

Abschlussbericht SUCAP

Success factors for training in virtual space: from simulation to practice



Dr. Oliver Christ und FHNW-Team

Olten, 30.11.2023

Inhaltsverzeichnis

Management Summary	5
Projektmitwirkende	8
1 Ausgangslage	10
1.1 SstA Kurs und Annäherungsdistanz/Checkliste	10
1.2 Erfahrungsbasiertes Lernen	11
1.3 Projektziel	11
2 Studie I – Bestimmung der Lernszene	13
2.1 Zielsetzung	13
2.2 Workshop Studie	14
2.2.1 Evaluation VR-Brillen	14
2.2.2 Fokusgruppe Lernszene und Elemente	14
2.2.3 Phase 1 Gruppendiskussion und Auswertung	15
2.2.4 Modell-Szenario und Auswertung	16
2.2.5 Workshop Ergebnisse	17
2.3 Fragenbogen Studie	20
2.3.1 Stichprobe und Materialentwicklung	20
2.3.2 Befragung und Auswertung	21
2.3.3 Ergebnisse Fragebogen	23
2.4 Integration Workshop und Befragung	25
2.5 Diskussion	27
2.6 Limitationen und Ausblick	30
3 Studie II – Bedienungsparadigma	32
3.1 Einleitung und Orientierung	32
3.2 Forschungsfragen	33
3.3 Methode und Sample	34
3.3.1 Material	34
3.3.2 Hardware	36
3.3.3 Fragebogen	37
3.3.4 Ablauf	38
3.3.5 Datenaufbereitung und -analyse	40
3.4 Ergebnisse	41
3.4.1 Quantitative Ergebnisse	41
3.4.2 Qualitative Ergebnisse	49
3.5 Diskussion	51

3.5.1	Virtuelle Agentin	52
3.5.2	Intrinsische Motivation	52
3.5.3	Cognitive Load	53
3.5.4	Fehlerhafte Interaktionen und Bearbeitungszeit	54
3.5.5	Anzahl gedrückte Buttons Buttonpräferenzen	55
3.5.6	Usability	56
3.5.7	User Experience	56
3.6	Fazit und Handlungsempfehlungen	57
3.7	Limitationen und Ausblick	58
4	Studie III – iVR-Multiplayer-Training mit Auszubildenden	60
4.1	Einleitung und Orientierung	60
4.1.1	CAMIL Modell	60
4.1.2	TAM-Modell	63
4.2	Forschungsfrage	64
4.3	Methode	65
4.3.1	Trainingskonzeption	65
4.3.2	Virtuelle Bedingungen	67
4.3.3	Stichprobe	70
4.3.4	Tutorial	70
4.3.5	Fragebogen	71
4.3.6	Versuchsablauf	71
4.4	Ergebnisse	72
4.4.1	Voraussetzungsprüfung der Daten	73
4.4.2	Selbstwirksamkeit (subjektiver Lernerfolg)	74
4.4.3	Nutzungsabsicht	74
4.4.4	Wahrgenommene Nützlichkeit	75
4.4.5	Wahrgenommene Benutzerfreundlichkeit	76
4.4.6	Prüfen nach Unterschieden	76
4.5	Diskussion	78
4.5.1	Präsenzempfinden	79
4.5.2	Situationsbezogenes Interesse	80
4.5.3	Wahrgenommene Kompetenz	80
4.5.4	Handlungsfähigkeit	81
4.5.5	Samplegrösse	81
4.6	Fazit	82
5	Studie IV – Ergänzung SstA mit iVR	83

5.1	Einleitung und Orientierung	83
5.1.1	Reglement RTE 20100	83
5.1.2	Checkliste Selbstschutz und Annäherungsdistanzen	83
5.1.3	Arbeitserfahrung	84
5.2	Forschungsfragen	85
5.3	Methode	86
5.3.1	Stichprobe/Stichprobengewinnung	87
5.3.2	Material	88
5.3.3	Virtuelle Lernumgebungen	89
5.3.4	Datenerhebung	89
5.3.5	Durchführung der Datenerhebung	89
5.3.6	Probleme bei der Datenerhebung	90
5.3.7	Ablauf der Datenerhebung	91
5.3.8	Fragebogen	93
5.4	Ergebnisse	95
5.4.1	Voraussetzungsprüfung der Daten	95
5.4.2	Punktzahl nach den Trainings	96
5.4.3	Arbeitseinschätzung TAM-Skalen	97
5.4.4	Arbeitseinschätzung Situationsbewusstsein und Gefahrenwahrnehmung	98
5.5	Diskussion	99
5.5.1	Probanden	101
5.5.2	Interaktion Trainer und Technik	101
5.5.3	iVR Classroom	102
5.6	Fazit	102
6	Ausblick	103
7	Literatur	105
8	Anhang	120

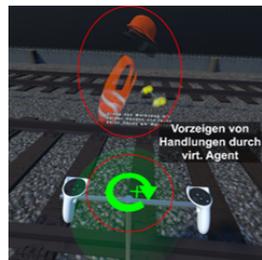
Management Summary

Das Ziel des vorliegenden Projekts war es, mögliche Erfolgs- oder Misserfolgskriterien von immersiver virtueller Realität (iVR) im Bereich des Gleissicherheitstrainings zu bestimmen. Immersive Virtuelle Realität steht hierbei für die Virtuelle Realität, in die man durch spezielle Brillen (Head Mounted Displays/HMDs) eintauchen kann. Im Projekt sollten Spezifikationen wie das Lernsetting (iVR- Einzelspieler, iVR-Mehrspieler, iVR-Mehrspieler mit retrospektivem Perspektivwechsel, iVR-Classroom) und Hardware (VR-Brille, Laptop) in der Gleissicherheitsschulung zum Kurs Selbstschutz Arbeit (SstA) überprüft werden. Die Notwendigkeit ergibt sich u.a. aufgrund der im Jahr 2020 beobachtbaren Durchfallquote im SstA von 17% (Zusatzkosten von 103k.-CHF). Ein möglicher Grund hierfür sehen die SBB-Experten in der Aufgabe zur Checkliste/ Annäherungsdistanz. Bei letzterer wird häufig nicht die volle Punktzahl erreicht und somit steigt die Chance die SstA Prüfung nicht zu bestehen. Durch Verbesserung der Situation mittels immersiver Realität könnten Einsparpotentiale von bis zu 103k.- CHF an Wiederholungsprüfungen erzielt werden. Um dies zu erreichen wurde eine virtuelle Schulungsumgebung in vier aufeinanderfolgenden Forschungsstudien evidenzbasiert entwickelt:

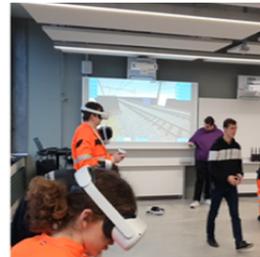
Studie 1: Lernszene wurde via Workshop (N=5) und Fragebogen umgesetzt (N=17) VR-Endgerät bestimmt und Tutorial angeregt



Studie 2: Entwicklung eines VR-Tutorials für Gleissicherheitstrainings mit oder ohne Mapping (N= 60) mit Elementen des Trainings



Studie 3: Messung von Einfluss verschiedener Multiplayer Settings auf menschliche Faktoren und subjektiven Lernerfolg von Auszubildenden N=69



Studie 4: Messung von Einfluss verschiedener Lernsettings und Hardware auf menschliche Faktoren und objektiven Lernerfolg von SstA Teilnehmende mit N=127



Die Umsetzung des Projekts begann in der ersten Studie mit Hilfe von Sicherheitstrainern und Sicherheitsexperten der SBB. Zusammen wurde eine Lernszene gestaltet und sicherheitskritische Elemente wie z.B. Sichtveränderung durch Wetterbedingung identifiziert. Gleichzeitig wurde ein iVR-System (Meta Quest 2) von den Trainern für die Handhabbarkeit im Kurs ausgewählt und die Notwendigkeit eines kürzeren Tutorials für die iVR-Brille (als das des Herstellers) erkannt.

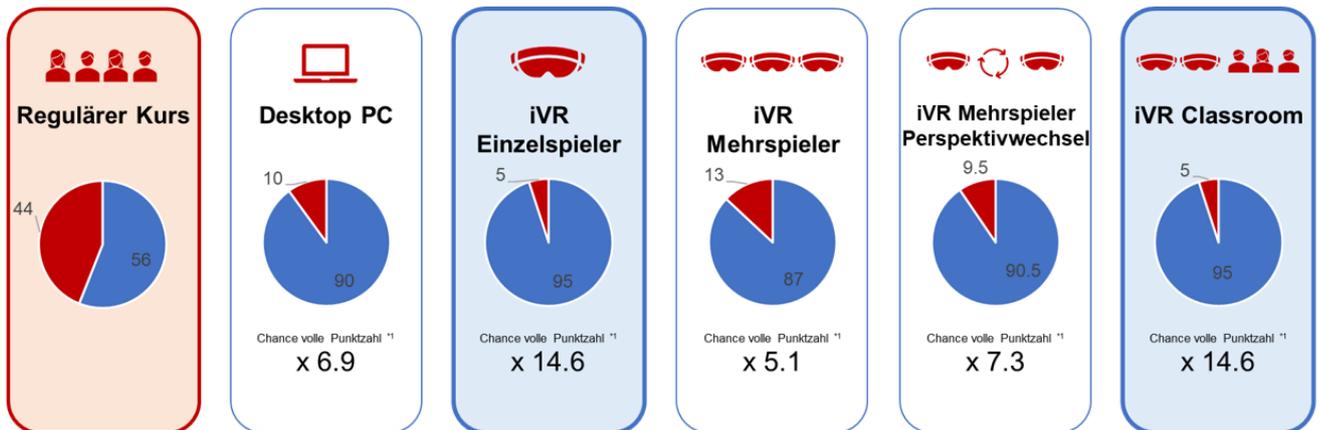
Während der Programmierung der Lernszene, wurde in der zweiten Studie das Tutorial entwickelt und evaluiert. Die Dauer konnte auf ca. 6-7 Minuten (im Vergleich zu 15 Minuten Herstellertutorial) gesenkt werden. Weiterhin konnte gezeigt werden, dass Erwachsene mehr von einer Zuordnung (Mapping) einer bestimmten Taste (z.B. Triggertaste am Controller) zu einem Ereignis (z.B. Teleport im virtuellen Raum) profitieren, als dass dies bei Auszubildenden der Fall war.

Mit diesem Wissen und der ersten Version der Trainingsumgebung konnte bei der login Berufsbildung AG in Altstetten die dritte Studie durchgeführt werden. Hierbei war die Fragestellung, welche psychologischen Erfolgsfaktoren des CAMIL Models (Cognitive Affective Model of Immersive Learning nach Makransky und Pettersen, 2021) die Nützlichkeit, Benutzerfreundlichkeit, Nutzungsintention und den subjektiven Lernerfolg in iVR-Mehrspielsettings vorhersagen können. Gleichzeitig wollten wir wissen, ob die identifizierten Erfolgsfaktoren sich in ihrer Ausprägung in drei iVR- Mehrspielsetting (iVR-Mehrspieler, iVR-Mehrspieler mit retrospektiven Perspektivwechsel und iVR-Classroom Setting) unterscheiden. Wir konnten zeigen, dass das Präsenzepfinden, die Handlungsfähigkeit, die Wahrgenommene Kompetenz und das Interesse statistisch bedeutsame Erfolgsfaktoren für die empfundene Nützlichkeit, Benutzerfreundlichkeit, Nutzungsintention und den subjektiven Lernerfolg sein können. Gleichzeitig gab es wenig bis keine Unterschiede in den drei iVR-Mehrspielsettings. Somit steht es der login Berufsbildung offen, sich für die ökonomisch sinnvollste Variante (iVR Classroom Setting) zu entscheiden, da nur 2 VR-Brillen und ein Bildschirm innerhalb der Sicherheitsinstruktionen benötigt werden würden. Am Ende der dritten Studie waren sowohl der iVR-Singleplayer, sowie die Desktopversion des Trainings fertig gestellt.

Die vierte und letzte Studie evaluierte schliesslich die gesamte Trainingsumgebung und die Lernsettings in echten SstA Kursen. Hierzu wurden SstA-Teilnehmende in eine von sechs Gruppen zugeteilt. Die interessierenden Variablen waren die erzielten Punkte, Nutzerfahrung und Arbeitseinschätzung:

Training as usual	Singleplayer- player VR	Singleplayer- player Desktop	iVR-Multiplayer	iVR-Multiplayer mit ret- rospektivem Perspektiv- wechsel	iVR- Classroom
					
Der SstA Kurs fand wie gewohnt statt. Die Teilnehmenden sehen Filme und arbeiten an einer Holzseilbahn. Am Ende des Tages erfolgte eine Befragung mittels Fragebogen	Zusätzlich zum regulären Kurs erlebten die Teilnehmenden allein mittels PC-Monitor vier Wetterbedingungen und Zugdurchfahrten. Mit Hilfe eines T-Schlüssels mussten sie blinkende Schrauben festdrehen. Unterstützt wurden sie durch eine Stimme respektive durch Text	Zusätzlich zum regulären Kurs erlebten die Teilnehmenden allein mittels HMD vier Wetterbedingungen und Zugdurchfahrten. Mit Hilfe eines T-Schlüssels mussten sie blinkende Schrauben festdrehen. Unterstützt wurden sie durch eine Stimme respektive durch Text	Zusätzlich zum regulären Kurs erleben die Teilnehmenden zusammen mit dem Trainer mittels HMD die virtuelle Welt. Der Trainer teilt die Personen zur Aufgabe ein, lässt Züge durchfahren und verändert das Wetter analog zum Singleplayer	Zusätzlich zum regulären Kurs erleben die Teilnehmenden zusammen mit dem Trainer mittels HMD die virtuelle Welt. Der Trainer teilt die Personen zur Aufgabe ein, lässt Züge durchfahren und verändert das Wetter analog zum Singleplayer. Am Ende einer Wetterbedingung erleben die Teilnehmenden sich selbst am Gleis 30 Sekunden vorher durch die Perspektive aus dem Zugführerstand	Zusätzlich zum regulären Kurs erleben die Teilnehmenden in Zweiergruppen mittels HMD die virtuelle Welt. Der Trainer teilt von der echten Welt aus die Personen zur Aufgabe ein, lässt Züge durchfahren und verändert das Wetter analog zum Singleplayer. Dies beobachtet der restliche Kurs von der echten Welt aus an einer Leinwand

Die Ergebnisse zeigten, dass sich in allen digitalen Interventionen überzufällige Verbesserungen im Vergleich zum regulären Kurs ergaben. Teilweise konnten starke Effekte beobachtet werden, auch wenn sich die digitalen Interventionen untereinander nicht signifikant voneinander unterschieden (in den Kreisdiagrammen ist blau =volle Punktzahl, rot= nicht volle Punktzahl):



Beispielsweise war die Chance volle Punktzahl in der Aufgabe zur Checkliste/Annäherungsdistanz zu bekommen in den iVR-Bedingungen Einzelspieler und Classroom im Vergleich zu regulärem Kurs um den Faktor 14,6 erhöht. Während im regulären Kurs 56% der Teilnehmenden die volle Punktzahl erzielten, waren es z.B. in der iVR-Bedingungen Einzelspieler und Classroom 95% der Teilnehmenden mit voller Punktzahl. Dies ist eine Senkung der Rate von Personen ohne volle Punktzahl von 44% im regulären Kurs auf 5%. Wenn wir durch den Einbezug von iVR in weitere Themen des SstA Kurses die allgemeine Durchfallrate auf 5% senken könnten, wäre das eine Reduktion von Zusatzkosten von 103.- auf 29k.- CHF und somit ein Einsparpotential von 74k.-CHF. Bei all diesen positiven Ergebnissen muss jedoch auch beachtet werden, dass wir aufgrund des Studiendesigns (keine randomisierte Zuteilung, wenig kontrollierte Bedingungen durch z.B. unterschiedliche Räume und Trainer) keine Ursache-Wirkung-Beziehung haben und erst am Anfang stehen, was das Verständnis über die Wirkung von iVR betrifft. Weiterhin fiel das positive Herausstechen der iVR-Classroom Bedingung in den subjektiven Bewertungen zur Nutzerfahrung und Arbeitseinschätzung auf, welche ein sehr ökonomisches Trainingssetting mit nur zwei iVR-Brillen darstellt. Wir wissen jedoch nicht, ob auch diese Bewertung in kontrollierten, randomisierten Studienbedingungen weiterbestehen würde. Darüber hinaus fehlen Interviewdaten von den Sicherheitstrainern, die zeigen, wie sie sich in den unterschiedlichen Bedingungen gefühlt haben und wo sie eventuell weitere Herausforderungen im Praxisalltag sehen. Gleichzeitig sind die Trainer eine sehr wertvolle Ressource, da (1) sie wussten, wo Handlungsbedarf besteht (siehe regulärer Kurs 44% keine volle Punktzahl), (2) sie wissen, was die iVR- Lernumgebung an Funktionen braucht (siehe Studie 1) damit ein Erfolg möglich ist, der in diesem Projekt definitiv nicht zum Nachteil der SstA-Teilnehmenden war.

Projektmitwirkende

Steering Board

Stephan Gut	SBB - Fachbereichsleiter Sicherheit, Kunde, Konzern
Dr. Oliver Christ	FHNW - Dozent ohne Gesamtauftrag
Stefan Adam	SBB - Fachexperte Bildung
Paul Hügli	SBB - Leiter Arbeitssicherheit und Arbeitsstellensicherheit
Roland Jakob	SBB - Teamleiter
Dr. Günhan Akarçay	SBB - Leiter XR
Matthias Hirschi	Login AG - CFO

SBB

Pascal Duchenne	Ausbilder Sicherheit
Markus Haas	Ausbilder Sicherheit
Ernst Müntener	Ausbilder Sicherheit
Stevo Keric	Ausbilder Sicherheit
Nicola Porreca	Ausbilder Sicherheit
Gottfried Herger	Ausbilder Sicherheit
Thorsten Imhof	Ausbilder Sicherheit
Hans Jürgen Sommer	Ausbilder Sicherheit

Login AG

Bea Walder	Login AG - Leiterin Basisausbildung Gleisbau
Charles Bögli	Berufsbildner Gleisbau
Fitim Sutaj	Berufsbildner Gleisbau

FHNW Team

Pascal Meier	Mitarbeiter
Andreas Papageorgiou	Mitarbeiter
Andreas Urech	Student MSc
Anja Schmid	Studentin MSc
Jessica Ujik	Studentin MSc
Aaron Ettin	Student MSc
Dimitry Wey	Student MSc
Kim Nina Oberholzer	Studentin MSc
Luca Niederhauser	Student MSc
Céline Abt	Studentin BSc
Moritz Aufdenblatten	Student BSc
Loredana Di Cesare	Studentin BSc
Sofia Harrer	Studentin BSc
Anita Knezevic	Studentin BSc

Anmerkung: Dieser Bericht ist mit Hilfe aller unter FHNW-Team aufgeführten Personen als Autoren entstanden, die im Zuge von Leistungsnachweisen einen wertvollen Beitrag zum Gelingen dieses Projekts beigetragen haben.

1 Ausgangslage

Die Digitalisierung schreitet voran und in vielen Unternehmensbereichen werden positive sowie negative Aspekte dieser Entwicklung deutlich. Im Bildungsbereich sind digitale Tools nicht mehr wegzudenken und helfen mittels Videos und interaktiver Software, das Wissen an Teilnehmende zu vermitteln. Neuere Methoden wie immersive (eintauchbare) Technologien (z.B.: immersives Virtual Reality), können mittels Head-Mounted-Displays (HMDs) im Menschen die Illusion erzeugen, sich an einem, rein computer-generierten Ort zu befinden. Diese als «Präsenzempfinden» beschriebene Ortsillusion scheint für viele Trainingsszenarien von Vorteil, da man in sicherer Umgebung, zeitlich und räumlich unabhängig, Situationen erleben kann, die in der echten Welt unmöglich, sehr aufwändig oder gefährlich wären. Gleichzeitig lassen sich die Erlebnisse zeitlich vorspulen, zurückspulen und durch die Augen von anderen Personen, die im virtuellen Raum anwesend sein können, aber nicht müssen, erleben. Auch zeigen wissenschaftliche Ergebnisse, dass die durch immersive Technologien vermittelte Lernleistung den durch herkömmliche Methoden überlegen sein können, aber nicht müssen (Wu et al., 2020; Checa and Bustillo, 2020; Di Natale et al., 2020; Radianti et al., 2020; Huang et al., 2021; Pellas et al., 2021; Hamilton et al., 2021; Coban et al., 2022; Rojas-Sánchez et al., 2023). Zusammenfassend kann man mittels immersiven Technologien eine umfassendere sensorische und perspektivische Lernerfahrung bieten, die in der echten Welt nicht möglich ist. Dies kann dann von Vorteil sein, wenn bestimmte herkömmlich Trainingstools in manchen Themen nicht den gewünschten Effekt erzielen. Nach einem Austausch mit Experten für Sicherheitstrainings bei der SBB, konnte ein solches Thema im Rahmen einer Sicherheitsschulung (SstA) identifiziert werden.

1.1 SstA Kurs und Annäherungsdistanz/Checkliste

Im von der SBB durchgeführten Kurs Selbstschutz Arbeit (SstA), werden interne wie externe Mitarbeitende hinsichtlich der Verhaltensweisen bei Arbeiten am Gleis geschult. Der Kurs dauert etwa zwei Tage und schliesst mit einer theoretischen Prüfung. Erst nach bestehen diese Prüfung darf man im Arbeitsauftrag das Gleis betreten. Als Methoden kommen im Kurs Videos (Beamer) und Holzmodelle von Eisenbahnen und Schienen auf einem Tisch zu Einsatz. Es wird auch eine Begehung im Gleis durchgeführt. Die Experten wiesen darauf hin, dass die Teilprüfungsleistung Checkliste/Annäherungsdistanz meist nicht vollständig gelöst wird/ weniger oft volle Punktzahl erreicht wird. Sie schliessen weiterhin darauf, dass diese Teilaufgabe einen wesentlichen Beitrag zum nicht bestehen der gesamten Prüfung liefert. Betrachtet man hierzu die Daten von 2020 zur Prüfungsleistung zeigt sich folgendes Bild (siehe Abbildung 1.1)

Test	Number of participants	Number passed	Number failed
Sst B Gk	1096	1062	34
Sst A Gk	1726	1539	187
Sst A Wk	1173	842	331

Abbildung 1.1: Ausschnitt aus dem SUCAP-Projektantrag. 2020 fielen bei 2899 SstA-Kursen 518 Personen durch die Prüfung (17%). Bei etwa 200 CHF.- pro Wiederholungsprüfung machen das 103000.-CHF zusätzlichen Aufwand. Würde man durch geeignete Massnahmen die Durchfallquote halbieren oder ganz verringern, wären Einsparpotentiale von 51500- 103000 CHF möglich.

In Abbildung 1.1 sieht man schnell, dass 17% die Prüfung nicht bestanden haben und dies mit Mehrkosten von ca. 103000.-CHF verbunden ist. Eine Verbesserung dieser Situation würde also nicht nur dem Menschen beim Lernen unterstützen, sondern könnte auch Kosten bis in den sechsstelligen Bereich einsparen. Eine neue Möglichkeit das Thema der Checkliste/Annäherungsdistanz im SstA Kurs zu vermitteln, ist der Einsatz immersiver virtueller Realität als erfahrungsbasiertes Lernmedium.

1.2 Erfahrungsbasiertes Lernen

Erfahrungsbasiertes Lernen ist nach Kolb (2014) eine Form des aktiven Lernens in dem die lernende Person Wissen durch das Beobachten, Agieren und Ausprobieren in realen oder simulierten Situationen erlangt. Dieses steht passivem Lernen gegenüber bei welchem die lernende Person Wissen durch zum Beispiel Texte oder Erzählungen vermittelt bekommt (Kolb, 2014). Erfahrungsbasiertes Lernen findet nach Kolb (2014) in einem Kreislauf von vier Schritten statt: Die interaktive Teilhabe an einer konkreten Erfahrung im hier und jetzt, dem kognitiven Verarbeiten und Reflektieren dieser Erfahrung, dem Bilden eines abstrakten Konzeptes nach welchen Regeln die erfahrene Situation abläuft und dem aktiven Experimentieren wie man in der erfahrene Situation besser agieren könnte. Bei jedem Durchlaufen der Situation wird somit mehr über diese gelernt (Kolb, 2014). Moon (2013) betont dabei die Wichtigkeit von Reflektion und der Notwendigkeit dedizierte Möglichkeiten für das Lernen durch Reflektion und Feedback zu schaffen. Obwohl erfahrungsbasiertes Lernen selbständig ablaufen kann, betonen Jacobson und Ruddy (2004) das Potenzial die Reflexion der Erfahrung durch Erleichterungen, wie gezieltes Fragen und Leiten der Auszubildenden, anzustossen und zu fördern.

1.3 Projektziel

Erfahrungsbasiertes Lernen kann somit allein oder in einer Gruppe erfolgen. Möchte man nun allein oder in Gruppen wiederholbare und teils in der echten Welt nicht mögliche Situation erfahrungsbasiert nutzen, braucht es die immersive virtuelle Welt. Das Ziel des vorliegenden Projekts ist somit mögliche

Erfolgs- oder Misserfolgskriterien von immersiver virtueller Realität (iVR) im Bereich Gleissicherheits- trainings zu bestimmen. Hierbei sollen Spezifikationen wie das Lernsetting iVR- Einzelspieler, iVR- Mehrspieler, iVR-Mehrspieler mit retrospektivem Perspektivwechsel (Aufnahme eigener Handlung am Gleis und Wiedergabe 30 Sek. in der Vergangenheit mit Sicht durch das Zugführeresstand) iVR-Class- room und Hardware (VR-Brille, Laptop) in der Gleissicherheitsschulung (Aufgabe der Checkliste/Annä- herungsdistanz) überprüft werden. Hierbei wurde mittels vier Studien eine virtuelle Trainingsumgebung evidenzbasiert (Grundlage aus Wissenschaft und Praxis) entwickelt und in der vierten Studie an Teil- nehmenden des SstA Kurses getestet. Die Übersicht des Projekts ist in Abbildung 1.2 dargestellt:

Studie 1: Lernszene wurde via Workshop (N=5) und Fragebogen umgesetzt (N=17) VR-Endgerät bestimmt und Tutorial angeregt



Studie 2: Entwicklung eines VR-Tutorials für Gleissicherheitstrainings mit oder ohne Mapping (N= 60) mit Elementen des Trainings



Studie 3: Messung von Einfluss verschiedener Multiplayer Settings auf menschliche Faktoren und subjektiven Lernerfolg von Auszubildenden N=69



Studie 4: Messung von Einfluss verschiedener Lernsettings und Hardware auf menschliche Faktoren und objektiven Lernerfolg von SstA Teilnehmende mit N=127



Abbildung 1.2: Übersicht des Projektverlaufs.

2 Studie I – Bestimmung der Lernszene

Diese Studie möchte untersuchen, welches iVR-System am geeignetsten für den Einsatz im Gleissicherheitsstraining ist und wie eine virtuelle Umgebung ideal gestaltet und werden kann. Dafür wurden Workshops mit Ausbildern durchgeführt und Fragebogen verschickt. Die Resultate zeigen die Meta Quest 2 als beste iVR-Lösung. Zudem konnten Gestaltungskategorien beschrieben und nach ihrer Relevanz geordnet werden.

2.1 Zielsetzung

Aus dem zuvor beschriebenen Projektrahmen ergeben sich in der vorliegenden Arbeit folgende Forschungsfragen:

- 1) Wie bewerten SBB-Sicherheitsauszubildende iVR-Trainingselemente für zukünftige iVR-Trainings?
- 2) Wie gelingt es konzeptionelle Hinweise für die Erstellung einer iVR-Umgebung zum Thema der Gleissicherheit zu erarbeiten?

Die erste Forschungsfrage geht den Gestaltungshinweisen für die Technik und das Lernszenario in der iVR nach. Die Gestaltungshinweise sollen dabei auf dem Wissen der Experten im Bereich der Gleissicherheit basieren. Die Forschungsfrage zwei stellt die Frage nach dem Weg bzw. der Methodik. Was ist die Methodik, die zu den Gestaltungshinweisen zu der Technik und dem idealen Szenario in der VR führt? Es geht in der vorliegenden Arbeit folglich um Gestaltungshinweise und gleichzeitig um die Methode, die zu diesen Gestaltungshinweisen führt. Aus den beiden Forschungsfragen ergab sich ein «exploratory design» nach Cresswell et al. (2003, zit. nach Kuckartz, 2014). Bei dem Zwei-Phasen-Design (siehe Abbildung 2.1) hatten die qualitativen Methoden (QUAL) zu Beginn des Projekts Priorität (Kuckartz, 2014). Ein explorativer Mixed Methods Zugang war nötig, da verteiltes Wissen über den fachlichen Inhalt der Schulungen für die Gleisbegehungen im Selbstschutz bei der SBB vorhanden waren. Zudem wurden in der Literatur keine methodischen Ansätze zur Erarbeitung von iVR-Umgebungen für Schulungszwecke gefunden. Die nachfolgende quantitative (quant) Befragung wurde basierend auf den qualitativen Ergebnissen erstellt. Ziel war es, das Datenmaterial mittels Zahlenangaben zu vertiefen, zu ergänzen und zu validieren. Zum Schluss wurden die qualitativen und quantitativen Ergebnisse zusammengeführt, um die Gestaltungshinweise zur Technik der iVR abzuleiten.

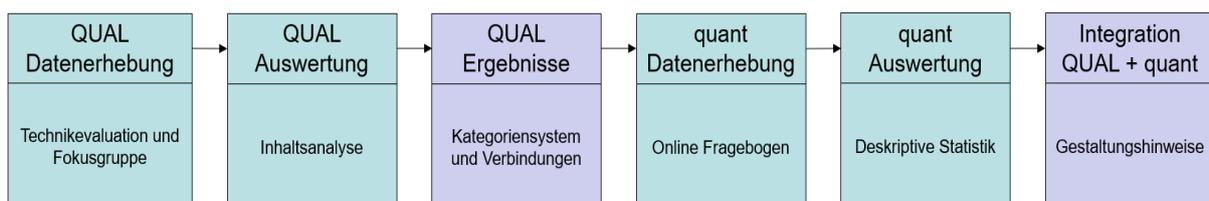


Abbildung 2.1: Mixed Methods Versuchsdesign der Studie 1

2.2 Workshop Studie

Für die Stichprobe wurden bewusst fünf Experten ausgewählt (Diekmann, 2007), die hauptberuflich Schulungen im Bereich der Gleisbegehung im Selbstschutz durchführen. Diese verfügten über das notwendige Expertenwissen zur Evaluation der Technik und der Erarbeitung der iVR-Inhalte.

