

DELFI2021

19. Fachtagung Bildungstechnologien der GI

Andreas Lingnau (Hrsg.)

**Proceedings of
DELFI Workshops 2021**

**13.09.2021
Dortmund (Online), Deutschland**

Editor: Dr. Andreas Lingnau, Hochschule Ruhr West, Lützowstraße 5, 46236 Bottrop

ISBN 978-3-946757-03-0

© Hochschule Ruhr West 2021



This book is licensed under a Creative Commons BY-SA 4.0 licence.

Vorwort

Als die Planungen zur DELFI Tagung 2021 begannen, hatten fast alle Lehrenden und Studierenden in Deutschland ihr erstes Semester unter COVID-19 Bedingungen hinter sich. Das *Emergency Remote Teaching*, welches die meisten Hochschulen und Universitäten in dieser Zeit eingeführt hatten, unterscheidet sich sehr von der Vorstellung, die wir normalerweise von einem gut geplanten, strukturierten, technologiegestützten Online-Lernen haben. Man konnte aber schon im Herbst 2020 feststellen, dass entgegen aller Schwierigkeiten und Diskussionen um mögliche Defizite bei der Digitalisierung an Schulen und Hochschulen, sehr viele neue und interessante Ansätze und Lösungen im Bereich *Digitale Lehre und Bildungstechnologien* entstanden sind. Unter diesem Eindruck wurde auch das Tagungsmotto für die DEFLI Tagung 2021 formuliert: „Digitale Lehre zwischen Präsenz- und Online-Betrieb“.

So ist es auch nicht verwunderlich, dass im Rahmen der Workshops auf der DELFI Tagung 2021 besondere Schwerpunkte im Zusammenhang mit den Bedingungen der Lehre während der Pandemie gesetzt wurden. Und auch die Durchführung der Workshops selber unterlag besonderen Bedingungen, nachdem das Programmkomitee die Entscheidung treffen musste, die gesamte Tagung komplett online durchzuführen. Um so erfreulicher ist es, dass es in diesem Jahr insgesamt elf Workshop-Einreichungen gab, von denen 9 akzeptiert wurden. Insgesamt konnten die Workshops so über 120 Teilnehmer*innen verzeichnen.

Der Workshop-Tagungsband zur DELFI 2021 erscheint, anders als in vergangenen Jahren, erst im Nachhinein. Dafür finden Sie hier neben den üblichen Workshop-Beschreibungen und den schriftlichen Beträgen der Workshop-Teilnehmer*innen zusätzlich rückblickende Zusammenfassungen einiger Workshop-Ergebnisse.

Mein Dank gilt sowohl den Organisationskomitees der einzelnen Workshops als auch allen aktiven Teilnehmer*innen, die gemeinsam für interessante und gelungene Online-Workshops im Rahmen der DELFI-Tagung 2021 gesorgt haben.

Bottrop, im September 2021

Andreas Lingnau

Inhaltsverzeichnis

Raphael Zender, Heinrich Söbke und Miriam Mulders	
<i>4. Workshop VR/AR-Learning.....</i>	9
Jana Höning, Marc Schnierle, Marc Hüttenberger, Christopher Polak, Sascha Röck	
<i>Mixed-Reality-in-the-Loop Simulation für Schulungen im Maschinen- und Anlagenbau</i>	11
Lena Florian	
<i>Virtuelle Welten der Geometrie</i>	23
Robin Horst, Ramtin Naraghi-Taghi-Off, Linda Rau, Ralf Dörner	
<i>Remote Emergency Teaching and Virtual Reality Education: A Case-Study Using VR Nuggets in Non-VR Courses.....</i>	35
Andrea Schmitz und Miriam Mulders	
<i>Institutionelle Rahmenbedingungen für den Einsatz von Virtual Reality als Lerntechnologie.....</i>	47
Jule M. Krüger und Daniel Bodemer	
<i>Drei Eigenschaften von Augmented Reality-Erfahrungen und ihre Relevanz beim Lernen</i>	59
Lisa Guth, Heinrich Söbke, Eva Hornecker, and Jörg Londong	
<i>An Augmented Reality-supported Facility Model in Vocational Training</i>	72
Madeleine Ratter, Johannes Klöckner und Mara Kaufeld	
<i>Augmented Reality in der Ausbildung an Hubschrauberkonsolen</i>	84
Sven Judel	
<i>Anreizsysteme für Studierendenpartizipation in Learning Analytics.....</i>	96
Birte Heinemann, René Röpke, Matthias Ehlenz und Ulrik Schroeder	
<i>Nuts & Bolts of Educational Research Projects</i>	102
Julia Hense, Ina Stein und Cordula Torner	
<i>Das didaktische “Support-KIT” des KI-Campus: Planungshilfe für den Einsatz digitaler Lernmaterialien in der Hochschullehre</i>	106

Clara Schumacher, Nathalie Rzepka und Niels Seidel	
<i>Workshop on Learning Analytics</i>	113
Daria Novoseltseva, Kerstin Wagner, Agathe Merceron, Petra Sauer, Nadine Jessel , Florence Sedes	
<i>Investigating the Impact of Outliers on Dropout Prediction in Higher Education</i>	120
Linda Mai, Alina Köchling, Lynn Schmodde and Marius Wehner	
<i>Teacher vs. Algorithm: Learners' Fairness Perception of Learning Analytics Algorithms</i>	130
Philipp Krieter, Michael Viertel and Andreas Breiter	
<i>Supporting Students' Privacy: How Does Learner Control over Their Data Affect the Dataset for Learning Analytics?</i>	146
Ian Wolff, David Broneske and Veit Köppen	
<i>FAIR Research Data Management for Learning Analytics</i>	158
Benjamin Weiher, Niels Seidel, Marc Burchart and Dirk Veiel	
<i>Indicators of group learning in collaborative software development teams</i>	164
Maximilian Karl and Niels Pinkwart	
<i>Using GitHub data to analyse student's teamwork in a programming course to prevent discrimination.....</i>	176
Areej Aldaghamin, Andreas Becker, Regina Brautlacht, Samira Herb- Cless, Natalie Kiesler, Alexander Knoth, Katrin Löhr, Ekaterina Mikhay- lova, Wolfgang Radenbach, Christian Reimann, Anne Sennhenn, Judith Venherm, Nina Wagenknecht und Carsten Wolff	
<i>Digitalisierung und Internationalisierung in der Hochschulbildung</i>	180
Gerd Kortemeyer	
<i>A User-Model for a Next Generation Learning Management System</i>	189
Susan Beudt, Berit Blanc, Marco Kähler, Nicole Vieregg and Rolf Feichtenbeiner	
<i>Workshop "Present and possible future(s) of inclusive EdTech and their applications in education and training"</i>	193

4. Workshop VR/AR-Learning

Raphael Zender¹, Heinrich Söbke² und Miriam Mulders³

Vorwort

Virtual Reality bzw. Virtuelle Realität (VR) ist ein Sammelbegriff für ein breites Spektrum interaktiver Computersimulationen. Diese erfassen einerseits die Aktivitäten von Nutzerinnen und Nutzern gegenüber dem IT-System und stimulieren andererseits durch ihr Feedback verschiedene Sinne der Nutzerinnen und Nutzer, so dass eine subjektive Wahrnehmung von Anwesenheit (Präsenz) in der Simulation entsteht. Eine mit VR assoziierte Form stellt die Augmented Reality (AR) dar, bei der der Fokus der Nutzerinnen und Nutzer zwar in der physischen Realität verbleibt, diese jedoch um virtuelle Artefakte und Informationen erweitert wird. Die Kombination der VR/AR-Technologien in so genannten Mixed Reality Anwendungen zur Mensch-Maschine-Interaktion eröffnet eine Vielzahl an Vorteilen für die mediengestützte Aus- und Weiterbildung.

Die technologischen und wirtschaftlichen Entwicklungen der vergangenen Jahre haben zu einem erheblichen Zuwachs an VR/AR-Anwendungsbereichen, Entwicklungsprozessen sowie Nutzungsinteresse geführt. Auf Seiten der Anwendungen resultierte dies in einer beinahe unüberschaubaren Masse voneinander isolierter VR/AR-Erfahrungen. Es fehlen nach wie vor wirksame und systematische Integrationskonzepte in bestehende Systemarchitekturen, etablierte Bildungsprozesse sowie andere mediale Konzepte, aber auch anerkannte Anforderungen zur Gestaltung virtueller und augmentierter Lernwelten und aussagekräftige Studien zu Lerneffekten von VR/AR-Lernanwendungen. Auch Fragen der Organisation derartiger Lehr-/Lernwerkzeuge im Rahmen institutioneller Lehr-/Lernprozesse sind bisher weitestgehend ungeklärt. Diese Defizite sind insbesondere bedenklich, wenn man sich vor Augen führt, dass VR/AR-Technologien sich selbst zunehmend einer massetauglichen Reife nähern.

Der Workshop thematisiert in seinem vierten Jahr unter anderem diese Herausforderungen. Erbeten wurden sowohl wissenschaftlich fundierte Beiträge, Studierendenbeiträge als auch Best-Practice-Beispiele und Fallstudien zu Themenbereichen des Lernens mit VR/AR-Technologien. Veranstaltet wird der Workshop vom gleichnamigen Arbeitskreis VR/AR-Learning⁴ der Gesellschaft für Informatik (GI).

Für den Workshop gab es insgesamt sieben Einreichungen. Jede wurde von mindestens

¹ Universität Potsdam, Komplexe Multimediale Anwendungsarchitekturen,
August-Bebel-Straße 89, 14482 Potsdam, raphael.zender@uni-potsdam.de

² Bauhaus-Universität Weimar, Bauhaus-Institut für zukunftsweisende Infrastruktursysteme (b.is),
Goetheplatz 7/8, 99421 Weimar, heinrich.soebe@uni-weimar.de

³ Universität Duisburg-Essen, Learning Lab, Universitätsstr. 2, 45141 Essen, miriam.mulders@uni-due.de

⁴ <https://ak-vrarl.gi.de>

drei Mitgliedern des Programmkomitees begutachtet, mit dem Ergebnis einer ausnahmslos sehr hohen fachlichen Qualität. Nach einer Überarbeitung konnten daher alle sieben Beiträge angenommen und veröffentlicht werden. Sie ließen durchgängig hochrelevante Vorträge sowie angeregte Diskussionen erwarten.

Zudem wurden auf dem Workshop die Preisträger des AVRiL-Wettbewerbs 2021 zu gelungenen VR/AR-Lernszenarien ausgezeichnet. Der Siegerbeitrag wurde im Rahmen des Workshops vorgestellt.

Im Weiteren möchten wir den Mitgliedern des Programmkomitees für ihr großartiges Engagement für die Qualitätssicherung des Workshops in Form der Begutachtung der Beiträge danken!

Programmkomitee

- Maria Bannert (Technische Universität München)
- Josef Buchner (Universität Duisburg-Essen)
- Mario Donick (vFlyteAir Simulations)
- Ralf Dörner (Hochschule RheinMain)
- Torsten Fell (Institute for Immersive Learning)
- Lena Florian (Universität Potsdam)
- Paul Grimm (Hochschule Fulda)
- Sebastian Habig (Universität Duisburg-Essen)
- Wibke Holtij (Technische Hochschule Mittelhessen)
- Andreas Kohn (Materna TMT GmbH)
- Felix Kretschmer (Technische Universität Berlin)
- Rolf Kruse (Fachhochschule Erfurt)
- Carsten Lecon (Hochschule Aalen)
- Anja Richert (Technische Hochschule Köln)
- Andrea Schmitz (ZWH-Dienstleistungs GmbH)
- Hartmut Seichter (Hochschule Schmalkalden)
- Pia Spangenberger (Technische Universität Berlin)
- Sven Strickroth (Ludwig-Maximilians-Universität München)
- Johannes Tümler (Hochschule Anhalt)
- Markus von der Heyde (vdH-IT)
- Matthias Weise (Universität Potsdam)

Mixed-Reality-in-the-Loop Simulation für Schulungen im Maschinen- und Anlagenbau

Jana Höning  ¹, Marc Schnierle ², Marc Hüttenberger ³, Christopher Polak ⁴, Sascha Röck ⁵

Abstract: In diesem Beitrag wird die Mixed-Reality-in-the-Loop Simulation (MRiLS) zur Schulung technischer Fachkräfte im Maschinen- und Anlagenbau vorgestellt. Die MRiLS kombiniert das im Entwicklungsprozess für die Steuerungsentwicklung eingesetzte Konzept der X-in-the-Loop Simulation (XiLS) mit Visualisierungs- und Interaktionsmethoden der Mixed Reality und ermöglicht damit eine vollständige Integration des Menschen in die Simulationskreisläufe. Zur Kopplung der XiLS mit Mixed Reality (MR) Umgebungen wird eine Digital-Twin-as-a-Service Plattform präsentiert. Die Plattform ermöglicht neben der flexiblen Bereitstellung der Digitalen Zwillinge auf verschiedenen Mixed Reality Endgeräten auch den Einsatz unterschiedlicher Ein- und Ausgabemöglichkeiten sowie die flexible Ankopplung von industrieller Steuerungstechnik. Dabei können mehrere Nutzer auf Basis einer Interaktionsabstraktion standortunabhängig mit einem gemeinsamen Digitalen Zwilling intuitiv und multimodal interagieren. Die Tragfähigkeit der vorgestellten Lösung wird anhand verschiedener Anwendungen belegt.

Keywords: Digitaler Zwilling, X-in-the-Loop Simulation, Mixed Reality, Augmented Reality, Virtual Reality, Maschinen- und Anlagenbau, Lehr-/Lernszenarien, Interaktion, Multiuser

1 Motivation

Im Maschinen- und Anlagenbau ist in den letzten Jahren ein steigender Automationsgrad sowie eine zunehmende Komplexität und Vernetzung der eingesetzten Maschinen und Anlagen zu beobachten (vgl. [VDI17]). Die Anforderungen an die Aus- und Weiterbildung der technischen Fachkräfte zur Bedienung oder Wartung der Anlagen steigen mit dem wachsenden Automatisierungsgrad und der zunehmenden Komplexität gleichermaßen an. Neben den inhaltlichen Anforderungen an die Schulung sind bei den Lernenden auch die Heterogenität des Wissenstandes sowie die verschiedenen Lerntypen zu berücksichtigen.

¹ Virtual Automation Lab (VAL), Fakultät Maschinen und Systeme, Hochschule Esslingen, Kanalstraße 33,

73728 Esslingen, jana.hoenig@hs-esslingen.de  <https://orcid.org/0000-0003-1564-8742>

² Virtual Automation Lab (VAL), Fakultät Maschinen und Systeme, Hochschule Esslingen, Kanalstraße 33, 73728 Esslingen, marc.schnierle@hs-esslingen.de

³ Virtual Automation Lab (VAL), Fakultät Maschinen und Systeme, Hochschule Esslingen, Kanalstraße 33, 73728 Esslingen, mahuette@hs-esslingen.de

⁴ Virtual Automation Lab (VAL), Fakultät Maschinen und Systeme, Hochschule Esslingen, Kanalstraße 33, 73728 Esslingen, chpolak@hs-esslingen.de

⁵ Virtual Automation Lab (VAL), Fakultät Maschinen und Systeme, Hochschule Esslingen, Kanalstraße 33, 73728 Esslingen, sascha.roeck@hs-esslingen.de

1.1 Aktuelle Formen der Wissensvermittlung im Maschinen- und Anlagenbau

Derzeit kommen bei der Schulung technischer Fachkräfte im Maschinen- und Anlagenbau verschiedene Formen der Wissensvermittlung zum Einsatz, welche entsprechend der Ausprägung des eingesetzten Lerngegenstandes kategorisiert werden können (siehe Abbildung 1). Es kann zwischen der Wissensvermittlung unter Einsatz realer Lerngegenstände (z.B. physisch existierende Maschine) und virtueller Lerngegenstände (z.B. Simulationsmodell) unterschieden werden.

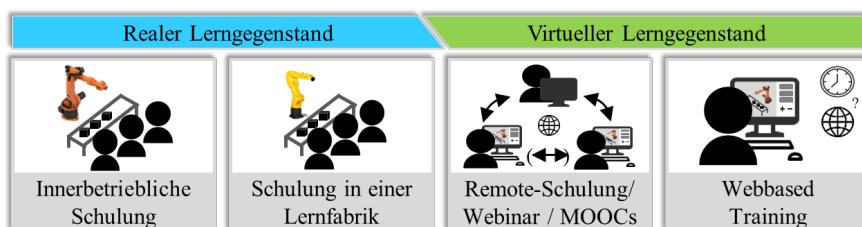


Abbildung 1: Aktuell eingesetzte Formen der Wissensvermittlung im Maschinen- und Anlagenbau

Im Folgenden werden die beiden Ausprägungsformen detaillierter analysiert:

- **Lernen mit realem Lerngegenstand:** Bei der Maschinen-Schulung unter Verwendung eines realen Lerngegenstandes kommt ein physisch existierender Lerngegenstand, wie beispielsweise die später im Einsatz befindliche oder eine artverwandte reale Anlage, zum Einsatz. Der Lernende trainiert das Verhalten sowie die Bedienung an der realen Maschine und kann dieses Wissen direkt auf die späteren Aufgaben transferieren. Die Schulung kann sowohl als innerbetriebliche Schulung vor Ort als auch in einer Lernfabrik stattfinden. An der später im Einsatz befindlichen Anlage kann es aufgrund der fehlenden Parallelisierungsmöglichkeit von Schulung und Betrieb zu einer Produktionsunterbrechung kommen. Im Gegensatz hierzu entsteht in der Lernfabrik durch die Entkopplung von Schulung und Produktion kein Produktionsausfall, jedoch muss eine zum Betrieb parallele Infrastruktur aufgebaut werden. Daher finden die Schulungen meist an externen Standorten statt, wobei der Lernende häufig nicht mit exakt der gleichen Maschine wie im späteren Betrieb lernen kann. Zudem sind an der realen Anlage das Trainieren des Verhaltens in Gefahrensituationen sowie explorative Lernformen aufgrund der Gefährdung von Mensch und Maschine nicht bzw. nur eingeschränkt möglich. Darüber hinaus ist der reale Lerngegenstand an den Standort gebunden und nicht flexibel skalierbar, wodurch nicht jedem Lernenden ein eigener Lerngegenstand zur Verfügung steht.
- **Lernen mit virtuellem Lerngegenstand:** Ergänzend zum Lernen an der realen Anlage kommen Formen der Wissensvermittlung mit virtuellem Lerngegenstand, wie bspw. Remote-Schulungen, Webinare, Massive Open Online Courses (MOOCs) und Webbased Trainings, zum Einsatz. Als Lerngegenstand werden Simulationen oder video- bzw. textbasierte Inhalte zur Verfügung gestellt. Die

Lerninhalte werden auf zweidimensionalen Computer-Monitoren visualisiert und der Lernende ist bei der Interaktion auf Eingaben mit Maus und Tastatur beschränkt. Bei diesen Schulungspraktiken wird ein ortsunabhängiges und teilweise ein zeitunabhängiges Lernen sowie eine flexible Skalierbarkeit des Lerngegenstandes ermöglicht. Allerdings können nicht alle Schulungsinhalte handlungsorientiert abgebildet werden.

Durch die Rahmenbedingungen der Corona-Pandemie (2020/2021) ist eine Reduktion von Präsenzveranstaltungen festzustellen, wodurch bei der Wissensvermittlung vermehrt eine Verschiebung zum Lernen mit virtuellen Lerngegenständen zu verzeichnen ist. Während dies bei einigen Schulungsinhalten, wie zum Beispiel der Programmierung, problemlos möglich ist, können andere Schulungsinhalte, wie bspw. bei der Maschinen-Bedienung, nur unzureichend abgebildet werden, wodurch nicht dieselbe Schulungstiefe wie an der realen Anlage erreicht werden kann.

1.2 Mixed Reality Methoden zur Wissensvermittlung

Eine Erweiterung der bislang eingesetzten Formen der Wissensvermittlung im Maschinen- und Anlagenbau stellt der Einsatz von modernen Visualisierungs- und Interaktionsmethoden der Mixed Reality (MR) dar. Bereits 1994 beschrieb Milgram in seinem Realitäts-Virtualitäts-Kontinuum die Mixed Reality als stufenlose Kombination realer und virtueller Komponenten (vgl. [Mi94]). Die Mixed Reality umfasst neben den beiden Extremen, der realen und der rein computergenerierten Umgebungen, auch die Kombinationen aus realen und virtuellen Komponenten in einer gemeinsamen Umgebung (siehe Abbildung 2). Entsprechend dem Verhältnis von virtuellen und realen Komponenten wird dabei zwischen der Augmented Reality (AR) und der Augmented Virtuality (AV) unterschieden.

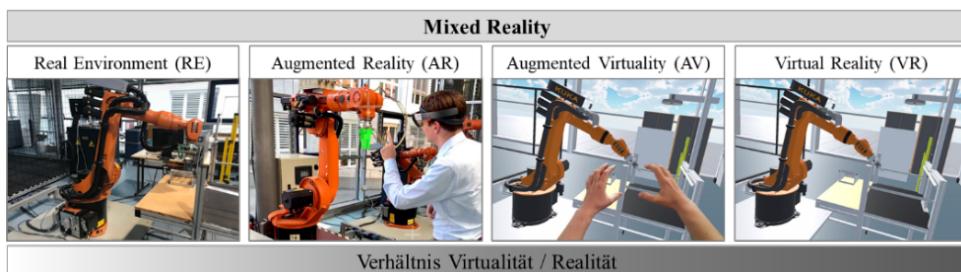


Abbildung 2: Realitäts-Virtualitäts-Kontinuum der Mixed Reality

Der Einsatz moderner Visualisierungsmethoden der Mixed Reality ermöglicht durch die vom Benutzer abhängige (egozentrische) Visualisierung sowie die intuitive multimodale Interaktion neue Möglichkeiten in der Wissensvermittlung bei der Schulung technischer Fachkräfte im Maschinen- und Anlagenbau. Zwar existieren Arbeiten, die den Einsatz von Mixed Reality Technologien in diesem Feld thematisieren (z.B. [Fe17], [Sh19]), jedoch integrieren diese Umsetzungen zumeist keine reale Steuerungstechnik sowie multimodale

Interaktionsmöglichkeiten, wodurch die mögliche Modell- und Schulungstiefe eingeschränkt ist.

2 Mixed-Reality-in-the-Loop Simulation (MRiLS)

Als Lösung der zuvor beschriebenen Defizite wird in diesem Beitrag der Einsatz der Mixed-Reality-in-the-Loop Simulation (MRiLS) zur Schulung technischer Fachkräfte im Maschinen- und Anlagenbau vorgeschlagen.

Im Entwicklungs- und Inbetriebnahmeprozess des Maschinen- und Anlagenbaus kommen derzeit drei verschiedene Ausprägungen der X-in-the-Loop Simulation (XiLS) zum Einsatz: Model-in-the-Loop Simulation (MiLS), Software-in-the-Loop Simulation (SiLS) und Hardware-in-the-Loop Simulation (HiLS) (vgl. [VD16]). Diese Testkonfigurationen sind in ihrer konzeptionellen Architektur identisch und integrieren je eine Ausprägung der Steuerung und des Digitalen Zwilling einer Anlage bzw. einer Anlagenkomponente in einen Simulationskreislauf. Sie unterscheiden sich jedoch in der Ausprägungsstufe der Steuerung sowie des Digitalen Zwilling. Der Digitale Zwilling bildet definierte Eigenschaften der realen Anlage ab und setzt sich zum Beispiel aus dem Geometriemodell, dem Verhaltensmodell sowie weiteren Metainformationen (z.B. Datenmodell) zusammen. Die Mixed-Reality-in-the-Loop Simulation (MRiLS) erweitert die bislang eingesetzten Ausprägungen der XiLS um moderne Visualisierungs- und Interaktionsmethoden der Mixed Reality (siehe Abbildung 3).

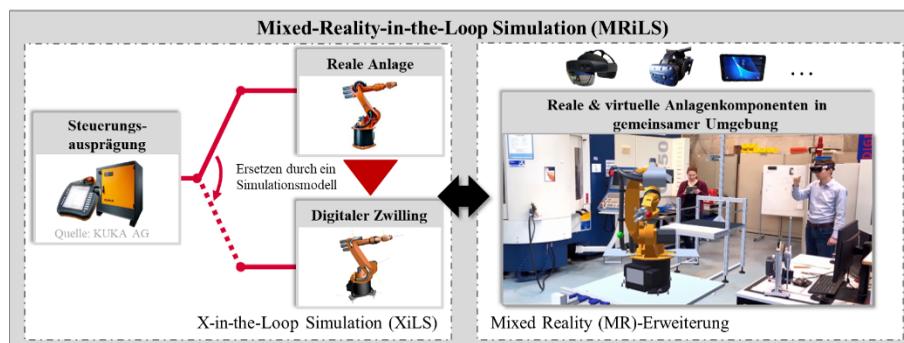


Abbildung 3: Aufbau der Mixed-Reality-in-the-Loop Simulation (MRiLS)

Die Kopplung der egozentrischen Visualisierung und der intuitiven multimodalen (mehrere Sinne ansprechende) Mensch-Modell-Interaktion der Mixed Reality mit den aus dem Engineering-Prozess vorhandenen XiLS-Modellen bildet die Grundlage für eine stufenlose Kombination der Virtualität und Realität. Durch das realitätsgetreue Abbild des Anlagenverhaltens im Digitalen Zwilling kann bei der Schulung mittels MRiLS ein hoher Detaillierungsgrad erreicht werden, wodurch die Lernenden anhand des späteren Anlagenverhaltens geschult werden können.

Zur Erreichung der vollständigen Integration des Menschen in den Simulationskreislauf und somit als Grundlage für die Schulung technischer Fachkräfte mittels MRiLS sind verschiedene Anforderungen an eine MRiLS zu stellen (siehe auch Abb. 4):

- **Egozentrische Visualisierung:** Die Grundlage der vollständigen Integration des Menschen in den Simulationskreislauf bildet eine egozentrische Visualisierung des Digitalen Zwillings.
- **Kopplung realer und virtueller Komponenten:** Aufbauend auf der egozentrischen Visualisierung bedarf es einer vollständigen Kopplung zwischen realen und virtuellen Komponenten in einer gemeinsamen Umgebung.
- **Intuitive multimodale Mensch-Modell-Interaktion:** Sowohl die Manipulation der virtuellen Komponenten als auch die multimodale Rückmeldung sollte möglichst intuitiv sein und die für die MRiLS relevanten Sinne des Menschen durch synthetisch erzeugte Stimuli ansprechen (z.B. visuell, auditiv, haptisch).
- **Multiuser-Kollaboration:** Für die kollaborative Schulung müssen mehrere Nutzende dasselbe MRiL-Simulationsmodell in einer gemeinsamen Umgebung betrachten und mit diesem interagieren können.



Abbildung 4: Anforderungen an die Mixed-Reality-in-the-Loop Simulation (MRiLS)

3 Digital-Twin-as-a-Service

Die Erfüllung der zuvor beschriebenen Anforderungen bedingt die flexible Kopplung der X-in-the-Loop Simulationsmodelle mit Visualisierungsmethoden der Mixed Reality und entsprechenden Interaktionsmöglichkeiten sowie der industriellen Steuerungstechnik. Im Folgenden wird ein Plattformansatz vorgestellt, der diese flexible Kopplung durch Integration eines Digital-Twin-as-a-Service Konzepts ermöglicht.

3.1 Middleware

Die Basis des Digital-Twin-as-a-Service Konzepts bildet eine Middleware, die auf einer Client-Server-Architektur basiert und eine flexible Verknüpfung von Mixed Reality Endgeräten als Clients sowie Datenquellen und -senken als Assets mit dem Digitalen

Zwilling ermöglicht. Reale Assets umfassen zum Beispiel industrielle Steuerungssysteme, während Offline-Programmiersysteme, Simulations-Tools oder Cloud-Dienste virtuelle Assets darstellen. Die Middleware selbst beinhaltet die Datenbasis des Digitalen Zwilling, ein Autorentool zur Erstellung des Digitalen Zwilling, zusätzliche Mehrwertdienste wie bspw. eine Multiuser-Umgebung sowie die Abstraktion der Client- und der Asset-Funktionalitäten (siehe Abbildung 5).

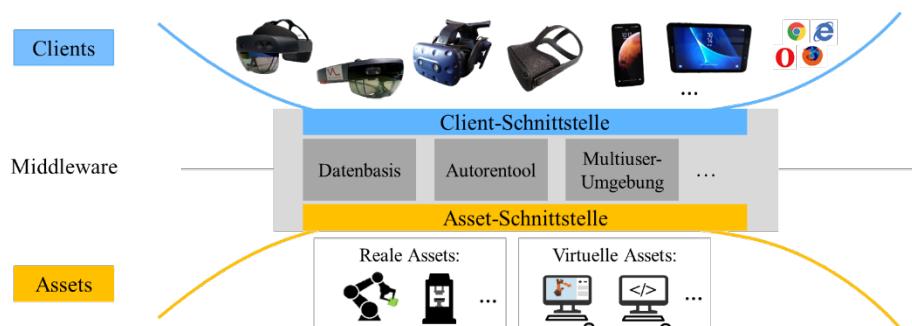


Abbildung 5: Konzept der Middleware

Die angebundenen Clients verbinden sich über die Client-Schnittstelle an die Middleware und können den Digitalen Zwilling endgeräteunabhängig instanzieren. Über die Asset-Schnittstelle binden sich die realen und virtuellen Assets an die Middleware an, wodurch die Erstellung eines realdatengetriebenen Digitalen Zwilling ermöglicht wird, welcher anschließend endgeräteunabhängig für die Nutzenden bereitgestellt wird. Die Nutzenden können den Digitalen Zwilling und die angebundenen Assets über die entsprechende vielfältige Gerätetechnik intuitiv und multimodal manipulieren.

3.2 Interaktionsabstraktion

Um die Vielzahl der verschiedenen multimodalen Ein- und Ausgabemöglichkeiten sowie die angebundenen Assets auf Basis eines gemeinsamen Digitalen Zwilling ohne wiederkehrende Adoptions- und Modellierungsaufwände koppeln zu können, ist eine Abstraktion der Interaktion notwendig. Die Abstraktion der Interaktion erfolgt entsprechend [Sc19] und ist in Abbildung 6 dargestellt.

Der Digitale Zwilling wird unabhängig von der eingesetzten Gerätetechnik sowie den angebundenen Mehrwertdiensten implementiert. Die Eingaben wie beispielsweise Nutzer-Eingaben zur Manipulation, Realdaten von angebundenen Assets sowie Eingaben von weiteren Mehrwertdiensten (z.B. Cloud-Diensten) werden abstrahiert an den Digitalen Zwilling übermittelt. Der Digitale Zwilling reagiert auf diese Eingaben entsprechend des implementierten geräteunabhängigen Verhaltens- und Interaktionsmodells und generiert geräteunabhängige Ausgaben. Die Ausgaben werden in einer Abstraktionsschicht in eine der Gerätetechnik entsprechenden Ausgabegerätefunktionalität überführt. Die Ausgabemöglichkeiten können verschiedene Wahrnehmungen des Menschen, wie

beispielsweise die visuelle, haptische oder auditive Wahrnehmung durch synthetisch erzeugte Reize ansprechen, sodass der Nutzende eine umfassende multimodale Illusion der virtuellen Komponenten erfahren kann. Der Nutzende kann über Eingabegeräte den Digitalen Zwilling manipulieren und somit dessen Verhalten aktiv beeinflussen. Die Manipulation kann über verschiedene Kommunikationskanäle wie bspw. durch Sprache oder über Gesten erfolgen. Die Eingaben werden in der Abstraktionsschicht in Events gewandelt und abstrahiert an den Digitalen Zwilling weitergeleitet. Über die multimodalen Ausgabegeräte erhält der Nutzende wiederum eine Reaktion des Digitalen Zwilling auf seine Eingabe. Die Aus- und Eingaben können entweder in eine Interaktionsschleife integriert oder solitär angebunden sein, wobei die Interaktionsschleife eine Ein- und Ausgabe mit demselben angebundenen Asset bzw. mit der verwendeten Gerätetechnik kombiniert. Die Abstraktion und Anbindung weiterer Ein- und Ausgaben erfolgt entsprechend des vorgestellten Beispiels.

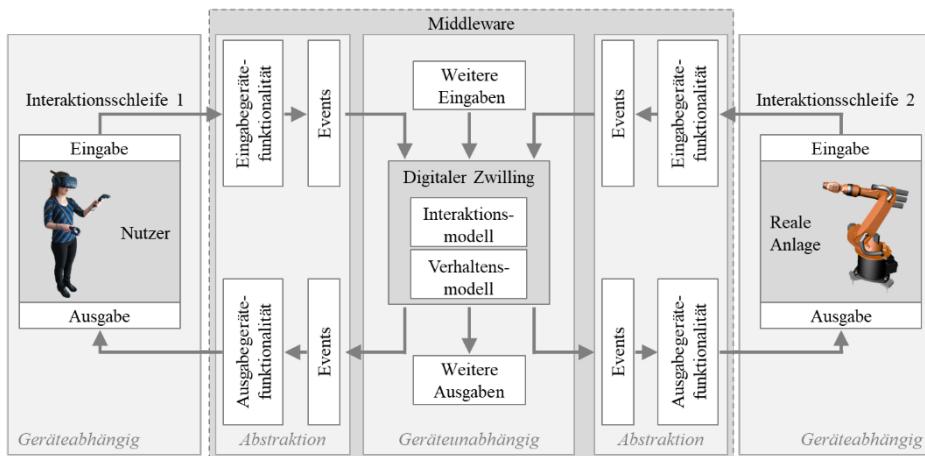


Abbildung 6: Abstraktion der Interaktion in Anlehnung an [Sc19]

Die Interaktionsabstraktion ermöglicht den Einsatz verschiedener multimodaler Ein- und Ausgaben auf Basis eines gemeinsamen Digitalen Zwilling, die flexible Kopplung der Assets sowie weiterer Mehrwertdienste mit dem Digitalen Zwilling. Der Digitale Zwilling muss somit nicht an die eingesetzte Gerätetechnik bzw. angebundenen Mehrwertdienste angepasst werden, sondern kann unabhängig von diesen implementiert werden.

3.3 Speech-Service

In den letzten Jahren hat das Thema Sprachverarbeitung, vor allem durch die Fortschritte im Bereich der künstlichen Intelligenz an großer Bedeutung gewonnen. Die Zuverlässigkeit der Sprachverarbeitung hängt jedoch stark von den dafür notwendigen KI-Technologien ab. Die Entwicklung einer eigenen Sprachverarbeitung mit Methoden des

maschinellen Lernens erfordert KI-Experten und große Mengen an Daten, was ein Hindernis für den flexiblen Einsatz dieser Technologien darstellt.

Aufbauend auf der zuvor vorgestellten Interaktionsabstraktion ist ein Speech-Service in die Middleware zur verbalen Kommunikation mit dem Digitalen Zwilling sowie den angebundenen Assets implementiert. Der Speech-Service umfasst einen webbasierten und frei konfigurierbaren Service, der als Softwareschnittstelle das Knowhow von mehreren großen Cloudanbietern bündelt und somit unabhängig von einem einzelnen Cloudanbieter ist. Die Sprachbefehle werden über ein Mikrofon aufgenommen, in der Middleware gespeichert, anonymisiert an den ausgewählten Cloudanbieter gesendet und dort entsprechend ausgewertet. Das Ergebnis des Cloudanbieters wird in der Middleware weiterverarbeitet und löst eine entsprechende Aktion im Digitalen Zwilling aus. Gleichermassen erfolgt die Sprachausgabe von Texten. Der Speech-Service ermöglicht dem Nutzenden somit eine intuitive Interaktion mit den virtuellen Komponenten sowie mit den an die Middleware angebundenen Assets. Abbildung 7 zeigt einen beispielhaften Einsatz des Speech-Service für die verbale Kommunikation mit einem Industrieroboter.

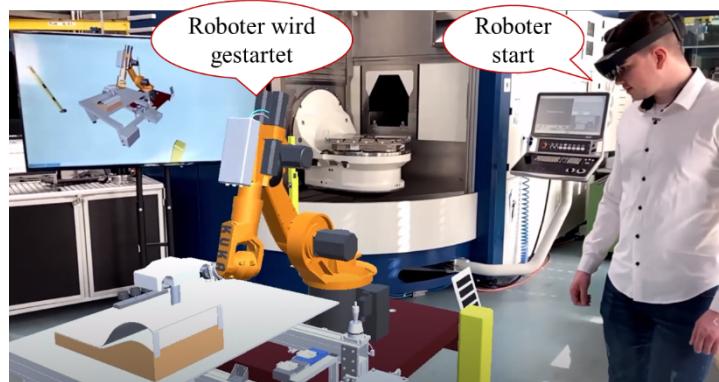


Abbildung 7: Beispielhafter Einsatz des Speech-Service in einer MRiLS eines Industrieroboters

Der Nutzende löst mittels eines Sprachbefehls ("Roboter Start") eine Interaktion aus, woraufhin die abstrahierte Interaktion verarbeitet und eine Sprachausgabe ("Roboter wird gestartet") für den Nutzenden als Bestätigung ausgegeben wird. Die flexible Integration von Sprachein- und ausgabe in die Middleware ermöglicht vielfältige Potentiale bei der Schulung mittels MRiLS. So können beispielsweise mit der Spracheingabe virtuelle Komponenten manipuliert, Funktionen gesteuert oder auch Informationen abgefragt werden. Über die Sprachausgabe können dem Lernenden zudem Hilfestellungen gegeben werden, um ihn bei seinem individuellen Lernprozess zu unterstützen. Die verbale Kommunikation kann jedoch nicht nur für die Interaktion zwischen Nutzenden und Maschinen bzw. Digitalem Zwilling eingesetzt werden, sondern auch für die Kommunikation zwischen mehreren Nutzenden (siehe Abschnitt 3.4).

3.4 Multiuser-Kollaboration

Ein wichtiger Bestandteil von Schulungen ist das kollaborative Lernen. Der Austausch zwischen Lernenden über Sprache, Gestik und Vorgehen an der Maschine gehören zu den Grundlagen eines erfolgreichen Schulungskonzepts in der MRiLS. Abbildung 8 zeigt einen beispielhaften Einsatz der MRiLS für die Multiuser-Kollaboration.



Abbildung 8: Multiuser-Kollaboration in Virtual Reality (links) und Augmented Reality (rechts)

Durch den Aufbau der Middleware als Client-Server-Modell und der Interaktionsabstraktion ist es möglich, mehrere Endgeräte der Mixed Reality an dieselbe Instanz einer Mixed-Reality-in-the-Loop Simulation zu koppeln und damit mehrere Nutzende in einer gemeinsamen Mixed Reality Umgebung zusammenzuführen. Je nach Endgerät muss dabei die bisherige Interaktionsschleife mit persönlichen Clientdaten ergänzt werden, um die Kommunikation zwischen den Teilnehmenden zu gewährleisten. Zu diesen Daten gehören beispielsweise Posen des Endgeräts (z.B. VR-Brille) und der Eingabegeräte (z.B. Hände), Interaktionseingaben, Sprache, personalisierte Avatare und der Gerätetyp. Die Multiuser-Umgebung ermöglicht den Nutzenden das gemeinsame Betrachten und Interagieren mit dem Digitalen Zwilling sowohl von Tablets und Smartphones als auch von VR- und AR-Brillen. Durch die Verwendung von Cloud-Technologien können sich die Nutzenden örtlich getrennt aufhalten und sich über das Internet in einer gemeinsamen Multiuser-Umgebung treffen und über personalisierte Avatare identifizieren und kommunizieren. Der Schulungsleitende kann somit beispielsweise eine Funktion vorführen, welche die Lernenden im Anschluss an dem eigenen Digitalen Zwilling die Funktion anwenden und explorativ erkunden können. Durch das gemeinsame Interagieren mit dem Digitalen Zwilling werden Schulungen im virtuellen Raum noch realitätsnäher und ermöglichen eine höhere Schulungstiefe.

4 Potentiale und Anwendungen der MRiLS

Die Kopplung der XiLS-Modelle mit modernen Visualisierungsmethoden der Mixed Reality lässt die Weiterverwendung der aus dem Engineering und der Inbetriebnahme vorhandenen XiLS-Modelle zu, so dass das Anlagenverhalten für die Schulung mittels MRiLS ohne umfangreiche Modellierungsschritte abgeleitet werden kann. Die MRiLS

kombiniert dabei ein ortsunabhängiges Lernen mit einer hohen Schulungstiefe. Die egozentrische Visualisierung und die multimodale intuitive Interaktion ermöglichen zudem neue Interaktionsformen mit der Maschine wie bspw. die maßstabsgerechte Visualisierung mit realitätsnaher Interaktion. Der Lerngegenstand der MRiLS kann flexibel skaliert und bereitgestellt werden, so dass jeder Lernende Zugang zu einem eigenen Lerngegenstand erhält. Darüber hinaus kann in dem XiLS-Modell der Schulungsinhalt reproduziert werden, so dass die Lernenden mit der MRiLS das Verhalten in Gefahren-Situationen ohne Gefährdung von Mensch und Maschine in einer sicheren Umgebung trainieren können. Die Entkopplung von Schulung und realer Anlage ermöglicht zudem eine parallelisierte Schulung ohne Produktionsunterbrechungen.

Die vorgestellten Methoden werden am Virtual Automation Lab (VAL) der Hochschule Esslingen in eine Digital-Twin-as-a-Service-Plattform (DTaaSP) (technische Details siehe [Sc18]) überführt und verschiedene Anwendungsszenarien für die sekundäre Bildung (z.B. Berufsschule), tertiäre Bildung (z.B. Hochschule) und quartäre Bildung (z.B. Weiterbildung) umgesetzt. Beispielhaft wurde ein Digitaler Zwilling der Smart Factory Testumgebung der Hochschule Esslingen erstellt, der die Maschinen und Anlagen, das Gebäude sowie die Infrastruktur der drei Maschinenhallen umfasst (siehe **Fehler! V erweisquelle konnte nicht gefunden werden.**).



Abbildung 9: Smart Factory Testumgebung an der Hochschule Esslingen

Der Digitale Zwilling kann auf Basis der zuvor vorgestellten Technologien ohne Adoptionsaufwand auf verschiedenen Endgeräten (z.B. AR-/VR-Brillen, Tablets) unter Einsatz multimodaler Ein- und Ausgabemöglichkeiten (z.B. Controller, Gesten, Sprache) verwendet werden. Die Daten des Digitalen Zwillings stehen zudem verschiedenen Mehrwertdiensten zur weiteren Verarbeitung zur Verfügung (z.B. Daten-Monitoring). Die Lernenden können mit Hilfe der Smart Factory Testumgebung das Anlagenverhalten gefahrlos erlernen und eigenständig explorativ erkunden. In Projektarbeiten haben die Studierenden die Möglichkeit, einzelne Funktionalitäten des Digitalen Zwillings zu erweitern und erhalten dadurch einen niederschwelligen Zugang zu MR Technologien.

Durch die Abstraktion können die Lernenden ihre Ergebnisse (z.B. Programmierung von Interaktionen) direkt auf den verschiedenen Endgeräten testen. Die Erkenntnisse aus der Forschung fließen in die Plattform ein und somit zurück in die Lehre, sodass sich ein iterativer Optimierungsprozess einstellen kann. Zudem können verschiedene Tools (z.B. Simulationsprogramme, Offline-Programmiersysteme) sowie reale Maschinendaten flexibel angebunden werden, sodass die Lernenden die bislang zweidimensional projizierten Ergebnisse dreidimensional und immersiv erleben können.

Abbildung 10 zeigt eine Übersicht weiterer ausgewählter Lernszenarien, die unter Verwendung der DTaaSP für die sekundäre und tertiäre sowie die quartäre Bildung als Mixed-Reality-in-the-Loop Simulation erfolgreich umgesetzt werden konnten.

Lern-szenario	Automatisierungs-system	Roboter-Programmierung	Regelungstechnik	Konstruktion und Kinematisierung	Umformtechnik	Robotik
Einsatz	Industrie	Industrie	Hochschule	Hochschule	Hochschule	Berufsschule
Impres-sion						
Lern-ziele	<ul style="list-style-type: none"> ▪ Inbetriebnahme Automatisierungssystem ▪ Programmierung ▪ Bedienung ▪ Verhalten in Störsituationen 	<ul style="list-style-type: none"> ▪ Programmierung mit Offline-Programmier-System ▪ Prozess- und Steuerungs-optimierung ▪ Roboterauswahl 	<ul style="list-style-type: none"> ▪ Grundlagen Regelungstechnik ▪ Regler-Entwurf und Simulation in Matlab/Simulink ▪ Parameter-identifikation 	<ul style="list-style-type: none"> ▪ Auslegung komplexer Kinematiken ▪ Vorwärts-/ Rückwärts-transformation ▪ Numerische Lösungsverfahren 	<ul style="list-style-type: none"> ▪ Grundlagen Umformtechnik ▪ Pressen-Kinematik ▪ Parameter-Auswirkungen auf Umformprozess ▪ Bedienung 	<ul style="list-style-type: none"> ▪ Grundlagen Robotik ▪ Roboterverhalten (z.B. Singularitäten) ▪ Programmierung

Abbildung 10: Überblick verschiedener MRiLS Lernszenarien unter Verwendung der DTaaSP

Eine detaillierte Betrachtung der Lernziele und des didaktischen Designs des Lernszenarios "Konstruktion und Kinematisierung" können [Hö20] entnommen werden. Details des Lernszenarios "Automatisierungssystem" sind in [Hö21] zu finden. Weiterführende Evaluationen hinsichtlich der Lerneffekte werden im weiteren Projektverlauf durchgeführt.

5 Ausblick

Im Rahmen dieses Beitrages wurde die Mixed-Reality-in-the-Loop Simulation (MRiLS) zur Schulung technischer Fachkräfte sowie ein Plattform-Ansatz zur technischen Umsetzung der MRiLS vorgestellt. Die flexibel anbindbaren Mehrwertdienste der DTaaSP werden in Zukunft weiterentwickelt und um weitere Funktionalitäten (z.B. Objekterkennung) erweitert. Insbesondere die Kopplung der Multiuser-Kollaboration mit einer Sprachübersetzung des Speech-Service wird vielfältige Potentiale für multilinguale Maschinenschulungen eröffnen. Nutzende mit unterschiedlichen Sprachkenntnissen können dann an einer gemeinsamen Schulung teilnehmen und über die Echtzeitübersetzung in ihrer jeweiligen Muttersprache kommunizieren. Dies ermöglicht einen schnelleren und direkteren Austausch zwischen Schulungsteilnehmenden im

Vergleich zur bisher üblichen begleitenden Übersetzung durch einen Dolmetscher. Der Einsatz der MRiLS bietet neben den zahlreichen Potentialen in der Schulung auch neue Einsatzgebiete im Monitoring, der Inbetriebnahme sowie in der Planung und Entwicklung.

6 Danksagung

Die Ergebnisse wurden im Rahmen des Forschungsprojektes MRiLS (Fkz.: 16SV8348, Förderung durch das Bundesministerium für Bildung und Forschung) und der Transferplattform BW Industrie 4.0 (Förderung durch das Ministerium für Wirtschaft, Arbeit und Wohnungsbau des Landes Baden-Württemberg) erarbeitet. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autoren.

Literaturverzeichnis

- [Fe17] Fehling, D.: Neue Lehr- und Lernformen in der Ausbildung 4.0: Social Augmented Learning in der Druckindustrie. In: Berufsbildung in Wissenschaft und Praxis, Vol.46 (2), S. 30-33. 2017.
- [Hö20] Hönig, J. et al.: DigiTwin:didact - Maschinen verstehen mit dem Digitalen Zwilling in AR/VR-Lernszenarien. Ausgezeichnet mit dem AVRiL 2020 Silber.
- [Hö21] Hönig, J. et al.: Mixed-Reality-in-the-Loop Simulation von Produktionssystemen zur Aus- und Weiterbildung. atp magazin, atp 63 (6-7), Vulkan-Verlag, 2021.
- [Mi94] Milgram, P. et al.: Augmented Reality: A class of displays on the reality-virtuality continuum. In: Telemanipulator and Telepresence Technologies. 2351. 1994.
- [Sc18] Schnierle, M. et al.: Plattform für die Mixed-Reality-in-the-Loop-Simulation, wt-online. Internet: www.werkstattstechnik.de. Düsseldorf: Springer-VDI-Verlag, 2018.
- [Sc19] Schnierle, M. et al.: Mensch-Roboter-Interaktion mit Mixed Reality auf Basis einer "Digital Twin as a Service"-Plattform, atp magazin 5/2019, Robotik und Digital Twin in der Smart Factory, Vulkan-Verlag, 2019.
- [Sh19] Shamsuzzoha, A.: Digital factory – virtual reality environments for industrial training and maintenance. Interactive learning environments. 2019.
- [VD16] VDI/VDE-Gesellschaft Mess- und Automatisierungstechnik: Virtuelle Inbetriebnahme - Modellarten und Glossar, VDI/VDE 3693 Blatt 1, 2016
- [VD17] Verband Deutscher Maschinen- und Anlagenbau e. V.: IT-Report „Simulation im Maschinenbau“, 2017

Virtuelle Welten der Geometrie

Planung, Durchführung und Auswertung einer praxisorientierten Lehrveranstaltung für angehende Mathematiklehrkräfte

Lena Florian  ¹

Abstract: Räumliches Vorstellungsvermögen ist für viele Lernprozesse im Mathematikunterricht grundlegend, jedoch nicht angeboren. Virtual Reality bietet neue Formen, die Förderung von Raumvorstellung zu unterstützen. Im Projektseminar „Virtuelle Welten der Geometrie“ wurden mit Lehramtsstudierenden des Fachs Mathematik Möglichkeiten erarbeitet, Raumvorstellung mithilfe von Virtual-Reality-Anwendungen im Mathematikunterricht der Primar- und Sekundarstufe zu fördern. Dabei arbeiteten die Lehramtsstudierenden mit Studierenden der Informatik sowie Lehrkräften zwei verschiedener Schulen zusammen, um Gestaltungsprinzipien zu erarbeiten, Lernumgebungen zu entwickeln und diese schließlich in der Praxis zu erproben. Begleitend wurde der Kompetenzzuwachs der Studierenden sowie abschließend ihre Einstellung erhoben, Virtual-Reality-Technologien im Mathematikunterricht einzusetzen. Daraus ergaben sich Gestaltungsprinzipien für Lernumgebungen aus Sicht der Studierenden. Der Beitrag berichtet von der Planung, Durchführung und Auswertung der Lehrveranstaltung und zeigt erste Ergebnisse der Evaluation.

Keywords: Virtual Reality, Lernumgebungen, Lehramtsstudierende, Professionalisierung, Mathematikdidaktik.

1 Virtual Reality im Mathematikunterricht

Forschende und Lehrende beschäftigen sich bereits seit den 1990er Jahren mit den Möglichkeiten von Virtual-Reality- und Augmented-Reality-Technologien in der schulischen und universitären Lehre der Mathematik. Noch bevor es marktreife Systeme gab, entwickelten beispielsweise Kaufmann et al. [KSW00] mit der Anwendung *Construct3D* einen ersten Prototypen, der Menschen die Interaktion mit mathematischen Körpern im dreidimensionalen Raum ermöglichte. 20 Jahre später sind Virtual-Reality-Anwendungen (VR) insbesondere für den Mathematikunterricht trotzdem immer noch rar gesät [FK21].

Angesichts der wenigen VR-Anwendungen, die derzeit gezielt für den Mathematikunterricht entwickelt wurden, besteht die Gefahr, dass VR-Technologien ohne mathematikdidaktische Fundierung im Unterricht eingesetzt werden. Dabei werden nicht nur Potentiale unzureichend ausgeschöpft, sondern im schlimmsten Fall unbeabsichtigt Fehlvorstellungen bei den Schüler:innen gefördert oder Lernprozesse erschwert – beispielsweise indem Lernmodelle genutzt werden, die nicht anschlussfähig an Handlungen im analogen Raum

¹ Universität Potsdam, Institut für Mathematik, Didaktik der Mathematik, Karl-Liebknecht-Straße 24/25, 14476 Potsdam OT Golm, lena.florian@uni-potsdam.de, <https://orcid.org/0000-0003-2090-8895>

sind. Dem kann begegnet werden, indem der Transfer mathematikdidaktischer, lerntheoretischer und medienpädagogischer Erkenntnisse und Theorien bereits im Lehramtsstudium durch die Entwicklung und Gestaltung passender Lernumgebungen durch die Studierenden mit bereits bestehenden VR-Technologien und -Anwendungen gefördert wird. Auf diese Weise ist es möglich, fachdidaktische Potentiale für den Mathematikunterricht didaktisch und lerntheoretisch fundiert bereits jetzt auszuschöpfen und die Qualität der eingesetzten Lernumgebungen zu sichern.

2 Virtuelle Welten der Geometrie – Planung und Durchführung

Auf der Grundlage der genannten Prämissen wurde ein Seminar zur Förderung entsprechender Kompetenzen bei Lehramtsstudierenden der Mathematik geplant und durchgeführt. Im Folgenden werden die Grundlagen der Planung sowie erste Ergebnisse der Evaluation des Seminars vorgestellt. Ziel ist, Möglichkeiten aufzuzeigen, einen passenden Umgang Studierender mit VR-Technologien im Schulunterricht im Rahmen des Lehramtsstudiums zu fördern.

2.1 Die Ausgangslage

Das Seminar wurde im Rahmen des Master of Education für Studierende der Mathematik für das gymnasiale Lehramt, die Sekundarstufe sowie die Primarstufe geplant und durchgeführt. Die Studierenden befinden sich noch am Anfang ihres Masterstudiums und haben bisher nur wenige Praxiserfahrungen sammeln können. Sie können jedoch bereits auf mathematikdidaktische Grundlagenkenntnisse aus dem Bachelor zurückgreifen. Die entsprechenden Lehrveranstaltungen liegen für die meisten allerdings zwei Jahre zurück.

Das Seminar wurde als Innovatives Lehrprojekt des Zentrums für Qualitätsentwicklung in Lehre und Studium der Universität gefördert. Dadurch standen Mittel für VR-Equipment und eine wissenschaftliche Hilfskraft (wHK) zur Verfügung. Um eine Ausleihe von Geräten an Studierende zu ermöglichen und den Zugang für Laien zu erleichtern, wurden Oculus Quest Headsets beschafft, die eine autarke Nutzung von VR-Technologie ermöglichen.

Im Rahmen des Seminars wurde einerseits mit einem Unity-Kurs eines Informatiklehrstuhls der Universität und andererseits im Rahmen eines Campusschulennetzwerks mit einem Gymnasium und einer Grundschule kooperiert.

2.2 Didaktische Grundlagen und Ziele

Ziel des Seminars war es, Lehramtsstudierende zu befähigen, virtuelle Lernwerkzeuge und Lernumgebungen in der Geometrie – insbesondere der Raumgeometrie

- stoffdidaktisch und lerntheoretisch zu analysieren,
- geeignete Lernaufgaben zu erstellen sowie Lernumgebungen zu gestalten
- und diese in der Praxis zu evaluieren.

Das Seminar ist projektorientiert organisiert und enthält Phasen der Erarbeitung von stoff- und fachdidaktischen sowie psychologischen und ethischen Grundlagen sowie Phasen der selbstgesteuerten Projektarbeit zur Entwicklung eigener Lernumgebungen durch die Studierenden. Der Fokus bei der Planung der Lehrveranstaltung wurde auf die folgenden Handlungsfelder gelegt, die im Leitbild Lehre der Universität verankert sind: Forschungsorientierung, Kompetenzorientierung und Medienbildung. Insbesondere Aspekte der Medienbildung sollen im Folgenden kurz didaktisch eingordnet werden.

Lehramtsstudierende haben oft das Gefühl, dass die praktische Ausbildung in ihrem Studium zu kurz kommt [Wa19]. Sie wünschen sich theoretische Grundlagen, die sie dabei unterstützen, die schulische Praxis zu verstehen. Zudem werden digitale Medien in der universitären Lehre der Naturwissenschaften zumeist nur konsumierend eingesetzt und weniger konstruierend. Studierende erhalten nur selten die Gelegenheit, selbst Lernprodukte mit digitalen Medien zu gestalten [Vo19]. Um dem zu begegnen, zielt die Konzeption des Seminars darauf ab, stoffdidaktische Inhalte direkt mit der schulischen Praxis zu verknüpfen und so den Transfer theoretischen Wissens in die Praxis zu erleichtern.

Die Studierenden sammeln im Rahmen des Seminars Erfahrungen mit neuesten Technologien, die erst seit wenigen Jahren Eingang in die schulische Praxis erfahren. Durch die Verknüpfung stoffdidaktischer Grundlagen mit der Entwicklung und Erprobung eigener digitaler Lernumgebungen erhalten die Studierenden einen Zugang zur reflektierten und fundierten Nutzung im Mathematikunterricht. Durch die interdisziplinäre Verbindung mit einem Unity-Kurs werden ihnen darüber hinaus Kenntnisse und Fertigkeiten vermittelt, mit denen sie die Entwicklung von digitalen Werkzeugen de- und rekonstruieren können. Auf diese Weise werden sie – orientiert an Studien zur Bedeutung von konstruktiven Gestaltungsprozessen und damit verbundenen Selbstwirksamkeitserfahrungen [TGH18, Vo19] sowie der Dagstuhl-Erklärung [GI16] – nicht nur befähigt, VR-Technologien zu nutzen, sondern auch zu bewerten und ihre Funktionsweise zu verstehen. Grundlage hierfür bildet das TPaCK bzw. DPaCK-Modell [Hu19]. Demnach wird der Fokus nicht nur auf technisches, pädagogisches und inhaltsbezogenes Wissen, sondern auch auf fachdidaktische Analysekompetenz gelegt. Ziel ist, dass die Studierenden nach dem Seminar den Einsatz von VR-Technologien im Mathematikunterricht kritisch diskutieren und entsprechende Lernumgebungen situationsadäquat entwickeln und einsetzen können.

2.3 Verlaufsplanung und Durchführung

Das Seminar wurde sowohl im Sommersemester 2020 als auch im Wintersemester 2020/21 durchgeführt. In beiden Durchgängen mussten Anpassungen aufgrund der Covid19-Pandemie vorgenommen werden. Ein Großteil der Lehre fand online statt. Dabei wurden verschiedene digitale Werkzeuge zur Unterstützung der Studierenden genutzt. Grundlage bildete ein Moodlekurs, in den per HTML externe Anwendungen wie Padlet

[Wa12] und Mentimeter [Me14] eingebunden wurden. Insgesamt bestand das Seminar aus vier Phasen (Abb. 1).

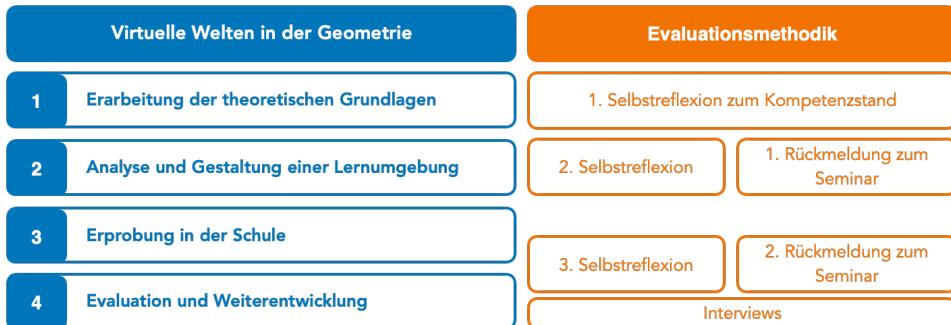


Abb. 1: Verlaufsplanung

Die erste Phase diente vor allem dazu, eine fachliche, aber auch eine gruppendifamische Grundlage für die folgenden projektorientierten Phasen zu schaffen. Sie war deshalb geprägt von synchronen Online-Sitzungen, in denen die Studierenden an die technischen und didaktischen Grundlagen herangeführt wurden. Aufgrund der Distanzlehre wurden autarke VR-Brillen an die Studierenden ausgeliehen, um ihnen ein Kennenlernen der neuen Technik zu ermöglichen und neue Formen der Interaktion durch SocialVR ausnutzen zu können. Darüber hinaus wurden im zweiten Durchgang des Seminars Holzwürfel ausgeliehen, um Grundlagen zur Raumvorstellung und damit verbundene Schwierigkeiten von Schüler:innen praktisch erarbeiten zu können. Der Hauptteil der Studierenden hatte bis zu diesem Seminar kaum Erfahrungen mit VR-Technologien sammeln können – zum Teil hatten sie zuvor noch nie eine VR-Brille aufgesetzt. Um die Studierenden nicht mit der für sie unbekannten Technik allein zu lassen, erhielten sie ein von der wHK erstelltes Video-Tutorial zur Einrichtung der VR-Brille sowie Hinweise und Videos zu ersten Schritten. Eine Sitzung des Seminars widmete sich darüber hinaus vollständig dem Kennenlernen der Technik. Die Studierenden arbeiteten zu zweit an einem Lernpfad, der sie durch verschiedene mathematikhaltige Lernaufgaben in unterschiedlichen VR-Anwendungen führte. Als Anwendungen wurden TiltBrush [Go16], SculptVR [Ro16], CubelingVR [Fl21] und AltSpaceVR [Mi15] verwendet. Durch die Arbeit zu zweit hatten sie immer eine Ansprechperson. Bei Problemen konnte die Seminarleitung individuell unterstützen. Begleitet wurden diese ersten Sitzungen von einem Informatikstudenten, der in enger Zusammenarbeit mit den Lehramtsstudierenden eine VR-Anwendung namens MatheMaler entwickelte, die mehrere Gruppen zur Gestaltung ihrer Lernumgebungen nutzten.

