller. Pocket6: A 6DoF controller based on a simple smartphone application. *SUI 2018 - Proceedings of the Symposium on Spatial User Interaction*, pages 2–10, 2018. doi:10.1145/3267782.3267785.

3 Hrvoje Benko, Eyal Ofek, Feng Zheng, and Andrew D. Wilson. FoveAR: Combining an optically see-through near-eye display with spatial augmented reality projections. *UIST 2015 - Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology*, (November):129–135, 2015. doi:10.1145/2807442.2807493.

4 Christoph W Borst and Arun P Indugula. Realistic virtual grasping. In *IEEE Proceedings. VR 2005. Virtual Reality, 2005.*, pages 91–98. IEEE, 2005.

5 Frederik Brudy, Christian Holz, Roman Rädle, Chi Jui Wu, Steven Houben, Clemens Nylandsted Klokmose, and Nicolai Marquardt. Cross-device taxonomy: Survey, opportunities and challenges of interactions spanning across multiple devices. *Conference on Human Factors in Computing Systems - Proceedings*, pages 1–28, 2019. doi:10.1145/3290605.3300792.

6 Ruei Che Chang, Chi Huan Chiang, Shuo Wen Hsu, Chih Yun Yang, Da Yuan Huang, and Bing Yu Chen. TanGo: Exploring Expressive Tangible Interactions on Head-Mounted Displays. *Proceedings - SUI 2020: ACM Symposium on Spatial User Interaction*, 2020. doi:10.1145/3385959.3418457.

7 Jan Gugenheimer, Mark McGill, Frank Steinicke, Christian Mai, Julie Williamson, and Ken Perlin. Challenges using head-mounted displays in shared and social spaces. *Conference on Human Factors in Computing Systems - Proceedings*, pages 1–8, 2019. doi:10.1145/3290607.3299028.

8 Jan Gugenheimer, Evgeny Stemasov, Julian Frommel, and Enrico Rukzio. ShareVR: Enabling co-located experiences for virtual reality between HMD and Non-HMD users. *Conference on Human Factors in Computing Systems - Proceedings*, 2017-May:4021–4033, 2017. doi:10.1145/3025453.3025683.

9 Keisuke Hattori and Tatsunori Hirai. Inside-out Tracking Controller for VR/AR HMD using Image Recognition with Smartphones. *ACM SIGGRAPH 2020 Posters, SIGGRAPH 2020*, 2020. doi:10.1145/3388770.3407430.

10 Valentin Heun, James Hobin, and Pattie Maes. Reality editor: Programming smarter objects. *UbiComp 2013 Adjunct - Adjunct Publication of the 2013 ACM Conference on Ubiquitous Computing*, pages 307–310, 2013. doi:10.1145/2494091.2494185.

11 Daniel Holz, Sebastian Ullrich, Marc Wolter, and Torsten Kuhlen. Multi-contact grasp interaction for virtual environments. *JVRB-Journal of Virtual Reality and Broadcasting*, 5(7), 2008.

12 Zhanpeng Huang, Weikai Li, and Pan Hui. Ubii: Towards seamless interaction between digital and physical worlds. *MM 2015 - Proceedings of the 2015 ACM Multimedia Conference*, pages 341–350, 2015. doi:10.1145/2733373.2806266.

13 Pascal Jansen, Fabian Fischbach, Jan Gugenheimer, Evgeny Stemasov, Julian Frommel, and Enrico Rukzio. ShARe: Enabling Co-Located Asymmetric Multi-User Interaction for Augmented Reality Head-Mounted Displays. pages 459–471, 2020. doi:10.1145/3379337.3415843.

14 Mohamed Khamis, Carl Oechsner, Florian Alt, and Andreas Bulling. VRPursuits: Interaction in virtual reality using smooth pursuit eye movements. *Proceedings of the Workshop on Advanced Visual Interfaces AVI*, 2018. doi:10.1145/3206505.3206522.

15 Pascal Knierim, Dimitri Hein, Albrecht Schmidt, and Thomas Kosch. The smartphone controller. *i-com*, 20(1):49–61, 2021. URL: https://doi.org/10.1515/icom-2021-0003, doi:doi:10.1515/icom-2021-0003.