2.2.1 Evaluation VR-Brillen

Die iVR-Systeme (zwei stationäre und zwei mobile Systeme) wurden bezogen auf die Anwendung in einer realen Schulungssituation bewertet. Forschungsgegenstand war das Expertenwissen über die Schulungssituationen. Die Evaluation wurde in drei Schritte unterteilt. In einem ersten Schritt wurden den Experten die iVR-Brillen und ihre Funktionen präsentiert. Die präsentierten Brillen waren: HP G2 Omnicep, Pimax8K X, Meta Quest 2 und HTC Focus 3. Die Präsentation der Funktionen umfassten technische Details sowie mögliche Einsatzoptionen wie bspw. iVR-Singleplayer, iVR-Multiplayer oder iVR-Classroom. In einem zweiten Schritt zogen die Experten die Brillen an und navigierten in einer iVR-Testumgebungen. Daraufhin erfolgte die eigentliche Evaluation der iVR-Brillen. Die Bewertung fand entlang der Kriterien Tragekomfort, Bildqualität, Steuerbarkeit (Steuerungsmodus im iVR), Erlernbarkeit (wie gut kann der Umgang mit der Brille von Schulungsleitenden erlernt werden) und benötigtes Zubehör (ist bspw. ein externer Computer nötig) statt. Die Kriterien sollten dabei immer auf das reale Schulungssetting bezogen werden. Die Bewertung fand anhand der Vergabe von Punkten statt. Alle Experten verfügten über eine Stimme, die sie ihrer bevorzugten iVR-Brille geben konnten. Zudem mussten die Experten angeben, ob sie die gewählte Brille in einem Singleplayer-, Multiplayer- oder Unterrichtsetting einsetzen würden.

2.2.2 Fokusgruppe Lernszene und Elemente

Die Erarbeitung der Bilder bzw. Szenarien für die VR fand in Fokusgruppen statt. Ziel war es, iVR-Modell-Szenarien zu erarbeiten, die die Schulung des Selbstschutzes bei Gleisbegehungen ergänzen. Die Modell-Szenarien sollten dabei eine Situation darstellen, welche die Experten mittels iVR in eine Schulung einbauen könnten. Die Fokusgruppe wird aus methodischer Sicht in zwei Phasen aufgeteilt, da während der Fokusgruppe zwei Forschungsgegenstände betrachtet wurden. Einmal war die Gruppendiskussion zwischen den Experten ein Forschungsgegenstand und das daraus entstandene Produkt eines Modell-Szenarios in seiner bildlichen Form bildete den zweiten Forschungsgegenstand. Diese Unterscheidung ist relevant, da die Forschungsgegenstände unterschiedlich ausgewertet und verarbeitet wurden. Die Methode der Fokusgruppe ermöglichte auf diese Weise den Diskurs der gemeinsamen

Erfahrungen der Experten und gleichzeitig die Herausarbeitung der kollektiven Wissensbestände als Ergebnis (Przyborski & Riegler, 2010).

2.2.3 Phase 1 Gruppendiskussion und Auswertung

In der Gruppendiskussion diskutierten die Experten ihr Wissen und begannen mit den Bastelmaterialien Modell-Szenarien zu erstellen (siehe Abbildung 2.2). Da dieser Prozess ein Abstimmen von Meinungen und Erfahrungen zwischen den Experten verlangte, war es wichtig, dass die Diskussionsleitung möglichst wenig in die Diskussion eingriff (Przyborski & Riegler, 2010). Es wurde daher zu Beginn der Auftrag und das Ziel kommuniziert und die verfügbaren Bastelmaterialien vorgestellt. Das Ziel war es, ein Modellszenario zu erstellen, das Schulungsteilnehmende in der VR erleben und dabei wichtige Inhalte der Schulung erlernen können. Forschungsgegenstand während der Gruppendiskussionen waren die kommunikativ generalisierten Sinngehalte. Was wird gesagt und diskutiert? Diese Daten wurden mittels Videoaufnahmen festgehalten.

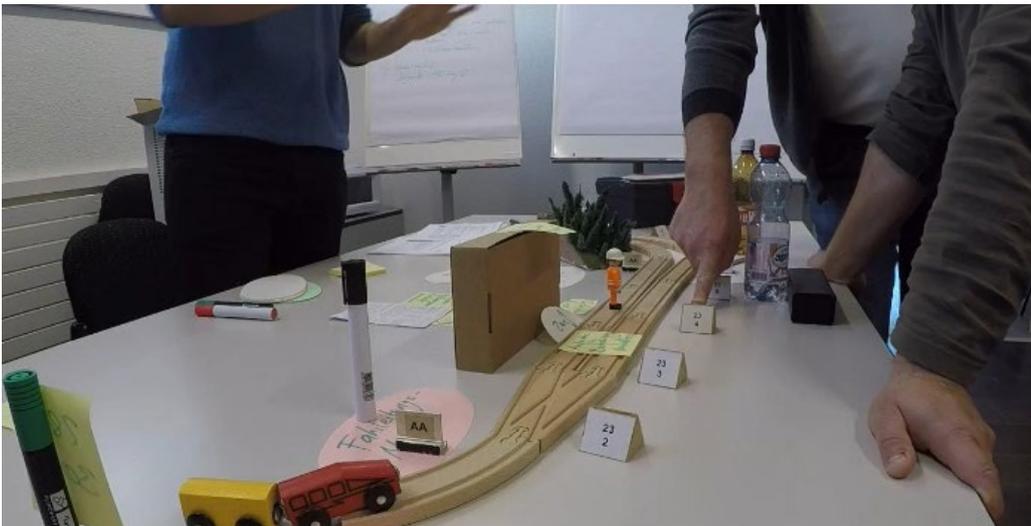


Abbildung 2.2: Bildausschnitt aus den Videoaufzeichnungen der Modell-Szenario-Erarbeitung

Die Videoaufnahmen aus der Gruppendiskussion wurden mittels inhaltlich strukturierender Inhaltsanalyse nach Kuckartz (2012) ausgewertet. Ziel der Auswertung war es, alle genannten Gegenstände der Diskussionen festzuhalten, da diese nicht alle in den finalen Modell-Szenarien dargestellt wurden. Zudem beugte es allfälligen Verzerrungseffekten vor, falls bspw. dominante Teilnehmende die Definition des finalen Modell-Szenarios zu stark prägten. Das Datenmaterial wurde mit dem Ziel codiert, Kategorien aus den Gegenständen, die von den Experten während der Erarbeitung der Szenarien genannt wurden, zu bilden. Aus zeitlichen Gründen wurde nicht das ganze Datenmaterial codiert, sondern nur

die Textpassagen, in denen die Gegenstände genannt wurden. Bei den Codes handelte es sich um Faktencodes nach Kuckartz (2014). Aufgrund des unbekanntes Themenfeldes und dem Ziel, Muster in den Daten zu finden, wurde das Datenmaterial entlang der In-Vivo-Kodierstrategie codiert und die Kategorien wurden rein induktiv gebildet. Dieses Vorgehen ermöglichte die Analyse von genannten Gegenständen, die für die Erarbeitung einer iVR-Umgebung relevant sein könnten, indem alle prominenten Begrifflichkeiten codiert wurden und in das Code-System einfließen. Die Kategorien wurden in einem zweiten Schritt thematisch geclustert und visualisiert, um ein Kategoriensystem mit Haupt- und Subkategorien bilden zu können. Die Verbindungen unter den Haupt- und Subkategorien wurden ebenfalls mittels der visualisierten Cluster exploriert. Die Verbindungen unter den Kategorien wurden dann in einem dritten Schritt mittels Kreuztabelle quantifiziert. Im Folgenden wird anhand einer Beispielaussage aus dem Videomaterial der Fokusgruppe verdeutlicht, wann von einer Verbindung zwischen zwei Kategorien gesprochen wurde: «Die Gleisabstände sind wichtig. Wenn es fünf Meter sind ist es kein Thema, wenn es aber nur dreieinhalb Meter sind, dann ist es kein Fluchtraum mehr.» In diesem Zitat ist einerseits die Kategorie «Abstand zwischen Gleisen» und andererseits die Kategorie «Fluchtraum» enthalten. Zudem ist mit der Angabe des Abstandes von fünf bzw. dreieinhalb Meter eine mögliche Ausprägung der Kategorie «Abstand zwischen Gleisen» enthalten. Mögliche Ausprägungen von Kategorien wurden im Kategoriensystem jeweils als Subkategorie zur dazugehörigen Hauptkategorie aufgeführt. Aufgrund der gleichzeitigen Nennung von Kategorien im gleichen Satz, wurde somit eine Verbindung identifiziert. In der Kreuztabelle wurde die Verbindung zwischen der Kategorie «Abstand zwischen Gleisen» und der Kategorie «Fluchtraum» eingetragen.

2.2.4 Modell-Szenario und Auswertung

Das Produkt aus der Bastel- bzw. Gruppendiskussion waren die physischen Darstellungen von Modell-Szenarien (siehe Abbildung 2.3). Nachdem die Experten in der Gruppendiskussion herausfanden, wo ihre gemeinsame Erfahrung lag, pendelten sich diese auf die gemeinsamen Erlebniszentren ein (Bohensack, 2000). In den Modell-Szenarien manifestieren sich die kollektiven Orientierungen der Experten. Forschungsgegenstand ist hierbei die Ebene des konjunktiven Wissens, die Sinnebene, auf der sich die Experten mit ihrer gemeinsamen Erfahrung verstehen (Przyborski & Riegler, 2010). Dies wurde in der Form von Bildern (siehe Abbildung 2.3) festgehalten.

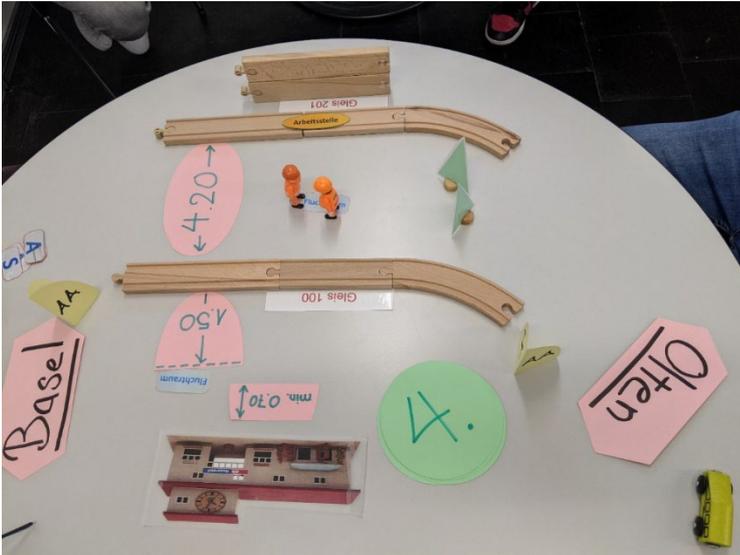


Abbildung 2.3: Bild eines Modell-Szenarios in der Phase der Erarbeitung im Gruppenprozess

Die Bilder der entstandenen Modellszenarien wurden miteinander verglichen und Gemeinsamkeiten, die auf allen oder dem Grossteil der Bilder vorkamen, wurden festgehalten. Basierend auf den Gemeinsamkeiten wurde ein abstrahiertes Modellszenario erstellt, das diese Gemeinsamkeiten in sich vereint. Dieses Modellszenario wird im Folgenden Grundkulisse genannt. Diese Grundkulisse enthält in stark verdichteter Form die verbindenden Wissensbestände der Experten.

2.2.5 Workshop Ergebnisse

In Tabelle 2.1 sind die Bewertungen der vier präsentierten VR-Brillenmodelle durch die Experten zusammengefasst. Am häufigsten wurde die Meta Quest 2 als Favorit genannt (drei von fünf Experten), die HTC Vive Focus 3 und die HP G2 Omnicept haben jeweils eine Stimme bekommen. Alle fünf Experten haben das iVR-Classroom als das geeignetste Format für ihre Schulungen bewertet. Die iVR-Multiplayer-Variante wurde zudem viermal als Alternative zum iVR-Classroom genannt.

Tabelle 2.1: Bewertung VR-Brillen bezüglich Umsetzbarkeit im Schulungsalltag

	Modell VR-Brille	Anzahl Stimmen	iVR-Single-Player	iVR-Multi-Player	iVR-Classroom
1	Meta Quest 2	3	0	alternativ	3
2	HTC Vive Focus 3	1	0	0	1
3	HP G2 Omnicept	1	0	alternativ	1
4	Primax 8kx	0	0	0	0

Bei der Einweisung in die Brillen viel auf, dass es bei der Quest etwa 15 Minuten dauerte bis alle Trainer die Bedienung der Controller und das Anlegen der Brille an sich, selbstständig durchführen konnten. Dies veranlasste uns zur Entwicklung und Evaluation eines Tutorials mit dem Ziel grundlegende

Interaktion in der Lernszene zu kennen und die Dauer der Einführung zu reduzieren. Diese Untersuchung ist unter Studie II Bedienungsparadigma im nächsten Kapitel aufgeführt. Insgesamt wurden 105 Minuten Videomaterial der Fokusgruppendifkussionen ausgewertet und daraus entstanden 11 Hauptkategorien. Tabelle 2.2 gibt einen Überblick über die Hauptkategorien sowie jeweils beispielhafte und oft genannte Subkategorien. Im Anhang A ist das gesamte Haupt- und Subkategoriensystem hinterlegt.

Tabelle 2.2: Haupt- und Subkategorien

Hauptkategorien	Subkategorien (nicht abschliessend)
Fluchtraum	Fluchtraum neben Lärmschutzwand, Bahnhof als Fluchtraum
Tätigkeiten zur Aufrechterhaltung der Sicherheit	Fluchtraum aufsuchen, fortlaufendes Einschätzen von Distanzen, Gleissperrung veranlassen
Sichtweite	Annäherungsdistanz, Kilometrierungs-Tafeln
Ort der Arbeitsstelle	Gleis mit Lärmschutzwand, zwischen den Gleisen
Arbeiten am Gleis	Kontrollen, Vermessungen, Gleisüberquerung
Wetter	Nebel, starker Regen/ Schneefall
Zug	Zuggeschwindigkeit, Fahrtrichtung
Hindernisse	Wald, Bagger, Baustelle, Lärmschutzwand
Ablenkungen	Telefon klingelt, Baustellenlärm
Zustand des Gleisabschnitts	Betriebsgleis, gesperrt/ungesperrt
Personenkonstellation	Allein, zu zweit

Aus den Diskussionen der Fokusgruppen wurde deutlich, dass viele Interaktionen zwischen den einzelnen Haupt- sowie Subkategorien bestehen. So hat zum Beispiel das Wetter Einfluss auf die Sichtweite, was wiederum Folgen für die Bestimmung des Fluchtraumes hat. Diese Interaktionen wurden in Form von Verbindungen zwischen den einzelnen Kategorien visualisiert. Eine Verbindung kommt dann zustande, wenn zwei Kategorien im gleichen Zusammenhang genannt wurden.

Die visuelle Auswertung führte zur Erkenntnis, dass durch die Vielzahl von Verbindungen unter den Kategorien kein Muster festzustellen ist. Die anschliessende Quantifizierung der Interaktionen bestätigte diese Erkenntnis. In der Tabelle 2.3 sind die zehn Haupt- oder Subkategorien mit den meisten Verbindungen zu anderen Haupt- oder Subkategorien aufgeführt.

Tabelle 2.3. Kategorien mit den meisten Verbindungen zu anderen Kategorien

Anzahl Verbindungen	Kategorie
17	Tätigkeiten zur Aufrechterhaltung der Sicherheit
12	Räumungszeit
10	Fluchtraum
9	Arbeiten am Gleis
9	Annäherungsdistanz
7	Ort der Arbeitsstelle
6	Sichtweite
6	Nebel
5	Gleis mit Lärmschutzwand
5	Zuggeschwindigkeit

Aus der qualitativen Auswertung der Modell-Szenarien ging die Grundkulisse hervor (siehe Abbildung 2.4). Die darin enthaltenden Objekte, kommen im Grossteil der entstanden Modellszenarien vor weshalb die Grundkulisse als Verdichtung des Expertenwissens betrachtet wird. Die Grundkulisse verfügt über zwei befahrene Gleise und auf der einen Seite eine Lärmschutzwand, die den direkten Zugang zum Gleis verhindert. Auf beiden Seiten der Gleise sind Gehwege eingezeichnet und in undefinierter Distanz biegt eine Kurve ab während gleichzeitig ein Waldabschnitt die die Sicht auf allfällige Züge aus

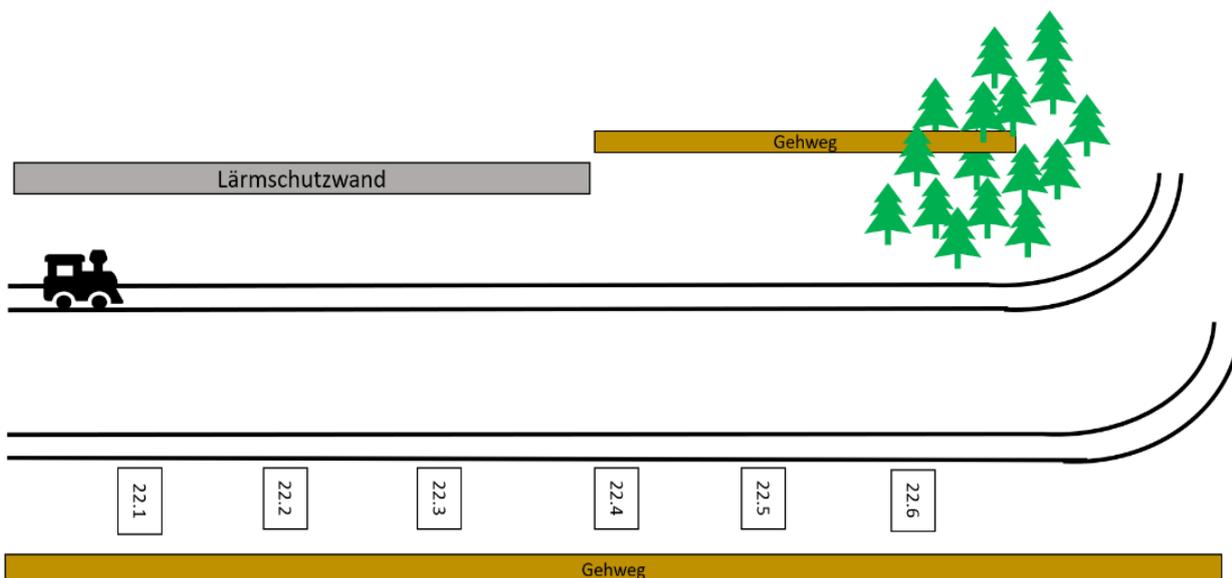


Abbildung 2.4: Modell-Szenario aus der Fokusgruppe

dieser Richtung verdeckt. Das Grundscenario verfügt zudem über Kilometrierungsangaben auf Schildern. Anhand dieser können bspw. Distanzen abgelesen werden.

2.3 Fragenbogen Studie

Die Vielzahl an entstandenen Haupt- und Subkategorien sowie die Erkenntnisse aus der Quantifizierung der Verbindungen unter den Kategorien zeigte deutlich auf, dass die Sicherheit im Selbstschutz bei Gleisbegehungen von einer Vielzahl von Interaktionen zwischen Gegenständen abhängt. Diese Interaktionen sind komplex und zeigen sich bei kleinsten Änderungen der Situation sehr unterschiedlich. Diese Komplexität ist schwer zu programmieren und es wird angenommen, dass ab einem gewissen Komplexitätsgrad eine Überforderung im Sinne der Cognitive load theory (Sweller, 2005; 2010) von iVR-Nutzenden eintreten könnten. Dies führt zu der Notwendigkeit einer Komplexitätsreduktion, was wiederum bedeutet, gewisse Kategorien auszuschliessen, damit es zu weniger Interaktionen kommen kann. Um Kategorien auszuschliessen, muss herausgefunden werden, über welche Relevanz bzw. welche Priorität diese für die Sicherheit im Gleisschutz verfügen. Aktuell ist nur klar, dass die Kategorien über eine Relevanz verfügen, jedoch nicht wie hoch diese Relevanz ist. Die Priorisierung der Kategorien erfolgte mittels quantitativer Befragung. Ist bspw. die Hauptkategorie Wetter ein wichtigerer Faktor als die Hauptkategorie Sichtweite? Oder ist bspw. die Wetterausprägung Starkregen wichtiger als Nebel? Gleichzeitig wurde die Validierung der Kategorien vorgenommen. Um die Befragung durchführen zu können, war es nötig, das entstandene Kategoriensystem in die Form eines Fragebogens zu übersetzen.

2.3.1. Stichprobe und Materialentwicklung

Die Stichprobe ($n=17$) für den Fragebogen wurde bewusst aus den Abteilungen Sicherheit & Qualität und den Sicherheitstrainern (Schulungen zur Gleisbegehungen) ausgewählt (Diekmann, 2007). Das war nötig, damit die Personen über Expertenwissen zum Selbstschutz in der Gleisbegehung verfügten. Um zu verstehen welche Kategorien die höchste Priorität haben wurden Haupt- und Subkategorien visualisiert, da dies gleichzeitig eine höhere Informationsdichte erlaubte und so insbesondere die Verbindung zwischen den Kategorien verdeutlicht wurden. In die im qualitativen Teil entstandene Grundkulisse wurden Kategorien (Haupt- und/oder Subkategorien) eingebaut. Eingebaut wurden die Haupt- bzw. Subkategorien mit den meisten Verbindungen gemäss den Erkenntnissen aus der Quantifizierung. Gewisse Kategorien wurden zudem aufgrund inhaltlicher Überlegungen ausgeschlossen. Bspw. wurde die Kategorie «Tätigkeiten zur Aufrechterhaltung der Sicherheit» ausgeschlossen, da es das Ziel der Schulung ist, diese Tätigkeiten zu erlernen, wobei das Erleben der iVR-Umgebung behilflich sein soll. Die so entstandene Visualisierung von ausgewählten Haupt- und Subkategorien ist in der Abbildung 2.5 ersichtlich. In der Visualisierung bilden alle gelben Felder sowie roten Pfeile Kategorien ab. Es konnten jedoch aufgrund des Umfangs nicht alle Haupt- und Subkategorien abgebildet werden. Die

Visualisierung kommuniziert ein Szenario, in dem ersichtlich ist, was fix ist (Grundkulisse) und was verändert werden kann (Haupt- und Subkategorien). Bezogen auf die Visualisierung konnte in der Umfrage nun gefragt werden, was die wichtigsten Kategorien in der Grundkulisse sind und was die wichtigsten Ausprägungen (Subkategorien) darstellen, wenn die Verbindung zu den anderen Kategorien in die Bewertung miteinbezogen wurde. Diese Visualisierung half somit die Verbindungen der Kategorien untereinander mental stabil zu halten. Dies wird als Weg gesehen, der es erlaubt, die Haupt- und Subkategorien bewerten zu lassen, ohne die Interaktionen abzubilden, gleichzeitig aber die Interaktionen nicht zu vernachlässigen.

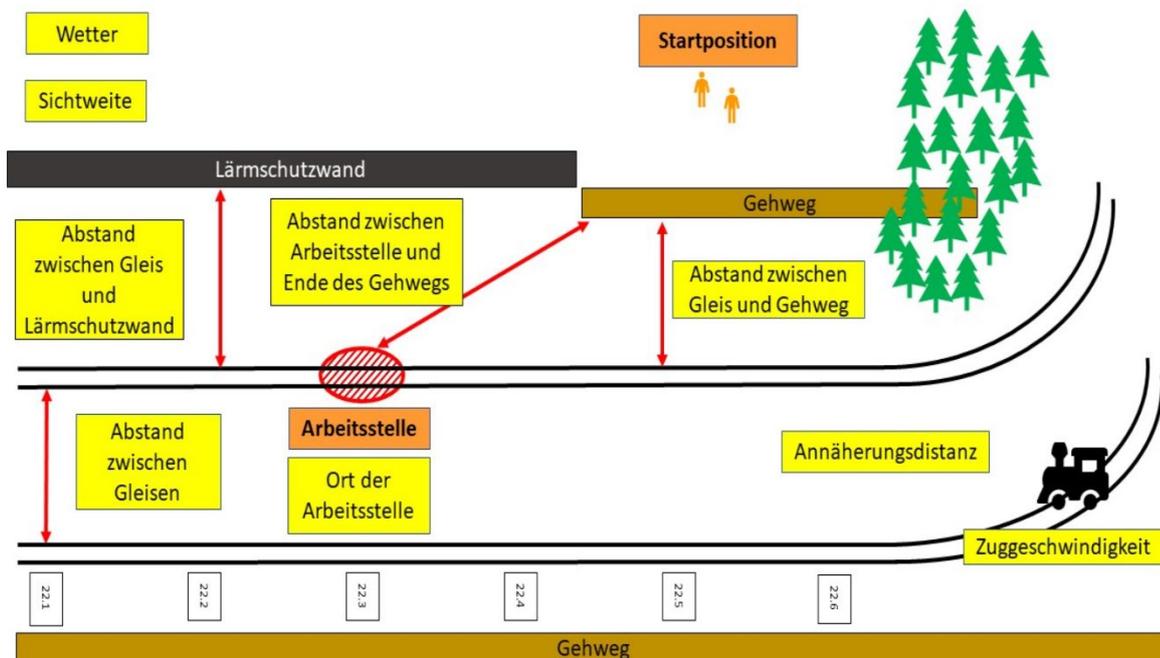


Abbildung 2.5. Grundkulisse angereichert mit den Hauptkategorien

2.3.2 Befragung und Auswertung

In der Befragung wurden die ausgewählten Kategorien nach ihrer Relevanz untereinander bewertet (siehe Fragetyp 1, Abbildung 2.6).

Fragetyp 1: Rating der Relevanz der Kategorien untereinander

- Wetter
- Sichtweite
- Abstand zwischen Gleis und Lärmschutzwand
- Abstand zwischen Arbeitsstelle und Ende des Gehwegs
- Abstand zwischen den Gleisen
- Ort der Arbeitsstelle
- Abstand zwischen Gleis und Gehweg
- Annäherungsdistanz
- Zuggeschwindigkeit

Fragetyp 2: Rating der Relevanz von Ausprägungen der Kategorien

- Nebel
- Saisonabhängiges Tageslicht (bspw. Blendung durch Sonne oder Dämmerung)
- starker Regen
- Schneefall
- Nacht / Dunkelheit

Fragetyp 3: Abfrage von relevanten Ranges von Abständen

	Abstand 1 in Meter	Abstand 2 in Meter	Abstand 3 in Meter	Abstand 4 in Meter	Abstand 5 in Meter
Abstand zwischen Gleis und Lärmschutzwand	<input type="text"/>				
Abstand zwischen Arbeitsstelle und Ende des Gehwegs	<input type="text"/>				
Abstand zwischen den Gleisen	<input type="text"/>				
Abstand zwischen Gleis und Gehweg	<input type="text"/>				

Abbildung 2.6. Fragetypen aus der quantitativen Umfrage

Unterschiedliche Ausprägungen von Kategorien (bspw. Wetterzustände) wurden ebenfalls untereinander bewertet (siehe Tabelle 2.4 für alle Ausprägungen, die untereinander bewertet wurden).

Tabelle 2.4: Unterschiedliche Ausprägungen der Kategorien

Rang	Kategorie	Ausprägungen der Kategorien
1	Ort Arbeitsstelle	Muss frei definiert werden können
2	Zuggeschwindigkeit	40 km/h bis 160 km/h
3	Annäherungsdistanz	100m bis 900m
4	Sichtweite	50m bis 1000m
5	Abstand zwischen Gleis und Lärmschutzwand	0.5m bis 5m
6	Abstand zwischen Gleis und Gehweg	0.5m bis 10m
7	Wetter	1) Nebel 2) Dunkelheit 3) Schneefall 4) Saisonales Tageslicht 5) Starker Regen
8	Abstand zwischen Gleisen	0.5m bis 5m
9	Abstand zw. Arbeitsstelle & Gehweg	0.5m bis 50m

Die Bewertung der Ausprägungen untereinander erfolgte in der Form des gleichen Fragetyps (siehe Fragetyp 2, Abbildung 2.6). Handelte es sich um intervallskalierte Ausprägungen (bspw. Abstände oder Geschwindigkeit) wurde mittels offenem Fragetyp nach den Werten gefragt, die eingestellt werden können sollten (siehe Fragetyp 3, Abbildung 2.6). Im Zentrum der Kategorien-Bewertung stand somit die Frage, was die relevanteste Hauptkategorie ist. Ist bspw. die Kategorie Wetter wichtiger als die Sichtweite für die Sicherheit bei einer Gleisbegehung? Diese Priorisierung sollte eine Reduktion der Kategorien und somit auch der Verbindungen untereinander auf der Grundlage den gewonnenen Zahlenwerten ermöglichen. Die Daten aus dem Fragebogen wurden deskriptiv ausgewertet. Hierfür wurde basierend auf den Summenscores aus den Werten der Bewertung eine Rangreihenfolge der Kategorien gebildet, um die Interpretationen entsprechend mit Zahlenwerten zu begründen. Bei den Fragen nach der Ausprägung von intervallskalierten Ausprägungen wurde der tiefste Wert und der höchste Wert abgelesen, um die Ranges zu bilden. Die Umfrage bot in der Form von freien Textfeldern zudem die Möglichkeit, fehlende Kategorien zu ergänzen oder auf unnötige Kategorien hinzuweisen. Diese qualitativen Ergänzungen flossen aufgrund ihrer Inkonsistenz nicht in die weitere Interpretation ein.

2.3.3 Ergebnisse Fragebogen

Die quantitativen Ergebnisse lassen sich in drei Teile gliedern, die sich an den drei unterschiedlichen Fragetypen orientieren. Insgesamt wurde der Online-Fragebogen von 17 Mitarbeitenden der SBB ausgefüllt. In Tabelle 2.5 sind die Ergebnisse aus Fragetyp 1 abgebildet. Die Kategorien wurden dabei nach dem Wert der Summenscores in eine Rangfolge gebracht. Die Summenscores setzen sich zusammen aus den summierten Rängen, die den Kategorien durch die Experten zugewiesen wurden. Dabei hat Rang 1 den Wert 1, Rang 2 den Wert 2 etc. Je kleiner der Summenscore, desto relevanter wurde die betreffende Kategorie von den Experten bezogen auf das vorgeschlagene Grundscenario bewertet. Wie in Tabelle 2.5 ersichtlich ist, wird der Ort der Arbeitsstelle als für den Selbstschutz am relevantesten eingeschätzt. Ebenso werden die Zuggeschwindigkeit, Annäherungsdistanz und Sichtweite als eher relevant eingeschätzt. Als weniger relevant wurde zum Beispiel das Wetter, der Abstand zwischen den Gleisen und der Abstand zwischen Arbeitsstelle und Gehweg bewertet.