Nach der Erarbeitung der Grundlagen erfolgte in der **zweiten Phase** eine mehrwöchige Projektarbeit, die zum Großteil von den Studierenden selbst gesteuert wurde. Einmal pro Woche hatten die Studierenden die Möglichkeit, im Rahmen einer Feedbackrunde zur ursprünglichen synchronen Seminarzeit eine Rückmeldung zu ihren Lernumgebungen von

den anderen Studierenden und der Seminarleitung zu erhalten. Den Fortschritt ihrer Lernumgebungen dokumentierten sie in geteilten Dokumenten, damit auch die beteiligten Lehrkräfte vor der Erprobung in der Schule auf einfacherem Weg Hinweise für die praktische Umsetzung der Lernumgebungen geben konnten. Die Erprobung in der Schule musste in der **dritten Phase** aufgrund der Covid-19-Lage in beiden Durchgängen anders ablaufen als geplant, dennoch hatte ein Großteil der Studierenden die Gelegenheit ihre Lernumgebungen in der Unterrichtspraxis zu erproben. In beiden Durchgängen bereiteten die Studierenden zusätzlich zu ihrer Lernumgebung Evaluationsinstrumente vor, mit denen sie ihre Erprobung im Nachhinein auswerten konnten. In der **vierten Phase** analysierten die Studierenden im Rahmen eines Blockseminars die erhobenen Daten aus den Erprobungen an den Schulen und gaben sich gegenseitig Rückmeldungen zu ihren Lernumgebungen. Die Lernumgebungen wurden mit TiltBrush, CubelingVR und der unveröffentlichten Anwendung MatheMaler erstellt. Ergänzt wurden diese Anwendungen zum Teil durch Aktivitäten mit Papier und Stift und physischen Modellen.

3 Evaluation

Grundlage für die Planung des Seminars bildeten, wie oben dargestellt, insbesondere Studien zur Selbstwirksamkeitserfahrung von Lehramtsstudierenden in der Medienbildung und zur fehlenden Kohärenz zwischen schulpraktischer und theoretischer Ausbildung. Darüber hinaus wurde mit der Einbindung von VR-Technologien ein Medium in die mathematikdidaktische Lehre integriert, das für die Studierenden bisher unbekannt war. Die Evaluation der Veranstaltung zielt aufgrund dieser Schwerpunkte auf Erkenntnisse zu den folgenden Fragen ab:

- **Kompetenzentwicklung:** Wie schätzen die Studierenden subjektiv ihre Kompetenzentwicklung im Verlauf des Seminars ein? Inwiefern sind sie in der Lage, Theorie und Praxis miteinander in Verbindung zu bringen?
- **Virtual Reality:** Welche Einstellung haben die Studierenden gegenüber dem Einsatz von VR-Technologien im Mathematikunterricht nach Besuch des Seminars?
- **Gestaltungsprinzipien für eine Lernumgebung:** Welche Kriterien legen die Studierenden einer Lernumgebung mit VR-Technologien für den Mathematikunterricht zugrunde? Inwiefern können die Studierenden als Expert:innen für solche Lernumgebungen dienen?

3.1 Methodik

Um die Forschungsfragen untersuchen zu können, wurden verschiedene Methoden der qualitativen Bildungsforschung genutzt (Abb. 1). Im Verlauf des Seminars reflektierten die Studierenden im Rahmen von kollaborativen Reflexionstabellen ihren Kompetenzzuwachs. Dazu schätzten sie in einer geteilten Tabelle orientiert an den Standards für die

Mathematiklehrkräftebildung [DGM08] ihre Kompetenzen in den Bereichen Stoffdidaktik und Lernumgebungen ein (vgl. Abb. 2). Die Einschätzung erfolgte anhand einer 5er-Likertskala, deren Antworten lernpsychologisch angepasst wurden, um die Studierenden im Reflexionsprozess zu unterstützen. Darüber hinaus wurde an zwei Zeitpunkten – in der Mitte des Semesters und nach Abschluss des Seminars – durch offene Fragen ein Feedback zum Seminar eingeholt. Nach Ende des Seminars wurden mit einzelnen Studierenden leitfadengestützte Interviews durchgeführt. Die Teilnahme an den Interviews war freiwillig. Insgesamt nahmen bisher 6 von 22 Studierenden an den Interviews teil – drei weibliche und drei männliche im Alter von 21 bis 30 Jahren. Aufgrund der pandemischen Situation wurden die Interviews per Videotelefonie durchgeführt, sodass bei Bedarf Arbeitsdokumente aus dem Seminar gezeigt werden konnten. Die Interviews gliederten sich orientiert an den Forschungsfragen in drei Schwerpunkte: 1. Kompetenzentwicklung und Theorie-Praxis-Bezug, 2. Virtual Reality und 3. Gestaltungsprinzipien für eine Lernumgebung. Insbesondere wurden Fragen zu den Vorkenntnissen sowie zu den ersten Erfahrungen mit Virtual Reality gestellt. Zudem wurde gefragt, wie die Studierenden den Einsatz von VR als für sie neue Technologie in der Schule erlebt haben und inwiefern sie zukünftig VR in der Schule einsetzen würden. Die Interviews und Rückmeldungen der Studierenden wurden mit einer qualitativen Inhaltsanalyse nach Mayring [Ma15] analysiert. Dabei wurden die deduktiv gebildeten Kategorien Kompetenzzuwachs, Theorie-Praxis-Bezug, Einstellung zu VR und Gestaltungsprinzipien durch induktive Kategorien verfeinert. Mit den Studierenden aus dem Wintersemester 2020/21 wurden bisher keine Interviews geführt.

3.2 Der Kompetenzzuwachs der Studierenden

Die Reflexionen der Studierenden zu Beginn und zum Ende des Semesters geben einen guten ersten Eindruck über ihre Kompetenzentwicklung. In Abb. 2 ist der Vergleich zwischen der ersten (I) und der letzten Sitzung (II) im Wintersemester dargestellt. Die Studierenden verfügten zu Beginn zum Teil lediglich über stoffdidaktische Grundlagen. Im Verlauf des Semesters entwickelten sie insbesondere ihre Kompetenzen in der Analyse und Gestaltung von Lernumgebungen. Im Mittel veränderte sich ihre subjektive Einschätzung in diesen Bereichen um 2 Punkte. Deutlich wird jedoch auch, dass viele Lernprozesse nach Ansicht der Studierenden noch nicht abgeschlossen sind, insbesondere in den theoretischen Grundlagen zur Raumvorstellung und Raumgeometrie gaben 7 von 12 Studierenden am Ende des Semesters an, sie müssten ihre Kenntnisse noch vertiefen. Zudem erläuterten Studierende aus dem Sommersemester in den Interviews, dass sie viele theoretische Grundlagen vom Beginn des Semesters wieder vergessen hätten. Sie waren allerdings in der Lage Kriterien, die für sie in einer Lernumgebung bedeutsam sind, mit Theorie und Erfahrungen aus der Praxis zu stützen. Dabei wurde vor allem auf Modelle der Raumvorstellung (Perspektivwechsel) und Grundlagen zur Kommunikation oder Binnendifferenzierung eingegangen. Darüber hinaus konnten sie Bezüge zu fachdidaktischen Veranstaltungen im Bachelor herstellen.

Es besteht also eine gewisse Diskrepanz zwischen der subjektiven Einschätzung der Studierenden und ihren Äußerungen in den Interviews sowie den Rückmeldungen zum Seminar. Diese Diskrepanz lässt sich auch innerhalb einer Kohorte feststellen und ist möglicherweise dadurch zu erklären, dass die Studierenden verschiedene Modelle bereits in ihre Wissensstrukturen integriert haben, aber noch nicht über die Kompetenz verfügen, diese aktiv zu benennen. Hinzu kommt, dass die Interviews und auch die letzte Reflexion kurz nach der Praxisphase erfolgten. Die Studierenden hatten sich demnach über einen längeren Zeitraum vor allem mit Fragen der unterrichtspraktischen Umsetzung beschäftigt. In der letzten Sitzung wurden die theoretischen Grundlagen nur kurz wiederholt. Eine letzte Erklärung könnte sein, dass sie nach der Praxiserfahrung einen erhöhten Bedarf zur Vertiefung erkennen können. Dies lässt sich aber mit den vorliegenden Daten nicht verifizieren.

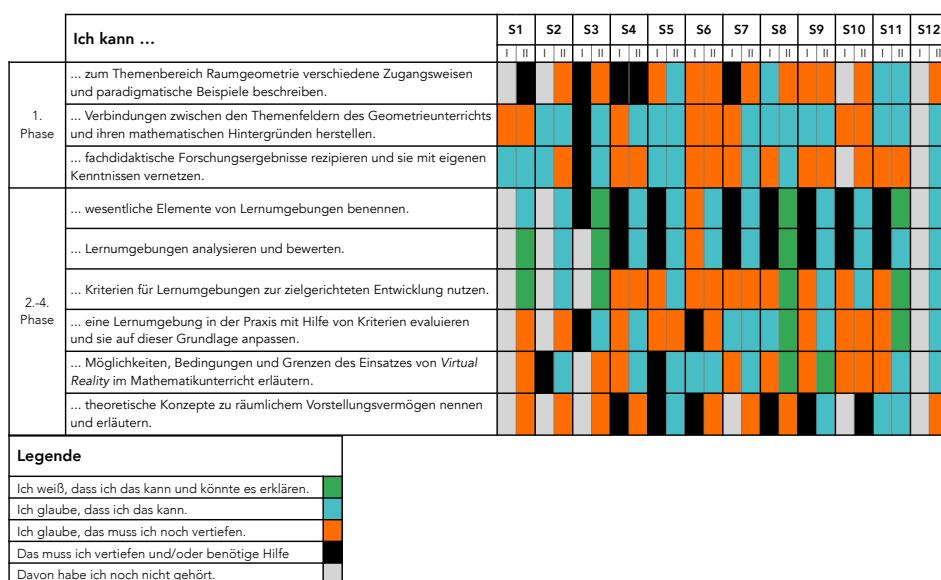


Abb. 2: Kompetenzstand in der ersten (I) und letzten Sitzung (II)

In den Interviews nannten die Studierenden über die Kompetenzen in der Reflexionstabelle hinaus weitere Fähigkeiten und Fertigkeiten, die sie im Rahmen des Seminars vertieften konnten. Darunter fiel die Entwicklung der eigenen Raumvorstellung durch die Nutzung der VR-Brille und durch die Gestaltung der Lernumgebung für die Schüler:innen. Darüber hinaus nannten sie vor allem Kompetenzen in der Planung von Unterricht und Medienkompetenzen.

Nach Abgleich der subjektiven Einschätzung der Studierenden und der Fähigkeiten, die sie im Interview zeigten, wurden vor allem Theorie-Praxis-Verknüpfungen bei der Erstellung von Lernumgebungen gefördert. Alle Studierenden waren in der Lage, nach Ende des Seminars Kriterien für eine gute Lernumgebung zu benennen und zu begründen. Besonders bedeutsam erschienen den Studierenden ihre Kompetenzentwicklungen im Bereich

der Raumvorstellung und in der Entwicklung und Analyse von multimedialen Lernumgebungen sowie der praktische Einsatz dieser. Hervorgehoben wurde die Bedeutung der praktischen Umsetzung in der Schule und der damit verbundene Erkenntnisgewinn.

3.3 Die Einstellung der Studierenden zu VR im Mathematikunterricht

Die Studierenden äußerten sowohl in den Interviews als auch in den Zwischenevaluierungen des Seminars anfängliche Skepsis in Bezug auf VR-Technologien und Technik im Allgemeinen. Sie bezeichneten sich selbst in den Interviews als „nicht technikaffin“ und „keine Zocker:in“. Umso erstaunlicher ist ihre Einstellung zu VR im Unterricht nach Abschluss des Seminars und die Beschreibung ihrer ersten Erfahrungen mit der VR-Brille in den eigenen vier Wänden. Die Studierenden zeigen Faszination und Begeisterung für die ihnen unbekannte Technik und keinerlei negative Erfahrungen bei der ersten Nutzung, obwohl diese allein zuhause stattfand. Einige erwähnen in den Interviews und den schriftlichen Rückmeldungen die Video-Tutorials und den Lernpfad als hilfreiche Ergänzungen zu ihren ersten eigenen Erfahrungen mit der VR-Brille.

Die Studierenden zeigen sich nach dem Seminar offen, VR-Technologien in der Schule einzusetzen, nennen jedoch auch Einschränkungen wie die Gesundheit (z.B. „die Brille sollte nicht zu lange von Kindern getragen werden, weil die gesundheitlichen Auswirkungen noch nicht feststehen“), technische Voraussetzungen in den Schulen oder fehlende VR-Anwendungen für den Mathematikunterricht. Sie sind in der Lage, konkrete Anwendungskontexte im Mathematikunterricht zu benennen – wie Rotationskörper oder die Einführung des räumlichen Koordinatensystems – und sogar Übertragungen in ihr jeweiliges anderes Fach (Optik im Physikunterricht, Moleküle im Chemieunterricht).

Diese Offenheit gegenüber VR lässt sich aus den Interviews heraus einerseits durch die positiven Erfahrungen im Seminar, andererseits aber auch die positiven Rückmeldungen der Schüler:innen erklären. Die Studierenden haben sich in der Aneignung einer neuen Technologie und deren Anwendung in der Praxis als selbstwirksam wahrgenommen. Manche Studierende sehen allerdings auch weiteren Bedarf zur Vertiefung und wünschen sich, weitere Erfahrungen mit VR im universitären Kontext machen zu können – insbesondere auch zu ethischen Fragen von VR in der Schule.

3.4 Gestaltungsprinzipien aus Sicht der Studierenden

Im Rahmen des Seminars haben die Studierenden vor der Entwicklung eigener Lernumgebungen literaturbasiert Kriterien für gelungene Lernumgebungen gesammelt. Zurückgegriffen wurde dabei zum Beispiel auf Literatur von Roth [Ro19] und van Randenborgh [Ra15]. Nach Abschluss des Seminars wurden die Studierenden aus dem Wintersemester im Rahmen der Interviews gebeten, Kriterien für eine aus ihrer Sicht gelungene Lernumgebung zu nennen. Diese bestand dabei für beide Gruppen grundsätzlich aus virtuellen

und physischen Materialien, wie Holzwürfeln oder einer VR-Anwendung, sowie Aufgabenstellungen und dem allgemeinen Klassenraumkontext. Sie orientierten sich damit unter anderem an Reinmann und Mandl [RM06]. Vor diesem Hintergrund wurden folgende Kriterien genannt:

	Sommersemester	Wintersemester
Mathematik	Zielgerichtete Nutzung von VR als Werkzeug und nicht als Spielzeug, Fachlichkeit	Mathematische Sinnhaftigkeit (insbesondere hinsichtlich der Strukturierung: Anknüpfen an Bekanntes, transparente Ziele der Lernumgebung), Auswertung, Reflexion und Evaluation der Lernumgebung
Klassenraumkontext	Offenheit, Berücksichtigung des Klassenraumkontextes bei der Planung des Einsatzes, Binnendifferenzierung, Planung von Kommunikation und Kooperation unter den Schüler:innen	Lernzielen entsprechende Arbeitsform (z.B. „mathematisches Kommunizieren stärken = Gruppenarbeit“), Lernklima (z.B. „Träger:in der Brille wird nicht angefasst“), Arbeitsmaterial: Verständlich, zielgerichtet, angemessen (zeitlich, Schwierigkeitsgrad), differenziert. ²
VR-Anwendung	Einfach und übersichtlich: Nicht zu viele Funktionen in einer Anwendung, einfache Steuerung, Nutzung von VR als Motivationsfaktor, Möglichkeit der Dokumentation von Ergebnissen	–
Organisatorisches	Möglichkeiten zum Streamen, klare Zeitangaben	Raumaufteilung (genug Platz für VR), klare Zeitangaben, Einführung in VR

Tab. 1: Kriterien der Studierenden für eine gelungene Lernumgebung

Ausgeschlossen wurden in der Gruppe im Wintersemester „Entdeckendes Lernen und Experimentieren“ als allgemeingültige Kriterien für jede Lernumgebung sowie die Nutzung von Hilfekarten als Maßnahme zur Binnendifferenzierung, da diese in der praktischen Er-

² Als Begründung wurde hier unter anderem angeführt, dass die Schüler:innen direkt und ungestört arbeiten können sollten, ohne dass Fragen auftreten. Dazu gehöre „die Einweisung in digitale/virtuelle Medien und jede:r kann in seinem Tempo und mit bevorzugten Mitteln lernen“.

probung kaum von den Schüler:innen genutzt worden seien. Auffällig ist, dass die Studierenden fast ausschließlich Kriterien nannten, die sie zuvor aus der Literatur erarbeitet hatten (Tab. 1 **Fehler! Verweisquelle konnte nicht gefunden werden.**). Zum Teil wurden diese Kriterien genannt, konnten aber nicht weiter ausgeführt werden, z. B. „Fachlichkeit“. Andere Kriterien wurden durch Praxiserfahrungen gestützt, wie beispielsweise Kommunikation, Kooperation oder auch Binnendifferenzierung.

Im zweiten Durchgang wurde deshalb eine Seminarstunde eingefügt, in der die Studierenden zunächst eigene Kriterien formulierten und danach mit Literatur konfrontiert wurden. Ziel war es, eine größere inhaltliche Auseinandersetzung mit den Kriterien zu erreichen. Sie erhielten die Aufgabe, ihre Kriterien mit Hilfe von Literatur zu stützen und im Anschluss auszuwählen, welche ihnen am wichtigsten und schlüssigsten erschienen (Tab. 1). Die Kriterien wurden im Dokument zum Teil explizit mit Literatur begründet und es ist insgesamt eine stärkere Durchdringung der theoretischen Grundlagen zu erkennen. Einschränkend sei erwähnt, dass die Studierenden im zweiten Durchgang bisher nicht zu den Kriterien interviewt wurden.

Nach dieser ersten Auswertung der Kriterien der Studierenden lässt sich feststellen, dass sie nach Abschluss des Seminars in der Lage sind, auf Grundlage von theoretischen Modellen aus der Literatur und ersten unterrichtspraktischen Erfahrungen sowie ausgiebiger Auseinandersetzung mit der bereitgestellten VR-Technologie, Kriterien für eine Lernumgebung zu benennen, die virtuelle Welten mit realen Handlungsräumen verbinden.

4 Fazit

Ziel des Seminars war es, dass die Studierenden den Einsatz von VR im Mathematikunterricht wissenschaftlich fundiert reflektieren und situationsadäquat begründen können. Das Seminar wurde insgesamt sehr positiv von den Studierenden bewertet. Dies hing zum einen mit den positiven Unterrichtserfahrungen zusammen, die die Studierenden während des Seminars sammeln konnten, aber auch mit der offenen und unvoreingenommenen Auseinandersetzung mit einer für sie neuen Technologie. Sie betonten in den Rückmeldungen, dass sie sich im Rahmen des Seminars sehr wohl gefühlt haben und dadurch ermutigt wurden, sich mit theoretischen Modellen und technischen Grundlagen auseinanderzusetzen. Besonders hervorgehoben wurden der Aufenthalt in der Schule und die eigene Entfaltungsmöglichkeit im Rahmen des Seminars. Diese positive Grundstimmung bildete die Basis für eine erfolgreiche Kompetenzentwicklung der Studierenden – insbesondere im Bereich der Raumvorstellung sowie der Analyse und Gestaltung von Lernumgebungen für den Mathematikunterricht. Dabei bewerteten sie selbst ihre Kompetenzen zum Teil recht kritisch und sehen weiteres Vertiefungspotential. Bezüge zwischen Theorie und Praxis konnten bei der Beschreibung von Praxiserfahrungen hergestellt, aber zum Großteil auf Nachfrage nicht explizit als Theorie-Praxis-Bezug von den Studierenden benannt werden. Hier besteht eine Diskrepanz zwischen Wahrnehmung der Studierenden und tatsächlicher Kompetenz. Denn anhand weiterer erhobener Daten wird deutlich, dass

die Studierenden durchaus in der Lage sind, verschiedene Anwendungskontexte für VR im Mathematikunterricht zu nennen und auch mit mathematik- und mediendidaktischen Grundlagen kritisch zu hinterfragen. Sie können Kriterien für eine gelungene Lernumgebung beschreiben und begründen – zum Teil mithilfe von Theorie, zum Teil mithilfe von Erfahrungen aus der Praxis. Eine Analyse der Qualität der genannten Kriterien steht jedoch noch aus. Die Studierenden können auf dieser Grundlage derzeit zwar noch nicht als Expert:innen für die Entwicklung von VR-haltigen Lernumgebungen gesehen werden – wohl aber als fortgeschrittene Noviz:innen. Da es sich um eine geringe Versuchspersonenzahl handelt, ist die Aussagekraft der Daten jedoch limitiert. Die Ergebnisse können lediglich als explorativ gewertet werden. In zukünftigen Erhebungen könnten unter anderem die Einstellung und Akzeptanz der Technologie näher untersucht werden – beispielsweise mit Modellen wie dem TAM/TAM2-Modell [VD00].

Insgesamt kann der Einsatz von VR in der mathematikdidaktischen Lehre – zumindest im Kontext eines ähnlichen Seminars – empfohlen werden. Die Studierenden haben nach Abschluss des Seminars eine positive Einstellung gegenüber dieser (für sie) neuen Technologie gewonnen und bezeichnen sich selbst als „offener gegenüber der Technik im Allgemeinen“. Im Sinne eines lebenslangen Lernens sind dies gute Voraussetzungen für die Ausbildung von engagierten und zugleich reflektierten Mathematiklehrkräften.

Literaturverzeichnis

- [DGM08] DMV, GDM & MNU: Standards für die Lehrerbildung im Fach Mathematik. Empfehlungen von DMV, GDM und MNU, Mitteilungen der DMV 16, 2008.
- [FK21] Florian, L.; Kortenkamp, U.: Virtuelle Welten im Mathematikunterricht. Lernumgebungen in erweiterter Realität. In (Walter, D.; Reinhold, F.; Schacht, F.; Pinkernell, G. Hrsg.): Digitales Lehren und Lernen von Mathematik in der Schule. Springer, angenommen, 2021.
- [Fl21] Florian, L.: CubelingVR, 2021.
- [GI16] GI, Gesellschaft für Informatik e.V.: Dagstuhl-Erklärung. Bildung in der digitalen vernetzten Welt. Eine gemeinsame Erklärung der Teilnehmerinnen und Teilnehmer des Seminars auf Schloss Dagstuhl, Leibniz-Zentrum für Informatik GmbH, 2016.
- [Go16] Google: TiltBrush, 2016
- [Hu19] Huwer, J.; Irion, T.; Kuntze, S.; Schaal, S.; Thyssen, C.: Von TPaCK zu DPaCK. Digitalisierung im Unterricht erfordert mehr als technisches Wissen. MNU Journal 5/19, S. 358-364, 2019.
- [KSW00] Kaufmann, H.; Schmalstieg, D.; Wagner, M.: Construct3D: A Virtual Reality Application for Mathematics and Geometry Education. Education and Information Technologies 4/2000, S. 263-276, 2000.
- [Ma15] Mayring, P.: Qualitative Inhaltsanalyse. Grundlagen und Techniken, Beltz, Weinheim, 2015.

- [Me14] Mentimeter, www.mentimeter.com, Stand: 10.08.2021.
- [Mi15] Microsoft: AltSpaceVR, 2015.
- [Ra15] van Randenborgh, C.: Instrumente der Wissensvermittlung im Mathematikunterricht, Springer Fachmedien, Wiesbaden, 2015.
- [RM06] Reinmann, G.; Mandl, H.: Unterrichten und Lernumgebungen gestalten. In (Krapp, A.; Weidenmann, B. Hrsg.): Pädagogische Psychologie. Ein Lehrbuch. 5. Aufl., Beltz, Weinheim u.a., S. 613-658, 2006.
- [Ro19] Roth, J.: Digitale Werkzeuge im Mathematikunterricht. Konzepte, empirische Ergebnisse und Desiderate. In (Büchter, A.; Glade, M.; Herold-Blasius, R.; Klinger, M.; Schacht, F.; Scherer, P. Hrsg.): Vielfältige Zugänge zum Mathematikunterricht. Konzepte und Beispiele aus Forschung und Praxis. Springer Fachmedien, Wiesbaden, S. 233-248, 2019.
- [Ro16] Rowe, N.: SculptVR, 2016.
- [TGH18] Tulodziecki, G.; Grafe, S.; Herzig, B.: Medienbildung in Schule und Unterricht. Grundlagen und Beispiele, 1. Aufl., UTB, Stuttgart, 2018.
- [VD00] Venkatesh, V.; Davis, F. D.: A Theoretical Extension of the Technology Acceptance Model. Four Longitudinal Field Studies. *Management Science* 2/2000, S. 186-204, 2000.
- [Vo19] Vogelsang, C.; Finger, A.; Laumann, D.; Thyssen, C.: Vorerfahrungen, Einstellungen und motivationale Orientierungen als mögliche Einflussfaktoren auf den Einsatz digitaler Werkzeuge im naturwissenschaftlichen Unterricht. *Zeitschrift für Didaktik der Naturwissenschaften* 1/2019, S. 115-129, 2019.
- [Wa12] Wallwisher: Padlet, 2012.
- [Wa19] Wagener, U.; Reimer, M.; Lüschen, I.; Schlesier, J.; Moschner, B.: „Kraß lehramtsbezogen“. Lehramtsstudierende wünschen sich mehr Kohärenz in ihrem Studium. Herausforderung Lehrer*innenbildung – Zeitschrift zur Konzeption, Gestaltung und Diskussion 1/2019, S. 210-226, 2019.

Remote Emergency Teaching and Virtual Reality Education: A Case-Study Using VR Nuggets in Non-VR Courses

Robin Horst¹, Ramtin Naraghi-Taghi-Off², Linda Rau³, Ralf Dörner⁴

Abstract: COVID19 posed several challenges to educators, for example, the abrupt transition into Remote Emergency Teaching (RET). One area of Computer Science education where such challenges are evident is the education about Virtual Reality (VR) development, normally involving specific hardware for testing changes to a virtual environment. This also makes a high-level introduction to VR challenging. In this paper, we describe a case-study conducted within two non-VR courses utilizing VR Forge - a facile VR development tool providing authors with ready-made VR experiences based on educational design patterns as a start for further authoring processes. We state lessons learned based on the use of the resulting short VR experiences within two semesters at a local university and contribute a practical insight into how we could implement the teaching of basic VR authoring processes in the last two semesters that were coined by challenges due to RET.

Keywords: Virtual Reality Education, Remote Emergency Teaching, Multimedia Content Creation, Development Frameworks and Environments, Virtual Reality Nuggets

1 Introduction

The sudden emergence of the COVID19 pandemic posed many unexpected challenges at educational institutions such as universities. As a precedent within the recent decades, many educators had to switch to a paradigmatic teaching methodology that coined the education within times of COVID19 – Remote Emergency Teaching (RET) [Ho20]. Universities had to close, and only virtual communication was conceivable to continue teaching. RET as a response to a crisis or disaster differs significantly from well-planned online learning experiences. However, in future, not all courses can additionally be planned as an online learning experience in case of a crisis like COVID19.

In some areas, the abrupt transition to RET might have caused more severe problems than in others. For example, introducing Computer Science students to VR and VR

¹ RheinMain University of Applied Sciences, Design Computer Science Media, Kurt-Schumacher-Ring 18, 65197 Wiesbaden, robin.horst@hs-rm.de, 

² RheinMain University of Applied Sciences, Design Computer Science Media, Kurt-Schumacher-Ring 18, 65197 Wiesbaden, Ramtin.naraghi-taghi-off@student.hs-rm.de,

³ RheinMain University of Applied Sciences, Design Computer Science Media, Kurt-Schumacher-Ring 18, 65197 Wiesbaden, linda.rau@hs-rm.de,

⁴ RheinMain University of Applied Sciences, Design Computer Science Media, Kurt-Schumacher-Ring 18, 65197 Wiesbaden, ralf.doerner@hs-rm.de 

development processes normally requires specific VR hardware for testing changes to a virtual environment and experiencing the vividness of VR. Such challenges might mitigate when VR is the overall topic of an entire course. Mobile VR hardware might be assigned and passed back and forth between all students long term over a semester, and exercises might be able to cover more specific tasks of VR development that can also be taught without hardware, such as VR system design, 3D math, etc. However, when VR is only a sub-aspect of a course, a more high-level introduction to VR development is problematic.

A recent VR authoring approach that aims to be suitable for and conducted by novices is called VR nuggets [HD19b]. Authoring tools based on VR nuggets provide authors with ready-made VR experiences that are based on design patterns derived from the application domain of a VR nugget. The VR nugget authoring concept draws from educational concepts, particularly Microlearning [Hu05a, Hu05b], where learning structures such as courses are divided into self-contained and reusable units.

In this paper, we explore whether such pattern-based VR experiences and their creation can in turn be used for teaching purposes. We conducted a case-study with Computer Science students in courses that were not particular VR courses. We evaluate the suitability of using VR nuggets to provide learners with crucial firsthand contact with VR and let them try out software that may otherwise only be executable with associated hardware only available within the university lab. The study took place during two semesters of RET due to the COVID-19 pandemic outbreak and utilized VR Forge [HD19b] as VR-nugget-based authoring environment. We address the following research question: Is pattern-based VR authoring suitable for being used in teaching within a Computer Science-related university course? We focus on answering the question with respect to a student's experience and investigate it with respect to RET.

This paper is organized as follows. The next section reviews related work about VR education and RET, and VR nuggets. Section 3 presents our case-study and discusses findings and lessons learned. The last section concludes our work and gives an outlook on future work.

2 Related Work

In this section, we briefly review the concept of VR nuggets and the VR Forge authoring tool that we are going to use within our case-study. Furthermore, we review related work about VR Education and RET.

2.1 VR Nuggets

VR nuggets are short VR experiences based on patterns derived from a domain of their application. Current research focused on educational patterns as an evident application

domain, since the usage of VR nuggets is based on methodology from the educational domain, namely from Microlearning [HD19b]. The authors also introduce a novel authoring approach for such short educational VR experiences based on VR nuggets that exploits the reusability of pattern-based authoring approaches. The authoring approach with VR nuggets was built to let non-programmers create VR experiences based on ready-made VR nuggets, where content can be adjusted to a certain degree to tailor a VR nugget towards a specific use-case of the authors. The authoring approach with VR nuggets distinguishes between three phases, each involving a different authoring role [Ho21]. An initial authoring system must be implemented in the first phase, conducted by programming and design professionals. The second phase entails deriving recurring educational scenarios suitable for VR. These are formulated as design patterns and for each pattern, a VR nugget is created within the authoring system. We illustrate this relation in Fig. 1. Educators themselves use the resulting VR authoring system to generate VR content in the final phase. This is accomplished by instantiating a VR nugget of the chosen pattern in the system and gradually substituting assets (e.g., 3D models, pictures, etc.) with their own material to tailor the already implemented VR nuggets for a certain use-case.

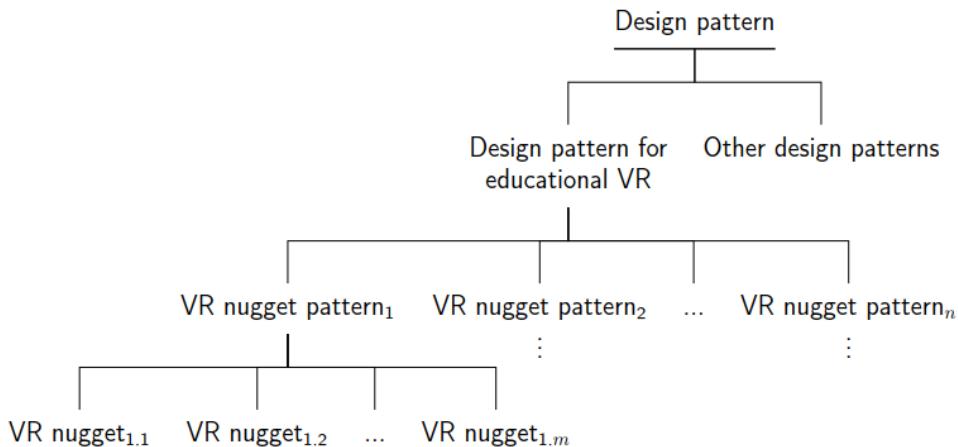


Fig. 1: The hierarchy graph shows the connection of VR nuggets and patterns. The term VR nugget pattern is used in this graph to illustrate that there may be design patterns for educational VR that are not suitable for VR nuggets, for example, when they are not self-contained or would not result in a short VR experience.

Studies indicate that educational experts and learners [HD19a] could successfully use the integration of short learning-nugget-like VR experiences. However, only the resulting VR applications were utilized. On a meta level, the actual facile VR content creation approach might also be suitable to be utilized for educational purposes itself, for example, to introduce VR technology and their development in courses where VR is only a sub-topic and cannot be taught in-depth. Still such investigations remain open.

Currently, there exist two authoring systems that utilize VR nuggets – VR Forge [HD19b] (Fig. 2) and IN Tiles. Since IN Tiles provides authors with an immersive authoring

[LKB05] interface, only VR Forge might be suitable for RET purposes, also since it provides a VR simulator that lets authors experience VR nuggets without VR hardware.



Fig. 2: A screenshot of the VR Forge authoring interface with a simple nugget active that annotates components of a fuel cell model.

2.2 VR Education and Remote Emergency Teaching

People can hardly put themselves into the mind of other people [De73]. This indicates that a VR experience and perceptual aspects related to such immersive technologies cannot be communicated adequately by only using textual or oral descriptions. More general educational research also highlights that students can benefit from hands-on-activities during their studies (e.g., [SH96, HGB10]). This is also underlined by different prominent instruction methods, for example, grounded in constructivism [Ma09]. Recent research that focuses on VR education supports the notion of including firsthand experiences for students. For example, student could benefit from actually using VR hardware, setting up a VR system, adjusting content of it, and then testing it again with other students in small mockup user tests [DH21]. An overview on the design of VR courses can be found in [SD17]. Overall, a certain immersion and experiences should be included in VR course curricula [Fo20].

Not only the learning experience, but also efficiency-related aspects matter when planning courses with VR technology [DH21]. Due to the complexity of the topic, VR development is often a part of longer courses (e.g., capstone courses [Ta16]). Still, VR and its development can also be targeted in shorter courses, for example, conducted within a single day, or short introductions to the topic within a single session of a non-VR course.

There are very different approaches of conveying knowledge about VR and VR development within courses, resulting in a varying latent difficulties that come with them. In the most simple case, ready-made applications can be utilized by students as black-box applications, so that students only have to set up the VR hardware. However, this will not raise much awareness with regard to content creation for a VR application. On the more complex side of the VR education spectrum, students could build applications with low-level graphic APIs such as Direct3D or Vulkan. This would need more time-intensive preparations of students and prior knowledge in the Computer Sciences. In the middle of this spectrum, again, we identified the VR nugget approach to be particularly interesting, with not many other solutions being available. Both the preparation of educators and students could be reduced to a minimum since students would create their own VR experiences with the lightweight VR nugget approach. High-level VR development tasks such as searching the web for appropriate assets (e.g., 3D models and textures) and the composition of a 3D environment from such assets could be part of the students' tasks. These are important tasks for students since VR development is becoming more popular due to the rising availability of low-cost VR systems within the last years [AH08], also for educational purposes.

With regard to RET, there does not exist much literature dealing with teaching VR development. At this point, RET should clearly be distinguished from other remote or online teaching methodology. RET is a response to an unexpected switch to remote teaching in the consequence of a disaster such as the COVID19 pandemic [Ho20]. As a result, different teaching methodologies are instantaneously not available [ASK20], such as group sessions, co-operative usage of hardware etc. While some courses may not be affected and impaired by such restrictions, VR-related teaching is surely influenced to a high degree. We conclude that this novel educational paradigm should be examined. For example, it might be beneficial to provide students with practical firsthand experiences of VR development using VR nuggets, particularly, when VR development is not a course's main topic and thus cannot be covered in all theoretical details

3 Case Study with VR Nuggets Authoring within RET

In a user study, we evaluated our pattern description scheme and explored the use of VR nuggets for teaching purposes. The user study involved 34 unpaid participants (2 female, aged between 20 and 30 years with \bar{O} 24.8 and SD 2.9). Their experience in the area of VR was captured on a 0-3-point scale with \bar{O} 1.1 and SD 0.1, where 0 means they never used VR technologies and 3 means they regularly use VR. On that basis, we classify the participants of our study as novices in the field of VR. All participants were computer science students at the local university and attended a computer games and entertainment computing course. The study was conducted during two semesters. Each course had 17 students. VR nuggets were utilized within both courses.

The user study was conducted as a moderated remote study and took place as follows: At first, participants were welcomed and then informed about the study's topic. They were briefly introduced to the user interface of the VR forge authoring tool [HD19b] that we kindly obtained from its developers for our study. Thereafter, the study's actual tasks started. We asked our participants to create VR nuggets with given intent and provided exemplary assets between which they could choose. After they created a set of VR nuggets with the tool, they were asked to explore their creations using the VR simulator that the VR forge incorporates. It uses WASD and mouse controls as in first-person computer games to navigate through the immersive learning experiences from a VR user's point of view and interact with the content.

Different evaluation tools were used within our study to draw conclusions on the suitability of VR nuggets for teaching purposes in remote emergency teaching (aspect 1; A1). Quantitative data was captured within a post-study questionnaire. Six items I1 to I6 were included concerning the use of the short and pattern-based VR experiences. The items utilized a 7-point Likert scale, where 0 depict negative and 6 positive answers. The questionnaire was translated into the participants' native language. The following were the items:

- I1. The brief use of patterns and VR nuggets gave me some insight into VR creation.
- I2. Briefly using the patterns and VR nuggets gave me a better understanding of VR.
- I3. I would like to look into VR nuggets beyond this study to create VR content myself.
- I4. I would like to experience my created VR nuggets with VR hardware.
- I5. I could easily imagine the use of VR nuggets in teaching (e.g., in a separate session of another lecture) to bring VR into other courses as well.
- I6. I could see the use of VR nuggets in teaching being well suited to courses in other disciplines (outside of computer science).

Furthermore, the abbreviated version of the AttrakDiff questionnaire [HBK03, Us21] was utilized to measure how the participants perceived the product character [Ha18] of the VR Forge authoring tool within an educational environment based on pragmatic qualities (usability) and hedonic qualities. Demographic questions and space for free text comments concluded the questionnaire. Furthermore, thinking-aloud methodology [Le82] was conducted during the sessions. A single session of the study was performed within 45 minutes.

3.1 Analysis of the Results

Figure 3 represents the single items' value distributions. The box-whisker plots show that all questions' mean values lie above 3, which is the neutral value of the utilized 7-point scale. Items I2, I3, and I4 show large a deviation ranging from over five or more values. I1 has a small variation including only one value (5). Outliers can be observed at I1, and I5. No item received a value of 0. Wilcoxon signed-rank tests [WW64] were conducted

on the individual items to analyze how the items were rated by the participants compared to a hypothetical neutral rating. A test was run for each item and the overall mean value (A1) compared to the neutral value of 3 points. With a threshold for statistical significance of 5%, 6 tests did confirm significant differences (all except for I3). Table 1 shows all p-values aside from the mean values and standard deviations for each question and each aspect.

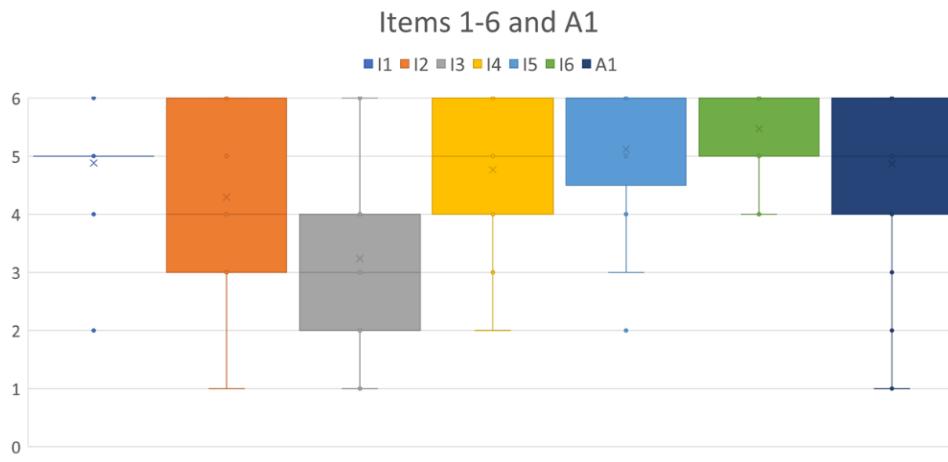


Fig. 3: Box-whisker plots for the questionnaire items I1-I6 and A1.

Figure 3 (right) also shows the value distributions of the aggregated aspect (A1) in a box-whisker. The mean value lies above 4. It shows a large deviation ranging from 6 to 1. No outlier could be detected. A Wilcoxon signed-rank test was conducted to test the aspect's outcome against a neutral rating. The result is shown in Tab. 1. It confirms a significant difference.

Questions/ aspects	$\bar{\theta}$ -values	Standard deviation	P-values
I1	4.8824	0.9275	$p \leq 0.0001$
I2	4.2941	1.5718	$p = 0.0068$
I3	3.2353	1.6405	$p = 0.6377$
I4	4.7647	1.3005	$p = 0.0003$
I5	5.1176	1.2187	$p \leq 0.0001$
I6	5.4706	0.7998	$p \leq 0.0001$
A1	4.8706	1.3035	$p \leq 0.0001$

Tab. 1: Mean values and output of the Wilcoxon signed-rank tests for I1 to I6 and A1. All values are rounded to four decimal places. 0 is the lowest, and 6 is the highest possible mean value.

The free-form comments from the questionnaire, observations, and statements from the thinking-aloud protocol were used to capture additional data about the usage of VR nuggets for remote VR teaching. Regarding A1, 21 participants addressed the tool's usability and specifically the movement of the 3D content while creating the 3D scene. Three participants noted it negatively with statements such as 'Sometimes it moves as expected, and sometimes it moves differently'. The participant referred to the movement of 3D objects. Another one said that 'Once you get used to the movement, it works'. Three stated positive comments such as 'All very intuitive with the movement of the mouse', whereby the participant referred to the camera movement centered on the currently selected object of the scene. Besides issues regarding the authoring tool, two participants addressed the VR simulator that was utilized during the remote teaching scenario for interactively experiencing what they created. They stated that the VR simulator helped them to get a 'Quick understanding to test VR remotely and without hardware'.

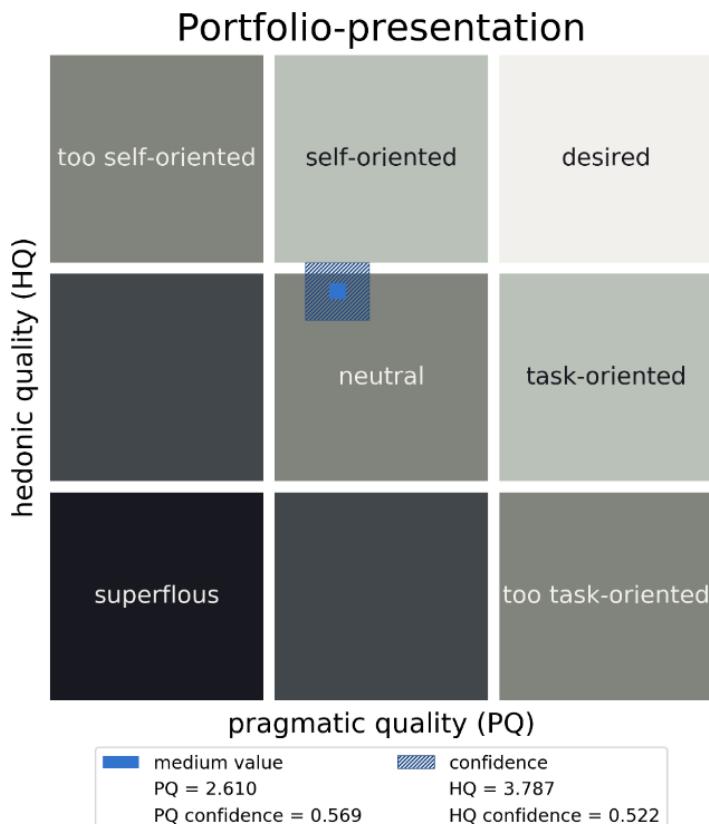


Fig. 4: Portfolio-presentation of the AttrakDiff values. The small dark blue square represents the classification of the VR Forge for our teaching purpose, and the light blue square represents the confidence.

The outcome of the AttrakDiff questionnaire was used to analyze the product character of VR Forge rated by our participants within the educational setting. The results are shown in the portfolio presentation in Fig. 4. It shows that the tool was assessed with higher hedonic than pragmatic quality. The visualization places VR Forge within the area ‘neutral’. Its confidence rectangle intersects with the area ‘self-oriented’.

The description of the word-pairs (Fig. 5) shows the mean value for each item of the AttrakDiff individually. Three of four values related to the pragmatic quality are below 3, and only one item has a mean value higher than 3, which is ‘confusing–clearly structured’. On the contrary, three items relating to the hedonic quality have a value higher than 3, and only one was lower than 3 (‘tacky–stylish’). The two items related to the tool’s overall attractiveness are placed near the 3-point mark (‘ugly–attractive’) and between the values 1 and 2 (‘bad–good’). ‘Bad–good’ also has the lowest mean value of all items. The item with the highest mean value is ‘unimaginative–creative’ with approximately 4.5.

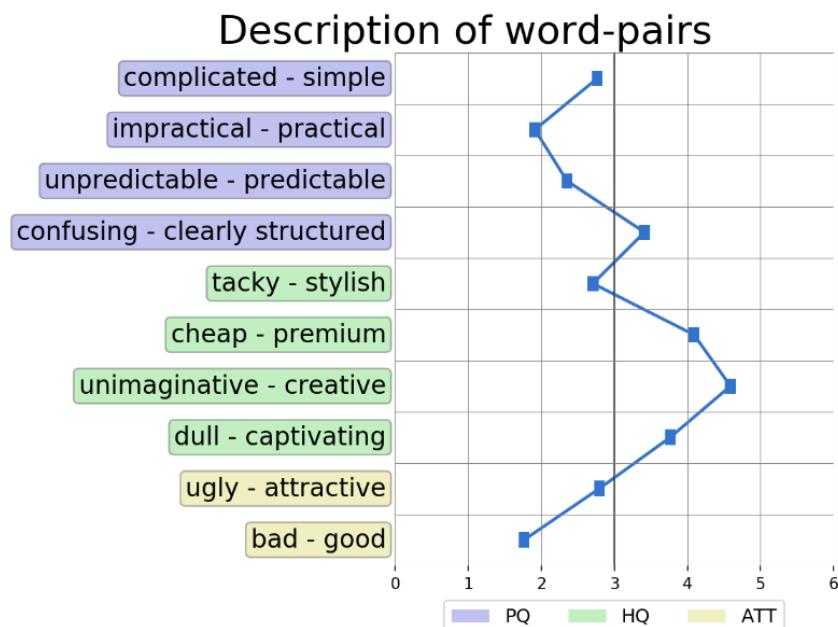


Fig. 5: Word-pair visualization of the mean values of the AttrakDiff questionnaire. Each word-pair represents a single item within the questionnaire. The first four items relate to the pragmatic quality, the second four items to the hedonic quality, and the last two items to the attractiveness of the VR Forge.

3.2 Discussion of the Results

The user study results show that the students could well imagine their use on the one hand within Computer Science-related courses where VR may only be a side aspect and on the other hand in courses that are even outside of computer science (the questions with the highest ratings I5 and I6). However, the VR nugget authoring and usage could not clearly motivate the students to continue engaging with it beyond the computer games course they took at this time (I3). This statement is supported by I6 being the only question that did not confirm a statistically significant difference between its rating and a neutral rating. Furthermore, the AttrakDiff questionnaire evaluation indicated that our participants found the general idea of using pre-made applications as a start for developing their own VR helpful within one session of the course suitable. However, the usability of the VR forge was criticized by the participants. Specific authoring actions within the tool, such as positioning and moving through the scene, were perceived as counter-intuitive and challenging. It was found acceptable after a short familiarization period for some participants, but the AttrakDiff outcome also supports that these actions can be improved when the tool is used briefly within one teaching session only. The VR simulator provided by the tool was perceived helpful, and our participants indicated that it helped them to get an idea of VR – even without being provided with hardware in this remote emergency course.

4 Conclusion and Future Work

In this paper, we have explored the authoring and usage of pattern-based VR concerning the concept of VR nuggets and their use in teaching Computer Science students. We have underlined that VR nuggets and the VR Forge authoring tool could be used for remote emergency teaching purposes. We have indicated that the VR Forge and its simulator component could provide students with a firsthand contact of VR software without using VR hardware. The absence of experiencing immersive experiences without suitable hardware is surely not optimal. Still, despite the challenges and unforeseen restrictions posed by the COVID-19 pandemic to education and the usability issues of the VR Forge we pointed out, the students indicated that the VR nuggets and their authoring could give them an idea of VR development.

In future work, we will use VR nuggets within courses other than Computer Science and related ones, as pointed out as a valuable option within our user study. Furthermore, particular teaching methodologies guidelines for educators that include pattern-based VR experiences should be explored. For example, how should educators integrate VR nuggets in their courses outside Computer Science? Can the students prepare their own VR content with VR nuggets? How can learning sessions with pattern-based VR experiences be designed? How many VR headsets are needed for such a session apart from remote emergency teaching?

Acknowledgment

The work is supported by the Federal Ministry of Education and Research of Germany in the project Innovative Hochschule (funding number: 03IHS071).

References

- [AH08] Adams, J. C.; Hotrop, J.: Building an economical VR system for cs education. ACM SIGCSE Bulletin, 40(3):148–152, 2008.
- [ASK20] Aouneh, S.; Salha, S.; Khlaif, N. Z.: Designing quality e-learning environments for emergency remote teaching in coronavirus crisis. Interdisciplinary Journal of Virtual Learning in Medical Sciences, 11(2):135–137, 2020.
- [DH21] Dörner, R.; Horst, R.: Conveying Firsthand Experience: The Circuit Parcours Technique for Efficient and Engaging Teaching in Courses about Virtual Reality and Augmented Reality. In Beatriz Sousa Santos and Gitta Domik, editors, Eurographics 2021 - Education Papers, The Eurographics Association, 2021.
- [Dr73] Dretske, F. I.: Perception and other minds. *Noûs*, pages 34–44. Wiley. 1973.
- [Fo20] Fominykh, M.; Wild, F.; Klamma, R.; Billinghurst, M.; Costiner, L. S.; Karsakov, A.; Mangina, E.; Molka-Danielsen, J.; Pollock, I.; Preda, M.; et al.: Developing a model augmented reality curriculum, In Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education, pages 508–509, 2020.
- [Ha18] Hassenzahl, M.; Burmester, M.; Koller, F.: Attrakdiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In Mensch & computer 2003, pages 187–196, Springer, 2003.
- [HD19a] Horst, R.; Dörner, R.: Integration of bite-sized virtual reality applications into pattern-based knowledge demonstration. In Proceedings of the 16th Work-shop Virtual and Augmented Reality of the GI Group VR/AR, pages 137–148. Gesellschaft für Informatik, Shaker Verlag, 2019.
- [HD19b] Horst, R.; Dörner, R.: Virtual reality forge: Pattern-oriented authoring of virtual reality nuggets. In 25th ACM Symposium on Virtual Reality Software and Technology, VRST ’19, New York, NY, USA, 2019, Association for Computing Machinery.
- [HGB10] Holstermann, N.; Grube, D.; Bögeholz, S.: Hands-on activities and their influence on students’ interest. Research in science education, 40(5):743–757, 2010.
- [Ho20] Hodges, C.; Moore, S.; Lockee, B.; Trust, R.; Bond, A.; et al.: The difference between emergency remote teaching and online learning. Educause review, 27:1–12, 2020.
- [Ho21] Horst, R.; Naraghi-Taghi-Off, R.; Rau, L.; Dörner, R.: Design Patterns and Author Roles for Pattern-Based Educational Virtual Reality Experiences. Proceedings of the 18th Workshop Virtual and Augmented Reality of the GI Group VR/AR, in press, 2021
- [Hu05a] Hug, T.: Microlearning: A new pedagogical challenge (introductory note). 2005.

-
- [Hu05b] Hug, T.: Micro learning and narration: exploring possibilities of utilization of narrations and storytelling for the design of “micro units” and didactical micro-learning arrangements. *Proceedings of Media in Transition*, 2005.
 - [Le82] Lewis, C.: Using the thinking-aloud method in cognitive interface design. IBM TJ Watson Research Center Yorktown Heights, NY, 1982.
 - [LKB05] Lee, G. A.; Kim, G. J.; Billingham, M.: Immersive authoring: What you experience is what you get (wyxiwyg). *Communications of the ACM*, 48(7):76–81, 2005.
 - [Ma09] Mayer, R. E.: Constructivism as a theory of learning versus constructivism as a prescription for instruction. *Constructivist instruction: Success or failure*, pages 184–200, Routledge 2009.
 - [SD17] Sousa Santos, B.; Dias, P.: What should a virtual/augmented reality course be? In *Proceedings of the European Association for Computer Graphics: Education papers*, pages 59–62. 2017.
 - [SH96] Stohr-Hunt, P. M.: An analysis of frequency of hands-on experience and science achievement. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 33(1):101–109, 1996.
 - [Ta16] Takala, T. M.; Malmi, L.; Pugliese, R.; Takala, T.: Empowering students to create better virtual reality applications: A longitudinal study of a vr capstone course. *Informatics in Education*, 15(2):287–317, 2016.
 - [Us21] User Interface Design GmbH, AttrakDiff questionnaire, <http://www.attrakdiff.de>, Retrieved: 21.07.2021
 - [WW64] Wilcoxon, F.; Wilcox, R. A.: Some rapid approximate statistical procedures. Lederle Laboratories, 1964.

Institutionelle Rahmenbedingungen für den Einsatz von Virtual Reality als Lerntechnologie

Systematisierung von Einflussfaktoren auf die Implementierung von VR

Andrea Schmitz¹ und Miriam Mulders²

Abstract: Virtuelle Realität in der beruflichen Aus- und Weiterbildung wird zunehmend häufiger und in vielfältigen Szenarien eingesetzt. Daher stellt sich die Frage nach den personellen, organisatorischen und infrastrukturellen Anforderungen für den Einsatz dieser Technologie in den verschiedenen Bildungsinstitutionen. Diese Fragestellung wird im interdisziplinären Forschungsprojekt *HandLeVR* adressiert und untersucht. In diesem Paper wird das Vorgehen im Projekt beschrieben und erste Resultate hinsichtlich institutioneller Rahmenbedingungen werden vorgestellt und diskutiert. Die Ergebnisse zeigen positive Tendenzen bezüglich der Technologieakzeptanz, jedoch sind insbesondere die Hardware- und Software-Ausstattung der Institutionen sowie die Vorerfahrungen innerhalb der Zielgruppen heterogen. Die überbetriebliche Lehrlingsunterweisung erweist sich als geeignetes Szenario für den Einsatz von VR in der beruflichen Aus- und Weiterbildung.

Keywords: Virtual Reality, virtuelle Lernanwendungen, Institutionen der beruflichen Aus- und Weiterbildung, Rahmenbedingungen, Gelingensbedingungen.

1 Einleitung

Virtual Reality (VR) als Bildungstechnologie bietet vielfältige Potenziale zur Optimierung des praxis- und arbeitsplatznahen Lernens in der beruflichen Aus- und Weiterbildung [TMN18]. Als kosten- und ressourcenschonende Ergänzung zum bestehenden Bildungsaltag wird diese Technologie immer öfter in unterschiedlichen Anwendungskontexten erprobt und untersucht. Auch das HPI³ schreibt in seinem Technologiemonitoring der Verwendung von VR eine wachsende Bedeutung für die handwerkliche Aus- und Weiterbildung zu. Trotz der vielseitigen und gewerkeunabhängigen Einsatzmöglichkeiten befindet sich das Handwerk noch in einer Sensibilisierungsphase, sodass die Potenziale von VR-Technologien noch nicht vollständig ausgeschöpft werden [MPS18], [HP20].

¹ Zentralstelle für die Weiterbildung im Handwerk e. V., Lehrgangsentwicklung und Neue Medien, Sternwartstraße 27–29, 40223 Düsseldorf, aschmitz@zwh.de.

² Universität Duisburg-Essen, Lehrstuhl für Mediendidaktik und Wissensmanagement, Universitätsstraße 2, 45141 Essen, miriam.mulders@uni-due.de.

³ Heinz-Piest-Institut für Handwerkstechnik an der Leibniz Universität Hannover.

Vor diesem Hintergrund wird im Forschungsprojekt *HandLeVR*⁴ untersucht, welche Rahmenbedingungen innerhalb der handwerklichen Bildungsinstitutionen für den Einsatz von VR vorliegen und welche Faktoren eine erfolgreiche Implementierung und Nutzung dieser Technologie in der beruflichen Aus- und Weiterbildung begünstigen.

Als Untersuchungsgegenstand dient eine VR-Lernanwendung, die handlungsorientiertes Lernen in der Berufsausbildung im Bereich des Fahrzeuglackierens unterstützt. Die im Projekt entstehende Lernanwendung beinhaltet authentische Aufgaben in Form eines Kundenauftrags, die nach dem empirisch validierten 4C/ID-Modell [MK18] aufbereitet werden. Zusätzlich zur VR-Lernanwendung bietet das dreiteilige Lernszenario noch ein Autorenwerkzeug für Lehrende zur Vorbereitung eigener Lernszenarien sowie eine Reflexionsanwendung für Lehrende und Lernende zum Nachbesprechen der Lernleistung.

Um einen nachhaltigen Transfer der Projektergebnisse zu begünstigen, wird eine niederschwellige, hersteller- bzw. betriebsunabhängige sowie flexible Nutzung des Lernsystems angestrebt. Die Projektergebnisse werden als frei zugängliche Bildungsressource veröffentlicht. Die Prototypen werden durch den betrieblichen Anwendungspartner (Mercedes-Benz Ludwigsfelde GmbH) sowie durch Bildungszentren nationaler Handwerkskammern validiert. Die technischen Anforderungen (bestehend aus einem handelstüblichen VR-System, Gaming-PC sowie einer einfach nachgebauten Lackierpistole)⁵ sind gering. Besonders seitens des Partners der ZWH e. V. wurden im Projektverlauf fortlaufend institutionelle Rahmenbedingungen untersucht, um den Einsatz von VR in der beruflichen Aus- und Weiterbildung auch über das Projektende hinaus zu ermöglichen. Das Paper präsentiert erste Ergebnisse zu den vorliegenden institutionellen Rahmenbedingungen und bietet Empfehlungen für den erfolgreichen Einsatz von VR in der beruflichen Aus- und Weiterbildung.

2 Grundlage und Methoden der Untersuchung

Der Untersuchung der Anforderungen zum Einsatz von VR-Lerntechnologie liegt die Prämissen zugrunde, dass VR als Lernmedium zwar spezielle und neue Anwendungsszenarien eröffnet [SB02], allerdings bei der Implementierung nicht anders zu behandeln ist als andere digitale Lernmedien. Entsprechend kann hier ein Vorgehensmodell aus der gestaltungsorientierten Mediendidaktik verwendet werden, das die Konzeption und Entwicklung von mediengestützten Lernangeboten und -lösungen im Kontext der konkreten Anwendungssituation und der bestimmten Lerninhalte und -ziele, Zielgruppen und Rahmenbedingungen betrachtet [Ke18].

⁴ Handlungsorientiertes Lernen in der VR-Lackierwerkstatt; <https://handlevr.de>.

⁵ Weitere Details zu den technischen Anforderungen können der Projekthomepage entnommen werden:
<https://handlevr.de/projekt/vrlackierwerkstatt/>.

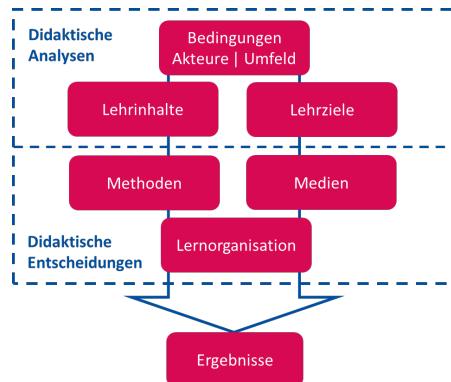


Abbildung 1: didaktische Analysen und Entscheidungen in der Mediendidaktik (entnommen aus Kerres, 2018, S. 229)

Wie in der Abbildung 1 dargestellt, werden grundlegende didaktische Entscheidungen und Analysen vorgenommen, die einerseits Bedingungsfaktoren wie vorgegebene und nicht gestaltbare institutionelle Voraussetzungen erfassen und andererseits Hilfestellungen für die Auswahl der Lehrinhalte, -methoden und -medien geben [Ke18].

Die Entscheidungsfaktoren beziehen sich auf die Auswahl des Lehrinhalts, der Lehrmethode sowie der Medien, die im Rahmen der didaktischen Konzeption und der technischen Umsetzung der VR-Lackierwerkstatt miteingeflossen sind und die Ausgangsbasis für die Analyse der Rahmenbedingungen bilden. Dabei werden didaktische Aspekte hinsichtlich der Lerninhalte und der Erreichung der Lehr- und Lernziele aus dem Blickwinkel der Mediendidaktik gesondert untersucht. Im Rahmen des vorliegenden Papers werden sowohl (*strukturelle*) *Bedingungen, Erfahrungen und Einstellungen der Zielgruppen* (Lehrende und Lernende) sowie die *Lernorganisation* als auch die *technische Ausstattung* erfasst und analysiert.