16 Chang Liu, Jason Orlosky, and Alexander Plopski. Eye Gaze-based Object Rotation for Head-mounted Displays. *Proceedings - SUI 2020: ACM Symposium on Spatial User Interaction*, 2020. doi:10.1145/3385959.3418444.

17 Paul Milgram and Fumio Kishimo. A taxonomy of mixed reality. *IEICE Transactions on Information and Systems*, 77(12):1321–1329, 1994.

18 Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. Augmented Reality: A class of displays on the reality-virtuality continuum. *Telemanipulator and Telepresence Technologies*, 2351:282–292, 1995. doi:10.1117/12.197321.

19 Peter Mohr, Markus Tatzgern, Tobias Langlotz, Andreas Lang, Dieter Schmalstieg, and Denis Kalkofen. TrackCap: Enabling smartphones for 3D interaction on mobile head-mounted displays. *Conference on Human Factors in Computing Systems - Proceedings*, (April), 2019. doi:10.1145/3290605.3300815.

20 Nelusa Pathmanathan, Michael Becher, Nils Rodrigues, Guido Reina, Thomas Ertl, Daniel Weiskopf, and Michael Sedlmair. Eye vs. Head: Comparing gaze methods for interaction in augmented reality. *Eye Tracking Research and Applications Symposium (ETRA)*, 2020. doi:10.1145/3379156.3391829.

21 Ken Pfeuffer, Jason Alexander, Ming Ki Chong, and Hans Gellersen. Gaze-touch: Combining Gaze with Multi-touch for Interaction on the Same Surface. pages 509–518, 2014. doi:10.1145/2642918.2647397.

22 Jing Qian, David A. Shamma, Daniel Avrahami, and Jacob Biehl. Modality and Depth in Touchless Smartphone Augmented Reality Interactions. *IMX 2020 - Proceedings of the 2020 ACM International Conference on Interactive Media Experiences*, pages 74–81, 2020. doi:10.1145/3391614.3393648.

23 Jie Ren, Yueting Weng, Chengchi Zhou, Chun Yu, and Yuanchun Shi. Understanding window management interactions in AR headset + smartphone interface. *Conference on Human Factors in Computing Systems - Proceedings*, pages 1–8, 2020. doi:10.1145/3334480.3382812.

24 Bernhard E. Riecke, Joseph J. LaViola, and Ernst Kruijff. 3D user interfaces for virtual reality and games: 3D selection, manipulation, and spatial navigation. *ACM SIGGRAPH 2018 Courses, SIGGRAPH 2018*, (August), 2018.

25 Boris B. Velichkovsky, Mikhail A. Rumyantsev, and Mikhail A. Morozov. New solution to the Midas Touch Problem: Identification of visual commands via extraction of focal fixations. *Procedia Computer Science*, 39(C):75–82, 2014. doi:10.1016/j.procs.2014.11.012.

26 Christian Winkler, Julian Seifert, David Dobbelstein, and Enrico Rukzio. Pervasive information through constant personal projection: The ambient mobile pervasive display (AMP-D). *Conference on Human Factors in Computing Systems - Proceedings*, pages 4117–4126, 2014. doi:10.1145/2556288.2557365.

27 Shengzhi Wu. Mergereality : Leveraging Physical Affordances for Multi-device Gestural Interaction in Augmented Reality from a tablet to AR environment. pages 1–4, 2020.

28 Li Zhang, Huidong Bai, Mark Billinghurst, and Weiping He. Is This My Phone? Operating a Physical Smartphone in Virtual Reality. *SIGGRAPH Asia 2020 XR, SA 2020*, (Figure 1):1–2, 2020. doi:10.1145/3415256.3421499.

29 Fengyuan Zhu and Tovi Grossman. BISHARE: Exploring Bidirectional Interactions between Smartphones and Head-Mounted Augmented Reality. *Conference on Human Factors in Computing Systems - Proceedings*, pages 1–14, 2020. doi:10.1145/3313831.3376233.