Tabelle 2.5: Rangfolge zur Relevanz der Kategorien im Grundszenario

Rang	Kategorie	Summenscore
1	Ort Arbeitsstelle	49
2	Zuggeschwindigkeit	55
3	Annäherungsdistanz	56
4	Sichtweite	57
5	Abstand zwischen Gleis und Lärmschutzwand	68
6	Abstand zwischen Gleis und Gehweg	93
7	Wetter	100
8	Abstand zwischen Gleisen	103
9	Abstand zw. Arbeitsstelle & Gehweg	115

In Tabelle 2.6 werden die Ergebnisse des zweiten Fragetyps abgebildet. Auch bei diesem Fragetyp war die Rangfolge nach Relevanz entscheidend. Zu lesen ist die Tabelle 5 von oben nach unten. So wurde innerhalb der Hauptkategorie "Wetter" Nebel als relevanteste Subkategorie bewertet, gefolgt von Dunkelheit und Schneefall. Die Rangfolgen entstanden wiederum durch die Auswertung via Summenscores, die in Klammer notiert sind. Neben den Rangfolgen hatten die Teilnehmenden die Möglichkeit, weitere wichtige Zustände in einem offenen Kommentarfeld hinzuzufügen bzw. darauf hinzuweisen, welche der vorgeschlagenen Zustände weggelassen werden sollten. Da es sich bei den Kommentaren um einzelne Sichtweisen und Vorschläge handelt, werden sie in den Ergebnissen nicht explizit berücksichtigt.

Tabelle 2.7 bietet abschliessend eine Übersicht über die Ergebnisse des Fragetyps 3. Unter «Rang» werden jeweils die tiefsten und höchsten Angaben von relevanten Abstands- bzw. Geschwindigkeitsmassen zusammengefasst. So sollen zum Beispiel beim Abstand zwischen dem Gleis und der Lärmschutzwand Werte zwischen 0.50 und 5.00 Meter eingestellt werden können, da sich in diesem Bereich wichtige Unterschiede bezogen auf die Beurteilung der Situation ergeben. Wenn der Abstand grösser als 5.00 Meter ist, dann hat das laut Experten keinen entscheidenden Einfluss mehr auf das Szenario.

Tabelle 2.6: Rangfolge zur Relevanz der Ausprägung der Kategorien im Grundszenario

Rang	Wetter	Gleis-Zustand	Hindernisse	Auftrag	Ablenkungen
1	Nebel (31)	Betriebsgleis (36)	Brücke (42)	Kontrolle / Messung an mehreren Stellen (37)	Smartphone / Tablet (40)
2	Dunkelheit (39)	Arbeitsgleis (45)	Lärmschutz-wand (43)	Anziehen von Befestigung / Schraube (45)	Arbeitskolleg*innen (42)
3	Schneefall (51)	Ungesperrt (49)	Tunnel (49)	Messung am Gleis (58)	Telefon klingelt (42)
4	Saisonales Tageslicht (57)	Doppelspur (55)	Wald (60)	längere Distanz dem Gleis folgen (59)	Baustellenlärm (53)
5	Starker Regen (62)	Gesperrt (59)	Baustelle (67)	Überqueren von mehreren Gleisen (63)	Gegenstand fällt zu Boden (70)
6	-	Rangiergleis (77)	Bäume mit Blättern (79)	Startposition zur Arbeitsstelle gelangen (83)	Regenfall auf den Helm (89)
7	-	-	Bagger (79)	Gleis verlassen (85)	-
8	-	-	Bäume ohne Blätter (101)	-	-

Anmerkung. In Klammer steht jeweils der Summenscore. Tiefer Score = Hohe Relevanz

Tabelle 2.7: Rang zur Einstellung von Abständen und der Zuggeschwindigkeit

	Range	
	von	bis
Abstand zwischen Gleis und Lärmschutzwand	0.50	5.00
Abstand zwischen Arbeitsstelle und Ende des Gehwegs	0.00	50.00
Abstand zwischen den Gleisen	0.50	5.00
Abstand zwischen Gleis und Gehweg	0.50	10.00
Sichtweite	50.00	1000.00
Annäherungsdistanz	100.00	910.00
Zuggeschwindigkeit	40.00	160.00

Anmerkung. Abstände und Sichtweite in Meter, Zuggeschwindigkeit in km/h

2.4 Integration Workshop und Befragung

Die folgende Integration der Ergebnisse soll als Katalog dienen, der bei der anschliessenden Programmierung des Szenarios Anhaltspunkte zur Auswahl, Priorisierung und Bedeutung der Ergebnisse geben soll, um die notwendige Komplexitätsreduktion auf die Ergebnisse der Studie stützen zu können. Aufgrund der Bewertung der iVR-Brillenmodelle durch die Schulungsexperten der SBB wird für die technische Umsetzung empfohlen, die Meta Quest 2 im Unterrichtsformat einzusetzen. Dies wird aus Sicht des FHNW-Teams auf den hohen Tragekomfort, die geringe Infrastruktur und die einfache Erlernbarkeit des Umgangs zurückgeführt.

Zum besseren Verständnis der Ergebnisse zum Inhalt bzw. dem Bild des VR-Szenarios wurde die Abbildung 2.7 erstellt. Anhand dieser Abbildung wird im Folgenden erklärt, welche Kategorie über welche

Relevanz verfügt und was wiederum deren relevanteste Ausprägungen sind (siehe Tabelle 2.8). Diese Priorisierung der zu programmierenden Objekte, als auch die Priorisierung deren Ausprägungen, soll eine effiziente und ressourcenschonende Programmierung des Szenarios erlauben, die sich auf den Erkenntnissen der vorliegenden Arbeit abstützt.

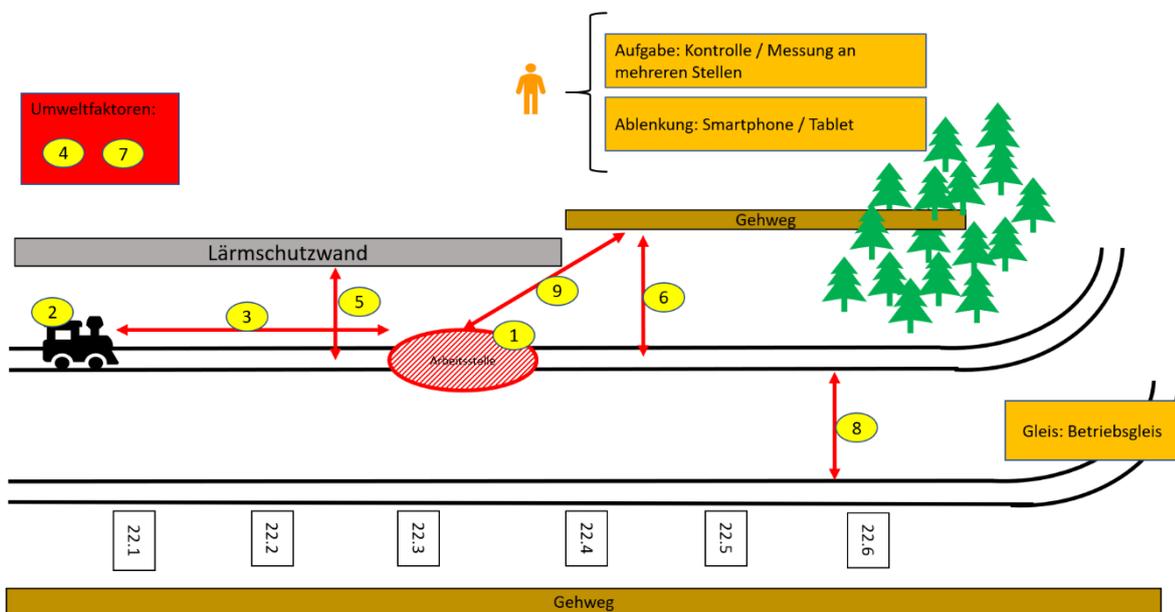


Abbildung 2.7. VR-Szenario basierend auf der Integration der qualitativen und quantitativen Ergebnisse

Tabelle 2.8: Relevanz der Kategorien und deren Ausprägungen

Rang	Kategorie	Ausprägungen der Kategorien
1	Ort Arbeitsstelle	Muss frei definiert werden können
2	Zuggeschwindigkeit	40 km/h bis 160 km/h
3	Annäherungsdistanz	100m bis 900m
4	Sichtweite	50m bis 1000m
5	Abstand zwischen Gleis und Lärmschutzwand	0.5m bis 5m
6	Abstand zwischen Gleis und Gehweg	0.5m bis 10m
7	Wetter	1) Nebel 2) Dunkelheit 3) Schneefall 4) Saisonales Tageslicht 5) Starker Regen
8	Abstand zwischen Gleisen	0.5m bis 5m
9	Abstand zw. Arbeitsstelle & Gehweg	0.5m bis 50m

Grundsätzlich können die in Abbildung 2.7 gelb markierten Elemente verschiedene Zustände einnehmen. Für die Programmierung wird empfohlen, mit Nummer 1 (relevanteste Kategorie) anzufangen und sich dann innerhalb der verfügbaren Ressourcen Rang für Rang runterzuarbeiten. Der Ort der Arbeitsstelle wurde von den Experten als relevanteste Kategorie bewertet. Da sich die Arbeitsstelle je nach Auftrag und Situation ändern kann. Die Arbeitsstelle muss frei verschoben werden können und verfügt aufgrund ihrer Natur nicht über klar definierbare Zustände. Die Zuggeschwindigkeit wird als zweitwichtigste Kategorie bewertet und soll zwischen 40 und 160 km/h variieren können. Die Annäherungsdistanz (100 bis 900 Meter) und Sichtweite (50 bis 1000 Meter) werden ebenfalls als eher relevant bewertet. Die verschiedenen Abstände auf den Rängen 5, 6, 8 und 9 können sich in den in Tabelle 2.8 zusammengefassten Ranges bewegen. Je nach Ressourcen im Programmierungsteam und im Hinblick auf die Komplexitätsreduktion (Sweller, 2005; 2010) könnten aber gewisse Abstände auch fix definiert werden. Zudem gibt es für das Wetter auf Rang 7 noch unterschiedliche Ausprägungen, die ebenfalls in einer Rangfolge nach Relevanz vorgeschlagen werden.

Aus Gründen der Komplexitätsreduktion wurden für die drei orangenen Kategorien Auftrag, Ablenkung und Gleiszustand jeweils die relevantesten Ausprägungen gewählt. Grund dafür war, dass die Festlegung einer Kategorie mit nominaler Ausprägung zu mehr Komplexitätsreduktion führt als die Festlegung von Kategorien mit intervallskalierter Ausprägung. So soll die Person damit beauftragt werden, eine Kontrolle oder Messung vorzunehmen. Ebenfalls soll virtuell ein Smartphone oder Tablet simuliert werden, um als Ablenkung zu dienen. Bei den abgebildeten Gleis soll es sich um Betriebsgleise handeln. Zwar wurde bei der Kategorie Hindernisse die Brücke als am relevantesten bewertet, da aber in Szenarien aus den qualitativen Fokusgruppen immer eine Lärmschutzwand als Hindernis gewählt wurde und diese bei der quantitativen Umfrage mit nur einem Punkt mehr auf den zweiten Rang gewählt wurde, wurde die Grundkulisse diesbezüglich nicht angepasst.

2.5 Diskussion

Im Rahmen des vorliegenden Projektauftrages wurde in dieser ersten Teilstudie der Frage nachgegangen, wie eine iVR-Schulungsumgebung gestaltet werden muss, damit sie den inhaltlichen Ansprüchen der Schulung zum Selbstschutz im Gleisbereich bei der SBB genügt. Ausserdem sollte das methodische Vorgehen zur Beantwortung dieser Frage Hinweise darüber geben, wie künftige Forschungen in diesem Bereich vorgehen sollen. Nachfolgend werden die Ergebnisse der vorliegenden Forschungsarbeit diskutiert sowie die Fragestellungen dazu beantwortet. Es werden danach die Limitationen der Arbeit aufgezeigt und Hinweise für künftige Forschungsarbeit in diesem Bereich gegeben.

Beim Einsatz von iVR im Schulungssetting wird versucht, eine gewünschte Situation möglichst realitätsgetreu virtuell zu simulieren. Dabei ist es möglich, mit der virtuellen Welt zu agieren (Mikropoulos &

Natsis, 2011) und verschiedene Sinnessysteme gleichzeitig anzusprechen (Webster, 2016), was aus biopsychologischer Sicht den Lerneffekt erhöht (Raithel et al., 2007). Dieser erhöhte Lerneffekt hat somit das Potenzial, Schulungen entsprechend anzureichern und im Fall der SBB zum Beispiel die bisherigen Spielzeug-Simulationen im Schulungszimmer mit realitätsgetreuen Simulationen zu untermauern. Ebenfalls ist es mit iVR möglich, besonders gefährliche Situationen zu simulieren (Burghofer & Lackner, 2014), die so bei echten Gleisbegehungen nicht zu verantworten wären. Das VR-Simulationstraining bietet zudem die Möglichkeit, sich sicherheitsrelevante Verhaltensweisen in der VR-Umgebung anzueignen, welche später in der Berufspraxis angewendet werden (Burghofer & Lackner, 2014). Das wiederum erklärt den Transfereffekt von Handlungsweisen, welche durch den Einsatz iVR-Trainings erzeugt werden (Hamilton et al., 2020). Gemäss den Experten, welche am Workshop der qualitativen Teilstudie teilgenommen haben, können jedoch reale Berufserfahrungen nicht durch virtuelle ersetzt werden. Obwohl sich diese Aussage der Experten mit den Erkenntnissen von Gaba (2004) widersprechen, ist trotzdem davon auszugehen, dass diese Art von Training die praktischen Erfahrungen von Personen ergänzen bzw. verstärken können. Ein weiterer Vorteil des iVR-Trainings ist auch beispielsweise die Wiederholbarkeit, geben von unmittelbarem Feedback sowie die Variierbarkeit der Szenarien (Burghofer & Lackner, 2014). Durch den Einsatz der Meta Quest 2 sind die Mitarbeitenden in der Lage, die Schulungen einerseits zu einem späteren Zeitpunkt und andererseits auch Zuhause auszuführen. Es wird kein Klassenzimmer mehr benötigt, sondern lediglich Internetzugang sowie die entsprechende Hard- und Software, um die virtuelle Lernumgebung betreten zu können. Die iVR-Technologie erlaubt es zudem den Personen sich allein der iVR-Simulation zu stellen oder sich virtuell zu begegnen und gemeinsam Lernszenarien zu bewältigen (Brehmer & Becker, 2017). Die Präferenz der Workshop-Teilnehmer zeigte, dass ein iVR-Classroom für das iVR-Simulationstraining gewünscht wird. So kann beispielsweise eine Person die Aufgaben des virtuellen Simulationstrainings absolvieren, während der Rest der Klasse sowie Kursleitung die Handlungen der aktiven Person verfolgen. Als eine Form der Rückmeldung kann das gezeigte Verhalten anschliessend gemeinsam reflektiert werden oder als Diskussionsplattform dienen. Innerhalb der iVR-Umgebung ist es für die SBB möglich, die einzelnen Elemente in ihrer Ausprägung zu variieren. Das heisst, von der einen zur anderen Sekunde kann Nebel in der virtuellen Umgebung erscheinen oder der Zug fährt mit einer schnelleren Geschwindigkeit an den Personen vorbei. Diese Variierbarkeit ermöglicht, dass ein Grundszenario trotzdem vielfältig präsentiert werden kann und die Mitarbeitenden ihre erlernten Handlungsweisen unter verschiedenen Bedingungen anwenden können. Auch dieser Aspekt spricht wieder für den Transfereffekt, welche durch den iVR-Einsatz generiert wird. Bei der Frage, wie iVR-Inhalte für das virtuelle Gleisbegehungstraining gewonnen werden können, wies die Literatur keinen klaren und konsistenten Weg auf. Was die Studien jedoch gemeinsam hatten, war der Miteinbezug von Expertenwissen. Solche

Fachpersonen können in Fokusgruppen oder Interviews aufzeigen, was für Faktoren und Elemente innerhalb eines Simulationstrainings relevant sind und wie sich diese wechselseitig beeinflussen. Anhand des explorativen Vorgehens im qualitativen Teil des Mixed Methods-Design wurde in einem ersten Schritt eine Fülle an Datenmaterial zu den inhaltlichen Ansprüchen der Experten an eine Schulungsumgebung gesammelt sowie Haupt- und Subkategorien gebildet (siehe Tabelle 2.6). Dabei waren es vor allem die Interaktionen unter den gefundenen Kategorien, welche die grosse Komplexität des Inhalts ausmachten. Um den Hinweisen aus der Literatur zur Cognitive Load Theory (Sweller, 2005; 2010) gerecht zu werden und den Rahmen der zur Verfügung stehenden Ressourcen für den Projektauftrag nicht zu sprengen, sollte das Datenmaterial verdichtet sowie die relevanten Inhalte herausgearbeitet werden. Diese Verdichtung von Inhalten dient auch der Verhinderung einer kognitiven Überforderung des Gehirns (Mayer & Moreno, 2003). Mit einer angemessenen Anzahl von präsentierten Inhalten innerhalb des iVR-Trainings, ist davon auszugehen, dass weniger Sinneseindrücke generiert sowie angesprochen werden und so der Lerneffekt nicht beeinträchtigt wird. Um die Inhalte zu reduzieren, wurde eine quantitative Online-Befragung durchgeführt. Anhand der so gewonnenen Zahlenwerte zur Relevanz der einzelnen Kategorien bezüglich des Selbstschutzes, konnten die wichtigsten Kategorien und deren Ausprägungen identifiziert werden. Die Integration der qualitativen und quantitativen Ergebnisse in ein zu programmierendes Szenario (siehe Abbildung 2.7) stellt das zentrale Ergebnis der Arbeit dar und soll als Gestaltungsanleitung bzw. Katalog zur Programmierung des Bildes der iVR unter Berücksichtigung der optimalen Technik dienen. Gleichzeitig beantwortet es die erste Fragestellung der vorliegenden Arbeit: Wie bewerten SBB-Sicherheitsauszubildende iVR-Trainingselemente für zukünftige iVR-Trainings? Bei der Bewertung des erarbeiteten (virtuellen) Grundschemas anhand der Gestaltungsprinzipien nach Jonassen (1994) sind bereits jetzt einige dieser Prinzipien umgesetzt worden. Einerseits ist es durch die iVR-Technologie sowie durch die verschiedenen Ausprägungen, welche die einzelnen VR-Elemente annehmen können, möglich, die Realität auf vielfältige Weise darzustellen. Andererseits erlaubt es das Szenario die Präsentation authentischer Aufgaben und die Schaffung einer realitätsnahen und fallbezogenen Lernumgebung. Ebenso bietet das Unterrichtsetting die Gelegenheit für Personen, ihr Handeln innerhalb der Unterrichtsklasse zu reflektieren und unterstützt dabei den kollaborativen Aufbau von Wissen durch soziale Prozesse (Jonassen, 1994). Die zweite Fragestellung widmete sich dem methodischen Vorgehen: Wie gelingt es konzeptionelle Hinweise für die Erstellung einer iVR-Umgebung zum Thema der Gleissicherheit zu erarbeiten? Das methodische Vorgehen in seiner präsentierten Form eignet sich gut zur Erarbeitung von Gestaltungshinweisen für die Programmierung einer iVR-Schulungsumgebung. Der explorative Zugang im qualitativen Teil ist jedoch nur dann nötig, wenn über die konkreten Inhalte des Bildes für den Zweck des Projektes nur wenig bekannt ist. Handelt es sich um ein Themenfeld, in dem eine umfangreiche Literatur verfügbar ist, könnte

der qualitative Teil der Methodik durch eine ausführliche Literaturrecherche ersetzt werden. Ebenfalls wären Überlegungen interessant, inwiefern ein Clustern oder Validieren von gefundenen Kategorien nicht bereits im qualitativen Teil im Rahmen des Workshops stattfinden könnten. Was jedoch als generelle Herausforderung eingeschätzt wird, ist der Umstand, dass es in einem Szenario, welches anpassbar sein soll und mit welchem auch interagiert wird, relativ schnell zu einer Vielzahl von Interaktionen kommen kann. Die vorliegende Arbeit leistet in dieser Hinsicht einen ersten methodischen Beitrag. Denn mit der Abfrage der gefundenen Kategorien mit den meisten Verbindungen in visualisierter Weise und der anschließenden Bewertung der Ausprägungen dieser Kategorien, wird eine Möglichkeit geschaffen, wie eine methodisch geleitete Reduktion von Haupt- und Subkategorien unter Berücksichtigung der Interaktionen zwischen den Kategorien vorgenommen werden kann. Die Reduktion der Kategorien führt folglich automatisch auch zur Reduktion der Interaktionen und damit zur Reduktion der Komplexität, weil weniger veränderbare Zustände im Szenario abgebildet werden. Gerade im Rahmen eines Schulungssettings sollte gezielt auf die relevantesten Faktoren eingegangen werden können.

2.6 Limitationen und Ausblick

Eine Limitation der gewählten Methodik war, dass aus der quantitativen Befragung lediglich Rangfolgen als Ergebnis hervorgingen. Zwar lieferten diese Rangfolgen wertvolle Hinweise zur Priorisierung der betreffenden Elemente, jedoch gab es durch die rein deskriptive Auswertung der Daten keine konkrete Aussage darüber, wo ein Schnitt gezogen werden kann und welche Elemente genügend relevant sind, um in das finale Szenario aufgenommen zu werden. Zudem entstanden durch die offenen Kommentare im Fragebogen zum Teil fachliche Rückfragen. Eine erneute Expertenbefragung im weiteren Verlauf des betreffenden SBB-Projektes im Anschluss an die vorliegende Arbeit sollte sicherlich noch mehr Klarheit bezüglich der quantitativen Ergebnisse und deren Bedeutung bringen. Des Weiteren wurden im vorgeschlagenen Grundscenario die veränderbaren Kategorien und deren Zustände jeweils einzeln beurteilt. In der Realität gibt es aber sicher Situationen, in denen sich gewisse Zustände überlagern (Interferenzen). Also bspw. könnten Nebel und Starkregen gleichzeitig auftreten. Um den Transfereffekt zu verbessern, müsste sicherlich in dieser Hinsicht noch einen Schritt weiter gegangen werden. Zum Schluss muss darauf hingewiesen werden, dass die vorgeschlagenen Gestaltungshinweise und das generelle Vorgehen sehr individuell auf die SBB und den Projektauftrag zugeschnitten und deshalb nicht generalisierbar sind. Weiterführende Forschung sollte sich dem Transfer und der Weiterentwicklung des methodischen Vorgehens zur Reduktion der Interaktionen annehmen. Insbesondere sollte der Fokus auch daraufgelegt werden, wann die Reduktionen zu stark sind und wichtige

Interaktionen verloren gehen. Wo befinden sich die Kippunkte bezogen auf die Interaktionen, die das Szenario überhaupt erst spannend machen, aber aufgrund ressourcentechnischer Limitationen nicht in ihrer realen Komplexität abgebildet werden können? Zur Überprüfung der hier präsentierten Methode wären weiterführende Untersuchungen denkbar.

3 Studie II – Bedienungsparadigma

3.1 Einleitung und Orientierung

Eine virtuelle Welt beinhaltet oft eine Interaktionsebene: Verschiedene Werkzeuge oder Gegenstände können benutzt werden, die Umgebung kann angepasst werden (Ziegler et al., 2020) und eine Teleportfunktion hilft Distanzen in iVR zurückzulegen. Dies, obwohl man sich gleichzeitig in der realen Welt nicht oder nur minimal bewegt (Buttussi & Chittaro, 2021). Wenn die Steuerung von Funktionen in iVR unklar oder unkomfortabel ist, kann dies zu Frustration führen (Faric et al., 2019). Somit kann der Steuerung eine wichtige Schlüsselrolle beim Erleben von iVR zufallen. Die am häufigsten verwendeten Steuerungen in iVR sind Handbewegungen (Gesten) oder über die Knöpfe der iVR-Controller (siehe Abbildung 3.3). Handgesten gelten als flexibler, weil man keinen Controller halten muss und nicht von der Anzahl der Knöpfe abhängig ist. Allerdings fehlt das haptische Feedback, die Präzision ist schlechter (Huang, Liu, Lee & Yeh, 2021) und die Bedienung ist schwieriger und weniger zuverlässig (Nyssönen, Helle, Lehtonen & Smed, 2022). Je nach Art der Gesten kann es zu einer Muskelermüdung kommen (Kamińska, Zwoliński & Laska-Leśniewicz, 2022). Während einige AutorInnen untersuchten, welche Gesten besser geeignet sind (siehe z.B. Rempel et al., 2014), untersuchten Huang et al. (2021) hybride Ansätze, indem sie eine hybride Eingabe, d.h. eine Kombination aus Controllern und Gesten, ausprobierten. Ihre Ergebnisse zeigten, dass Bewegungen, die Präzision und Stabilität erfordern, mit Controllern durchgeführt werden sollten. Daraus lässt sich schliessen, dass Teleportfunktionen, bei denen man ein Ziel anvisieren muss, und Funktionen, bei denen man Objekte halten muss, besser mit einem Controller funktionieren. Ausserdem brauchen die Nutzer zusätzliche Hinweise, damit klar ist, ob etwas über eine Geste oder den Controller gesteuert wird. Kangas et al. (2022) untersuchten, ob sich bestimmte Eingabearten (nur Handbewegungen, Controller und Trigger-Button, Controller und Grab-Button) unterschiedlich verhalten. Es gab signifikant mehr Fehlerhafte Interaktionen in der Aufgabe zwischen den beiden Controller-Eingabe-Typen und sie unterschieden sich auch in der subjektiven Bewertung. Zum Beispiel fühlten sich Controller und Grab-Button natürlicher an und gaben den Benutzenden mehr Freiheit, während Controller und Trigger-Button klarer und genauer bewertet wurden. Dies legt die Vermutung nahe, dass die Position einer Taste auf dem Controller sowie subjektive Präferenzen einen Einfluss auf die Aufgabenerfüllung haben. Es stellt sich nun die Frage, ob die Zuordnung eines Buttons auf dem Controller zu einer Funktion (Mapping) sich von einer Bedingung «Ohne Mapping» unterscheidet, in welcher ein beliebiger Button gedrückt werden kann und die Funktion durch das System erkannt wird. Denkbar wäre beispielsweise, dass beim Drücken eines beliebigen Buttons in der Nähe von oder beim Zielen auf interaktionsfähige Objekte, eine Interaktion mit dem Objekt ausgelöst wird. Andernfalls wäre eine Navigationsfunktion wie Teleport aktiv. Die

Handlungstheorie (Frese & Zapf, 1994) erlaubt es, Annahmen über diese beiden Bedingungen zu treffen, die mit Usability-Massnahmen untersucht werden können.

Um die iVR-Controller nutzen zu können, müssen die Benutzenden das Steuerungssystem erlernen, was auf unterschiedliche Weise geschehen kann: Anleitungen, Erklärungen, Trial & Error und Tutorials sind einige von ihnen. Da Zeit ein begrenzter Faktor beim späteren Training sein wird, fällt Trial & Error weg. Ein geführtes Tutorial scheint am geeignetsten zu sein, da es Bilder beinhalten kann und die Lernleistung gegenüber Wörtern in einer Anleitung oder Erklärung besser ist (siehe z.B. Schumacher & Martin, 2013). In dem Tutorial sollten die Nutzer vor allem lernen, wie sie mit den iVR-Controller im virtuellen Raum navigieren und interagieren können. Daher ist es wichtig, herauszufinden, was dieses Lernen negativ oder positiv beeinflussen kann. Im Hinblick auf das motorische Lernen ist es sinnvoll, durch das Vorführen von Bewegungen zu lernen (Witte, 2018). Ein animierter virtueller Charakter (folgend virtueller Agent genannt) kann diese Aufgabe im virtuellen Raum übernehmen (Lampen, Liersch & Lehwald, 2020) und den Nutzern zusätzlich verbale Informationen abgeben (Graesser, Moreno, Marineau, Adcock & Olney, 2003). Auf der Ebene der Verarbeitung versucht die Theorie der kognitiven Belastung zu erklären, wie die Belastung der Informationsverarbeitung bei Lernaufgaben reduziert werden kann (Sweller, van Merriënboer & Paas, 2019). Auch die Motivation mit ihren intrinsischen und extrinsischen Komponenten hat einen Einfluss auf den Lernprozess (Deci, Koestner & Ryan, 2001) und sollte berücksichtigt werden. User Experience kann schliesslich dazu verwendet werden, den gesamten Prozess der iVR-Erfahrung zu bewerten, einschliesslich der Frage, wie unterhaltsam oder grafisch ansprechend sie war.

3.2 Forschungsfragen

Aus der Einleitung ergibt sich ein klareres Bild, mit was sich die Studie befasste. Vorausgehend sollten aus bestehender Literatur Hinweise zur Gestaltung eines intuitiven iVR-Tutorials herausgezogen und umgesetzt werden. Operationalisierbare theoretische Konstrukte, welche die Grundlage der Gestaltungshinweise darstellten, sollten dabei als messbare Ebene im Menschen verwendet werden. Anhand zwei zu variierender Einflussgrössen sollten dann in einem Feldexperiment Unterschiede im Erleben eines iVR-Tutorials festgestellt werden. Als eine dieser Einflussgrössen wurde die Steuerung der Controller verwendet. Eine Ausprägung dieser Steuerung stellte dabei eine pro Button zugeordnete Funktion («Mit Mapping») dar, während in der anderen Ausprägung ein beliebiger Button gedrückt werden konnte («Ohne Mapping»). Die breite Streuung und unterschiedliche Erfahrung der Schulungsabsolvierenden stellte die zweite Einflussgrösse dar, Auszubildende und Berufstätige wurden unterschieden.

Folgende Forschungsfragen wurden abgeleitet, wobei das Vorgehen bei der Beantwortung von F1 explorativ ist.

F1: Auf welchen Ebenen wird der Einfluss von Berufstatus und Mapping vs. kein Mapping bei Menschen am besten abgebildet?

F2: Welches Mapping auf den VR-Controllern ist für die Nutzenden intuitiver und effektiver zu erleben?

F3: Unterscheiden sich die Zielgruppen Berufstätige und Auszubildende hinsichtlich des in F2 beschriebenen Erlebens?