Zur Erhebung der oben beschriebenen Rahmenbedingungen wurden verschiedene Maßnahmen durchgeführt: Im ersten Projektjahr wurden intensive Recherchen zu den institutionellen Strukturen und Grundlagen der beruflichen Aus- und Weiterbildung betrieben, wie z. B. zu beteiligten Bildungsstätten und ihrer Organisationsstruktur sowie zu den Grundlagen der inhaltlichen Unterrichtsgestaltung (Rahmenlehrpläne etc.). Zur Ermittlung der technischen Ausstattung in den Bildungszentren flossen die Ergebnisse aus dem Projekt *ARSUL*⁶ im Bereich Sanitär-Heizung-Klima ein [AR19]. Ausgehend von dem in *ARSUL* genutzten Fragebogen zur technischen Ausstattung wurden die technischen Rahmenbedingungen mit Vertreterinnen und Vertretern im Maler- und Lackiererhandwerk erneut erhoben. Der Fragebogen wurde über das Lernmanagementsystem (LMS) der ZWH e. V. an die Bildungszentren von sieben Handwerkskammern aus Mittel-, Nord- und Ostdeutschland distribuiert. Die dabei

⁶ <http://projekt-arsul.de/>

gewonnenen quantitativen Daten wurden mit Excel statistisch ausgewertet. Darüber hinaus wurden im Sommer 2020 fünf Leitfadeninterviews mit Lehrkräften der beteiligten Bildungszentren sowie des Praxispartners durchgeführt. Die Interviews wurden (teil-) transkribiert und durch eine qualitative Inhaltsanalyse [Ku18] ausgewertet. Mithilfe der Interviews konnten Daten über das aktuelle Lernszenario inklusive der Nutzung von digitalen Medien im Unterricht erhoben werden. Die Ergebnisse der Interviews flossen in eine Analyse der Zielgruppe Auszubildende ein. Ferner wurden weitere Daten zur Zielgruppenanalyse durch einen Fragebogen im Rahmen der ersten formativen Evaluierung der VR-Lackierwerkstatt erhoben. Dabei lag der Fokus auf der Ermittlung der Vorerfahrung mit und Akzeptanz von VR-Lernanwendungen. Der dabei verwendete Fragebogen basiert auf Vorarbeiten des Forschungsprojekts *ARSUL* sowie einer Studie von Manis und Choi [MC18]. Abgefragt wurden Erfahrungen mit virtuellen Welten sowie mit Computern bzw. Handyspielen im privaten und beruflichen Umfeld. Zusätzlich wurde die Nutzungsintensität (z. B. regelmäßig) und die Art der VR-Anwendung (z. B. Abenteuerspiele) abgefragt. Zur Ermittlung der Technologieakzeptanz wurde ein weit verbreitetes Instrument (TAM, siehe Abschnitt 3.2) genutzt. Die dazugehörige Erhebung wurde mit weiteren Bildungsverantwortlichen der Handwerkskammern sowie den Auszubildenden des Praxispartners durchgeführt.

Untersuchte Dimension	Erhebungsquelle
(Strukturelle) Bedingungen	<ul style="list-style-type: none"> - Sichtung der gesetzlichen Grundlage - Strukturen und Institutionen der handwerklichen Ausbildung
Erfahrungen und Einstellungen der Zielgruppen	<ul style="list-style-type: none"> - Befragung der Dozierenden und Auszubildenden zur Vorerfahrung, Technologieakzeptanz - Leitfadeninterviews
Lernorganisation	<ul style="list-style-type: none"> - Sichtung Rahmenlehrpläne und Unterweisungspläne aus dem Bereich Fahrzeuglackierung - Leitfadeninterviews

Tab. 1: Übersicht Datenquellen der untersuchten Dimensionen

Die folgenden Ausführungen beschreiben die Ergebnisse der einzelnen Untersuchungsgegenstände aus den Recherchen, Interviews und Befragungen. Bei den vorgestellten Resultaten handelt es sich um ein Zwischenergebnis (Dezember 2020). Im Projektverlauf sind weitere Datenerhebungen geplant.

3 Rahmenbedingungen für den Einsatz der VR-Lernszenarien im Bildungskontext

3.1 (Strukturelle) Bedingungen

Die Ausgangsbasis der Untersuchung ist eine Analyse der Bedingungen, in denen eine Lernlösung implementiert wird. Neben den beteiligten Personen und den verfügbaren Ressourcen sind im institutionellen Kontext strukturelle Begebenheiten relevant [SE03]. Die dreijährige Ausbildung zum/zur Fahrzeuglackierer/in wird im dualen System realisiert und findet hauptsächlich im Betrieb und in der Berufsschule statt. Ein weiterer Bestandteil der dualen Ausbildung im Handwerk ist die überbetriebliche Lehrlingsunterweisung (ÜLU) als Teil der betrieblichen Ausbildung [AW17]. Sie findet an überbetrieblichen Ausbildungsstätten (ÜBS) statt. Getragen werden die ÜBS u. a. von Kammern und Innungen. Auf Basis von flexiblen und gestaltungsoffenen Unterweisungsplänen⁷ werden die Lehrgänge von den Lehrkräften vor Ort ausgestaltet. Dadurch können regionale und betriebliche Besonderheiten berücksichtigt werden, um ein einheitliches Niveau sicherzustellen und neue technologische und wirtschaftliche Entwicklungen in der Ausbildung zu berücksichtigen [Fr19]. Die Gestaltungsoffenheit der ÜLU dient als erster Ansatzpunkt für die Erprobung von VR in der beruflichen Aus- und Weiterbildung und kann perspektivisch auf andere Lernorte ausgeweitet werden. Die beschriebene Verteilung der Lernorte ist typisch und lässt sich auf weitere Ausbildungsberufe übertragen.

3.2 Erfahrungen und Einstellungen innerhalb der Zielgruppen (Lehrende und Lernende)

Eine Zielgruppenanalyse beschreibt demografische Daten, Vorwissen und Motivation sowie Einstellungen und Erfahrungen der beteiligten Personen in Bezug auf die verwendete Lerntechnologie [Ke18]. Zentral ist dabei die Akzeptanz von Lerntechnologien, da sie im Kontext von Industrie 4.0 als kritische Größe zählt [PZ18]. Studien belegen, dass die Nutzung einer Technologie nicht automatisch mit deren Verfügbarkeit einhergeht und somit eine Analyse von hemmenden und fördernden Faktoren zur Akzeptanz und Nutzung dieser Technologie notwendig ist, damit deren Potenziale ausgeschöpft werden können [VM03], [DGW89].

Das prominenteste Modell zur Technologieakzeptanz (TAM) stammt von Davis. In seinen Arbeiten identifiziert er zwei Hauptindikatoren der Akzeptanz: zum einen die wahrgenommene Nützlichkeit des Systems zur Verbesserung der eigenen Arbeitsleistung, zum anderen die wahrgenommene Benutzerfreundlichkeit in Bezug auf die leichte Bedienbarkeit des neuen Systems [Da89]. Das TAM liefert die Basis für Untersuchungen in unterschiedlichen Kontexten, wobei Einflussfaktoren erweitert und angepasst werden

⁷ Diese entstehen in Zusammenarbeit zwischen den Fachverbänden des Handwerks und dem Heinz-Piest-Institut (HPI).

[KH06], [LKL03]. So wurden auch bei diesem Forschungsvorhaben unterschiedliche Einflusskriterien für die Technologieakzeptanz bei den Bildungsverantwortlichen und Dozierenden der Handwerkskammern ($N=11$) sowie bei den Auszubildenden des Praxisbetriebs ($N=14$) erhoben und zwischen den Zielgruppen verglichen. Die Fragebogenitems wurden dabei mithilfe einer fünfstufigen Likert-Skala von „*Stimme überhaupt nicht zu*“ bis „*Stimme voll zu*“ bewertet. Untersucht wurden die Faktoren wahrgenommene Nützlichkeit ($M = 3.28$, $SD = .65$), wahrgenommene Benutzerfreundlichkeit ($M = 3.82$, $SD = .59$), Verwendungsabsicht ($M = 3.79$, $SD = .62$) sowie wahrgenommener Genuss ($M = 4.18$, $SD = .59$). Die Ergebnisse zeigen insgesamt eine positive Bewertung der Technologie, wobei der Faktor Genuss die höchsten Werte erhielt. Auch in anderen Studien gilt Genuss als wichtiger Faktor [AW16]. Er beschreibt, inwieweit die Nutzung der Technologie als angenehm bzw. unterhaltsam empfunden wird – unabhängig von dem Nutzen bzw. der Leistung [MC19]. Ein Vergleich zwischen Auszubildenden und Bildungsverantwortlichen der Handwerkskammern zeigt keine signifikanten Unterschiede hinsichtlich der Akzeptanz. Lediglich die Benutzerfreundlichkeit wird von den Auszubildenden höher eingestuft, jedoch immer noch knapp im nicht signifikanten Bereich.

Die Untersuchung der Vorerfahrung der Beteiligten weist darauf hin, dass bislang wenige Kontaktpunkte mit VR-Technologien vorhanden sind, was einen möglichen Schulungsbedarf innerhalb der Zielgruppen aufzeigt. Des Weiteren begünstigt die Vorerfahrung die positive Wahrnehmung der Technologie sowie deren tatsächlichen Einsatz, wie es im Forschungsprojekt *VASE*⁸ aufgezeigt wurde [PZ18].

Zuletzt sind unterstützende Bedingungen zentrale Erfolgskriterien für die Implementierung von VR-Lerntechnologien. Sie werden im Rahmen der Unified Theory of Acceptance and Use of Technology (UTAUT) als Grad der Wahrnehmung einer organisatorischen und technischen Infrastruktur zur Unterstützung beschrieben [VM03]. Dazu zählen im Kontext der institutionellen Bildung neben Ressourcen, Zeit und technischer Ausstattung auch der Aufbau von Medienkompetenzen bei den Lehrkräften und die Verfügbarkeit von technischen Supportstrukturen [Ba16], [Kr11], die im Rahmen der Bereitstellung der Projektergebnisse, in Form von Schulungskonzepten und Leitfäden für die Nutzung und Wartung des Lernsystems, mitaufgebaut werden müssen.

Das Konzept zur Befähigung von Lehrenden baut auf dem Europäischen Rahmen für die digitale Kompetenz Lehrender (DigCompEdu) auf [RP17]. Neben Materialien für die Nutzung der VR-Lackierwerkstatt wird ebenfalls ein Schulungskonzept zur allgemeinen Nutzung von VR-Lernszenarien erarbeitet. Die Konkretisierung der Materialien ist für das letzte Transferjahr geplant.

⁸ Virtual and Analytics Service im Maschinen- und Anlagenbau.

3.3 Lernorganisation

Die Lernorganisation beschreibt die Organisation und Durchführung des Lernangebots sowie das Arrangement der Lernelemente. Analysiert werden die zeitlichen, räumlichen und sozialen Komponenten [Ke18].

Zeitliche Organisation	Räumliche Organisation	Soziale Organisation
• Zeitraum der Nutzung	• Präsenz- und Onlinephasen	• Individuelles Lernen, mit anderen lernen, in einer Gruppe oder Gemeinschaft lernen
• Gesamte Lernzeit des Angebotes	• Lernorte und technische Ausstattung	
• Startpunkte (u. a. feste Termine, flexibler Einstieg)	• Durchführung von rechtssicheren Prüfungen	• Betreuung der Teilnehmenden
• Geplante Anzahl der Durchgänge		
• Taktung		

Tabelle 1: Übersicht Lernorganisation in der Mediendidaktik nach Kerres [Ke18]

Im Anwendungsfall der VR-Lackierwerkstatt wird mit der Analyse der Lernorganisation in den Bildungszentren ein Abgleich zwischen aktuellem Lernszenario und geplanten VR-Szenarios vorgenommen.

Die ÜLU wird im Rahmen von Lehrgängen mit bis zu zwölf Teilnehmenden durchgeführt, wobei Planung und Taktung der Lehrgänge den Bildungszentren obliegen und durchaus unterschiedlich gestaltet sind. Sie dauern in der Regel ein bis zwei Arbeitswochen. Zeitpunkte der Lehrgänge werden bedarfsgerecht für die Lehrjahre definiert. Während die zeitliche Organisation variiert, ist die Unterrichtsgestaltung im Sinne der sozialen Organisation eher homogen. Zur Förderung der beruflichen Handlungskompetenz gemäß den Rahmenlehrplänen steht die Bearbeitung von Kundenaufträgen im Zentrum der ÜLU. Dies ermöglicht, dass „*Handlungen (...) von den Lernenden möglichst selbstständig geplant, durchgeführt, überprüft, ggf. korrigiert und schließlich bewertet werden*“ [KMK03]. Dabei findet die Ausführung nicht in Gruppen statt, sondern wird einzeln bearbeitet. Die Rolle der Dozierenden ist begleitend. Digitale Medien werden zur Unterstützung als Informationsquelle oder zur Dokumentation eingesetzt. Die Gestaltung und Bereitstellung digitaler Medien ist durchaus heterogen.



Abbildung 2: Darstellung des aktuellen Lernszenarios in der ÜLU

In der räumlichen Aufteilung sind die Bildungszentren meist ähnlich aufgebaut: Neben Schulungsräumen stehen Werkstätten und Lackierkabinen zur Verfügung. Zusätzlich gibt es vereinzelt PC-Räume.

Im Rahmen der Workshops und Interviews mit den Dozierenden wurden erste Lernszenarien erarbeitet, in denen VR implementiert wird. Unter anderem bietet sich der Einsatz während Wartezeiten an, die bedingt durch begrenzte räumliche Kapazitäten (z. B. in der Lackierkabine) und individuelles Arbeitstempo zustande kommen. Während der Wartezeit können die Auszubildenden entweder allein oder in Kleingruppen Aufgaben in der VR-Lackierwerkstatt bearbeiten. Ein weiteres Einsatzszenario ist die Bearbeitung zu zweit, in dem eine Person die VR-Anwendung nutzt und die zweite beobachtet. Die Dozierenden betonten, dass die VR-Lackierwerkstatt kein Ersatz für die bestehende Ausbildungsumsetzung darstellt, sondern als Ergänzung eingesetzt werden soll.

3.4 Technische Ausstattung

Neben mangelnder Medienkompetenz der Lehrkräfte zählen IT-Ausstattung und technischer Support zu den wesentlichen hemmenden Faktoren für den Einsatz von digitalen Bildungsformaten [Ba16], [TMN18]. Dementsprechend wurden im Projekt Daten über den aktuellen Stand der technischen Ausstattung erhoben. Wie auch im Projekt *ARSUL* zeigt sich, dass die technische Ausstattung in den verschiedenen Bildungszentren heterogen ist [AR19]. Letzteres resultiert aus dezentralen Trägerschaften. Dabei werden die Kosten zur Unterhaltung und Nutzung der ÜBS sowohl durch die Beiträge der Betriebe als auch durch öffentliche finanzielle Förderung getragen. Bildungseinrichtungen können die Beschaffung von technischer Ausstattung entweder über Eigenmittel oder Förderprogramme finanzieren⁹ [Wi20]. Unterschiede bezüglich der technischen Ausstattung zeigen sich insbesondere bei der Nutzung und Bereitstellung von digitalen Materialien für die Lehrgänge – angefangen bei der Nutzung eigener Endgeräte bis hin

⁹ Hier sind in den letzten Jahren insbesondere für den Aufbau und die Modernisierung der ÜBS verschiedene Programme ins Leben gerufen worden.

zur Bereitstellung von Rechnerpools. Auch die Distribution der Materialien erfolgt über verschiedene Kanäle: Apps, Cloud-Dienste oder lokale Server. Die klassischen Lernmanagementsysteme sind zwar größtenteils vorhanden, finden aber speziell im Bereich Fahrzeuglackierung selten Anwendung.

4 Zusammenfassung

Das Paper beschreibt die Untersuchung der institutionellen Rahmenbedingungen für das Projekt *HandLeVR* und präsentiert erste Resultate. Hierfür wurden verschiedene Faktoren aus dem mediendidaktischen Vorgehensmodell [Ke18] analysiert:

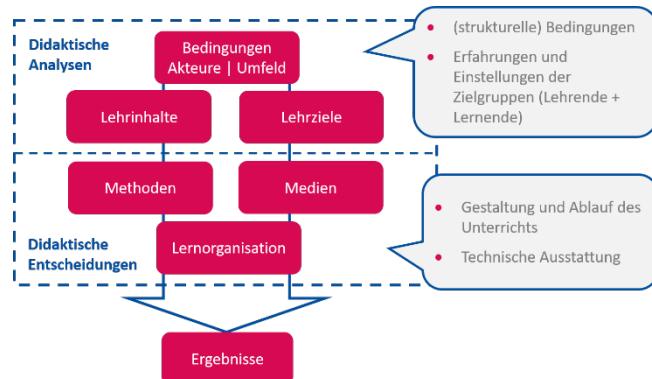


Abbildung 3: untersuchte Faktoren der institutionellen Rahmenbedingungen in Anlehnung an Kerres' mediendidaktische Analysen [Ke18]

Die erhobenen Erkenntnisse hinsichtlich der strukturellen Bedingungen, der Zielgruppe, der Unterrichtsgestaltung und der technischen Ausstattung der ÜLUs indizieren, dass die derzeitigen infrastrukturellen Rahmenbedingungen in deutschen Bildungsinstitutionen der beruflichen Aus- und Weiterbildung einen Einsatz von VR-Lernszenarien ermöglichen, zumindest im Bereich des Fahrzeuglackierens. Die Gestaltungsoffenheit der ÜLUs bietet einen Ansatzpunkt, um die VR-Lackierwerkstatt nach Projektende im Berufsalltag ergänzend einzusetzen und langfristig zu erproben. Seitens der Lehrkräfte besteht Akzeptanz und Bereitschaft, das Medium zu integrieren. Limitiert werden die Ergebnisse durch eine kleine Stichprobe und die Erhebung in der Entwicklungsphase.

Offen bleibt die Frage, wie VR in die Infrastruktur der Bildungsstätten einzubetten ist. Es ist zu klären, wie Nutzungsdaten und -ergebnisse gespeichert und welche Schnittstellen zu vorhandenen Systemen (z. B. LMS) aufgebaut werden. Die Verknüpfung mit einem Verwaltungssystem zur Synchronisation von Nutzungsdaten sowie eine zentrale Speicherung von Lerndaten wäre wünschenswert. Die Verwendung von standardisierten Schnittstellen (z. B. LTI) sollte geprüft werden. Weitere Faktoren wie Datenschutz oder

Hygienestandards spielen eine Rolle. Ebenso sollen neben dem Schulungskonzept handlungsweisende Materialien zum adäquaten Umgang mit und in VR erstellt werden.

Die vorliegende Untersuchung bezieht sich auf den Bereich Fahrzeuglackierung. Das duale Ausbildungssystem ist in Deutschland verbreitet und für einen Großteil der Ausbildungsberufe gültig [MAB20]. Bildungszentren sind oft multifunktional mit mehreren Gewerken in einem Haus ausgerichtet und obliegen derselben Infrastruktur. Somit sind die gewonnenen Erkenntnisse auf andere handwerkliche Berufe und Institutionen übertragbar.

Insgesamt skizzieren die bisherigen Ergebnisse ein optimistisches Bild für den Einsatz von VR in der beruflichen Aus- und Weiterbildung. Es bedarf jedoch weiterer Forschung und vor allem Erprobung im Berufsalltag, um infrastrukturelle Schwierigkeiten bei der Implementierung von VR-Lernszenarien zu identifizieren. Auch die Ergebnisse dieser Untersuchung sind ein Zwischenstand, der im Entwicklungsprozess eines VR-Lernszenarios entstanden ist. Innerhalb der Projektlaufzeit kann lediglich das System entwickelt und ansatzweise etabliert werden. Infrastrukturelle Schwierigkeiten werden erst im laufenden Ausbildungsbetrieb ersichtlich, was nicht mehr in der Projektlaufzeit umsetzbar ist. An dieser Stelle wird die Problematik von zeitlich gebundenen Forschungsprojekten, die einen langfristigen Transfer von funktionierenden Lernanwendungen im Bildungsalltag anstreben, deutlich.

Literaturverzeichnis

- [AR19] ARSUL: Bericht über organisatorische Rahmenbedingungen des Lernens und Lehrens, Düsseldorf, 2018.
- [AW16] Abdullah, F.; Ward, R.: Developing a General Extended Technology Acceptance Model for E-Learning (GETAMEL) by analysing commonly used external factors. *Computers in Human Behavior*, 56, S. 238–256, 2016.
- [AW17] Asmuth, S.; Werner, U.: Duale Berufsausbildung sichtbar gemacht. 3. aktualisierte und erweiterte Auflage, Bundesministerium für Bildung und Forschung, Bonn, 2017.
- [Ba16] Bach, A.: Nutzung von digitalen Medien an berufsbildenden Schulen – Notwendigkeit, Rahmenbedingungen, Akzeptanz und Wirkungen. In Seifried, J.; Seeber, S. et al. (Hrsg.): *Jahrbuch der Berufs- und wirtschaftspädagogischen Forschung*, S. 107–123, 2016.
- [Da89] Davis, F.: Perceived Usefulness, Perceived Ease of Use and User Acceptance of Information Technology. *MIS Quarterly*, 13, S. 319–339, 1989.
- [DGW89] Davis, F.; Bagozzi, R. et al.: User Acceptance of Computer Technology: A Comparison of two Theoretical Models. *Management Science*, 35(8), S. 982–1003, 1989.

- [Fr19] Franke, D.: Überbetriebliche Unterweisung im Handwerk. Zahlen und Analysen zur Inanspruchnahme im Jahr 2018. Verein zur Förderung des Heinz-Piest-Instituts für Handwerkstechnik an der Leibniz Universität Hannover e. V., Hannover, 2019.
- [HP20] HPI Technologiemonitoring: Steckbrief AR/VR, https://hpi-hannover.de/veroeffentlichungen/Steckbriefe/HPI-Technologie-Steckbrief_ARVR.pdf, (16.02.2021).
- [Ke18] Kerres, M.: Mediendidaktik. Konzeption und Entwicklung digitaler Lernangebote. 5. Auflage, De Gruyter, Berlin, 2018.
- [KH06] King, W.; He, J.: A meta-analysis of the Technology Acceptance Model. *Information & Management*, 43, S. 740–755, 2006.
- [KMK03] Rahmenlehrplan für den Ausbildungsberuf Fahrzeuglackierer/Fahrzeuglackiererin, Beschluss der Kultusministerkonferenz vom 16.05.2003, 2003.
- [Kr11] Kreidl, C.: Akzeptanz und Nutzung von E-Learning-Elementen an Hochschulen. Gründe für die Einführung und Kriterien der Anwendung von E-Learning. Medien in der Wissenschaft, Waxmann Verlag GmbH, Münster, 2011.
- [Ku18] Kuckartz, U.: Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung. 4. Auflage. Beltz Juventa, Weinheim, 2018.
- [LKL03] Lee, Y.; Kozar, K. et al: The Technology Acceptance Model: Past, Present, and Future. *Technology*, 12, 2003.
- [MAB20] Maaz, K.; Artelt, C.; Brugger, P.; Buchholz, S.; Kuehne, S.; Leerhoff, H.; Rauschenbach, T.; Rockmann, U.; Roßbach, H.-G.; Schrader, J.; Seeber, S.: Bildung in Deutschland 2020. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung in einer digitalisierten Welt, 2020.
- [MC19] Manis, K.; Choi, D.: The virtual reality hardware acceptance model (VR-HAM): Extending and individualizing the technology acceptance model (TAM) for virtual reality hardware. *Journal of Business Research*, 100, S. 503–513, 2019.
- [MK18] Van Merriënboer, J. J. G.; Kirschner, P. A.: 4C/ID in the Context of Instructional Design and the Learning Sciences. *International Handbook of the Learning Sciences*, S. 169–179, 2018.
- [MPS18] Mahrin, B.; Pfetsch, J.; Stoll, C.: Mobiles Lernen im Handwerk. *Handbuch Mobile Learning*. Springer VS, Wiesbaden, S. 943–970, 2018.
- [PZ18] Pletz, C.; Zinn, B.: Technologieakzeptanz von virtuellen Lern- und Arbeitsumgebungen in technischen Domänen. *Journal of Technical Education (JOTED)*, 6(4), S. 86–105, 2018.
- [RP17] Redecker, C.; Punie, Y. (2017): European Framework for the Digital Competence of Educators: DigCompEdu; deutsche Übersetzung durch das Goethe-Institut 2019.
- [SE03] Seufert, S.; Euler, D: Nachhaltigkeit von eLearning-Innovationen. SCIL, 2003.
- [SB02] Schwan, S.; Buder, J.: Lernen und Wissenserwerb in virtuellen Realitäten. In Bente, G.; Krämer, N. C.; Petersen A. (Hrsg.): *Virtuelle Realitäten*, S. 109–129, Hogrefe Verlag für Psychologie, Göttingen, 2002.

- [TMN18] Thomas, O.; Metzger, D.; Niegemann, H.: Digitalisierung in der Aus- und Weiterbildung: Virtual und Augmented Reality für Industrie 4.0. Springer-Verlag, 2018.
- [VM03] Venkatesh, V.; Morris, M. et al: User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27, S. 425–478, 2003.
- [Wi20] Wittig, W.: Evaluation der Förderung überbetrieblicher Berufsbildungsstätten und ihrer Weiterentwicklung zu Kompetenzzentren, Bundesinstitut für Berufsbildung, Bonn, 2020.

Drei Eigenschaften von Augmented Reality-Erfahrungen und ihre Relevanz beim Lernen

Jule M. Krüger ¹, Daniel Bodemer  ²

Abstract: Augmented Reality (AR) bietet viele Möglichkeiten, Lernprozesse durch die Darstellung virtueller Elemente in Kombination mit realen Elementen anzuregen. Für die experimentelle Erforschung der zugrundeliegenden Wirkmechanismen ist jedoch eine Systematisierung verschiedener AR-spezifischer Faktoren und Erfahrungen sinnvoll. In diesem Beitrag werden Kontextualität, Interaktivität und Räumlichkeit als drei Eigenschaften von AR-Erfahrungen aus Perspektive der Nutzenden beschrieben und ihre Anwendungsmöglichkeiten in Lernsettings erläutert. Weiter werden verschiedene Studien aus einer eigenen Forschungsreihe und ihre Ergebnisse, die Erkenntnisse über das Potenzial dieser Eigenschaften in Lernsettings liefern, beschrieben. Zum Abschluss werden mögliche Themen für zukünftige Studien in diesem Bereich benannt.

Keywords: Augmented Reality, Interaktivität, Kontextualität, multimediales Lernen, Räumlichkeit, technologiegestütztes Lernen

1 Einleitung

Augmented Reality (AR) ist eine Form der Informationsvisualisierung, in der virtuelle und reale Elemente kombiniert werden, so dass informatives Material unterschiedlich dargestellt werden kann. Im Lernkontext kann instruktionsgebendes oder -unterstützendes virtuelles Material sowohl in informelle (z. B. Waldspaziergang) als auch in formelle (z. B. Schulexkursion) Lernkontakte in der echten Welt eingebettet werden. Auch können virtuell nachgebaute Elemente aus der Außenwelt oder dem Anwendungsfeld in instruktionale Umgebungen eingebettet werden und Lernen sowohl in formellen (z. B. Universität) als auch informellen (z. B. Museum) Lernkontexten unterstützen. Die Erkenntnisse aus der Forschung deuten darauf hin, dass AR grundsätzlich positiven Einfluss z. B. auf Lernmotivation, -prozesse und -ergebnisse haben kann [z. B. AA17, CT13, DD14, Ra14, Wu13]. Um AR-Systeme genauer zu definieren, formulierte Azuma (1997) drei Eigenschaften: sie kombinieren echte und virtuelle Elemente, sie sind interaktiv in Echtzeit, und sie registrieren die Objekte in drei Dimensionen (3D) [Az97]. Da in AR nicht die Realität selbst, sondern die Wahrnehmung dieser erweitert wird [HFN11], besteht eine direkte Interaktion mit der menschlichen Erfahrungswelt und ist es wichtig, AR in Forschung und Entwicklung nicht nur aus technologischer Sicht, sondern auch aus Sicht der nutzenden Personen

¹ Universität Duisburg-Essen, Medienbasierte Wissenskonstruktion, Lotharstr. 65, 47057 Duisburg,

jule.krueger@uni-due.de,  <https://orcid.org/0000-0002-1141-8803>

² Universität Duisburg-Essen, Medienbasierte Wissenskonstruktion, Lotharstr. 65, 47057 Duisburg,

bodemer@uni-due.de,  <https://orcid.org/0000-0003-2515-683X>

zu betrachten. Drei Eigenschaften, die auf Basis der von Azuma beschriebenen technologischen Merkmale von AR-Systemen aus Perspektive der Erfahrung bei der Nutzung von AR formuliert wurden, sind Kontextualität, Interaktivität und Räumlichkeit [KBB19]. Diese spielen in vielen AR-basierten Lernsettings einzeln oder in Interaktion eine Rolle und können helfen, die Forschung weiter zu systematisieren. Forschung und Erkenntnisgewinn in Bezug auf verschiedene Wirkmechanismen im Rahmen dieser Eigenschaften kann relevant sein, um weitere Erkenntnisse für den effizienten und effektiven Einsatz von AR-Erfahrungen in der Bildung zu sammeln. Aus diesem Grund werden im weiteren Verlauf die drei Eigenschaften anhand einer beispielhaften AR-Anwendung erläutert und ihre mögliche Bedeutung für das Lernen mit AR beschrieben. In der beispielhaften Anwendung (s. Abb. 1) können Papier-Marker (echte Objekte), an die virtuelle 3D-Modelle verschiedener Komponenten eines Kraftwerks (virtuelle Elemente) geknüpft sind, zusammengebaut werden. Es können Stromoutput und Wirkungsgrad des aktuell und des vorherig gebauten Kraftwerks angezeigt werden. Die jeweilige Bedeutung der drei Eigenschaften im Kontext dieser Anwendung werden in Abschnitt 2 beschrieben. Im Anschluss werden in Abschnitt 3 erste Studien aus einer Forschungsreihe, die auf Basis dieser Eigenschaften durchgeführt wird, vorgestellt. In Abschnitt 4 wird zum Abschluss eine Zusammenfassung der Ergebnisse und ein Ausblick auf weitere Forschung gegeben.

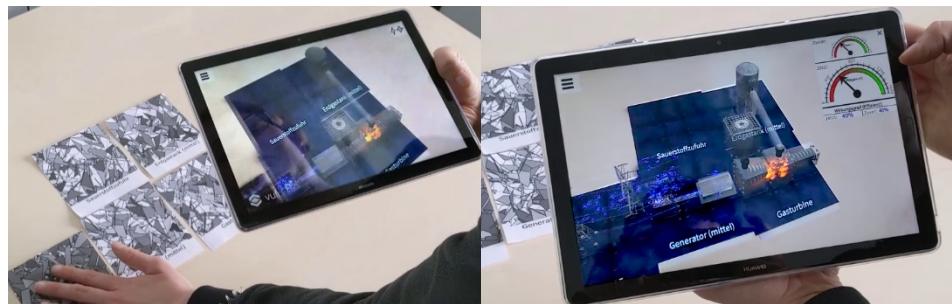


Abb. 1: AR-Anwendung zum Thema Gas-und-Dampf-Kombikraftwerke (links: Interaktion mit Papier-Markern; rechts: Anzeige zum Stromoutput und Wirkungsgrad)

2 Drei Eigenschaften von AR

2.1 Kontextualität

Bei der Eigenschaft *Kontextualität* geht es um die kombinierte Wahrnehmung echter und virtueller Elemente in AR [vgl. KBB19]. Da die virtuellen Elemente den Kontext nicht überdecken (wie z. B. bei VR), ist dieser parallel wahrnehmbar. Auch sind die Realitäten nicht voneinander getrennt (wie z. B. bei normaler Tablet-Nutzung), sondern können integriert dargestellt sein. Somit liegt der Fokus hier auf Merkmalen der realen Welt, in der AR-basiertes Lernen stattfindet, und der thematischen Verbindung dieser mit virtuellen

Elementen. Hierbei können zwei Ebenen realer Elemente unterschieden werden. Einerseits ist das AR-basierte Lernen in der realen Umgebung kontextualisiert, andererseits sind die virtuellen Elemente an realen Elementen (z. B. AR-Marker oder spezifische Formen) verankert. Kontextualität in AR-Erfahrungen kann u.a. durch die thematische Relevanz der realen (Umgebung und Anker) für die virtuellen Elemente oder durch die Wahrnehmbarkeit des Kontexts variieren. Im Beispiel in Abb. 1 haben weder die Anker-Elemente (Papier-Marker) noch die Umgebung (Tisch in einem Raum) direkte thematische Relevanz für die virtuellen Elemente und sind nicht Teil des Lerninhalts. Während die Umgebung einfach wahrgenommen werden kann, werden die Papier-Marker durch die virtuellen Elemente verdeckt. Die Anwendung kann in jeglicher Umgebung mit genügend Fläche genutzt werden. Dies kann basierend auf [Re05] als willkürliche Verknüpfung (arbitrary linkage) beschrieben werden, bei der weder der genaue Ort (particular location), noch die allgemeinen Eigenschaften der Umgebung (physicality) relevant sind.

Kontextualität kann lernförderliches Design von AR-Lernerfahrungen ermöglichen [vgl. KBB19]. Durch die parallele Wahrnehmung virtueller Elemente und echter Umgebung kann das Lernen in relevante Umgebungen eingebettet werden [DD14], so dass Authentizität und Verknüpfung mit der Realität unterstützt werden [Wu13] und instruktionale Unterstützung z. B. im Anwendungsfeld situiert wird [Bo14]. Situiertes Lernen wurde als ein für AR relevantes theoretisches Framework herausgestellt [z.B. CT13, DD14, Bo14, Wu13]. In Bezug auf die Situierung von AR-Erfahrungen in der echten Welt ist aus Perspektive der Lernenden Immersion ein relevantes Konzept [GK17]. Immersion beschreibt das Gefühl, eine bestimmte Umgebung um sich herum zu haben [WS98], wobei in AR speziell Kontext Immersion (context immersion) eine Rolle spielt [Ki13]. Das Konzept der Presence hängt hiermit zusammen und kann in AR als das Gefühl definiert werden, in einer virtuell-real gemischten Umgebung anwesend zu sein [Be12]. Immersion und Presence können einen Einfluss auf das Vergnügen der Lernenden haben [Sy10], was motivational relevant ist. Durch Kontextualität können die mentale Verknüpfung und Integration der virtuellen Elemente, echten Elementen und echten Umgebung unterstützt werden.

2.2 Interaktivität

Die Eigenschaft *Interaktivität* behandelt die Interaktionsmöglichkeiten zwischen Nutzenden und AR-Elementen, die in AR sehr unterschiedlich ausfallen können [vgl. KBB19]. Der Fokus liegt hier somit auf der Verbindung der Nutzenden mit den (interaktiven) Elementen. Hierbei können drei Ebenen der Interaktion Teil einer AR-Erfahrung sein [vgl. KB20]: Lernende können rein mit realen Elementen oder rein mit virtuellen Elementen interagieren, können aber auch in vermittelte Interaktion durch die Manipulation realer Elemente (z. B. AR-Marker) virtuelle Elemente beeinflussen (vgl. tangible interface metaphor [BD12]). Die Interaktion kann dabei mehr oder weniger ausgeprägt sein, so dass z. B. ein einfaches Umrunden oder das Bewegen eines virtuellen Objekts möglich ist, aber auch ein Objekt komplett neu erstellt werden kann. Hierbei kann außerdem mehr oder weniger Bewegung nötig sein und die körperliche Interaktion kann mehr oder weniger relevant für die mentale Interaktion mit den Lerninhalten sein. Im Beispiel in Abb. 1 ist

auf verschiedenen Ebenen verschiedene Interaktion möglich. Hierbei können Lernende durch Verschieben der echten Elemente (Papier-Marker) die virtuellen Elemente (3D-Modelle der Kraftwerkkomponenten) bewegen, welche sich verbinden und aufeinander reagieren können. Auch kann durch Bewegung des Tablets die Perspektive auf die Modelle geändert werden. Außerdem können Lernende im App-Interface mit rein virtuellen Elementen (Button mit Blitz und Plus oben rechts) interagieren, was weitere Informationen über aktuellen Stromoutput und Wirkungsgrad des gebauten Kraftwerks liefern kann.

Auch Interaktivität kann lernförderliches Design von AR-Lernerfahrungen ermöglichen [vgl. KBB19, KB20]. Die Embodied Cognition Theory beschreibt, dass Lernen durch körperliche Interaktion unterstützt werden kann [Wi02]. Auch beschreibt das ICAP (interactive - constructive - active - passive)-Framework, dass offenkundiges, aktives Lernverhalten auf kognitive Verarbeitung des Lernmaterials schließen lässt und diese unterstützen kann, wenn das Verhalten relevant für die Lernaufgabe ist [CW14]. Des Weiteren kann selbstgesteuerte statt passiver Perspektivwechsel um Objekte zu erfolgreicherem Lernen führen [HNS18]. Forschung im Bereich des multimedialen Lernens beschreibt, dass mentale Interaktion bedeutender für das Lernen ist als körperliche Interaktion [CM16], so dass es wichtig ist, mit körperlicher Interaktion auch relevante mentale Interaktion anzuregen. In Bezug auf die verschiedenen Möglichkeiten interaktiver AR können aus Perspektive der Lernenden somit körperliche Anforderungen und kognitive Verarbeitung der Inhalte eine große Rolle spielen, insbesondere in Bezug auf lernrelevante und -irrelevante Verarbeitung. Wenn eine körperliche Interaktion nur zu weiterer Belastung, aber nicht zur lernrelevanten Verarbeitung führt, kann diese lernhinderlich sein, statt Lernen zu fördern.

2.3 Räumlichkeit

Bei der Eigenschaft *Räumlichkeit* geht es um die räumliche Darstellung und Anordnung virtueller Elemente, die durch ihre Platzierung in der echten 3D-Welt räumliche Eigenschaften ähnlich derer echter Elemente erhalten können [vgl. KBB19]. Hiermit liegt der Fokus auf den räumlichen Eigenschaften der virtuellen Elemente und der räumlichen Verknüpfung dieser mit realen Elementen. Virtuelle Elemente können reale Elemente nachbilden und scheinen in AR räumliche Eigenschaften und eine Ausdehnung im Raum zu haben. Hier spielt die Dimensionalität eine große Rolle, da z. B. sowohl 2D-Abbildungen als auch 3D-Objekte nachgebaut werden können. Bei der räumlichen Verbindung zwischen realen und virtuellen Elementen ist es wichtig zwischen zwei Ebenen zu unterscheiden. Auf der kleinteiligeren Ebene befindet sich die direkte Verknüpfung von realem Anker-Element (z. B. AR-Marker oder echtes Objekt) und virtuellem Element, durch welche vermittelte Interaktion und eine Beachtung der räumlichen Kontiguität [Ma09] möglich sind. Auf der weitläufigeren Ebene befindet sich die räumliche Platzierung virtueller Informationen an bestimmten Orten der realen Welt, was oft mit thematischer Relevanz einhergeht. Im Beispiel in Abb. 1 äußert sich die Räumlichkeit sowohl in der direkten räumlichen Koppelung der virtuellen Modelle mit den Papier-Markern als auch der Dreidimensionalität der virtuellen Modelle, die Platz im Raum einzunehmen scheinen.

Auch Räumlichkeit kann lernförderliches Design von AR-Lernerfahrungen ermöglichen [vgl. KBB19]. Die räumliche Integration echter und virtueller Elemente kann hierbei durch räumliche Kontiguität und somit die Nähe zueinander gehörender Elemente mentale Integrationsprozesse unterstützen [Ma09]. Hier spielt insbesondere auch die Cognitive Load Theory eine Rolle, welche kognitive Belastung in Intrinsic, Extraneous und Germane Cognitive Load unterteilt [SVP98, SVP19]. Die Theorie postuliert, dass lernförderliche, germane Prozesse vom Lernmaterial gefördert, während lernhinderliche, extraneous Prozesse vermieden werden sollten. Auch Darstellung in 3D kann eine wichtige Rolle bei der Lernunterstützung spielen [vgl. KB21, KPB21]. Hierbei können 3D-Objekte realistischer und mit mehr Tiefenhinweisen (z. B. bewegungsbasiert, [vgl. Cr13]) dargestellt werden, was authentischeres Lernen mit virtuellen Elementen ermöglicht. Die Nutzung von 3D-Modellen kann zum Aufbau vollständigerer mentaler Modelle führen [CHS15], da nicht erst mental eine 2D- in eine 3D-Repräsentation umgewandelt werden muss. In Bezug auf die verschiedenen Möglichkeiten der Nutzung von Räumlichkeit in AR sind hier somit aus Perspektive der Lernenden kognitive Verarbeitung und Belastung relevant.

3 Forschung zu AR-basierten Lernerfahrungen

In der AR-Forschung werden häufig traditionelle und AR-basierte Lernsettings verglichen, so dass es u.a. durch mögliche Konfundierung verschiedener Variablen häufig nicht möglich ist, Rückschlüsse in Bezug auf die Wirkmechanismen spezifischer Attribute, Faktoren und Designentscheidungen von AR-Lernerfahrungen zu ziehen (s. Kritik an Medienvergleichsstudien nach Surry und Ensminger [SE01]). Um aussagekräftige Grundlagen und Richtlinien für die Entwicklung und den Einsatz effektiver und effizienter AR-Erfahrungen zu erhalten, sollten AR-spezifische Wirkmechanismen, (Design-)Elemente, und der Einfluss relevanter Personenvariablen systematisch, experimentell untersucht werden. Hierbei sollte bei der Analyse welche AR-Eigenschaften welche Rolle spielen könnten auch das Lernziel beachtet werden. Sollen Lernende z. B. Wissen erlangen, welches in direkter Verbindung zur echten Welt steht, ist hier vermutlich die Kontextualität ein wichtiger Faktor, während für das Erlernen räumlicher Strukturen die Räumlichkeit relevant scheint. Hierbei können die drei beschriebenen Eigenschaften von AR-Erfahrungen für die Strukturierung verschiedener AR-spezifischer Faktoren genutzt werden, um die Erforschung zu unterstützen. Im Folgenden werden beispielhaft erste Studien aus einer eigenen Forschungsreihe, die auf Basis der drei Eigenschaften durchgeführt wird, vorgestellt.

3.1 Forschung zu Kontextualität

In verschiedenen Studien wurde gezeigt, dass relevanter Kontext positive Effekte auf Lernprozesse und -ergebnisse haben kann [z. B. GK21, Ka13]. Eine in Bezug auf Kontextualität in dieser Forschungsreihe durchgeführte Studie fokussierte deswegen spezifisch die Relevanz der echten Umgebung einer AR-Erfahrung. Hierbei sollte Lernen mit einer AR-Anwendung an mehr und weniger relevanten Orten verglichen werden, was jedoch

auf Grund der Corona-Pandemie nicht vor Ort möglich war. Zur Untersuchung relevanter Wirkmechanismen wurde die Studie online mit fotobasierter simulierter AR umgesetzt und die Forschungsfrage lautete „Wie beeinflussen Relevanz des Kontexts und räumliche Integration virtueller und echtweltlicher Elemente Immersion und Lernergebnis?“. In Bezug auf die Relevanz des Kontexts waren Annahmen in dieser Studie, dass eine höhere Relevanz zu höherer Immersion und Wissen führen würde. In einem 2x2 Between-Subjects Design wurden Relevanz des Kontexts (Hintergrund relevant: Baum in der Natur; nicht relevant: Baum auf einem Arbeitsblatt) und räumliche Anordnung (Darstellung integriert: see-through Tablet vor dem Baum; getrennt: ohne see-through neben Baum) manipuliert. Siehe Abb. 2 für ein Beispiel der fotobasierten Umsetzung. $N = 136$ Personen zwischen 18 und 65 Jahren ($M = 27.83$, $SD = 8.81$) nahmen an der Studie teil; 39 gaben an männlich zu sein, 90 weiblich, 3 divers und 4 machten keine Angabe. Immersion wurde durch den AR Immersion Fragebogen (ARI [GK17]) erfasst und für die Erfassung des Wissens wurde am Ende der Erhebung ein selbsterstellter Wissenstest durchgeführt.

In den Ergebnissen zeigte sich eine signifikante Interaktion in Bezug auf Immersion, die unerwartet die höchste Score in der Bedingung mit nicht relevantem Kontext und geteilter Darstellung und die niedrigste Score bei relevantem Kontext und geteilter Darstellung zeigte. Dies kann u.a. mit der Vertrautheit des Settings zusammenhängen, da AR noch nicht so verbreitet ist. Die Ergebnisse im Wissenstest deuten ein interessantes Muster an, auch wenn keine signifikanten Effekte gefunden wurden. Fragen, die sich mit einer Verknüpfung der echten und virtuellen Elemente beschäftigten, wurden am besten in der Gruppe mit relevanter und integrierter Darstellung beantwortet, während Fragen, die sich mit Verknüpfungen innerhalb der virtuellen Elemente beschäftigten, am besten in der nicht-relevanten und getrennten Gruppe beantwortet wurden. Bei der Interpretation der Ergebnisse muss jedoch auch die Art der Studie (Simulation von AR) beachtet werden. Eine weitere Umsetzung in einem echten AR-Setting sollte durchgeführt werden, um die Ergebnisse zu überprüfen. Auch die weitere Untersuchung der verschiedenen Arten von Wissen in Bezug auf die Verknüpfung realer und virtueller Elemente scheint relevant.



Abb. 2: Bedingungen der Studie zur Kontextualität (oben: relevant; unten: nicht relevant; links: integriert; rechts: getrennt)

3.2 Forschung zu Interaktivität

In verschiedenen Studien wurde gezeigt, dass Interaktion mit AR-Lernmaterial positive Effekte auf Lernprozesse und -ergebnisse haben kann [z. B. JM17, Li16]. Eine in dieser Forschungsreihe in Bezug auf Interaktivität durchgeführte Studie fokussierte deswegen die verschiedenen Möglichkeiten der Interaktion mit AR-Lernmaterial. In dieser Laborstudie ging es hier spezifisch um den Einfluss mentaler und körperlicher Interaktion, mit Fokus auf die Forschungsfrage „Welche Rolle spielen mentale und körperliche Interaktion mit AR-Lernmaterial für kognitive Belastung und Lernergebnisse?“ [vgl. KB20 für diese Studie]. Es wurde angenommen, dass höhere mentale Interaktion zu höherem Germane Cognitive Load und Wissen führen würde, während höhere körperliche Interaktion hauptsächlich höheren Extraneous Cognitive Load und geringfügig höheren Germane Cognitive Load und Wissen hervorrufen würde. In einem 2x2 Between-Subjects Design wurden mentale Interaktion (hoch: wenige Vorgaben in den Instruktionen; niedrig: viele Vorgaben) und körperliche Interaktion (hoch: eigenständiges Zusammenbauen der Marker; niedrig: Marker schon zusammengebaut) manipuliert. Versuchspersonen beantworteten mithilfe der Anwendung in Abb. 1 vorgegebene Hypothesen zu Kraftwerken. $N = 128$ Personen zwischen 18 und 40 Jahren ($M = 22.55$, $SD = 3.90$) nahmen an der Studie teil; 39 gaben an männlich zu sein, 89 weiblich. Cognitive Load wurde durch einen Fragebogen von Klepsch, Schmitz und Seufert (2017; [KSS17]) erfasst und für die Erfassung des Wissens wurde am Ende der Erhebung ein selbsterstellter Wissenstest durchgeführt.

In den Ergebnissen [vgl. KB20] zeigte sich in Bezug auf das Wissen eine signifikante Interaktion, die eine höhere Score in den Gruppen zeigte, in denen nur entweder mentale oder körperliche Interaktion hoch war, als in den Gruppen, in denen beides hoch oder beides niedrig war. Hieraus kann geschlossen werden, dass sowohl zu viel Unterstützung und somit wenig eigene Aktivität und Unterbelastung als auch zu wenig Unterstützung und somit viel eigene, ungesteuerte Aktivität und Überbelastung zu Schwierigkeiten beim Lernen in einer interaktiven AR-Umgebung führen können. Die Ergebnisse in Bezug auf kognitive Belastung waren in dieser Studie nicht eindeutig. Bei der Beobachtung einiger der Versuchspersonen stellte sich jedoch heraus, dass weniger Unterstützung teilweise zu mehr Fehlern und somit eventuell Miskonzeptionen führte. Vielleicht nahm hier die Konzentration auf die Umsetzung auch Kapazitäten weg, die eigentlich zum tieferen Lernen hätten genutzt werden können. Etwas Unterstützung (mental oder körperlich) kann vermutlich gerade bei der ersten Verwendung einer interaktiven AR-Anwendung hilfreich sein, was auch mit Forschung zu Computersimulationen übereinstimmt [De06].

3.3 Forschung zu Räumlichkeit

In verschiedenen Studien wurde gezeigt, dass die 3D-Darstellung von Objekten positive Effekte auf Lernprozesse und -ergebnisse haben kann [z. B. HP16, SH16], wobei räumliche Fähigkeiten hierbei eine wichtige Rolle spielen können [z. B. KB21, SH16]. Eine in dieser Forschungsreihe in Bezug auf Räumlichkeit durchgeführte Studie fokussierte deswegen die Dimensionalität der Darstellung eines virtuellen Objekts. In dieser Laborstudie

ging es spezifisch um den Vergleich einer 3D- und 2D-Darstellung mit Fokus auf die Forschungsfrage „Wie beeinflusst die Dimensionalität einer Darstellung in AR kognitive Belastung und Lernergebnis, und welche Rolle spielen räumliche Fähigkeiten dabei?“ [eingereichtes Paper: KPB21]. Annahmen in der Studie waren, dass eine 3D-Darstellung zu einem höheren Germane Cognitive Load und Wissen und einem niedrigerem Extraneous Cognitive Load führen würde und dass Lernende mit geringeren räumlichen Fähigkeiten mehr von 3D profitieren würden. In einem 2-Gruppen Design wurde die Dimensionalität der Darstellung eines virtuellen Objekts (3D; 2D) manipuliert. Versuchspersonen erhielten zu einem Text über das menschliche Herz einen 3D- oder eine 2D-Querschnitt in einer tabletbasierten AR-Anwendung (s. Abb. 3). $N = 150$ Personen zwischen 17 und 31 Jahren ($M = 21.98$, $SD = 2.98$) nahmen an der Studie teil; 41 gaben an männlich zu sein, 109 weiblich. Cognitive Load wurde durch den bereits genannten Fragebogen [KSS17] erfasst, räumliche Fähigkeiten durch einen Mental Rotation Test [Pe95], und für die Erfassung des Wissens wurde am Ende der Erhebung ein selbsterstellter Wissenstest durchgeführt.

In den Ergebnissen [KPB21] zeigten sich signifikante Unterschiede bei Germane Cognitive Load und Wissen über die räumlichen Zusammenhänge der Komponenten des Herzens mit höheren Scores in der 3D-Bedingung. Extraneous Cognitive Load unterschied sich nicht signifikant, wies aber deskriptiv in die erwartete Richtung. Unerwartet profitierten Lernende mit hohen räumlichen Fähigkeiten von der 3D-Darstellung, Lernende mit geringen Fähigkeiten jedoch nicht. Hieraus kann geschlossen werden, dass ein gewisser Grad an räumlichen Fähigkeiten nötig sein könnte, um 3D-Darstellungen in AR lernförderlich verarbeiten zu können, während 2D-Darstellungen vermutlich nicht automatisch zu mentalen Transformationen in 3D-Darstellungen führen. Die Ergebnisse einer weiteren, explorativen Pilot-Studie [vgl. KB21] deuten darauf hin, dass verschiedene räumliche Fähigkeiten unterschiedlichen Einfluss auf verschiedene Aufgaben- und Lernergebnisse haben können. Ihre Bedeutung für das Lernen mit AR sollte weiter untersucht werden.

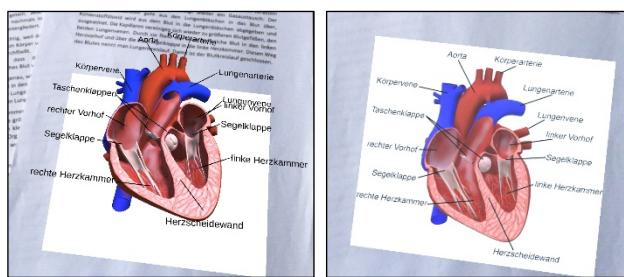


Abb. 3: Bedingungen der Studie zur Räumlichkeit (links: 3D; rechts: 2D)

3.4 Interaktion der Eigenschaften in der Forschung

Die drei beschriebenen Eigenschaften scheinen gerade in ihrem Zusammenspiel Potential für AR-Lernerfahrungen zu haben. Die räumliche Nähe von virtuellen und echten Elementen ist z. B. direkt verknüpft mit der thematischen Nähe und Relevanz dieser. In der

unter 3.1 genannten Studie wurde die Interaktion von räumlicher Nähe und Relevanz der Umgebung untersucht und interessante Interaktionseffekte in Bezug auf Immersion gefunden (s. oben). In einer momentan laufenden Studie wird die Interaktion verschiedener Grade Interaktivität und verschiedener Grade Kontextualität untersucht, um mehr über das Zusammenspiel dieser beiden Variablen zu erfahren. Auch hier sollten in Zukunft weitere Forschungsarbeiten durchgeführt werden.

4 Zusammenfassung und Ausblick

Alles in allem zeigen die drei aus der eigenen Forschungsreihe präsentierten Studien, dass es von Interesse sein kann, auf Basis der drei Eigenschaften Kontextualität, Interaktivität und Räumlichkeit experimentelle Studien mit systematisch manipulierten Variablen durchzuführen. Hierbei sind die Ergebnisse zur Kontextualität aus dieser Forschungsreihe noch nicht sehr aussagekräftig, da eine Umsetzung in einem echten Kontext zur Überprüfung der Ergebnisse noch erfolgen muss. Allgemein zeigt sich in der Literatur, dass AR-Anwendungen, die an einem bestimmten, relevanten Ort eingesetzt werden, zu Lernerfolgen und positiven Erfahrungen der Lernenden führen können [z. B. GK21, Ka13], so dass eine Verfolgung dieses Forschungsstrangs zu weiteren Erkenntnissen zur genauen Umsetzung ortsbezogener AR-Anwendungen in verschiedenen Settings und mit verschiedenen Lernzielen führen kann. In Bezug auf Interaktivität deuten die Ergebnisse an, dass ein bestimmtes Maß an Interaktion zwar lernförderlich sein kann, jedoch nicht alles den Lernenden überlassen werden sollte, sondern gerade bei der ersten Nutzung Hilfestellungen wichtig sein können. Allgemein zeigen sich in der Literatur positive Ergebnisse von Interaktion in AR auf Lernen [z. B. JM17, Li16], jedoch wird weitere Forschung in Bezug auf die genauen Faktoren und Arten von Interaktion, die in interaktiven AR-Anwendungen eine Rolle spielen, benötigt. Die Ergebnisse zur Räumlichkeit zeigen, dass eine 3D-Umsetzung visueller Elemente beim Lernen räumlicher Strukturen relevant sein kann, es jedoch möglich ist, dass die räumlichen Fähigkeiten der Lernenden nicht immer ausreichen, um 3D-Darstellungen zu verarbeiten. Auch in weiteren Studien hat sich der Einfluss räumlicher Fähigkeiten beim Lernen mit 3D-Visualisierungen gezeigt [z. B. KB21, SH16], der jedoch nicht immer eindeutig ist und weiter erforscht werden sollte, um Rückschlüsse für den effektiven Einsatz von AR-Anwendungen ziehen zu können. Auch die mögliche Trainierbarkeit der Fähigkeiten sollte hierbei betrachtet werden. Die Ergebnisse der Studien deuten insgesamt darauf hin, dass Variablen wie kognitive Belastung, Immersion und verschiedene Arten von Wissen durch das verschiedene Design von AR-basierten Lernanwendungen beeinflusst werden können und auch Personenvariablen wie räumliche Fähigkeiten in diesem Zusammenhang betrachtet werden sollten.

Die systematische Manipulation und experimentelle Analyse einzelner, kleinteiliger, AR-spezifischer Faktoren und Attribute ermöglicht die getrennte, unkonfundierte Untersuchung einzelner Wirkmechanismen, die für den Erkenntnisgewinn in der Erforschung AR-basierter Lernerfahrungen entscheidend ist. Im weiteren Verlauf sollten die Ergebnisse aus dieser Forschung in komplexeren, authentischeren Lernsettings geprüft werden, um

eine Übertragbarkeit zu gewährleisten. Zukünftige Studien können hier auch auf die Kombination und Interaktion verschiedener Eigenschaften schauen, da AR-Anwendungen meist auf allen drei Eigenschaften aufbauen und um eine Basis für die sinnvolle Zusammenstellung lernrelevanter Designvariablen und -kriterien für komplexe AR-Anwendungen bereitzustellen. Des Weiteren kann die Erforschung der Beeinflussung verschiedener Lernziele, -ergebnisse und -aufgaben in der Praxis relevant sein, speziell um einzelne Faktoren der Eigenschaften entsprechenden Anwendungsbereichen zuordnen zu können, da die verschiedenen Eigenschaften besonders für jeweils passende Lernziele interessant zu sein scheinen. Hierbei sollten aus anderen Forschungsbereichen bekannte Wirkmechanismen (z. B. 3D-Darstellung kann räumliches Lernen unterstützen) betrachtet und auf AR-Lernerfahrungen übertragen werden. Zusätzlich zu individuellen Lernsettings sind auch kollaborative Lernsettings, die in der Bildung immer verbreiteter sind, und ihre Umsetzung auf Basis der drei Eigenschaften interessant, da es auch hier ein großes Potential gibt (z. B. Lernpartner als Teil des Kontexts; gemeinschaftliche Interaktion mit Elementen; räumliche Positionierung um ein Objekt herum; s. auch [KBB19]). Auch der Fokus auf verschiedene Personenvariablen wie Motivation oder bestimmte Fähigkeiten ist relevant, da diese Einfluss auf die Verarbeitung von Inhalten und die Lernerfahrung in AR haben können. Im weiteren Verlauf sollten außerdem andere Technologien betrachtet werden, die eine bessere Umsetzung der AR-Eigenschaften mit sich bringen können (z. B. AR-Brille, um Hände für Interaktion frei zu haben). Alles in allem bietet AR viele neue Lernmöglichkeiten, die weitere empirische Erforschung benötigen, um eine Basis für die Umsetzung und den Einsatz effektiver und effizienter AR-Anwendungen bereitzustellen.

Literaturverzeichnis

- [AA17] Akçayır, M.; Akçayır, G.: Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review* 20/17, S. 1-11, 2017.
- [Az97] Azuma, R.: A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments* 6(4)/97, S. 355-385, 1997.
- [Be12] Benyon, D.: Presence in blended spaces. *Interacting with Computers*, 24(4)/12, S. 219-226, 2012.
- [BD12] Billinghurst, M.; Dünser, A.: Augmented reality in the classroom. *Computer*, 45(7)/12, S. 56-63, 2012.
- [Bo14] Bower, M. et al.: Augmented reality in education—Cases, places and potentials. *Educational Media International*, 51(1)/14, S. 1-15, 2014.
- [CHS15] Chen, S.-C.; Hsiao, M.-S.; She, H.-C.: The effects of static versus dynamic 3D representations on 10th grade students' atomic orbital mental model construction: Evidence from eye movement behaviors. *Computers in Human Behavior*, 53/15, S. 169-180, 2015.
- [CM16] Clark, R.C.; Mayer, R. E.: Engagement in e-Learning. In R. C. Clark & R. E. Mayer (Hrsg.): *E-Learning and the Science of Instruction: Proven Guidelines for Consumers*

- and Designers of Multimedia Learning. John Wiley & Sons, Hoboken, USA, S. 219-238, 2016.
- [Cr13] Craig, A.B.: Chapter 2—Augmented reality concepts. In A. B. Craig (Hrsg.), Understanding augmented reality. Morgan Kaufmann, Boston, USA, S. 39-67, 2013.
- [CT13] Cheng, K.-H.; Tsai, C.-C.: Affordances of augmented reality in science learning: Suggestions for future research. *Journal of Science Education and Technology*, 22(4)/13, S. 449-462, 2013
- [CW14] Chi, M.T.H.; Wylie, R.: The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist*, 49(4)/14, S. 219-243, 2014.
- [de06] de Jong, T: Technological Advances in Inquiry Learning. *Science*, 312(5773), S. 532-533, 2006.
- [DD14] Dunleavy, M.; Dede, C.: Augmented reality teaching and learning. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Hrsg.), *Handbook of Research on Educational Communications and Technology*. Springer, New York, USA, S. 735-745, 2014.
- [GK17] Georgiou, Y.; Kyza, E.A.: The development and validation of the ARI questionnaire: An instrument for measuring immersion in location-based augmented reality settings. *International Journal of Human Computer Studies*, 98/17, S. 24-37, 2017.
- [GK21] Georgiou, Y.; Kyza, E.A.: Bridging narrative and locality in mobile-based augmented reality educational activities: Effects of semantic coupling on students' immersion and learning gains. *International Journal of Human-Computer Studies*, 145/21, Artikel 102546, 2021.
- [HFN11] Hugues, O.; Fuchs, P.; Nannipieri, O.: New augmented reality taxonomy: Technologies and features of augmented environment. In B. Furht (Hrsg.), *Handbook of Augmented Reality*. Springer, New York, USA, S. 47-63, 2011.
- [HNS18] Holmes, C.A.; Newcombe, N. S.; Shipley, T. F.: Move to learn: Integrating spatial information from multiple viewpoints. *Cognition*, 178/18, S. 7-25, 2018.
- [HP16] Hackett, M.; Proctor, M.: Three-dimensional display technologies for anatomical education: A literature review. *Journal of Science Education and Technology*, 25(4), S. 641-654, 2016
- [JM17] Johnson-Glenberg, M.C.; Megowan-Romanowicz, C.: Embodied science and mixed reality: How gesture and motion capture affect physics education. *Cognitive Research: Principles and Implications*, 2(1)/17, S. 1-28, 2017.
- [Ka13] Kamarainen, A.M. et al.: EcoMOBILE: Integrating augmented reality and probeware with environmental education field trips. *Computers and Education*, 68/13, S. 545-556, 2013.
- [KB20] Krüger, J.M.; Bodemer, D.: Different Types of Interaction with Augmented Reality Learning Material. In D. Economou et al. (Hrsg.), 2020 6th International Conference of the Immersive Learning Research Network (iLRN), S. 78-85, 2020.