3.3 Methode und Sample

Es wurde ein 2x2 Between-Subjects-Design mit zwei unterschiedlichen Tastenbelegungen auf dem iVR-Controller («Mit Mapping» vs. «Ohne Mapping») und unterschiedlichen Ausprägungen im Alter bzw. im Ausbildungsstatus verwendet (Berufstätig vs. Ausbildung suchend/Auszubildend) verwendet. Als abhängige Variablen wurden Bearbeitungszeit, fehlerhafte Interaktionen, Anzahl gedrückte Buttons und Items aus den Fragebögen von SUS, GUESS, API, KIM und MLCSV (siehe Abschnitt 3.3.3) verwendet. Vorerfahrung mit iVR diente als Kontrollvariable. Die benötigte Anzahl Teilnehmende wurde a priori durch die Software G*Power (Faul, 2020) berechnet. Für das gewählte statistische Verfahren (siehe 3.5) wurde mit einer geschätzten moderaten Effektstärke von $f^2(V) = .15$, einer Teststärke von $.80$, vier Gruppen und acht abhängigen Variablen eine Stichprobengröße von $N = 60$ berechnet. Diese Stichprobengröße konnte mit einer ausgeglichenen Gruppengröße (jeweils $n = 15$) erreicht werden. Von den 60 Teilnehmenden waren 30 Berufstätige, die eine Sicherheitsschulung bei der SBB absolviert hatten, und 30 Auszubildende ohne Berufserfahrung, die meisten von ihnen wurden auf einer Berufsmesse für die Studie rekrutiert. Das Durchschnittsalter der Berufstätigen betrug 33.97 Jahre mit einer Spanne von 22 bis 56 Jahren, während das Durchschnittsalter der Auszubildenden 14.57 Jahre mit einer Spanne von 13 bis 20 Jahren betrug. Drei Teilnehmende wurden nicht in die Stichprobe inkludiert, da sie aufgrund technischer Probleme das Tutorial vorzeitig beendeten. Alle Teilnehmenden nahmen freiwillig an der Studie teil und erhielten keine finanzielle Entschädigung. Von allen Teilnehmenden wurde vor der Teilnahme das Einverständnis eingeholt.

3.3.1 Material

Nach der Handlungstheorie (Frese & Zapf, 1994) braucht eine Handlung immer ein Ziel. Um dieses Ziel Schritt für Schritt zu erreichen, werden Handlungen in kleinere, hierarchisch und sequenziell aufgebaute Teilhandlungen unterteilt. In den meisten Fällen werden zunächst kognitive Ressourcen angesprochen, die eine Handlung planen, bevor die Handlungen dann muskulär ausgeführt werden. Nach

wiederholter Ausführung einer Handlung kann diese automatisiert werden, so dass sie weniger kognitive Ressourcen erfordert. Um Handlungen in iVR zu erlernen, basiert der Lernprozess im Tutorial auf diesem Prinzip. Um Prozesse auf der intellektuellen Ebene zu reduzieren, wurde den Nutzern ein Ziel vorgegeben. Dies unterscheidet sich von einer Lernumgebung, in welcher Benutzende Dinge frei ausprobieren und sich selbst Ziele setzen können. Durch begrenzt verfügbare Zeit wurde jedoch dieses Vorgehen gewählt. Das Ziel war, insgesamt vier Funktionen zu erlernen: Teleportieren, Werkzeuge greifen, ein Werkzeug mit beiden Händen bedienen (blinkende Schrauben am Gleis festdrehen) und mit einem Funkgerät einen Funkspruch beantworten. Für jede dieser vier Funktionen wurden sechs Teilziele definiert, die eine Schritt-für-Schritt-Anleitung darstellten (siehe Abbildung 3.1).

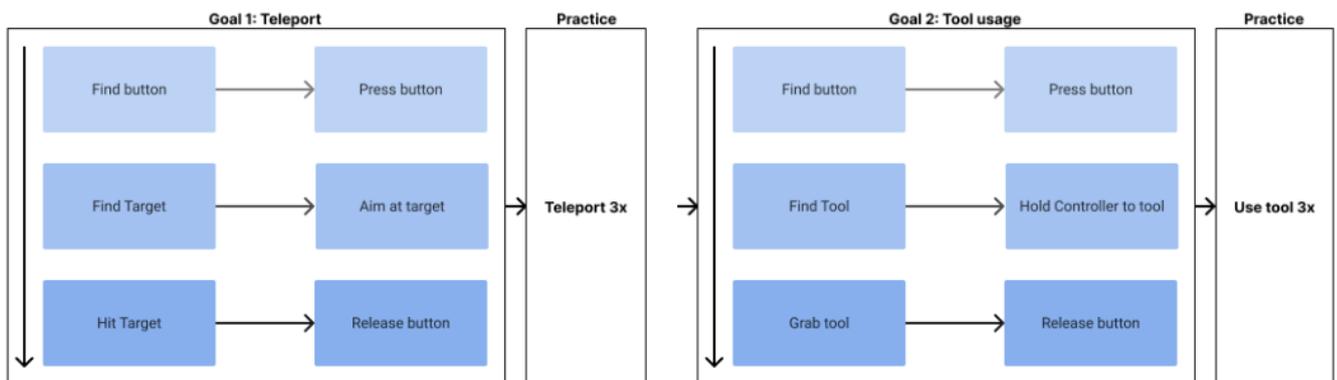


Abbildung 3.1: Hierarchisch-sequenzielle Unterteilung der Ziele

Es wurde darauf geachtet, dass die Aufgaben der Teilziele zu Beginn so einfach wie möglich sind (z.B. Button finden, Button drücken), keine Kombinationen beinhalten und das Ziel klar vorgegeben ist. So sollen Fehler auf höheren intellektuellen Ebenen (Frese & Zapf, 1994) ausgeschlossen werden. Auch nach der Cognitive Load Theory (CLT) sollte diese Aufgabenteilung Vorteile bringen: niedrige CL Werte sollte durch niedrige Komplexität (Klepsch et al., 2017) erreicht werden und gleichzeitig konnte so Vorwissen (Klepsch et al., 2017) aufgebaut werden. Anschliessend wurde die Aufgabenschwierigkeit schrittweise erhöht (z.B. Ziel finden und anvisieren, dann anvisieren und Knopf drücken) und dann die Aufgabe als Ganzes dreimal geübt, bis eine neue Funktion als nächstes Ziel präsentiert wurde. Durch die dreimalige Wiederholung wurde eine Verinnerlichung bzw. Routinierung der Abläufe angestrebt (Frese & Zapf, 1994). Abbildung 3.1 zeigt die beschriebene hierarchisch-sequenzielle Struktur der Ziele am Beispiel von zwei Zielen. Weitere Inputs aus der Literatur kamen hinzu: Die Kombination von Medien und Modalitäten (Noetel et al., 2022) wurde durch Textbausteine und einer virtuellen Agentin, die den Text verbal wiedergab (Graesser et al., 2003), umgesetzt. Ausserdem zeigte sie den Benutzenden parallel alle Aufgaben, einschliesslich der zu drückenden Knöpfe, wodurch eine Imitation (Lampen et al., 2020) angestrebt werden sollte. Die Agentin wurde als Entität mit Warnweste und Helm

dargestellt, vor denen die beiden VR-Controller schweben (siehe Abbildung 3.2). Auf einen Körper wurde bewusst verzichtet, damit die Benutzenden nicht annehmen, sie befänden sich in einer 3rd-Person-Ansicht. Für die virtuelle Agentin wurde eine weibliche Stimme verwendet, die standardmässig von einer Text-to-Speech-Schnittstelle erzeugt wird, den Benutzenden wurde sie als Coach Emma vorgestellt.

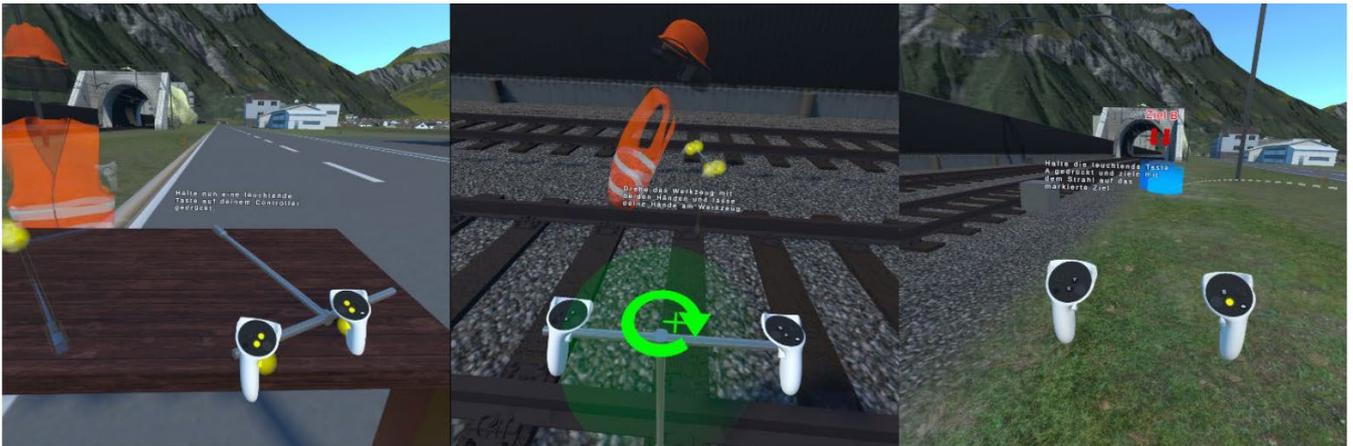


Abbildung 3.2: Der virtuelle Agent, visuelle Hinweise, Textbausteine, hervorgehobene Schaltflächen und Controller

Dekorative Elemente in der Umgebung wurden, entgegengesetzt den Empfehlungen von Noetel et al. (2022) nicht entfernt, da das Tutorial eine reale Situation in der Nähe der Bahngleise nachstellen sollte. Noetel et al. (2022) entsprechend wurden daher visuelle Hinweise implementiert, z.B. wurden wie von Sayer (2004) vorgeschlagen Navigationsmarkierungen eingeblendet, wenn sich die Benutzenden bewegen sollten oder interaktionsfähige Objekte wurden farblich hervorgehoben. Auf eine Karte als Navigationshilfe wurde aufgrund der überschaubaren Umgebung verzichtet (siehe auch Sayer, 2004).

3.3.2 Hardware

Es wurde eine Meta Quest 2 im Standalone-Modus (ohne Verbindung zu einem PC) als HMD verwendet. Zusätzlich wurden Brillen-Spacer und Elite-Straps verwendet, um den Komfort zu erhöhen und das Tragen einer Brille für Brillentragende zu unterstützen. Die Displayauflösung der Quest 2 beträgt pro Auge 1832 x 1920 Pixel, mit einer Bildwiederholungsrate von 90 Hertz. Die beiden Linsen können in drei Distanzen zueinander eingestellt werden: 58mm, 63mm, 68mm (Facebook, n. d.). Bei allen Teilnehmenden wurde die mittlere Einstellung eingestellt. Audio nahmen die Teilnehmenden über die eingebauten Raumklang-Lautsprecher wahr. Die Controller, die die Benutzenden in der Hand hielten, waren in der virtuellen Umgebung sichtbar (siehe Abbildung 3.3). Wie von Domin et al. (2017) vorgeschlagen wurde darauf verzichtet, virtuelle Hände abzubilden. Die Tasten für jede zu verwendende

Bedingung sind in Abbildung 3.3 zu sehen und wurden in der virtuellen Welt in heller Farbe hervorgehoben (ersichtlich in Abbildung 3.2). Dies diente der Verbesserung der Interaktionstechnik und der Hervorhebung von Anweisungen, beides zur Senkung der CL. Über die Hardware konnten die Usability-Messgrößen Zeit, Tastendrucke und fehlerhafte Interaktionen direkt erfasst und exportiert werden. Als Fehlerhafte Interaktion wurde geloggt, wenn Aufforderungen durch die virtuelle Agentin bzw. das System wiederholt werden mussten. Dies geschah beispielsweise, wenn man auf eine markierte Stelle zum Teleportieren zielte, aber den gedrückten Button nicht losliess oder wenn man in der Aufgabe beim Werkzeug greifen das Werkzeug nach Aufforderung nicht losliess. Was nicht als fehlerhafte Interaktion geloggt wurde, waren falsch gedrückte Buttons. Da dies nur in einer Bedingung möglich war, hätte dies zu einer Verzerrung der Daten geführt bzw. wäre nicht auswertbar gewesen.

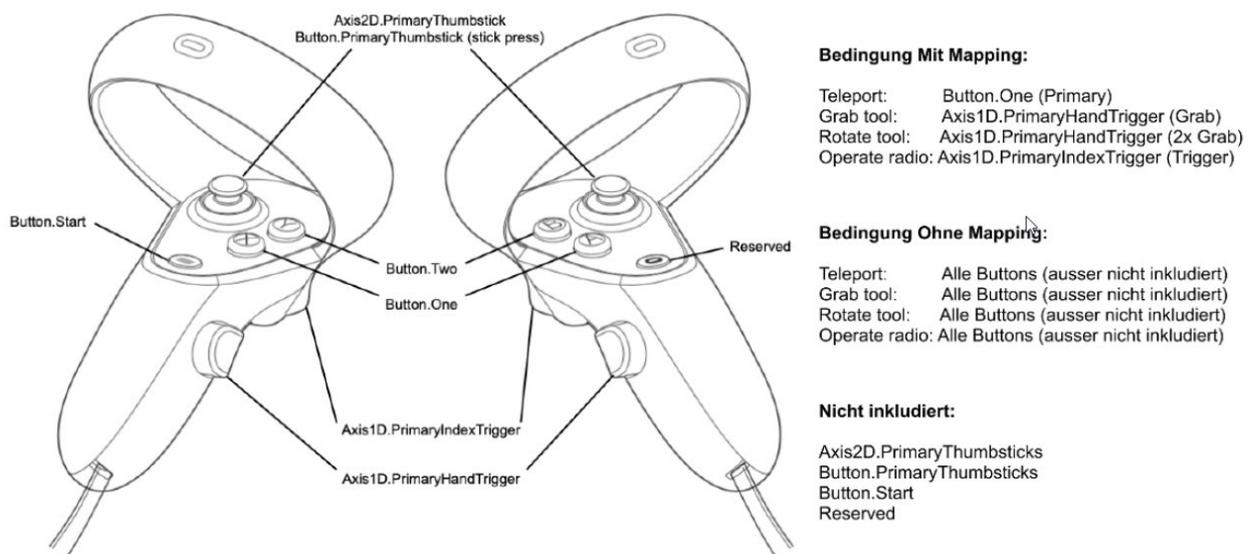


Abbildung 3.3: Controller und Mapping für jede Bedingung. Angepasst an: Meta. (o.D.). [Meta Controller]. Abgerufen am 19. August 2022, von <https://developer.meta.com/documentation/unity/unity-ovrinput/>.

3.3.3 Fragebogen

Mit der Software Enterprise Feedback Suite (Tivian XI GmbH, 2022) wurde ein Fragebogen erstellt, den die Nutzer nach der Erledigung der Aufgaben in der VR ausfüllten. Er enthielt 56 Items zu den Konstrukten Usability, CL, Motivation und UX, zwei offene Fragen zu Vorlieben und Abneigungen sowie fünf Fragen zu Vorerfahrungen und demografischen Daten. Alle Items wurden, sofern nötig, auf Deutsch übersetzt und an den Kontext angepasst. Um den Fragebogen kürzer zu halten, wurden einige Skalen ausgeschlossen. Folgend wird ein Überblick über die Literatur gegeben, aus der die Fragen entnommen wurden und begründet, weshalb einige Items ausgeschlossen wurden. Für die Messung der Usability wurde der 10-Item-SUS verwendet, der eine Reliabilität von Cronbachs $\alpha = 0.91$

aufweist und in der Forschung weit verbreitet ist (Lewis, 2018). Zur Messung der kognitiven Belastung wurden 11 Items des MCLSVE-Fragebogens verwendet, der eine Reliabilität von .85 aufweist (Andersen & Makransky, 2021). Es wurden nur Items zur Messung der unterschiedlichen CL (ECL Instructions, Interactions und Environment) verwendet, da diese für die Studie als relevant empfunden wurden. Beispiele der Items wären «The instructions and/or explanations used in the simulation were very unclear. », «The interaction technique used in the simulation was very unclear. » und «The virtual environment was full of irrelevant content.». Zur Bewertung des virtuellen Agenten wurden die Unterkonstrukte "Facilitating Learning" und "Credible" aus dem Fragebogen Agent Persona Instrument (Baylor & Ryu, 2003) herangezogen, die eine Reliabilität von Cronbachs $\alpha = .85$ bzw. $.65$ aufweisen und die Nützlichkeit der vermittelten Information bewerten (Baylor & Ryu, 2003). Weil mit der virtuellen Agentin nicht auf eine affektive Interaktion abgezielt wurde, wurden die Items "Human-Like" und "Engaging" ausgeschlossen. Ein Beispiel-Item für Facilitating Learning aus dem Fragebogen lautet «The agent focused me on the relevant information», für Credible «The agent was knowledgeable.». Für die Messung der Motivation wurden «Interesse / Vergnügen», «Wahrgenommene Kompetenz» und «Druck / Anspannung» des KIM (Wilde et al., 2009) verwendet, das eine Reliabilität von Cronbachs $\alpha = .79$ hat, mit Ausnahme des Konstrukts «Druck / Anspannung», das nur eine Reliabilität von Cronbachs $\alpha = .54$ hat. Trotzdem wurde dieses Konstrukt aufgenommen, da das Training der SBB, in welchem das Tutorial stattfinden soll, in kurzer Zeit durchgeführt wird und abschliessend eine Prüfung stattfindet - eine angespannte Situation ist also gut vorstellbar. Das Konstrukt "Wahrgenommene Wahlmöglichkeit" wurde hingegen weggelassen, da das Tutorial angeleitet wurde, die Nutzer keinen Einfluss auf den Prozessablauf hatten und somit keine Varianz in den Antworten erwartet wurde. Ein Beispielitem aus dem KIM betreffend Interesse lautet «Ich fand die Tätigkeit in der Ausstellung sehr interessant». Für die UX wurden 10 Items der Unterkonstrukte «Usability / Playability», «Enjoyment», «Personal Gratification» und «Play Engrossment» des GUESS-18 Fragebogens gewählt, welcher eine Reliabilität von Cronbach's $\alpha = .79$ hat (Keebler et al., 2020). Die Konstrukte "Narratives", "Audio Aesthetics", "Social Connectivity" und "Creative Freedom" wurden nicht einbezogen, da sie für das Tutorial als nicht relevant empfunden wurden. Der vollständige Fragebogen ist in Anhang B einsehbar.

3.3.4 Ablauf

Die Studie fand an insgesamt sieben Tagen an drei Standorten (siehe Abbildung 3.4) statt. An Standorten A und B nahmen Berufstätige Personen teil, welche am selben Tag eine Sicherheitsschulung der SBB besuchten. Die Räumlichkeiten boten Platz für neun (Standort A) bzw. zehn (Standort B) gleichzeitig Teilnehmende, was allerdings nie der Fall war. Standort C befand sich an einer Berufswahlmesse, hier nahmen Schülerinnen und Schüler sowie Lernende teil, welche die Messe besuchten. Der

Standort unterschied sich von A und B durch begrenztere Platzverhältnisse (max. zwei gleichzeitige Teilnehmende) sowie einen höheren Lärmpegel. Technische Einschränkungen wie schlechte Internetverbindung, zu helles Licht oder reflektierende Materialien, welche die Sensoren des HMD stören, gab es an keinem der drei Standorte.



Abbildung 3.4: Durchführungsstandorte A, B und C

An allen Standorten erfolgte derselbe Ablauf. Alle Teilnehmenden wurden über den Projekthintergrund und das Ziel der Studie informiert. Bezüglich der iVR wurde erklärt, dass es eine räumliche Begrenzung gibt und wie diese aussieht, wenn sie überschritten wird, wie das HMD aufgesetzt wird und dass das HMD bei Übelkeit abgenommen werden sollte. Die Teilnehmenden zogen dann per Los eine der beiden Testbedingungen «Mit Mapping» oder «Ohne Mapping» und konnten sich danach einen der Plätze frei aussuchen, an welchem ein HMD platziert war. In Bedingung «Mit Mapping» wurden alle Funktionen einer festen Taste zugewiesen (siehe Abbildung 5), während es in Bedingung «Ohne Mapping» keine Rolle spielte, welche der Tasten gedrückt wurde: In der Standardeinstellung war die Teleportfunktion aktiv und das System erkannte, ob man sich in der Nähe eines interaktionsfähigen Objekts befand und ob man auf dieses Objekt zielte. Wenn dies der Fall war, wurde die Funktion, je nach Objekt, automatisch auf Greifen oder Funkgerät bedienen umgestellt. Beim Starten der Applikation unterstützte die Studienleitung die Teilnehmenden, anschliessend bearbeiteten die Nutzer die Aufgaben im iVR-Tutorial selbstständig und ohne Zeitlimit. In der ersten Aufgabe mussten sich die Teilnehmenden über die Teleportfunktion fortbewegen. Dazu musste der Primary Button bzw. ein beliebiger Button gedrückt, gleichzeitig auf ein Ziel gezielt und dann der Button losgelassen werden. Danach folgte die zweite Aufgabe, in welcher ein auf einem Tisch liegendes Werkzeug gegriffen und wieder losgelassen werden musste. Dafür musste einer der Controller an das Werkzeug gehalten und der Grip Button bzw. ein beliebiger Button gedrückt (bzw. wieder losgelassen) werden. In der dritten Aufgabe musste das zuvor kennengelernte Werkzeug eingesetzt werden. Dazu musste es auf dieselbe Weise wie zuvor gegriffen werden, dann an eine Schraube gehalten und, während beide Hände am Werkzeug griffen, gedreht werden. Schlussendlich musste in der letzten Aufgabe über ein Funkgerät ein Funkspruch

erwidert werden. Das Funkgerät war um den virtuellen Körper gehängt, es wurde sichtbar, wenn man nach unten schaute. Um es zu bedienen, musste einer der beiden Controller an das Funkgerät gehalten werden und der Trigger Button bzw. ein beliebiger Button gedrückt werden. Das System erkannte automatisch, wenn Teilnehmende eine Aufgabe erfolgreich abgeschlossen hatte, und löste die nächste Aufgabe aus. Falls eine fehlerhafte Interaktion gemacht wurde bzw. keine Interaktion stattfand wurde die Aufgabeninstruktion wiederholt. Nach der letzten Aufgabe forderte das System die Teilnehmenden auf, das HMD abzunehmen. Abschliessend füllten die Nutzer den Fragebogen auf ihren eigenen Smartphones oder an einem zur Verfügung gestellten Laptop aus.

3.3.5 Datenaufbereitung und -analyse

Die Rohdaten aus dem HMD-Export wurden in Excel (Microsoft, 2022) durch den Einsatz von Formeln so aufbereitet, dass *Bearbeitungszeit (s)*, *Anzahl gedrückte Buttons* und *Fehlerhafte Interaktionen* als Variable ausgelesen und einer Session-ID zugeordnet werden konnten. In der Entwicklungsumgebung RStudio (RStudio, 2022) und der Statistiksoftware R (R Core Team, 2021) mit den R-Paketen tidyverse (Wickham et al., 2019) und haven (Wickham, Miller & Smith, 2022), wurden die HMD-Daten und die Daten aus dem Fragebogen zusammengeführt, bereinigt und Scores für die Konstrukte der abhängigen Variablen aus dem Fragebogen berechnet. Die Daten wurden danach für die weiteren Auswertungen exportiert. Für die Datenanalyse wurde eine zwei faktorielle MANOVA (Multivariate Varianzanalyse) gewählt, welche nach linearen Trends in mehreren abhängigen Variablen sucht. Ihr Vorteil gegenüber mehreren separaten ANOVA's besteht darin, dass die Kumulation von alpha-Fehlern nicht berücksichtigt werden muss, sowie ersichtlich wird, welchen kombinierten Effekt die abhängigen Variablen als Folge der unabhängigen Variablen haben (Hahs-Vaughn, 2016). Field (2017) weist darauf hin, dass nicht einfach beliebige abhängige Variablen in eine MANOVA einbezogen werden sollten. Die in dieser Studie einbezogenen abhängigen Variablen haben alle einen Einfluss auf das Erleben des iVR, weshalb dieses Verfahren als geeignet angesehen wird. Durch das Verfahren der MANOVA können Gruppenunterschiede festgestellt werden, welche bei der Betrachtung einzelner abhängiger Variablen eventuell nicht erkennbar wären – trotzdem können signifikante Effekte einzeln untersucht werden (Hahs-Vaughn, 2016). Da die MANOVA mit zwei unabhängigen Variablen berechnet werden wird, lassen sich sowohl Haupt- als auch Interaktionseffekte feststellen. Zudem bieten mehrere unabhängige Variablen in einer MANOVA den Vorteil, dass die Fehlervariation zumindest versucht wird zu reduzieren (Hahs-Vaughn, 2016). Die Datenanalyse wurde mit der Statistiksoftware SPSS 27 (IBM Corp., 2020) durchgeführt. Statistische Abbildungen wurden aus darstellungstechnischen Gründen mit der Statistiksoftware jamovi (jamovi, 2021) oder dem R-Paket tidyverse (Wickham et al., 2019) umgesetzt.

3.4 Ergebnisse

3.4.1 Quantitative Ergebnisse

Die deskriptive Übersicht über die Daten dient zur Exploration der einzelnen Aufgaben des Tutorials und bestimmter Buttons. Diese Daten können nicht mit statistischen Verfahren analysiert werden, da Vorbedingungen nicht gegeben sind. In der Bedingung «Mit Mapping» ist beispielsweise davon auszugehen, dass nicht beliebige Buttons gedrückt werden, sondern die im Tutorial erwähnten. Weiter sind die einzelnen Aufgaben zwar ähnlich aufgebaut, aber nicht in einer Art identisch in welcher Vergleiche Sinn machen würden. Trotzdem können diese Daten für mögliche Erklärungen oder Handlungsempfehlungen einbezogen werden.

Tabelle 3.1 zeigt Mittelwerte und Standardabweichungen der abhängigen Variablen, aufgeteilt nach Bedingung und Berufsstatus. Für *Bearbeitungszeit*, *Fehlerhafte Interaktionen* und *Anzahl gedrückter Buttons* gibt es keinen festgelegten Wertebereich, d.h. theoretisch sind Werte von 0 bis unendlich möglich. Die Gruppe der Berufstätigen mit der Bedingung «Ohne Mapping» fällt durch grössere Standardabweichungen in *Bearbeitungszeit* und *Fehlerhafte Interaktionen* auf. Auch interessant sind die unterschiedlichen Mittelwerte in der *Bearbeitungszeit* (s) zwischen den Gruppen: während die Auszubildenden in der Bedingung «Mit Mapping» höhere Werte erreichen, ist es bei den Berufstätigen die Gruppe der Bedingung «Ohne Mapping». Für die restlichen abgebildeten Variablen sind begrenzte Wertebereiche möglich. In Score SUS können Werte zwischen 0 und 100 erreicht werden. Da der SUS in der Forschung verbreitet ist, existieren Studien, welche bei der Interpretation des Scores helfen (siehe z.B. Bangor, 2009). Ein Wert von ungefähr 70 wird als «Good» eingestuft, während 85 «Excellent» und 100 «Best Imaginable» ist (Bangor, 2009, S. 121). Während in Score *KIM* und Score *MLCSV* Werte zwischen 3 und 15 möglich sind, sind es in Score *API* 2 bis 10. Die Werte in Score *GUESS* bewegen sich zwischen minimal 5 und maximal 25. Anzumerken ist, dass im ungekürzten Fragebogen Werte von 9 bis 63 (siehe Zusatzdokument «GUESS-18 Scoring Guidelines» zu Phan et al., 2016) möglich sind, was für allfällige Vergleiche mit anderen Studien berücksichtigt werden sollte. In diesen Variablen fällt auf, dass die Werte in Score SUS mit über 73 als eher hoch einzuschätzen sind. Auch die Werte in Score *GUESS* und Score *API* sind mit 18.6 respektive 7.24 in der Gruppe mit dem tiefsten Wert, als eher hoch einzustufen. Zusätzlich zu entnehmen ist die Vorerfahrung mit iVR, welche als Kontrollvariable aufgenommen wurde. In ihr zeigten sich ähnlich ausgeprägte Häufigkeitswerte über alle vier Gruppen.

Um herauszufinden welche Buttons im Mittelwert in welchem Teil des Tutorials wie oft gedrückt wurden, wurde eine Aufteilung nach den einzelnen Aufgaben gemacht. Trotz extremer

Standardabweichungen im Mittelwert wurde in Abbildung 3.5 versucht ein Überblick über etwaige Trends zu geben und zusätzlich der Median abgebildet.