- [KB21] Krüger, J.M.; Bodemer, D.: Space, a central frontier – the role of spatial abilities when learning the structure of 3D AR objects. In D. Economou et al. (Hrsg.), 2021 7th International Conference of the Immersive Learning Research Network (iLRN), S. 258-265, 2021.
- [KBB19] Krüger, J.M.; Buchholz, A.; Bodemer, D.: Augmented reality in education: Three unique characteristics from a user's perspective. In M. Chang et al. (Hrsg.), Proceedings of the 27th International Conference on Computers in Education, S. 412-422, 2019.
- [KPB21] Krüger, J.M.; Palzer, K.; Bodemer, D.: Learning with augmented reality: Impact of dimensionality and spatial abilities. [Zur Veröffentlichung eingereichtes Paper]
- [Ki13] Kim, M.J.: A framework for context immersion in mobile augmented reality. Automation in Construction, 33/13, S. 79-85, 2013.
- [KSS17] Klepsch, M.; Schmitz, F.; Seufert, T.: Development and validation of two instruments measuring intrinsic, extraneous, and germane cognitive load. Frontiers in Psychology, 8/17, Artikel 1997, 2017.
- [Li16] Lindgren, R. et al.: Enhancing learning and engagement through embodied interaction within a mixed reality simulation. Computers and Education, 95/16, S. 174-187, 2016.
- [Ma09] Mayer, R.E.: Spatial Contiguity Principle. In R. E. Mayer (Hrsg.), Multimedia Learning, Cambridge University Press, Cambridge, S. 135-152, 2009.
- [Pe95] Peters, M. et al.: A redrawn Vandenberg and Kuse mental rotations test - Different versions and factors that affect performance. Brain and Cognition, 28(1)/95, S. 39-58, 1995.
- [Ra14] Radu, I.: Augmented Reality in Education: A Meta-Review and Cross-Media Analysis. Personal and Ubiquitous Computing 18(6)/14, S. 1533-1543, 2014.
- [Re05] Reid, J. et al.: Experience Design Guidelines for Creating Situated Mediascapes. Report from Mobile and Media Systems Laboratory, HP Laboratories Bristol, S. 1-71, 2005.
- [SE01] Surry, D.W.; Ensminger, D.: What's wrong with media comparison studies? Educational Technology, 41(4)/01, S. 32-35, 2001.
- [SH16] Stull, A.T.; Hegarty, M.: Model manipulation and learning: Fostering representational competence with virtual and concrete models. Journal of Educational Psychology, 108(4)/16, S. 509-527, 2016.
- [SVP98] Sweller, J.; van Merriënboer, J.J.G.; Paas, F.: Cognitive architecture and instructional design. Educational Psychology Review, 10(3)/98, S. 251-296, 1998.
- [SVP19] Sweller, J.; van Merriënboer, J.J.G.; Paas, F.: Cognitive architecture and instructional design: 20 years later. Educational Psychology Review, 31(2)/19, S. 261-292, 2019.
- [Sy10] Sylaiou, S. et al.: Exploring the relationship between presence and enjoyment in a virtual museum. International Journal of Human Computer Studies, 68(5)/10, S. 243-253, 2010.
- [Wi02] Wilson, M.: Six views of embodied cognition. Psychonomic Bulletin and Review, 9(4)/02, S. 625-636, 2002.

- [WS98] Witmer, B.G.; Singer, M.J.: Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3)/98, S. 225-240, 1998.
- [Wu13] Wu, H.-K. et al.: Current status, opportunities and challenges of augmented reality in education. *Computers & Education*, 62/13, S. 41-49, 2013.

An Augmented Reality-supported Facility Model in Vocational Training

Lisa Guth¹, Heinrich Söbke², Eva Hornecker³, and Jörg Londong⁴

Abstract: Augmented reality (AR) is recognized as a powerful support for learning processes, although it is not yet state of the art within a wide variety of learning contexts. This exploratory study hence investigates a tablet-based AR app for visualizing additional information of a facility model in a formal learning context of vocational training. Specifically, the facility model demonstrates water supply and wastewater disposal processes in vocational training for water supply and wastewater treatment technology specialists. Specifically, learning scenarios related to two processes of similar technical complexity are examined. In one learning scenario, apprentices receive information via an AR app, using a tag-based tracking. In the other learning scenario, the information must be acquired from a paper information sheet. Apprentices (N=14) work through both learning scenarios in groups of two, in random order. Data are collected via questionnaires before and after each learning scenario, as well as through a semi-structured interview at the end of both learning scenarios. The results suggest a higher level of interest in the AR-assisted learning scenario. Furthermore, the study reveals non-negligible efforts to customize the AR app and provides hints for further development of the learning scenarios.

Keywords: augmented reality, wastewater disposal, water supply, facility model, vocational training.

1 Introduction

Augmented reality (AR), an extension of real-world objects using computer-generated stimuli, is considered a valuable support for learning [AA17], [SM18], [Ga20]. In this way, various approaches conducive to learning are advocated using AR, such as the contiguity principle, i.e., bringing the object and additional information, such as the name of the object, together spatially [MF14], or communication [Za13] on the object. Motivating learners is one of the most reported benefits of AR in learning scenarios [GPB19], [AA17]. With the advancement of mobile devices, such as smartphones and tablets, and AR software libraries, AR is being opened up for many more learning scenarios. One key application area is the training of craft skills [SPJ21]. For example, assembly training is conducted [BMW19], fault analysis skills in wind turbines are trained [Ma21], or

¹ Ernst-Abbe-Hochschule Jena, Carl-Zeiss-Promenade 2, 07745 Jena, Germany, lisa.guth@eah-jena.de

² Bauhaus-Universität Weimar, Bauhaus-Institute for Infrastructure Solutions (b.is), Goetheplatz 7/8, 99423 Weimar, Germany, heinrich.soebke@uni-weimar.de,  <https://orcid.org/0000-0002-0105-3126>

³ Bauhaus-Universität Weimar, Human Computer Interaction Group, Bauhausstraße 11, 99423 Weimar, Germany, eva.hornecker@uni-weimar.de

⁴ Bauhaus-Universität Weimar, Bauhaus-Institute for Infrastructure Solutions (b.is), Goetheplatz 7/8, 99423 Weimar, Germany, joerg.londong@uni-weimar.de,  <https://orcid.org/0000-0002-2861-056X>

visualizations are created as a basis for communication [FMA16] — all using AR technology. Despite its wide-ranging potential, AR as learning tool in vocational training is still in its infancy [LH20], [Sö21]. In particular, AR as learning tool is mostly still limited to mobile devices such as tablets or smartphones [Le21]. A meta study [Ka21] even failed to find significant effects of trainings using Extended Reality (XR), a superset of AR. Thus, the meta study concludes that among the advantages of XR is the opportunity of training even when using real objects is not feasible or involves significant risks. This condition in particular is true for the water supply and wastewater disposal facilities considered in this study. Rather than being simulated purely digitally, for example by Virtual Reality (VR), the facilities in this study are represented by a real facility model, which are supplemented by AR for facilitation of learning. The scientific evaluation of AR-supplemented real and tangible models, such as a model facility, is not adequately covered in the literature so far, e.g. [La20], [Ji17]. This exploratory study aims to evaluate to what extent AR might replace the conventional paper-based provision of information in learning scenarios (RQ 1), what challenges arise during practical use of AR (RQ 2), and whether motivational effects of AR might be observed (RQ 3). The remainder of the article is organized as follows: the next section describes in more detail the methodology including the facility model as well as the learning scenarios used and the demography of the participants sample. Section 3 presents the results, which are discussed in the following section 4 and summarized in section 5.

2 Methodology

The facility model is used in vocational training for providing apprentices with a basic understanding of technical facilities for water supply and wastewater disposal. With the help of a customizable AR app, additional information, such as the name of specific facility model components or the flow direction of the water, is provided to the apprentices. AR is considered particularly valuable for this model facility, since the model facility initially requires a high level of familiarization, the flow of water in the individual components is not self-explanatory, and the mental transfer of the abstract flow diagram to the model facility is considered as challenging. This study investigates effects of how the additional information is provided: in one learning scenario, the additional information is delivered via AR app on an iPad tablet; in the second learning scenario, the additional information is provided using paper information sheets. The subjects who participated in the study in June 2020 were selected by random drawing from all first-year apprentices for water supply technology and wastewater technology of a vocational school in Weimar, Germany. All participants went through both learning scenarios in groups of two, consecutively (within-subjects design), whereby the order of the learning scenarios was determined randomly for excluding order effects. Methodologically, the subjects first received a short introduction outlining the setup of the learning scenario. The participants subsequently completed questionnaire 1 on demographic data, prior knowledge, and motivation. Afterwards, the first learning scenario was carried out under the supervision of a teacher. The participants had one hour to complete the learning scenario. Afterwards, questionnaire

2 on learning outcomes and motivation was completed. Subsequently, the next learning scenario was presented and questionnaire 3 on motivation was completed. The second learning scenario was also given one hour to complete. After the subsequent questionnaire 4 on learning outcomes and motivation, individually with each participant a semi-structured interview was conducted by the study head. The feasibility of the study was initially validated using a test group of two computer science bachelor students. One of the teachers also supervised the participants during the practical part of the study. The observations during the learning scenarios as well as the analyses of the semi-structured interviews contribute to answering RQ 2. All participants or the parents of the underage participants gave their informed consent to the study and were informed that they were allowed to terminate their participation in the study at any time without adverse consequences.

2.1 Facility model

The facility model [AD21] is a modular training device that can be used to simulate all core processes of water supply and wastewater disposal (Figure 1). On the left side of the figure the water purification in a waterworks is shown. The used drinking water becomes wastewater, which is discharged to a wastewater treatment plant. Sensors, such as level gauges, and actuators, such as pumps and valves, which are controlled by software on a notebook, are used for process simulation. The facility model is described in a training manual, in which several learning scenarios for a range of core processes are presented. The vocational school acquired the facility model in 2018, set it up in a dedicated location, and is currently integrating the facility model into its educational activities. As part of the integration into the educational activities, the supplementation of the facility model with a customizable AR app for learning purposes [Fe21b] is also being explored. The use of the AR app requires the manual setup of so-called AR *scenes*, as described below.

Study location. The study was carried out at the facility model, setup in the vocational school.

AR app customization. For customization of the AR app, a configuration file in XML format has to be created [Fe21a]. In this file, the tags used for tracking and their positions on the facility model are described. Furthermore, the relative positions of all components of the facility model have to be specified, as well as the additional information and their relative positions. Multiple AR scenes may be described for each learning scenario. The configuration file is to be stored on a web server for making it accessible to the AR app.

Learning scenarios. For the study, the learning scenario "Sedimentation at varying stream velocity (EDS 2)" was selected to be supported by the AR app. With the help of granulate added to the water, sedimentation (settling of granulate in a flowed basin) is observable in this scenario (Figure 2). Based on the ability to change the flow rate, the effect of flushing pulses may be shown. Furthermore, the learning scenario "Water Storage and Transport (EDS 1)" was used for the paper-based learning scenario. The objective of this learning scenario is controlling a reservoir to ensure a constant water supply in case of time-varying

consumptions. For both learning scenarios, there exists on the one hand a *worksheet* that guides the work on the facility model with instructions and questions to be answered. Secondly, each worksheet is accompanied by an *information sheet* that provides information necessary to complete the worksheet. The technical complexity of both learning scenarios was judged by two teachers involved in this study being of almost equal complexity, since in both learning scenarios a system consisting of several valves, pumps, pipes and tanks has to be controlled.

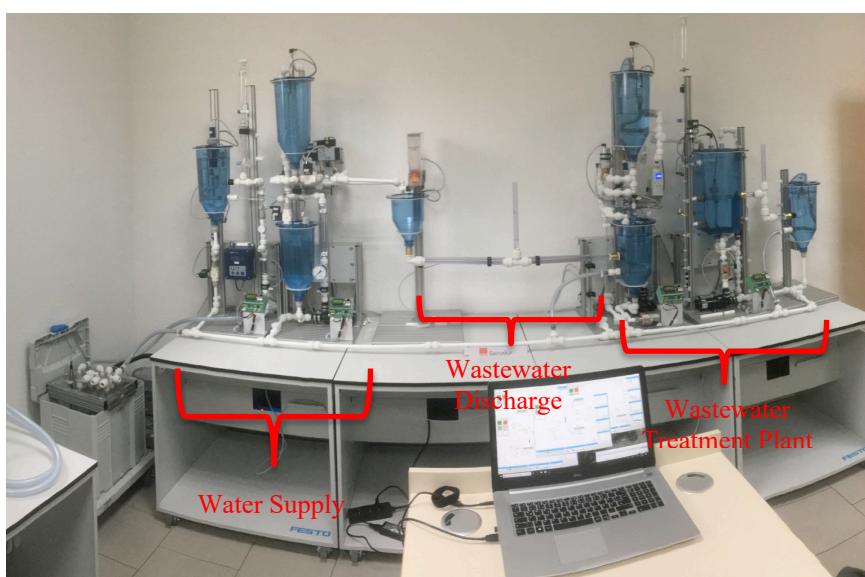


Figure 1 Facility model and notebook for control

AR app-provided additional information. The scenes for the AR-supported learning scenario created as part of this study accommodated all, but not more, the information from the information sheet. Thus, the requirement of information equivalence between AR-supported and paper-supported learning scenario has been met. The most important information was the flow diagram in which all components of the facility model were represented and named. In the scenes, this information was spatially mapped to the facility model (Figure 2). The information conveyed through the AR app included location and naming of the facility model components, such as valves, pumps, tanks, and pipes, as well as the direction of flow of the water in the pipes. The visualization elements introduced into the scenes included text and graphic symbols. Only static information was integrated into the scene, i.e., no values from sensors that might be queried dynamically were displayed, for example.

2.2 Study sample: demographics and prior knowledge

The participants were 14 apprentices in their first year of apprenticeship. 12 participants were in the apprenticeship program to become wastewater technology specialists, and one participant each was being apprenticed to become a water supply technology specialist and a plant mechanic for piping system technology. The average age of the participants was 18.4 years with a range of 17 to 21 years. 13 of the participants were male, 1 participant was female.

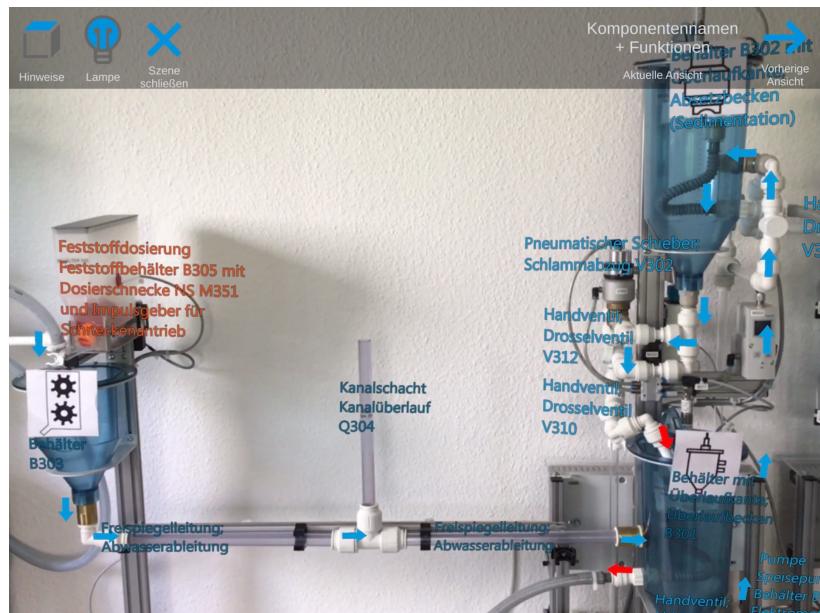


Figure 2 AR app screenshot: annotations visualized by the AR app; blue arrows indicate the flow direction; in the left and right: marker for tracking

The apprenticeship of the participants had recently begun, so a rather low level of prior knowledge was to be assumed. Thus, 6 of the 14 participants stated that they had already heard of AR, and 4 participants stated some experience with AR on their own. However, a total of 8 of the 14 participants indicated that they had installed the AR game app Pokémon GO [Ni16] on their smartphone before. This larger number suggests gaps in knowledge about AR as a technology. For example, only one participant also stated being able to explain the term AR to a person being unfamiliar with AR. 6 participants denied this ability, while the 7 remaining participants claimed not being sure if they could explain it. Only 2 participants had worked with the facility model itself before. In contrast, 13 of the 14 participants had already been able to gain knowledge about wastewater treatment plants (WWTPs). On a 7-point Likert scale, participants rated their knowledge of WWTPs at an average of 5.0 points (SD: 1.33) (1: no knowledge, 7 very good knowledge). Interest

in WWTPs was also rated on a 7-point Likert scale with 5.9 points ($SD: 0.74$). While the subjective expert knowledge interest is to be rated as high, the intentional experiences with AR as well as with the facility model itself are to be regarded as low.

3 Results

3.1 Questionnaire on motivation

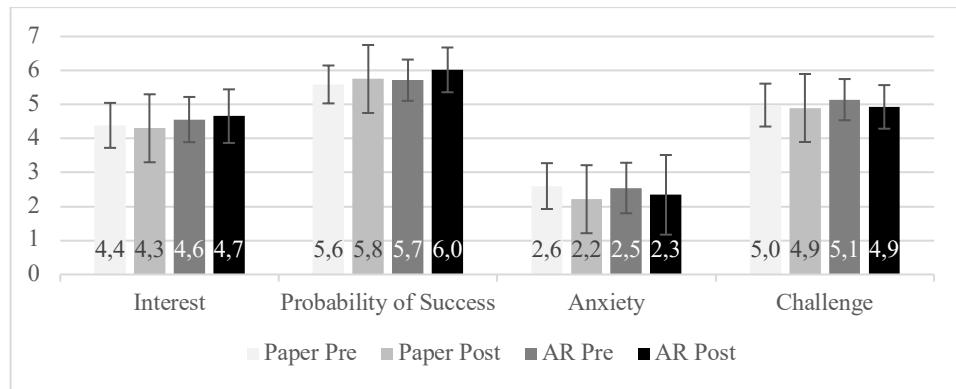


Figure 3 QCM subscales (N=14)

Among the major reasons for applying AR in learning scenarios is its effect on motivation [AA17]. Accordingly, motivation was selected for evaluation in this study. A frequently used questionnaire to survey motivation is the *Questionnaire to Assess Current Motivation in Learning Situations* (QCM) by Rheinberg et al. [RVB01]. The QCM consists of 18 items assessing the subscales *Interest*, *Probability of Success*, *Anxiety* and *Challenge*. Figure 3 shows the subscale values depending on the AR or paper variant of the learning scenarios and on the point of measurement (pre or post). The subscale *Interest* shows higher values in both points of measurement for the AR variant. *Interest* may be driven by interest in a technology that is new to most participants. The value for the subscale *Probability of Success* reaches its highest value in the post-measurement for the AR variant. It is conceivable that AR is expected by the participants to have a positive influence on task performance. At the same time, the value for the subscale *Challenge* is slightly higher in the pre-measurements, i.e., the additional technology also appears more demanding. The subscale *Anxiety* is almost equally low for both learning scenarios. Overall, the measured values can be seen as very positive for the learning scenario: a high level of *Interest* is surpassed by a high level of *Challenge*, with *Probability of Success* being very high facing a low level of *Anxiety*. Additionally, a MANOVA analysis was completed, but no significant effects were found. Furthermore, t-tests were conducted for pre- and post-measurement values regarding all four subscales of the two variants AR and paper. For the post-measurement values of the subscale *Interest*, a significant effect showing a higher value

in the AR variant was identified. However, the effect was not confirmed by a Mann-Whitney U test.

3.2 Semi-structured interviews

In the semi-structured interviews, participants were asked about the learning scenarios conducted. Topics asked in the semi-structured interview included challenges in completing the worksheets, usability and usefulness of the AR app, differences between the two learning scenarios, learning outcomes, enjoyment, and motivation. The semi-structured interviews were recorded and transcribed. The transcriptions were further processed based on the qualitative content analysis according to Mayring [Ma04]. The software ATLAS.ti [AT21] was used. The study head classified the transcribed interviews and obtained the following topic categories after two classification passes:

New experiences. All participants indicated having enjoyed the learning scenarios and especially using the AR technology. Some participants drew parallels to other new technologies they were aware of and mentioned Virtual Reality (VR) in particular. At this point, the sometimes imprecise delimitation of the term Augmented Reality became apparent again. Various statements made clear that the conscious use of AR technology is generally still felt to be exceptional. Especially in the field of vocational training the application of AR for learning purposes felt very innovative for the participants. (*"I would not have thought that your project has already come so far"* [T7]).

Comparison of the AR-supported and the paper-supported learning scenario. The advantages and disadvantages as well as difficulties and simplicity of the two scenarios were asked, first separately from each other and then in comparison. Both learning scenarios had aspects that participants highlighted as pleasant. The immediate overview of the experiment without AR and the entire facility model was rated as positive by all participants. Most of them had no problems understanding how the facility model works and how it is set up. The information sheets provided made the information required easily accessible. However, most of the participants only used the flow diagram and neglected the other material. (*"I have to admit I hardly used the material except for the flow diagram to recognize where which component is, but otherwise, I found it went well and it was relatively convenient to work with the flow diagram only."* [T9]). Because of the corresponding labelling on the flow diagram and on the facility model, orientation was easy for the participants, even if it took partially longer. In the AR-based learning scenario, the participants considered the work speed as very convenient regarding the workflow and named various aspects. When comparing the AR to the paper variant, the participants liked the easy-to-recognize flow direction using virtual arrows. (*"[...] was simply pictured and arrows were used to represent this flow direction"* [T3]). The participants considered the visualization and label on top of each component of the facility model as very helpful, as the schematic drawing was partly deviated from the shape of the facility model component. (*"[...] with AR and with the tablet one could see and understand it better, it was shown pictorially and without it was just harder to find all the valves, it wasn't quite as*

understandable" [T3]). However, there were also some aspects in both experiments that the participants found rather difficult or were seen as a disadvantage. All participants confirmed that finding components in the facility model took significantly longer if the information were available only on paper. Although the labels were the same here, the facility model components and the flow diagram appeared quite different. The AR-based learning scenario was also rated difficult in parts. The weight of the tablet was too heavy for two participants and was therefore described as unwieldy. ("*[...] always a bit tedious to walk around with the iPad [...]*" [T4]) A display size that allows all components to be clearly visualized is reasonable. On the other hand, attention is to be paid for ensuring that the device does not interfere with work requirements. A balance has to be found regarding the size and kind of device.

Task difficulty. The answers of the participants regarding the difficulty of the paper-based learning scenario were very heterogeneous. Answers ranged from "*it was straight-forward because all the necessary information were given*" via "*it was about the same difficulty*" to "*it was harder because you need to find your way without support*". Overall, however, tendentially, the paper-based learning scenario was considered more difficult. This tendency confirms previous statements by participants that it takes way more time to fully understand the system by only using the flow diagram on paper. Opinions also differ about the difficulty of the AR-based learning scenario. Few participants regarded the AR-based learning scenario more difficult, because of possible challenges in using the AR app on top of the technical problems. However, the majority of the participants preferred the AR-based learning scenario, as the AR simplified and accelerated understanding of the facility model.

Added value of new technologies: The responses of the participants abstracting from the experiences made in these learning scenarios revealed that the added value of an AR app is motivating apprentices better to participate in lectures and to shorten the time required to convey complex tasks. The participants describe this as more intensive learning, for example, in the course of exam preparation. ("*Well, that you can then perhaps work through more learning material and that you then perform better when you discover such innovative things for yourself and also learn and use them.*" [T6]) One participant stated that the early use of AR app can significantly increase the interest of the students in their professional field. Another participant added that when learning is fun, especially through the use of modern technology, more people are interested in the specific professional field.

Future requests and future work. In their requests, the participants clearly refer to the use of VR. While the extension of real-world objects using virtual content and additional information was perceived as very positive, the participants want a, in the perfect scenario, completely accessible virtual world. The participants see the added value of a virtual world in the option of repeating work processes and thus being able to memorize the work processes better. Also, obviously, the participants requested more technological support, such as using a beamer to illustrate complex issues by showing drawings, pictures, or videos. Especially when attending presentations, the participants requested handouts for taking personal notes and markings. Small projects and short presentations were also requested.

Two participants mentioned using tablets for digital work. On the other hand, these statements revealed that it will still take some time before AR is sufficiently established in the minds of the participants.

4 Discussion

The results of the explorative study suggest the potential of AR as learning tool. However, deficits also become apparent that generally pose challenges in the context of the digitization of learning activities. The **high setup effort** for the preparation of the learning activity is to be noted: the creation of the XML configuration file is to be described as time-consuming, even minor changes, such as changing the position of an augmentation only a little, entail a considerable amount of work, such as updating the XML configuration file on the web server. Having a comprehensive authoring system would ease the work here to allow technically non-versed teachers to update the XML configuration file on the web server. The **technical requirements** might also compromise the viability of the learning scenario. For example, a Wi-Fi router first had to be set up in the room of the facility model for enabling the operation of mobile devices. The facility model itself requires **high efforts for maintenance**: Granulate was added to the water as a solid to simulate sedimentation. After completion of the learning scenarios, the facility model needs to be tediously cleaned up. Here, AR itself might help by visualizing the sedimentation of solids based on sensors reporting the flow velocity. Yet, this would entail the visualization of dynamic processes, which is technically feasible. Critical seems to be the **high level of supervision** required. During the entire learning scenarios (both paper variant and AR variant), teachers were available to assist in case of difficulties. Compared to the pre-study group, the participants acted less independent and required more supervision. For the pre-study, test persons should have been recruited who were closer to the target group allowing to tailor the learning scenarios to the target group better. The results of the study are to be interpreted with the restriction of the **small number of participants**. The small sample size probably caused non effects to be found. Likewise, potential effects might not be explained by solely using an AR app, but could be due to various factors, including novelty effect, small group size with good supervisory ratio, or even the model facility itself. Also, no measurement of learning outcomes was conducted, which should be the subject of further studies as well as the consideration of the teachers' reflections on learning outcomes. The information equivalence of the two scenarios studied should be more accurately determined by a detailed analysis of the information conveyed than by the more general statements of the teachers. In future studies, the sampling procedure should also ensure that all participants have the same level of prior knowledge about the model facility. Further, almost exclusively male apprentices took part in the study. However, this phenomenon may be observed more frequently in craft-related vocational training programs [Fu19], [De21]. Furthermore, the QCM instrument was administered after a learning activity, which does not correspond to the recommended usage assuming a measurement before the learning activity. Nevertheless, to influence the measurement results as little as

possible, participants were asked to hypothetically assume an upcoming learning unit of equal structure. In summary, AR is not yet an established learning tool in vocational training. Moreover, AR cannot replace teachers for learning activities, such as employed in this study. In more detail, the texts stored in the AR app were reported as too long in some cases. This feedback may be seen as a request to increase the interactivity of the AR app with further prompts. Also, the adaptivity of the AR app ought to be increased for supporting the needs of learners better. Undisputed is the need to incorporate AR to a greater extent in vocational training, and beyond just as an educational tool: Guided by the requirements of modern workplaces, where, for example, AR is increasingly used in facility maintenance [Pa18], in maintenance in general [FDB16] or in production lines [Wh20], often as foundation of assistance systems [Ha20], the need for the integration of AR technologies into vocational training arises even stronger. For successful integration, it is necessary that AR technology is further developed for enhancing accessibility and ease of use, as the use cases of aircraft manufacturers illustrate [DL14]. Further, organizational preconditions need to be met, such as the reorganization of vocational training planned in the field of environmental professions aims at increasing the digital literacy of learners [HCF19].

5 Conclusions

In this exploratory study, two learning scenarios using a facility model were investigated, one of which was supported by Augmented Reality (AR), while the second required additional information to be retrieved from a paper information sheet. Thereby, AR was found being suitable to convey additional information to the learners (RQ 1). However, there are significant technical challenges to overcome, which currently still require skills from teachers that are not part of their training (e.g., updating web server content) (RQ 2). Both AR variant and paper variant exhibited decent values for motivation suggesting mild advantages for the AR variant. However, AR represents only one among multiple possible causative factors. Thus, further studies need to clarify the actual influence of AR on motivation in settings of AR-supplemented model facilities (RQ 3). Overall, the study suggests that the AR-supported facility model has led to an enrichment of teaching and, in particular, to positive learning experiences that cannot be provided in reality, as it is mostly too risky to use corresponding facilities in educational scenarios.

Bibliography

- [AA17] Akçayır, M.; Akçayır, G.: Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review* 20, 2017.
- [AD21] ADIRO Automatisierungstechnik GmbH: EDS\textregistered Water Management for Festo Didactic, 2021.

-
- [AT21] ATLAS.ti: ATLAS.ti Qualitative Data Analysis, 2021.
 - [BMW19] Beckmann, J.; Menke, K.; Weber, P.: AR/VR-Trainings in the ARSul-Project: Results From a Prototype Evaluation. INTED2019 Proceedings March/1, pp. 4353–4362, 2019.
 - [De21] Deutscher Industrie- und Handelskammertag: Zahlen, Daten und Fakten: Die wichtigsten Informationen über den Beitrag der IHKs zur Ausbildung in Deutschland, 2021.
 - [DL14] Davies, P.; Lee, D.: Augmented Reality in Manufacturing at the Boeing Company: Lessons Learned and Future Directions, 2014.
 - [FDB16] Frigo, M. A.; Da Silva, E. C. C.; Barbosa, G. F.: Augmented Reality in Aerospace Manufacturing: A Review. Journal of Industrial and Intelligent Information 2/4, pp. 125–130, 2016.
 - [Fe21a] Festo Didactic: AR Scene Programming Reference, 2021.
 - [Fe21b] Festo Didactic SE: Festo Didactic AR, 2021.
 - [FMA16] Fehling, C. D.; Müller, A.; Aehnelt, M.: Enhancing Vocational Training with Augmented Reality. Proceedings of the 16th International Conference on Knowledge Technologies and Data-driven Business October, 2016.
 - [Fu19] Funk, J. et al.: Implementation of a 3D-360 ° - lesson in the practical training of craftsmen. In (Schulz, S. Ed.): Proceedings of DELFI Workshops 2019 Berlin, Germany, September 16, 2019. Gesellschaft für Informatik eVz, pp. 161–172, 2019.
 - [Ga20] Garzón, J. et al.: How do pedagogical approaches affect the impact of augmented reality on education? A meta-analysis and research synthesis. Educational Research Review, p. 100334, 2020.
 - [GPB19] Garzón, J.; Pavón, J.; Baldiris, S.: Systematic review and meta-analysis of augmented reality in educational settings. Virtual Reality 0123456789, 2019.
 - [Ha20] Haase, T. et al.: Learning in the process of work - wish or reality? An interdisciplinary approach to designing technology-based learning and assistance systems to promote learning. IMSCI 2020 - 14th International Multi-Conference on Society, Cybernetics and Informatics, Proceedings January, pp. 219–224, 2020.
 - [HCF19] Herb, S.; Czichy, C.; Fassnacht, A.: Berufsbildung für eine Wasserversorgung 4.0: Die Neuordnung der Umweltberufe im Hinblick auf die Digitalisierung wird vorbereitet. energie | wasser-praxis 4, pp. 2–5, 2019.
 - [Ji17] Jin, W. et al.: Augmented and virtual reality: Exploring a future role in radiation oncology education and training. Applied Radiation Oncology December, pp. 13–20, 2017.
 - [Ka21] Kaplan, A. D. et al.: The Effects of Virtual Reality, Augmented Reality, and Mixed Reality as Training Enhancement Methods: A Meta-Analysis. Human Factors 4/63, pp. 706–726, 2021.
 - [La20] Larson, K.: Augmenting the Corporate Training Reality: A Review of Literature on AR in the Corporate Training Environment: Innovate Learning Summit 2020. Association for the Advancement of Computing in Education (AACE), pp. 241–250, 2020.

-
- [Le21] Lee, M. J. W. et al.: State of XR & Immersive Learning 2021 Outlook Report, Walnut, CA, 2021.
 - [LH20] Lester, S.; Hofmann, J.: Some pedagogical observations on using augmented reality in a vocational practicum. *British Journal of Educational Technology* 3/51, pp. 645–656, 2020.
 - [Ma04] Mayring, P.: Qualitative content analysis. In (Flick, U.; Karsdorff, E. von; Steinke, I. Eds.): *A Companion to Qualitative Research*. SAGE Publications, London, pp. 159–176, 2004.
 - [Ma21] Matthes, N. et al.: Trainieren der Fehlerdiagnosekompetenz in der Ausbildung. Qualitative Studie mit Lehrenden im Bereich Metall- und Elektrotechnik. *Journal of Technical Education (JOTED)* 1/9, pp. 31–53, 2021.
 - [MF14] Mayer, R. E.; Fiorella, L.: 12 Principles for Reducing Extraneous Processing in Multimedia Learning: Coherence, Signaling, Redundancy, Spatial Contiguity, and Temporal Contiguity Principles: *The Cambridge Handbook of Multimedia Learning*. Cambridge University Press, pp. 279–315, 2014.
 - [Ni16] Niantic Inc.: *Pokémon GO*, 2016.
 - [Pa18] Palmarini, R. et al.: A systematic review of augmented reality applications in maintenance. *Robotics and Computer-Integrated Manufacturing February/49*, pp. 215–228, 2018.
 - [RVB01] Rheinberg, F.; Vollmeyer, R.; Burns, B. D.: QCM A questionnaire to assess current motivation in learning situations. *Diagnostica* 47, pp. 57–66, 2001.
 - [SM18] Sommerauer, P.; Müller, O.: Augmented reality for teaching and learning - A literature review on theoretical and empirical foundations. *26th European Conference on Information Systems: Beyond Digitization - Facets of Socio-Technical Change, ECIS 2018 July*, 2018.
 - [Sö21] Söbke, H. et al.: MR-Lernwerkzeuge Ergebnisse einer Online-Umfrage des Arbeitskreises VR/AR-Learning, 2021.
 - [SPJ21] Stender, B.; Paehr, J.; Jambor, T. N.: Using AR/VR for Technical Subjects in Vocational Training – Of Substantial Benefit or Just Another Technical Gimmick?: *2021 IEEE Global Engineering Education Conference (EDUCON)*, pp. 557–561, 2021.
 - [Wh20] White, G.: BMW Uses Google Glass on its Vehicle Production Lines, 2020.
 - [Za13] Zarraonandia, T. et al.: An augmented lecture feedback system to support learner and teacher communication. *British Journal of Educational Technology* 4/44, pp. 616–628, 2013.

Augmented Reality in der Ausbildung an Hubschrauberkonsolen

Ein empirischer Lernmittelvergleich zu Nutzungserlebnis und Lernerfolg

Madeleine Ratter¹, Johannes Klöckner¹ und Mara Kaufeld¹

Abstract: AR-Anwendungen bieten neue Möglichkeiten Wissen zu vermitteln und Lernprozesse autodidaktisch zu gestalten. In dieser Arbeit wird die Verwendung von AR-Lernmitteln zum Erlernen von Prozeduren in der Ausbildung von Hubschrauberpiloten genauer betrachtet. Dazu wurde ein prototypisches AR-Lernsystem aufgebaut, das durch räumlich verankerte Informationen und individuelle Einstellungsmöglichkeiten für Nutzende gekennzeichnet ist. In einer Nutzerstudie ($n = 32$) wurde das entwickelte Lernmittel interindividuell mit einer Anleitung aus Papier verglichen. Die Analyse von Nutzungsdaten, Lernergebnissen und Befragungen zeigte, dass Probanden mit beiden Lernmitteln erfolgreich die Prozedur erlernten. Es zeigten sich keine Unterschiede in den Lernergebnissen zwischen den Gruppen, allerdings deuten die Ergebnisse darauf hin, dass AR-Technologie im Vergleich zur Papieranleitung die Lernaufgabe erleichterte. Zudem wurde das Lernen mit AR von den Probanden als origineller, attraktiver und stimulierender bewertet.

Keywords: Augmented Reality; Hubschrauber; virtuelle Ausbildung; Nutzungserlebnis; Lernerfolg.

1 Einleitung

Der Einsatz von Augmented Reality (AR) zu Trainings- und Lernzwecken bietet medienspezifische Vorteile. Dazu gehören ein hoher Grad an Anpassbarkeit sowie die Darstellung von in der Realität nicht beobachtbaren Objekten und eine direkte räumliche Koppelung von digitalen Inhalten und realen Objekten. Deswegen profitieren insbesondere Anwendungen, in denen es eine starke Beziehung von digitalen Inhalten zu realen Objekten gibt, von AR [CBM12].

Im Bereich der Hubschrauberpilotenausbildung hat digital unterstütztes Training besonderes Potenzial, da das Realtraining nicht immer möglich, kostenintensiv sowie von einer Vielzahl von Faktoren (z.B. Wetterbedingungen) abhängig ist. Insbesondere Fluganfänger könnten von zusätzlichen Trainingsmöglichkeiten profitieren und so einen leichteren Einstieg in das Realtraining erhalten. Ein vielversprechendes Anwendungsfeld

¹ Fraunhofer FKIE, Mensch-Maschine-Systeme, Zanderstraße 5, 53177 Bonn, Deutschland
{madeleine.ratter@fkie.fraunhofer.de, johannes.kloeckner@fkie.fraunhofer.de,
mara.kaufeld@fkie.fraunhofer.de}

für AR-Anwendungen bietet dabei das Erlernen von Bedienelementen im Hubschrauber. Aufgrund der technischen Komplexität von Hubschraubern enthalten Cockpits typbedingt über 80 Eingabe- und Instrumentenkonsolen. Die Konsequenz ist eine raumeffiziente Gestaltung der Eingabe-Konsolen, die jedoch mit kleinen Tasten und stark abgekürzten Beschriftungen einhergeht. Wird die Bedienung der einzelnen Konsolen erst im Full Flight Simulator unter Einweisung von Fluglehrern geübt, entstehen hohe Kosten und wichtige Übungszeit unter professioneller Anleitung geht verloren. Deswegen ist es sinnvoll, die praktische Bedienung schon früh während der theoretischen Ausbildung zu erlernen und zu trainieren. Dafür wären Lernmittel, die das selbständige Lernen forcieren und Lernende motivieren, besonders wichtig.

Diese Arbeit soll das Potenzial von AR für die Ausbildung an Hubschrauber-Konsolen genauer untersuchen. Dafür wurde ein AR-System für die visuelle Erweiterung einer beispielhaften Konsole im Hubschraubercockpit entwickelt. In einer Nutzerstudie wurde evaluiert, inwiefern sich Lernergebnisse und Bewertungen für dieses AR-Lernmittel von denen eines herkömmlichen Lernmittels (z.B. einer Papieranleitung) unterscheiden.

2 Theoretischer Hintergrund

Das wissenschaftliche Interesse am Lernen mit AR ist seit einigen Jahren stark gestiegen, was möglicherweise darauf hindeutet, dass AR neben einer gewissen technischen Reife auch Fuß im Bereich der Ausbildung fasst [GPB19]. In diesem Kapitel sollen bisherige Erkenntnisse zusammengefasst und insbesondere der Kontext der Flugausbildung betrachtet werden. Aus den Erkenntnissen der vorangegangenen Arbeiten wird anschließend die Fragestellung dieser Studie abgeleitet.

2.1 Lernen mit AR

Erkenntnisse zum Lernen mit AR finden sich in unterschiedlichen Review-Studien. In einem Review von AR-Studien zum Erwerb von natur- und ingenieurwissenschaftlichen Kompetenzen primär bei Schülern [ID18] wurde gezeigt, dass durch den Einsatz von AR-Anwendungen vor allem affektive Zustände (z.B. Motivation) und Einstellungen positiv beeinflusst werden. Ähnliche Effekte der gesteigerten Motivation beim Lernen mit AR zeigten sich in einem Review [GPB19], das ebenfalls größtenteils Studien mit Schülern im naturwissenschaftlichen Bereich einbezog. Beide Reviews stellten aber auch Herausforderungen für den Einsatz von AR für das Lernen heraus, vor allem für Lehrende, die mit der Technik nicht vertraut sind, sowie durch technische Probleme oder schlecht gestaltete Systeme [GPB19, ID18].

Die stärkere affektive und motivationale Wirkung könnte ein großes Potenzial von Lernen mit AR darstellen. In einem Modell des selbstregulierten Lernens [GP91] ist intrinsische Motivation ein starker Einflussfaktor für selbstreguliertes Lernen und kognitive Leistungsbereitschaft. Intrinsisch motivierte Lernende bringen demnach mehr Ressourcen

zum Lernen auf und lernen selbstständig. In Anwendungskontexten in denen das selbstregulierte Lernen im Mittelpunkt steht, wie der Erwachsenenbildung, könnten diese Effekte genutzt werden, um langfristig bessere Lernergebnisse zu erzielen.

In einer aktuellen Review-Studie [GSD20] wurde ein anderes Anwendungsfeld, nämlich der medizinische Bereich, ausgenommen Operationen, betrachtet. Eine ältere Zielgruppe, Aufgabenstellungen aus dem beruflichen Kontext sowie das Erlernen von Prozeduren stehen hier im Fokus. Es zeigen sich Hinweise darauf, dass AR-Technologie Lernprozesse im medizinischen Bereich unterstützen kann und in vielen Fällen zu besseren Ergebnissen führte als traditionelle Lernmittel [GSD20]. Allerdings merken die Autor*innen auch an, dass die betrachteten Studien methodische Schwächen aufweisen und die Ergebnisse daher kritisch zu betrachten sind. Es zeigt sich also, dass noch Forschungsbedarf besteht.

Auch im speziellen Anwendungsbereich des Einsatzes von AR in der Flugausbildung gibt es neue Forschungsergebnisse. In einer aktuellen explorativen Studie [Sc20] wurde das Einsatzpotenzial von AR-Anwendungen für die (Flugzeug-)Pilotenausbildung unter Beachtung von geschlechterspezifischen Unterschieden untersucht. Dazu wurden 60 Pilot*innen und Fluglehrer*innen interviewt und gebeten, den Einsatz von AR-Anwendungen in verschiedenen Bereichen zu bewerten. Als besonders sinnvolle Einsatzmöglichkeiten von AR wurden die theoretische Ausbildung und die Vorflugkontrolle genannt. Die meistgegebene Antwort auf die Frage, für welche Inhalte AR am besten einzusetzen sei, war „Prozeduren innerhalb des Cockpits“. Damit liefert die Studie [Sc20] wertvolle Hinweise darauf, welche Anwendungen für AR-Training im Hubschrauber priorisiert betrachtet werden sollten.

2.2 Fragestellung

Vor dem Kontext der bereits thematisierten Herausforderungen in der Flugausbildung sowie den Erkenntnissen zu möglichen Einsatzgebieten von AR [Sc20], sollen Lerneffekte und Nutzungspräferenzen für eine Anwendung zur Erlernung von Prozeduren an Cockpit-Elementen untersucht werden.

Dafür wurde ein prototypisches AR-Lernmittel entwickelt, mit dessen Hilfe eine Prozedur an einer Cockpitkonsole erlernt werden kann. Dieses soll mit einem Papier-Lernmittel verglichen werden, das die gleichen Informationen zur Prozedur enthält. Der Fokus liegt damit explizit auf den unterschiedlichen Präsentationsformen. Dabei zeichnet sich das AR-Lernmittel durch in der Realität verankerte Inhalte und deren Adoptionsmöglichkeiten für die Nutzenden aus.

Basierend auf den Erkenntnissen bisheriger Forschung zu AR und Lernen werden in einem direkten Vergleich positive Effekte des AR-Lernmittels gegenüber einem konventionellen Lernmittel erwartet. Insbesondere werden Auswirkungen auf die Motivation der Proband*innen und die User Experience antizipiert. Das könnte wiederum auch

Akzeptanz und potenzielle freiwillige Nutzung des Lernmittels beeinflussen. Folgende Hypothesen sollen im Rahmen dieser Studie überprüft werden:

1. Die Lernleistung von Lernenden nach einer autodidaktischen Lernphase mit dem AR-Lernmittel ist mindestens so gut wie mit dem herkömmlichen Lernmittel.
2. Lernende mit dem AR-Lernmittel sind motivierter als Lernende mit dem herkömmlichen Lernmittel.
3. Lernende mit dem AR-Lernmittel bewerten die User Experience positiver als Lernende mit dem herkömmlichen Lernmittel.
4. Lernende mit dem AR-Lernmittel geben häufiger an, ihr Lernmittel zukünftig (autodidaktisch) nutzen zu wollen als Lernende mit dem herkömmlichen Lernmittel

Im Folgenden wird die Nutzerstudie vorgestellt, die zur Untersuchung der Fragestellung durchgeführt wurde.

3 Methode

Auf den Hypothesen aufbauend wurde eine Nutzerstudie geplant, in der das Lernen mit einem AR-Lernmittel mit dem Lernen mit einer inhaltlich äquivalenten Papier-Anleitung interindividuell verglichen wurde. Die Teilnehmenden waren gleichmäßig (je 16 Teilnehmende) auf die zwei Bedingungen aufgeteilt und wurden nicht informiert, dass es sich um einen Vergleich unterschiedlicher Lernmittel handelt.

Insgesamt nahmen 35 Freiwillige am Versuch teil, aus denen 32 vollständige Datensätze in die Analyse eingingen. Die Stichprobe hatte ein Durchschnittsalter von 22.03 Jahren ($SD = 9.20$) und identifizierte sich größtenteils als männlich (28 männlich, 4 weiblich). Die Teilnehmenden waren hauptsächlich technische Auszubildende und hatten keine Erfahrung mit dem entsprechenden Cockpitelement sowie nur in wenigen Fällen (drei innerhalb der AR-Gruppe) Erfahrung mit AR.

3.1 Aufbau der AR-Anwendung

Bei dem für die Studie gewählten Bedienelement handelt es sich um ein CMA-3000, ein verbreitetes Navigational Management System (NMS), das als Eingabekonsole unter anderem in Trainingshubschraubern wie dem H135 Verwendung findet. Ein solches NMS bietet Funktionen zur Anpassung des Autopiloten und des Funksystems über verschiedene Tasten, Warnleuchten und ein Display.

Der Hardware-Aufbau des AR-Demonstratorsystems bestand aus einer MS HoloLens 2, die mit einem Clicker der HoloLens 1 kombiniert wurde. Außerdem wurde eine originalgetreue Replik eines CMA-3000 NMS verwendet, die es den Probanden ermöglichte reale Eingaben zu tätigen. Für die Entwicklung der Software wurde die Unity

3D Game Engine mit der Version 2019.4.10 [Un19] eingesetzt. Zusätzlich wurde das Mixed Reality Toolkit der Version 2.4.0 [Mi21] für HoloLens-spezifische Funktionalitäten und die Vuforia Engine der Version 8.5.9 [PT21] für Bilderfassung in Unity integriert.

Mithilfe der Vuforia Engine wurde über die HoloLens 2 die räumliche Position des NMS erfasst und ermöglicht virtuelle Inhalte nah am Gerät zu verankern. Für die Interaktion mit dem System wurde eine Eingabemethode verwendet, bei der die Orientierung des Kopfes zur Selektion und ein separates Medium, hier der HoloLens Clicker, zur Bestätigung verwendet werden. Diese Interaktionsweise wurde gewählt, da Vorstudien gezeigt haben, dass Nutzende Schwierigkeiten mit der nativen Gestensteuerung der HoloLens hatten. Um die Auswahl kleinerer digitaler Bedienelemente zu erleichtern, wurde die IntenSelect Methode [HKP05], eine Variante der Flashlight-Selektion, gewählt. Hierdurch ist die Selektion mittels Kopforientierung weniger anfällig gegenüber Ungenauigkeiten und kleineren unbeabsichtigten Kopfbewegungen. Auch die IntenSelect-Parameter wurden durch Vorversuche optimiert.

3.2 Inhalt und Funktionalität der Lernmittel

Mit dem entwickelten Demonstrator sollten Nutzende lernen, wie man eine Route mit verschiedenen Wegpunkten, Flugkorridoren und Details zu Start und Landung korrekt in das NMS eingibt. Damit auch unerfahrene Nutzende sich schnell im System zurechtfinden, wurde zu Beginn der Anwendung ein Tutorial integriert, das die Nutzenden erfolgreich abschließen mussten, bevor ihnen die eigentlichen Inhalte gezeigt wurden. In dem Tutorial wurden unter anderem die Navigation und Interaktion mit der AR-Anwendung erklärt und abgeprüft.

Die AR-Anwendung führt Nutzende durch die einzelnen Schritte der Routeneingabe und unterstützt sie dabei durch verschiedene Hilfestellungen. Links neben dem NMS wird ein Anwendungsfester eingeblendet (vgl. Abb. 2). Hier können Nutzende einsehen, was das nächste Zwischenziel der Routeneingabe ist und welche einzelnen Handlungsschritte (i.d.R. Drücken bestimmter Tasten am NMS) dazu notwendig sind. In manchen Fällen werden außerdem schriftliche Hintergrundinformationen bereitgestellt. Außerdem werden die aktuell relevanten (realen) Tasten des NMS durch eine virtuelle weiße Umrandung hervorgehoben (vgl. Abb. 1). Darüber hinaus kann der Nutzende eine Animation abspielen, in dem eine stilisierte Hand (vgl. Abb. 1) die erforderlichen Handlungsschritte demonstriert.

Die Anwendung wechselt automatisch zwischen zwei Modi. Während der ersten Routeneingabe erhalten Nutzende alle verfügbaren Hilfestellungen, sodass sie keine wichtigen Informationen verpassen. Nachdem die komplette Routeneingabe einmal durchlaufen wurde, geht das System in den Trainingsmodus über. Hier können die Nutzenden nun selbstständig bestimmen, ob und welche Hilfestellungen sie angezeigt bekommen möchten. Zu Beginn einer neuen Route sind alle Hilfestellungen ausgeschaltet

können sie unabhängig voneinander und zu jeder Zeit durch Drücken von Buttons in der Anwendung an- oder ausschaltet werden. Dadurch soll den Nutzenden die Möglichkeit gegeben werden, die Schwierigkeit der Aufgabe selbst zu steuern, sowie das eigene Wissen zu überprüfen. Dabei standen fünf unterschiedliche Trainingsrouten zur Verfügung.

Analog zur AR-Anwendung wurde eine Papieranleitung entwickelt, deren Textanteile identisch zu denen der AR-Anwendung waren. Die Tastenhervorhebungen und Animationen der AR-Anwendungen wurden im Papier-Lernmittel durch eine schematische Zeichnung des NMS mit Tastenbeschriftungen ersetzt (vgl. Abb. 1). Dadurch war die inhaltliche Äquivalenz der Lernmittel gegeben.

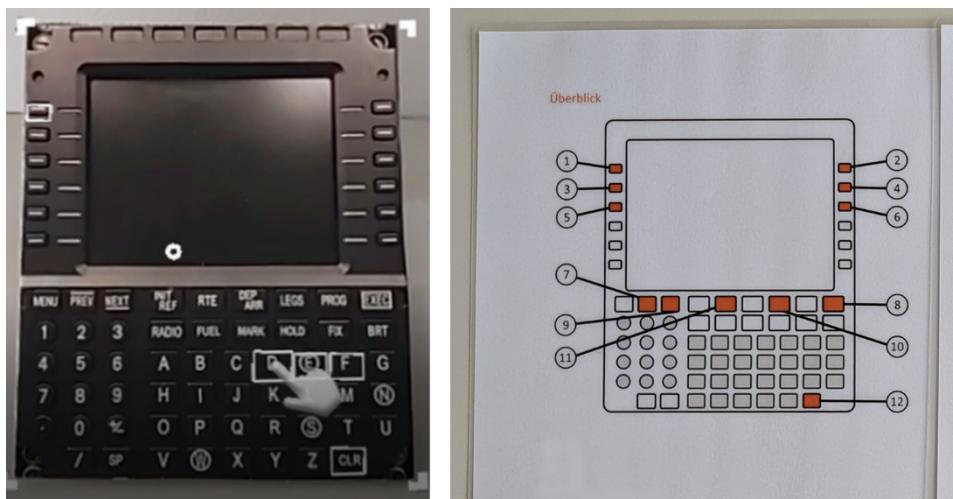


Abb. 1: Demonstration der gegebenen Hilfestellung zum Auffinden relevanter Tasten mithilfe der AR-Brille (links) bzw. der Papieranleitung (rechts).

3.3 Versuchsablauf und Messgrößen

Nach der Begrüßung und Einführung durch den Versuchsleiter bekamen die Teilnehmenden die Instruktion, solange mit dem präsentierten (AR- oder Papier-) Lernmittel zu lernen, bis sie sich sicher bei der Routeneingabe fühlten. Für diese Lernphase hatten die Probanden kein Zeitlimit. Nach der Lernzeit füllten die Teilnehmenden zunächst die Fragebögen aus. Anschließend wurden sie gebeten, ohne Unterstützung durch ihr Lernmittel eine Route einzugeben. Abschließend führte der Versuchsleiter ein halbstrukturiertes Interview mit den Teilnehmenden.



Abb. 2: Versuchsaufbau mit exemplarischer Darstellung der AR-Elemente.

Während des Versuchs wurden verschiedene Messgrößen zum Vergleich der Lernmittel erhoben. Für beide Lernmittel wurden die Anzahl der bearbeiteten Übungen sowie die Lernzeit als Nutzungsdaten gemessen. Als Lernergebnisse wurden die Anzahl der Fehler in der anschließenden Testaufgabe sowie die Zeit zur Bearbeitung der Testaufgabe gemessen. Darüber hinaus wurden subjektive Bewertungen erfasst. Zur Bewertung der empfundenen Aufgabenschwierigkeit wurde die Single-Ease-Question (SEQ) gestellt. Die Bewertung erfolgt auf einer 7-stufigen Likert-Skala mit den Polen sehr einfach und sehr schwierig [SD09]. Die User Experience wurde über den User Experience Questionnaire (UEQ) [LSH06] erfasst. Zusätzlich wurden selbst konstruierte Items verwendet, bei denen vor allem motivationale Aspekte und die Unterstützung beim selbstständigen Lernen im Vordergrund standen. Hier wurden die Teilnehmenden gebeten, ihre Zustimmung bezüglich verschiedener Aussagen auf einer 7-stufigen Skala einzuordnen. Außerdem sollten sie das Lernmittel mit einer Schulnote bewerten. Im abschließenden Interview wurden zusätzliche qualitative Daten aufgenommen, welche die gemachte Erfahrung der Teilnehmenden mit den Lernmitteln widerspiegeln.

4 Ergebnisse

Die Datenverarbeitung und –analyse wurde mit der Software R in der Version 4.0.3 durchgeführt [R 20]. Für Vergleiche zwischen den Gruppen wurde für mindestens intervallskalierte Messgrößen pauschal der Welch-T-Test verwendet. Dieses Vorgehen basiert auf Empfehlungen von [RKM11]. Als Signifikanzniveau wurde ein Alphafehler von 0,05 festgelegt. Bei der statistischen Analyse der Fragebogendaten bestand aufgrund des multiplen Testens die Problematik der Alphafehler-Kumulierung. Um das allgemeine

Signifikanzniveau von 0,05 beizubehalten, wurden deshalb die p -Werte angepasst. Dazu wurde die Bonferroni-Holm-Korrektur verwendet.

4.1 Nutzung

Bis auf eine Ausnahme bearbeiteten alle Teilnehmenden mindestens eine zusätzliche Trainingsaufgabe. Dabei hatte das Lernmittel keinen Effekt auf die Anzahl der bearbeiteten Übungen ($t(27,95) = 1,26; p = 0,218$). Im Durchschnitt lernten die Teilnehmenden der AR-Gruppe 21,92 min ($SD = 7,91$) und der Papier-Gruppe 20,69 min ($SD = 5,53$). Auch hier lässt sich kein Unterschied feststellen ($t(26,84) = 0,51; p = 0,614$).

In dem abschließenden Interview äußerten sich alle Teilnehmenden der AR-Gruppe positiv über die bereitgestellten Übungen, in der Papier-Gruppe bewerteten drei Teilnehmende sie als nicht nötig. Die AR-Gruppe wurde außerdem gebeten den Trainingsmodus (die Möglichkeit Hilfestellung an- und abzuschalten) zu bewerten. Fast alle Teilnehmenden (94%) bewerteten den Trainingsmodus dabei als positiv beziehungsweise sinnvoll.

Während der Durchführung des Versuchs und im Interview merkten die Teilnehmenden wiederholt an, dass die virtuellen Darstellungen die realen Tasten ungünstig überlagerten und die Bearbeitung der Aufgabe so erschwerten. Dies konnte teilweise durch Korrektur des Sitzes der Brille behoben werden. Allerdings ist davon auszugehen, dass die Darstellung der AR-Anwendung nicht für alle Teilnehmenden optimal war. Alle Probanden konnten die Anwendung jedoch nutzen und berichteten abseits der genannten Darstellungsschwierigkeiten keine Beeinträchtigungen durch die AR-Brille.

4.2 Lernergebnisse

In den Lernergebnissen konnten keine Unterschiede zwischen den beiden Lernmitteln festgestellt werden. Bei der Überprüfung der Lernleistung konnten die meisten der Teilnehmenden (69%) die Route fehlerfrei eingeben. Insgesamt zehn Teilnehmende machten bei der Eingabe mindestens einen Fehler, vier davon hatten vorher mit dem AR-Lernmittel gelernt und sechs mit dem Papier-Lernmittel. Ein exakter Test nach Fisher zeigt, dass diese Verteilung keine geringere Fehlerrate durch Lernen mit der AR-Anwendung zeigt (Odds Ratio = 1,767; $p = 0,352$). Auffällig war dabei, dass die meisten dieser Teilnehmenden (80%) nur einen Fehler im letzten Schritt der Eingabe machten. Nur ein Teilnehmender, der mit dem Papier-Lernmittel gelernt hatte, brach die Aufgabe ohne vollständige Eingabe ab. Bei der Analyse der Bearbeitungsdauer der Prüfungsaufgabe zeigt sich kein signifikanter Unterschied ($t(19,61) = 0,21; p = 0,839$).

4.3 Subjektive Bewertung

Die Bewertung der SEQ lag bei der AR-Gruppe bei 2,50 ($SD = 0,82$) und bei der Papier-Lernmittel-Gruppe bei 3,13 ($SD = 0,89$), was einer eher leichten Aufgabenschwierigkeit entspricht. Dabei bewerteten die Teilnehmenden in der AR-Lernmittel-Gruppe die Aufgabe als signifikant leichter als die der Papier-Lerngruppe ($t(29,81) = 2,08; p = 0,047$).

Bei der Bewertung der User Experience erhielt die AR-Anwendung durchschnittlich positive Bewertungen (Skalenwert > 0,8) auf allen Skalen. Das Papier-Lernmittel wurde etwas schlechter, aber ebenfalls eher positiv bewertet. Abb. 3 zeigt die Verteilung der Bewertungen nach Skalen und Lernmittel aufgeschlüsselt. Auf den Skalen Attraktivität ($t(28,10) = 3,27; p = 0,003; p_{cor} = 0,017$), Stimulation ($t(23,16) = 2,75; p = 0,011; p_{cor} = 0,045$) und Originalität ($t(25,64) = 3,62; p = 0,003; p_{cor} = 0,016$) wird das AR-Lernmittel als signifikant besser bewertet. Auf den anderen Skalen konnten keine signifikanten Unterschiede festgestellt werden ($p_{cor} > 0,05$).

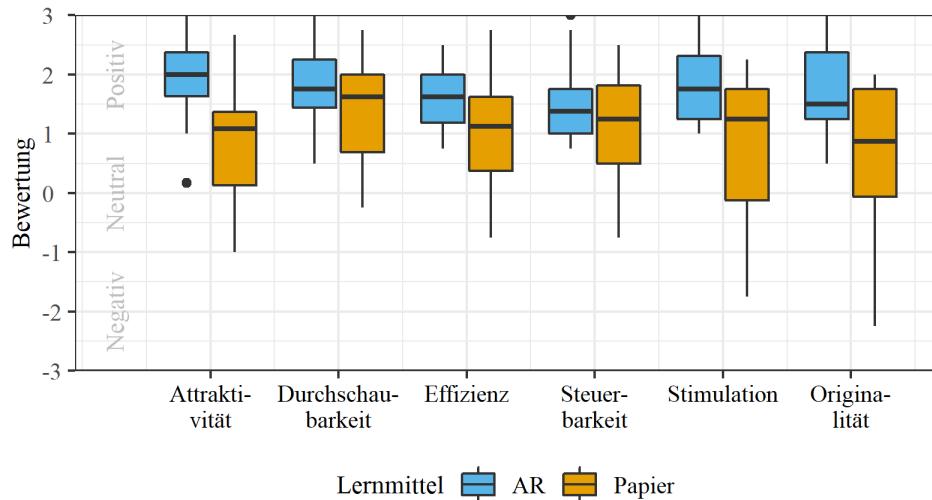


Abb. 3: Übersicht der Bewertungen auf den Skalen des UEQ.

Positive Bewertungen der AR-Anwendungen zeigen sich auch in den Ergebnissen der selbstkonstruierten Frageitems (vgl. Abb. 4). Alle Mittelwerte des AR-Lernmittels liegen deskriptiv über dem des Papier-Lernmittels. Die größten Abweichungen bei der Bewertung der Lernmittel waren bei den Items zu Motivation ($t(21,06) = 1,90; p = 0,071; p_{cor} = 0,286$), strukturiertem Lernen ($t(21,24) = 1,97; p = 0,062; p_{cor} = 0,312$) und dem berichteten Willen das Lernmittel wieder zu benutzen ($t(22,56) = 2,19; p = 0,039; p_{cor} = 0,236$) zu finden. Die statistische Auswertung zeigt aber, dass bei keinem Item nach

Korrektur um die Anzahl durchgeföhrter Tests signifikante Unterschiede vorliegen ($p_{\text{cor}} > 0,05$).

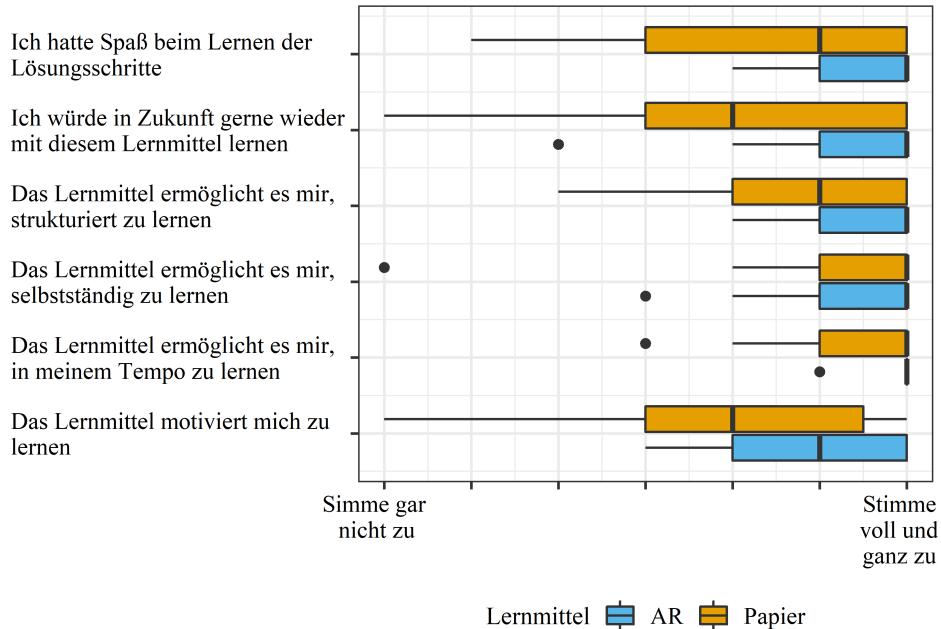


Abb. 4: Frageitems zu Motivation und Lernerfahrung.

Auch bei der Gesamtbewertung durch die Teilnehmenden in Schulnotenform konnte kein signifikanter Unterschied festgestellt werden ($t(26,61) = 0,82; p = 0,418$). Hier lagen die Mittelwerte bei 2,00 ($SD = 0,52$) in der AR-Gruppe und bei 2,19 ($SD = 0,75$) in der Papier-Gruppe, was einer guten Bewertung entspricht.

5 Diskussion und Fazit

Die Ergebnisse der Studie zeigen, dass beide Lernmittel dazu geeignet waren, die Prozedur zur Routeneingabe zu erlernen. Nach der Lernzeit war ein Großteil der Probanden in der Lage, eine Route fehlerfrei einzugeben, und die Gesamtbewertung der Lernmittel waren in beiden Gruppen positiv.

Hypothese 1 lässt sich dahingehend bestätigen, dass es keinen Unterschied in den gemessenen Lernergebnissen gab. Jedoch zeigte sich, dass Nutzende des AR-Lernmittels die Aufgabe als leichter empfanden. Vor dem Hintergrund, dass die Aufgabe insgesamt als eher leicht wahrgenommen wurde, könnte es sein, dass sich Unterschiede in den Lernergebnissen erst bei komplexeren oder längeren Aufgabenstellungen zeigen. Hier

könnte weitere Forschung ansetzen, um mehr über den Zusammenhang von Lernmittel und Aufgabenschwierigkeit und Lerndauer zu erfahren.

Im Rahmen der Bewertungen der Studienteilnehmenden konnte weder gezeigt werden, dass das AR-Lernmittel die Lernenden mehr motivierte (Hypothese 2), noch, dass die Lernenden das AR-Lernmittel zukünftig wieder verwenden wollen würden (Hypothese 4). Es zeigten sich zwar hohe Motivationswerte und Zustimmung bezüglich erneuter Verwendung, allerdings unterscheiden sie sich nicht signifikant von den Werten des Vergleichslernmittels. Auch bei diesem waren berichtete Motivation und Willen zur erneuten Verwendung eher hoch. Signifikante Unterschiede zwischen den Lernmitteln finden sich auf einigen Skalen des UEQ. Hier wurde das AR-Lernmittel als attraktiver, stimulierender und origineller bewertet. Diese berichtete bessere User Experience entspricht den Annahmen der Hypothese 3.

Die bereitgestellten zusätzlichen Übungen wurden von den Teilnehmenden genutzt und als positiv bewertet. Gleiches gilt für den implementierten Trainingsmodus der AR-Anwendung und die Möglichkeit Hilfestellungen an- und abzuschalten. Probleme bereiteten die visuellen Hilfestellungen in Form von Einblendungen über den realen Konsolentasten. Hier kam es zu optischen Verschiebungen, die bei den kleinen Tasten schnell zu Überlagerung der realen Information führten. Für die Gestaltung von ähnlichen Anwendungen mit kleinen Zielobjekten, sollte sichergestellt werden, dass entsprechende Visualisierungen stets sichtbar und robust gegen eine nicht optimale Einstellung der Hardware durch den Nutzende sind.

Es lässt sich feststellen, dass beide Lernmittel als gut bis sehr gut bewertet wurden. Die gute Bewertung des Papier-Lernmittels könnte dabei mit der Gewöhnung der Teilnehmenden an ähnliche Anleitungen zusammenhängen. Gleichzeitig könnten verschiedene Aspekte die Verallgemeinerbarkeit der Bewertung des AR-Lernmittels einschränken. So hat die Interaktion mit einer für die Teilnehmenden noch unbekannten Technologie gegebenenfalls einen positiven Einfluss auf ihre Bewertungen, während aufgetretene technische Probleme bei einigen Probanden die Urteile negativ beeinflusst haben könnten. Letztendlich könnte auch die Studiensituation, mit der einhergehenden Beobachtung, die Studienteilnehmenden dahingehend beeinflusst haben, dass sie sich länger mit dem Lernmittel beschäftigten, als sie es sonst getan hätten oder weniger Kritik übten.

Insgesamt deuten die Ergebnisse darauf hin, dass AR-Anwendungen in der Ausbildung an Cockpitelementen eine sinnvolle Trainingsergänzung darstellen können. In dieser Studie lernten die Teilnehmenden selbstständig und ohne Probleme eine Prozedur durch eine AR-Anwendung. Entscheidend ist dabei, dass die Probanden vertraut mit der Technik sind und eigenständig kleinere Probleme, wie den Sitz der AR-Brille, korrigieren können. Unter diesen Bedingungen könnten AR-Anwendungen ein besseres Benutzungserlebnis und eventuell weitere motivationale Vorteile beim Lernen bieten.

Literaturverzeichnis

- [CBM12] Carmichael, G.; Biddle, R.; Mould, D.: Understanding the Power of Augmented Reality for Learning: E-Learn 2012--World Conference on E-Learning; S. 1761–1771.
- [GP91] Garcia, T.; Pintrich, P. R.: Student motivation and self-regulated learning: A LISREL model, 1991.
- [GPB19] Garzón, J.; Pavón, J.; Baldiris, S.: Systematic review and meta-analysis of augmented reality in educational settings. In Virtual Reality, 2019; S. 447–459.
- [GSD20] Gerup, J.; Soerensen, C. B.; Dieckmann, P.: Augmented reality and mixed reality for healthcare education beyond surgery: An integrative review. In International Journal of Medical Education, 2020, 11; S. 1–18.
- [HKP05] Haan, G. de; Koutek, M.; Post, F. H.: IntenSelect: Using Dynamic Object Rating for Assisting 3D Object Selection: IPT/EGVE, 2005; S. 201–209.
- [ID18] Ibáñez, M.-B.; Delgado-Kloos, C.: Augmented reality for STEM learning: A systematic review. In Computers & Education, 2018; S. 109–123.
- [LSH06] Laugwitz, B.; Schrepp, M.; Held, T.: Konstruktion eines Fragebogens zur Messung der User Experience von Softwareprodukten. In Mensch und Computer 2006: Mensch und Computer im Strukturwandel, 2006; S. 125–134.
- [Mi21] Microsoft: Mixed Reality Toolkit. Microsoft, 2021.
- [PT21] PTC Inc.: Vuforia. PTC Inc., 2021.
- [R 20] R Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2020.
- [RKM11] Rasch, D.; Kubinger, K. D.; Moder, K.: The two-sample t test: pre-testing its assumptions does not pay off. In Statistical Papers, 2011, 52; S. 219–231.
- [Sc20] Schaffernak, H. et al.: Potential augmented reality application areas for pilot education: An exploratory study. In Education Sciences, 2020, 10; S. 86.
- [SD09] Sauro, J.; Dumas, J. S.: Comparison of three one-question, post-task usability questionnaires. In (Olsen, D. R. Hrsg.): Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, New York, NY, 2009.
- [Un19] Unity Technologies: Unity 3D. Unity Technologies, 2019.

Anreizsysteme für Studierendenpartizipation in Learning Analytics

Sven Judel¹

Abstract: In diesem Beitrag werden die Ergebnisse des Workshops zur Ermittlung von Anreizsystemen für Studierendenpartizipation in Learning Analytics vorgestellt. Ziel des Workshops war es eine Antwort auf die Frage zu finden, wie Learning Analytics für Studierende attraktiv gemacht werden kann. Dazu wurden Erfahrungen und Ideen zu fünf Unterfragen über die Interessen der Studierenden, die Kommunikation der Vorteile von Learning Analytics, dem Datenschutz, der Notwendigkeit von personalisierten Daten und wie Studierenden schon mit wenig gesammelten Daten hilfreiches Feedback gegeben werden kann, gesammelt. Daraus abgeleitet wurde festgehalten, dass ein Learning Analytics Tool, welches die Studierenden gut im Studium unterstützt, selbst den größten Anreiz darstellen kann. Um eine erste Nutzergruppe aufzubauen, welche das Tool später weiterempfehlen könnte, wurden Wege der Kommunikation der Vorteile und des Mehrwerts sowie die Einhaltung der Datenschutzregularien diskutiert und hier beschrieben.

Keywords: Learning Analytics, Anreizsysteme, Motivation, Studierende

1 Einleitung

Learning Analytics (LA) soll das Lernen verbessern [Du12]. Dafür müssen allen Personen, die am Lernprozess beteiligt sind, Informationen bereitgestellt werden, um gute, den Lernprozess fördernde Entscheidungen treffen zu können [Oc15]. Für Lehrende meint dies z. B. die Anpassung der Lehre oder Lehrmaterialien, für Studierende die Förderung des aktiven Lernens [Jo16]. Damit Studierende Feedback durch ein LA Tool erhalten können, müssen diese aktiv am LA Prozess teilnehmen. Dies beinhaltet unter anderem das Sammeln von Lerndaten, die Analyse dieser Daten und die Repräsentation der Analyseergebnisse mit der Ableitung von nächsten Lernschritten durch die Studierenden [Ch12].

Das Ziel des Workshops war das Finden einer Antwort auf die Leitfrage: *Wie kann Learning Analytics für Studierende attraktiv gemacht werden?* Dies zielte darauf ab, Ideen zu generieren, wie Studierende zur Teilnahme am LA Prozess motiviert werden können, damit diese von den Vorteilen, wie der Unterstützung des Lernprozesses und des Studienverlaufs profitieren können.

¹ RWTH Aachen, Informatik 9 (Learning Technologies), Ahornstr. 55, 52074 Aachen, judel@informatik.rwth-aachen.de

Dazu wurden zunächst 5 Unterfragen definiert, zu denen während des Workshops in je einer Brainstorming Session, gefolgt von einer Diskussionsrunde, Erfahrungen, Gedanken und Ideen gesammelt wurden:

1. Was interessiert Studierende?
2. Wie können die Vorteile von Learning Analytics Studierenden vermittelt werden?
3. Wie vermittelt man Studierenden ein sicheres Gefühl bei der Nutzung?
4. Wo reichen pseudonyme Datensammlungen?
5. Welches Feedback kann Studierenden schon mit wenig gesammelten Daten gegeben werden?

Die Antworten auf diese Fragen wurden abschließend zusammengefasst und eine Antwort auf die Leitfrage abgeleitet. Im Folgenden werden die gesammelten Antworten für die einzelnen Unterfragen sowie die Ableitung einer Antwort auf die Leitfrage vorgestellt.

2 Brainstorming und Diskussion

Zu jeder der fünf gestellten Unterfragen wurden zunächst in einer Brainstorming Session die Erfahrungen, Gedanken und Ideen der Teilnehmenden in einem Onlinedokument kollaborativ zusammengetragen. In zehn bis fünfzehn Minuten wurden Stichpunkte gesammelt und anschließend in der Gruppe weiter ausgeführt und diskutiert.

2.1 Was interessiert Studierende?

Aus den gesammelten Gedanken wurden fünf Interessenskategorien abgeleitet:

Statische Informationen: Neben dem Interesse am Thema, kann eine Einschätzung des Aufwands für Studierende ausschlaggebend für die Belegung eines Kurses sein. Daher wäre eine erfahrungsorientierte Angabe des Aufwands von Interesse, die z. B. durch Evaluationen der Lehrveranstaltungen erhoben werden kann. Die Angabe und Erinnerung an ausstehende Deadlines, wie z. B. Wahlen von Seminaren oder Praktika, die für das kommende Semester relevant sind, kann die Studienplanung weiter unterstützen. Der Frage nach dem sicheren Gefühl bei der Nutzung von Learning Analytics (LA) vorgreifend wurde auch eine Angabe der Rollenrechte vorgeschlagen. Studierende sollten einsehen können, welche Systemrollen (Administratoren, Lehrende, Tutoren, Studierende) auf welche Daten zugreifen können.

Übersicht des aktuellen Standes: Ein Überblick über den Studienverlauf durch z. B. die Auflistung der bisher erhaltenen Credits und der geplanten Credits für das aktuelle Semester kann den Studierenden bei der Orientierung helfen. So können Studierenden für sich selbst reflektieren, ob sie den für sich selbst definierten Studienplan einhalten. Durch die Kombination mit der Angabe von ausstehenden Deadlines (siehe oben), kann die

Evaluation des aktuellen Standes und damit verbunden die Planung des nächsten Semesters erleichtert werden.

Vorschläge und Prognosen: Vorschläge im Bezug auf den Studienverlauf können auf den persönlichen und fachlichen Interessen der Studierenden, deren bewältigbaren Workload pro Semester oder den Pfaden früherer Studierender beruhen. So sollen sich Studierende Fragen nach der Studiumsbewältigung bei einem Nebenjob oder der Reaktion auf ein nicht bestandenes Fach beantworten können. Auf Kursebene können Materialien zum Erreichen konkreter Lernziele vorgeschlagen werden, bei Bedarf auch externe Materialien wie z. B. Lernspiele, die interessierten Studierenden Kursinhalte durch ein anderen Medium vermitteln, oder tiefer in die Materie eintauchen. Innerhalb eines Kurses könnte die Prognose des Erhalts der Klausurzulassung sowie ein möglicher Weg dahin interessant sein. So können Studierende die Erfolge bei der Erarbeitung der Zulassung von Beginn an erkennen und bei eventuellen Rückschlägen, wie dem schlechten Abschneiden in einer von mehreren Abgaben, eine Handlungsempfehlung erhalten, um Ausgleiche zu schaffen. Dies setzt voraus, dass die Zulassungsbedingungen zu Beginn des Kurses modelliert sind.

Soziale Bezugsnormen: Zuzüglich zur Übersicht des eigenen Fortschritts wurde die Möglichkeit der Vergleiche diskutiert. Studierenden die Möglichkeit zu geben, sich mit anderen Studierenden bzw. dem Kursdurchschnitt zu vergleichen, könnte überdurchschnittliche Studierende in ihrem Lern- und Arbeitsverhalten bestärken, bei unterdurchschnittlichen aber negative Auswirkungen haben. Untersuchungen von Bestenlisten (Leaderboards) im Kontext von Gamification zeigen einen positiven Einfluss auf die Motivation der Studierenden, sich dort gut platzieren zu wollen, [Fo16]. Unterdurchschnittlich abschneidende Studierende geben aber auch an, dass sie sich durch die Bestenliste sozialem Druck ausgesetzt fühlten, was einen negativen Einfluss auf die Teilnahme in und an der Lehrveranstaltung hatte [Li19].

Selbstbestimmung: Zur Unterstützung des Selbstregulierten Lernens könnte Studierenden die Möglichkeit gegeben werden, sich in einem LA Tool eigene Lernziele zu definieren. Das Erreichen der Ziele kann durch das Tool verfolgt, durch den Vorschlag von hilfreichen Materialien unterstützt und abschließend geprüft werden.

2.2 Wie können die Vorteile von Learning Analytics Studierenden vermittelt werden?

Als wichtigster Punkt der Kommunikation der Mehrwerte und Vorteile von Learning Analytics (LA) und LA Tools wurde der Mehrwert selbst hervorgehoben. Ein LA Tool soll die Studierenden in Ihren Kursen und dem Studienverlauf unterstützen. Wenn es dieser Aufgabe gut nachkommt, könnte es Nutzende dazu verleiten, es ihren Kommilitonen zu empfehlen.

Darüber hinaus und um eine erste Gruppe von Nutzenden aufzubauen, sollten die Vorteile durch separate Demos und Beispiele in Form von Folien oder Erklärvideos aufbereitet

werden. Es sollte klar vermittelt werden, was das Tool den Studierenden bietet und was nicht. Dabei sollten die Demos auf verschiedene Studierendenprofile ausgerichtet und entsprechend gekennzeichnet werden. Überdurchschnittlich arbeitende Studierende, die einem Kurs gut folgen, erhalten durch das LA Tool anderes und ggf. auch weniger Feedback, als Studierende, die mehr Unterstützung benötigen. Für jedes Tool, dessen Vorteile präsentiert werden sollen, muss eine Möglichkeit gefunden werden, die grundlegenden Funktionalitäten allgemein zu beschreiben und anhand von Beispielen verschiedener Nutzergruppen angepasste Demos zu konzipieren.

Um Studierende nicht mit zu viel Feedback auf einmal zu konfrontieren, wurde über die Möglichkeit, Feedback zu staffeln und erst auf Anfrage mehr anzuzeigen, diskutiert. Dies könnte zum einen eine Überforderung der Studierenden verhindern und zum anderen bei Demos die aufzubereitenden Szenarien begrenzen, da mehr auf die Art von möglichem Feedback eingegangen werden kann, als auf die mögliche Menge.

Neben den Demos durch externe Materialien wurden auch integrierte Touren diskutiert. Diese sollten Studierende auf die wichtigsten Funktionalitäten hinweisen ohne zu viel Zeit bei der Erstnutzung zu fordern. Sie sollten ebenfalls wiederholbar sein, falls sie zunächst weggeklickt wurden oder später wiederholt werden sollen. Stellt das LA Tool Funktionen zur Unterstützung des Selbstregulierten Lernens bereit, sollten die Studierenden in diese eingewiesen, ihnen aber auch Raum zum Selbstentdecken und Kennenlernen gegeben werden.

Des Weiteren wurde auch bei dieser Diskussion der Datenschutz aufgegriffen und im Workshop als Überleitung zur nächsten Unterfrage genutzt.

2.3 Wie vermittelt man Studierenden ein sicheres Gefühl bei der Nutzung?

Diese Frage sollte Ideen sammeln, wie Studierenden die Einhaltung sämtlicher Datenschutzregularien kommuniziert werden kann. Dabei wurde auch darüber diskutiert, wie Studierenden die wichtigsten Aspekte zusammengefasst werden können, sodass die vollständigen Datenschutzerklärungen nur bei Bedarf und mehr Interesse konsultiert werden müssen. Dafür wurden dedizierte Merkblätter vorgeschlagen, die einzelne Teilaspekte nicht nur rein textuell sondern, wenn möglich, auch visuell aufbereitet auf einer Seite wiedergeben. Hervorgehoben wurden die Aspekte *Datentransparenz* (welche Daten werden aus welchem Grund erhoben, wie werden die Daten verarbeitet), *Datensicherheit* (wo werden die Daten gespeichert, wer kann auf die Daten zugreifen), *Anonymität* (sind Rückschlüsse auf einzelne Personen möglich und wenn ja, warum) und *Datenhoheit* (expliziter Ausschluss bestimmter Nutzungen, Recht auf Löschung). Den Studierenden ein Verständnis der Datenverarbeitung zu vermitteln und wie ihnen damit Feedback oder Vorschläge durch ein Learning Analytics Tool gemacht werden, könnte ebenfalls die Akzeptanz der Rückmeldungen steigern. In diesem Kontext wurde auch die Veröffentlichung des Quellcodes diskutiert. Dies bietet jedoch nur den Studierenden einen Mehrwert, die Quellcode lesen können. Zudem ist durch die Veröffentlichung von Code per se nicht gewährleistet, dass exakt dieser ausgeführt wird.

2.4 Wo reichen pseudonyme Datensammlungen?

Ziel dieser Frage war es herauszufinden, ob für bestimmte Analysen das Erheben personalisierter Daten notwendig ist. Nach erneutem Durchsehen der in Abschnitt 2.1 beschriebenen, gesammelten Interessen wurde keine solche Notwendigkeit festgestellt. Es reicht aus, Interaktions- und Leistungsdaten anhand eines Pseudonyms miteinander in Verbindung zu bringen.

2.5 Welches Feedback kann Studierenden schon mit wenig gesammelten Daten gegeben werden?

Um Studierenden schon früh den Mehrwert von Learning Analytics (LA) aufzuzeigen, wurden Ideen für frühes Feedback, bzw. Feedback mit wenigen erhobenen Daten gesammelt. Die Nutzung von Kursmaterialien kann beispielsweise verwendet werden um zu prüfen, ob Studierende die Informationen zur Kursorganisation oder die Basisliteratur angesehen oder runtergeladen haben. Abhängig vom ausgegebenen Material kann ab der zweiten Kurswoche Feedback zum Fortschritt im Kurs gegeben werden. Damit wären auch Vergleiche mit dem Kursdurchschnitt oder anderen Teilnehmenden möglich. Die Diskussion über mögliche Einflüsse solcher Vergleiche wurde in Abschnitt 2.1 beschrieben.

Wenige Daten können auch dann vorliegen, wenn Studierende länger nicht mehr auf ein Hochschulsystem zugegriffen haben (wobei nicht die Menge der Daten, sondern konkret die letzten Zugriffszeiten analysiert werden sollten). Das LA Tool könnte sich z. B. per Mail bei solchen Studierenden nach dem Verbleib erkundigen und beim nächsten Zugriff anzeigen, was verpasst wurde, bzw. wie man sich auf den aktuellen Stand bringen kann. Durch den Vorschlag einer Bearbeitungsreihenfolge könnte die Wiedereinstiegshürde gesenkt werden.

3 Zusammenfassung und offene Punkte

Aus den Ideen und Gedanken zu den fünf Unterfragen lässt sich als eine Antwort auf die Frage, wie Learning Analytics (LA) für Studierende attraktiv gemacht werden kann, ein gutes LA Tool, welches die Interessen der Studierenden bedient, ableiten. Dies beinhaltet mit guten Analysen und passendem, hilfreichem Feedback die Planung des Studiums bzw. die Durchführung eines Kurses zu unterstützen. Dabei soll den Studierenden klar kommuniziert werden, dass alle Datenschutzvorgaben eingehalten werden und sie die Kontrolle über ihre eigenen Daten haben. Die Idee ist, dass je besser ein LA Tool ist, Studierende eher geneigt sind, es Kommilitonen zu empfehlen. Um eine erste Nutzergruppe aufzubauen, können durch Demos und Beispiele, angepasst auf verschiedene potentielle Studierendenprofile, die Vorteile des LA Tools dargestellt werden. Durch kurze Merkblätter, welche die wichtigsten Aspekte des umgesetzten

Datenschutzes zusammenfassen, könnte Studierenden ein sichereres Gefühl bei der Nutzung gegeben werden.

Die praktische Umsetzung dieser Punkte muss für jedes LA Tool selbst erarbeitet werden. Dies beginnt der bei Erhebung der konkretisierten Interessen der Studierenden, über die Ableitung der Funktionalitäten des Tools, möglicher Profile Studierender und der Ausgestaltung der Demos. Einer guten Umsetzung der Analysen sollte eine Evaluation des durch das Tool gegebenen Feedbacks folgen um dieses bei Bedarf anzupassen und so eine gute Qualität sicherzustellen.

Literaturverzeichnis

- [Ch12] Chatti, M. A.; Dyckhoff, A. L.; Schroeder, U.; Thüs, H.: A reference model for learning analytics. In International Journal of Technology Enhanced Learning, vol. 4, 5, 318–331, 2012.
- [Du12] Duval, E.: Learning Analytics and Educational Data Mining. <https://erikduval.wordpress.com/2012/01/30/learning-analytics-and-educational-data-mining/>, 19.09.2021
- [Fo16] Fotaris, P.; Mastoras, T.; Leinfellner R.; Rosunally, Y.: Climbing Up the Leaderboard: An Empirical Study of Applying Gamification Techniques to a Computer Programming Class. In The Electronic Journal of e-Learning, vol. 14, 2, 94 - 110, 2016.
- [Jo16] Johnson, L.; Adams Becker, S.; Cummins, M.; Estrada, V.; Freeman, A.; Hall, C.: NMC Horizon Report: 2016 Higher Education Edition. The New Media Consortium, 2016.
- [Li19] Lim, L.; Dawson, S.; Joksimović, S.; Gašević, D.: Exploring students' sensemaking of learning analytics dashboards: Does frame of reference make a difference? In Proceedings of the 9th International Conference on Learning Analytics & Knowledge (LAK19). Association for Computing Machinery, New York, NY, USA, 250–259, 2019.
- [Oc15] Ochoa, X.: Visualizing Uncertainty in the Prediction of Academic Risk. In Proceedings of the First International Workshop on Visual Aspects of Learning Analytics collocated with 5th International Learning Analytics and Knowledge Conference (LAK 2015), Poughkeepsie New York, USA, 7 – 13, 2015.

Nuts & Bolts of Educational Research Projects

Die Technik hinter den Kulissen der digitalen Bildungsforschung

Birte Heinemann¹, René Röpke¹, Matthias Ehlenz¹, und Ulrik Schroeder¹

Abstract: Dieser Workshop beschäftigt sich mit der technischen Seite der Lernforschung und angrenzender Felder, denn der Forschungsprozess im Bereich des technologiegestützten Lernens ist gekennzeichnet von technischen Entscheidungen, Hindernissen und Stolpersteinen. Die Wahl der Technologie, ebenso wie die Vernetzung und die Entwicklung werden selten in Publikationen erläutert, sodass viele Gruppen vor ähnlichen Hürden stehen. Das Ziel dieses Workshops ist eine offene Diskussion über Konstruktionen, Ideen und technische Probleme anzuregen und zu ermöglichen mit der Absicht ein Netzwerk im Gebiet der Bildungstechnologien aufzubauen.

Keywords: Lernumgebungen, Architekturen, Anwendungen, Multimodale Daten, Infrastrukturen.

1 Motivation

Bildungstechnologien bieten nicht nur didaktisch eine große Bandbreite an Möglichkeiten, auch technisch lassen sich Entwicklungen auf verschiedene Arten umsetzen. Von der Integration verschiedener Komponenten über die Planung von Schnittstellen und Architekturen werden die Herausforderungen immer größer je weiter ein Projekt fortschreitet und je umfangreicher es wird. Doch die Dokumentation zu solchen Fragestellungen ist nur selten aufbereitet, schlüssig erfasst und detailliert beschrieben. Im Gegensatz dazu werden Methoden, Auswertungen und Experimentaldesigns in Publikationen sauber dokumentiert erfasst und diskutiert. Der häufig beschränkte Seitenumfang in Veröffentlichungen rückt andere Ergebnisse statt technischer Details in den Fokus. Selten werden Argumentationen und Überlegungen für bestimmte technische Entscheidungen umfassende Anteile in Veröffentlichungen zugeschlagen. Plattformen und Möglichkeiten für einen Austausch und eine Vernetzung sind rar.

In Abb. 1 ist ein Beispiel für eine komplexe Architektur im Bereich der Learning Analytics dargestellt. Doch nicht nur in diesem Bereich stellen wir die zuvor beschriebenen Lücken fest, sondern auch bei der Entwicklung in anderen Projekten werden Entscheidungen getroffen, Hindernisse gefunden und überwunden. Forscher arbeiten daran Forschungs-instrumente einzubinden, Tools und Labore zu vernetzen und die Möglichkeiten der Digitalisierung für die Bildung zu nutzen. Welches

¹ RWTH Aachen, Informatik 9 (Learning Technologies), Ahornstr. 55, 52074 Aachen, {heinemann, roepke, ehlenz, schroeder}@informatik.rwth-aachen.de

Netzwerkprotokoll eignet sich am besten für die Kommunikation zwischen bestimmten Geräten, welche Möglichkeiten haben wir um verschiedene Labore, Sensoren und Komponenten miteinander zu vernetzen? Welche Programmiersprache eignet sich für einen Anwendungsfall? Welche Anforderungen haben unterschiedliche Nutzer- und Lernergruppen? Welche Datenformate gibt es im Bereich Bildungstechnologien?

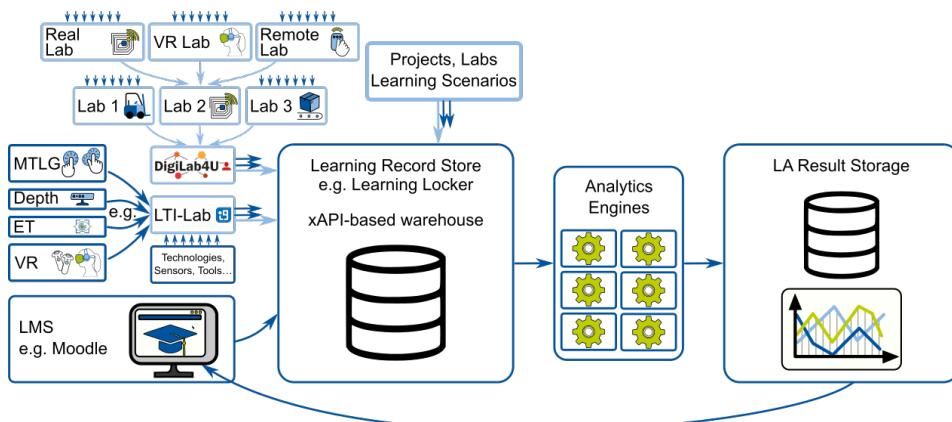


Abbildung 1. Komplexe Learning Analytics Architektur im Forschungsprojekt DigiLab4U

Viele Institutionen stehen vor denselben Hürden, Aufgaben und Möglichkeiten. Dieser Workshop soll eine Plattform zur Vernetzung anbieten, die Möglichkeit geben über technische Herausforderungen zu diskutieren oder Entscheidungsprozesse, Argumentationen und eine systematische Auswahl der technischen Werkzeuge zu reflektieren und mit anderen Wissenschaftlerinnen und Wissenschaftlern zu diskutieren. Auf diese Weise können sich Best Practices etablieren, Lösungen ausgetauscht und neue Herausforderungen identifiziert, diskutiert und bearbeitet werden.

2 Ziele des Workshops

In Anklang an das diesjährige Thema der Konferenz „Digitale Lehre zwischen Präsenz- und Online-Betrieb“ sollen die Aspekte der Vernetzung, Digitalisierung von Lernsettings und Multimodalität auf technischer, aber auch konzeptioneller Ebene im Fokus des Workshops stehen. Ziel des Workshops ist der Austausch zu Herausforderungen, Entscheidungsprozessen und Erfahrungen in der Entwicklung komplexer und multimodaler Architekturen für Forschungsprojekte im Bereich der Bildungstechnologien.

Das übergeordnete Ziel des Workshops ist es, die technische Black Box der Forschungsprozesse im Feld der Lerntechnologien zu öffnen und Projekte, Probleme und Ideen auf technischer Ebene ausführlich diskutieren zu können. Dies soll insbesondere vier Zielen dienen:

- **Dokumentation des Prozesses der Toolentwicklung:** Doktorandinnen und Doktoranden entwickeln auf der Suche nach Antworten auf ihre Forschungsfragen „ganz nebenbei“ spannende Tools und Werkzeuge. Hier sollen sie Gelegenheit erhalten, diese und deren Entwicklungsprozess vorzustellen. Dabei geht es insbesondere auch, um getroffene Technologieentscheidungen und deren Begründungen sowie die Darstellung von Irrwegen und Sackgassen.
- **Lerneffekte und Erkenntnisse auf dem Weg:** Durch die o.g. Vorträge soll Doktorandinnen und Doktoranden, die noch am Anfang ihres Weges stehen gleich auf dreierlei Weise geholfen werden. Sie werden darauf vorbereitet hinreichend zu durchdenken, welche Fragen sie sich selbst bei der Entwicklung des persönlichen Toolsets stellen sollten, sie lernen aus den Fehlern anderer und erfinden das Rad nicht jedes Mal aufs Neue und schließlich wird ihnen die Angst genommen, es einfach anzugehen und mit der Planung einzusteigen.
- **Diskussion und Austausch:** Durch die Vorträge zu akzeptierten Beiträgen ergeben sich Einstiegspunkte für den Austausch zwischen den Teilnehmenden. Forschende, deren Prozess noch nicht so weit gereift ist, dass sie ihn hier präsentieren möchten, erhalten so die Gelegenheit eben jene Zweifel zu diskutieren und sich über Stolpersteine auszutauschen.
- **Nachhaltigkeit und Offenheit:** Durch die aktive Kommunikation über Tools und die Open-Source-Veröffentlichung erhöht sich die Chance auf deren Wiederverwendung drastisch. Ein Community-Bewusstsein könnte hierdurch langfristig zu dem Effekt führen, dass Forschungswerkzeuge auch mit der Weiterverwendung in anderen Forschungsarbeiten im Hinterkopf modular und gut dokumentiert entwickelt werden. Der Austausch und die Öffnung führen zu potenziellen Kooperationen.

Insgesamt soll dieser Workshop nicht nur Neulingen in der Forschung helfen, sich zu finden und einen (praktischen) Plan zu entwickeln, sondern auch Fortgeschrittenen die Gelegenheit verschaffen, Zweifel und Entscheidungsprobleme im eigenen Prozess mit Gleichgesinnten zu reflektieren und diskutieren und jenen in der Endphase die Möglichkeit verschaffen ihre Erfahrungen zu teilen und einen Teil ihrer Arbeit zu präsentieren, der sonst häufig nur im Hintergrund bleibt.

3 Inhaltliche Ausgestaltung

Der Workshop findet als halbtägiges Angebot mit zwei 90-minütigen Sessions statt. In der ersten Session werden unterschiedliche Impulse der Fachcommunity eingeholt und anschließend gemeinsam diskutiert. Mittels verschiedener Kurzvorträge bekommen eingeladene Akteurinnen und Akteure die Chance ihre Erfahrungen mit bildungstechnologischen Forschungsprojekten vorzustellen und von wertvollen Lesson's learned zu berichten. Die zweite Session des Workshops befasst sich mit der Untersuchung unterschiedlicher Lernsettings und beginnt mit einer interaktiven Demo eines Frameworks zur Untersuchung von Kollaboration an großen Multitouch-Tabletop-Displays. Im

Anschluss erfolgt ein Austausch zu Ansätzen, Ideen, Problemen und Fragestellungen rund um die Beforschung möglicher Lernsettings.

Eingeladene Vorträge

Lessons learned aus dem Projekt ELLI – Warum technische Innovationen auch fest in die Lehre integriert werden müssen

von Tobias Ortelt, TU Dortmund

Im Projekt ELLI/ ELLI 2 (Exzellentes Lehren und Lernen in den Ingenieurwissenschaften) wurden verschiedene technische Innovationen (Remote-Labore, Virtual-Reality-Anwendungen) entwickelt und in der Lehre erprobt und eingesetzt. Leider zeigte sich nun nach Projektende, dass diese neuen Lern-Lehr-Ansätze nur weiter genutzt werden, wenn diese fest in die Lehrpläne integriert wurden. Im Vortrag werden dazu verschiedene Ansätze gezeigt, wie man Lern-Lehr-Innovationen gewinnbringend und auch nachhaltig in die Lehre integrieren kann.

Ein JupyterHub für jede Gelegenheit - Szenarien, Potentiale und Probleme

von Andreas Gödecke, Universität Paderborn

Jupyter Notebooks erlangen nicht nur durch Microsofts hauseigenes Zertifizierungsprogramm und Googles Machine Learning Kurse einen immer größeren Stellenwert. Auch an Universitäten spielen sie eine immer zentralere Rolle - beispielsweise als vorbereitete Programmierumgebungen oder bei der Abgabe und Bewertung von Heimübungen. Doch welche Hürden sind beim Betrieb einer eigenen Infrastruktur zu bewältigen? Welche use-cases müssen bedacht und wie kann Forschung damit kommuniziert werden? In dem Vortrag sprechen wir über unsere Erfahrungen im Betrieb an der Universität Paderborn und über künftige Anwendungsmöglichkeiten.

Höhen und Tiefen der technischen Entwicklung von VR-Lernanwendungen am Beispiel des Virtuellen Klassenzimmers

von Axel Wiegke, Universität Potsdam und Birte Heinemann, RWTH Aachen

Cross-Plattform Entwicklung für Virtual Reality Headsets ist ein wichtiges Ziel für den Bildungsbereich. Nur so kann eine nachhaltige Nutzung der mühsam entwickelten Lernszenarien stattfinden. In dem Vortrag sprechen wir über den interdisziplinären Entwicklungsprozess, die technischen und didaktischen Perspektiven und den Übergang zu OpenXR – endlich ein robustes und zuverlässiges Ökosystem für die Erstellung von VR-Lernanwendungen?

Das didaktische “Support-KIT” des KI-Campus: Planungshilfe für den Einsatz digitaler Lernmaterialien in der Hochschullehre

Julia Hense, Ina Stein, Cordula Torner

Abstract: Wie lassen sich digitale Lernmedien sinnvoll in hochschulische Lehrveranstaltungen integrieren? Diese Frage ist nicht erst seit der Corona-Pandemie relevant, hat dadurch jedoch eine neue Dringlichkeit erreicht. Auf dem KI-Campus, einer Lernplattform für Künstliche Intelligenz (KI), werden verschiedene digitale Lernmaterialien – von Micro-Content bis MOOCs zum Thema KI- und Datenkompetenzen unter offener Lizenz angeboten. Als Forschungs- und Entwicklungsprojekt wird auf dem KI-Campus auch gezielt die Nutzung der Angebote zur Vermittlung von KI-Kompetenzen in verschiedenen Bildungsbereichen erprobt. Für den Einsatz von digitalen Lernmaterialien des KI-Campus in hochschulischen Lehrveranstaltungen wurde ein didaktisches “Support-KIT” entwickelt, das Lehrenden Hilfestellung und praktisches Wissen darüber vermittelt, wie sich digitale OER-Inhalte in eine analoge oder hybride Veranstaltung integrieren lassen. Der bisherige Arbeitsstand wurde bei der DELFI-Tagung 2021 vorgestellt und auf Basis des Erfahrungsschatzes der Workshop-Teilnehmenden diskutiert und weiterentwickelt.

Keywords: Didaktik, Hochschullehre, hybrides Lernen, constructive alignment, OER

1 Integration digitaler Lernmaterialien in die Hochschullehre

Nicht erst seit der Corona-Pandemie ist das Lernen mit digitalen Medien ein hochaktuelles Thema. Seither stellt sich die Frage, wie gutes Lernen und Lehren mit digitalen Medien gelingen kann, noch einmal in einer erhöhten Dringlichkeit. Digitale Lernmedien können viel ermöglichen, was in rein analogen Settings so nicht möglich ist. Die Bandbreite der didaktischen Einsatzszenarien ist groß.

In diesem Workshoppaper werden deshalb Einsatzmöglichkeiten digitaler OER-Lernmedien in der Hochschullehre aufgezeigt. Im Mittelpunkt der Betrachtung steht dabei das praktische Beispiel des KI-Campus, einer Online-Lernplattform zum Thema Künstliche Intelligenz (KI), die als Forschungs- und Entwicklungsprojekt aktuell vom Bundesministerium für Bildung und Forschung gefördert wird, und die Möglichkeiten, exemplarische Lerninhalte des KI-Campus in die Hochschullehre zu integrieren, sowohl fachgebunden als auch fachfremd [KI21].

KI-Kompetenzen zu vermitteln ist nicht nur fachliche Aufgabe, sondern auch gesellschaftliche Verpflichtung, um einen mündigen und selbstbestimmten Umgang mit KI-Technologien zu erzielen. Vielerorts fehlt es jedoch noch an entsprechenden Lehr- und Lernangeboten. Das legt den Gedanken nahe, bereits bestehende, digitale Angebote in die eigene Hochschullehre zu integrieren. Die Nutzung entsprechender OER können

in dieser Situation ein effektiver Ausweg sein [Br17].

Als Forschungsprojekt untersucht und erprobt der KI-Campus daher gemeinsam mit Hochschullehrenden, neue didaktische Lehr-Lernsettings, die genau darauf abzielen. Digitale Medien werden bereits in einer oder anderen Art in die hochschulische Lehre integriert. Selten jedoch werden ganze Kursangebote und fachfremde Inhalte gezielt ausgewählt und didaktisch integriert. Hier setzt das didaktische „Support-KIT“ des KI-Campus an.

2 Das didaktische “Support-KIT” des KI-Campus

Immer wieder zeigt sich in der Praxis, dass es nicht sinnvoll ist, analoge Konzepte des Lernens eins zu eins in die digitale Lernwelt zu übersetzen. Umso wichtiger ist es, gute Praxisbeispiele für die Integration digitaler Lernmedien in die eigene Hochschullehre zu finden und Hinweise dazu zu bekommen, wie diese Integration sich didaktisch gut gestalten und bewerkstelligen lässt.

Aus diesem Gedanken ist das didaktische “Support-KIT” des KI-Campus entstanden. Hier war die Überlegung ausschlaggebend, dass es wichtig ist, einen Leitfaden anzubieten, wie eine Lehrveranstaltung an der Hochschule digital angereichert werden kann und hier sowohl vollständig digitale als auch hybride Veranstaltungstypen mitzudenken.

Dabei steht der ganzheitliche Blick aus der Perspektive der Lehrenden im Vordergrund. Berücksichtigt werden die Spezifika der Hochschullehre, etwa in Form verschiedenster gesetzter Veranstaltungstypen und Vorgaben, die sich aus Studien- und Prüfungsordnungen ergeben, als auch der gesamte Planungsprozess, der mit einer Lehrveranstaltung einhergeht. Hier setzt das didaktische “Support-KIT” des KI-Campus an.

2.1 Was leistet das “Support-KIT”?

Das didaktische “Support-KIT” des KI-Campus ist eine praktische Planungshilfe für Hochschullehrende, die sich mit dem Gedanken tragen, digitale Lernmedien zum Thema KI in ihre Hochschullehre einzubinden. Da der KI-Campus ausschließlich Lernmaterial zum Thema KI und Datenkompetenzen anbietet, geht damit die Entscheidung einher, das Thema KI in die eigene Hochschullehre einzubinden. Hier ist sowohl eine fachbezogene als auch eine fachfremde Einbindung denkbar. Letzteres erfordert neben der Kompetenz, Lehrveranstaltungen, um digitale Lernmedien zu erweitern, auch fachfremde Inhalte zu vermitteln.

Das didaktische “Support-KIT” des KI-Campus liefert konkrete, praktisch erprobte

Hinweise dazu, wie diese Integration gelingen kann und worauf es zu achten gilt. Erarbeitet wurde es gemeinsam von den Instructional Designerinnen und den Lehr-Fellows des KI-Campus, einer Gruppe von Hochschullehrenden unterschiedlichster Fachdisziplinen, die erstmals Inhalte des KI-Campus in ihrer eigenen Hochschullehre eingesetzt haben. Entstanden ist das didaktische “Support-KIT” auf der Basis der Erfahrungen dieser einjährigen Erprobungsphase und richtet sich an Lehrende aller Fachdisziplinen, die nach konkreten Möglichkeiten suchen, ihre Lehrveranstaltungen gezielt mit digitalen Lernmedien zum Thema KI anzureichern.

2.2 Wie ist das didaktische “Support-KIT” aufgebaut?

Das didaktische “Support-KIT” des KI-Campus besteht aus drei Teilen. Teil I umfasst eine kurze, praxisorientierte Handreichung, die grundsätzliche Informationen zu verschiedenen Lernsettings liefert. Es geht hier darum, zunächst aufzufächern, welche Möglichkeiten es gibt, Lehrveranstaltungen didaktisch zu konzipieren und wie sich digitale Medien jeweils integrieren lassen.

Die Handreichung beinhaltet zusätzlich Empfehlungen für die praktische Umsetzung, gibt Hinweise zu weiteren Ressourcen für die vertiefte Lektüre und bietet anhand von Fallbeispielen aus der Hochschulpraxis Ideen und Anregungen dazu, wie sich die vorgestellten Ansätze praktisch umsetzen lassen [Bi96] [Ho15].

Die Fallbeispiele kommen dabei von den Fellows, die bereits Lerninhalte aus dem KI-Campus eingesetzt haben und ihre Erfahrungen dazu in der Handreichung teilen. Die weiterführenden Ressourcen und Praxistipps kommen aus der täglichen Arbeit am und um den KI-Campus.

Der zweite Teil des “Support-KIT” ist eine Checkliste für die Planung und Konzeption von Lehrveranstaltungen an der Hochschule, die mit digitalen Lernmedien aus dem KI-Campus angereichert werden sollen. Anhand von Fragensets, die sich in einen grundsätzlichen Teil zur Planung von Lehrveranstaltungen und einen vertiefenden Teil aufsplitten, der am Modell des constructive alignment orientiert ist, können Lehrende systematisch und strukturiert ihre eigenen Lehrveranstaltung planen und sichergehen, dass sie alle relevanten Aspekte berücksichtigt haben [AK01] [Bl56] [Ho21].

Der dritte Teil des “Support-KIT” besteht aus einer Legende, die an die Lerninhalte des KI-Campus geknüpft ist. Diese grafisch gestaltete Legende gibt Lehrenden, die daran interessiert sind, Lerninhalte aus dem KI-Campus in ihre Lehre einzubinden, auf einen Blick Informationen dazu, was sie von diesem Lernangebot erwarten können, an welche Zielgruppe das Angebot sich richtet, ob es für Anfänger oder Fortgeschrittene geeignet ist, eher theorie- oder eher anwendungsbezogen aufgebaut ist, welche Medien sie vorrangig vorfinden werden etc.

Aktuell wird die Legende final ausgestaltet und soll künftig für alle Kurse, die das Label “KI-Campus Original” tragen verfügbar sein. Sukzessive soll die Legende für weitere

Lernmaterialien auf dem KI-Campus verfügbar sein. Künftig ist geplant, sie direkt auf den Kursinformationsseiten zu verlinken und dort zusammen mit der Checkliste und der Handreichung, sowie weiteren didaktischen Leitfäden und Tipps des KI-Campus für den Download bereitzustellen.

3 Das “Support-KIT” in der Erprobung

Aktuell befindet sich das didaktische “Support-KIT” in der Erprobungsphase. Der aktuelle Stand ist eine Beta-Version, die von den Lehrenden des Fellowship-Programms in ihrer eigenen Hochschullehre getestet wird und im nächsten Durchlauf des Fellowship-Programms weiterentwickelt werden wird. Auf dem Workshop der DELFI2021 wurden erste Erfahrungen mit dem didaktischen Support-Kit geteilt und aus der Diskussion im Workshop Ansätze für die künftige Weiterentwicklung abgeleitet.

3.1 Erste Rückmeldungen zum “Support-KIT”

Auf der Basis der von den Fellows artikulierten Bedürfnisse wurde das “Support-KIT” zunächst in seiner Grundstruktur entwickelt und in einem zweiten Schritt im aktuellen Arbeitsstand an die Fellows zurückgespielt. Die Rückmeldungen zum Support-KIT waren hier durchgehend positiv. Struktur, Aufbau und Inhalte wurden als schlüssig beschrieben.

Die Fellows regten an, die Handreichung des “Support-KIT” um einen konkreten Abschnitt zur Formulierung von Lernzielen anzureichern und hier darauf einzugehen, wie Lernziele in ihrer Formulierung und in ihrer taxonomischen Einordnung zu verfassen sind.

Außerdem wurde angeregt, einen weiteren Abschnitt einzufügen, der sich damit befasst, aus Lernzielen Prüfungsfragen abzuleiten. Beide Aspekte wurden bereits ergänzt und werden jetzt weiter in die praktische Erprobung gegeben.

Auf Anraten der Fellows wurde die Legende des “Support-KIT” ebenfalls im Hinblick auf Visualisierungen angepasst, so dass nun anhand einer Spinnennetzgrafik sofort erkennbar ist, wie sich die verschiedenen Medien im Verhältnis auf den Kurs aufteilen und was genau erwartet werden kann (siehe Abbildung 1).



Legende: Einführung in die KI

Das Lernangebot richtet sich an:



Dieses Lernangebot ist ein:



Dieses Lernangebot ist:



Dieses Lernangebot beinhaltet:

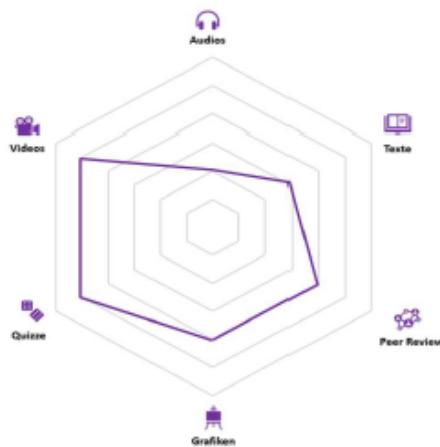


Abb. 1.: Legende des „Support-KITS“ am Beispiel des Kurses „Einführung in die KI“

Auf dem Workshop auf der DELFI2021 wurde das Support-KIT ebenfalls positiv aufgenommen. Insbesondere die Legende als Werkzeug für einen schnellen Überblick, inwiefern ein Kursangebot für die Einbindung in die eigene Lehre geeignet sein kann, wurde gut aufgenommen.

Hier wurde konkret angeregt, die Legende um den Hinweis der CC-Lizenz des Lernangebots zu ergänzen und die einzelnen Hinweise nicht nur auf Kursebene herunterzubrechen, sondern perspektivisch auch auf Modulebene, was hilfreich ist, wenn nicht ganze Kurse in die eigene Lehre integriert werden sollen, sondern lediglich einzelne Module ausgewählt werden. Eine Legende, die hier bis auf die Modulebene heruntergebrochen wird, kann dabei gute Dienste leisten.

3.2 Nächste Schritte für die Weiterentwicklung

Das didaktische “Support-KIT” wird in einem nächsten Schritt sowohl als PDF-Set zum Download auf dem KI-Campus zur Verfügung gestellt, als auch konkret auf dem KI-Campus bei den einzelnen Angeboten eingebunden werden. Auf dem Workshop anlässlich der DELFI 2021 wurde das “Support-KIT” vorgestellt. Das Feedback der Workshopteilnehmenden soll bei der Umsetzung auf dem KI-Campus ebenfalls eingearbeitet werden.

Das “Support-KIT” soll perspektivisch sukzessive weiterentwickelt werden. Denkbar ist z.B. die Legende für möglichst viele Kurse anzulegen und ggf. langfristig auch einen Fragebogen zu entwickeln, der es Lehrenden erleichtert die passenden Kurse zu finden, z.B. gekoppelt an ein Vorschlagswesen, das anhand der Antworten aus diesem Fragebogen Empfehlungen für Kurse, Kursmodule oder sogar einzelne Lernangebote auflistet, die für den Einsatz in der eigenen Lehre interessant sein könnten.

Bibliographie

- [AK01] Anderson, Lorin; Krathwohl, David Reading.: A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. Pearson education, 2001.
- [Bi96] Biggs, John: (Enhancing teaching through constructive alignment. High Education 32, 347–364. 1996.
- [Bl56] Bloom, Benjamin; Engelhart, Max; Furst, Edward; Hill, Walker; Krathwohl, David: Taxonomy of educational objectives. The classification of educational goals. Handbook I: The cognitive domain. David McKay, 1956.
- [Br17] Braßler, Mirjam; Holdschlag, Arnd; Berk, Ivo van den: Nachhaltige Zukunftsperspektiven. Erstellung von Open Educational Resources (OER) in der

Hochschullehre, 2017.

[Ho15] Hochschulrektorenkonferenz (Hrsg.) (2015): Lernergebnisse praktisch formulieren, 2015.

[Ho21] Hochschulforum Digitalisierung (Hrsg.): Digitale Prüfungen: (Digitale) Aufgaben konstruieren, 2021.

[KI21] KI-Campus, <https://www.ki-campus.org>, 13.09.2021.

Workshop on Learning Analytics

Considering student diversity with regard to assessment data and discrimination.

Clara Schumacher¹, Nathalie Rzepka² und Niels Seidel³

Abstract: This workshop focuses on current topics in the domain of learning analytics (LA). In particular, the increasing diversity of students in higher education also needs to be considered in LA. Offering adaptive support is a major aim of LA. Assessments are considered having a major impact on student learning. However, to date LA still do not sufficiently make use of assessment data related to context, activity and results. In contrary to the aim of giving the individual learner the support needed, LA are also criticized to foster prejudices. This becomes even more important in the light of a high diversity of the student cohorts. Thus, this workshop includes submissions in the domain of LA with emphasize on the impact of outliers on dropout prediction, students' perceptions of algorithms regarding grading, students' control over data collection, the application of the FAIR principles for data management to learning analytics, and the identification of indicators of group learning in collaborative software development. Furthermore, Professor Ryan Baker from University of Pennsylvania will hold a keynote on "Algorithmic Bias in Education".

Keywords: learning analytics, diversity, assessment analytics, discrimination

1 Introduction

Due to higher enrollment rates, student diversity is gaining relevance in higher education [HMU17]. The term diversity in higher education is used in the context of underrepresented student groups [Bo15] but refers also to students' individual differences regarding for instance prior education, learning strategies and motivation, and preferences [ZM14].

To offer increasingly diverse students the support needed, LA might be a suitable means as they aim at offering feedback and identifying students at-risk for timely interventions. However, to offer valid feedback on learning processes and on how to improve data collection and analysis need to be guided by theory [Se20]. Assessments are considered to determine what students are learning. Hence, assessments should not focus on tasks that are easy to analyze such as single- or multiple-choice but on tasks assessing valued learning outcomes [SB10]. Considering the debate on 21st century skills the valued

¹ Humboldt-Universität zu Berlin, Department of Computer Science, Unter den Linden 6, Berlin, 10099, clara.schumacher@hu-berlin.de

² Hochschule für Technik und Wirtschaft Berlin, Treskowallee, 8, 10318 Berlin, nathalie.rzepka@htw-berlin.de

³ FernUniversität in Hagen, Chair of Cooperative Systems, Universitätsstraße 1, 58084 Hagen, niels.seidel@fernuni-hagen.de

competences are complex to measure [We18]. In that respect, [Lo17] emphasize the need for further interdisciplinary efforts for integrating assessment models and research with LA. As a first attempt, [NCB19] developed an ontological model enhancing current xAPI approaches with contextual assessment data (e.g., assessment type; prior knowledge; current assessment performance; difficulty level; assessment results and attempts). Only when the data collected are valid, interpreted with relation to theory, and their context meaningful inferences and interventions can be deduced from LA. The validity is increasingly relevant as LA are used to identify students at-risk which might negatively affect their motivation and self-efficacy beliefs. Therefore, the validation of the analyses must consider potential discrimination of the increasingly diverse students. Neglecting the potential of discrimination can reinforce it by confronting individuals with inappropriate interventions [AP20]. However, just as serious as an inappropriate intervention can be a lack of intervention, whereby a student does not receive the adjustments needed to increase learning success.

For at-risk predictions, the technology applied as well as the variables used are important [SHR15]. Particularly, with regard to discrimination it is vital which variables were used for the algorithms. Most significant are students' GPA and internal assessment, but also demographic data (e.g., age, gender, disability, family background) are used commonly [SHR15]. [GBB19] found significant differences of model fairness for gender which depends on both, the algorithm as well as the feature set used. [RSS20] compared data of disabled and non-disabled students and found, that underrepresentation of minorities can lead to unfair classification. Further research indicated that not only unfair classification can be discriminatory, but also that the quality of predictive models can vary among sub-groups [RS19].

LA are far from being perfect and the context and processes they focus on are complex [KSG18]. Thus, the stakeholders facing LA and the related interventions need to be capable of understanding the underlying basic concepts, the possibilities but also the associated limitations and biases.

Current research might investigate how available assessment models need to be adapted to be integrated into holistic LA and what this means for the learning design of courses. How do current approaches deal with incomplete data in their models and potential self-selection biases of students agreeing to divulge personal data? How are biases of algorithms and discrimination considered in current LA systems in different countries? To what extent are the stakeholders of LA aware of biases and how can this be fostered? Is diversity considered in LA interventions without fostering discrimination but still supporting individual needs? Which variables are necessary to model student diversity without discrimination in LA? How can interdisciplinary communication and research be fostered?

2 Goals of the workshop

The overall goals of this interdisciplinary workshop are: (i) Networking of the LA community in the German speaking countries to initiate research and joint projects as well as fostering network activities beyond this workshop; (ii) Presentation of current research projects and results; (iii) Enhancing the interdisciplinarity in the working group: e.g., computer sciences, artificial intelligence, math, learning sciences and psychology as well as ethics and philosophy; (vi) Developing a joint outcome (e.g., paper, digital learning resources, event) based on the workshop discussions.

3 Submissions to the workshop

The call for papers for the workshop has been published online: <http://akla.f4.htw-berlin.de/workshop-2021/> and allowed submissions presenting work in progress and preliminary results to the community. Submissions have been peer-reviewed. In total, 6 submissions have been accepted for publication. The accepted papers deal with topics related to control over data collection, influence of data quality and completeness on prediction, students' perceptions of learning analytics and data collection as well as analysis of students' collaborative software development.

A predominant aim of learning analytics is dropout prediction. However, this might be affected by outliers of student characteristics. Hence, Daria Novoseltseva, Kerstin Wagner, Agathe Merceron, Petra Sauer, Nadine Jessel, and Florence Sedes investigate what kinds of student-outliers can be detected and if they affect dropout prediction by analysing different models using three types of algorithms. Their results indicate that removing outliers could improve dropout prediction.

Grading is work intensive and learning analytics might be a supportive means. In their study Linda Mai, Alina Köchling, Lynn Schmodde and Marius Wehner investigate students' perceptions of three scenarios (a) teacher, (b) teacher supported by learning analytics and (c) learning analytics grading. Their results indicate that students' stated satisfaction is lower with the decrease of teacher involvement in the grading process, but the perception of fairness is higher when grading is based on an algorithm. Furthermore, they found that perceived fairness is a mediator for satisfaction when grading is based on an algorithm.

In the context of adult education Philipp Krieter, Michael Viertel and Andreas Breiter used interviews, log file data and screen recordings to investigate how student privacy can be supported by giving students control over their data collection. Students used tablets for learning music and were allowed to turn on and off the permanent recording of their screen but not the additionally collected system log files. Findings of this small-scale study indicate that this approach does not affect the usefulness of the data set and that students consider it meaningful for preventing the collection of activities not related to learning.

As learning analytics are relying on a great variety of different types of data but common practices for data handling are missing, Ian Wolff, David Brioneske and Veit Köppen relate the FAIR principles for data management to the domain of learning analytics.

In their paper Benjamin Weiher, Niels Seidel, Marc Burchart and Dirk Veiel focus on indicators for identifying group learning in collaborative software development. They found a total of 31 indicators for group learning: 7 indicators for code quality, 16 indicators for participation in the group work and 8 indicators on group cohesion. These indicators can be made available on dashboards for supporting teachers in keeping track of group processes and students' individual contributions as well as using the indicators for the provision of feedback.

As programming in teams is considered essential for professional programmers this is also part of higher education courses. However, students' participation and performance in group projects is difficult to measure. Thus, Maximilian Karl and Niels Pinkwart use GitHub data from programming projects to classify students' type of teamwork (e.g., collaboration, cooperation or solo). Their classification method revealed that in their sample most projects were done by one individual student. Such additional information on student contribution can be made available for supporting grading or offering feedback.

The keynote of Professor Ryan Baker, Associate Professor at University of Pennsylvania and Director of the Penn Center for Learning Analytics, is about "Algorithmic Bias in Education". As announced by Ryan Baker his keynote will include the following topics: "The advanced algorithms of learning analytics and educational data mining underpin modern adaptive learning technologies, for assessment and supporting learning. However, insufficient research has gone into validating whether these algorithms are biased against historically underrepresented learners. In this talk, I briefly discuss the literature on algorithmic bias in education, reviewing the evidence for how algorithmic bias impacts specific groups of learners, and the gaps in that literature – both in terms of „known unknowns“ and „unknown unknowns“. I conclude with potential directions to move the research community towards better understanding how bias impacts educational algorithms, and how to address these problems so that learning systems better promote fairness and equity."

4 Workshop summary

The workshop took place as an all-day online event with up to 25 participants. In the morning the workshop organizers presented the schedule and introduced the topic of the workshop followed by the first paper presentations and their discussion. In order to stimulate further debates, a group phase with a subsequent discussion format using shared editors and breakout rooms guided by statements on current issues of learning analytics took place. After that, the remaining papers were presented and discussed. During the lunch break, there was the opportunity to follow the short paper presentations. In the

afternoon, Professor Ryan Baker (University of Pennsylvania) hold a keynote on “Algorithmic Bias in Education”. A stimulating Q&A with Ryan Baker was followed by a summing up discussion with all participants. Furthermore, joint projects and the organization of the working group were discussed.

During the workshop, many discussions and starting points for further research developed. It became clear in many discussions how much impact the rationale for collecting data for LA has on users. If collected data are used for anonymized analysis, this has a very different impact on perceptions of privacy and fairness than if the data are also used to assign grades. In the debate about fairness and satisfaction around grading, there was also an appeal to see greater opportunities in LA applications. Even the subjective grading of teachers is not without its faults, but it seems that researchers in the LA field nevertheless have a particularly cautious view of LA applications, even though they can achieve a lot.

In the group phase, it was discussed to what extent relevant indicators can be defined and measured and how these measurement models are developed. Here it becomes apparent that the application area of LA is so broad and diverse that such questions are difficult to answer globally. It is therefore important to look at the exact task in order to be able to create applications well. Furthermore, indicators should be defined from day-to-day practice and continuously evaluated.

Furthermore, two paper presentations considered collaboration in software development teams presenting different measures for monitoring group activity and progress using data from GitLab and GitHub. In order to provide formative feedback to the learner and to achieve an equal participation in the learning process teachers require support for a wide range of didactic group scenarios and collaborative programming tasks. An open question is the standardization of data formats and sharing of data sets to study group collaboration in small and large software development teams.

Ryan Baker showed in his keynote that fairness and equity can only be driven if one also knows which subgroups are affected by discrimination. He presented examples from literature in which models worked less well for subgroups than for groups with whose data they were trained or tested. He said that there has been no research on algorithmic bias using data from European educational institutions. Apart from this, the research on discrimination in LA mainly focuses on gender and ethnic discrimination, but there are many groups that have not been studied at all. Here he gives the examples of non-binary people, transgender people, international students, language dialects, different educational backgrounds of parents or people with disabilities. In order to move from this unknown bias to known bias, he suggests tackling two main obstacles: lack of data on group membership and the lack of transparency on bias and group-specific outcomes. The lack of data on group membership often results from privacy concerns and legal restrictions. While the students’ privacy should be well taken care of, Baker suggests balancing the right to privacy by the right not to be discriminated against. He concludes that much more data and information should be available on the students’ demographic in order to avoid unconscious discrimination. The subsequent discussion showed here that the

transferability of his idea in Europe is only partly realistic, as the differences in the education system and in the handling of personal data are great. The second obstacle, lack of transparency on bias and group-specific outcome should be tackled by defining standards for demonstrating effectiveness, for example in algorithmic bias reviews.