Tabelle 3.1: Gruppengrößen, Mittelwerte und Standardabweichungen der abhängigen Variablen sowie gleichzeitig gedrückte Buttons und Vorerfahrung mit iVR, aufgeteilt nach Berufsstatus und Bedingung mit N=60

Variable	Berufsstatus	Bedingung					
		Mit Mapping			Ohne Mapping		
		<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
<i>Bearbeitungszeit (s)</i>							
	Auszubildend	15	401.27	88.90	15	371.27	80.60
	Berufstätig	15	387.80	82.40	15	439.60	143.10
<i>Fehlerhafte Interaktionen</i>							
	Auszubildend	15	31.60	8.50	15	27.8	9.89
	Berufstätig	15	23.30	6.18	15	34.1	16.18
<i>Anzahl gedrückte Buttons</i>							
	Auszubildend	15	164.0	61.12	15	133.8	44.95
	Berufstätig	15	84.33	33.37	15	94.50	38.94
<i>Anzahl gleichzeitig gedrückte Buttons</i>							
	Auszubildend	15	66.30	40.06	15	61.50	30.73
	Berufstätig	15	56.6	32.90	15	53.60	38.60
<i>Score GUESS</i>							
	Auszubildend	15	20.27	2.24	15	18.57	2.27
	Berufstätig	15	19.60	2.68	15	20.03	1.19
<i>Score SUS</i>							
	Auszubildend	15	73.67	14.63	15	78.33	12.42
	Berufstätig	15	84.50	12.86	15	78.00	7.75
<i>Score MLCSV</i>							
	Auszubildend	15	6.91	2.95	15	5.32	2.14
	Berufstätig	15	4.53	1.36	15	4.85	1.15
<i>Score KIM</i>							
	Auszubildend	15	11.07	1.33	15	11.07	1.56
	Berufstätig	15	12.31	1.22	15	11.73	0.88
<i>Score API</i>							
	Auszubildend	15	7.86	1.26	15	7.65	1.06
	Berufstätig	15	7.43	0.90	15	7.24	0.89
<i>Vorerfahrung mit iVR (dichotom)</i>							
		<i>n</i>	<i>Ja (n)</i>	<i>Nein (n)</i>	<i>n</i>	<i>Ja (n)</i>	<i>Nein (n)</i>
	Auszubildend	15	10	5	15	9	6
	Berufstätig	15	9	6	15	7	8

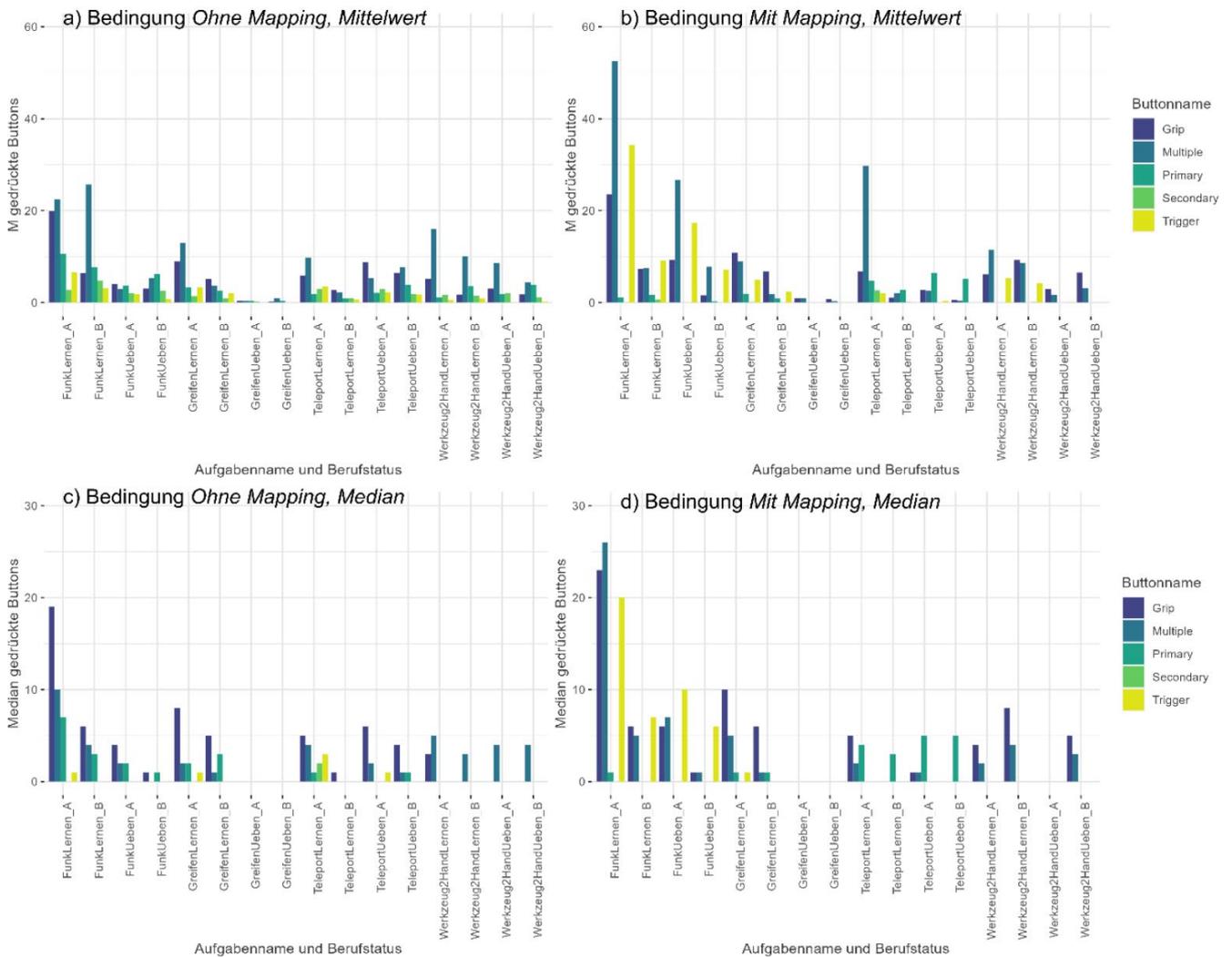


Abbildung 3.5: Mittelwerte und Median Anzahl gedrückte Buttons, aufgeteilt nach Aufgabe und Buttonname. Multiple = mehrere gleichzeitig gedrückte Buttons. A = Auszubildend, B = Berufstätig.

Es fällt auf, dass es über beide Bedingungen im Mittelwert eine augenscheinliche Abnahme in der Betätigung der Buttons vom Erlernen zum Ausprobieren gibt. Weiter werden von Auszubildenden beim Erlernen der Funkgerätbedienung besonders viele Tastendrucke vorgenommen. Interessanterweise finden dabei in der Bedingung «Mit Mapping» bei den Auszubildenden mehr Betätigungen des Grip-Button ($M = 23.53$) statt als in der Bedingung «Ohne Mapping» ($M = 19.86$), obwohl dieser Button dort nicht funktional war. Zusätzlich gibt es in derselben Gruppe und Bedingung einen Höchstmittelwert in der Betätigung des Trigger-Buttons von $M = 34.26$. Dies ist dreieinhalbmal höher als bei den Berufstätigen derselben Bedingung und Aufgabe ($M = 9.20$). Es fällt auch auf, dass Trigger- und Grip-Button oft in derselben Aufgabe beide vermehrt gedrückt wurden. Weiter fällt auf, dass in der Bedingung «Ohne Mapping», also wenn der Button selbst wählbar war, Auszubildende bei allen Aufgaben den Grip-

Button am meisten verwenden. Berufstätige verwenden vermehrt den Primary-Button, nur beim Lernen und Üben von Teleport sowie beim Lernen des Greifens wurden mehr Grip-Betätigungen festgestellt – beim Üben des Greifens wurde dann wieder häufiger der Primary-Button gedrückt. Der Secondary-Button wird in der Bedingung «Ohne Mapping» im Mittel in keiner Aufgabe am häufigsten verwendet, in der anderen Bedingung wurde er beinahe nie benutzt.

Folgend wird das Vorgehen und die Ergebnisse der Prüfung verschiedener Voraussetzungen der MANOVA berichtet. Die MANOVA setzt intervallskalierte abhängige Variablen, nominalskalierte unabhängige Variablen, Unabhängigkeit der Messungen, multivariate Normalverteilung der abhängigen Variablen, Linearität und Homogenität der Kovarianzmatrizen der abhängigen Variablen voraus (Hahs-Vaughn, 2016). Unabhängigkeit der Messungen sowie vorausgesetzte Skalenniveaus der abhängigen bzw. unabhängigen Variablen sind gegeben und werden nicht weiter erläutert. Die Stichprobengröße von $N = 60$ mit vier gleichgrossen Gruppen entspricht statistisch gesehen eher einer kleineren, aber ausgeglichenem Sampling. In kleinen Stichproben haben Variablen mit leptokurtischer Verteilung (Kurtosis > 3.00 ; «Heavy Tails») eine geringe Teststärke (Finch, 2005). Zudem weisen Variablen mit leptokurtischer Verteilung mehr Ausreisser auf. Das Vorhandensein univariater Ausreisser wurde visuell durch Box-Plots überprüft. Es wurden einzelne extreme Ausreisser in den abhängigen Variablen Anzahl gedrückte Buttons und Bearbeitungszeit gefunden. Die Ausreisser wurden vorerst im Datensatz belassen, Messfehler wurden ausgeschlossen. Die Daten wurden visuell mit Q-Q -Plots und über den Shapiro-Wilk-Test auf Normalverteilung geprüft. Der Shapiro-Wilk-Test zeigt sich im Vergleich zu anderen Tests (Kolmogorov-Smirnov, Lilliefors, Anderson-Darling) für alle Verteilungsarten und Stichprobengrößen am aussagekräftigsten (Razali & Wah, 2011). Die Ergebnisse zeigten, dass in den abhängigen Variablen *Bearbeitungszeit*, *Fehlerhafte Interaktionen*, *Anzahl gedrückte Buttons*, *Score SUS* und *Score MLCSV* in einigen Gruppen die Normalverteilungsannahme ($p > .05$) verletzt wurde. Die Verwendung eines parametrischen Tests sollte bei der Verletzung der Normalverteilung kein Problem darstellen, solange die Kovarianzmatrizen gleich sind, und auch beim Vorhandensein von Ausreissern weisen parametrische und nichtparametrische Testverfahren eine ähnliche Teststärke vor (Finch, 2005). Trotzdem wurden die von der Verletzung der Normalverteilungsannahme betroffenen Variablen transformiert, was eine Möglichkeit bei dieser Verletzung darstellt (Hahs-Vaughn, 2016). Durch die Johnson-Transformation (Johnson, 1949) wurden die Daten einer Normalverteilung nähergebracht, die Transformation erfolgte über eine Onlinesoftware (Hemmerich, 2016). Nach der Transformation wies keine Variable mehr eine Verletzung der Normalverteilungsannahme ($p < .001$) vor. Die weiter oben erwähnten Ausreisser wurden durch die Transformationen bzw. den Ausschluss der Variable entfernt. Eine Prüfung auf Multikollinearität ergab Korrelationen zwischen den abhängigen Variablen von maximal $r = .636$, was auf eine Erfüllung dieser Voraussetzung ($r < .80$) hindeutet (siehe z.B. Pituch &

Stevens, 2016 oder Verma, 2015). Als nächstes wurde durch die Berechnung der Mahalanobis-Distanz und den dazugehörigen p-Werten auf das Vorhandensein multivariater Ausreisser geprüft. Es wurden keine multivariaten Ausreisser ($p > .001$) gefunden. Die Prüfung auf eine lineare Beziehung zwischen allen abhängigen Variablen für jede Faktorstufe wurde mittels Streudiagrammen vorgenommen. Nicht alle Punktwolken entsprechen einem eindeutigen linearen Zusammenhang. Im Falle der Verletzung der vorausgesetzten Linearität können Daten transformiert oder ausgeschlossen (Hahs-Vaughn, 2016), oder eine nichtlineares Testverfahren durchgeführt werden (Nimon, 2012), da eine Abweichung der Linearität die Teststärke einer MANOVA reduziert (Tabachnick & Fidell, 2013). Eine Transformation der noch nicht transformierten Variablen wurde geprüft, ergab augenscheinlich aber keine Verbesserung der linearen Beziehungen. Es erfolgte kein Ausschluss von Daten. Um auf Varianzhomogenität zu prüfen, wurde der Levene-Test für Gleichheit der Fehlervarianzen verwendet. Die Homogenität der Fehlervarianzen war in allen abhängigen Variablen gegeben ($p > .05$). Zuletzt wurde auf die Homogenität der Kovarianzmatrizen geprüft, welche laut dem Box-Test gegeben war ($p > .001$). Ateş et al. (2019) schlagen bei einer MANOVA mit zwei abhängigen Variablen, ausgeglichenem Sampling und Verletzung der Normalverteilung aber gegebener Varianzhomogenität die Pillai-Spur als Teststatistik vor. Trotzdem wird auf Unterschiede in Signifikanzwerten zwischen den von SPSS ausgegebenen Testverfahren (Pillai-Spur, Wilks-Lambda, Hotelling-Spur, Grösste charakteristische Wurzel nach Roy) geachtet und bei Vorhandensein berichtet. Die Ergebnisse der MANOVA zeigten einen signifikanten Haupteffekt im Berufsstatus, $F(8, 49, p < .001, \text{partielles } \eta^2 = .481, \text{ Pillai-Spur} = .481)$ sowie einen Interaktionseffekt zwischen Berufsstatus und Bedingung $F(8, 49, p < .034, \text{partielles } \eta^2 = .275, \text{ Pillai-Spur} = .275)$. Für alle abhängigen Variablen wurde Post-hoc eine einfaktorielle ANOVA berechnet. Die Ergebnisse mit Mittelwerten und Standardabweichungen pro Gruppe sind Tabelle 3.2 zu entnehmen. Im Berufsstatus gab es signifikante Unterschiede in den abhängigen Variablen Anzahl gedrückte Buttons JS, Score KIM sowie im Score MLCSV JS. In der Interaktion Berufsstatus*Bedingung gab es in drei abhängigen Variablen, Fehlerhafte Interaktionen JS, Score SUS JS und Score MLCSV JS, signifikante Unterschiede. Abbildung 3.6 zeigt die einzelnen gefundenen signifikanten Haupteffekte grafisch aufbereitet. Abbildung 3.6A zeigt einen Haupteffekt im Berufsstatus in der abhängigen Variable Anzahl gedrückte Buttons JS. Die Gruppe der Auszubildenden drückte signifikant mehr Buttons als die Gruppe der Berufstätigen ($1.10, \text{ptukey} < .001$). In Abbildung 3.6B ist ebenfalls ein Haupteffekt im Berufsstatus in der abhängigen Variablen Score KIM. Hier erreichten Auszubildende signifikant tiefere Werte als Berufstätige ($-0.957, \text{ptukey} = 0.005$).

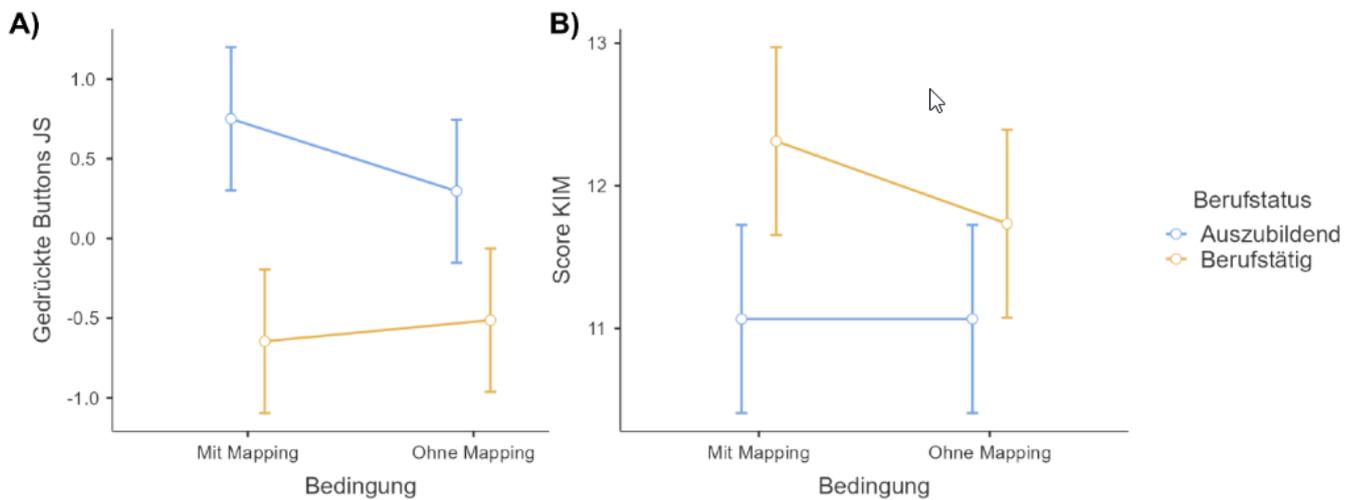


Abbildung 3.6A-B: Ergebnisse der Post-hoc ANOVA. Haupteffekte in A) Gedrückte Buttons JS und B) Score KIM. N = 60. Fehlerbalken entsprechen 95% Konfidenzintervall des Mittelwerts.

Abbildung 3.7 zeigt die gefundenen Interaktionseffekte. Abbildung 3.7A zeigt eine disordinale Interaktion in der abhängigen Variable Fehlerhafte Interaktionen JS. Die Auszubildenden der Bedingung «Mit Mapping» unterscheiden sich durch mehr fehlerhafte Interaktionen von den Berufstätiger der Bedingung «Mit Mapping», allerdings geht durch die Tukey-Korrektur die Signifikanz verloren (0.927 , $p_{\text{tukey}} = .062$). Dasselbe gilt auch für beide Gruppen der Berufstätigen, wobei diejenigen der Bedingung «Mit Mapping» weniger fehlerhafte Interaktionen machen als diejenigen der Bedingung «Ohne Mapping» (-0.93 , $p_{\text{tukey}} = .063$). In Abbildung 3.7B wird eine disordinale Interaktion der abhängigen Variable Score SUS JS dargestellt. Obwohl die Auszubildenden der Bedingung «Mit Mapping» einen tieferen Score als Berufstätige der Bedingung «Mit Mapping» haben, geht auch hier durch die Tukey-Korrektur die Signifikanz verloren (-0.98 , $p_{\text{tukey}} = .053$).

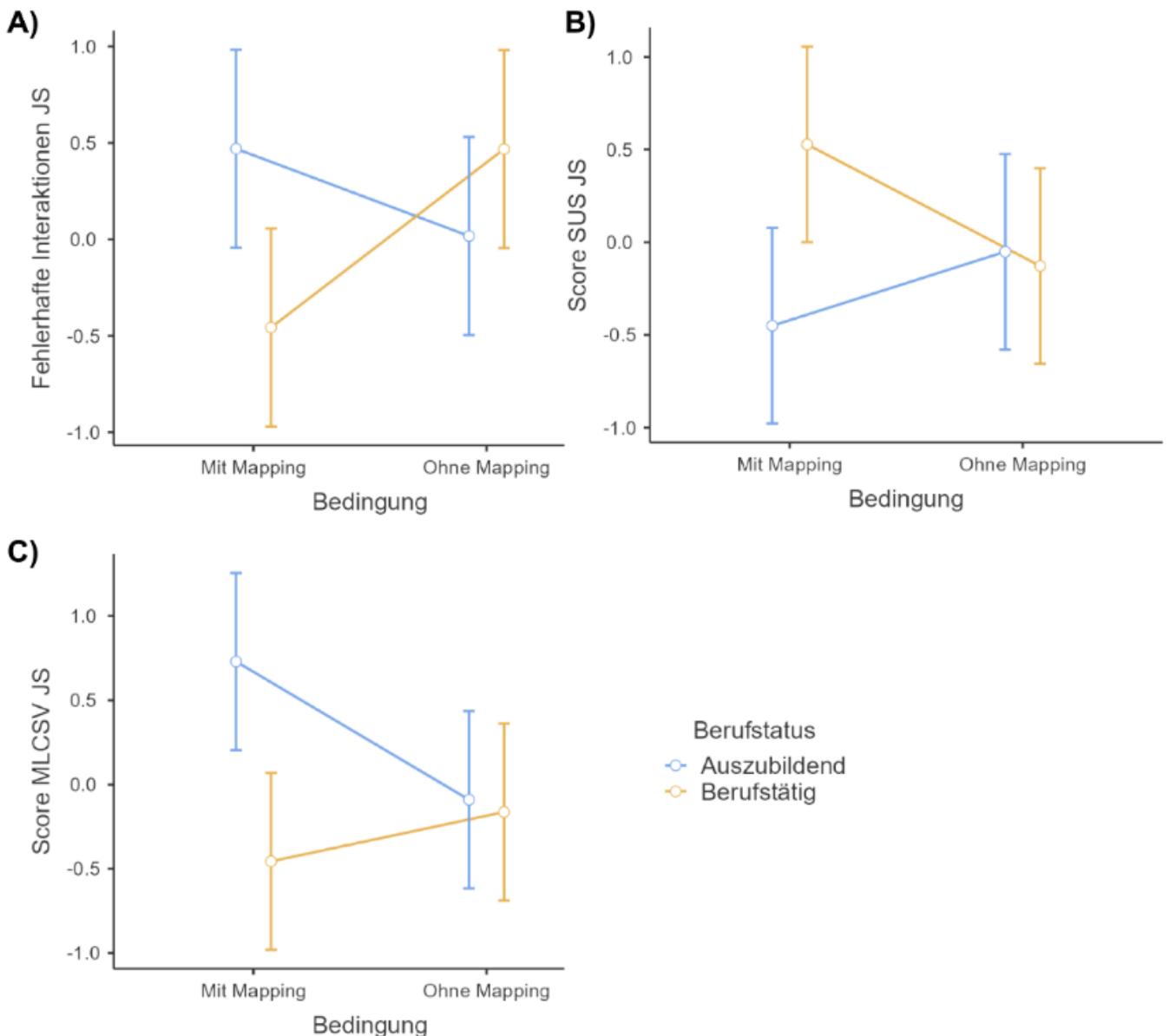


Abbildung 3.7A-C: Ergebnisse der Post-hoc ANOVA. Interaktionseffekte in A) Fehlerhafte Interaktionen JS, B) Score SUS JS und Haupt- als auch Interaktionseffekt in C) Score MLCSV JS. N = 60. Fehlerbalken entsprechen 95% Konfidenzintervall des Mittelwerts.

Die Abbildung 3.7C zeigt schliesslich eine hybride Interaktion in der abhängigen Variable Score MLCSV JS. Dies bedeutet, dass der Haupteffekt (im Berufsstatus) grösser ist als der Interaktionseffekt. Für den Haupteffekt heisst dies, Auszubildende haben einen signifikant höheren Score im MLCSV als Berufstätige (0.63, $p_{\text{tukey}} = .020$). Betrachtet man den Interaktionseffekt, zeigt sich ein signifikanter Unterschied: Auszubildende der Bedingung «Mit Mapping» erreichen einen höheren Score als Berufstätige der Bedingung «Mit Mapping» (1.19, $p_{\text{tukey}} = .012$). Durch die Tukey-Korrektur entfällt die Signifikanz anderer Unterschiede: Auszubildende der Bedingung «Mit Mapping» erreichen einen höheren

Score als die Auszubildenden der anderen Bedingung (0.82, $p_{\text{tukey}} = .133$) und auch als Berufstätige der Bedingung «Ohne Mapping» (0.89, $p_{\text{tukey}} = .087$).

Tabelle 3.2: MANOVA Post-hoc Ergebnisse (einzelne ANOVA) mit N = 60. ANOVA = analysis of variance; BS = Berufsstatus; BD = Bedingung, JS = Johnson-Transformation.

Variable	Mit Mapping			Ohne Mapping			ANOVA			
	n	M	SD	n	M	SD	Effekt	p-Wert	F (1,56)	η^2_p
<i>Bearbeitungszeit JS</i>										
Auszubildend	15	-0.05	1.15	15	-0.39	1.06	BS	.213	1.589	.028
Berufstätig	15	-0.11	0.87	15	0.34	1.04	BD	.833	.045	.001
							BS x BD	.147	2.162	.037
<i>Anzahl gedrückte Buttons JS</i>										
Auszubildend	15	0.75	1.01	15	0.30	0.69	BS	<.001	24.153	.301
Berufstätig	15	-0.65	0.75	15	-0.51	0.98	BD	.476	.514	.009
							BS x BD	.196	1.709	.030
<i>Fehlerhafte Interaktionen JS</i>										
Auszubildend	15	0.47	0.76	15	0.02	0.98	BS	.357	.862	.015
Berufstätig	15	-0.46	0.83	15	0.47	1.31	BD	.361	.850	.015
							BS x BD	.009	7.224	.114
<i>Score SUS JS</i>										
Auszubildend	15	-0.45	1.16	15	-0.05	1.01	BS	.093	2.295	.050
Berufstätig	15	0.53	1.21	15	-0.13	0.59	BD	.629	.237	.004
							BS x BD	.050	4.013	.067
<i>Score GUESS</i>										
Auszubildend	15	20.27	2.24	15	18.58	2.27	BS	.477	.513	.009
Berufstätig	15	19.60	2.68	15	20.03	1.19	BD	.262	1.285	.022
							BS x BD	.061	3.645	.061
<i>Score KIM</i>										
Auszubildend	15	11.07	1.34	15	11.07	1.56	BS	.005	8.465	.131
Berufstätig	15	12.31	1.22	15	11.73	0.88	BD	.383	.772	.014
							BS x BD	.383	.772	.014
<i>Score MLCSV JS</i>										
Auszubildend	15	0.73	1.03	15	-0.09	1.24	BS	.020	5.945	.093
Berufstätig	15	-0.46	0.96	15	-0.16	0.79	BD	.319	1.011	.018
							BS x BD	.038	4.503	.074
<i>Score API</i>										
Auszubildend	15	7.86	1.26	15	7.65	1.06	BS	.120	2.494	.043
Berufstätig	15	7.43	0.90	15	7.24	0.89	BD	.466	2.038	.035
							BS x BD	.970	.001	.000

3.4.2 Qualitative Ergebnisse

Im Fragebogen wurde mit zwei offenen Fragen nach positiven und negativen Eigenschaften des Tutorials gefragt. Insgesamt wurden 73 Antworten gesammelt. Die Antworten wurden mit der Software MAXQDA Analytics Pro 2022 (VERBI GmbH, 2022) ausgewertet. In einem ersten Schritt wurden die Antworten nach Bedingung und Berufsstatus in Gruppen unterteilt. Die Antworten wurden dann exploriert und codiert (markieren von Abschnitten in Textdokumenten), wobei die Codierungen den beiden deduktiven Kategorien Positiv und Negativ zugeteilt wurden. Anschliessend wurden die Codierungen durch deduktive Kategorienbildung weiter unterteilt. Tabelle 3.3 gibt einen Überblick über die Kategorien, die verschiedenen Gruppen und die zugehörige Anzahl an Codierungen.

Tabelle 3.3: Auswertung der qualitativen Antworten nach Kategorien, aufgeteilt nach Gruppen mit Total Anzahl Antworten = 73. Total Anzahl Codierungen = 76.

Kategorie	Codierungen pro Gruppen			
	Mit Mapping Berufstätig	Mit Mapping Auszubildend	Ohne Mapping Berufstätig	Ohne Mapping Auszubildend
Andere/ Allgemein/ Negativ	3	2	7	3
Andere/ Allgemein/ Positiv	5	4	3	5
Steuerung/ Allgemein/ Negativ	2	1	1	0
Steuerung/ Teleport/ Positiv	0	2	0	0
Steuerung/ Bedienung Werkzeug/ Negativ	2	1	0	0
Steuerung/ Bedienung Werkzeug/ Positiv	2	2	3	0
Steuerung /Funkgerät/ Negativ	2	0	0	0
Erklärung/ virtuelle Agentin/ Negativ	1	0	1	1
Erklärung/ virtuelle Agentin / Positiv	3	3	1	1
Visuelle Effekte/ Grafik/ Negativ	1	0	1	1
Visuelle Effekte/ Grafik/ Positiv	5	0	2	0
Soundeffekte/ Audio/ Negativ	0	1	0	1
Soundeffekte/ Audio/ Positiv	1	0	1	1

Allgemein das Tutorial betreffend gab es die meisten negativen Rückmeldungen. Die Aussagen bezogen sich vor allem auf den Inhalt bzw. Umfang des Tutorials: Die Personen hätten gerne noch weitere Tätigkeiten oder Funktionen kennengelernt, folgende Zitate veranschaulichen dies: «Ein bisschen mehr Werkzeug» «mehr verschiedene Situationen» «Etwas weniger oft die gleiche Aufgabe repetieren dafür mehr Verschiedenes machen».

Zwei Personen hatten das HMD nicht korrekt aufgesetzt und hatten deswegen verschwommene Sicht. Das Anziehen des HMD wurde zwar instruiert, einer Aussage nach besteht dort aber noch

Verbesserungspotential: «Meine Vision war z.T. unscharf und /oder doppelt. Erst am Schluss habe ich festgestellt, dass es an der Anpassung der Brille [HMD] an meinen Kopf / Augen lag. Besser instruieren am Anfang!» Die Instruktion fand im selben Raum wie das Experiment statt, die Personen hatten also jeweils ein HMD vor sich. Da die Geräte für viele neu waren, ist eine verminderte Aufmerksamkeit der Instruktion gegenüber vorstellbar. Für zukünftige Studien sollte dies berücksichtigt werden. Auffällig viele negative Rückmeldungen gab es in der Kategorie Allgemein / Negativ von Berufstätigen in der Bedingung «Ohne Mapping». Die meisten positiven Rückmeldungen finden sich auch in allgemeinen Rückmeldungen zum Tutorial. Viele fanden alles gut («Mir hat alles gefallen»), konkretere Äusserungen zeigten, dass die Interaktivität gut ankam («Das man selber etwas machen konnte»; «Tolle Gestaltung in Bezug auf die Aufgaben»), dass die Technik einwandfrei funktionierte («Technik hat sehr gut funktioniert - Gute Grafik - Verständliches Tutorial») oder auch dass Personen Freude hatten, die iVR-Technologie kennenzulernen («Mir hat gefallen das mal da reinzusehen, ist was neues für mich»). Eine eigene Kategorie beinhaltete Äusserungen zur Steuerung. Neben zwei Einzelaussagen zur Ungenauigkeit und Nicht-Funktion eines Buttons, meldeten zwei Personen der Bedingung «Mit Mapping», dass die Buttons, welche in der virtuellen Welt angezeigt werden, teilweise nicht auf der gleichen Position liegen wie auf dem realen Controller: «Mit der Steuerung ist es für unerfahrene User nicht gleich klar welche Buttons zu bedienen sind, da die virtuelle Version nicht exakt dieselbe ist. Ansonsten Top» Die Teleportfunktion betreffend wurden zwei positive Aussagen gemacht, beide von Auszubildenden («Ich fand die Teleportation ziemlich cool»). Zur Werkzeugbedienung (Werkzeug greifen und zweihändig bedienen) wurden sieben positive und drei negative Antworten abgegeben. Positiv wurde die Handhabung bzw. die Interaktion mit dem Werkzeug angesehen: «Etwas zu greifen, das Werkzeug an die Schraube zu halten oder zu drehen funktionierte sehr einfach» «Bedienung Werkzeug mit zwei Händen war sehr gut synchronisiert» Eine negative Rückmeldung aus der Bedingung «Mit Mapping» war, dass gerne eine andere Taste verwendet werden würde: «Das man das „A“ [Primary] auch während den Schrauben zu ziehen verwenden könnte.» Andere Einzelaussagen fielen zum Task Werkzeuge zweihändig bedienen, in welchem Schrauben angezogen werden mussten. Eine Person drehte sich selbst anstelle des Werkzeugs (« Schrauben drehen, man sollte sich nicht drehen müssen») und eine andere verlor das Werkzeug («Ich habe das Werkzeug fallen gelassen und war dann auf mich gestellt, es wieder aufzuheben»). Zur Funkgerätbedienung äusserten sich zwei Berufstätige aus der «Mit Mapping»-Bedingung. Offenbar war die Steuerung nicht klar: «Die Steuerung des Funkgeräts war nicht klar. Musste man nur die Taste drücken oder die gelbe Kiste mit dem Werkzeug berühren und dann drücken?» «Funkgerätbenutzung, so dass es einer Hand zugewiesen wird.»

Für die virtuelle Agentin und Erklärungen im Tutorial wurde ebenfalls eine eigene Kategorie erstellt. Hier sammelten sich acht positive und drei negative Antworten. Bei den negativen Antworten handelt es sich um Einzelaussagen: Das Aussehen der Agentin könne verbessert werden («Als Coach eine Person nicht nur eine Weste»), sie schränkte im Vorgehen ein («Ich durfte erst das Werkzeug in die Hand nehmen als der Virtuelle Coach mir es erlaubte. Das mochte ich nicht wirklich») und der visuelle Textblock überlappte sich. Die positiven Antworten sprachen vor allem die gute Verständlichkeit und Einfachheit der Anweisungen an («Die Anweisungen waren klar und einfach zu befolgen»), aber auch die virtuelle Agentin direkt («Der virtuelle Coach hat mir mit der Erklärung sehr gut geholfen.»; «Dass ich vom Coach gut geleitet wurde.»). Interessanterweise erfolgten die konkreten Antworten zur virtuellen Agentin von Auszubildenden (beide Bedingungen). Ebenfalls von einer auszubildenden Person stammt ein Hinweis, dass der eingeblendete Text hilfreich war: «Ich fand es hilfreich, dass der gesprochene Text auch aufgeschrieben wurde. So konnte ich, wenn ich etwas nicht verstand, nachlesen.»