At the end of the event, the importance of interdisciplinarity was highlighted and discussed. In addition, the great responsibility that comes with having an influence on (young) people in the field of education was emphasized. Learning analytics should not be an end in itself, but should support and promote the development of students. This is realized through interdisciplinary projects, which are composed of both the technical as well as the educational practice and pedagogical perspectives.

5 Organizing team (OT) and program committee (PC)

Dr. Clara Schumacher, Humboldt-Universität zu Berlin (OT)
Nathalie Rzepka, Hochschule für Technik und Wirtschaft Berlin (OT)
Dr. Niels Seidel, FernUniversität in Hagen (OT)
Prof. Dr. Niels Pinkwart, Humboldt-Universität zu Berlin (PC)
Prof. Dr. Dirk Ifenthaler, Universität Mannheim (PC)
Prof. Dr. Katharina Simbeck, Hochschule für Technik und Wirtschaft Berlin (PC)
Prof. Dr. Josef Guggemos, Universität St. Gallen (PC)
Dr. Jakub Kuzilek, Humboldt-Universität zu Berlin (PC)
Dr. Martin Hlosta, The Open University UK (PC)
Maximilian Karl, Humboldt-Universität zu Berlin (PC)

Bibliography

- [AP20] Archer, E.; Prinsloo, P.: Speaking the unspoken in learning analytics: troubling the defaults. In *Assessment & Evaluation in Higher Education*, 2020, 45; pp. 888–900.
- [Bo15] Bosse, E.: Exploring the role of student diversity for the first-year experience. In *Zeitschrift für Hochschulentwicklung*, 2015, 10.
- [GBB19] Gardner, J.; Brooks, C.; Baker, R.: Evaluating the Fairness of Predictive Student Models Through Slicing Analysis: Proceedings of the 9th International Conference on Learning Analytics and Knowledge. ACM, New York, NY, USA, 2019.
- [HMU17] Holmegaard, H. T.; Madsen, L. M.; Ulriksen, L.: Why should European higher education care about the retention of non-traditional students? In *European Educational Research Journal*, 2017, 16; pp. 3–11.
- [KSG18] Kitto, K.; Shum, S. B.; Gibson, A.: Embracing imperfection in learning analytics: Proceedings of the 8th International Conference on Learning Analytics and Knowledge. ACM, 2018.

-
- [Lo17] Lodge, J. M. et al.: Inferring Learning from Big Data: The Importance of a Transdisciplinary and Multidimensional Approach. In *Technology, Knowledge and Learning*, 2017, 22; pp. 385–400.
 - [NCB19] Nouira, A.; Cheniti-Belcadhi, L.; Braham, R.: An ontology-based framework of assessment analytics for massive learning. In *Computer Applications in Engineering Education*, 2019, 27; pp. 1343–1360.
 - [RS19] Riazy, S.; Simbeck, K.: Predictive Algorithms in Learning Analytics and their Fairness. In 1617-5468, 2019.
 - [RSS20] Riazy, S.; Simbeck, K.; Schreck, V.: Fairness in Learning Analytics: Student At-risk Prediction in Virtual Learning Environments: Proceedings of the 12th International Conference on Computer Supported Education. SCITEPRESS - Science and Technology Publications, 2020.
 - [SB10] Shute, V. J.; Becker, B. J.: *Prelude: Assessment for the 21st Century: Innovative Assessment for the 21st Century*. Springer, Boston, MA, 2010; pp. 1–11.
 - [Se20] Sedrakyan, G. et al.: Linking learning behavior analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation. In *Computers in Human Behavior*, 2020, 107; 105512.
 - [SHR15] Shahiri, A. M.; Husain, W.; Rashid, N. A.: A Review on Predicting Student's Performance Using Data Mining Techniques. In *Procedia Computer Science*, 2015, 72; pp. 414–422.
 - [We18] Webb, M. E. et al.: Challenges for IT-Enabled Formative Assessment of Complex 21st Century Skills. In *Technology, Knowledge and Learning*, 2018, 23; pp. 441–456.
 - [ZM14] Zervakis, P.; Mooraj, M.: Der Umgang mit studentischer Heterogenität in Studium und Lehre. Chancen, Herausforderungen, Strategien und gelungene Praxisansätze aus den Hochschulen. In *Zeitschrift für Inklusion*, 2014.

Investigating the Impact of Outliers on Dropout Prediction in Higher Education

Daria Novoseltseva ¹, Kerstin Wagner², Agathe Merceron², Petra Sauer², Nadine Jessel³, Florence Sedes¹

Abstract: Many institutions of higher education seek to reduce the dropout rate through the development of models which can detect students with a high risk of dropping out to provide specific advice for them. Classical models usually ignore the students-outliers with uncommon and inconsistent characteristics although they may show significant information to domain experts and affect the prediction models. The present paper provides an analysis of students' performance and aims to answer the following research questions: What kinds of students-outliers can be detected? Do outliers affect dropout prediction models? To answer the first question, students-outliers have been detected and their characteristics have been analyzed. To address the second question, the dropout prediction models have been compared in terms of different algorithms and the presence of outliers in the data. The results of the work indicate that the performance of prediction models, particularly in terms of recall, can be improved by removing outliers.

Keywords: Outlier Detection, Dropout Prediction, Machine Learning, Models' Comparison.

1 Introduction

Nowadays, one of the main research lines in the educational field is predicting students with a high risk of dropping out from their studies. As mentioned in [Au19], university attrition negatively affects social and economic aspects, such as misuse of public and private resources or social stigmatization. Therefore, domain experts and researchers seek to develop models to predict students with a high risk of dropping out to provide specific advice for them. Classical models tend to perform better for the majority of students with common and consistent characteristics. However, they ignore the students which are not aligned with the majority - outliers. Students-outliers may show significant information to domain experts and affect the prediction models.

The recent studies in dropout prediction [Au19, Be19, BRGS19] achieve good results, however, they do not mention how they deal with students-outliers with deviating characteristics. Thus, to close this gap, we investigate the impact of outliers on dropout prediction in higher education. This study aims at responding to research questions such

¹IRIT, Paul Sabatier University, 118 Route de Narbonne, F-31062, Toulouse, France, {daria.novoseltseva, florence.sedes}@irit.fr,  <https://orcid.org/0000-0002-2987-1403>

² Beuth University of Applied Sciences Berlin, Fachbereich VI, Luxemburger Straße 10, 13353 Berlin, Germany, {kerstin.wagner,merceron, sauer}@beuth-hochschule.de

³ IRIT, Jean Jaurès University of Toulouse, 5 allées Antonio Machado, 31058 Toulouse, nadine.baptiste@irit.fr

as: RQ1. What kinds of outliers can be detected? Herein, we examine which students' characteristics are outlying; RQ2. Do outliers affect dropout prediction models? In this part, we investigate the impact on dropout prediction in two directions: the impact of outliers in general and the impact of specific classes of outliers.

The paper is organized as follows. Section 2 describes related work. Section 3 describes data collection. Sections 4 and 5 respond to the research questions and present experimental results. Finally, Section 6 concludes the paper and discusses the future work.

2 Related work

Outliers exist in almost every real-world dataset and can arise due to a variety of reasons [HA04]. An outlier is a data point that is very different from the rest of the data based on some measure. This measure strongly depends on the defined similarity/dissimilarity context. Many researchers try to avoid problems coming from outliers by removing them from the data. Meanwhile, another group of researchers believes that outliers are an important part of reality. Therefore, they try to retain such observations in their learning and testing samples [NV19].

In the educational field, outlier detection techniques attend to analyzing abnormal behavior or noisy reactions [BGS16]. According to [AASF19], only 2.25% of investigated papers are focused on outlier detection. For instance, in [Ca19] the authors perform outlier detection for student assessment in distance learning programs (e-learning) based on face-to-face exams. They define outliers as students, which do not use resources from the learning platform but still pass face-to-face exams, and detect them applying the isolation forest technique. In [ATEH12] the authors use unsupervised outlier detection algorithms (k-Nearest Neighbors and Local outlier factor) as part of educational data mining pipeline to improve graduate students' performance and overcome the problem of low grades of students who graduated. Their findings show that the detected outliers are students with excellent results in some degrees. In this current work, outliers have been defined as students who differ from the majority of students according to a certain set of features related to students' performance.

Dropout prediction in higher education is an active research topic that has been recently investigated in many studies [OM18]. The increasing rate of dropout negatively affects social aspects, such as reducing the number of people with higher education, and economic aspects, such as financing students who do not complete their studies as mentioned in [Au19, Be19]. To produce better results, existing works in dropout prediction examine various features (e.g., sociodemographic or performance), machine learning algorithms models (e.g., classification or clustering), and study cases (e.g., online courses or specific courses). Aulck et al. [Au19] predict graduation and re-enrollment of students in a US university. The authors use data of students after the first calendar year and achieve the best results with logistic regression. Furthermore, this study shows that including sociodemographic features in the prediction model hardly

improves the results. Berens et al. [Be19] develop an early detection system to predict students' dropout in German universities. AdaBoost is implemented to improve the results of logistic regression, random forest, and neural networks. Baneres et al. [BRGS19] apply naive Bayes, decision tree, k-Nearest Neighbors, and support vector machines to performance features of students from a fully online university in order to identify at-risk students as soon as possible. All these works have good prediction results, however, they do not mention whether they deal with outliers. This work aims to understand which kinds of outliers can be detected and whether outliers impact dropout prediction models.

3 Data collection

This work uses data of students' performance collected during their study in a six-semester bachelor's degree program at a German university. The data includes students from three educational programs, who started their study from winter 2012 to summer 2019, and have dropped out or graduated. Each degree program has a curriculum that contains a list of planned courses for each semester. Students may follow this curriculum or not. For instance, they can enroll in courses from the 1st semester when they are in the 2nd semester, and vice versa. The data records contain courses taken by each student including the earned grade and the respective semester as well as the semesters of graduation or dropout for each student. The grade scale for passing a course is [1.0; 1.7; 2.0; 2.3; 2.7; 3.0; 3.3; 3.7; 4.0], where the best grade is 1.0 and the worst is 4.0. If a student fails the examination, the grade is 5.0. Students may enroll in courses without taking the exam. To graduate, students must complete all mandatory courses and a program-specific number of electives with a maximum of three exam attempts.

To make the analyses applicable across study programs, we used global features, i. e. features that can be generated independently of study programs, as opposed to program-specific features e.g. achieved grade in the course *B01*. We considered features such as average grades and numbers of courses passed, failed, and enrolled per semester, similar to [Au19, Be19]. In terms of the number of courses passed, we made a further distinction: whether students passed them as planned or earlier, or later than planned. Empty values that have occurred during aggregation have been handled as follows: if no grade was obtained in the respective semester, the value of the grade is set to 6.0; if no course was passed, failed and/or enrolled in, the respective total is set to 0. Since the number of planned courses is program-specific (e.g. 5 planned courses in S2 for program 1, and 6 planned courses for program 2), the features related to courses were converted to proportions by division by the respective number of planned courses.

Outlier analysis and dropout prediction have been conducted for students after their 1st and 2nd semesters of study. Fig.1a depicts the analyzed features and their descriptive statistics. The codes of the features are presented as follows: **S1** and **S2** correspond to the 1st and 2nd semesters, **Av_grade** is the mean semester grade for passed courses, **F_ex** is the proportion of failed exams, **En_ex** is the proportion of courses which students enrolled in but did not take the exam, **P_ex_p**, **P_ex_d**, **P_ex_a** are the proportions of

passed exams according to the plan, with delay, and in advance respectively.

In contrast to the data after the 1st semester, data after the 2nd semester reveals the development of performance over time, e.g. whether more or fewer courses were taken or whether exam grades improved. We removed students who dropped out after the 1st semester due to the absence of data for them in the 2nd semester. The considered dataset includes 1 809 students, among them, 1 007 students are labeled *Graduate*, and 802 students are labeled *Dropout*. The distribution of labels according to the educational program is shown in Fig. 1b.

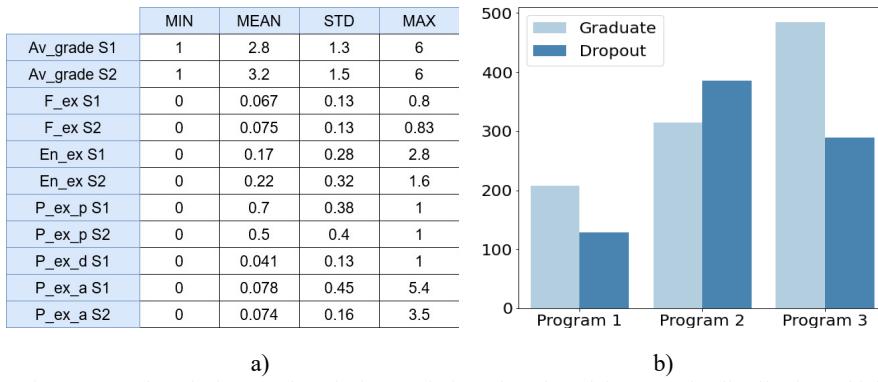


Fig. 1: Data description: a) descriptive statistics of analyzed features; b) distribution of labels Dropout/Graduate for three educational programs

4 RQ1: What kinds of outliers can be detected?

4.1 Methodology

The definition of outliers strongly depends on the context of the analysis. However, outliers always correspond to the two characteristics: (i) they are different from the norm with respect to their features; (ii) they are rare in a dataset compared to normal instances [GU16]. Therefore, such aspects as norm and rareness should be determined.

To specify the norm, the similarity metrics between objects have to be chosen. In the current work, we defined outliers as students which differ from the majority of students according to the set of numerical features described above. This is the task of finding global outliers - objects which are outlying for the entirety of the dataset in which it is contained. Furthermore, the ground truth indicating which students are outliers is absent and non-evident. Hence, only unsupervised algorithms can be used. Thus, we considered one of the most commonly used unsupervised algorithms to detect global outliers: distance-based k-Nearest Neighbors (kNN). The kNN outlier detection algorithm is described in [AP02] and is based on the idea that outliers are data points are distant from their neighbors and have sparse neighborhoods. The algorithm requires one

predefined parameter – the number of nearest neighbors k , which should be in the range $10 \leq k \leq 50$ [GU16]. In our case, the outliers obtained with various values of k significantly intersected, therefore we focused on $k = 50$. Concerning the rareness aspect, we investigated two assumptions: (i) 5% of students are outliers; (ii) 10% of students are outliers. The outliers scores obtained by the kNN algorithm were ranked, and 5% or 10% of the students with top ranks were considered as outliers.

To understand the nature and the kind of the detected outliers, k-means clustering has been run, and the number of clusters has been chosen according to the elbow curve. The mean values of features for each cluster have been examined and compared, which allowed emphasizing the outlying characteristics.

4.2 Results

The distribution of labels (Dropout/Graduate) of detected outliers is as follows: 47/53% for the 5% assumption, 62/38% for the 10% assumption. Hence, the number of outliers with the dropout label is increasing when increasing the threshold of the assumption.

Running the k-means clustering algorithm, eight main clusters were obtained for both assumptions. The clusters for the two assumptions overlap significantly, and clusters 1 to 3 for the 5% assumption case are subsets of clusters 1 to 3 for 10% assumption. Therefore, we focus on the clusters obtained with the 10% assumption in this study. These clusters are presented in Tab.1. The table includes information about the number of outliers in each cluster (N), their labels (Dropout/Graduate), the number of outliers that are also detected as outliers in 5% assumption ($N_{5\%}$), and mean cluster values for each feature. Although the features for outlier detection and clustering were standardized, Tab. 1 includes average values for non-standardized features of the clusters to help the interpretation:

- **Cluster 1 - Intense S1 and missed S2.** Students with an extremely high number of passed exams in S1, but without any passed exams in S2: the proportion of passed exams in S1 according to the plan is 0.64, while the proportion of passed exams ahead of the plan is 2.39. Meanwhile, the average grade in S2 is 6, which points out the fact that students from this cluster have not passed any exams in S2.
- **Cluster 2 - Intense S1.** Students with an extremely high number of passed exams in S1: the proportion of passed exams according to the plan is 0.99 and passed exams ahead of the plan is 2.99. Unlike cluster 1, this cluster contains students who have performance in S2. The majority of students from this class have the label graduate.
- **Cluster 3 - Intense S2.** Students with an extremely high number of passed exams in S2. The proportion of passed exams according to the plan is 1, the proportion of passed exams ahead of the plan is 2.75, the proportion of passed exams behind the plan is 0.83. In this small cluster all students have the label graduate.

- **Cluster 4 - Procrastination in S1.** Students with a bad performance in S1, which try to reestablish in S2: they pass exams from S1 in S2 with average marks. The labels of students in this cluster are mixed.
- **Cluster 5 - Average performance and intention to pass exams in advance.** Students with average grades during all semesters, who prefer to pass exams ahead of the plan in each semester instead of exams that correspond to the plan. The labels of students in this class are mixed.
- **Clusters 6-8 - Various types of bad performance.** These students belong to different clusters, and have various characteristics: some of them have low grades in all semesters, some have a high number of failed exams, others have a high number of enrollments without attending the exam. The majority of them have a dropout label.

Cluster	1	2	3	4	5	6	7	8
N	14	22	2	33	35	29	25	21
Graduate	3	22	2	17	17	2	3	3
Dropout	11	0	0	16	18	27	22	18
N 5%	8	22	2	21	17	5	8	8
Av_grade_S1	2.31	2.28	4.58	5.85	2.28	4.08	4.01	4.69
Av_grade_S2	6.00	2.04	2.50	2.61	2.09	4.65	3.13	3.45
F_ex_S1	0.06	0.01	0.00	0.04	0.03	0.14	0.49	0.13
F_ex_S2	0.00	0.02	0.00	0.07	0.02	0.51	0.09	0.14
En_ex_S1	0.24	0.08	0.00	0.17	0.24	0.15	0.14	0.87
En_ex_S2	0.20	0.08	0.00	0.04	0.20	0.37	0.61	0.36
P_ex_p_S1	0.64	0.99	0.00	0.00	0.31	0.34	0.34	0.05
P_ex_p_S2	0.00	0.08	1.00	0.04	0.22	0.07	0.19	0.05
P_ex_d_S2	0.00	0.01	0.83	0.71	0.11	0.04	0.22	0.33
P_ex_a_S1	2.39	2.99	0.20	0.00	0.55	0.10	0.03	0.02
P_ex_a_S2	0.00	0.66	2.75	0.02	0.41	0.05	0.11	0.06

Tab. 1: Characteristics of clusters of outliers detected by kNN algorithm with 10% assumption

Students-outliers from clusters 1-3 have a high number of passed exams in advance. This can be explained by the recognition of courses from a former study when students changed their educational program. Thus, we denoted these clusters as specific classes and investigated their impact on dropout prediction as a separate case.

5 RQ2: Do outliers affect dropout prediction models?

5.1 Methodology

Using the data from the three undergraduate programs, we built 15 cross-program models, which predict dropout, where independent variables are global features (Fig.1a) and the dependent variable is the feature with *Dropout/Graduate* labels. The models

based on different approaches to handle the outliers from the previous step (a. keep all outliers, b. remove all outliers from training data, c. remove only specific outlier clusters from training data) with different types of algorithms: decision tree (DT) and logistic regression (LR) to train interpretable models and random forest (RF) as an ensemble method that usually performs well. The analysis has been implemented using the Python scikit-learn library.

The dataset that keeps the outliers in the training data (approach a) served as the baseline to evaluate the outlier handling approaches. For approach b, regarding each assumption (5%, 10%) all outliers have been removed (datasets 2, 3), and for approach c, the outliers from specific clusters have been removed (datasets 4, 5). The detailed description of the analyzed datasets with the size of train and test sets is shown in Tab. 2.

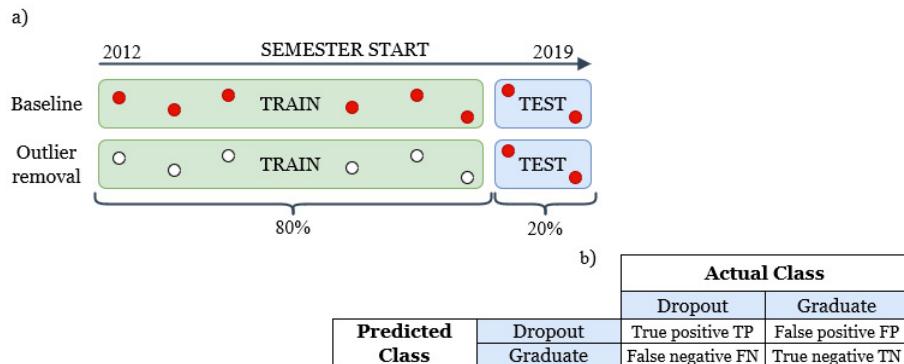


Fig. 2: Dropout prediction outlines: a) time-aware train/test splits b) the confusion matrix

As argued in [KMA19], time-aware nested cross-validation has been used to build the models to predict drop out: firstly, we sorted the students' data by their program start semester, then split without shuffling into training and test disjoint sets (80/20%), so that the test set contains the most recent students. Finally, we used the 10-fold cross-validation for the training set and took the best model in each case to predict dropout. The test set remained the same for each model and contained both inliers and outliers since outliers cannot be excluded from the prediction in a real-life scenario (362 students). The evaluation of the models included the following measures: recall, accuracy, precision, and area under the ROC curve. The detailed outlines of dropout prediction are depicted in Fig.2. In the context of our work, the most important metric is recall, since it returns a ratio of correctly predicted dropouts and the actual dropouts $TP/(TP+FN)$. The higher the recall score, the more dropouts are correctly predicted.

5.2 Results

The results of the prediction models are presented in Tab. 2. The rows show the results for each considered dataset. The columns correspond to the prediction metrics and set size. The numbers in bold show best values per metric among considered datasets for

each algorithm (DT, LR, RF). The colored numbers in italic correspond to the best values per metric across all models.

Models		Prediction metrics				Set size		
Alg	Dataset	Rec	Acc	Prec	Auc	Train	Test	
DT	1 All data	82.07%	83.15%	92.79%	83.83%	1447	362	
	2 w/o 5%		83.27%	81.77%	89.70%	80.82%	1384	362
	3 w/o 10%	80.88%	81.22%	91.03%	81.43%	1312	362	
	4 w/o 5% clusters 1-3	80.88%	81.49%	91.44%	81.88%	1432	362	
	5 w/o 10% clusters 1-3	80.88%	81.77%	91.86%	82.33%	1428	362	
LR	1 All data	79.28%	84.53%	98.03%	87.84%	1447	362	
	2 w/o 5%		80.48%	85.36%	98.06%	88.44%	1384	362
	3 w/o 10%	80.08%	84.81%	97.57%	87.79%	1312	362	
	4 w/o 5% clusters 1-3	79.28%	84.25%	97.55%	87.39%	1432	362	
	5 w/o 10% clusters 1-3	79.68%	84.53%	97.56%	87.59%	1428	362	
RF	1 All data	81.27%	85.08%	96.68%	87.48%	1447	362	
	2 w/o 5%	82.87%	83.70%	92.86%	84.23%	1384	362	
	3 w/o 10%		84.86%	86.46%	95.09%	87.48%	1312	362
	4 w/o 5% clusters 1-3	81.67%	84.53%	95.35%	86.33%	1432	362	
	5 w/o 10% clusters 1-3	82.87%	85.36%	95.41%	86.93%	1428	362	

Tab. 2: Results of dropout prediction models for datasets with outliers (All data), without outliers (w/o), and for algorithms (decision tree (DT), logistic regression (LR), random forest (RF))

Training a model without outliers improves the performance of the model on some metrics. Among 12 columns (each metric for each algorithm), for 8 the best result is somewhere else in the column, not in the first line with dataset 1 (all data). Comparing with the model trained with all data (dataset 1), recall has improved: 1.2% for decision tree and for logistic regression, and 3.59% for random forest. The improvement of the other metrics is more modest. The best results are never obtained by removing the outliers from specific classes (datasets 4, 5). These outliers are characterized by a large number of courses passed ahead of the plan in semester 1 or 2. Most probably, these courses have been passed before enrolling formally in the degree and recognized after formal enrollment. Concerning the other kinds of outliers, there is no dataset that works better for all metrics. In terms of algorithms, one winner is random forest which has the best recall (84.86%) and accuracy (86.46%), another winner is logistic regression that shows the best values for precision (98.06%) and AUC ROC (88.44%).

6 Conclusion and future work

The present study investigates the impact of outliers on dropout prediction models in higher education. Analyzing the students' performance after the two first semesters, we aimed to answer the following research questions: RQ1. What kinds of outliers can be detected? RQ2. Do outliers affect dropout prediction models?

To answer the first research question, the students-outliers have been detected by an unsupervised outlier detection algorithm (kNN) with two assumptions (5% or 10% of

students considered as outliers) and clustered (k-means algorithm). The first three clusters are characterized by high values of passed exams in the first two semesters. This can be explained by the recognition of courses that have been passed before enrolling formally (e.g. students who changed their educational program). Clusters 4 and 5 show other examples of rare performance: students who failed the first semester and then tried to reestablish as well as students who prefer courses from semesters going beyond the current curriculum. These clusters are interesting since they contain students who seem to have the same probability to graduate or to drop out. Some targeted intervention could be designed for such students. Students from clusters 6 to 8 have low marks, a high number of enrollments without attending exams, or a high number of failed exams. They tend to drop out.

To address the second research question, we build 15 cross-program dropout prediction models based on different approaches: a. keep all outliers (dataset 1), b. remove all outliers from training data (datasets 2, 3), c. remove only specific outlier clusters from training data (datasets 4, 5), using three different types of algorithms (decision tree, logistic regression, random forest). The models trained without all outliers have improved the performance with respect to some metrics, particularly for the recall, where the highest improvement is 3.59%. The best recall value is always achieved by models from approach b - removing all outliers from the training dataset.

Our results suggest that removing outliers could improve dropout prediction. Several limitations are noteworthy regarding this study, and they point out directions for future work. Firstly, the outlier detection has been performed using one unsupervised algorithm and with two threshold assumptions. The choice of the threshold is the most challenging task of outlier detection when the ground truth is unknown. Therefore, in future work, other thresholds as well as other outlier detection algorithms should be considered. Another limitation concerns the specific classes of outliers that have been removed from the training sets. The results of this study show that removing outliers from the clusters 1 to 3 does not improve dropout prediction. Thus, we intend to investigate how other kinds of outliers may impact dropout prediction. Regarding the dropout prediction, only three classification algorithms have been trained (decision tree, logistic regression, and random forest). However, in future work, more advanced algorithms such as AdaBoost or artificial neural networks can be applied. Moreover, we have investigated the impact of outliers on dropout prediction taking the scikit-learn standard values for the hyper-parameters of the models without feature preselection. The encouraging results of this work motivate us to improve the results through optimizing the hyper-parameters and selecting relevant features. Furthermore, similar analyses should be undertaken for data after the 1st, 3rd, and 4th semesters. These findings may enhance existing practices of dropout prediction in higher education. Furthermore, the outlier detection might be a part of the pipeline in the early warning system of detecting students with high risk of dropout.

Bibliography

- [AASF19] Aldowah, Hanan; Al-Samarraie, Hosam; Fauzy, Wan Mohamad: Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics*, 37:13–49, 4 2019.
- [AP02] Angiulli, Fabrizio; Pizzuti, Clara: Fast outlier detection in high dimensional spaces. In: European conference on principles of data mining and knowledge discovery. Springer, pp.15–27, 2002.
- [ATEH12] Abu Tair, Mohammed M; El-Halees, Alaa M: Mining educational data to improve students' performance: a case study. *International Journal of Information*, 2(2), 2012.
- [Au19] Aulck, Lovenoor; Nambi, Dev; Velagapudi, Nishant; Blumenstock, Joshua; West, Jevin: Mining University Registrar Records to Predict First-Year Undergraduate Attrition International Educational Data Mining Society, 2019.
- [Be19] Berens, Johannes; Schneider, Kerstin; Gortz, Simon; Oster, Simon; Burghoff, Julian et al.: Early Detection of Students at Risk-Predicting Student Dropouts Using Administrative Student Data from German Universities and Machine Learning Methods. *JEDM| Journal of Educational Data Mining*, 11(3):1–41, 2019.
- [BGS16] Bansal, Rashi; Gaur, Nishant; Singh, Shailendra Narayan: Outlier Detection: Applications and techniques in Data Mining. 2016 6th International Conference - Cloud System and Big Data Engineering (Confluence), pp. 373–377, 2016.
- [BRGS19] Baneres, David; Rodríguez-Gonzalez, M Elena; Serra, Montse: An early feedback prediction system for learners at-risk within a first-year higher education course. *IEEE Transactions on Learning Technologies*, 12(2):249–263, 2019.
- [Ca19] Carneiro, Rubens E; Drapal, Patrícia; Fagundes, Roberta AA; Maciel, Alexandre MA; Rodrigues, Rodrigo L: Anomaly Detection on Student Assessment in E-Learning Environments. In: 2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT). volume 2161. IEEE, pp. 168–169, 2019.
- [GU16] Goldstein, Markus; Uchida, Seiichi: A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data. *PloS one*, 11(4):e0152173, 2016.
- [HA14] Hodge, Victoria J.; Austin, Jim: A Survey of Outlier Detection Methodologies. *Artificial Intelligence Review*, 22(2):85–126, Oct 2004.
- [KMA19] Krauss, Christopher; Merceron, Agathe; Arbanowski, Stefan: The timeliness deviation: A novel approach to evaluate educational recommender systems for closed-courses. In: Proceedings of the 9th International Conference on Learning Analytics & Knowledge. pp.195–204, 2019.
- [NV19] Nyitrai, Tamás; Virág, Miklós: The effects of handling outliers on the performance of bankruptcy prediction models. *Socio-Economic Planning Sciences*, 67:34–42, 2019.
- [OM18] Ochoa, Xavier; Merceron, Agathe: Quantitative and Qualitative Analysis of the Learning Analytics and Knowledge Conference 2018. *Journal of Learning Analytics*, 5(3):154–166,2018.

Teacher vs. Algorithm: Learners' Fairness Perception of Learning Analytics Algorithms

Linda Mai, Alina Köchling, Lynn Schmodde and Marius Wehner¹

Abstract: The use of Learning Management Systems (LMS) is increasing at schools and universities and, hence, a large amount of data is available for the assessment, recommendation, and final grading of learners. Learning analytics (LA) algorithms can either support or even substitute teachers in the grading process. However, knowledge is limited of how learners react to the algorithmic examination of personal learning data. We examine the effects of the grading entity (i.e., teacher, teacher supported by LA, and LA) on learners' reactions. Using a scenario-based between-subject design of a common grading situation at the end of the school year (N=118), we investigate whether a decrease in teachers' involvement in the grading decision diminishes learners' perceived satisfaction regarding the grading process mediated by fairness. Surprisingly, our preliminary results show that with a decrease of the teacher's involvement the perceived satisfaction is lower, but the perceived fairness is higher. Finally, fairness mediated the influence of the grading entity on satisfaction. Overall, this study offers important theoretical and practical implications when designing and applying an LMS.

Keywords: Learning Management System, Learning Analytics, Scenario-based Study, Fairness, Satisfaction

1 Introduction

Since Learning Management Systems (LMS) can promote personalized learning processes and support learners [LP15], their use is increasing in schools and universities. LMSs provide a wide range of possibilities to support individualized learning like adaptive learning [MW16], providing feedback [Pa19], predicting student dropouts [JGC10]; [De20] and support collaboration [Ma21], as well as differentiated teaching [SF17]. Each action like reading files or sending messages on the platform leaves traces of data, which can be analysed for education-related decisions [LP15], [GDS15]. LMSs can collect this huge amount of data and generate a range of highly accurate analyses. Therefore, LMSs offer a range of possible applications like early-warning systems [Ak19], prediction systems based on machine learning methods [LOM21], systems to improve education systems [RV10], and models to improve student learning and success [JRS17]. These techniques can support teachers and instructors in their daily work. While assessments and grading are common time-intensive tasks, algorithms can make a fully automated final grading possible, which is already employed in some countries, for instance in the UK [Sa20].

¹ Junior Professorship for Business Administration, esp. Corporate Management, Faculty of Business Administration and Economics, Heinrich-Heine-University Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf

Graduation of school and university attendance requires assessments and judgments by teachers. These human judgments can be biased. At the same time, an accurate and unbiased assessment is crucial for a fair grading process [UW21]. In this context, learning analytics (LA) can support the processes of judgments and final grading.

Concurrently, different stakeholders have varying perspectives on LA and objectives for its usage [WHH08]; [ECN18]. Learners (i.e., pupils and students), who are important stakeholders alongside teachers, parents, and institutions [GD12], aim for a personalized learning experience and recommendations, while educators are interested in objective feedback and classification [RV10]. Students can improve their learning by peer feedback [MWE21] and perceive LA-based feedback as positive under certain conditions [Li20]. While LA-based feedback can support learners, teacher-based feedback still plays an important role in schools [ISS21]. Though the different objectives of the stakeholders do not necessarily contradict each other, both groups must be considered when developing LMSs. Teachers, for instance, benefit from the support by LA and are hence able to respond to the needs of their students in a personalized and adaptive way. Due to the development of information and communications technologies, teachers and instructors need to integrate technology with pedagogy and subject content in their lessons and lectures [TL19]. However, taking a learner-centred approach, research has fallen behind the practice, and there is little research on the acceptance of LA applications and learners' reactions [Fa20]. This requires the clarification of the question to what extent the recommendation of learning analytics should be incorporated into the final grading decision.

To fill this void, we analyze the possibilities of using an LMS from the learner's perspective and examine in particular how learners from schools and universities react to different involvements of automated decisions. Using an experimental between-subject scenario study ($N=118$), we investigate whether a decrease in teacher involvement triggers adverse learners' reactions. In more detail, we investigate how the deciding entity of (a) an exclusive teacher decision, (b) a teacher decision supported by a LA, (c) and an exclusive LA decision influences a learner's reaction to the grading process. We argue that decreased teacher influence in the grading decision is associated with adverse learner reactions, specifically with lower perceived fairness of the process and lower satisfaction with the grading process. This can have important consequences for institutions since maintaining a good reputation is essential for schools and universities.

With this study, we contribute to the literature on LA in several ways. First and foremost, we shift our focus from an institutional perspective to the perspective of the learners. Additionally, we introduce and provide a deeper understanding of the role of the deciding entity. Addressing the learners' reactions to an LA-based grading process, this study provides important theoretical as well as practical implications that can help in the development and use of LA.

2 Theory

LMSs are developed to support learners and instructors and the learning processes [OCN16]. Analyzing the learners' traces of data left while using the LMS can improve learning success and the grading processes. There is evidence, that personalized feedback based on the used material provides important information about the students' progress [Li20]. At the same time, we know little about how learners perceive the usage of LMSs regarding the final grading process. Especially when it comes to grading, fairness and satisfaction are very sensitive issues, because grading often determines learners' opportunities for further education and career opportunities. When there are discrepancies in students' expectations and perceptions with regard to the grading process, it can lead to dissatisfaction [TLP12]. Therefore, it is necessary to understand how learners in schools and universities perceive algorithmic grading.

2.1 Fairness and Satisfaction in educational contexts

We draw on justice theory to answer our research questions [CBW07]; [CFK15]; [Le80]. According to justice theory, learners examine the justice of the outcome (distributive justice), the adequacy of the justice process (procedural justice), and the appropriateness of the treatment learners receive from teachers (interactional justice) [CBW07]; [CFK15]; [Le80]. In this context, justice is more descriptive and subjective rather than objective reality, and therefore, a personal evaluation of the teachers' behavior [CBW07].

Fairness

Traditional grading processes are teacher-centred. This means teachers and instructors are responsible for the learners' grades based on exams and other forms of examination. There is evidence, that teachers' assessments are sometimes not objective and there are demands for relying on observable and valid indicators to ensure correct and fair judgments [UW21]. The relationship between students' effort and performance can explain the perceived fairness, while especially a combination of high effort and low performance can be found to be least likely to result in a fair grade [TLP12].

Decisions made with help of algorithms have the advantage, that they lead to the same result for a data set and are free of random human fluctuations [DFM89]. While algorithmic decisions can be more accurate, people react differently to algorithmic advice, depending on expert or lay status [LMM19].

While LA is generally expected to increase the correctness and fairness of judgments, it is crucial that the underlying algorithms are accurate and fair themselves. In general, decision processes with help of algorithms are more consistent, more objective, more accurate, and less biased than a human decision [LS21]. Therefore, the consistency could also be higher when using LA as decision support or when an LA system is the decision-maker since the system evaluates each student with the same attention and with the same

time investment. While the systems are often not free of bias (e.g., the misevaluations of England's exam regulator [Pa20]), it is not clear, how learners react to automated final grading.

However, even if the grading process is fairer from a statistical point of view, it is important how the learners perceive the fairness of the grading process. The use of LMSs is often associated with a lack of knowledge and transparency. Learners do not know how the algorithmic evaluation works and they may perceive the grading process as unfair since the learners are used to a grading process operated by a human. For this reason, they may think that it is unethical to surrender the process to an algorithmic LA system.

Satisfaction

Satisfaction with the grading process reflects a person's perception of being valued and esteemed during the grading process [WGK15]. The satisfaction with the process is influenced by the interpersonal treatment. Learners may be not satisfied with the process since the interpersonal interaction is missing when the grading process is operated with help of an LA system or the LA system operates the process alone. Moreover, learners may not deem the treatment of the teachers to be appropriate when leaving the decision to an LMS and, in turn, they may not feel valued by their teachers, which leads to a decrease in satisfaction with the grading process.

Non-human grading could lead to lower perceived fairness because of the lack of interpersonal contact. We expect that fairness and satisfaction are interrelated, meaning that lower perceived fairness can lead to lower satisfaction with the grading process. This is due to the assumption that learners are satisfied with a grading process when they perceive it as fair. Thus, we are interested in the following research question:

Will exposure to teacher/instructor evaluation lead to a more positive evaluation of satisfaction with the grading process mediated by fairness compared to LA support and LA?

3 Method and Study

To understand how algorithmic grading can influence the learners' perceived fairness and satisfaction, we developed three different scenarios, asking participants to put themselves in a specific grading situation. The use of written scenarios in an experimental design is an internally valid method if the participants are confronted with realistic situations [MD99]. We chose a grading situation because it reflects a common decision in school and university life and is therefore realistic.

3.1 Pilot Study and Scenarios

First, we conducted a pilot study to test our research questions and to validate our written

scenarios. The development of a pilot study is in line with current recommendations and best practice approaches when employing experimental scenarios [Ra15]. Based on an extensive literature search, we manipulated the decision-maker in the grading process (teacher, teacher supported by LA, LA) for the grading process. To ensure that participants can identify the different types of decision making, we used think-aloud protocols before the empirical examination with 10 participants to test the scenarios [Dr03].

Participants were asked to read the text thoroughly and imagine themselves in the described situations. They were supposed to read about a grading process at the end of the school year/end of the semester. Participants were randomly assigned to one of our three between-subject scenarios and completed a questionnaire that included all measures (see 3.2) [AB14], [CGK12].

In scenario 1, the teachers determined the grade, evaluating data manually. In scenario 2, the data was evaluated algorithmically, and the recommendation of LA served as decision support for the teacher. In scenario 3, the LMS system was the decision-maker without human influence and generated grades automatically (see Appendix for all descriptions). The scenarios covered the spectrum from traditional teacher grading (1) to an LA-oriented approach (2), and finally to a fully data-driven grading (3).

For our main study, we recruited 120 pupils and students from Germany with 40 participants per scenario using an online panel of an ISO 20252:2019 certified online sample provider. After excluding participants due to the age of over 35 years, because these were outliers. The final sample consisted of 118 participants. 60 of them were pupils, 58 were students. On average, the students were 23.3 years old, and the pupils were 18.1 years old. 76 per cent were female, 27 per cent were male, and 1 per cent said they were diverse.

3.2 Measures

To measure perceived fairness and satisfaction, we adopted scales from existing measures to ensure the reliability and validity of our measures. We measured both scales with items ranging from 1 (strongly disagree) to 7 (strongly agree).

Fairness with the grading process. We measured fairness with three items from Bauer et al. [Ba01]. The items were “Overall, the grading is fair”, “I think that the grading system described is a fair way to grade students”, and “Grades were assigned equally for all students”. Cronbach’s alpha of the fairness scale was .90.

Satisfaction with the grading process. Overall satisfaction with the grading process was measured with two items adapted from Wehner et al. [WGK15]. The items were measured with “All in all, I am satisfied with the grading process”, “The grading system corresponds to my ideas of an ideal way of grading”. Cronbach’s alpha of the satisfaction scale was .87.

3.3 Quality check

To understand how participants perceived the scenarios and to ensure that the scenarios were sufficiently realistic, we asked participants to rate the realism of the scenarios on a 7-point Likert scale (1: unrealistic to 7: realistic), as well as the valence, that is, how well they were able to put themselves into the situation (1: very bad to 7: very good). (Tab. 1). As reflected by the high means, our participants could put themselves in the scenario well and perceived all scenarios as realistic (Tab. 1). These results show that the different scenarios can be used for our analysis.

	Scenario 1 (n=37)	Scenario 2 (n=39)	Scenario 3 (n=42)
Valence	M = 5.41 SD = 1.54	M = 5.10 SD = 1.17	M = 5.31 SD = 1.33
Realism of the scenario	M = 4.8 SD = 1.34	M = 4.49 SD = 1.48	M = 4.43 SD = 1.47

Tab. 1: Implementation checks (7-point Likert scale)

4 Results

In the following, we present the preliminary results of our study. Our results reveal how participants reacted to the different scenarios in terms of their assessment of fairness and satisfaction.

4.1 Reactions

Our results show that the assessment context has a strong influence on the acceptance of LA. The fully automated grading as represented in scenario 3 was perceived as fairer on average than scenarios 1 and 2 (Tab. 2).

	Scenario 1 (n=37)	Scenario 2 (n=39)	Scenario 3 (n=42)	F value Pr (>F)
Fairness	M = 4.38 SD = 1.59	M = 4.37 SD = 1.31	M = 5.03 SD = 1.28	4.589 0.0343*
Satisfaction	M = 4.31 SD = 1.31	M = 3.67 SD = 1.33	M = 3.88 SD = 1.45	1.757 0.188

Tab. 2: Descriptive Statistics and ANOVA Results

A differentiation by school and university also shows these differences. Regardless of the institution, automated assessment is perceived as fairer. We observed differences in perceived fairness between scenario 1, scenario 2, and scenario 3. These differences show

that LA in the grading process increases the perceived fairness for most of the learners. We observed differences in perceived satisfaction between scenario 1 and scenario 2. To use the scenarios as variables, we have coded them as dummy variables. Running a regression to predict satisfaction with the grading process by LA support and LA, no significant effect could be found. Only when fairness is considered with LA support and LA does a significant influence of all three variables become apparent. The differentiation between student and learner and gender had no impact (Tab. 3). A regression to predict perceived fairness using LA support and LA shows a significant relationship for LA (Tab. 4). Differentiations as before also did not influence the result.

We found that our outcome variables fairness and satisfaction were indeed interrelated. Therefore, perceived fairness plays an important role when it comes to the implementation of new systems in schools and universities. Fairness can be the basis for the acceptance of new systems. Using the R package *mediation* [Ti14], we found that perceived fairness has a significant impact on satisfaction but is not relevant as a mediator. Instead, the LA support (scenario 2) has only a direct impact on satisfaction (Fig. 1). Perceived fairness acts as a mediating mechanism in scenario 3, where an algorithm is the deciding entity (Fig. 2). The reference category was scenario 1. For scenario 3, we can only identify an indirect effect, which means that we have to focus on the theoretical framework [ZLC10].

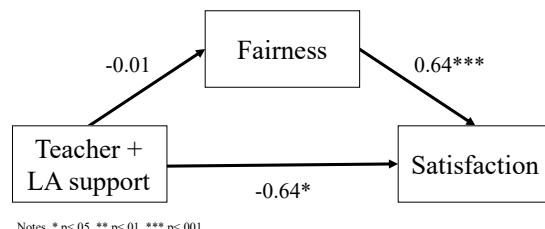


Fig. 1: Mediating Mechanisms Scenario 2

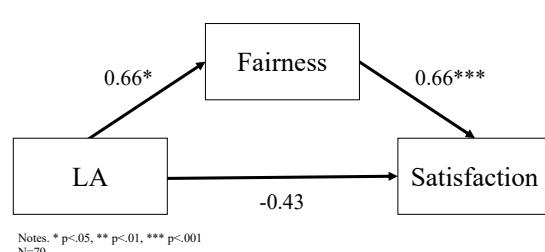


Fig. 2: Mediating Mechanisms Scenario 3

4.2 Summary

Our results reveal that learners perceived their grading process as fairer when an LA algorithm was involved, but only when there was no additional teacher influence in the grading process. This finding is surprising, since studies on the usage of algorithms in other contexts (e.g., HR recruitment and development), indicated that individuals prefer a human rather than an automated assessment, and that individuals generally perceive algorithmic processes as unfair [LKP19]; [Le18].

Diametrically opposite to these findings, our study shows that the implementation of algorithms in grading systems is not harmful and even can lead to a more positive evaluation of the grading process by learners in terms of perceived fairness when the process is conducted by purely LA (see Fig. 2). Looking closer to our constructs and a possible framework [ZLC10], a possible explanation for our results is that learners might have experienced unfair treatment of their teachers before and therefore perceive an algorithmic grading system as more objective and therefore fairer.

At the same time, the satisfaction with the combination of a teacher and LA-based grading process is lower than a teacher-based grading process (see Fig. 1). Learners probably perceive that they lack interpersonal contact with teachers when the process is supported by an LA system and thus may not feel esteemed by the teachers and instructors. Only when the use of algorithms is individually perceived as fairer, it leads to more satisfaction with the process.

The grading process is a very sensitive issue since it implies opportunities for further education and career chances. As teachers' assessments are often not objective and involve humans bias and personal prejudices [UW21], the support of LA can lead to a more consistent, fairer, and less biased grading process since each learner is rated by the algorithmic system applying the same criteria, the same amount of time, and the same attention. Thus, the use of technologies in LMSs can increase the accuracy of at-risk prediction and grading. This not only applies from an objective perspective but is also reflected in our findings from the learner's perspective, as they appreciated the automated grading process in terms of fairness, rather than the traditional grading by humans, namely teachers and instructors.

Still, the negative effects on satisfaction with the process when algorithms are involved in the grading process point at the adverse effects of the implementation of LA. Pupils and students hence apparently still prefer a teacher as a final deciding entity rather than the pure evaluation by an algorithm.

Our study contributes to the literature in several ways. First and foremost, we shift our focus from an institutional perspective on the perspective of the learners and shed light on the learner's perspective, which is extremely important when implementing LA algorithms. Second, we provide a deeper understanding of the role of the deciding entity and show that an LA-based decision is considered fairer in school and university contexts, but also leads to lower satisfaction when not mediated by fairness.

4.3 Implications

Grading in school or universities means chances for social and financial mobility. Therefore, all stakeholders should be involved in the implementation process. Especially the learner's perspective needs to be considered when implementing LA systems in school. It is important that they are considered as important stakeholders and accept newly implemented algorithmic-based grading tools.

Based on our results, LA as decision support can increase the learners' perceived fairness. This can improve the school's or university's reputation in the way, that human biases in assessments are minimized and final grading is perceived more objectively. However, also algorithmic systems are not free of bias, as extant research showed. This tension between both forms of biases (i.e., human and algorithmic) is not yet fully understood, but as our findings point at the positive evaluation of algorithmic decision-making, schools and universities should not refrain from implementing LMSs as a support for their grading processes. Still, as long as algorithmic bias cannot be ruled out, the final deciding and responsible entity should be a human in the very sensitive context of grading processes. While there is acceptance of LA-based grading, institutions must be aware of additional negative effects. If the teacher is no longer involved in the grading process, then important information about the learners is also missing here.

Schools and universities should also be aware of possible discrimination by using algorithmic decision-making. Hence, transparency about the used data and the procedures is important. Our results, which point out the acceptance of LA, is a positive sign for implementing new designs of LMS, but still, all stakeholders should be critical of algorithms' outputs.

4.4 Limitations and future research

First, due to data protection regulations, we focused on learners who are 16 years or older. In schools, also younger learners use LMSs and it is not clear, how they react to automated feedback and grading. Moreover, some LMSs also provide access to the parents. There is also no research on parents as stakeholders. Thus, it is still unclear, which influence the reports have on the parents and their parental support. Therefore, research on both parents, other persons entitled to custody, and younger learners is needed.

Second, there is no systematic education about data literacy or computational thinking [SSA17] in schools and universities, yet. If a large amount of personal data is used, then it might be useful to also develop data literacy [Fr19]. Computational thinking might increase the critical view on automated decision making because learners can understand how systems work [Wi08], [SSA17]. Computer science lessons have a positive impact on computational thinking [Gu20]. This could be useful to create a deeper understanding of the way LMSs work.

Furthermore, our study was conducted in Germany. While digitization in schools and universities is increasing, other countries are more advanced in digitization [Ga20]. Improving pedagogical and technological knowledge [TL19] is an important task for educators and, additionally, research on how learners understand LA is needed. A solely LA-based grading could reduce the instructional quality. The information from the LMS is then not used to improve learning progress and learning and grading are no longer perceived as belonging together. Using LA data only for grading ignores many positive and supportive ways of learning.

Third, employing a scenario means that participants had to imagine a hypothetical situation. Participants may not have been aware of all the consequences as in a real grading setting. Furthermore, none of the learners knew his or her future grading results and the study had no impact on their real lives. Moreover, as attitudes and experiences can change over time, we highly recommend conducting panel studies in schools and universities that implemented different LMS designs.

Future research should focus on grading processes in general because our results show the need for an objective evaluation of performance. Yet, it is not fully understood if learners are open-minded toward new technologies or if they are dissatisfied with their teachers. Additionally, the influence of other unknown factors can be important to understand the relationship between learners and LA. Maybe the lack of interpersonal contact can influence the perception of automated grading and could be the mediator, we have not included in our models.

As discussed before, teachers and instructors are vulnerable to various forms of bias [UW21]. There is evidence, that algorithmic evaluation leads to negative reactions in human resources contexts. Therefore, further research should focus on experiences of discrimination. A link could be suspected between unfair treatment by teachers and lecturers and LA preference.

5 Conclusion

Through a unique scenario-based study, we showed how learners react to automated and supported grading, which offers important implications for theory and practice. Learners both in schools and at universities should be taken into consideration as equal stakeholders. The benefit of an LMS for the whole institution depends largely on the learners' acceptance, but the relationship between LA support and LA and satisfaction is not fully understood yet. We hope to encourage other researchers with this study to conduct further research in this field.

6 Acknowledgment

We gratefully acknowledge financial support from the Federal Ministry of Education and Research 01JD1812A/B.

7 References

- [AB14] Aguinis, H.; Bradley, K. J.: Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational research methods* 17/4, pp. 351-371, 2014.
- [Ak19] Akçapınar, G.; Hasnine, M. N.; Majumdar, R.; Flanagan, B.; Ogata, H.: Developing an early-warning system for spotting at-risk students by using eBook interaction logs. *Smart Learning Environments* 6/4, pp. 1-15, 2019.
- [Ba01] Bauer, T. N.; Truxillo, D. M.; Sanchez, R. J.; Craig, J. M.; Ferrara, P.; Campion, M. A.: Applicant reactions to selection: Development of the selection procedural justice scale (SPJS). *Personnel Psychology* 2/54, pp. 387-419, 2001.
- [CBW07] Cropanzano, R.; Bowen, D. E.; Gilliland, S. W.: The management of organizational justice. *Academy of Management Perspectives* 21/4. pp. 34-48, 2007.
- [CFK15] Cropanzano, R.; Fortin, M.; Kirk, J. F.: How do we know when we are treated fairly? Justice rules and fairness judgments. *Research in personnel and human resources management*: Emerald Group Publishing Limited 38, pp. 279-350, 2015.
- [CGK12] Charness, G.; Gneezy, U.; Kuhn, M. A.: Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization* 81/1, pp. 1-8, 2012.
- [De20] Del Bonifro, F.; Gabbielli, M.; Lisanti, G.; Zingaro, S. P.: Student Dropout Prediction. In: I. I. Bittencourt, Cukurova, M.; Muldner, K.; Luckin, R.; Millán E. (Eds.): *Artificial Intelligence in Education*, 21st International Conference, AIED 2020, LNNAI 12163, pp. 129–140, 2020.
- [DFM89] Dawes, R. M.; Faust, D.; Meehl, P. E.: Clinical Versus Actuarial Judgment. *Science* 243, pp. 1668-1674, 1989.
- [Dr03] Drennan, J.: Cognitive interviewing: verbal data in the design and pretesting of questionnaires. *Journal of advanced nursing* 42/1, pp. 57-63, 2003.
- [ECN18] Ekuase-Anwansedo, A.; Craig, S. F.; Noguera, J.: How to Survive a Learning Management System (LMS) Implementation? A Stakeholder Analysis Approach. In: *Proceedings of the 2018 ACM SIGUCCS Annual Conference*, pp. 165-168, 2018.
- [Fa20] Falcão, T. P.; Mello, R. F.; Rodrigues, R. L.; Diniz, J. R. B.; Tsai, Y. S.; Gašević, D.: Perceptions and expectations about learning analytics from a brazilian higher education institution. In: *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, pp. 240-249, 2020.
- [Fr19] Fraillon, J.; Ainley, J.; Schulz, W.; Duckworth, D.; Friedman, T.: IEA international

- computer and information literacy study 2018 assessment framework. Springer Nature, 2019.
- [Ga20] Gapsalamov, A.; Bochkareva, T.; Vasilev, V.; Akhmetshin, E.; Anisimova, T.: Comparative analysis of education quality and the level of competitiveness of leader countries under digitalization conditions. *Journal of Social Studies Education Research* 11/2, pp. 133-150, 2020.
- [GD12] Greller, W.; Drachsler, H.: Translating Learning into Numbers: A Generic Framework for Learning Analytics. *Educational Technology & Society* 15/3, pp. 42–57, 2012.
- [GDS15] Gašević, D.; Dawson, S.; Siemens, G.: Let's not forget: Learning analytics are about learning. *TECHTRENDS TECH TRENDS* 59, pp. 64–71, 2015.
- [Gu20] Guggemos, J.: On the predictors of computational thinking and its growth at the high-school level. *Computers & Education* 161, 2020.
- [ISS21] Ifenthaler, D.; Schumacher, C.; Sahin, M.: System-based or Teacher-based Learning Analytics Feedback – What Works Best? International Conference on Advanced Learning Technologies (ICALT), 2021.
- [JGĆ10] Jadrić, M.; Garača, Ž.; Ćukušić, M.: Student Dropout Analysis with Application of Data Mining Methods. *Management: Journal of Contemporary Management Issues* 15/1, pp. 31-46, 2010.
- [JRS17] Javidi, G.; Rajabion, L.; Sheybani, E.: Educational Data Mining and Learning Analytics: Overview of Benefits and Challenges. From 2017 International Conference on Computational Science and Computational Intelligence (CSCI), pp. 1102-1107, 2017.
- [Le18] Lee, M. K.: Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society* 5/1, pp. 1-16, 2018.
- [Le80] Leventhal, G. S.: What should be done with equity theory? In *Social exchange*, pp. 27-55, Springer, 1980.
- [Li20] Lim, L.-A.; Dawson, S.; Gašević, D.; Joksimović, S.; Pardo, A.; Fudge, A.; Gentili, S.: Students' perceptions of, and emotional responses to, personalised learning analytics-based feedback: an exploratory study of four courses. *Assessment & Evaluation in Higher Education*, 2020.
- [LKP19] Langer, M.; König, C. J.; Papathanasiou, M.: Highly automated job interviews: Acceptance under the influence of stakes. *International Journal of Selection and Assessment* 27/3, pp. 217-234, 2019.
- [LMM19] Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes* 151, pp. 90-103, 2019.
- [LOM21] Lwande, C.; Oboko, R.; Muchemi, L.: Learner behavior prediction in a learning management system. *Education and Information Technologies* 26, pp. 2743–2766, 2021.
- [LP15] Liñán, L. C.; Pérez, Á. A. J.: Educational Data Mining and Learning Analytics: differences, similarities, and time evolution. *International Journal of Educational Technology in Higher Education* 12/3, pp. 98–112, 2015.

-
- [LS21] Leyer, M.; Schneider, S.: Decision augmentation and automation with artificial intelligence: Threat or opportunity for managers? *Business Horizons*, 2021.
- [Ma21] Martinez-Maldonado, R.; Gašević, D.; Echeverria, V.; Fernandez Nieto, G.; Swiecki, Z.; Buckingham Shum, S.: What Do You Mean by Collaboration Analytics? A Conceptual Model. *Journal of Learning Analytics* 8/1, pp. 126-153, 2021.
- [MD99] Maute, M. F.; Dubés, L.: Patterns of emotional responses and behavioural consequences of dissatisfaction. *Applied Psychology* 48/3, pp. 349-366, 1999.
- [MW16] Martin, F.; Whitmer, J. C.: Applying Learning Analytics to Investigate Timed Release in Online Learning. *Technology, Knowledge and Learning* 21, pp. 59–74, 2016.
- [MWE21] Misiejuk, K.; Wasson, B.; Egelandsdal, K.: Using learning analytics to understand student perceptions of peer feedback. *Computers in Human Behavior* 117, 2021.
- [OCN16] Oliveira, P. C. de; Cunha, C. J. C. D. A.; Nakayama, M. K.: Learning Management Systems (LMS) and E-Learning Management: an integrative review and research agenda. *JISTEM-Journal of Information Systems and Technology Management* 13/2, pp. 157–180, 2016.
- [Pa19] Pardo, A.; Jovanovic, J.; Dawson, S.; Gašević, D.; Mirriahi, N.: Using learning analytics to scale the provision of personalised feedback. *British Journal of Educational Technology* 50/1, pp. 128-138, 2019.
- [Pa20] Paulden, T.: A cutting re-mark. *Significance* 17, pp. 4-5, 2020.
- [Ra15] Raaijmakers, A. G.; Vermeulen, P. A.; Meeus, M. T.; Zietsma, C.: I need time! Exploring pathways to compliance under institutional complexity. *Academy of Management Journal* 58/1, pp. 85-110, 2015.
- [RV10] Romero, C.; Ventura, S.: Educational Data Mining: A Review of the State of the Art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 40/6, pp. 601-618, 2010.
- [Sa20] Satariano, A: British Grading Debacle Shows Pitfalls of Automating Government. *New York Times*, <https://www.nytimes.com/2020/08/20/world/europe/uk-england-grading-algorithm.html>, Retrieved on 22.06.2021.
- [SF17] Silva C.; Fonseca J.: Educational Data Mining: A Literature Review. In: *Europe and MENA Cooperation Advances in Information and Communication Technologies* (pp.87-94), Springer, 2017.
- [SSA17] Shute, V. J.; Sun, C.; Asbell-Clarke, J.: Demystifying computational thinking. *Educational Research Review* 22, pp. 142–158, 2017.
- [Ti14] Tingley, D.; Yamamoto, T.; Hirose, H.; Keele, L.; Imai, K.: mediation: R Package for Causal Mediation Analysis. *Journal of Statistical Software* 59/5, 2014.
- [TL19] Taimalu, M.; Luik, P.: The impact of beliefs and knowledge on the integration of technology among teacher educators: A path analysis. *Teaching and Teacher Education* 79, pp. 101-110, 2019.
- [TLP12] Tippin, G. K.; Lafreniere, K. D.; Page, S.: Student perception of academic grading: Personality, academic orientation, and effort. *Active Learning in Higher Education* 13/1,

pp. 51-61, 2012.

- [UW21] Urhahne, D.; Wijnia, L.: A review on the accuracy of teacher judgments. *Educational Research Review* 32, 2021.
- [WHH08] Wagner, N.; Hassanein, K.; Head, M.: Who is responsible for E-Learning Success in Higher Education? A Stakeholders' Analysis. *Educational Technology & Society* 11/3, pp. 26-36, 2008.
- [Wi08] Wing, J. M.: Computational thinking and thinking about computing. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences* 366, pp. 3717–3725, 2008.
- [WGK15] Wehner, M.; Giardini, A.; Kabst, R.: Recruitment Process Outsourcing and Applicant Reactions: When Does Image Make a Difference? *Human Resource Management*, 54/6, pp. 851-875, 2015.
- [ZLC10] Zhao, X.; Lynch Jr., J. G.; Chen, Q.: Reconsidering Baron and Kenny: Myths and Truths about Mediation Analysis. *Journal of Consumer Research* 37, pp. 197-206, 2010.

8 Appendix

Description of Scenario 1

Your teachers give you a grade for each subject at the end of the school year. The final grade is made up of your participation, classwork and other achievements (presentations, etc.). Your teachers also look at the stored data of the learning platform manually. Other criteria are not important for your grades. After the teachers' conference, you will receive your transcript and can apply.

Description of Scenario 2

Your teachers give you a grade for each subject at the end of the school year. The final grade is made up of your participation, classwork and other achievements (presentations, etc.). An algorithm evaluates the stored data from the learning platform. Your teachers receive a grade suggestion for each subject on this basis. You can accept this suggestion or not. Other criteria are not important for your grades. After the teachers' conference, you will receive your transcript and can apply.

Description of Scenario 3

An algorithm evaluates the data stored on the learning platform and automatically assesses your performance for each subject. The final grade is made up of your participation, class tests and other achievements (presentations, etc.). Your teachers use this grade for your transcript. Other criteria are not important for your grades. You will receive an automatic email at the end of the school year and can view your grades on the learning platform. For your application, you will also receive a printed transcript with the grades from the learning platform.