Betreffend Grafik und visueller Effekte wurden drei negative und sieben positive Antworten in einer eigenen Kategorie erfasst. In den negativen Antworten finden sich die weiter oben besprochenen Personen, welche Probleme mit verschwommener Sicht hatten, da sie das HMD nicht korrekt angezogen hatten. Weiter wurde von einer berufstätigen Person angemerkt, man könnte mehr dynamische Elemente einbauen: «...und etwas mehr visuelle Komponenten einbauen wie z.B. die vorbeifahrenden Züge». Alle der sieben positiven Antworten in dieser Kategorie können berufstätigen Personen zugeordnet werden. Sie sprachen die ansprechende Grafik («Die Grafik war ansprechend und es lief sehr flüssig.») oder auch den als gut empfundenen visuellen Effekt des vorbeifahrenden Zuges an.

In einer letzten Kategorie wurden Antworten zum Audio und Soundeffekten untergebracht. Negativ erfolgten zwei Aussagen von Auszubildenden: der Umgebungslärm war störend bzw. die Stimme der virtuellen Agentin zu leise («Einige Male war es sehr leise die Stimme. Gefallen hat mir das man den Zug gut hörte»). Positiv wurde, wie dem vorherigen Zitat zu entnehmen ist, den Soundeffekt des vorbeifahrenden Zuges wahrgenommen. Eine andere Person unterstützte diese Aussage und eine weitere merkte an, dass die im HMD verbauten Lautsprecher gut waren: «Visuell super. Der Ton war auch beeindruckend ohne Kopfhörer, ohne dass man die Nachbarn hört»

3.5 Diskussion

Die Ergebnisse der Auswertung zeigten, dass es in der Bedingung («Mit Mapping» vs. «Ohne Mapping») keinen Haupteffekt gab, sich *Auszubildende* und *Berufstätige* (Berufstatus) aber signifikant unterscheiden. Zudem gab es einen signifikanten Interaktionseffekt zwischen Berufstatus und Bedingung.

3.5.1 Virtuelle Agentin

Die virtuelle Agentin wurde eingesetzt, um Teilnehmende durch das Tutorial zu leiten, in dem sie die über Textbausteine eingeblendeten Erklärungen verbal von sich gab und Funktionen vorzeigte. Die Bewertung der virtuellen Agentin war durch Auszubildende insignifikant höher als durch Berufstätige. Es wurde von einem signifikanten Unterschied ausgegangen, doch den als hoch zu bewertenden Scores beider Gruppen des Berufstatus nach, war die virtuelle Agentin auch für die Berufstätigen hilfreich. Interessanterweise sind trotz des fehlenden statistischen Unterschiedes die meisten der positiven offenen Antworten der Bedingung «Mit Mapping» zuzuordnen (z.B. «Der virtuelle Coach hat mir mit der Erklärung sehr gut geholfen»). Nicht auszuschliessen ist ein negativer Einfluss der virtuellen Agentin auf den Cognitive Load von Personen mit geringem Vorwissen oder einem interaktiven Avatar (Schöbel et al., 2019). Als Interaktivität kann die Sprachausgabe des Avatars angesehen werden. Diese wurde verwendet, um kognitive Belastung nur durch über Textbausteine eingeblendete Instruktionen zu verringern, was mindestens in Einzelfällen funktionierte («Der virtuelle Coach hat mir mit der Erklärung sehr gut geholfen.»). Einer Aussage nach funktionierte die Entlastung auch andersrum, was die Wichtigkeit der Ansprache mehrerer Modalitäten (Noetel et al., 2022) unterstreicht («Ich fand es hilfreich, dass der gesprochene Text auch aufgeschrieben wurde. So konnte ich, wenn ich etwas nicht verstand, nachlesen.»). Von einer Person bemerkt und negativ angemerkt wurde die Unähnlichkeit der Agentin zu einem Menschen. Laine et al. (2022) berichten Vorteile menschenähnlicher Gestaltung. In dieser Studie wurde bewusst auf eine solche Gestaltung verzichtet, für zukünftige Studien wäre dies jedoch eine Überlegung wert, um beispielsweise die Motivation zu stärken.

3.5.2 Intrinsische Motivation

Die tieferen Motivationswerte bei Auszubildenden lassen sich spekulativ erklären. Erwartet wurde der gegenteilige Effekt, tiefere Werte bei den Berufstätigen, da angenommen wurde, dass sich diese im Rahmen der Sicherheitsschulung in einer angespannten Situation befinden würden (siehe Grolnick & Ryan, 1987). Während dem Durchführen des Feldexperiments zeigten sich aber folgende Punkte: Die Teilnahme am iVR-Tutorial bei der SBB erfolgte freiwillig nach der Schulung und die Personen zeigten sich entspannt. Die Auszubildenden hingegen hielten sich im Rahmen eines ausserschulischen Aufenthalts und durch eine Lehrperson betreut an der Berufsmesse auf. Während der Messe fiel auf, dass sie Informationen sammeln und Aufgaben erledigen mussten (z.B. Interviews führen), um für die Schule eine Arbeit zu absolvieren

Weiter war an der Berufsmesse die Geräuschkulisse im Vergleich zu den Räumlichkeiten bei der SBB erhöht – dies wurde auch im Fragebogen negativ angemerkt. Durch diesen Störfaktor hätte es zu Ablenkung kommen können, woraus die eigene wahrgenommene Kompetenz leiden könnte.

Ein weiterer möglicher Grund ist, dass Auszubildende die Tätigkeiten im Tutorial weniger interessant oder unterhaltsam fanden, da der Arbeitskontext, in welchem das Tutorial ansiedelte, für sie unbekannt war. Novelty-Effekte (siehe Jenö et al., 2019) traten vermutlich auf («Mir hat gefallen das mal da reinzusehen, ist was neues für mich»), ausgeschlossen werden aber Gruppenunterschiede, welche auf einen Novelty-Effekt bzw. Vorerfahrungswerte mit iVR zurückzuführen sind, da die Anzahl an Personen mit iVR-Vorerfahrung zwischen Auszubildenden und Berufstätigen personenmässig ungefähr gleich verteilt war.

3.5.3 Cognitive Load

Der Cognitive Load, also die kognitive Belastung bezieht sich auf die verwendete, limitierte Kapazität des Arbeitsgedächtnis. Auszubildende erreichten einen signifikant höheren Wert im Cognitive Load, sie beanspruchten also mehr der limitierten Ressourcen. Gemessen wurden speziell für iVR unterteilte Masse der ECL (siehe Andersen & Makransky, 2021), welche sich auf bei der Verarbeitung von Lernanweisungen beanspruchte Ressourcen bezieht (Sweller et al., 2019). Eine mögliche Erklärung für die höhere kognitive Belastung ist fehlendes Vorwissen hinsichtlich dem Arbeitskontext. Verschiedene Methoden zur Reduktion kognitiver Belastung beruhen auf dem Vorwissen (siehe Klepsch et al., 2017 oder Noetel et al., 2022). Schöbel et al. (2019) erwähnten, dass virtuelle Agenten bei fehlendem Vorwissen einen hohen Cognitive Load erzeugen können und Noetel et al. (2022) wiesen zudem auf dekorative oder ablenkende Elemente hin, welche ECL steigern können. Für das iVR-Tutorial wurde eine möglichst realistische Gleisumgebungssituation erzeugt. Auszubildende sind höchstwahrscheinlich unvertraut mit Gleisumgebungen und den im Tutorial verwendeten Werkzeugen und Gegenständen, Berufstätige bei der SBB sind damit eher vertraut. Durch die fehlenden Vorkenntnisse könnten die Umgebung und die Werkzeuge bzw. Gegenstände selbst (nicht deren Bedienung) als ablenkende Elemente wahrgenommen worden sein, welche es dann durch kognitive Verarbeitung kennenzulernen galt und somit eine höhere ECL erzeugt wurde. Diese Annahme wird dadurch gestützt, dass von den Berufstätigen vermehrt rückgemeldet wurde, das Tutorial könnte noch umfangreicher hinsichtlich gezeigter Funktionen oder Aufgaben sein, was nicht auf eine maximale kognitive Auslastung hindeutet. Durch Hinweisreize, konkret Markierungen und Hervorheben wichtiger Gegenstände bzw. Buttons, wie durch Noetel et al. (2022) vorgeschlagen, konnte der negative Effekt der Umgebungselemente nicht genügend reduziert werden. Der Interaktionseffekt zwischen Bedingung und Berufsstatus zeigt zusätzlich,

dass die Gruppe der *Auszubildenden* der Bedingung «Mit Mapping» nochmals signifikant höhere CL-Werte vorwies. Dies lässt darauf schliessen, dass die zugeordnete Tastenbelegung, die in den Auszubildenden sowieso schon hohe kognitive Last zusätzlich erhöhte. Zu dieser Gruppe zugeordnete Aussagen aus den offenen Fragen im Fragebogen stützen diese Annahme («Mit der Steuerung ist es für unerfahrene User nicht gleich klar welche Buttons zu bedienen sind, da die virtuelle Version nicht exakt dieselbe ist. Ansonsten Top»). Dies legt auch nahe, dass es durch eine unklare Steuerung zu mehr fehlerhaften Interaktionen kam. Wie Cognitive Load und fehlerhafte Interaktionen zusammenhängen könnten, wird folgend erläutert.

3.5.4 Fehlerhafte Interaktionen und Bearbeitungszeit

Sowohl Fehler als auch Zeit können als objektive Masse der Usability verwendet werden und geben Aufschluss über Effektivität und Effizienz. Der signifikante Interaktionseffekt in der abhängigen Variable *Fehlerhafte Interaktionen* zeigte, dass *Berufstätige* in der Bedingung «Mit Mapping» weniger Fehler machen als die anderen Gruppen. Die Ergebnisse von Plechatá et al. (2019) zeigten weniger Fehler bei jüngeren Personen, weshalb dieses Ergebnis überraschte. Betrachtet man nun den Interaktionseffekt, welcher den Cognitive Load darstellt, fällt die Gruppe der Auszubildenden der Bedingung «Mit Mapping» auf, welche eine höhere Belastung aufweist. Eine mögliche Überlegung wäre nun zu sagen, dass durch den höheren Cognitive Load mehr fehlerhafte Interaktionen getätigt wurden.

Es wird aber aufgrund von mehreren Punkten als wahrscheinlicher angesehen, dass ein umgekehrter Effekt stattfand und durch mehr fehlerhafte Interaktionen ein höherer Cognitive Load erzeugt wurde: Erstens kann erhöhte Cognitive Load zu Spannungen in der Muskulatur führen (Leyman et al., 2004), weshalb von damit einhergehenden Fehlern auf der sensomotorischen Ebene ausgegangen wird. Diese Fehler könnten sich durch eine Varianz in der Anzahl an gedrückten Buttons äussern - ein signifikanter Effekt, welcher mit der Gruppe der erhöhten Cognitive Load in Verbindung gebracht werden könnte, existiert allerdings nicht.

Zweitens wird anhand der Literatur davon ausgegangen, dass Fehler – besonders auf höheren Regulationsebenen – kognitiven Aufwand verursachen (Frese & Zapf, 1994). Die Art der Fehler konnte nicht vom System erfasst werden (siehe 3.3), Fehler höherer Regulationsebenen würden sich nach Frese und Zapf (1994) aber über eine erhöhte Fehlerbehebungszeit äussern. In der Bearbeitungszeit finden sich zwar keine signifikanten Effekte, aber ginge man, wie in der Literatur genannt, von allgemein höheren Bearbeitungszeit durch Berufstätige aus (Plechatá et al., 2019), wäre eine Tendenz zu mehr benötigter Zeit durch Berufstätige in der Bedingung «Ohne Mapping» erkennbar. Diese erhöhte Bearbeitungszeit könnte ein Hinweis auf Fehler höherer Regulationsstufen sein, welche Cognitive Load

erzeugen. Drittens deuten Aussagen aus dem Fragebogen darauf hin, dass durch mangelnde Instruktion Fehler nicht nur auf tieferen, sondern auch auf höheren Regulationsebenen auftraten – eine Person hatte offenbar Mühe, ein fallengelassenes Werkzeug zu finden und eine andere bewegte sich im Kreis, um eine Schraube anzuziehen, statt nur das Werkzeug zu drehen. Beide sind einer höheren Regulationsebene als der sensomotorischen zuzuordnen (siehe Frese & Zapf, 1994).

3.5.5 Anzahl gedrückte Buttons Buttonpräferenzen

Die durch Auszubildende signifikant höhere Anzahl an gedrückten Buttons überraschte, da anhand gelesener Literatur vom Gegenteil ausgegangen wurde. Um einer Erklärung näher zu kommen, wurde hier die nach Aufgabe des Tutorials und nach einzelnen Buttons aufgeteilten deskriptiven Ergebnisse betrachtet.

Es lässt sich ein Grossteil der durch die Auszubildenden gedrückten Buttons der Aufgabe «Funkgerät kennenlernen» bzw. «Funkgerät üben» zuordnen. Auffällig ist, dass bei dieser Aufgabe auch der Wert der gleichzeitig gedrückten Buttons am höchsten war. Schenkt man zusätzlich den Kommentaren aus dem Fragebogen Beachtung, wird klar, dass diese Aufgabe nicht ausreichend erklärt wurde. Eine mögliche Erklärung, weshalb der Unterschied zwischen den jüngeren Auszubildenden und den älteren Berufstätigen zustande kam, bezieht sich auf die Altersunterschiede zu implizitem und explizitem motorischem Lernen (siehe Verneau et al., 2014): Auszubildende profitieren durch explizite Instruktionen. Bei der Funkaufgabe könnte es durch unzureichende Instruktion dazu gekommen sein, dass versucht wurde, Gelerntes aus der vorherigen Aufgabe (Werkzeug mit Grip Button bedienen) anzuwenden. Der benötigte Schritt war jedoch, dass der iVR-Controller an das Funkgerät gehalten werden musste. Berufstätige könnten durch ihre Arbeit bereits Erfahrungen mit Funkgeräten gesammelt haben und daher auf Vorwissen zurückgreifen. Um Trends hinsichtlich bevorzugter Buttons pro Aufgabe in der Bedingung «*Ohne Mapping*» zu erkennen, wurden in 4.1 deskriptive Ergebnisse vorgestellt. Es zeigte sich, dass extreme Standardabweichungen vorhanden waren, was auf unterschiedliche Vorlieben hindeutet. Zum Mittelwert wurde zusätzlich der Median betrachtet. Der Grip Button wurde über alle Aufgaben am häufigsten verwendet, aufgrund seiner ergonomischen Lage am Controller und der Tatsache, dass oft ähnlich hohe Werte für mehrfach gedrückte Buttons einhergingen, lässt keinen eindeutigen Schluss darauf zu, ob er tatsächlich verwendet wurde oder aus Versehen gedrückt wurde.

Für zukünftige Gestaltung von Aufgaben sollte dies beachtet werden – gerade in Funktionen wie Teleportieren, wo man einen Button gezielt loslassen muss, könnte eine versehentlich gedrückter Button das Auslösen der Funktion blockieren. Für die Funkaufgabe und die Greifaufgabe wurde bei Wahlmöglichkeit neben den Grip-Button auch öfters der Primary-Button verwendet (in der Bedingung «Mit

Mapping» war es der Primary-Button). Die Vorliebe für den Primary-Button für die Greifaufgabe wurde zwar auch im Fragebogen angemerkt («[Verbesserungswunsch:] Das man das „A“ auch während den Schrauben zu ziehen verwenden könnte»).

3.5.6 Usability

Neben objektiven Messgrößen der Usability kann diese auch subjektiv erfragt werden. Die subjektiv bewertete Usability wurde durch den Score aus dem SUS gemessen. Die Auszubildenden der Bedingung «Mit Mapping» berichteten einen tieferen Score als die Berufstätigen derselben Bedingung. Die Ergebnisse passen gut in das Bild, welches aus den Effekten der fehlerhaften Interaktionen und des Cognitive Load entstand. Es ist wichtig hier nochmals anzumerken, dass alle vier Gruppen Scores erreichten, welche zwischen «Good» und «Excellent» einzustufen sind (Bangor, 2009). Dies sollte nicht als selbstverständlich empfunden werden, man betrachte die Erkenntnisse aus Domin et al. (2019), dass alle Personen Probleme mit iVR hatten.

3.5.7 User Experience

Die User Experience sollte das gesamte Erlebnis mit iVR, die Wahrnehmungen, Reaktionen und Zufriedenheit der Nutzenden mit iVR erfassen. Es wurden keine signifikanten Unterschiede gefunden, weder in der Bedingung noch im Berufsstatus. Nach der Literatur hätten Auszubildende wie in der Studie von Domin et al. (2019) höhere Werte erreichen sollen. Zur Erklärung des Effekts durch Domin et al. (2019) wurden Erfahrungswerte mit ähnlichen Controllern seitens der jüngeren Personen angegeben. In dieser Studie zeigte sich ein homogenes Bild in den Gruppen hinsichtlich Vorerfahrung in iVR, wodurch die fehlende Signifikanz erklärt wird. Es fehlen allerdings Angaben zu Vorerfahrungen zu Controllern, beispielsweise aus dem Gamingbereich, was in Folgestudien berücksichtigt werden sollte. Trotz fehlender Unterschiede ist positiv anzumerken, dass die UX aller Gruppen Werte erreichte, welche als hoch einzustufen sind (je nach Gruppe 18.57 - 20.27 von maximal 25 möglichen Punkten).

Einige der Antworten auf die offene Frage, was positiv empfunden wurde zeigen, dass das gesamte Erlebnis positiv war («Fand alles gut»; «Mir hat alles gefallen»; «Al-les»). Besonders positiv hervorgehoben wurden die visuellen und auditiven Effekte.

3.6 Fazit und Handlungsempfehlungen

Ziel der Studie war es festzustellen, welche Mapping-Bedingung der iVR-Controller - für sich im Berufsstatus unterscheidende Personen - intuitiver und effektiver war, und ob es Unterschiede im Erleben eines iVR-Tutorials im Berufsstatus der Zielgruppen bei der SBB gab. Dazu wurde ein iVR-Tutorial entwickelt, welches Nutzenden unterschiedliche Funktionen in einem Zeitraum von 6-7 Minuten näherbringt, indem es sie Schritt für Schritt anleitet und auf verschiedene Weisen unterstützt. Mit Hinweisen aus der Handlungstheorie wurden zu lernende Handlungen so unterteilt, dass sie schrittweise gelernt, geübt und schliesslich zur Routine werden konnten, womit der kognitive Aufwand geringgehalten werden sollte. Weiter wurden Hinweise zur Informationsgestaltung wie das Ansprechen mehrerer Modalitäten oder Verwenden von Hinweisreizen aus der CLT verwendet, um ebenfalls die kognitive Last zu minimieren. Der Einsatz einer virtuellen Agentin sollte schliesslich motorisches Lernen durch Vorzeigen und Erklärungen fördern.

Signifikante Unterschiede in der kognitiven Belastung und objektiver sowie subjektiver Usability zeigten, dass für Auszubildenden eine Steuerung mit Mapping ungeeignet ist, wogegen sie für Berufstätige empfohlen wird. Im Erleben des iVR-Tutorials wurden, von der Steuerung unabhängig, Unterschiede in der intrinsischen Motivation zwischen den beiden Zielgruppen gefunden.

Folgende Handlungsempfehlungen lassen sich aufgrund der gewonnenen Daten ableiten:

1. Die Bedingung «Ohne Mapping» ist für Auszubildende geeigneter, da die Cognitive Load geringer ist und weniger fehlerhafte Interaktionen vorkommen
2. Möglich wäre vermutlich auch eine weniger realistische Umgebung, wovon die Berufstätigen aber nicht profitieren.
3. Für Berufstätige ist die Bedingung «Mit Mapping» geeigneter, da weniger fehlerhafte Interaktionen vorkommen. Der Cognitive Load ist in beiden Bedingungen tief.
4. Grip-Button führt aufgrund seiner ergonomischen Lage zu mehrfach gedrückten Buttons, was bei der Bedingung «Ohne Mapping» hinsichtlich versehentlicher Funktionen berücksichtigt werden sollte
5. Für die Funkaufgabe ist der Primary-Button vermutlich geeigneter
6. Für die Teleportaufgabe ist der Trigger-Button vermutlich geeigneter

7. Das Anziehen des HMD sollte besser begleitet und getrennt vom HMD durchgeführt werden. Vorstellbar wäre auch ein Einführungsvideo.
8. Eine Verbesserung der Anleitung in der Funkaufgabe ist anzustreben
9. Ansprache von mehreren Modalitäten beibehalten (Verbal und über Textbausteine)
10. Schrittweisen Aufbau des Lernens der Funktionen beibehalten. Nur Einzelpersonen beschwerten sich über die Repetition oder die Eingeschränktheit durch Anweisungen der virtuellen Agentin
11. Menschenähnlichere Gestaltung der virtuellen Agentin ist möglich, aber nicht zwingend
12. Audio- und Videoeffekte beibehalten, jedoch gezielt einsetzen, um kognitive Belastung nicht unnötig zu erhöhen

3.7 Limitationen und Ausblick

Die Studie ist nicht frei von Limitationen. Als erstes ist anzumerken, dass Wert daraufgelegt wurde, die tatsächlichen Zielgruppen der SBB einzubeziehen. Dadurch wird von einer hohen zielgruppenspezifischen Validität ausgegangen, hat aber auch zur Folge, dass die Erkenntnisse nicht zwingend für einen anderen Kontext oder andere Personengruppen übertragbar sind.

Weiter stellt sich für die abhängigen Variablen die Frage, ob andere Messinstrumente nicht genauer gewesen wären. Es mag der Fall sein, dass ein anderes Messinstrument genauere Daten geliefert hätte, durch den Einsatz an mehreren Standorten der SBB und einer Messe wäre aber beispielsweise der Einsatz eines fMRT-Geräts nicht möglich gewesen. Im Gegensatz zu einem kontrollierten Experiment im Labor wurde es als wichtiger angesehen, ein Feldexperiment mit den tatsächlichen Nutzungsgruppen durchzuführen, also mit Personen aus dem Berufsumfeld der SBB und Auszubildenden.

Trotzdem hätten vorhersehbare Störvariablen wie beispielsweise Lärm, erhoben und berücksichtigt werden können.

Eine weitere Limitation dieser Studie bezieht sich auf die intrinsische kognitive Belastung. ICL wurde im Rahmen dieser Studie nicht berücksichtigt, da die zu lernenden Aufgaben im Tutorial als trivial eingeschätzt wurden und ECL in der Literatur als wichtigste Einflussgrösse auf CL benannt wurde. Bei kritischer Betrachtung im Nachhinein hätte die ICL vermutlich auch die Komplexität der Steuerungsbedingungen gut abbilden können. In Folgestudien, welche die eigentliche Sicherheitsschulung oder andere, komplexere Aufgaben behandeln, sollten sowohl ICL als auch GCL berücksichtigt werden. In zukünftigen Studien könnte zudem der Grad der Realität variiert werden, um die Interaktion von Bedingung und Berufstatus kontrollierter zu betrachten. Zum sei noch die Konfundierung des

Durchführungsraumes zu erwähnen, da die (potenziellen) Auszubildenden nur auf der Messe und die Berufstätigen nur in SstA-Kursen erhoben wurden und nicht zusätzlich auch umgekehrt. Dies sollte bei weiterführenden Studien auch Beachtung finden.

4 Studie III – iVR-Multiplayer-Training mit Auszubildenden

4.1 Einleitung und Orientierung

Im Projekt GENESUS (Christ & Hirschi, 2021) wurde überprüft ob und wie iVR einen Mehrwert in der Berufsbildung bei der login Berufsbildung/ SBB haben kann. Ein Teilergebnis der Studie war der Wunsch der Lernenden nach einem Mehrspielertraining (Multiplayer) in der Berufsbildung. Jedoch bestehen mehrere didaktische Ansätze des Multiplayer Trainings, die analog zu Einzeltrainings, erfolgreich sein können (z.B. Petersen et al., 2023). Vergleichende Studien sind jedoch noch hierzu noch nicht publiziert. Ein weiterer Ansatz findet sich z.B. bei der Österreichischen Bundesbahnen (ÖBB). Diese haben ein Projekt zum Einsatz von iVR in der Lehrlingsausbildung lanciert. In ihrem Video zum Projekt 100 Meter Meidling ist ein Setting ersichtlich bei dem nur eine Person mit einem HMD in iVR trainiert, während die anderen Lernenden von aussen über einen Bildschirm zusehen (unsere OEGB, 2021). Dieses Classroom iVR hat den Vorteil, dass weniger Ressourcen (nicht jeder braucht ein HMD) beim Training eingesetzt werden müssen. Neben den technischen Ressourcen sind auch menschliche Faktoren für die Bewertung von Nützlichkeit und Lernvorteilen ein wichtiges Mass für die Entscheidung ein iVR-Training im Unternehmen etablieren zu können. Zur Beschreibung relevanter menschlicher Faktoren und der Erfassung von Nützlichkeit und Lernvorteil wurden Fragebogenskalen aus zwei theoretischen Modellen abgeleitet: Das Cognitive Affektive Modell of Immersive Learning (CAMIL nach Makransky & Petersen, 2021) in Abb. 4.1 und das TAM Model (nach Davis, 1989; Özgen, Afacan & Sürer, 2021).

4.1.1 CAMIL Modell

Makransky und Petersen (2021) versuchten relevante Einflüsse aus den Erkenntnissen bestehender Forschung zum Prozess des VR-Lernens zu identifizieren und diese in einem Modell zusammenzuführen. Grundgedanke des Modells ist, dass die zwei technisch bedingten Konstrukte, Immersion und Interaktivität, welche bei VR besonders hoch ausgeprägt sind und die Technologie definieren (Johnson-Glenberg, 2018; Makransky & Petersen, 2021), in der VR die psychologischen Konstrukte *Präsenz* und *Handlungsfähigkeit* (en. Agency) hervorrufen. Diese wirken sich ihrerseits auf sechs kognitive und affektive Konstrukte aus: *Interesse*, *Motivation*, *Selbstwirksamkeit* (en. Self-Efficacy), *Verkörperung* (en. Embodiment), *kognitive Belastung* (en. Cognitive Load) und *Selbstregulation*. Diese sechs Konstrukte wirken sich ihrerseits auf den deklarativen und prozeduralen Wissenszuwachs, sowie den Lerntransfer aus (siehe Abbildung 2). Makransky und Petersen (2021) führen an, dass das Medium (iVR) nicht unabhängig von der Lernmethode auf den Lernerfolg einwirkt und daher für eine optimale Nutzung des

Mediums immer die Kombination aus Medium und Methode betrachtet werden muss. Daher sollte bei iVR-Lernsequenzen die erhöhte Immersion und Interaktivität im Rahmen der didaktischen Gestaltung spezifisch berücksichtigt und angesprochen werden (Makransky & Petersen, 2021). Es sollte auch immer mindestens eines der psychologischen Konstrukte Präsenz oder Handlungsfähigkeit durch den Übungsaufbau erzwungen werden, um die Interaktion zwischen Medium und Methode anzustossen (Müser & Fehling, 2022). Im Folgenden werden die einzelnen Konstrukte und ihre Auswirkungen auf den Lernerfolg aus Sicht von Makransky und Petersen (2021) kurz erläutert.

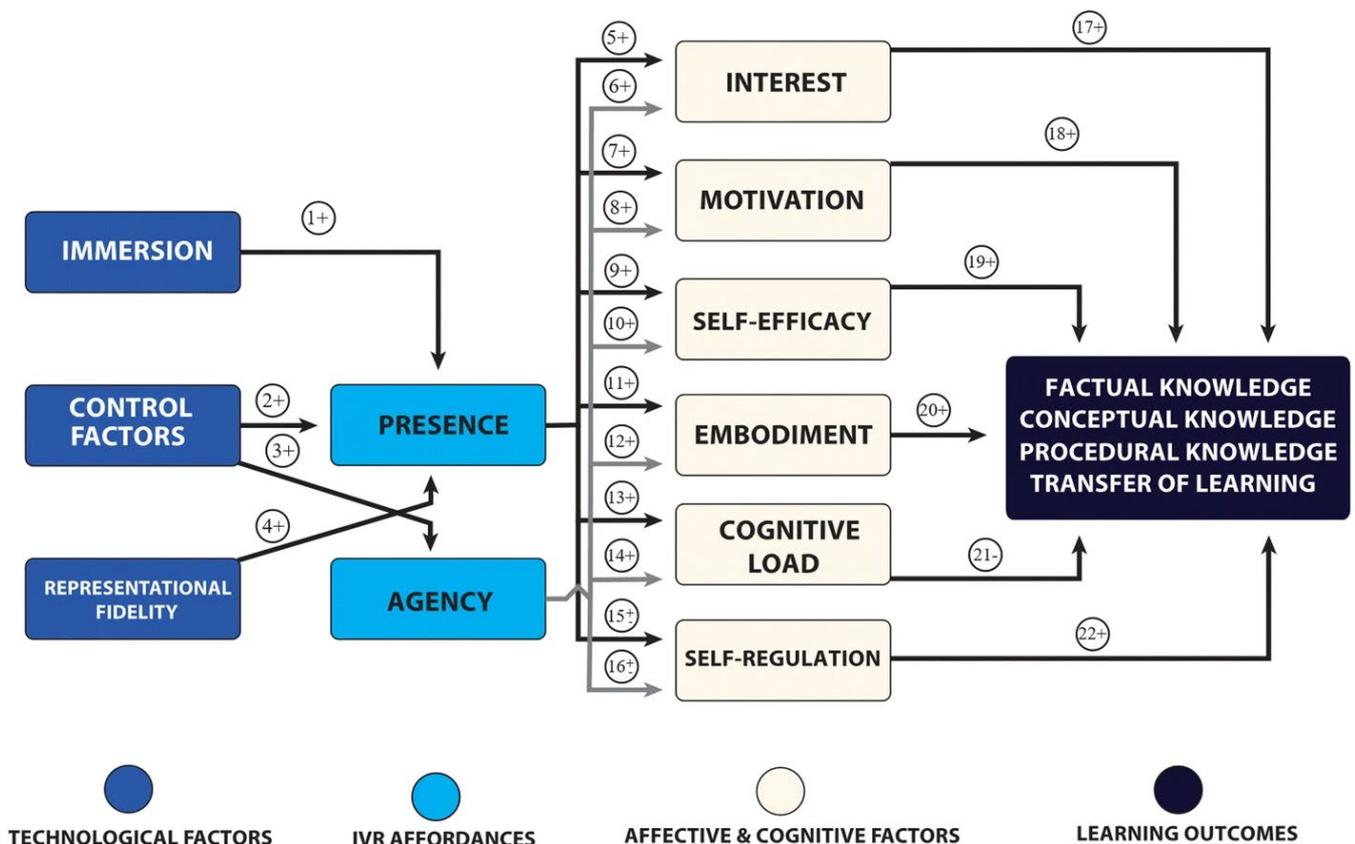


Abbildung 4.1. Übersicht CAMIL Model nach Makransky & Petersen (2021)

Technische Immersion gilt im Rahmen des CAMIL als Grad an sensorischer Realitätsnähe, die in der VR erfahren wird (Bowman & McMahan, 2007; Mütterlein, 2018). Interaktivität kann als Grad der Freiheit verstanden werden, mit dem die Lernenden die Lernerfahrung in der VR kontrollieren können (Makransky & Petersen, 2021).

Präsenz beschreibt das Gefühl an einem Ort zu sein und kann in physische, soziale und Selbstpräsenz unterteilt werden (Makransky, Lilleholt & Aaby, 2017). Physische Präsenz ist die Ortsillusion, bzw. die

Wahrnehmung wirklich in der virtuellen Welt physisch präsent zu sein, sich räumlich und zeitlich von der realen Welt abgeschnitten zu fühlen und mit der Umgebung und den Objekten der VR interagieren zu können, als ob diese real wären (Makransky & Petersen, 2021). Soziale Präsenz, oder Ko-Präsenz, ist die Wahrnehmung anderer virtueller Avatare als wirkliche soziale Akteure (Lee, 2004). Selbstpräsenz ist die Wahrnehmung des eigenen virtuellen Selbst als tatsächliches Selbst (Lee, 2004). Präsenz wird nach Makransky und Petersen (2021) durch ein hohes Mass an Immersion, Interaktivität und Repräsentationstreue gefördert.

Handlungsfähigkeit beschreibt das Gefühl selbständig Handlungen in der VR ausführen und kontrollieren zu können. Das Ausmass an Handlungsfähigkeit ist davon abhängig wie stark Lernende mit der VR und ihren Objekten interagieren und auf diesen Einfluss nehmen können (Makransky & Petersen, 2021).

Situatives Interesse beschreibt die fokussierte Aufmerksamkeit und affektive Reaktion im Moment, welche durch die spezifischen Reize der Umgebung ausgelöst werden (Hidi & Renninger, 2006). Situatives Interesse kann im Moment wissenssuchendes Verhalten auslösen (Knogler, Harackiewicz, Gegenfurtner & Lewalter, 2015). Nach Harackiewicz, Smith und Priniski (2016) kann situatives Interesse die Aufmerksamkeit und das Engagement der Lernenden fördern, was das Lernen vereinfachen kann.

Intrinsische Motivation liegt vor, wenn eine Person durch die Ausübung einer Tätigkeit an sich Befriedigung empfindet und nicht aufgrund äusserer Einflüsse (Ryan & Deci, 2000). Das fördert die Neugierde und Ausdauer beim Lernen (Dev, 1997) und die beim Lernen entstehende Freude, was zu flexibleren und kreativeren Lernstrategien führt (Pekrun, 2006).

Selbstwirksamkeit beschreibt die Überzeugung einer Person, dass sie eine bestimmte Sache oder Handlung gelernt hat oder ausführen kann (Schunk & DiBenedetto, 2016). Die Überzeugung einer Person, dass sie ein gewünschtes Ergebnis durch ihr eigenes Handeln erreichen kann, fördert die Zielsetzung, zielführendes Verhalten, die Aufwandbereitschaft und die Ausdauer, um das gewünschte Ergebnis zu erreichen (Eccles & Wingfield, 2002).

Verkörperung beschreibt das Gefühl sich in einem Körper zu befinden und diesen zu steuern (Kilteni, Groten & Slater, 2012). In VR bezieht sich dies auf die Inbesitznahme eines virtuellen Avatars und wird durch dessen visuelle Repräsentativität und die Fähigkeit dessen Handlungen zu kontrollieren, sowie die an diesen gerichteten Sinneseindrücken wahrzunehmen, gefördert (Kilteni et al., 2012; Stolz, 2015). Die Konzepte Handlungsfähigkeit und Verkörperung sind sich sehr nahe und entgegen des CAMIL sehen einige Forschende Handlungsfähigkeit als Teil von Verkörperung (Kilteni et al., 2012; Longo, Schüür, Kammers, Tsakiris & Haggard, 2008; Dewe et al., 2022).

Kognitive Belastung bezeichnet nach der Theorie der kognitiven Belastung (CLT; Sweller, Ayres & Kalyuga, 2011) die kognitive Beanspruchung, welche durch das Verarbeiten von Informationen entsteht.

Jeder Mensch kann in kurzer Zeit nur eine begrenzte Menge an Informationen verarbeiten. Wird diese Grenze überschritten, kommt es zu einer kognitiven Überlastung (Sweller et al., 2011; Mayer, 2014). Die CLT unterscheidet zwischen intrinsischer kognitiver Belastung (ICL), welche durch die Information, bzw. den Lerninhalt, an sich entsteht, extrinsischer kognitiver Belastung (ECL), welche durch die Art der Präsentation der Information entsteht und relevanter kognitiver Belastung (GCL), welche durch das korrekte Verarbeiten von Information und aufbauen mentaler Konstrukte entsteht (Sweller et al., 2011). In der iVR ist bei ECL wichtig zwischen ECL, welche durch Interaktion entsteht und ECL welche durch die iVR-Umgebung selbst entsteht zu unterscheiden (Andersen & Makransky, 2021). Extrinsische kognitive Belastung durch Interaktion (ECL-I) entsteht dabei durch die Methode wie in der iVR mit der Umwelt interagiert oder die Lernsequenz vorangetrieben wird (Andersen & Makransky, 2021). Extrinsische kognitive Belastung durch die Umgebung (ECL-E) entsteht dabei durch die Objekte und andere Akteure in der iVR. Umso mehr Objekte und Details die iVR beinhaltet, desto grösser ist die ECL-E (Andersen & Makransky, 2021). Eine höhere ECL reduziert die verfügbaren kognitiven Ressourcen, um sich mit dem Lerninhalt zu befassen, und hemmt dadurch das Lernen (Andersen & Makransky, 2021; Parong & Mayer, 2018; 2020).

Selbstregulation bezeichnet in diesem Fall die Fähigkeit einer Person ihr eigenes Verhalten so zu steuern, dass sie trotz möglichen Ablenkungen ihre Aufmerksamkeit und ihr Handeln auf die Erfüllung der Aufgabe konzentrieren kann (Boyd, Barnett, Bodrova, Leong & Gomby, 2005). Lernende welche während der Sequenz eine gute Selbstregulation zeigen, handeln, denken und fühlen eher in einer Art und Weise die für die Erreichung des Lernziels förderlich ist (Zimmermann, 2013) und können dadurch ihre Lernleistung und den Lerntransfer verbessern (Sitzmann & Ely, 2011).

4.1.2 TAM-Modell

Nach dem TAM-Modell hat die Einstellung gegenüber einer Technologie einen Einfluss auf die Nutzung dieser (Ajzen, 1985; Ajzen & Fishbein, 1980). Davis (1989) interpretiert, dass der Stimulus der Systemfunktionen und -fähigkeiten in der nutzenden Person gegenüber der Technologie eine positive Einstellung und Nutzungsmotivation auslösen, was schlussendlich zur tatsächlichen erneuten Benutzung führt. Die affektive Einstellung zur Verwendung einer Technologie wird nach Davis (1989) hauptsächlich durch die kognitiven Konstrukte der *wahrgenommenen Nützlichkeit* (PU) und der *wahrgenommenen Benutzerfreundlichkeit* (PEOU) vorhergesagt. Die affektive Einstellung zur Verwendung wirkt sich wiederum massgeblich auf die *Nutzungsabsicht* (IU) und das effektive Verhalten aus (Davis, 1989). Die kognitiven Konstrukte des TAM weisen dabei Parallelen zu denen der Usability auf (Niklas, 2014). Dabei entspricht wahrgenommene Nützlichkeit der Nützlichkeit (Utility) von Usability und

wahrgenommene Benutzerfreundlichkeit der Benutzbarkeit (Usefulness) von Usability. Usability, die Gebrauchstauglichkeit, stellt daher einen wichtigen Faktor der Nutzungsabsicht und Technikakzeptanz an sich dar (Niklas, 2014). Ein leistungsunabhängiges Konstrukt, welches die Nutzungsabsicht vorher-sagt, ist das wahrgenommene Vergnügen (PE) das bei der Verwendung der Technologie entsteht (Venkatesch et al., 2003). Als wichtigster Prädiktor für wahrgenommenes Vergnügen wird von Venka-tesch et al. (2003) Motivation angesehen. Des Weiteren kann eine hohe wahrgenommene Benutzer-freundlichkeit auch zu einer Reduktion physischer und kognitiver Belastung führen (Tenemaza, Ramírez & de Antonio, 2016). Marangunić und Granić (2015), sowie Legris, Ingham und Collette (2003) berichten eine Notwendigkeit nicht nur die Kernkomponenten des TAM zu untersuchen, son-der auch deren Zusammenhang mit anderen Konstrukten, wie Faktoren der Nutzenden oder der spe-zifischen Technologie. Entsprechend bietet sich eine Untersuchung mit den psychologischen Konstruk-ten des CAMIL an, welche wichtige Faktoren des iVR-Lernens abdecken sollten.

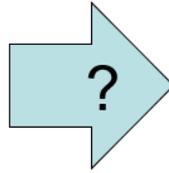
Das TAM wurde bereits erfolgreich zur Untersuchung der Technikakzeptanz von iVR in verschiedenen Bereichen angewandt, wobei iVR eine gute Technikakzeptanz aufwies: Betriebliche Bildung im techni-schen Bereich (Pletz & Zinn, 2018), in der Medizin (Chen, Chen, Or & Lo, 2022) und im akademischen Bereich für Lernende (Hill & du Preez, 2021; Horota et al., 2022; Liu et al., 2020; Özgen, Afacan & Sürer, 2021) und Lehrende (Manero, Álvarez, Romero & Cárdenas, 2022). Pletz und Zinn (2018) stel-len zusätzlich fest, dass Personen die Vorerfahrung mit iVR haben die wahrgenommene Nützlichkeit und Nutzungsabsicht signifikant höher einschätzen. Vergleiche mit dem TAM zwischen konventionel-len Methoden und iVR scheinen hingegen spärlich gesät. Özgen et al. (2021) verglichen iVR-Training mit Paper-Based-Training und konnten ein signifikant höheres wahrgenommenes Vergnügen und Nut-zungsabsicht bei iVR feststellen, während wahrgenommene Benutzerfreundlichkeit und wahrgenom-mene Nützlichkeit keine signifikanten Unterschiede aufwiesen.

4.2 Forschungsfrage

In der wissenschaftlichen Literatur gibt es bis heute keine Übersichtsarbeiten oder Einzelstudien, die verschiedene didaktische iVR-Mehrspieltrainings untersucht und verglichen haben. Petersen et al. (2023) haben zwar in ihrer neusten Studie einen Vorteil von Multiplayer (kollaborierenden Training in iVR) beschrieben, jedoch sind Biologiestudenten, die etwas zusammen über die tierische Zelle lernen, thematisch/ inhaltlich und praktisch weit weg von den sicherheitsbezogenen Tätigkeiten im öffentlichen Verkehr. Innerhalb des Forschungsteams wurden nach Diskussionen der oben beschriebenen Litera-tur, die in Abbildung 4.2 aufgezeigten Variablen als die relevanten menschlichen Faktoren und Erfolgs-kriterien bestimmt.

Prädiktoren:

Vergnügen
Wahrgenommene Kompetenz
Handlungsfähigkeit
Präsenz
Kognitive Belastungen (Interaktion)
Kognitive Belastungen (Umfeld)
Interesse



Kriterien:

Selbstwirksamkeit (subjektiver Lernerfolg)
Nutzungsabsicht
Wahrgenommene Nützlichkeit
Wahrgenommene Benutzerfreundlichkeit

Abbildung 4.2. Übersicht der auf mögliche Vorhersagekraft zu überprüfenden menschlichen Erfolgsfaktoren (Prädiktoren) auf die Erfolgskriterien (Kriterien).

Nach theoriegeleiteter Identifizierung möglicher menschlicher Erfolgsfaktoren kann überprüft werden, ob die Ausprägung dieser menschlichen Faktoren unterschiedlich gross in den einzelnen iVR-Mehrspielertrainings ausgeprägt ist. Somit stellt sich für diese Teilstudie mit Auszubildenden folgenden Forschungsfragen:

F1: Welche für ein iVR-Training relevanten menschlichen Faktoren (Vergnügen, Wahrgenommene Kompetenz, Handlungsfähigkeit, Präsenz, Kognitive Belastung (Interaktion), Kognitive Belastung (Umfeld) und Interesse) können überzufällig den Erfolg, gemessen als Selbstwirksamkeit/subjektiver Lernerfolg, Nutzungsabsicht, Wahrgenommene Nützlichkeit und Wahrgenommene Benutzerfreundlichkeit vorhersagen?

F2: Unterscheiden sich die Ausprägungen relevanter menschlicher Faktoren in den verschiedenen unterschiedlichen iVR-Multiplayer Gleissicherheitstrainings?

4.3 Methode

4.3.1 Trainingskonzeption

Die Basis für die Entwicklung der Trainingsumgebung bildeten die in «Studie 1- Bestimmung der Lernszene» eruierten und für die Gleissicherheit bei der SBB relevanten, Elemente und Variationen. Entsprechend der Rahmenbedingungen der SBB sollten vier Situationen mit unterschiedlichen Sichtverhältnissen (Nebel 250m und 350m Sicht, Regen 250m und 350m Sicht siehe Abbildung 4.3) und Zugdurchfahrten von 80km/h durchlaufen werden.

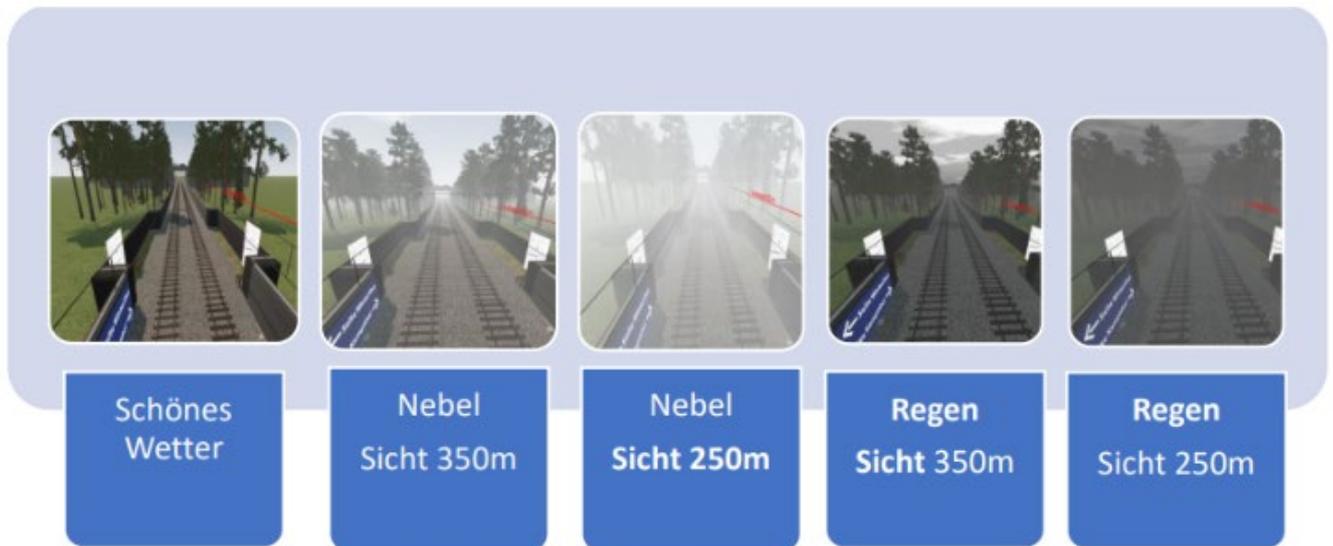


Abbildung 4.3: Verschiedene Beeinträchtigungen der Sicht durch unterschiedliche Wetterbedingungen. Der Abstand der Masten (Oberleitungen) betrug 60m. zwischen 5-6 Masten sollte somit die Sicht betragen. Bei Sicht 250m (egal ob Nebel oder Regen) musste das Gleis verlassen werden und der Schutzraum aufgesucht werden.

Entsprechend der Vorgaben entwickelte das SUCAP-Team den genauen Ablauf des Trainings für die Auszubildenden, dass sich nur in einem Punkt von der Studie mit SStA Teilnehmenden (siehe Studie IV – Ergänzung SstA mit iVR) unterschied: Begrüßung und Einführung, Tutorial, Sensibilisierung, Arbeitsauftrag und Checkliste (bei den Auszubildenden erfolgte kein gemeinsames Ausfüllen der Checkliste und deren Rekapitulation durch den Trainer), iVR-Lernsequenz, Fragebogen, Debriefing. Bei der Sensibilisierung begaben sich alle Teilnehmenden an die Gleise. Die Trainer (zwei Berufsbildner aus dem Gleisbau der login Berufsbildung AG) stellten Regen mit 250m Sicht ein und liessen einen Zug mit 160km/h durchfahren. Dies diente der Selbsterfahrung, wie spät ein Zug erst wahrnehmbar ist. Der Inhalt der Lernsequenz war das Anziehen einer Reihe von Schrauben (siehe Abbildung 4.4)

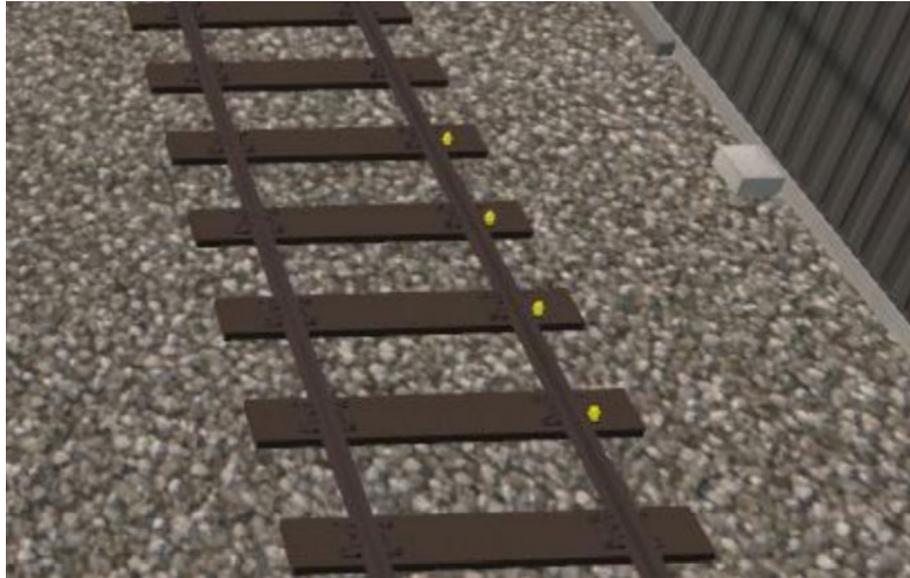


Abbildung 4.4: Gelb blinkende Schrauben mussten mit einem T-Schlüssel analog zum Tutorial in Studie II - Bedienungsparadigma angezogen werden.

mit einem T-Schlüssel, während die notwendige Mindestsichtweite (320m) und durchfahrende Züge beachtet werden mussten. Ist die Sicht schlecht oder nähert sich ein Zug, muss die Arbeit unterbrochen und sich in den Fluchraum begeben werden. Als Hilfestellung zur Einschätzung der Sichtweite waren in der virtuellen Welt alle 60 Meter Strommasten platziert. Jede Situation endete, nachdem alle Schrauben angezogen wurden und die Teilnehmenden wieder im Fluchraum angekommen waren. Unabhängig von der Teilnehmerzahl wurden bei jeder Sequenz alle vier Bedingungen je einmal durchlaufen. Bei wenigen Teilnehmenden in den Mehrspielersettings haben Zweiergruppen mehrere Sequenzen durchlaufen, während sie bei mehr Teilnehmenden eventuell nur eine Bedingung als aktive Akteure erlebt haben. Das Feedback zur Arbeit, die Zuteilung der Versuchspersonen zu der Arbeitsaufgabe und sämtlichen Abfolgen (wann kommt der Zug, welches Wetter, etc.) wurden vom Trainer festgelegt.

4.3.2 Virtuelle Bedingungen

Die Feldstudie wurde mit einem between-group Design aufgestellt. Die Versuchspersonen wurden jeweils einer Versuchsbedingung (iVR-Multiplayer, iVR-Multiplayer mit retrospektivem Perspektivenwechsel und iVR-Classroom) zugeteilt. Die Bedingung iVR-Multiplayer ist dadurch gekennzeichnet, dass sich alle Versuchspersonen inkl. Trainer im virtuellen Raum befinden. Der Trainer beobachtet die Szene, gibt Anweisungen und ändert die Bedingungen direkt im virtuellen Raum. Zwei Versuchspersonen nehmen abwechselnd die Rolle der Reparatur und Kontrolle wahr, der Rest der Versuchspersonen

schauf im virtuellen Raum zu bis gewechselt wird. Dabei nehmen alle Versuchspersonen die Rollen der Reparatur, Kontrolle und des Publikums abwechselnd wahr (siehe Abbildung 4.5).



Abbildung 4.5. Modus Multiplayer. Zwei Personen sind aktiv. Die restlichen Versuchspersonen betrachten die Szene aus der Perspektive des Fluchtraums.

Die zweite Bedingung iVR-Multiplayer mit retrospektiven Perspektivenwechsel (iVR-Multiplayer mit RPW) unterscheidet sich darin, dass die Instruktoren oder der Instruktor hier zusätzlich zum bereits beschriebenen Multiplayer Modus alle im virtuellen Raum anwesenden Personen in die Perspektive der Lokführenden eines durchfahrenden Zuges teleportieren kann. (siehe Abbildung 4.6). Dies geschieht, nachdem eine Zugdurchfahrt vom Trainer ausgelöst wird. Das Ziel dabei ist, den Versuchspersonen ihr eigenes vergangenes Verhalten 30 Sekunden vorher aus dem Blickwinkel einer anderen Berufsgruppe zu zeigen zu können. Dies wurde nach jeder fertigen Tätigkeit (Schrauben andrehen und kontrollieren, ob sich das Wetter ändert (s.u.) oder ein Zug angefahren kommt), durchgeführt.

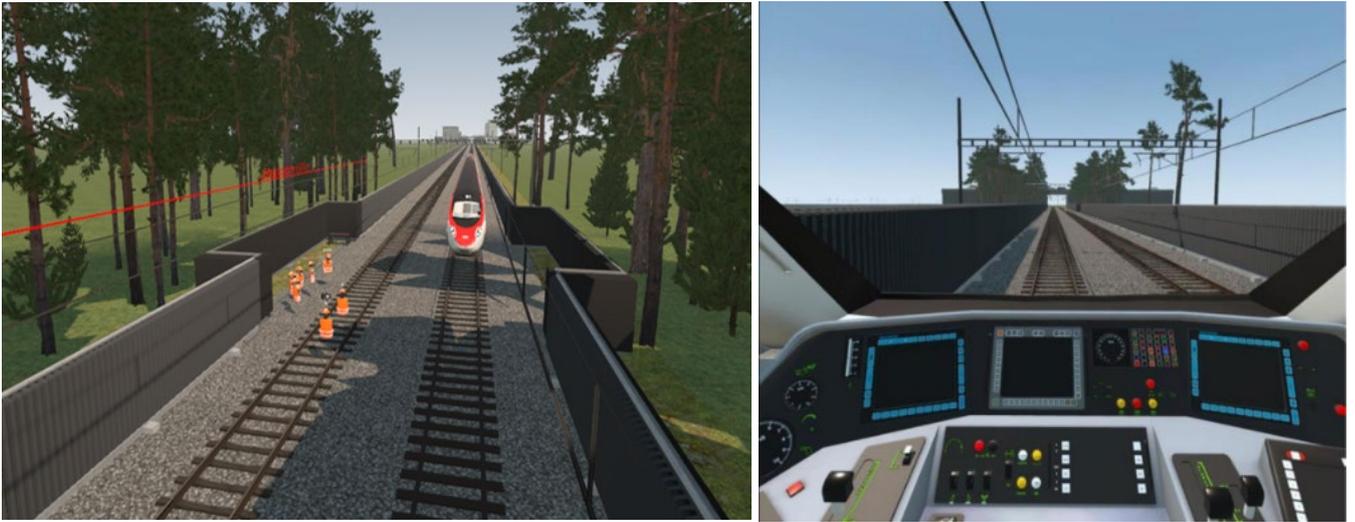


Abbildung 4.6. Modus Multiplayer mit RPW. Versuchspersonen im Bild links wechseln in Kürze (nach Abschluss der Tätigkeit und Zugdurchfahrt) in die Perspektive eines Zugführers versetzt (Bild rechts). Anschliessend fährt der Zug aus 30 Sekunden in der Vergangenheit an den Versuchspersonen vorbei. Die Versuchspersonen sehen sich somit selbst am Gleis, wie sie vor 30 Sekunden dort gearbeitet/sich verhalten haben.

iVR-Classroom bezeichnet, dass sich zwei Versuchspersonen im virtuellen Raum befinden und abwechselnd die Rolle der Reparatur und Kontrolle übernehmen, sie sind somit aktiv. Der Trainer beobachtet die Szene über die Desktopansicht, gibt Anweisungen und ändert die Umwelteinflüsse und Zugdurchfahrten. Der Rest der Versuchspersonen schaut der Szene ausserhalb des virtuellen Raums über einen Beamer oder Bildschirm zu. Sie bilden den passiven Teil. Alle Versuchspersonen nehmen abwechselnd die aktiven und die passive Rolle ein (siehe Abbildung 4.7).



Abbildung 4.7. Bei iVR-Classroom befinden sich nur zwei Versuchspersonen im virtuellen Raum Während die anderen auf einer Leinwand den virtuellen Raum und die Spieler betrachten können.

4.3.3 Stichprobe

Die Rekrutierung der Versuchspersonen fand durch die login Berufsbildung AG statt. Die Zielgruppe waren Jugendliche bis Erwachsene, welche eine Ausbildung bei login Berufsbildung AG absolvieren. Insgesamt nahmen 69 Versuchspersonen an der Studie teil. Davon gaben 68 Personen an, männlich zu sein und eine Person gab an, weiblich zu sein. Das Durchschnittsalter der Versuchspersonen betrug 18,1 Jahre (Median = 17 Jahre, Modalwert = 16 Jahre) mit einer Spanne von 15 bis 37 Jahren (88,4% waren zwischen 15-20 Jahre). Keine Teilnehmenden wurden von der Stichprobe exkludiert. Alle Teilnehmenden wurden aufgeklärt und haben ihre Einverständniserklärung unterschrieben.

4.3.4 Tutorial

Für die Studie wurde das entwickelte Tutorial der vorherigen Studie - Bedienungsparadigma - verwendet. Das Ziel war es, den Lernenden die Ausführung von vier Funktionen beizubringen: Teleportieren, Greifen der Werkzeuge, Handhabung der Werkzeuge mit beiden Händen und Beantwortung eines Funkspruchs mit einem Funkgerät. Das Tutorial stellt eine Situation in der Nähe einer Bahnstrecke nach. Die Anweisungen wurden in der virtuellen Welt eingeblendet und gleichzeitig vom virtuellen

Coach Emma wiedergegeben. Coach Emma wurde mit einer Warnweste und Schutzhelm ohne Körper dargestellt.

4.3.5 Fragebogen

Für die Erhebung wurde ein zweiteiliger Fragebogen verwendet, der zum einen demografische Daten erfasst und zum anderen die Konstrukte analog zur Studie II - Bedienungsparadigma misst. Der Fragebogen besteht aus insgesamt 44 Items zur Bewertung des Experiments, zwei demografische Fragen sowie die Unterteilung des vorhergehenden Modus (Multiplayer, Multiplayer mit RPW oder Classroom) und einer offenen Frage, bei welcher die Versuchspersonen ein optionales Feedback schreiben durften.

Aus den Modellen CAMIL & TAM wurden bereits verwendete und getestete Items ausgewählt. Um die Ergebnisse nicht zu verfälschen, wurden bei fast allen Konstrukten die gesamten Items übernommen. Betreffend intrinsische Motivation wurde anhand der deutschsprachigen Kurzskala intrinsische Motivation (KIM) die Items zu den Konstrukten Interesse und wahrgenommene Kompetenz übernommen (Wilde et al., 2009). Aus dem Cognitive Affective Model of immersive Learning (CAMIL) wurde in Anlehnung an die Ausführungen von (Makransky & Petersen, 2021) die Items zu Agency, Presence, Situational Interest, Self-efficacy, Extraneous cognitive load interaction und Extraneous cognitive load environment vollständig übernommen. Für die Motivation wurden die Subskalen Enjoyment und Wahrgenommene Kompetenz des KIM-Fragebogens (Wilde et al., 2019) ausgewählt. Aus dem Technology Acceptance Model (TAM) wurden Items zu den Dimensionen Perceived Ease of Use, Perceived Usefulness und Intention to use übernommen (Özgen, Afacan & Sürer, 2021).

Items mit Originalsprache Englisch wurden eigenständig durch das Projektteam ins Deutsche übersetzt und durch die Kontrolle von zwei externen englischsprachigen Personen überprüft. Das Projektteam verwendete eine 5-Punkte-Likert-Skala (Likert, 1932), die von 1 (stimme überhaupt nicht zu) bis 5 (stimme voll und ganz zu) reicht. Aufgrund der Formulierung der Items zum Konstrukt Situational Interest wurde hier eine 4-Punkte-Skala (1 = keineswegs, 4 = sehr) verwendet. Der vollständige Fragebogen ist in Anhang C einsehbar.

4.3.6 Versuchsablauf

Die Erhebung fand im Rahmen einer Feldstudie an drei verschiedenen Tagen bei der login Berufsbildung in Zürich Altstetten statt. Angeleitet wurde das Experiment von zwei Berufsbildenden des Gleisbaus, welche als Trainer fungieren. Diese wurden am 28. November 2022 in die Funktionsweise der Software und ihre Rolle als Leitung eingewiesen. Die Räumlichkeit bot Platz für acht gleichzeitig

teilnehmenden Personen. An den Erhebungsdaten wurden zur Sicherheit der Lernenden die Spiel- und Bewegungsbereiche eingegrenzt. Dies physisch anhand von Klebestreifen am Boden und virtuell in der Meta Quest 2 anhand vom Sicherheitsbereich auch *Guardian* genannt.