Table 3: Regression results using satisfaction as the criterion

Predictor	<i>b</i>	<i>b</i>		<i>beta</i>		<i>s²</i>	<i>r</i>	Fit	Difference
		95% CI	[LL, UL]	95% CI	[LL, UL]				
<hr/>									
(Intercept)	3.35**		[2.24, 4.45]						
learner	0.29		[-0.22, 0.80]	0.10	[-0.08, 0.29]	.01	[-0.5, 0.6]	.10	
gender	0.14		[-0.43, 0.70]	0.04	[-0.14, 0.23]	.00	[-0.1, 0.2]	.04	<i>R</i> ² = .012 95% CI[0.0, 0.7]
<hr/>									
(Intercept)	3.79**		[2.61, 4.98]						
learner	0.25		[-0.25, 0.76]	0.09	[-0.09, 0.28]	.01	[-0.2, 0.4]	.10	
gender	0.10		[-0.46, 0.66]	0.03	[-0.15, 0.22]	.00	[-0.1, 0.1]	.04	
LA support	-0.61		[1.24, 0.01]	-0.21	[-0.42, 0.00]	.03	[-0.3, 0.6]	-.14	
LA	-0.43		[1.05, 0.19]	-0.15	[-0.36, 0.07]	.02	[-0.3, 0.6]	-.03	$\Delta R^2 = .033$
<hr/>									
(Intercept)	1.10*		[0.05, 2.14]						
learner	0.00		[-0.38, 0.39]	0.00	[-0.14, 0.14]	.00	[-0.0, 0.0]	.10	
gender	0.23		[-0.19, 0.65]	0.08	[-0.06, 0.21]	.01	[-0.1, 0.3]	.04	
LA support	-0.63**		[-1.09, -0.16]	-0.21	[-0.37, -0.05]	.03	[-0.1, 0.8]	-.14	
LA	-0.84**		[1.31, -0.38]	-0.29	[-0.46, -0.13]	.06	[-0.0, .12]	-.03	
<hr/>									
fairness	0.60**		[0.23, 0.80]	0.68	[0.34, 0.82]	.43	[30, 56]	.63**	<i>R</i> ² = .475** 95% CI[.32, .56]
<hr/>									

Note. A significant b-weight indicates the beta-weight and semi-partial correlation are also significant. b represents unstandardized regression weights, beta indicates the standardized regression weights, s² represents the semi-partial correlation squared, r represents the zero-order correlation, LL and UL indicate the lower and upper limits of a confidence interval, respectively. * indicates p < .05. ** indicates p < .01.

The variables LA support and LA are dummy variables.

Table 4: Regression results using fairness as the criterion

Predictor	<i>b</i>	<i>b</i>	<i>beta</i>	<i>beta</i>	<i>sr</i> ²	<i>sr</i> ²	<i>r</i>	Fit	Difference
		95% CI [LL, UL]		95% CI [LL, UL]		95% CI [LL, UL]			
(Intercept)	4.30**	[3.18, 5.43]							
learner	0.42	[-0.10, 0.94]	0.15	[-0.03, 0.33]	.02	[-.03, .07]	.16		
gender	-0.25	[-0.83, 0.32]	-0.08	[-0.26, 0.10]	.01	[-.02, .04]	-.09		
								<i>R</i> ² = .031	
									95% CI[.00, .10]
(Intercept)	4.07**	[2.86, 5.27]							
learner	0.38	[-0.14, 0.89]	0.13	[-0.05, 0.31]	.02	[-.03, .06]	.16		
gender	-0.20	[-0.77, 0.37]	-0.06	[-0.24, 0.12]	.00	[-.02, .03]	-.09		
LA support	0.02	[-0.62, 0.66]	0.01	[-0.20, 0.22]	.00	[-.00, .00]	-.12		
LA	0.63*	[0.00, 1.25]	0.21	[0.00, 0.42]	.03	[-.03, .09]	.23*		
								<i>R</i> ² = .074	<i>ΔR</i> ² = .043
									95% CI[.00, .15]
									95% CI[-.03, .11]

Note. A significant b-weight indicates the beta-weight and semi-partial correlation are also significant. b represents unstandardized regression weights. beta indicates the standardized regression weights. sr2 represents the semi-partial correlation squared. r represents the zero-order correlation. LL and UL indicate the lower and upper limits of a confidence interval, respectively. * indicates p < .05. ** indicates p < .01.

Supporting Students' Privacy: How Does Learner Control over Their Data Affect the Dataset for Learning Analytics?

Philipp Krieter¹, Michael Viertel² and Andreas Breiter³

Abstract: An increasing amount of data for learning analytics is available due to the spread of digital learning environments. Collecting data on student behavior and action in these environments raises the question of how personal privacy can be addressed, as students are the most important stakeholder in the process. In our study we propose a simple but powerful approach to enhance student privacy: a switch to turn the data collection on and off. We let the students decide for themselves when they want to stop and go on with the collection of their data. In this paper, we take a closer look at how our students make use of this option and when and how this influences the collected data by exemplarily investigating the data of two of our participants in detail. In our analysis we combine system log files, screen recordings and qualitative interviews.

Keywords: privacy, data collection, mobile digital learning environments, screen recordings, student empowerment.

1 Introduction

While it is technically possible to collect large amounts of data from virtual learning environments, it is still a challenge to ensure students' rights of privacy and give them control over their data while at the same time enable researchers to collect necessary data. It is not easy to transfer, social, ethical, and legal aspects into system requirements. Students are important stakeholders in the learning analytics process [DG12]. This raises the question, how students in research studies can be made aware of which data they are sharing for what purpose and how they can control what they share. On the other hand, it is important to enable research and to maintain a balance between the effort and benefits of privacy ambitions, for participants and researchers alike. This is a frequent challenge in current Human-Computer Interaction (HCI) research on ethics and privacy [Br16]. Informed consent is necessary for most Learning Analytics studies unless they are

¹ University of Bremen, Institut for Information Management Bremen GmbH, Am Fallturm 1, 28359 Bremen, Germany, pkrieter@ifib.de,  <https://orcid.org/0000-0003-4421-4425>

² University of Bremen, Institut for Information Management Bremen GmbH, Am Fallturm 1, 28359 Bremen, Germany, mviertel@ifib.de,  <https://orcid.org/0000-0003-1144-7070>

³ University of Bremen, Institut for Information Management Bremen GmbH, Am Fallturm 1, 28359 Bremen, Germany, abreiter@ifib.de,  <https://orcid.org/0000-0002-0577-8685>

completely anonymous. In general, providing and using personal data in the EU is regulated by the GDPR. Hence, having a concept of privacy that can convince participants, and not just fulfilling the legal expectations, should be the goal of ethical research.

The learning setting of the presented work are digital music classes in Adult Learning Centers (ALC). We collected data on tablet devices using multiple data sources. The main data source of this study are permanent screen recordings in the background and additionally log files from a Learning Management System (LMS). The large data set of video recordings (320 hours) is analyzed using computer vision and machine learning methods [KB18a, KB18b]. Screen recordings can be compared to looking over the shoulder of a student: they are highly useful to follow in detail how students interact with a digital learning environment which contains of multiple applications and services to collect data for learning analytics [KB18b]. We chose this data source because we want to track all student actions in all relevant applications, not just in the LMS. Screen recordings are a highly detailed data source but at the same time very privacy invasive [Ta06, KVB20]. The students were in control of turning off the screen recording whenever they felt like it. Some participating students (6) agreed on having another logging framework (AWARE [FKD15]) installed on their tablet. This software tracks system log files on general usage in the background all the time and could not be turned off like the screen recordings. The log files from this framework do not provide the same level of detail as screen recordings, but it keeps track of which apps are opened and how long the screen is active. Using this setup, we can keep track of what this group of participants are doing when they turn off the screen recording. In addition to this data we conducted interviews with all participants on their perspective on the data collection and the given options of controlling it.

The proposed approach to ensure privacy is simple but powerful: we give the students the option to turn off the data collection at any time. In this short paper we explore how the students make use of this option and how this influences the collected data in our study. When and how often do they decide to stop the collection? What is happening in these breaks? How can we explain these breaks and how does this influence doing research?

We contribute to the growing discussion and interest in how we can address privacy and empowerment of students in the learning analytics data collection process. We advocate that putting the students in control of what they share with a research team, does not necessarily lower the research opportunities of the collected data. In a mixed-methods approach we evaluate an example of giving control to students over their privacy and what they share for learning analytics and how this affects the data we get. We compare and analyze the data from three different sources (screen recordings, system logs, interviews). In this short paper, we present a detailed report on the data of two of our

participants and examine at which points the students decide to stop the screen recording, what happens afterwards and how the interviews can put this data into context [HBF18]. The goal is to present a new strategy on how we can methodically address privacy.

2 Background and Related Work

With the growth of the learning analytics research and practice community ethical and privacy related considerations are gaining more attention (see e.g. work by the LACE EU-project⁴). Universities and schools that collect data and introduce learning analytics for feedback are increasingly concerned with the ethical implications of these processes [PS13] and the idea of improving privacy-control for students [Jo19]. Students are the main actors and beneficiaries of learning analysis [DG12], which makes taking their perspective and views on this process into account highly relevant. This requires adapted institutional policies for the integration of learning analytics in education [PS13]. Ifenthaler and Schumacher note that research on the student perspective on privacy issues related to learning analytics is underpinned by little empirical evidence [IS16]. They carried out a study on the perception of data protection principles in learning analytics by students. In their online study with 330 participants, they investigate this by examining how willing the participants are to share what data and for what type of learning analytics. They name transparency as one of the most important success factors. Previous research on the self-management of students' privacy points in a similar direction. For example, the work by Prinsloo and Slade [PS15] investigated how different providers of massive open online courses (MOOCs) informed students about the use and collection of their data and the options available to the students. They found that "increasing transparency in learning analysis activities will help higher education institutions gain the trust and cooperation of students" [PS15]. Drachsler and Greller [DG16] stress the importance of privacy issues and trust in learning analytics and developed a checklist (DELICATE) which can guide policy makers and institutional managers in the process. In relation to the approach of this paper, we address the point "consent" of the checklist by offering the possibility to opt-out and -in of the data collection at any time. For the successful implementation of learning analytics processes Tsai et al. [Ts18] presented the SHEILA framework for assisting policy and planning strategies for learning analytics in higher education. The option to control or opt-out of the data collection as part of the privacy concept for students is promoted by several previous work [Su18, PS14a, PS14b].

In terms of privacy, there are several concepts to address how personal data can be processed. McPherson et al. argue that there is a tendency towards a "one-size-fits-all" approach to learning analytics, across different disciplines and practices [Mc16]. A general but flexible concept which can guide software development is the concept of

⁴ <http://www.laceproject.eu/articles>

“Privacy by Design” which emerged in the ’90s. It consists of seven principles to be taken into account: Proactive not reactive, Privacy as the default setting, Privacy embedded into the design, Full functionality, End-to-end security, Visibility and transparency, Respect for user privacy [La01, Ca09]. Complete anonymization of the learner data is hardly achievable as learning analytics is strongly based on individual feedback. Also the concept of differential privacy (see e.g. [DR14]) is standing against most usage of data for learning analytics and individual feedback for students, as its goal is to hide the individual data and create more general aggregations. On the other hand, at least the data instructors or institutions see can be anonymized and still helpful for learning analytics experts as Gursoy et al. show in a privacy-preserving learning analytics experiment with a synthetic data set [Gu17].

Screen recordings raise obvious privacy issues when used for data collection in user studies, which can lead to rejection as a method of data collection by participants [Kr19]. As they reveal everything a user sees while interacting with a device, many abuse scenarios are possible. From an ethical point of view, it is therefore of interest to consider how privacy can be ensured. While anonymization can work for log files, the case is different when it comes to permanently recording the screens of participants. Giving control of what is shared directly to students follows the idea of privacy by design. A similar approach with screen recordings is pursued by Brown et al. [BMM14]. They collect mobile screen recordings and give their participants the option to directly check and delete screen recordings before sharing them with the research team. Heuer und Breiter can show that even when removing (personal) information from a learning analytics data set and just using the click frequency many prediction models still reach similar results on predicting student learning success [HB18].

Blind Review et al. [KVB20] investigate students’ perspectives on data collection via screen recording and log files in an learning analytics context and conclude transparency, extensive information, and distinction between private and learning use are important for privacy invasive data collections. In this paper, we look at the resulting data from a researcher’s perspective of how the impact on the data set is when students are given control over the collection.

3 Methods and data collection

In the following, we give an overview of the setup of the study and participants, our data sources and how we approached the analysis and combination of these multiple data sources.

3.1 Course Setting and Participants

In our project with adult education centers (ALC) in Germany, we offered a course that supports students in studying popular music in digital learning environments with

extensive self-determined learning time. The idea of the project from a learning analytics perspective, is to follow as close as possible how the students learn to make digital music. Four courses took place in late 2018 and late 2019 and lasted three to four months, depending on the schedule of the ALC. There was a weekly meeting for two hours with the course instructor. The learning setting allows room for learner initiative, combined with the possibilities of digital learning environments. The program was free and without age restrictions. The number of participants per course was limited to eight, due to the complex technological infrastructure required. All participants were provided with a tablet for the duration of the course, allowing them to connect to the Internet both at the ALC and at home. Part of the learning environment was the LMS Moodle. The LMS was set up especially for the course and offered individually designed learning content. Parts of the course communication were organized via the LMS. In total, we recruited a diverse group of 16 participants, five female and eleven male, ranging in age from 16 to 74 years. It is important to ensure that all participants understand the characteristics of our research design so that they can make an informed and reasonable decision about whether to agree to participate in our courses. To this end, we informed interested students in detail about the characteristics of data collection and analysis in our research project before the project started.

3.2 Data Sources

We collected data via permanent screen recording on all tablets in the background, which the participants could turn off. The recordings contain everything happening on the screen, which makes them a very high detailed data source to follow student actions at detailed level. In addition, we collected system log files using the AWARE framework [FKD15] on six tablets, which collected general system data (start of an application, screen on or off etc.) even if the screen recording was turned off.

3.3 Interviews

After the end of the study, we conducted interviews with all participants. The two main topics included questions of how they perceived the data collection and their attitude to data privacy and on the other hand music learning related questions. The interview results are used in the discussion part of this paper to interpret the participants' behavior in more detail. The focus and contribution of this work is on the comparison of the other data sources of the project, but the interviews generate background knowledge which makes discussing the data easier. All interviews were transcribed into text documents and analyzed using qualitative content analysis. Parts of this interview analysis have also been presented in previous work [KVB20].

3.4 Experimental Setup

In this paper we concentrate on the data of two of our participants. These two produced the highest durations of tablet usage and thus the largest amount of data. Additionally, they used the option to turn off the screen recordings most often and frequently. From this data we explore in detail what is happening in the system log files, after turning off the screen recordings and compare to the activities we can follow in the screen recordings. We present examples in form of timeline visualizations and describe student specific patterns we can find in the recording breaks.

After the combined description and analysis of screen recording and system log files in the results section, we discuss and interpret these results by adding insights from the interviews with both participants. By analyzing the interviews, we are able to put the digital traces of our students into context [HBF18] and can make assumptions about the reasons behind these data points.

4 Results and Discussion

4.1 Comparison: Screen recordings and Log Files

In the results section we take a closer look at the data of two participants in detail. We chose these two students because they actively turned off the recording and used the tablet in these breaks. The other participants (4) only rarely used this option. One never used the off-switch, two other participants only one time and two times. Another participant used the off-switch 9 times, but only to turn it on again right after (probably to just make sure recording is really turned on). Of this group we present average numbers (see table 1).

Table 1. Overview of relevant participants' data: The time frame of the considered data is from the 30th of October until the 16th of December. The average numbers contain the data of all 6 students which collected additional data with the AWARE framework.

Participant	Screen Recorded	Screen time	Not recorded	Rec. turned off
C2P1	33,6 hours	70,7 hours	52,5 %	9 times
C2P4	29,3 hours	53,6 hours	45,3 %	11 times
Average (n=6)	22,6 hours	40,8 hours	49,1 %	5 times

For student C2P1 and C2P4 we compare the collected screen recording data to the system log files we collected. We specifically pay attention to the moments when the participants actively turn off the screen recording and what happens after this or before. Table 1 shows an overview of the collected data. Both participants recorded their screen around 30 hours and used the tablet for 53 to 70 hours during the considered time frame

of 6 weeks. Considering the collection of the complete learner data via screen recording as the goal, we have a loss of around 50% for both students comparing the amount of collected video material to the system log files. This is comparable with the average in the group of these six students. The difference of recorded screen video and screen time results from the option to turn off the recording at any time. Another factor is that after starting or restarting the tablet (e.g. because the battery was low) the students had to turn on the screen recording again, due to Android security restrictions. Although the recording application starts automatically, recording was not mandatory of course, leading to less recording time. One participant reported technical issues, which resulted in crashes of the recording application, resulting in lower recording time as well.

During the collection of data one participant (C2P1) decided nine times to turn off the recording of the screen, the other one eleven times (C2P4). Both participants log data in the breaks of screen recording show participant-specific usage patterns. In the following we show two visualizations of usage sessions after or in-between recording has been stopped. We filtered out very small (milliseconds) pauses in the data. For example, if the screen goes off but gets activated immediately again by the user. Also, just activating the display for a couple of seconds but not unlocking the screen has been removed from the data.

For student C2P1 the use of the applications WhatsApp, TikTok and a video effect application are characteristic. These applications are not or only rarely present in the times screen recording is active. Another application which is present in the logs is a child-mode tool for Android.

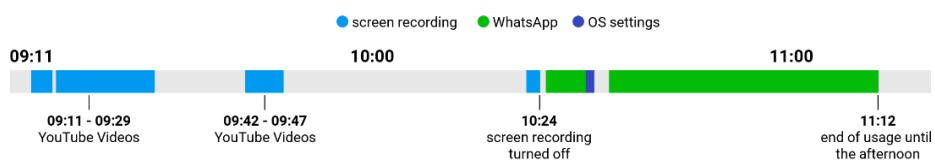


Fig. 1. Timeline visualization of C2P1 on the morning of the 3rd of November.

Figure 1 shows the tablet usage of student C2P1 on the morning of the 3rd of November. C2P1 turns on screen recording in the morning and then turns it off again at 10:47 PM. The screen recordings show that the participant is watching different music videos on YouTube. The videos range from current rap music, to 80th classics. After the screen recording is turned off, a period of WhatsApp usage follows right after, only interrupted by a short presence of the tablet's operating system (OS) settings. On this day the screen recording stays turned off. In the afternoon, we can see some more WhatsApp and also TikTok occurrences. In the evening the logs also show two usage sessions containing YouTube and the use of a piano practice application, but the screen recording is not

active. The data of this day matches the pattern of recording breaks of student C2P1, mentioned above.

For the other student (C2P4) a typical usage pattern in the recording breaks is using the tablet's internet browser or notebook application. These are the only relevant applications the student is using in these cases. In most cases this student just turns on the recording again right after a break. This pattern repeats in most cases, only two times after turning off the recording a longer period of tablet usage follows.

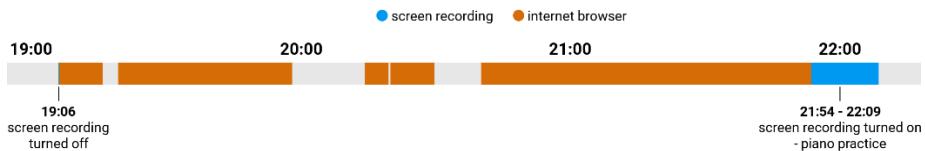


Fig. 2. This timeline visualization shows one of the usage behavior of participant C2P1 on the 21st of November.

In the timeline visualization (fig. 2) we can follow the tablet activities of C2P1 in the evening of the 21st of November, which represents all usage for the whole day. In the beginning (19:06) the student unlocks the tablet and searches for the screen recording app and deactivates the recording. After that multiple sessions of internet browser usage follow, these lasted a total of 2 hours and 16 minutes. The system log files do not show details about usage, like which website has been viewed etc. but from the screen recordings after the breaks we can tell that the browser usage in the breaks seems to be not course related. For example, from some opened browser tabs we can tell that the student seems to be using the tablet to check on soccer results and sports news in the recording breaks. At the end of the day the participant turns on screen recording again, and we can follow some piano practice in the screen recordings for around 15 minutes.

4.2 Contextualization with Interview Data

With the transfer of control over the data collection process to the student side, we saw a potential for two undesirable scenarios, that are to be avoided. On the one hand the unintended recording of private activities. On the other hand, that research relevant learning activities - accidentally or not - were not recorded. For this reason, system logs data were collected and evaluated with the screen recording and the interviews of the participants.

Generally, our analysis of the screen recordings and log data shows that private activities were only recorded in a few exceptional cases. First possible indications for this circumstance are given in the interviews. Aspects of data protection and privacy play an important role for all participants. A detailed analysis of the interviews regarding the perspectives of data collection via screen recording [KVB20] showed that for a big part

of the students the aspect of control over their data is in the center of attention. They emphasized that the physical separation of private and course-relevant hardware is for them the central moment of control. For this reason, the participants underlined in the interviews, that it was not necessary for them to switch off the screen recording on the course tablet.

In this paper we took a closer look at the two participants who actually used the opportunity to stop the recording frequently. We wondered what the impact of using the recording control on the collected data would be. Referring the relationship between their time of tablet use and the recording of use, the first thing that emerges is that most of the usage time was also captured via screen recording. In the case of our small experimental data collection, the system log data confirms that the opportunity for participants to control the timing and duration of screen recordings does not lead to the recording of substantially less data for our research. For a better understanding we start a closer examination of the use of the tablet before and after the times of on and off recording. There were worries that the participants might forget to end the recording function before private use or to reactivate it after. But this happened very rarely, which is also shown by the results of the automatic analyzed of the screen recordings. The evaluation of the screen recordings shows (as demonstrated) that the participants are aware of the screen recording and actively use the control function to protect their privacy. This awareness was supported by constant reminders during the course by the course trainer, as well as the visualization of the running and stopped recording by an icon on the tablet screen. It can be seen here that beyond the usual data protection and research-ethical clarifications and consents at the beginning of a research project, the constant personal and technical reminders of data collection during the study do not seem to have a negative effect on data collection. In the interviews, these reminders are evaluated very positively by the participants and compared to the experiences of everyday movement in the digital world. In their view, too often their personal data is collected in an inscrutable and uncontrollable way.

In a few cases, the log files show that course-related applications were used outside the screen recordings. Also, here the analysis of the interviews gives possible reference points for the reason for this behavior. For a group of participants screen recording during the course focuses primarily on the quality and scope of their own recorded activities. The aspect of performing music was of high relevance and possible assessments and evaluations by the course instructor are described which refer to content-related requirements, expectations and personal motivations concerning the course and its goals. Participants repeatedly emphasize the desire to include particularly good and presentable activities and to make them available to course instructors or researchers. For this purpose, the recordings on the tablet were deliberately stopped or other devices were used. It is shown here that it makes sense to sensitize this group in the preparation and accompaniment that not only the (perfect) results are of importance. Rather, the exercises and failures are also of great importance for research, especially for analyzing learning processes.

5 Conclusion and Future Work

The contribution of this paper is the presentation of a novel strategy to privacy in learning analytics and its results in form of the analysis of two exemplary cases. By combining data from system log files, screen recordings, and interviews we address the question of how to empower students by giving them control over the collection of screen recording influences the results of the video data set. We contribute to the growing interest and discussion of how effective privacy mechanisms can find their way into learning analytics research.

- Although our sample is limited, the two examples can show that our approach of giving control over their data to students does not influence the data set (at least in these two cases) in a way which would make it less useful for our research on learning analytics in digital music making.
- The students highlight the privacy mechanisms of the study as convincing and use it with student specific patterns, but in general to prevent private activities which are not learning related to get into the data set for research.
- The example shows that powerful privacy control and enabling high detailed research at the same time can be possible (at least in this experimental study).

Future work should include more research in similar control options for data sources in learning analytics. Especially as more data is being collected out-side of just a LMS, privacy mechanisms need to be addressed before rolling out the analytics in practice.

6 ACKNOWLEDGMENTS

We would like to thank our participants for working together with us and letting us collect data for our research, and for their time and effort doing interviews with us. This work was funded by the German Ministry of Education and Research (reference number: 01JKD1709B).

References

- [BMM14] Brown, B., McGregor, M., & McMillan, D. 2014. 100 Days of iPhone Use: Understanding the Details of Mobile Device Use. In Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services (pp. 223–232). New York, NY, USA: ACM. Retrieved from <http://doi.acm.org/10.1145/2628363.2>
- [Br16] Brown, Barry, Alexandra Weilenmann, Donald McMillan, und Airi Lampinen. 2016. „Five Provocations for Ethical HCI Research“. S. 852–863 in Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, CHI ’16. New York, NY, USA: ACM.

-
- [Ca09] Cavoukian, A. 2009. Privacy by design: The 7 foundational principles. Information and privacy commissioner of Ontario, Canada, 5.
- [DG12] Drachsler, H., & Greller, W. 2012. The pulse of learning analytics understandings and expectations from the stakeholders. Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, 120–129. <https://doi.org/10.1145/2330601.2330634>
- [DG16] Drachsler, H., & Greller, W. (2016). Privacy and analytics: It's a DELICATE issue a checklist for trusted learning analytics. Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, 89–98. <https://doi.org/10.1145/2883851.2883893>
- [DR14] Dwork, C. and Roth, A., 2014. The algorithmic foundations of differential privacy. Foundations and Trends in Theoretical Computer Science, 9(3-4), pp.211-407.
- [FKD15] Ferreira, Denzil, Vassilis Kostakos, and Anind K. Dey. 2015. „AWARE: Mobile Context Instrumentation Framework. Frontiers in ICT 2. doi: 10.3389/fict.2015.00006.
- [Gu17] Gursoy, M. E., A. Inan, M. E. Nergiz, und Y. Saygin. 2017. „Privacy-Preserving Learning Analytics: Challenges and Techniques“. IEEE Transactions on Learning Technologies 10(1):68–81. doi: 10.1109/TLT.2016.2607747.
- [HBF18] Hepp, A., Breiter, A., & Friemel, T. N. 2018. Digital Traces in Context — An Introduction. International Journal of Communication, 12, 439–449.
- [HB18] Heuer, H. and Breiter, A., 2018. Student success prediction and the trade-off between big data and data minimization. DeLFI 2018-Die 16. E-Learning Fachtagung Informatik.
- [IS16] Ifenthaler, D., & Schumacher, C. 2016. Student perceptions of privacy principles for learning analytics. Educational Technology Research and Development, 64(5), 923–938. <https://doi.org/10.1007/s11423-016-9477-y>
- [Jo19] Jones, Kyle M. L. 2019. „Learning Analytics and Higher Education: A Proposed Model for Establishing Informed Consent Mechanisms to Promote Student Privacy and Autonomy“. International Journal of Educational Technology in Higher Education 16(1):24. doi: 10.1186/s41239-019-0155-0.
- [KB18a] Krieter, P., & Breiter, A. (2018, September). Analyzing mobile application usage: generating log files from mobile screen recordings. In Proceedings of the 20th international conference on human-computer interaction with mobile devices and services (pp. 1-10).
- [KB18b] Krieter, P., & Breiter, A. 2018. Track every move of your students: log files for Learning Analytics from mobile screen recordings. DeLFI 2018-Die 16. E-Learning Fachtagung Informatik.
- [Kr19] Krieter, P. 2019. Can I record your screen? Mobile screen recordings as a long-term data source for user studies. In Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia (pp. 1-10).
- [KVB20] Krieter, P., Viertel, M., & Breiter, A. 2020. We Know What You Did Last Semester: Learners' Perspectives on Screen Recordings as a Long-Term Data Source for Learning Analytics. In European Conference on Technology Enhanced Learning (pp. 187-199). Springer, Cham.
- [La01] Langheinrich, M. 2001. Privacy by Design — Principles of Privacy-Aware Ubiquitous Systems. S. 273–91 in Ubicomp 2001: Ubiquitous Computing, Lecture

- Notes in Computer Science, published by G. D. Abowd, B. Brumitt, und S. Shafer. Berlin, Heidelberg: Springer.
- [Mc16] McPherson, J., Tong, H. L., Fatt, S. J., & Liu, D. Y. T. 2016. Student perspectives on data provision and use: Starting to unpack disciplinary differences. Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, 158–167. <https://doi.org/10.1145/2883851.2883945>
- [PS15] Prinsloo, P., & Slade, S. 2015. Student privacy self-management: Implications for learning analytics. Proceedings of the Fifth International Conference on Learning Analytics And Knowledge, 83–92. <https://doi.org/10.1145/2723576.2723585>
- [PS14a] Slade, S., & Prinsloo, P. (2014, October). Student perspectives on the use of their data: Between intrusion, surveillance and care. In EDEN Conference Proceedings (No. 2, pp. 291-300).
- [PS14b] Pardo, A., & Siemens, G. 2014. Ethical and privacy principles for learning analytics. British Journal of Educational Technology, 45(3), 438-450.
- [PS13] Prinsloo, P., & Slade, S. 2013. An evaluation of policy frameworks for addressing ethical considerations in learning analytics. Proceedings of the Third International Conference on Learning Analytics and Knowledge, 240–244. <https://doi.org/10.1145/2460296.2460344>
- [Ta06] Tang, John C., Sophia B. Liu, Michael Muller, James Lin, und Clemens Drews. 2006. „Unobtrusive but Invasive: Using Screen Recording to Collect Field Data on Computer-mediated Interaction“. S. 479–482 in Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work, CSCW '06. New York, NY, USA: ACM.
- [Ts18] Tsai, Y.-S., Moreno-Marcos, P. M., Jivet, I., Scheffel, M., Tammets, K., Kollom, K., & Gašević, D. (2018). The SHEILA Framework: Informing Institutional Strategies and Policy Processes of Learning Analytics. *Journal of Learning Analytics*, 5(3), 5–20–25–20. <https://doi.org/10.18608/jla.2018.53.2>
- [Su18] Sun, K., Brooks, C., Mhaidli, A. H., Schaub, F., & Watel, S. (2018). Taking student data for granted? A multi-stakeholder privacy analysis of a learning analytics system. In EDM 2018 Workshop on Policy and Educational Data Mining.

FAIR Research Data Management for Learning Analytics

Ian Wolff ¹, David Broneske  ², and Veit Köppen  ³

Abstract: Learning analytics is a data-driven science in which scientists collect manifold data of different types and in different formats. Still, research-data-management culture does not exist so far. This means, concepts for a common data handling do not exist in this domain. We counter this problem by using the FAIR principles for data management in learning analytics research projects. The principles seek to improve Findability, Accessibility, Interoperability, and Reuse of digital assets. Our core contribution is, thus, to show how the FAIR principles is applicable in a discipline specific learning-analytics-data repository and what is necessary for a good data handling.

Keywords: research data management, learning analytics, guidelines, metadata, data repository

1 Introduction

A lot of learning analytics methods produce digital data by using various e-learning platforms. When research data is generated, a feasible data management for the discipline is required. Learning analytics does not possess a research-data-management culture yet, even though the discipline is driven by digital data. Therefore, a holistic infrastructure for supporting the reuse of learning analytics data across single studies is urgently necessary, because the data bare enormous potential for reuse and secondary analysis of third-party researchers, to run pretests or even build upon already existing data. Nonetheless, studies have shown that there are certain concerns about learning analytics data sharing from researchers all around the world [BP21]. Accordingly, infrastructure and guidelines for supporting the reuse of learning analytics data are needed.

One starting point to allow for reuse of data could be the FAIR principles for scientific data management and stewardship [WDA16]. These principles provide a guideline for the publication of digital resources to make them findable, accessible, interoperable, and reusable. A learning analytics domain-specific repository can benefit from applying the FAIR principles, because they imply community standards in the point of metadata, vocabulary, and data formats to increase the reusability of data sets. This comes with several advantages for learning analytics scientists, when they search for appropriate data for their research question.

¹ Otto-von-Guericke-University Magdeburg, Universitätsplatz 2, Magdeburg, 39106, ian.wolff@ovgu.de, <https://orcid.org/0000-0002-0413-0035>

² German Centre for Higher Education Research and Science Studies, Lange Laube 12, Hannover, 30159, broneske@dzhw.eu, <https://orcid.org/0000-0002-9580-740X>

³ Zentral- und Landesbibliothek Berlin, Breite Str. 30 – 36, Berlin, 10178, veit.koeppen@zlb.de, <https://orcid.org/0000-0002-6068-3275>

An important first step for research data management is a domain-specific data repository and metadata schema, which currently do not exist for the learning analytics community.^{4,5} The neighboring discipline, educational data mining, already possesses a research data repository for the community in the American area, but the repository does not follow the FAIR principles [Ko11].⁶ There is also the option to store learning analytics data in generic repositories, e.g., Zenodo, which adopts most of the FAIR principles, but it does not meet community standards, like a discipline-specific metadata model, vocabulary or restricts the data set regarding data formats used.

The superior goal is the reuse of research data and the citation of research data and metadata by meeting community standards. Therefore, the principles apply for people and machines in the same way. Additionally, funding organizations like the new Horizon Europe funding program (2021–2027) refer to the FAIR principles regarding the open access of research results,⁷ as well as the German Research foundation in their code of conduct – guidelines for safeguarding good research practice –.⁸ As a result of the benefits of the FAIR principles, we propose to apply the principles for learning analytics projects to be able to share learning analytics data across single studies. To this end, we present actions that are necessary to set up a learning-analytics-research-data repository by considering the FAIR principles.

2 FAIR Principles in Learning Analytics Repository Usage

As the FAIR principles play a vital role in research data management, it is of utmost importance to apply them in any research project. However, especially in interdisciplinary research projects with little expertise in data management (e.g., in higher education research), understanding and applying the FAIR principles could be a big effort. To lower this effort, we review single components of the FAIR principles, and we show how to apply them within a learning analytics scenario.

2.1 Findability

To make research data findable for humans and machines in the same way, a persistent identifier (PID) has to be used. That means for example the usage of a digital object identifier (DOI), which is a permanent link to a digital object and gives access to information about the linked object.⁹ We achieve this by using metadata. DOIs refer for instance to the DataCite schema, that inherits a metadata for a URL to indicate where the

⁴ The metadata directory of the research data alliance lists no specific metadata schema for learning analytics, metadata directory, <https://rd-alliance.github.io/metadata-directory/>, accessed: 21/06/2021.

⁵ Registry of Research Data Repository, <https://www.re3data.org/>, accessed: 21/06/2021.

⁶ DataShop, <https://pslcdatashop.web.cmu.edu/>, accessed: 21/06/2021.

⁷ Horizon Europe, <https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/9570017e-cd82-11eb-ac72-01aa75ed71a1>, accessed: 28/07/2021.

⁸ Guidelines for Safeguarding Good Research Practice, Code of Conduct, <https://doi.org/10.5281/zenodo.3923602>, accessed: 27/07/2021.

⁹ DOI, <https://www.doi.org/index.html>, accessed: 27/06/2021.

object can be found.¹⁰ The repository must allow the usage of different metadata schemes, ranging from general schemes, like the Dublin Core (DC)¹¹, to a discipline specific one. Important for learning analytics is the use of a rich and accurate description of relevant attributes for the learning-analytics-research process to meet community standards. Here, we develop a discipline-specific metadata schema that provides information about the *learner* as well as all *measurement instruments* used for data collection and in what *environment* the collection takes place, accompanied by observed *activities* and *measurement results* like variables.¹² Therefore, already existing metadata schemas can be used, like the aforementioned DataCite, but also parts of other schemas can be used like the DC for general description and the Research Data Repository RADAR,¹³ the Learning Objects Metadata (LOM)¹⁴ for discipline-specific description, and Preservation Metadata Implementation Strategies (PREMIS)¹⁵ for technical aspects.

If the data model is well-structured, a standard protocol, like XML, can be used to make the export of metadata possible for discovery tools. Furthermore, standard protocols for the consumption and distribution of metadata are realized through the Open Archive Initiative-Protocol for Metadata Harvesting (OAI-PMH) [DPG11]. On the one hand, the protocol allows collecting metadata from other repositories and, on the other hand, the repository's metadata can be collected. Furthermore, the repository can be listed as a discovery tool in the world registry of research data repositories re3data.¹⁶

2.2 Accessibility

To ensure human and machine access to the learning analytics data, a repository should use a standard protocol, like HTTP/HTTPS, such that metadata are accessible with proper access rights. Therefore, a system of rights and roles must be defined and applied to the repository. This is important for learning analytics and one of the biggest hurdles due to privacy issues. The role system can adapt the data curation continuum of Treloar, which defines differences of data usage in private, collaborative – research project team –, and public domain [TK19]. For each domain, access rights and a level of anonymization for the data must be modified. Furthermore, metadata should be accessible, even if the data set is not available anymore, which could be realized by saving the metadata in another institution.

¹⁰ DataCite, <https://datacite.org/>, accessed: 27/06/2021.

¹¹ Dublin Core Metadata Initiative, <https://www.dublincore.org/specifications/dublin-core/dces/>, accessed: 27/06/2021.

¹² Project DiP-iT, <http://dip-it.ovgu.de/>, accessed 06/08/2021.

¹³ RADAR, <https://radar.products.fiz-karlsruhe.de/en/radarfeatures/radar-metadatenschema>, accessed: 28/07/2021.

¹⁴ IEEE Standard for Learning Object Metadata, https://standards.ieee.org/standard/1484_12_1-2020.html, accessed: 27/06/2021.

¹⁵ PREMIS, <https://www.loc.gov/standards/premis/>, accessed: 27/06/2021.

¹⁶ Registry of Research Data Repository, <https://www.re3data.org/>, accessed: 21/06/2021.

2.3 Interoperability

Data must be provided in such a way that exchange and interpretation is possible within other systems. Therefore, the usage of the human and machine-readable standard protocol XML is recommended to make metadata import and export possible. For data interoperability and extensibility, standard protocols like JavaScript Object Notation (JSON) are usable to allow one machine to store or process their data on other machines. Furthermore, the vocabulary should be documented according to domain-specific standards to make datasets easily findable and accessible for everyone. One starting point could be the glossary of learning analytics developed by the LAEP project.¹⁷ When data sets are linked with other data sets, the metadata should include qualified references to other data, e.g. DOI.

2.4 Reusability

Optimizing the reuse of data is an ultimate goal of the FAIR principles. As aforementioned, data should be well-described in the point of their provenance, such that they can be replicated and combined in different settings. A license must be linked to the dataset to define the rights attached to the data and how they can be (re-)used. What is also important for automated searches is to clarify the conditions for machines and humans. Creative commons, for example, offer a wide range of different licenses.¹⁸ This point must be proven for further usage, because it depends on the level of anonymization and on the characteristics of the collected data. If the data is fully anonymized, open access of the data can be guaranteed as public use file. If the data includes personal information, the repository needs an access concept and an internal rights system to provide access to the data as a scientific use file. Nonetheless, researchers are also responsible to take all relevant actions for their data curation into consideration [BAL18] [GPB14]. This includes for instance the preparation of their research in form of a data management plan [Mi15]. Learning analytics-research-data management only functions, if all relevant privacy matters are taken into account.

Digital data are fragile, depending on the type of medium on which they are stored, they may be exposed to damage over time. A preservation strategy can mitigate risks and ensures long-term access to data. Besides these infrastructure concerns, stored data of a learning analytics repository should meet criteria for sustainable data formats. For instance, sustainability factors of the library of congress can be adapted, which proofs the disclosure, adoption, transparency, self-documentation, external dependencies, impact of patents, and technical protection mechanisms of data formats.^{19,20} Basic requirements for

¹⁷ LAEP Glossary, <https://iet-ou.github.io/cloudworks-ac-uk/cloud/view/9781/links.html>, accessed: 27/06/2021.

¹⁸ Creative Commons, <https://creativecommons.org/licenses/?lang=de>, accessed: 27/06/2021.

¹⁹ Library of Congress, <https://www.loc.gov/preservation/digital/formats/sustain/sustain.shtml>, accessed: 27/07/2021.

²⁰ Library of Congress, <https://www.loc.gov/preservation/resources/rfs/RFS%202020-2021.pdf>, accessed: 27/07/2021.

long-term access to data are the use of non-proprietary open standard formats with a software code free of charge.

3 Discussion

The use of FAIR principles and data sharing in general comes with several advantages for researchers. Publication of data leads to more citations, exchange with colleagues leads to new insights. Scientists who make data available to peers for reuse, receive feedback on the quality of their research and sharing prevents cost for gathering the same data a second time. Nonetheless, research data sharing in learning analytics comes with two main hurdles that also limit the use of FAIR principles. (1) A very general point of learning analytics is data sharing, and this is related to privacy issues. Usually, learning analytics data sharing is starting before the release of a research project. Hence, a discipline-specific data management plan is necessary with an informed consent between research subjects, in order to store data in a repository and make these data accessible later on. This includes all stakeholders – from the researcher to the institution of the data repository [Mi15] [GPB14]. (2) Full open access sharing of learning analytics data is only possible with anonymization of the data sets. When personal information is involved, the data sets can only be made accessible as scientific use file, which requires skilled personal from a data center to verify the person who wants to use the data and to control the usage of data. Therefore, depending on the research question, an automatic access verified by a machine, like the FAIR principles recommend, is not possible in all cases of learning analytics data management. Even in a smaller context of a research team, a role and access concept for a repository is required that considers subject's privacy by making the research possible in the same way. Already existing generic research data repositories are lacking regarding the discipline-specific characteristics needed from the FAIR principles, like a metadata schema, used vocabulary, or file formats. They could adapt in these points, but most research data repositories, Zenodo for instance, are specifically claiming not to be a domain-specific repository.²¹ DataShop could pick up FAIR principles by implanting a persistent identifier and a discipline-specific metadata schema with mandatory metadata recommendations.

4 Conclusion

We have shown how the FAIR principles could be used for a learning analytics research data management and which discipline-specific features are needed for a learning analytics repository. Most important for the findability of data is the use of a PID and rich metadata to describe the data sets. Learning analytics accessibility of data can be realized by defining access rights for data users depending on the persons' context to the data.

²¹ ZENODO, <https://about.zenodo.org/principles/>, accessed: 28/07/2021.

Depending on the type of data, skilled personal and data management infrastructure is required. By using standard protocols, like XML, and domain-specific vocabulary, learning analytics data can be linked to each other and can be made interoperable in this way. Reusability is another goal of the FAIR principles and can be ensured in the learning analytics context by giving an adequate license for a data set and using sustainable file formats. We urgently claim for the development of a discipline-specific research data repository, that provides the infrastructure and information for a FAIR orientated research data management.

Acknowledgements

Funding: This work was supported by the German Federal Ministry of Education and Research [grant number 16DHB 3008]

Bibliography

- [BAL18] Borghi, J.; Abrams, S.; Lowenberg, D. et al.: Support Your Data: A Research Data Management Guide for Researchers. *Research Ideas and Outcomes* 4, 2018.
- [BP21] Biernacka, Katarzyna; Pinkwart, Niels: Opportunities for Adopting Open Research Data in Learning Analytics, In (Tomei, L.; Azevedo, A.; Azevedo, J.M. et. al. ed.): Advancing the Power of Learning Analytics and Big Data in Education, IGI Global, Hershey, pp. 29-60, 2021.
- [DPG11] Devarakonda, R.; Palanisamy, G.; Green, J. et al.: Data sharing and retrieval using OAI-PMH. *Earth Science Informatics* 4/1, pp. 1-5, 2011.
- [GPB14] Goodman, A.; Pepe, A.; Blocker, A. et al.: Ten simple rules for the care and feeding of scientific data. *PLoS computational biology* 10/4, 2014.
- [Ko11] Koedinger, K.R.: A Data Repository for the EDM Community: The PSLC DataShop. In (Romero, C., Ventura, S., Pechenizkiy, M., Baker, R., ed.): Handbook of Educational Data Mining, CRC Press, Boca Raton, pp. 43-55, 2011.
- [Mi15] Michener, W.: Ten Simple Rules for Creating a Good Data Management Plan. *PLoS computational biology* 11/10, 2015.
- [MSM19] Miksa, T.; Simms, S.; Mietchen, D. et al.: Ten principles for machine-actionable data management plans. *PLoS computational biology* 15/3, 2019.
- [WDA16] Wilkinson, M. D.; Dumontier, M.; Aalbersberg, I. J. et. al.: The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data* 3, 2016.
- [TK19] Treloar, A.; Klump, J.: Updating the Data Curation Continuum, *IJDC* 14/1, 87-101, 2019.

Indicators of group learning in collaborative software development teams

Benjamin Weiher¹, Niels Seidel¹ , Marc Burchart¹  and Dirk Veiel² 

Abstract: The supervision of collaborative software projects is a great challenge for teachers. All learners involved must be able to participate in the learning process and group collaboration must be ensured, while the program code can take on a considerable size. In this paper, we identified, defined, and validated a total of 32 indicators for collaborative learning in software development teams. The resulting model describes collaborative interactions in programming teams considering 7 indicators for code quality, 16 group participation indicators, and 9 indicators for group cohesion. In addition, we presented a data processing pipeline for extracting, calculating, and visualizing these indicators on a teacher dashboard. This approach enables teachers to keep track of complex group activities and individual contributions, and subsequently provide targeted formative feedback to the groups.

Keywords: Teaching Collaborative Programming; Learning Analytics; Group Assessment

1 Introduction

Software development takes place almost exclusively in teams, so it is especially important to be a team player, to be organized, and to communicate properly [Ah13, RTLr19]. Learning such competencies is crucial for employability and has great relevance for the job market. In higher education, these skills are therefore often trained in groups, within the context of computer-supported collaborative learning (CSCL) [SMG20]. In computer science education, students must be prepared for this collaboration by developing their competencies in the required methods (e.g. Scrum or Adaptive Software Development), tools, and the programming languages used. Hence version control systems (e.g. Git, CVS, and SVN) and issue tracking systems (e.g. GitLab, Bugzilla, and Zendesk) are widely used, both in education and in the software industry (e.g. [Gl19]). These systems generate data that can be used not only for risk analysis of projects (e.g. [MGM19]) but also to support students working together on a software project.

¹ FernUniversität in Hagen, Faculty of Mathematics and Computer Science, Universitätsstraße 1, 58084 Hagen, benjamin.weiher@studium.fernuni-hagen.de, niels.seidel@fernuni-hagen.de, marc.burchart@fernuni-hagen.de,  Niels Seidel <https://orcid.org/0000-0003-1209-5038>,  Marc Burchart <https://orcid.org/0000-0001-5668-7137>, 

² FernUniversität in Hagen, Research Cluster D²L², Universitätsstraße 27, 58084 Hagen, dirk.veiel@fernuni-hagen.de,  Dirk Veiel <https://orcid.org/0000-0003-0228-103X>

In this paper, we argue that teachers can benefit from monitoring tools that represent students' contributions in collaborative software development tasks. When supervising student teams in these settings, teachers need to maintain the learning situation for the students, ensure that all team members acquire knowledge in the different domains (e.g. project management, design, development, and testing), support completely heterogeneous students (e.g. regarding skills, pre-knowledge, and interests) when problems and questions arise, and provide them appropriate feedback during the development phases. This guidance and support is an enormous challenge. In software project teams, the lines of code increase enormously over time. The code and its quality can often only be analysed with great effort. Furthermore, the individual contribution of each student across multiple files, commits, and branches is not easily to identify. Hence, the degree of collaboration can only be examined through student reports. Difficulties of individual students or even free-riding effects can remain hidden over a long period of time, so that the participation of all students in the learning process cannot be ensured.

Our presented approach aims to support teachers in the challenges mentioned above. The goal of this work is to identify, define, and validate indicators for collaborative learning in software development teams. To address the particularities of collaborative learning and software engineering, we pose the following three research questions: **(RQ1)** What indicators can be used to describe learner participation in collaborative programming teams? **(RQ2)** Which indicators provide insights into collaborative software development? **(RQ3)** What indicators are suitable for capturing code quality in dedicated programming languages? Finally, a fourth research question is posed combining the answers of the previous questions: **(RQ4)** How can teachers be supported in using these indicators? The indicators are derived from existing literature and adapted to the subject matter. Before the indicators are prepared for use by teachers, a validation is performed using real data sets from four learning groups. As a result of this work, teachers will be able to track complex group activities, code quality and individual contributions, and subsequently provide targeted formative feedback to the groups.

With this paper, we contribute to the field of CSCL and learning analytics. Our contribution to CSCL consists of a model describing collaborative interactions in programming teams considering coding, group participation, and group cohesion. Regarding learning analytics, we present a processing pipeline for analysing data from version control systems and issue tracking systems. From this pipeline, we gain data for a teacher dashboard for monitoring individual and group-related progress across multiple iterations of software development.

2 Related Works

In terms of teaching and learning, there is already some work that investigates collaborative software development [Ri19, SA20, Bu20]. However, these only consider individual contributions, not the group collaboration. [Gi20] analyzed commit messages

and classified team members as being collaborative, cooperative, or solo-submitters. In contrast to our work, the authors only considered data from the version control system, but did not include the communication and discussion necessary for project management and design. Furthermore, code metrics have not been examined. In a case study [TWM20] investigated the use of GitHub as a teaching tool for individual assignments. By analyzing the commit history and evaluating the comment quality the authors tried to classify students in order to find proxies for grading. This attempt was not very successful. Unlike our research, [TWM20] did not intend to support teaching on collaborative software development through formative but data-driven feedback provided by the teacher. Apart from the commits and the code comments, only a very small set of data was used for the modeling of indicators. The resulting source code and the communication among the students was not considered. More advanced analytics approaches consider, for example, co-editing networks [GSS19], commit quality [AAM15], refactoring detection [Ts18], change patterns [MM19], and risky commit prediction [RGS15], but without addressing aspects of learning. Beside that, automatic methods are still lacking for specific problems such as common errors in the use of Git [Er20]. Personal assistance systems support collaboration only on a low-level [ČS05] or even hinder it [We20], which is why instructors still play a central role in guiding student groups. In our approach we wanted to use a comprehensive set of indicators to flexibly support different didactic scenarios, team and project structures, and software-technical possibilities.

3 Model for group learning in software development

Before we can start modeling, we need to know the data that is available for modeling and can later be extracted automatically from the systems used. In this case, we rely on a version control system and an issue tracker. Version control systems provide three types of information: (1) code-related data regarding the quantity and quality of the program (e.g. code smells, security hotspots) as well as (2) logs and (3) content of commits, branches, and merges. Since the content data requires a qualitative analysis in view of the respective task, we focus on quantitatively exploitable data from the source code and the logs which are described in the following subsections. Issue trackers complement this data with created and commented issues, merge requests, and project plans (e.g. Kanban board). The presented model aims at selecting a set of variables from version control system and source code that can be translated into quantitative indicators that are easy to acquire and process by teachers. The indicators should describe individual efforts as well as the cooperation in the team. The chosen subset helps teachers to provide regular individual and group feedback in terms of learning practical skills and increasing employability. Indicators focusing on efficiency gains and risk avoidance appeared to be more relevant in professional and economic contexts and have not been considered here. The resulting indicators that are relevant for learning can be later used to present an overall picture of the collaborative process within a student-led software development team. This picture is intended to ensure and promote appropriate peer teaching with the

focus on supporting students in solving practical tasks or problems and developing programming skills. The model is based on the previous work of [Ca10] on effective group models and the considerations of [HG01] about teamwork.

3.1 Indicators for software code maintainability (RQ3)

Usually, as the size of a software project increases, so does the number of program errors, security vulnerabilities, and the maintenance effort. Developers therefore try to use tools and selected programming languages to detect certain types of errors at an early stage or even to exclude them completely. Even novice programmers can make use of these tools and improve the quality of their code, as long as they have configured their development environment accordingly and as long as they can understand and implement the advice. The same tools for the analysis of the software code maintainability (cf. [Ar20]) are suitable for the formative analysis and assessment of learning achievements in programming. We consider the indicators S1–S7 to be relevant in order to answer RQ 3: (S1) *Programming languages*: Number of programming languages including style sheet languages in use. In web development, for instance, this indicator can help to identify full-stack developers compared to the one that stick to a single programming language. (S2) *Code smells*: [Fo98] introduces the metaphor “code smells” to describe the patterns in the code that indicate the need for code refactoring. The external behavior of the code remains the same when refactoring, while the internal structure improves, i.e., it appears tidier, more traceable, and easier to maintain [EM02]. (S3) *Cognitive complexity*: Cognitive complexity is a measure of the understandability of a given piece of code, the complexity of which is determined by the number and order of control structures [Ca18]. (S4) *Security hotspots*: The security hotspot describes the sensitive areas where security is more important than in other areas. The most important security hotspots are authentication, storage, cryptography, logic errors, synchronization and timing, and validation [Se19]. (S5) *Vulnerabilities*: The number of vulnerabilities that refers to problems in the source code identified from poor coding patterns [Fo98] that lead to bugs, security vulnerabilities, performance problems, design flaws, and other difficulties [So21]. (S6) *Bugs*: Number of errors due to a specification that was not adhered to or incorrectly implemented (e.g. typing of return values). (S7) *Duplicated lines*: The absolute number of physical lines (not just lines of code) of source code that are involved in at least one additional location [CP13].

3.2 Indicators for participation (RQ1)

Extent of participation The extent of participation is a simple but essential aspect of collaboration. For collaboration to occur at all, the student must participate in the project. Following [Ta19], *Commit Count* (P1), *Opened Merge Request Count* (P2), *Branch Count* (P3), *Comment Count* (P4), and *Issue Count* (P5) has to be recorded.

Equal participation In a collaborative software project, individual project members may do most of the work. In an effective collaborative group, all members should participate to a similar degree without monopolizing behavior [Ca10]. Equal participation can be considered by the absolute deviation from the mean. In this way, the scope of individual contributions within a group can be quantified and compared using the indicators *Equal Commits* (P6), *Equal Merges* (P7), *Equal Issues* (P8), and *Equal Comments* (P9).

Extent of roles Related to participation, another factor may regard the variety of roles taken on by the members of the group. A good group should be one in which roles are played flexibly with participants rotating their roles during an iteration. This seems to be indicative of the attention paid to the whole group's process of planning tasks, developing code, and reviewing the quality. In small groups, the extent to which different roles are performed should be independent of assigned roles in agile projects (e.g. product owner) or traditional roles like team leader or quality engineer. In this sense, we define an *Active Reviewer* (P10) as someone who has left a comment on a merge request. This indicator is the ratio of active reviewers to students who have created a merge request. The share of *Active Developer* (P11) in a team considers those who created merge requests. The relative number of *Active planner* (P12) created or modified an issue.

Rhythm The rhythm of interactions is a metric that allows conclusions to be drawn about the synchronicity of activities. Regular and constant participation can be considered as an indicator of the individual's ability to deal primarily with the needs of the group rather than with personal problems, which also avoids the risk of distraction and a decline in cognitive tension. Considered for this purpose are the indicators *Coding Days* (P13), *Review Days* (P14), *Testing Days* (P15), and *Planning Days* (P16). Coding Days, for instance, counts the number of days a developer has contributed code to the project. Similarly, this applies to other important team tasks such as planning, reviewing each other's work, and testing.

3.3 Indicators for cohesion (RQ2)

Caring for one another among team members is critical to strengthening mutual trust and a sense of belonging, as well as the perception of positive interdependence among members of a group. (C1) *Commit-Comment-Ratio*: Describes the ratio of comments in commits to the number of commits created. In a collaborative project, the proportion of own commits and the number of comments should be balanced to ensure that the student both contributes their own code to the project and is willing to look at other students' code. (C2) *Reaction Time*: Response time is the time it takes a reviewer to respond to a comment directed to them. The other team members should respond to a comment on time so that the questioner does not lose time. (C3) *Responsiveness*: The response time is the average time taken to respond to a reviewer's comment with either another comment or a code revision. It shows the time between the last comment of the reviewer and the

response of the creator of a merge request. (C4) *Follow on Commits*: The number of code revisions added to a pull request after it was opened for review. Knowing the number of follow-up commits added to an open pull request will give you insight into the strength of your code review process. If you see a trend where many follow-up commits are added, further planning and testing may be required [Pl21]. The *Receptiveness* (C6) is the frequency with which the creator of the merge request takes comments as a reason to change the code (cf. [Pl21]). In addition to that, the *Time to first Comment* (C6) specifies the time between opening a pull request and the first reviewer commenting (cf. [Pl21]) and the *Time to resolve* (C7) is the time it takes to close a pull request (cf. [Pl21]).

Mutual help in programming can be expressed in continuing or improving the work of others. Therefore, the indicator *Helping others* (C8) describes the percentage of commits that a developer has used to modify the code of his team members (cf. [Pl21]). *Reactivity to proposals* (C9) Suggestions for new code contributions, which are called merge or pull requests, are essential for progress in a project. Therefore, suggestions from a team member must be immediately considered by the team.

4 Realisation and validation

Building on the indicators defined in the last section, we describe the technical processing steps for determining the indicators, present a dashboard for monitoring groups, and validate the indicators using data from four student groups.

4.1 Analytics environment

The analytics environment is responsible for storing the transformed raw GitLab data, their analysis, and presentation to the teachers. The component consists of four modules: First, the Extract, Load, Transform (ELT) [Go10] module performs the daily job of extracting the data from the GitLab projects, loading it to the Data Lake module, and pre-processing it for the analysis. Using SonarQube [So21] a code analysis is performed with the GitLab data, before the described indicators are computed. The Data Lake, as the second module, stores the raw data and their transformed version in a relational database schema. For data Analysis the third module performs additional analytics, including data validation as described in section 4.2. Finally, the fourth module consists of a Analytics Dashboard (AD) used by teachers. It is based on the R and Shiny³ package. For the AD the data from the Analytics and the Datalake is processed to a web-based interactive visualization. Fig. 1 shows the resulting dashboard for an exemplary group of 6 students that have been collaborated over 3 iterations. The indicators' values have been normalized for the sake of comparison. The visualization as small multiples

³ See <https://shiny.rstudio.com/> (last accessed: 2021-06-27).

enables a comparison between students on each iteration and for all indicators. For instance, user 33 was commendably able to reduce the cognitive complexity of his code over the three iterations. User 34's comparatively low participation (e.g., P1, P2, P4) should prompt the teacher to ensure learning success. Regarding group cohesion, the teacher could discuss with the team how to achieve mutual help during the programming task, since indicators C1, C2, C3, and C8 have low values.

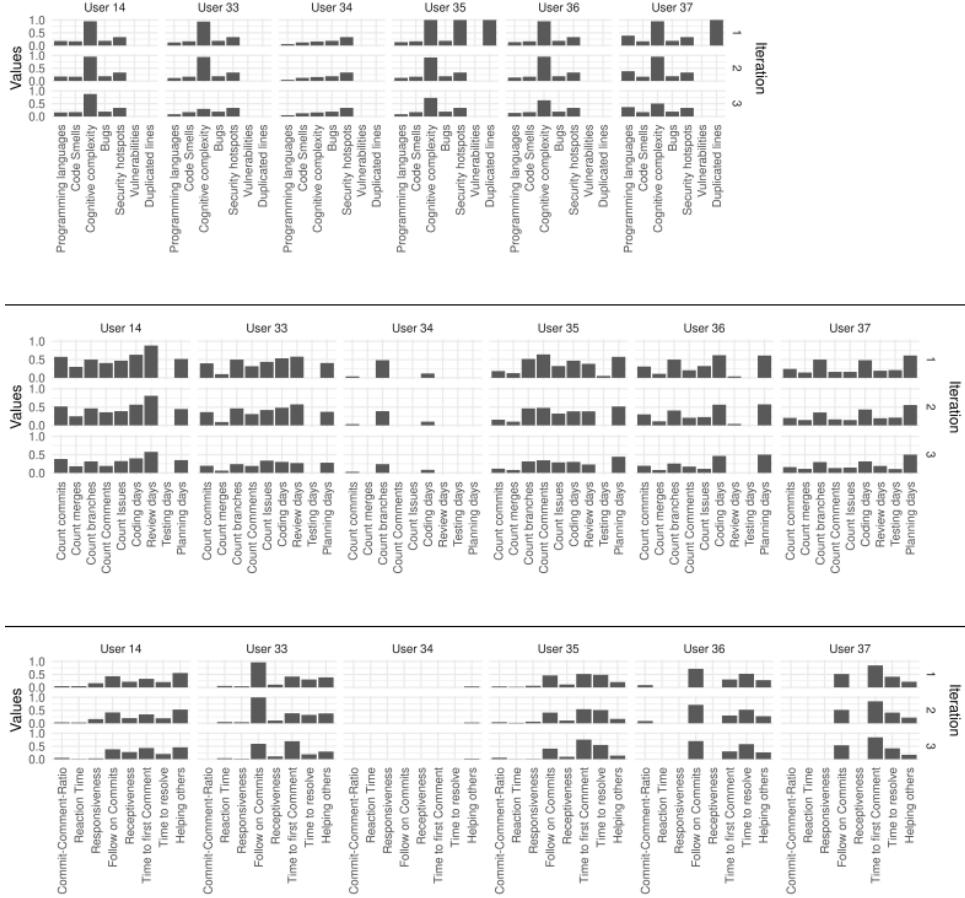


Fig. 1: Example output for the first three iterations of group C including the dimensions code quality (top), participation (middle), and cohesion (bottom)

4.2 Validation of indicators

Our first attempt to validate the model described in section 3 aims at testing the appropriateness, feasibility, and expressiveness of the indicators. For the validation, we used data from four mixed software development teams that worked on a complex task over 7 iterations à 2 weeks as part of the course “Fachpraktikum CSCW” (cf. Tab. 1). The agile teams worked on different tasks, which were however comparable in terms of effort and demand, in regular dialog with a role-played customer. The teams consisted of 5 to 6 students from Bachelor’s and Master’s programs in Computer Science. GitLab was used for collaboration. Each group worked in at least one GitLab project using provided issue tracker, version control, etc.

Group	Semester	Participants	Repositories	Issues	Commits
A	WS 2019/20	5	1	191	4402
B	WS 2020/21	5	1	192	2962
C	WS 2020/21	6	1	145	2419
D	WS 2020/21	5	3	128	1441

Tab. 1: Overview of the datasets used for validating the model

The indicators can be considered appropriate because they were piloted against the data of the teams in Tab. 1 and discussed with experienced teachers. In addition, the indicators have been adopted from existing tools and research. In terms of feasibility, it must be ensured that all indicators in the model can be calculated using data from real courses. The tests have shown that the calculation is feasible for the datasets at hand. To validate the expressiveness of the indicators as a formative feedback instrument in the supervision of student teams, the indicators should discriminate students to estimate differences among group members. This can be achieved by calculating the standard deviation (SD) among group members and the discrimination index (DI, see [De10]) across all indicators. The SD indicates heterogeneous values of an indicator within a group. If, on the other hand, the values within a group were always the same, the indicator would have little significance for a teacher. The number of review and planning days appeared to be often the same for all group members and across all iterations. Certain indicators showed to be less expressive in the last iteration (e.g. C1, C2, C5, and C6) when the groups were mainly focused on bug fixing (e.g. S6). The DI is the correlation of a particular indicator to the total score of all other indicators of a student. Values above .3 are considered good, between .2 and .3 acceptable, between .1 and .2 marginal, and below .1 poor. Indicators that are not sufficiently different from others, i.e., have poor DI, should be removed from the model. Within groups and iterations, no indicators had a poor DI.

4.3 Discussion

In this section, we briefly present and discuss the findings, limitations, and shortcomings of our current approach. We will first address the measured parameters, then the

processing, and finally the use of analytics. For reasons of space, we will limit ourselves to a brief listing regarding the measured parameters. (1) Since we used GitLab as the main environment for collaboration, the team lacked an integrated communication option. Therefore, they communicated outside of GitLab so we could not track it. (2) Commits (e.g. P1, P6, and C4) often contained only small and rarely major changes. (3) In pair programming, usually only the driver made changes, while the observer or navigator did not show up in the commit log (e.g. C8). (4) There was a natural fluctuation in team performance over the course of a semester. (5) The previous point was also influenced by external factors, e.g. holidays, vacations, sick leave. (6) Other study-related tasks or exams and associated workloads were not considered. (7) It was assumed that all students have the same preposition (e.g., skills, experiences, personality [NTC20], health), i.e., related diversity aspects must be taken into account by the teacher. (8) The measured quantities were considered comparable atomic units, i.e. all commits were considered equal even if they differed in terms of number of changes, level of difficulty, and their relevance to the final program (e.g. P1, P6, and C4). (9) Students may have used code snippets or libraries from others (e.g., from the web) without acknowledging or referencing the sources. (10) Since students had to perform all roles at least once, role changes occurred after one iteration, which then led to different behaviours that were not yet considered in the validation of our approach. (11) Compared to many related works (e.g. [Gi20, TWM20]) we did not classify groups or individuals to provide a flexible tool for different learning and teaching scenarios. Consequently, the teacher has to decide which indicators are relevant for judging groups or learners at a given time.

In terms of the data processing, SonarQube (version 8.9 LTS) covers 27 languages, including CSS, JavaScript, and TypeScript as well as the corresponding analysis rules. Although this is a very good basis, false positives can still occur (e.g. S2, S4, S5, and S6). In this case, the rules used must be understood and adapted if necessary (i.e. by the teachers). Also, the coding conventions defined by the teams (e.g. in ESLint, JSLint) were not considered. For the planning activities, we did not consider the structure of the Kanban board (columns, movement of issues) and the degree of planning (e.g. defining an issue, assigning the issue, estimating the workload, setting a due date) as it could be derived from the data.

Besides, solutions are still being sought to avoid subjective assessments due to teachers' implicit bias against females and other underrepresented minorities in teams (e.g. [Be10]) in combination with teacher dashboards.

5 Conclusion and outlook

In this paper, we posed 4 research questions, identified, defined, and validated a total of 32 indicators for collaborative learning in software development teams. The resulting model describes collaborative interactions in programming teams considering 16 group

participation indicators (RQ1), 9 indicators for group cohesion (RQ2) and 7 indicators for code maintainability (RQ3). In addition, we presented a data processing pipeline for extracting, calculating, and visualizing these indicators on a teacher dashboard (RQ4). This approach enables teachers to keep track of complex group activities and individual contributions and then provide targeted formative feedback to the groups.

Next semester, we plan to use the tool as a monitoring instrument in a programming course. We hope that the tool will help teachers to provide better guidance and support regarding programming performance, participation in learning as well as group climate. So far, we focused on an iteration-based analysis. We plan to update the model and analysis to support three additional features: analysis of all previous iterations, relative comparisons between dedicated iterations either per person or per team, and individual trend analysis as well as enabling the teacher to dive into representative code segments.

Acknowledgments

This research was supported by the Research Cluster “Digitalization, Diversity and Lifelong Learning – Consequences for Higher Education” (D²L²) of the FernUniversität in Hagen, Germany.

Bibliography

- [AAM15] Agrawal, Kapil; Amreen, Sadika; Mockus, Audris: Commit Quality in Five High Performance Computing Projects. In: Proceedings of the 2015 International Workshop on Software Engineering for High Performance Computing in Science. SE4HPCS '15. IEEE Press, pp. 24–29, 2015.
- [Ah13] Ahmed, Faheem; Capretz, Luiz Fernando; Bouktif, Salah; Campbell, Piers: Soft skills and software development: A reflection from software industry. International Journal of Information Processing and Management, 4(3):171–191, 2013.
- [Ar20] Ardito, Luca; Coppola, Riccardo; Barbato, Luca; Verga, Diego: A Tool-Based Perspective on Software Code Maintainability Metrics: A Systematic Literature Review. August 2020. Publication Title: Scientific Programming Type: Review Article.
- [Be10] van den Bergh, Linda; Denessen, Eddie; Hornstra, Lisette; Voeten, Marinus; Holland, Rob W: The Implicit Prejudiced Attitudes of Teachers: Relations to Teacher Expectations and the Ethnic Achievement Gap. American Educational Research Journal, 47(2):497–527, jun 2010.
- [Bu20] Buffardi, Kevin: Assessing Individual Contributions to Software Engineering Projects with Git Logs and User Stories. In: Proceedings of the 51st ACM Technical Symposium on Computer Science Education. SIGCSE '20, ACM, New York, NY, USA, pp. 650–656, 2020.
- [Ca10] Calvani, Antonio; Fini, Antonio; Molino, Marcello; Ranieri, Maria: Visualizing and monitoring effective interactions in online collaborative groups. British Journal of Educational Technology, 41(2):213–226, March 2010.

-
- [Ca18] Campbell, G. Ann: Cognitive complexity: an overview and evaluation. In: Proceedings of the 2018 International Conference on Technical Debt. TechDebt '18, ACM, New York, NY, USA, pp. 57–58, May 2018.
- [CP13] Campbell, G. Ann; Papapetrou, Patroklos P.: SonarQube in Action. Manning Publications Co., USA, 1st edition, 2013.
- [ČS05] Čubranić, Davor; Storey, Margaret Anne D: Collaboration Support for Novice Team Programming. In: Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting Group Work. GROUP '05, ACM, New York, NY, USA, pp. 136–139, 2005.
- [De10] De Champlain, André F: A primer on classical test theory and item response theory for assessments in medical education. *Medical education*, 44(1):109–117, jan 2010.
- [EM02] Emden, E. van; Moonen, L.: Java quality assurance by detecting code smells. In: Ninth Working Conference on Reverse Engineering, 2002. Proceedings. pp. 97–106, November 2002.
- [Er20] Eraslan, Sukru; Rios, Julio César Cortés; Kopec-Harding, Kamilla; Embury, Suzanne M; Jay, Caroline; Page, Christopher; Haines, Robert: Errors and Poor Practices of Software Engineering Students in Using Git. In: Proceedings of the 4th Conference on Computing Education Practice 2020. CEP 2020, ACM, New York, NY, USA, 2020.
- [Fo98] Fowler, Martin: Refactoring: Improving the Design of Existing Code [Book]. 1998.
- [Gi20] Gitinabard, Niki; Okoilu, Ruth; Xu, Yiqao; Heckman, Sarah; Barnes, Tiffany; Lynch, Collin: , Student Teamwork on Programming Projects: What can GitHub logs show us?, 2020.
- [Gl19] Glassey, Richard: Adopting Git/Github within Teaching: A Survey of Tool Support. In: Proceedings of the ACM Conference on Global Computing Education. CompEd '19, ACM, New York, NY, USA, pp. 143–149, 2019.
- [Go10] Gour, Vishal; Sarangdevot, SS; Tanwar, Govind Singh; Sharma, Anand: Improve performance of extract, transform and load (ETL) in data warehouse. *International Journal on Computer Science and Engineering*, 2(3):786–789, 2010.
- [GSS19] Gote, Christoph; Scholtes, Ingo; Schweitzer, Frank: Git2net: Mining Time-Stamped Co-Editing Networks from Large Git Repositories. In: Proceedings of the 16th International Conference on Mining Software Repositories. MSR '19. IEEE Press, pp. 433–444, 2019.
- [HG01] Hoegl, Martin; Gemunden, Hans Georg: Teamwork Quality and the Success of Innovative Projects: A Theoretical Concept and Empirical Evidence. *Organization Science*, 12(4):435–449, August 2001.
- [MGM19] Menezes, Júlio; Gusmão, Cristine; Moura, Hermano: Risk factors in software development projects: a systematic literature review. *Software Quality Journal*, 27(3):1149–1174, 2019.
- [MM19] Martinez, Matias; Monperrus, Martin: Coming: A Tool for Mining Change Pattern Instances from Git Commits. In: Proceedings of the 41st International Conference on Software Engineering: Companion Proceedings. ICSE '19. IEEE Press, pp. 79–82, 2019.

- [NTC20] Nunes, Ingrid; Treude, Christoph; Calefato, Fabio: The Impact of Dynamics of Collaborative Software Engineering on Introverts: A Study Protocol. In: Proceedings of the 17th International Conference on Mining Software Repositories. MSR '20, ACM, New York, NY, USA, pp. 619–622, 2020.
- [Pl21] Pluralsight: Flow metrics Pluralsight Help Center. 2021.
- [RGS15] Rosen, Christoffer; Grawi, Ben; Shihab, Emad: Commit Guru: Analytics and Risk Prediction of Software Commits. In: Proceedings of the 2015 10th Joint Meeting on Foundations of Software Engineering. ESEC/FSE 2015, ACM, New York, NY, USA, pp. 966–969, 2015.
- [Ri19] Rios, Julio César Cortés; Kopec-Harding, Kamilla; Eraslan, Sukru; Page, Christopher; Haines, Robert; Jay, Caroline; Embury, Suzanne M: A Methodology for Using GitLab for Software Engineering Learning Analytics. In: Proceedings of the 12th International Workshop on Cooperative and Human Aspects of Software Engineering. CHASE '19. IEEE Press, pp. 3–6, 2019.
- [RTLr19] Rosli, Marshima; Tempero, Ewan; Luxton-reilly, Andrew: A Systematic Mapping Study on Data Quality in Software Engineering Data Sets. *Journal of Universal Computer Science*, 25(1):16–41, 2019.
- [SA20] Sandee, Jan Jaap; Aivaloglou, Efthimia: Git Canary: A Tool for Analyzing Student Contributions in Group Programming Assignments. In: Koli Calling '20: Proceedings of the 20th Koli Calling International Conference on Computing Education Research. Koli Calling '20, ACM, New York, NY, USA, 2020.
- [Se19] Sensaoui, Abderrahmane; Aktouf, Oum-El-Kheir; Hely, David; Di Vito, Stephane: An In-depth Study of MPU-Based Isolation Techniques. *Journal of Hardware and Systems Security*, 3(4):365–381, December 2019.
- [SMG20] Silva, Leonardo; Mendes, Antonio Jose; Gomes, Anabela: Computer-supported collaborative learning in programming education: A systematic literature review. *IEEE Global Engineering Education Conference, EDUCON, 2020-April(May)*:1086–1095, 2020.
- [So21] SonarQube: , Code Quality and Code Security, 2021.
- [Ta19] Tamburri, Damian A.; Palomba, Fabio; Serebrenik, Alexander; Zaidman, Andy: Discovering community patterns in open-source: a systematic approach and its evaluation. *Empirical Software Engineering*, 24(3):1369–1417, June 2019.
- [Ts18] Tsantalis, Nikolaos; Mansouri, Matin; Eshkevari, Laleh M; Mazinanian, Davood; Dig, Danny: Accurate and Efficient Refactoring Detection in Commit History. In: Proceedings of the 40th International Conference on Software Engineering. ICSE '18, ACM, New York, NY, USA, pp. 483–494, 2018.
- [TWM20] Tushev, Miroslav; Williams, Grant; Mahmoud, Anas: Using GitHub in large software engineering classes. An exploratory case study. *Computer Science Education*, 30(2):155–186, apr 2020.
- [We20] Wessel, Mairiel: Leveraging Software Bots to Enhance Developers' Collaboration in Online Programming Communities. In: Conference Companion Publication of the 2020 on Computer Supported Cooperative Work and Social Computing. CSCW '20 Companion, ACM, New York, NY, USA, pp. 183–188, 2020.

Using GitHub data to analyse student's teamwork in a programming course to prevent discrimination.

Maximilian Karl¹ and Niels Pinkwart²

Abstract: Programming in a team is a basic and essential skill and it is important to train teamwork in higher education with software development project courses. At the end of a course the tutor must grade every student individually based on the team performance. To prevent any discrimination towards students a classification was developed to measure teamwork. Such a classification gives the tutor a fast feedback about the teamwork if the team worked collaborative, cooperative or did not work as a team and how much each person contributed to the project.

Keywords: Learning Analytics, GitHub, Teamwork, Discrimination

1 Introduction

Teamwork is essential for professional programmer [Ik15] and is taught for software development courses in education. In some programming courses students have to work as a team to solve mutiple tasks or to create their own projects. The tutor can give in such a programming course grades based on the teamwork. It is challenging and takes a lot of effort for a tutor to grade each team member individually [SJ20]. Discrimination based on a personal subjective view of an instructor could have an effect on how the instructor interprets the teamwork of a group [Wi06]. Learning Analytics can improve teamwork assessment [Fi15] to analyse data objectively in a short time. A version control system like GitHub which is based on git provides objective data. When a team uses GitHub almost every user interaction is logged. The git commits and the GitHub issues and pull requests provide the most information of what every person has contributed to a project. A commit describes the action whenever a person saves local changes to git. One commit can cover multiple files which were edited, created or deleted by one person. Issues and pull requests can be used to exchange information and knowledge in a project. An issue can be used to describe a task like a new feature or an error produced by the code. Pull requests are useful to discuss the changed code to solve one or more issues. A code review can be requested in a pull request in order to let a different person review the new code. This helps to ensure the quality of the code and to reduce coding bugs. A classification to analyse GitHub data about the teamwork in a project was designed.

¹ Humboldt University Berlin, Mathematisch-Naturwissenschaftliche Fakultät, maximilian.karl@hu-berlin.de,

² Humboldt University Berlin, Mathematisch-Naturwissenschaftliche Fakultät, pinkwart@hu-berlin.de

2 Classification Method

A classification method was designed to classify whether a team worked together or not. A python package called `github2pandas` was developed to acquire the GitHub data which is uploaded to pypi.org [G2P21].

The classification is based on a study from 2020 where only on the commits of a GitHub project are analysed [Gi20]. The commit data about what every person uploaded can be used to classify the teamwork into collaborative, cooperative and solo work [Gi20]. Our classification extends this study by using commits and additionally using issues and pull requests from GitHub to evaluate the teamwork of a student group. The communication between team members can be analysed from issues and pull requests which is an advantage because commits only show what one person has done. The quantity and quality of commits, issues and pull requests are very different on every GitHub project data set, therefore the classification first classifies commits, issues and pull requests by itself and then compares them equally.

Every commit of a project has an impact on the classification because every commit includes changed files with edited lines of code or documentation. One edit of a commit is compared with all other edits from all commits. This implies that one commit will be split into multiple commits for every modified file and takes the edited lines as weighting. A commit is collaborative when the changed file is also be modified by a different person in a different commit. A commit is classified as solo work when only one person changed this file. Every issue and pull request is analysed by the number of contributors. If more than two people are contributing, it is marked as collaborative otherwise it is marked as solo work.

The teamwork of the GitHub repository is calculated for collaboration and solo work by the mean value of commits, issues and pull requests. The cooperation in the team can be calculated for the project with the lowest mean of solo work percentage by one person multiplied by the number of team members to get the cooperation percentage. The solo work percentage gets decreased by the cooperation. The teamwork is finally classified based on which teamwork classifier has the highest value. If the sum of collaboration and cooperation is higher than the solo work, the project will be classified as collaboration or cooperation depending on which has the higher percentage.

3 Results

One semester of a bachelor software development course generated 65 useful student projects with two team members for the classification. The total number of projects is 99 but unfortunately 32 projects do not have any GitHub data. Either the students did not do anything or they worked completely offline. There are also two projects without any commits but with issues. This can imply that the students programmed offline or never started programming because they did not understand the task. The results of the

classification are 10 collaborative projects, 7 cooperative projects and 48 solo work projects. The classification revealed that the most projects were solo work projects where only one team member did everything or most of the work.

An example of a solo work project is shown in table 1. The results show that student A did most of the work in pull requests, issues and commits, but when student B was working on the project then the students worked collaboratively. In all solo work projects the cooperation and collaboration in pull requests and issues is very low which shows that the teams did not communicate enough. A low communication indicates that one student will commit more alone.

Team member	Pull requests	Issues	Commits	Mean	Total
Collaboration	0%	33%	25%	19.3%	19.3%
Cooperation	-	-	-	-	0%
Solo Work (Student A)	100%	67%	75%	80.7%	80.7%
Solo Work (Student B)	0%	0%	0%	0%	0%

Tab. 1: Classification of project No. 1

In one student project the case occurred that the sum of collaboration and cooperation is higher than the solo work, but the solo work has the highest value. The results of the project are shown in table 2. The students split the task almost in similar parts and committed both about the same amount of code and documentation. The high solo percentage is caused by the issues because student A created only one issue and asked the tutor something. Student A closed the issue after the feedback of the tutor. The team did not use pull requests and issues because they probably communicated outside of GitHub.

Team member	Pull requests	Issues	Commits	Mean	Total
Collaboration	50%	0%	2.6%	17.5%	17.5%
Cooperation	-	-	-	-	38.6%
Solo Work (Student A)	50%	100%	39.5%	63.2%	43.9%
Solo Work (Student B)	0%	0%	57.9%	19.3%	0%

Tab. 2: Classification of project No. 2

4 Limitations

The analysis on GitHub data is limited on the logged data from GitHub. Whenever students work as a team or alone completely offline then the classification cannot detect anything. When students did pair programming then it could be classified as a solo work project because only one person will upload the changed code at the end. In this case the

students probably did not use pull requests and issues in GitHub because they already exchanged information and knowledge while pair programming. This limitation can be avoided in a future study when the tutor defines a rule whenever a team is pair programming then they should write pair programming in the commit message.

5 Conclusion

The classification detects teamwork from a GitHub project which can be used for feedback to the students and tutors. The classification feedback is very objective and represents what the team and each student has done in the project. This feedback can be used to prevent any discrimination and can help the tutor to grade each student more objectively. The tutor would save a lot of time to look deeply into every project to check how much every student participated in the project. A feedback system based on the teamwork classification will be designed for students and tutors and used in a next study. The classification feedback can also be used to identify free-riding early in a task or project. A free-riding behaviour is when only one student produces most of the group work [Sw12] which applies for 48 of 65 evaluated student projects.

Bibliography

- [Fi15] Fidalgo-Blanco, Á.; Sein-Echaluce, M. L.; García-Peña, F. J.; Conde, M. Á.: Using Learning Analytics to improve teamwork assessment. *Computers in Human Behavior* 47/, S. 149–156, 2015.
- [G2P21] Maximilian Karl & Sebastian Zug, 2021, url: <https://pypi.org/project/github2pandas/>, Stand: 23. 06. 2021.
- [Gi20] Gitinabard, N.; Okoilu, R.; Xu, Y.; Heckman, S.; Barnes, T.; Lynch, C.: Student Teamwork on Programming Projects: What can GitHub logs show us? arXiv preprint arXiv:2008.11262/, 2020.
- [Ik15] Ikonen, J.; Knutas, A.; Wu, Y.; Agudo, I.: Is the world ready or do we need more tools for programming related teamwork? In: Proceedings of the 15th Koli Calling Conference on Computing Education Research. S. 33–39, 2015.
- [Wi06] Willcoxson, L. E.: “It’s not fair!”: assessing the dynamics and resourcing of teamwork. *Journal of Management Education* 30/6, S. 798–808, 2006.
- [SJ20] Shishavan, H. B.; Jalili, M.: Responding to student feedback: Individualising teamwork scores based on peer assessment. *International Journal of Educational Research Open* 1/, S.100019,2020, issn:2666-3740.
- [Sw12] Swaray, R.: An evaluation of a group project designed to reduce free-riding and promote active learning. *Assessment & Evaluation in Higher Education* 37/3, S. 285–292, 2012.

Digitalisierung und Internationalisierung in der Hochschulbildung

Areej Aldaghamin¹, Andreas Becker², Regina Brautlacht³, Samira Herb-Cless⁴, Natalie Kiesler⁵, Alexander Knoth⁶, Katrin Löhr⁷, Ekaterina Mikhaylova⁸, Wolfgang Radenbach⁹, Christian Reimann¹⁰, Anne Sennhenn¹¹, Judith Venherm¹², Nina Wagenknecht¹³, Carsten Wolff¹⁴

Abstrakt: Die digitale Transformation verändert die internationale Kooperation der Hochschulen massiv. Über die Möglichkeiten der virtuellen Mobilität hinaus entstehen neue Themenfelder, die internationale Lern- und Lehrerlebnisse mit digitaler Unterstützung verändern, ergänzen oder neu ermöglichen. Dazu sind im Bereich der Förderung der Internationalisierung (DAAD, Erasmus+, BMBF u.a.) Projekte und Förderformate entstanden, die Digitalisierung und Internationalisierung kombinieren und die neuen Themenstellungen adressieren, z.B. didaktische Formate, administrative Prozesse (auch im Kontext OZG und DSGVO), virtuelle und hybride Mobilität, internationale Projekt- und Teamformate sowie schlussendlich auch Inhalte, die internationale, interkulturelle und interdisziplinäre Kompetenzen mit digitalen Kompetenzen verbinden. Der vorgeschlagene Workshop soll entsprechende Projekte zusammenbringen und die Themen strukturieren, um einen Überblick der Entwicklungen zu schaffen und somit einen Beitrag zur Definition des Themenfelds „Digitalisierung & Internationalisierung“ zu leisten.

Keywords: Digitalisierung & Internationalisierung, virtuelle und hybride Mobilität, digitale internationale Kooperation.

¹ Fachhochschule Dortmund, IDiAL, Otto-Hahn-Str. 23, 44227 Dortmund, areej.aldaghamin@fh-dortmund.de

² Fachhochschule Dortmund, Ruhr Master School, Fachbereich Informationstechnik, Sonnenstr. 96, 44139 Dortmund, andreas.becker@fh-dortmund.de

³ Hochschule Bonn-Rhein-Sieg, Grantham-Allee 20, 53757 Sankt Augustin, regina.brautlacht@h-brs.de

⁴ Deutscher Akademischer Austauschdienst (DAAD), P44 – Internationalisierung digital, Fachhochschulen/HAW, Kennedyallee 50, 53175 Bonn, herb@daad.de

⁵ Justus-Liebig-Universität Gießen, Hochschulrechenzentrum, Gruppe Medien und E-Learning, Heinrich-Buff-Ring 44, 35392 Gießen, natalie.kiesler@hrz.uni-giessen.de

⁶ Deutscher Akademischer Austauschdienst (DAAD), Chief Digital Officer (CDO), Markgrafenstr. 37, 10117 Berlin, knoth@daad.de

⁷ Fachhochschule Dortmund, Rektorat, Sonnenstr. 96, 44139 Dortmund, katrin.loehr@fh-dortmund.de

⁸ Fachhochschule Dortmund, IDiAL, Otto-Hahn-Str. 23, 44227 Dortmund, ekaterina.mikhaylova@fh-dortmund.de

⁹ Georg-August-Universität Göttingen, Abteilung Studium und Lehre, Platz d. Göttinger Sieben 5, 37073 Göttingen, wolfgang.radenbach@uni-goettingen.de

¹⁰ Fachhochschule Dortmund, IDiAL, Otto-Hahn-Str. 23, 44227 Dortmund, christian.reimann@fh-dortmund.de

¹¹ Georg-August-Universität Göttingen, Abteilung Göttingen International, Von-Siebold-Straße 2, 37075 Göttingen, anne.sennhenn@zvw.uni-goettingen.de

¹² Deutscher Akademischer Austauschdienst (DAAD), P44 – Internationalisierung digital, Fachhochschulen/HAW, Kennedyallee 50, 53175 Bonn, venherm@daad.de

¹³ Georg-August-Universität Göttingen, Abteilung Studium und Lehre, Platz d. Göttinger Sieben 5, 37073 Göttingen, nina.wagenknecht@uni-goettingen.de

¹⁴ Fachhochschule Dortmund, IDiAL, Otto-Hahn-Str. 23, 44227 Dortmund, carsten.wolff@fh-dortmund.de

1 Einleitung

Das vorliegende Positionspapier wurde von den Autorinnen und Autoren im Rahmen eines Workshops am 13.09.2021 im Zuge der 19. Fachtagung Bildungstechnologien der GI Fachgruppe Bildungstechnologien (DELFI 2021) in Dortmund erarbeitet. Der Workshop fand virtuell statt und wurde mittels Miro-Board dokumentiert (siehe Anhang).

Der Workshop wurde von einer Reihe von Fachvorträgen eingeleitet, mit dem Ziel, einen Überblick über die laufenden Projekte und Programme zu schaffen und die nachfolgende Diskussion zu stimulieren. Die Agenda des Workshops war wie folgt:

- 9:00-9:20: Begrüßung und Einführung in die Thematik, Katrin Löhr, Carsten Wolff
- 9:20-9:40: Keynote zum Thema "Digitalisierung & Internationalisierung", Alexander Knoth
- 9:40-10:00: Die DAAD-Programme IMKD, IP Digital und IVAC [IMK20, IPD20, IVA20]: Samira Herb-Cless, Judith Venherm
- 10:00-10:20: Das Projekt "Multidisciplinary Virtual Collaboration: Developing a Didactical Approach" [COV20, FJH20], Regina Brautlacht
- 10:20-10:30 Kaffeepause
- 10:30-10:50: Vorstellung des Student Journey Configurator [MAE20], Ekaterina Mikhaylova, Areej Aldaghamin, Christian Reimann
- 10:50-11:10: Das BMBF-Projekt PIM [PIM21], Nina Wagenknecht, Wolfgang Radenbach

An die Vorträge schloss sich eine ca. einstündige Diskussion der Teilnehmenden an.

2 Themen und Forschungsfragen

Ein Kernproblem für die Strukturierung und Definition des Themenfelds „Digitalisierung & Internationalisierung“ liegt darin, dass zwei sehr umfangreiche und schon seit vielen Jahren bearbeitete, aber zum Teil disparate Themenfelder der Hochschulbildung miteinander verbunden werden: die Digitalisierung und die Internationalisierung der Hochschulen.

In der „Digitalisierung der Hochschulen“ werden z.B. die folgenden Themen behandelt:

- Digitalisierung der Lehre durch mediendidaktische Ansätze, E-Learning etc.
- Digital Literacy als Kompetenzziel und Learning Outcome
- Die Digitalisierung der Geschäftsprozesse einer Hochschule

- Hochschul-IT, Campus Management Systeme

In der „Internationalisierung der Hochschulen“ geht es u.a. um:

- Sprachenpolitik und Sprachlehre
- Internationalisierung der Curricula
- Mobilität, internationale Kooperation und Netzwerke
- Interkulturelle und internationale Kompetenzen

Das *Themenfeld „Digitalisierung & Internationalisierung“* soll nicht etwa diese Themen neu bearbeiten oder ggf. wiederholen. Stattdessen geht es darum, die Schnittmenge der beiden Themen „Digitalisierung der Hochschulen“ und „Internationalisierung der Hochschulen“ neu zu betrachten und herauszuarbeiten:

- Was ist in der „Digitalisierung der Hochschulen“ neu, anders oder speziell zu berücksichtigen, wenn es aus dem Blickwinkel der Internationalisierung betrachtet wird? Was ist in der Digitalisierung anders, wenn sie international bzw. grenzüberschreitend betrieben wird? Wo sind digitale Kompetenzen anders ausgeprägt, wenn sie auch international funktionieren sollen?
- Was ist im Gegenzug in der „Internationalisierung der Hochschulen“ neu oder anders geworden durch die Digitalisierung? Wie verändern sich z.B. internationale und interkulturelle Kompetenzen in digitalen Kontexten?
- Welche Themen entstehen neu durch die Verbindung von Digitalisierung und Internationalisierung? Und wie begegnen den die Hochschulen, z.B. in der Europäischen Union (EU)?

Ein wesentliches Ziel des Workshops ist es, das Forschungsfeld „Digitalisierung und Internationalisierung in der Hochschulbildung“ zu strukturieren und im Austausch mit den Teilnehmenden den Status Quo sowie die künftigen Entwicklungen zu erarbeiten. Als Ergebnis soll ein *Überblick der wissenschaftlichen Forschungsfragen* entstehen, z.B.:

- Inwieweit hängen Digitalisierung und Internationalisierung in der Hochschulbildung zusammen? Wie bedingen und befördern sie sich gegenseitig?
- Welche Themengebiete sind im Forschungsfeld relevant?
- Welche Fragen werden bearbeitet, welche sind offen?
- Wie lassen sich die Fragen strukturieren und priorisieren?
- Wo liegen wesentliche Problemstellungen? Wo sind Zusammenhänge ungeklärt? Wo liegt empirischer oder auch experimenteller Forschungsbedarf?
- Wo ergeben sich Kooperations- und Kombinationsmöglichkeiten?

- Welche Formate sind relevant, sowohl bzgl. der Projekte als auch bzgl. der Didaktik, der Technologie und der Lernszenarien?

Internationalität als Teil der Global Competences [Raj12] sowie die digitalen Kompetenzen als Teil der Professional Competences bilden zusammen mit den fachlichen Kompetenzen, weiteren Soft Skills sowie der Befähigung zur internationalen, interkulturellen und interdisziplinären Zusammenarbeit ($3 \times „i“$) einen Zusammenhang, der vorzugsweise nicht isoliert in einzelnen Lehrelementen, sondern stattdessen in holistischen Lernerlebnissen vermittelt wird. Diese Lehr- und Lernsituationen bilden z.T. realistische Situationen aus einem späteren Berufsfeld ab, so dass besonders authentische Lernerlebnisse entstehen, z.B. aus der internationalen Projektarbeit in virtualisierten Arbeitsumgebungen. Digitalisierung und Internationalisierung gehen also in vielen Berufsfeldern Hand in Hand, so dass die Verbindung beider Bereiche in einem gemeinsamen Ansatz in der Hochschulbildung naheliegend und notwendig ist, z.B. adressiert in den Handlungsempfehlungen der HRK zur Internationalisierung der Curricula [HRK17]. Entsprechend werden – nicht nur seit Ausbruch der Corona Pandemie – die Themenfelder Digitalisierung und Internationalisierung auch von der EU in der Novelle des Digital Education Action Plan [EU20] als Priorität für die Entwicklung des europäischen Hochschulraums definiert und mit entsprechenden Initiativen miteinander integriert. Der DAAD hat zusammen mit dem Deutschen Institut für Entwicklungspolitik (DIE) die Digitalisierung als wesentlichen „Enabler“ für eine nachhaltigere Internationalisierung und Entwicklungszusammenarbeit definiert [DD18]. In einer umfassenden Betrachtung wird die Weiterentwicklung von Bologna in einem digitalen, virtuellen europäischen Hochschulraum als Basis der Strategien zur Internationalisierung der Hochschulbildung elaboriert [ROK19]. Prinzipiell sind dabei drei Dimensionen der Entwicklung relevant:

- Die *digitale Unterstützung der Kooperation* und der wissenschaftlichen Zusammenarbeit zur Schaffung grenzüberschreitender Verbünde und Projekte.
- Der Einsatz *digitaler Werkzeuge und virtualisierter Methodik in der Didaktik* und zur Vermittlung von digitalen und internationalen Lehr- und Lernerlebnissen.
- Die Vermittlung *digitaler und internationaler Kompetenzen* in einem umfassenden und ganzheitlichen Lernerlebnis.

Dabei kann es sowohl um entsprechende Fachkompetenzen gehen als auch um inhärente Lernerlebnisse und damit verbundene übergreifende, transversale Kompetenzen, z.B. definiert in den Overarching Learning Outcomes (OLO) des European Institute of Innovation and Technology [EIT17].

3 Ergebnisse des Workshops

Im Workshop wurde zunächst mit Hilfe eines Miro-Boards durch die Teilnehmerinnen und Teilnehmer eine Sammlung von Stichpunkten und Themen vorgenommen, die unter

dem Eindruck der Vorträge und der nachfolgenden Diskussionen stattfand. Danach wurden die Beiträge sortiert und es wurden Themencluster gebildet, zu denen Oberbegriffe gesucht wurden. Das Ergebnis ist im Anhang dargestellt.

Als wesentliche *Themen-Cluster* wurden identifiziert:

- *Virtuelle Mobilität als neues Internationalisierungswerkzeug:* Die Ergänzung der weiterhin für die Zielsetzungen der Internationalisierung sehr wichtigen physischen Mobilität (mit ihren zwischenmenschlichen Aspekten und Erlebnissen) durch Formate der virtuellen Mobilität wird als ein zentrales Themenfeld betrachtet, das in vielerlei Hinsicht ausgearbeitet werden muss:
 - Virtuelle Mobilität verbindet die Internationalisierung durch Studierenden- und Lehrendenaustausch mit der internationalization@home, so dass ein neues Arbeitsfeld entsteht, das unterschiedliche Organisations- und Kompetenzbereiche vereint.
 - Virtuelle Mobilität ermöglicht Gruppen internationale Lernerlebnisse, die bisher z.B. aus familiären oder finanziellen Gründen diesen Zugang nicht hatten oder nicht niederschwellig wahrnehmen konnten.
 - Es entstehen niederschwellige Angebote bzw. Einstiegsangebote in die Internationalität und den internationalen Austausch.
 - Offen sind Probleme der Anerkennung, der (finanziellen) Förderung, der vermittelbaren Kompetenzen, der Kompetenz bei Lehrenden und in der Organisation sowie des Zugangs zu entsprechender Technologie.
- *Virtuelle internationale Zusammenarbeit:* Die Digitalisierung eröffnet für Studierende (z.B. in cross border projects), Lehrende (z.B. in der Entwicklung gemeinsamer Module) und die Hochschulen als Organisation (z.B. in Projekttreffen) neue Möglichkeiten für eine niederschwellige und ggf. auch häufigere und intensivere Zusammenarbeit. In diesem Themenfeld wurden folgende Aspekte besonders diskutiert:
 - Die Herausforderungen und Best Practices der internationalen, virtuellen Teamarbeit müssen von den Hochschulen adressiert werden.
 - Durch unterschiedliche Zugänge zu digitalen Infrastrukturen (z.B. Internetbandbreite) können Partnerhochschulen benachteiligt werden im Sinne eines „digital divide“, so dass Kompensationsmaßnahmen geschaffen werden müssen.
 - Die Aspekte der virtuellen internationalen Zusammenarbeit in anderen Organisationen (z.B. internationalen Unternehmen) sind in vielerlei Hinsicht auch für die Hochschulen relevant, z.B. bzgl. der den Studierenden zu vermittelnden Kompetenzen, der Best Practices und der verwendeten Technologien und Plattformen.

- *Digitale und internationale Kompetenzen:* Eine Kernfrage betrifft die Kompetenzen, die durch die Verbindung der Themenfelder Digitalisierung und Internationalisierung erforderlich bzw. neu geschaffen werden, z.B. in Bezug auf:
 - Die Kompetenzen, die den Studierenden durch die Hochschulen vermittelt werden müssen, um z.B. digitale Zusammenarbeitsformen auch international einsetzen zu können.
 - Die Fragestellung, inwiefern sich interkulturelle und internationale Kompetenzen in einer zunehmend digitalen Welt verändern und inwieweit diese Kompetenzen digital und virtuell vermittelt werden können. Dabei ist der Wert bzw. die Qualität digitaler Lern- und Lehrerlebnisse für den Erwerb internationaler und interkultureller Kompetenzen zu betrachten.
 - Die Kompetenzen, die Lehrende und Hochschulangehörige sowohl im internationalen Bereich als auch im digitalen Bereich aufbauen müssen, um in ihren jeweiligen Arbeitsfeldern oder auch im Bereich der Schnittmenge aus Digitalisierung und Internationalisierung arbeiten zu können.
- *Digitale Technologien im internationalen Kontext:* Die Digitalisierung der Hochschulen erfolgt nur teilweise organisationsübergreifend und noch weniger länderübergreifend, da die digitalen Infrastrukturen aus vielfältigen Gründen vor allem auf eine einzelne Hochschule bezogen gedacht werden. Die internationale Interoperabilität digitaler Technologien und Infrastrukturen ist daher ein relevantes Themenfeld. Hier wurden vor allem folgende Aspekte diskutiert:
 - Cross-border identity management, d.h. der Zugang zu digitalen Infrastrukturen an internationalen Partnerhochschulen mit den Zugangsdaten der eigenen Hochschule.
 - Datenschutz in z.T. unterschiedlichen rechtlichen Systemen.
 - Organisations- bzw. grenzüberschreitende IT workflows (z.B. Erasmus without Paper, EWP).
- *Digitale Lehre und Didaktik im internationalen Kontext:* Dieses Themenfeld wird als sehr umfangreich eingeschätzt mit erheblicher Überlappung mit den Entwicklungen und Erfahrungen im Bereich der digitalen Lehre und Didaktik generell. Besondere Fragestellungen für das Themenfeld „Digitalisierung und Internationalisierung“ wurden in folgenden Bereichen identifiziert:
 - Hybride Lehr- und Lernsettings mit Studierenden bzw. Lehrenden vor Ort in der Kombination mit virtuell teilnehmenden Personen aus anderen Ländern, u.a. auch in hybriden cross border projects.
 - Die Qualität internationaler Lern- und Lehrerlebnisse sowie der Kompetenzerwerb in online-Formaten bzw. in hybriden Formaten.

- Das digitale Prüfen und Bewerten von Kompetenzen in internationalen Lehr- und Lern-Settings, u.a. auch bzgl. unterschiedlicher rechtlicher Rahmenbedingungen.
- *Organisationsfragen im Kontext der Digitalisierung & Internationalisierung:* In diesem Themenfeld wurden eine ganze Reihe von Fragestellungen der Digitalisierung besprochen, die sich in einem internationalen Kontext anders darstellen als auf eine einzelne Hochschule bezogen:
 - Anerkennung von Lehrleistungen: digital bzw. an einer Partnerhochschule.
 - Zeitliche Planung (Scheduling) und curriculare Einbindung gemeinsamer (geteilter) digitaler Lehrmodule.
 - Zeitliche Asynchronität aufgrund verschiedener Zeitzonen.
 - Organisationsübergreifendes digitales Prüfen von Kompetenzen.
 - Zugang zu neuen virtuellen Angeboten, im Zusammenhang mit rechtlichen Fragen rund um Studierendenstatus.
- *Qualität und Evaluation:* Die Qualitätssicherung und die Evaluation internationaler, digitaler Lehr- und Lernsettings ist eine besondere Herausforderung, vor allem auch bzgl. virtueller Betreuungskonzepte für internationale Studierende und Lehrende [KDF21].
- *Übergang der akademischen Student Journey hin zum Life-long Learning:* Durch das „Hinauswachsen“ der Student Journey aus der einzelnen Hochschule heraus und hin zu einem Life-long Learning ergeben sich für die langfristige internationale Zusammenarbeit und Netzwerkbildung weitere Chancen und Herausforderungen.

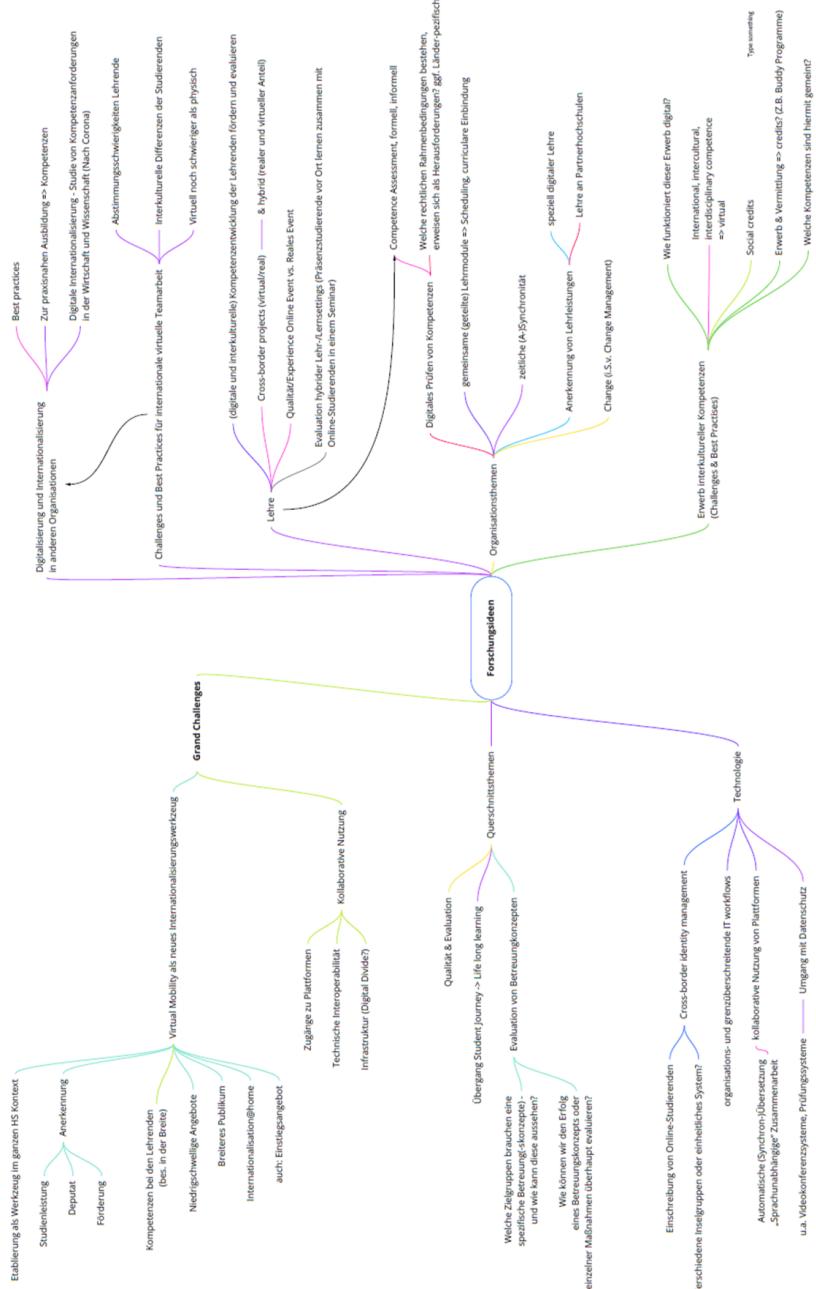
Die Themencluster sind im Zuge des Workshops nur in Ansätzen diskutiert worden, so dass eine Validierung bzw. Reflektion der Strukturierung nicht möglich war. Auch eine Priorisierung und Bewertung konnte nur in ersten Ansätzen diskutiert werden. Insofern ist die Zusammenstellung aus Sicht der Teilnehmenden nur ein Diskussionsbeitrag, der bei weiteren Gelegenheiten reflektiert und erweitert werden muss.

Literaturverzeichnis

- [COV20] Consortium of Virtual Exchange (COVE): <https://www.cove.education/>, letzter Zugriff 19.09.2021
- [DD18] DAAD & DIE, Digital Transformation: Higher Education and Research for Sustainable Development, https://static.daad.de/media/daad_de/der-daad/was-wir-tun/2018-12_daad_die_expertenpapier_digitaltransformation.pdf, 2018.
- [EIT17] European Institute of Innovation and Technology (EIT): “Quality for learning” EIT Quality Assurance and Learning Enhancement Model, 2017

- [EU20] EU: Digital Education Action Plan 2021-2027, 2020.
- [FJH20] Fonseca, P., Julian, K., Hulme, W., Martin, L., Brautlacht, R. (2020). The Multi-disciplinary Approach to an Interdisciplinary Virtual Exchange. IVEC 2020 – Book of Abstracts. P.68. http://iveconference.org/wp-content/uploads/2020/08/IVEC-2020-Book-of-Abstracts_19-August-2020-1.pdf
- [HRK17] HRK Hochschulrektorenkonferenz (2017). Zur Internationalisierung der Curricula. Empfehlung der 22. Mitgliederversammlung der HRK am 9. Mai 2017 in Bielefeld
- [IMK20] DAAD Programm „Internationale Mobilität und Kooperation Digital (IMKD)“: <https://www.daad.de/de/infos-services-fuer-hochschulen/weiterfuehrende-infos-zu-daad-foerderprogrammen/imkd/>, letzter Zugriff 19.09.2021
- [IVAC20] DAAD Programm „International Virtual Academic Collaboration (IVAC)“: <https://www.daad.de/de/infos-services-fuer-hochschulen/weiterfuehrende-infos-zu-daad-foerderprogrammen/ivac/>, letzter Zugriff 19.09.2021
- [IPD20] DAAD Programm „Internationale Programme Digital (IP Digital)“: <https://www.daad.de/de/infos-services-fuer-hochschulen/weiterfuehrende-infos-zu-daad-foerderprogrammen/ip-digital/>, letzter Zugriff 19.09.2021
- [KDF21] Kiesler, N., Dresing, K., Faber, S., Lizarazo, L., Nierste, W., & Volz, J. (2021). Internationale digitale Lehre unter Pandemie-Bedingungen. <http://dx.doi.org/10.22029/jlupub-71>
- [MAE20] Mikhaylova, E.; Aldaghamin, A.; Ebberg, F.; Tokanov, O.; Wolff, C.; Reimann, C.: Digital Education Ecosystem (DEE): User-Centred Design of the Student Journey Configurator, 2021 IEEE International Conference on Smart Information Systems and Technologies (2021 IEEE SIST), Nursultan, Kazakhstan, 2021
- [PIM20] PIM – Plattform für Internationale Studierendenmobilität, <https://pim-plattform.de/>, letzter Zugriff 19.09.2021
- [Raj12] Rajala, S.A.: Beyond 2020: Preparing Engineers for the Future. Proceedings of the IEEE, Vol. 100, May 13th, 2012, pp. 1376-1383, DOI 10.1109/JPROC.2012.2190169
- [ROK19] Rampelt, F.; Orr, D.; Knoth, A.: Bologna Digital 2020: White Paper on Digitalisation in the European Higher Education Area, Hochschulforum Digitalisierung, 2019.

Anhang: Im Zuge des Workshops erstelltes Miro-Board



A User-Model for a Next Generation Learning Management System

Gerd Kortemeyer  ¹

Abstract: Traditionally, learner data is stored and maintained locally within educational institutions, and there frequently within disjunct systems. Learners moving through different educational experiences at a number of different institutions are leaving behind a scattered breadcrumb trail of data, which is neither benefitting the learner when moving on to the next experience nor necessarily in compliance with data protection and security goals. The workshop investigates an alternative model, where learners are managing their own educational data in a self-sovereign fashion, while at the same time, the integrity of these data is guaranteed.

Keywords: Learning Management System, Self-Sovereignty, Social Linked Data, SOLID, Self-Sovereign Identity, SSI, Blockchain, Gaia-X.

1 Introduction

Currently, learner data are highly distributed and duplicated: academic records are traditionally held by the respective institutions that the learner attended. Within those institutions, some data are in institution-wide databases, while other data are stored in particular learning platforms like course management systems, and frequently associated with particular courses.

Leaners usually receive paper copies of their certified transcripts (showing earned course credits) and degrees, and they need to make photocopies or scans of those documents to transmit their credentials to other institutions or potential employers. Not surprisingly, fraud abounds, particularly in an increasingly global education and work space.

Learners leave behind a breadcrumb trail of educational data across platforms and institutions, over which they have very little control, and often they cannot even access it – the user lacks sovereignty. An alternative view would be to see the learner's educational experience as one contiguous, lifelong journey, and that data gathered along the way gets added to one continuous transcript ("ledger") of achievements, credits, certifications, and degrees.

¹ ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland, gerd.kortemeyer@let.ethz.ch,

 <https://orcid.org/0000-0001-6643-9428>

Also, along the way, learning analytics data builds up, which could enable learners to take better advantage of each station along their educational journey. While a perfect human tutor or mentor would accompany a learner over an extended amount of time and get to know him or her, the status quo is that any kind of AI-mentor would be rather scatter-brained and frequently suffering from memory-loss.

The result of the status quo is a lack of verifiability, coherence, and personal, user-centric data sovereignty (“data self-sovereignty”) of educational data. The workshop explores next generation user models to support migrant learners who move from educational experience to educational experience.

2 Players

When considering a new model for learner data, there are a number of entities at play.

- **User:** this is a lifelong learner, who carries his or her data from educational institution to educational institution, and who makes data selectively available to the other players.
- **Federated platform:** federated systems in which users and content interact with each other as part of educational experiences.
- **Federated educational institution:** a school, college, university, or training program, which would be member of some larger federation. On an administrative level, federation members acknowledge course credits and certifications (subject to mutually agreed-upon pairwise equivalency rules). On a technical level, the federated institutions build on a common infrastructure.
- **External educational institution:** an educational institution that is not federated, but which a user might want to apply to, and where a user might participate in educational experiences.
- **External stakeholder institution:** a non-educational external organization as a “consumer” of educational data, for example a potential employer, a scholarship organization, or some other funding agency.
- **External stakeholder user:** a person or system who contributes additional educational data or artefacts, such as letters of recommendation, internship evaluations, reviews, etc.
- **Service Providers:** these may be external providers of data space, cloud providers running campus IT systems, etc.

3 User Data

We need to consider different kinds of education-relevant data, including:

- **Personal data:** name, date-of-birth, basic demographics, student ID number or numbers (some countries have lifelong student IDs), etc. These data need to be disclosed or opened up in whole or in part to any other player that the user interacts with.
- **Certified transcript data:** academic credentialling, for example course credits or degrees. Currently, these data need to be transferred between educational institutions, where “the next chapter” of the learner’s education is written (as evidenced by terms such as “transfer student” or “transfer credit”), but in principle, this is one continuous record of the learner’s accomplishments.
- **Transactional data:** data routinely produced while interacting with learning platforms, which are potentially useful for personalization and guiding of learning, analytics, and content metrics. This is where the system “gets to know” the user.
- **Portfolio data:** a collection of artefacts generated during educational experiences, such as presentations, posters, artwork, compositions, videos, CAD-files, etc., which may be useful for applying for admission to other educational institutions or employment.
- **End-to-end confidential data:** Letters of recommendation, reviews, and the like, which a user needs to hand over from the author to the recipient without him- or herself being able to see them or to switch them out.

All of these data will need agreed-upon data formats, which is going to be a challenge that is outside the scope of this workshop. It is clear, though, that these will need to be structured, clear-text data formats (which could be represented for example in JSON) with possible binary attachments (e.g., images, PDFs, etc.).

4 Technologies

A new, distributed model for user-data will require a “trust nobody”-approach, particularly when it comes to identity and transcript data. There are a number of enabling technologies to be considered:

- **OpenID Connect:** When dealing with different players, first of all their identities need to be verified. A highly promising technology is OpenID Connect, which builds on the popular OAuth.
- **SSI:** Self-Sovereign Identify gives users control over their own identity, using the concept of decentralized identifiers.

- **SOLID:** Social Linked Data is a concept of storing data with the user instead of the application or service, under the control of the user. The user gets to choose which data they release to which platforms, i.e., the user has self-sovereignty over his or her data.
- **Federated Blockchain:** To establish trust, a key technology is that of a federated blockchain. Blockchains are usually associated with cryptocurrencies, energy-consuming Proofs-of-Work, and global entities like Ethereum, but none of that is needed here; likely, a federated blockchain, for example using Hyperledger Fabric, and simple Proof-of-Stake will do.
- **Gaia-X:** is a project to develop a secure and trustworthy federation of data infrastructure and service providers.

5 The Bigger Picture

User-data interacts with educational experiences, which to a large degree depend on educational content. A “Next Generation Learning Management System” basically becomes a platform where users (represented by their data pods) and content (modelled as Learning Objects) “meet up” for an educational experience. Figure 1 illustrates such a platform.

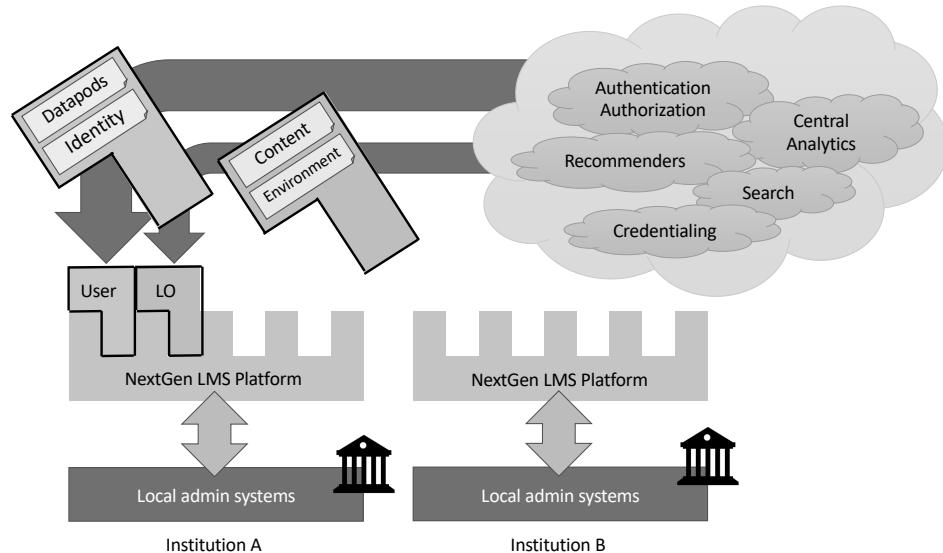


Fig. 1: A next generation learning platform

The user-model discussed here is thus part of a bigger learning ecosystem with pluggable components.

Workshop “Present and possible future(s) of inclusive EdTech and their applications in education and training”

Susan Beudt¹, Berit Blanc², Marco Kähler³, Nicole Vieregg⁴ and Rolf Feichtenbeiner⁵

Abstract: AI-based assistance and learning technologies can provide support for persons with various disabilities. An important question is what role AI-based technologies can play in the future, who could benefit from their application in vocational training and the workplace, and in which ways. To address these questions, a workshop concept was created based on the scenario methodology. The aim was to design and evaluate different foresight scenarios based on the topics AI, inclusion, and education. The workshop is divided into different phases and includes first research findings from the research project KI.ASSIST.

Keywords: AI, assistive technology, inclusion, EdTech, foresighting, scenario method

The COVID-19 pandemic in 2020 has made it particularly clear that efforts must be made to ensure that even persons without special needs have access to education. The use of digital technologies represents a great potential for the inclusion of persons with disabilities in learning and teaching processes [Ma18] [MH19]. However, the rapid developments in 2020 seen in online education and particularly in “emergency remote teaching” call even more for technologies that consider the needs of persons with disabilities.

Researchers at HU Berlin and DFKI work on various inclusion- and accessibility-related questions ranging from education in schools and universities to vocational training and rehabilitation. This workshop concept was created to further develop future scenarios about the use of digital and inclusive learning technologies in the context of vocational rehabilitation. The following questions guide the workshop. What developments are currently emerging in the field of assistive AI technologies? Which long-term implications can be derived for possible future developments?

To develop possible futures this workshop uses the scenario methodology. This method is widely used in future studies to depict various future developments and to create meaningful future images based on important key factors that can help to structure current actions and plan possible further steps [BH07], [KG08].

¹ Deutsches Forschungszentrum für Künstliche Intelligenz, Educational Technology Lab (EdTec Lab), Berlin, susan.beudt@dfki.de

² Deutsches Forschungszentrum für Künstliche Intelligenz, EdTec Lab, Berlin, berit.blanc@dfki.de

³ Deutsches Forschungszentrum für Künstliche Intelligenz, EdTec Lab, Berlin, marco.kaehler@dfki.de

⁴ Humboldt-Universität zu Berlin, Department of Computer Science, Berlin, vieregni@informatik.hu-berlin.de

⁵ Deutsches Forschungszentrum für Künstliche Intelligenz, EdTec Lab, Berlin, rolf.feichtenbeiner@dfki.de

The interest in topics related to inclusive EdTech seems to have grown significantly over the last years. This development can also be noticed regarding the DELFI when considering previous key topics (e.g., accessibility, inclusion), conference themes (2019: *Participation in education and science - access and barriers in the digital educational world*), and hosted workshops. With this workshop, we are following up on our last years' DELFI workshop *Digitally supported inclusive practices in education and training*. The aim is to set up a workshop series in this field that provides a platform for discussing current research, identifying relevant trends, and developing potential future collaborations.

This workshop generally aims to facilitate the interdisciplinary exchange between researchers working on topics related to the workshop theme and identification of potential future research topics. The aim is to develop and discuss future scenarios for the application of (AI-based) educational technologies for persons with disabilities, specifically in vocational rehabilitation in 2030, including a discussion of the most important influencing factors. Unfortunately, the workshop could not take place at the DELFI2021-conference as originally planned due to too few participants. Nevertheless, we would like to encourage you to use this contribution as an inspiration for future workshop concepts that aim to explore possible futures and identify opportunities for action.

Workshop format

Participants are invited to work on future scenarios for AI-based educational technologies for persons with disabilities within a foresight workshop conceptualized and hosted by the organizers. Based on initially presented results of the KI.ASSIST project and guided by the workshop hosts, participants will then have the opportunity to develop future scenarios for applying educational technologies in vocational education and training of persons with disabilities [BB20], [KI21].

The workshop was planned as an online event due to the given pandemic situation, but the workshop format could be easily adapted to a face-to-face situation. The content and procedure of the workshop are structured according to the following steps:

INTRODUCTION: Presentation of workshops targets and the procedure.

PRESENT: Examples of AI-based inclusive educational technologies – research goals and results of the project KI.ASSIST and its monitoring part

To get insight into current developments of AI-based assistance systems and to develop a common understanding of them, current results from the research project KI.ASSIST, especially the monitoring part of the project, are presented in this part. Results are based on scanning and analysis of AI-based assistance systems for the support of persons with disabilities [BB20]. In addition, concrete examples of such systems are shown, so that it becomes clear what is currently technically possible and what actual assistance can look like [KI21].

FUTURE I: Foresighting procedure and developed scenarios – presentation and discussion

In the course of the foresighting process that has already taken place in the KI.ASSIST research project, four scenarios were developed that examine the role of AI-based assistance systems in 2030. For this purpose, important key factors were determined with the help of AI experts, evaluated, and compiled into four scenarios. These scenarios were then evaluated in a Real-Time Delphi survey by a larger group of AI experts according to plausibility, probability, and desirability, and they were supplemented. Based on this, the four scenarios were then further developed. The exact steps of the foresighting process are presented and discussed in the workshop. The question regarding alternative options available in the course of the foresighting process will also be explored.

Presentation of the most important influencing factors (key factors) found in KI.ASSIST, projections of the future for each key factor and presentation of four created scenarios

Since not all influencing factors can be considered in the scenario creation, the most important influencing factors - key factors - were identified together with the AI experts in the previous workshop. For each of these key factors, the AI experts formulated four possible projections of the future during the workshop. Four scenarios were then created based on those projections. Finally, these scenarios were evaluated by a larger group of AI experts in a standardized Real-time Delphi survey.

Discussion about plausibility of key factors, their projections, and the developed scenarios

After the presentation of the key factors and the scenarios in this workshop, the next step is to understand their content and to discuss together whether the workshop participants also consider them to be relevant, plausible, and comprehensible.

FUTURE II: Revision of four possible scenarios

After the key factors and the scenarios have been presented and discussed, the next step is to critically examine the scenarios with regard to relevant discussion points and then to add to or modify them.

Breakout session (two groups) to discuss, complement or remove scenario descriptions (two scenarios per group)

To better work on the four scenarios described, the group is then divided into two subgroups. Each group is to evaluate two scenarios and then add to them if necessary. The aim is to use the shared knowledge of the groups to further supplement the scenarios at relevant points and to ensure that no important points are disregarded. Afterward, the groups shortly present their revised scenarios.

FUTURE III: Desired and probable scenarios – voting, discussion, and recommendations for action

The last step of the workshop focusses the following question: Which scenarios are most likely to become a reality in 2030 and which are most desirable? In order to make a better forecast, workshop participants should discuss and then vote on the most possible scenarios.

Afterward, the voting results are then discussed based on the following questions: What are the main reasons for the votes? How far apart are desire and reality or probability, respectively? How do we align wishes and reality, what recommendations for action can be formulated to achieve the desired scenarios?

At the end of the workshop, there should be a distinct scenario that participants think is most likely to happen, as well as a scenario of what they would like to see happen. Ideally, the workshop ends with the identification of the steps necessary to match the desired scenario with the probable scenario.

Bibliography

- [BB20] Beudt, S. et al.: Critical reflection of AI applications for persons with disabilities in vocational rehabilitation. Proceedings of DELFI Workshops. Bonn: Gesellschaft für Informatik e.V. 2020.
- [BH07] Bishop, P. et al.: The current state of scenario development: An overview of techniques. Foresight - The journal of future studies. 9. 10.1108/14636680710727516. 2007.
- [KG08] Kosow, H. & Gaßner, R: Methoden der Zukunfts- und Szenarioanalyse Überblick, Bewertung und Auswahlkriterien. 2008.
- [KI21] KI.ASSIST (English: AI.ASSIST): Overview of current examples of AI-based assistive technologies for persons with disabilities, Project-website, URL: <https://www.ki-assist.de/wissen/kuenstliche-intelligenz/ki-basierte-assistenztechnologien>. Last visit: 20.09.2021.
- [Ma18] Marzin, C.: Plug and Pray? A disability perspective on artificial intelligence, automated decision-making and emerging technologies. European Disability Forum. Brussels 2018.
- [MH19] Mark, B.G. et al.: Inclusion of Workers with Disabilities in Production 4.0: Legal Foundations in Europe and Potentials Through Worker Assistance Systems. Sustainability, 11(21), 5978, 2019.