Abbildung 4.8. Unterstützung der Teilnehmenden beim Start

Die iVR-Brillen und Controller wurden gestartet und vorbereitet, damit die Lernenden gleich starten konnten. Der virtuelle Raum wurde gleichzeitig mit einem Beamer an eine Leinwand projiziert. Als Einführung mussten die Lernenden mit dem erwähnten Tutorial starten, in welchem sie die iVR-Brille kennenlernten und die für das Training benötigten Bewegungen und Controller-Bedienung lernten. Den Lernenden wurde erklärt, dass sie die iVR-Brillen bei Übelkeit zu jederzeit abnehmen dürfen. Das Übertreten des Guardians wurde ebenfalls erklärt. Nach dem Tutorial konnte das iVR-Training vom Versuchsleitenden gestartet werden. Beim Starten des Systems wurden die Teilnehmenden unterstützt (siehe Abbildung 4.8). Die Erhebung und Zuteilung ist in Tabelle 4.1 dargestellt.

Tabelle 4.1: Gruppeneinteilung an den Erhebungsdaten und Summe der Versuchspersonen

Modus	Erhebung am 01.12.2022	Erhebung am 08.12.2022	Erhebung am 25.01.2023	Total
iVR-Multiplayer	6	9	7	22
iVR-Multiplayer mit RPW	8	7	8	23
iVR-Classroom	13	3	8	24
Total	27	19	23	69

4.4 Ergebnisse

Zur Beantwortung der Frage welche menschlichen Faktoren die Selbstwirksamkeit (subjektiver Lernerfolg), die wahrgenommene Nützlichkeit, die Nutzungsabsicht und die Nutzerfreundlichkeit vorhersagen können, wurden nach Prüfung der Voraussetzung der Daten vier multivariate Regressionen gerechnet. In

einem zweiten Schritt wurde dann überprüft, ob die Variablen, die eine überzufällige Vorhersage ermöglichen sich in den Bedingungen Multiplayer, Multiplayer mit RPW und Classroom voneinander überzufällig unterscheiden. Dies wurde je nach Voraussetzung mit Kruskal-Wallis-Test oder einer ein-faktoriellen Anova umgesetzt.

4.4.1 Voraussetzungsprüfung der Daten

Die Multiple-Regressionsanalyse setzt intervallskalierte abhängige Variablen, intervallskalierte unabhängige Variablen, zufällige Stichprobe, Homoskedastizität, Unabhängigkeit des Fehlerwerts, Normalverteilung des Fehlerwerts und keine Multikollinearität voraus (Heimsch, Niederer & Zöfel, 2018). Die vorausgesetzten Skalenniveaus der abhängigen Variablen und die zufällige Stichprobe sind gegeben und werden nicht weiter erläutert.

Homoskedastizität: Die Prüfung auf Homoskedastizität wurde anhand eines Breusch-Pagan-Tests durchgeführt. Die Ergebnisse dessen ergaben bei keiner der abhängigen Variablen ein signifikantes Ergebnis ($p > .05$), weshalb von Homoskedastizität ausgegangen werden kann.

Unabhängigkeit des Fehlerwerts: Die Unabhängigkeit des Fehlerwerts wurde anhand vom Durbin-Watson-Test überprüft. Es wird angenommen, dass Werte zwischen 1.5 und 2.5 akzeptabel sind. Werte unter 1 und über 3 weisen eindeutig auf eine Autokorrelation hin (Field, 2017). Die Werte der Durbin-Watson-Tests liegen bei minimal 1.639 und maximal 2.384, was für unabhängige Fehlerwerte spricht und weisen somit auf keine Autokorrelation hin.

Normalverteilung des Fehlerwerts: Die Normalverteilungen der Fehlerwerte wurden mit dem Shapiro-Wilk-Test geprüft. Im Vergleich zu anderen Normalverteilungstests (Kolmogorov-Smirnov, Lilliefors, Anderson-Darling) ist der Shapiro-Wilk-Test am aussagekräftigsten (Razali & Wah, 2011). Die Ergebnisse waren nicht signifikant, weshalb angenommen werden kann, dass die Fehlerwerte näherungsweise normalverteilt sind.

Multikollinearität: Dieses Kriterium setzt voraus, dass die unabhängigen Variablen nicht zu stark miteinander korrelieren. Überprüft wurde die Multikollinearität anhand einer Kollinearitätsstatistik. Die Ergebnisse unterschreiten den Toleranzwert ($r < .10$) nicht (Pituch & Stevens, 2016). Daraus lässt sich schliessen, dass keine Multikollinearität vorliegt. Zur Überprüfung der ersten Forschungsfragen wurden vier Multiple Regressionen gerechnet. Da das Konstrukt Situational Interest Skale mit einer 4-Punkte-Skala umgesetzt wurde und alle anderen Konstrukte mit eine 5-Punkte Skala werden sowohl der unstandardisierte (β) sowie der standardisierte Betakoeffizient (B) berichtet.

4.4.2 Selbstwirksamkeit (subjektiver Lernerfolg)

Das Regressionsmodell war statistisch signifikant ($F= 9.584$, $p < .001$) und erklärte ungefähr 46,9% der Varianz der Selbstwirksamkeit ($R^2=.524$, bereinigtes $R^2=.469$). Die Selbstwirksamkeit wurde in erster Linie durch das Präsenz-Empfinden ($\beta = 0.364$, $B= 404$, $t = 3.47$, $p < .001$) und Wahrgenommene Kompetenz ($\beta = 0.283$, $B= 376$, $t = 3.272$, $p = .002$) vorhergesagt. Alle anderen Prädiktoren zeigten keine überzufälligen Vorhersagekraft (siehe Abbildung 4.9)

Prädiktoren:

Vergnügen	
Wahrgenommene Kompetenz	*.283 *.376
Handlungsfähigkeit	
Präsenz	** .364 ** .404
Kognitive Belastungen (Interaktion)	
Kognitive Belastungen (Umfeld)	
Interesse	

Kriterien:

Selbstwirksamkeit (subjektiver Lernerfolg)
Nutzungsabsicht
Wahrgenommene Nützlichkeit
Wahrgenommene Benutzerfreundlichkeit

Abbildung 4.9: Signifikante Prädiktoren von Selbstwirksamkeit waren in unserer Studie die Wahrgenommene Kompetenz und das Präsenz-Empfinden im virtuellen Raum. Der obere Wert ist unstandardisiert, der untere Wert ist standardisiert mit $*p>.05$, $**p>.001$.

4.4.3 Nutzungsabsicht

Auch bei der Nutzungsabsicht war das Modell statistisch signifikant ($F= 12.869$, $p < .001$) und erklärte ungefähr 55% der Varianz der Nutzungsabsicht ($R^2=.596$, bereinigtes $R^2=.550$). Die Nutzungsabsicht wurde in erster Linie durch das Situative Interesse ($\beta = 1.102$, $B= .499$, $t = 4.368$, $p < .001$) und Präsenz-Empfinden ($\beta = 0.286$, $B=.248$, $t = 2.313$, $p = .024$) vorhergesagt. Alle anderen Prädiktoren zeigten keine überzufälligen Vorhersagekraft (siehe Abbildung 4.10).

Prädiktoren:

Vergnügen	
Wahrgenommene Kompetenz	
Handlungsfähigkeit	
Präsenz	*.286
Kognitive Belastungen (Interaktion)	*.248
Kognitive Belastungen (Umfeld)	
Interesse	**1.102
	** .499

Kriterien:

Selbstwirksamkeit (subjektiver Lernerfolg)
Nutzungsabsicht
Wahrgenommene Nützlichkeit
Wahrgenommene Benutzerfreundlichkeit

Abbildung 4.10: Signifikante Prädiktoren von Nutzungsabsicht waren in unserer Studie das situative Interesse und das Präsenz-Empfinden im virtuellen Raum. Der obere Wert ist unstandardisiert, der untere Wert ist standardisiert mit *p>.05, **p>.001.

4.4.4 Wahrgenommene Nützlichkeit

Ein weiteres signifikantes Modell ($F = 15.984, p < .001$) erklärte ungefähr 60% der Varianz der Wahrgenommenen Nützlichkeit ($R^2 = .647$, bereinigtes $R^2 = .607$). Die Wahrgenommene Nützlichkeit wurde in erster Linie durch das Situative Interesse ($\beta = .920, B = .434, t = 4.069, p < .001$) gefolgt von Presence Empfindung ($\beta = .459, B = .415, t = 4.139, p < .001$) und Handlungsfähigkeit ($\beta = 0.251, B = .197, t = 2307, p = .024$) vorhergesagt. Alle anderen Prädiktoren zeigten keine überzufälligen Vorhersagekraft (siehe Abbildung 4.11).

Prädiktoren:

Vergnügen	
Wahrgenommene Kompetenz	
Handlungsfähigkeit	*.251
Präsenz	*.197
Kognitive Belastungen (Interaktion)	** .459
Kognitive Belastungen (Umfeld)	** .415
Interesse	** .920
	** .434

Kriterien:

Selbstwirksamkeit (subjektiver Lernerfolg)
Nutzungsabsicht
Wahrgenommene Nützlichkeit
Wahrgenommene Benutzerfreundlichkeit

Abbildung 4.11: Signifikante Prädiktoren von Wahrgenommene Nützlichkeit waren in unserer Studie das Situative Interesse, das Präsenz-Empfinden und Handlungsfähigkeit im virtuellen Raum. Der obere Wert ist unstandardisiert, der untere Wert ist standardisiert mit *p>.05, **p>.001.

4.4.5 Wahrgenommene Benutzerfreundlichkeit

Das letzte statistisch signifikante Modell ($F= 5.358$, $p < .001$) erklärte ungefähr 31% der Varianz der Wahrgenommenen Benutzerfreundlichkeit ($R^2=.381$, bereinigtes $R^2=.310$). Diese wurde nur durch die Wahrgenommene Kompetenz ($\beta = .323$, $B=.378$, $t = 2.883$, $p = .005$) vorhergesagt. Alle anderen Prädiktoren zeigten keine überzufällige Vorhersagekraft (siehe Abbildung 4.12)

Prädiktoren:

Vergnügen
Wahrgenommene Kompetenz
Handlungsfähigkeit
Präsenz
Kognitive Belastungen (Interaktion)
Kognitive Belastungen (Umfeld)
Interesse

Kriterien:

Selbstwirksamkeit (subjektiver Lernerfolg)
Nutzungsabsicht
Wahrgenommene Nützlichkeit
Wahrgenommene Benutzerfreundlichkeit

** .323
** .378



Abbildung 4.12: Signifikante Prädiktoren von Wahrgenommene Nützlichkeit waren in unserer Studie das Situative Interesse, das Präsenz-Empfinden und Handlungsfähigkeit im virtuellen Raum. Der obere Wert ist unstandardisiert, der untere Wert ist standardisiert mit $*p>.05$, $**p>.001$.

4.4.6 Prüfen nach Unterschieden

Um zu verstehen, ob sich die Ausprägungen relevanter menschlicher Faktoren in den verschiedenen unterschiedlichen iVR-Multiplayer Gleissicherheitstrainings unterscheiden. Wurden die Werte der Variablen Situative Interesse, Handlungsfähigkeit, Wahrgenommene Kompetenz und Präsenzepfinden auf Normalverteilung mittels Shapiro Wilk-Test geprüft (siehe Tabelle 4.2)

Tabelle 4.2: Voraussetzungsprüfung auf Normalverteilung aller Gruppenvariablen. Mit * markierte Konstrukte sind überzufällig verschieden von einer Normalverteilung

Konstrukt	iVR-Trainingsbedingung	Statistik	dF	Signifikanz
Interesse*	Multiplayer	.863	22	.006
Handlungsfähigkeit*	Multiplayer	.872	22	.009
Wahrgenommene Kompetenz*	Multiplayer	.885	22	.015
Präsenz	Multiplayer	.937	22	.169
Interesse*	Multiplayer mit RPW	.862	23	.005
Handlungsfähigkeit	Multiplayer mit RPW	.950	23	.289
Wahrgenommene Kompetenz*	Multiplayer mit RPW	.857	23	.004
Präsenz	Multiplayer mit RPW	.946	23	.237
Interesse*	Classroom	.818	24	<.001
Handlungsfähigkeit*	Classroom	.775	24	<.001
Wahrgenommene Kompetenz*	Classroom	.866	24	.004
Präsenz	Classroom	.918	24	.053

Da acht von zwölf Variablen nicht normalverteilt sind wurden zur Prüfung auf Unterschiede drei Kruskal-Wallis-Tests gerechnet. Dabei zeigte nur die Variable von Wahrgenommene Kompetenz überzufällig unterschiedliche Werte ($p < .05$), dass durch den Vergleich von iVR-Multiplayer mit RPW und iVR-Classroom (höhere Werte im iVR-Classroom) erklärt werden kann. Alle anderen Variablen zeigten keine überzufällig unterschiedlichen Werte in den iVR-Trainingsbedingungen (siehe Tabelle 4.3). Die Höhe der Ausprägungen ist in Abbildung 4.13 dargestellt.

Tabelle 4.3: Vier Kruskal-Wallis Tests. Mit * markierte Variablen zeigen überzufälligen Einfluss der iVR-Bedingung.

Variable	N	Teststatistik	dF	Signifikanz
Interesse	69	.059	2	.971
Handlungsfähigkeit	69	.675	2	.714
Wahrgenommene Kompetenz*	69	6.741	2	.034
Präsenz	69	3.479	2	.176

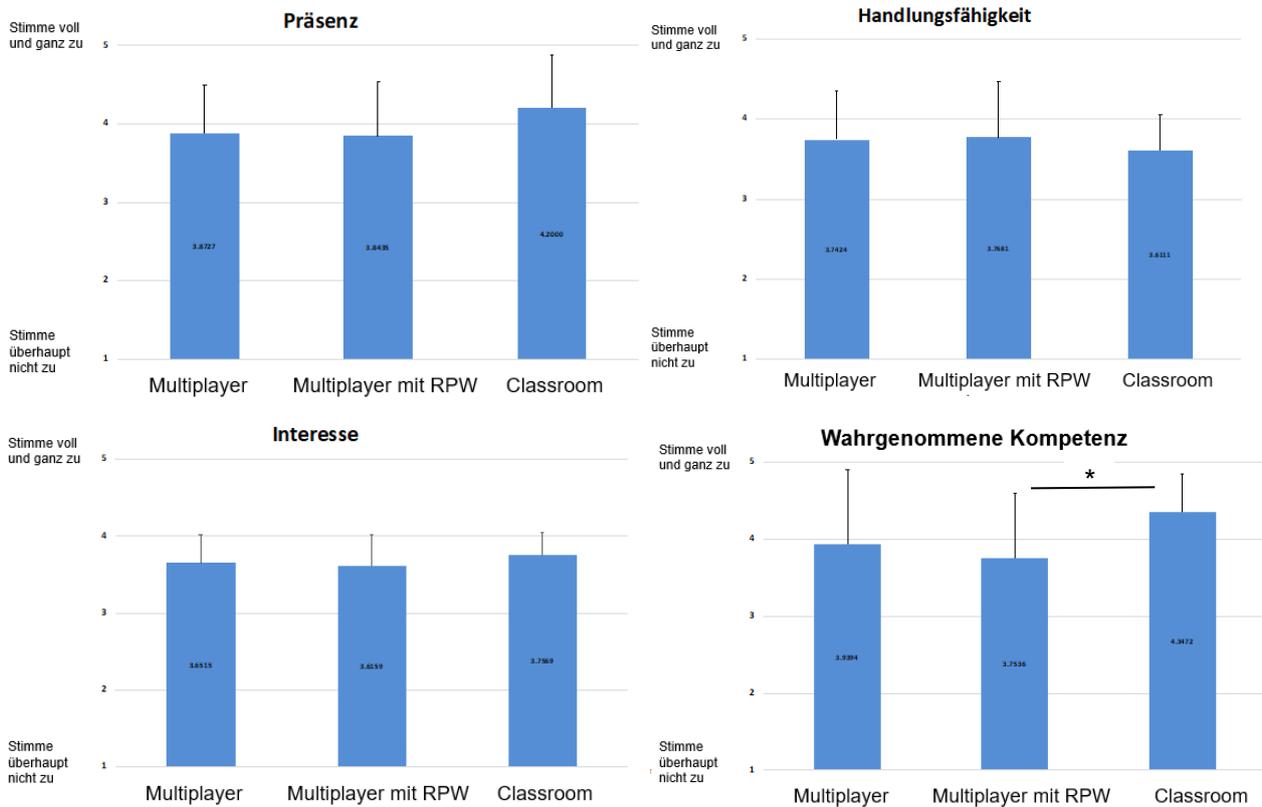


Abbildung 4.13: Ausprägung der Skalen zu Messung der Konstrukte Präsenz, Handlungsfähigkeit, Interesse und Wahrgenommene Kompetenz in unterschiedlichen IVR-Settings. Der überzufällige Unterschied ist mit * markiert. Cohens d betrug beim überzufälligen Unterschied $d=0.93$. Eine Analyse mit G*Power zeigte, dass mit 6 Personen mehr der Unterschied zwischen Classroom und Multiplayer ebenfalls signifikant geworden wäre. Im Vergleich der Variablen Präsenzepfinden wären 15/ 9 Personen mehr nötig gewesen um Unterschiede zwischen Classroom und Multiplayer/Multiplayer mit RPW signifikant werden zu lassen.

4.5 Diskussion

In dieser Teilstudie wollten wir Antworten auf zwei Forschungsfragen finden: (1) Welche für ein iVR-Training relevanten menschlichen Faktoren (Vergnügen, Wahrgenommene Kompetenz, Handlungsfähigkeit, Präsenz, Kognitive Belastung (Interaktion), Kognitive Belastung (Umfeld) und Interesse) können überzufällig den Erfolg, gemessen als Selbstwirksamkeit/subjektiver Lernerfolg, Nutzungsabsicht, Wahrgenommene Nützlichkeit und Wahrgenommene Benutzerfreundlichkeit vorhersagen? (2) Unterscheiden sich die Ausprägungen relevanter menschlicher Faktoren in den verschiedenen unterschiedlichen iVR-Multiplayer Gleissicherheitstrainings?

Wir konnten vier menschliche Faktoren identifizieren, die bei iVR- Gleissicherheitstrainings mit Auszubildenden wichtige Einflussgrößen zu sein scheinen. Dies waren Präsenzepfinden,

Situationsbezogenes Interesse, Wahrgenommene Kompetenz und Handlungsfähigkeit. Anhand der in den Modellen beschriebenen Erfolgsfaktoren werden diese Zusammenhänge nachfolgend diskutiert.

4.5.1 Präsenzepfinden

Präsenz als das Gefühl an einem Ort zu sein scheint überzufällig Selbstwirksamkeit, Nutzungsabsicht und Wahrgenommene Nützlichkeit vorherzusagen. Letzteres Erfolgskriterium wird mit einem Beta von .459 am stärksten gewichtet. Steigt das Präsenzepfinden um einen Punkt, so verändert sich die Wahrgenommene Nützlichkeit um 0.459 Punkte. Eine Massnahme, die beispielsweise durch Interaktivität, Narration oder graphische Realisierung das Präsenzepfinden erhöht, würde somit eine Zunahme der Wahrgenommenen Nützlichkeit in Gleissicherheitstrainings bei Auszubildenden bedeuten. Diese Zunahme ist stärker als die Zunahme in den anderen beiden Erfolgsfaktoren. Wenn auch hier das Präsenzepfinden um 1 Punkt steigt, würde die Selbstwirksamkeit (subjektiver Lernerfolg) um 0.364 Punkte und die Nutzungsabsicht um 0.286 Punkte zunehmen. Bei jungen Auszubildenden lässt sich aufgrund dieser Zusammenhänge feststellen, dass durch die Erhöhung der Präsenz Verhaltensänderungen beobachtbar sein sollten. Da die wahrgenommene Nützlichkeit der Usability (die Gebrauchstauglichkeit) entspricht, kann man durch Erhöhung des Präsenzepfinden auch eine gesteigerte Nutzungsabsicht und Technikakzeptanz erreichen (Niklas, 2014). Die Technologie sollte, unter der Voraussetzung hohe Präsenzwerte auszulösen, somit in der Lage sein die Offenheit neuer Technologie im allgemeine zu erhöhen. Gleichzeitig sollte verhindert werden, dass die Investition in die Technologie im Unternehmen versandet. Dies nicht nur durch Erhöhung der Wahrgenommenen Nützlichkeit, sondern auch durch den Einfluss auf die Nutzungsabsicht. Durch den Einfluss auf die Selbstwirksamkeit, sollte die Überzeugung eines Auszubildenden, dass ein gewünschtes Ergebnis durch eigenes Handeln erreichbar ist, gesteigert werden. Gleichzeitig fördert die Erhöhung von Präsenz ziel führendes Verhalten, die Aufwandbereitschaft und die Ausdauer, um das gewünschte Ergebnis zu erlangen (Eccles & Wingfield, 2002). Abschliessend bedeutet die Steigerung des Präsenzepfindens auch eine erhöhte Überzeugung eine Sache oder Handlung gelernt zu haben und oder ausführen zu können. (Schunk & DiBenedetto, 2016). In dem entwickelten iVR- Gleissicherheitstraining sind physische, soziale und Selbstpräsenz so gegeben, dass keine Unterschiede in den didaktischen Bedingungen (iVR-Multiplayer, iVR-Multiplayer mit RPW und iVR-Classroom) beobachtbar waren (Unterschiede, vor allem der Vergleich mit der Classroom Bedingung wären ab einem mehr von 15 bzw. 9 Versuchspersonen signifikant geworden, siehe dazu Abbildung 4.13). Auch wenn die gefundenen Effekte eher kleiner Natur sind, ist die effektive Gestaltung ein hohes Mass an Präsenz zu erreichen dennoch wichtig, da Präsenz breit auf drei von vier Erfolgsfaktoren wirkt. Nach Makransky und Petersen (2021) wird

Präsenz durch ein hohes Mass an Immersion, Interaktivität und Repräsentationstreue gefördert. Dies gilt es somit für zukünftige Entwicklungen bei iVR von Auszubildenden zu berücksichtigen.

4.5.2 Situationsbezogenes Interesse

Situatives Interesse übersetzt als fokussierte Aufmerksamkeit und affektive Reaktion im Moment (Hidi & Renninger, 2006) sagt nach unseren Daten überzufällig die Nutzungsabsicht und die Wahrgenommene Nützlichkeit vorher. Die Steigerung von 1 Punkt *Situatives Interesse* bewirkt die Steigerung von 1.1 Punkten Nutzungsabsicht und 0.92 Punkte bei der Wahrgenommenen Nützlichkeit. Der Aufwand der betrieben werden muss, um das *Situative Interesse* zu steigern verhält sich im Ertrag, den man in der Nutzungsabsicht und Wahrgenommenen Nützlichkeit findet, etwa gleich.

Wenn man das verwendete iVR-Training dahingehend verändern würde, dass es eine noch höhere Anzahl an Momenten wissenssuchenden Verhaltens auslösen würde (Knogler, Harackiewicz, Gegenfurtner & Lewalter, 2015) wären die Folgen eine erhöhte Nutzungsabsicht und bessere Wahrgenommene Nützlichkeit. Hier muss man nochmals erwähnen das sämtliche Situationen und Feedbacks von den beiden Berufsbildner gegeben wurden. Beide Berufsbildner konnten in der iVR-Bedingung die Lernsituation so attraktiv darstellen, dass sich die fokussierte Aufmerksamkeit in allen iVR-Bedingungen gleich positiv ausgeprägt war. Ist dies der Fall, wird nicht nur die Technologie genützt mit der Folge, dass sich das Investment gelohnt hat, sondern auch die Aufmerksamkeit und das Engagement der Lernenden kann so aktiviert werden, dass nach Harackiewicz, Smith und Priniski (2016) der Lernvorgang als einfacher wahrgenommen werden kann. Abschliessend darf man nicht vergessen, dass die standardisierten Effekte kleiner sind als die unstandardisierten. Hinsichtlich der Erfolgsfaktoren Nutzungsabsicht und Wahrgenommene Nützlichkeit bleibt das *Situative Interesse*, das mit einer Skala von 4 Antwortdimensionen umgesetzt wurde, dennoch die Variable mit der stärksten Vorhersagekraft im Vergleich zu den Skalen mit 5 Antwortdimensionen.

4.5.3 Wahrgenommene Kompetenz

Wahrgenommen Kompetenz als Teil der intrinsischen Motivation liegt vor, wenn eine Person durch die Ausübung einer Tätigkeit an sich Befriedigung empfindet und nicht aufgrund äusserer Einflüsse (Ryan & Deci, 2000). Das fördert die Neugierde und Ausdauer beim Lernen (Dev, 1997) und die beim Lernen entstehende Freude, was zu flexibleren und kreativeren Lernstrategien führt (Pekrun, 2006). Dies scheint in der iVR Classroom Bedingung stärker der Fall gewesen zu sein als in den anderen Bedingungen. In einer Subtheorie der Selbstbestimmungstheorie, der Cognitive Evaluation Theory, legen die Autoren dar, dass Kompetenzerleben während einer Handlung intrinsisch motivierend wirkt (Ryan & Deci, 2002). Dies scheint besser zu funktionieren, wenn sie für ein Teil der Aufgaben aus der virtuellen

Welt geholt werden und die Rolle des Beobachters einnehmen. Wenn man die Wahrgenommene Kompetenz um 1 Punkt erhöht, so bewirkt das die Steigerung der Wahrgenommene Benutzerfreundlichkeit um 0.32 Punkte. Dies ist ein kleiner Effekt und könnte damit erklärt werden, dass durch die Sinnhaftigkeit des eigenen Verhaltens am Gleis die Aufgaben zu erfüllen, die Leichtigkeit das System zu bedienen stärker ausgeprägt ist. Würde die Aufgabe, die zu erfüllen wäre, keinen Sinn machen, würden Mängel in der Bedienung oder unnatürliche Gebrauchstauglichkeit (z.B. Bedienung mit der Hand um laufen zu können -> Teleport) wohlmöglich schlechter bewertet werden.

4.5.4 Handlungsfähigkeit

Handlungsfähigkeit beschreibt das Gefühl selbständig Handlungen in der VR ausführen und kontrollieren zu können. Das Ausmass an Handlungsfähigkeit ist davon abhängig wie stark Lernende mit der VR und ihren Objekten interagieren und auf diesen Einfluss nehmen können (Makransky & Petersen, 2021). Kann man die Handlungsfähigkeit um 1 Skaleneinheit erhöhen steigt die Wahrgenommene Nützlichkeit um 0.251 Punkte an. Eine trainierende Person bewertet also die Nützlichkeit aufgrund der empfundenen Kontrolle, die sie über die Interaktionsmechanismen im virtuellen Raum hat. Schlechte Kontrolle auf Grund von zeitlicher Verzögerung oder Fehler im Programm haben also einen Einfluss auf die empfundene Nützlichkeit der Probanden, wenn auch nicht so stark wie das Präsenzepfinden, dass einen fast doppelt so starken Einfluss hat. Dies bestätigt nochmals die Notwendigkeit ausreichender Testungen der Interaktionsmechanismen im virtuellen Raum.

4.5.5 Samplegrösse

Mit 70 wären wir genau der Anforderung zehn Personen pro Prädiktor nach Harrel (2001) nachgekommen. Die Anforderungen an die minimale Samplegrösse bei späterer linearer Regressionsrechnung werden jedoch diskutiert (Green, 1991; Schmidt, 1971). Die kleinste gefundene Grösse beschreiben Austin & Steyerberg (2015) mit zwei Personen pro Prädiktor diskutiert. Insofern scheint unsere Samplegrösse ausreichend. Die zufällige Zuordnung der Lernenden auf die Gruppen erfolgte durch die login Berufsbildung AG. Dies sollte mögliche einflussnehmende Merkmale wie räumliche Vorstellung, motorische Fähigkeiten, etc. weitgehendst ausgeglichen haben, auch wenn man diesbezüglich eine vorherige Prüfung hätte anstreben können. Dies war jedoch aufgrund zeitlicher Limitationen nicht möglich. Die Auswahl der Trainer erfolgte ebenfalls durch die login Berufsbildungs AG. Hierbei waren keine technischen Schwierigkeiten beobachtbar. Einzig etwas nachlassende Konzentration/Disziplin im Multiplayer-Modus auf. Hier wurde seitens der Auszubildenden gerne mehr exploriert (schmeissen des T-Bars, auf den Gleisen stehen, etc.), dass häufigeres Ermahnen durch den Trainer zur Folge hatte.

4.6 Fazit

Mit der Studie konnten vier wichtige menschliche Faktoren identifiziert werden, welche für die Entwicklung von iVR-Gleissicherheitstrainings deswegen wichtig sind, da sie einen überzufälligen Einfluss auf die Nutzungsabsicht, die wahrgenommene Nützlichkeit, die Selbstwirksamkeit (subjektiver Lernerfolg) und die wahrgenommene Benutzerfreundlichkeit haben. Kann man ein Training so konzipieren, dass sich bei Trainierenden hohe Werte im/in Präsenzepfinden, wahrgenommene Kompetenz, Situatives Interesse und Handlungsfähigkeit einstellen, wird dieses Training nachgefragt werden, als bedienerfreundlich eingestuft und als sinnvoll bezeichnet werden. Darüber hinaus werden Trainierende von sich überzeugt sein etwas im iVR-Raum lernen zu können. Da es zwischen den iVR-Lernsettings keine eindeutigen Unterschiede in den menschlichen Faktoren gibt, steht es der login-Berufbildung AG frei, mit welcher Form des Multiplayer es später die Lernenden unterrichten will. Sollte es jedoch zum unternehmensweiten Einsatz und zur Weiterentwicklung kommen, sind die Faktoren Präsenzepfinden, wahrgenommene Kompetenz, Situatives Interesse und Handlungsfähigkeit regelmässig auf hohe Ausprägungen zu prüfen respektive die Weiterentwicklung hinsichtlich der genannten Faktoren zu gestalten.

5 Studie IV – Ergänzung SstA mit iVR

5.1 Einleitung und Orientierung

Sobald eine Person sicherheitsrelevante Tätigkeiten im Gleisbereich ausübt, braucht es eine Sicherheitsausbildung wie sie auch in entsprechenden Funktionsausbildungen, wie zum Beispiel der Sicherheitswärter, der Sicherheitschef oder die Sicherheitsleitung, vorhanden sein müssen. Diese werden im Reglement RTE 20100 beschrieben.

5.1.1 Reglement RTE 20100

Beim Reglement RTE 20100 handelt es sich um das Regelwerk Technik Eisenbahn (RTE). Ein Grundsatz des RTE ist die sinnvolle Standardisierung durch praxisbezogene und bewährte Lösungen. Des Weiteren handelt es sich beim RTE um die Sicherung, Sammlung und Verteilung von Erfahrung und Wissen von Fachleuten der Schweizer Bahn (Verband öffentlicher Verkehr, n.D.). Das RTE 20100 wird für die Sicherheit bei Arbeiten im Gleisbereich durch Funktionsausbildungen eingesetzt. Zu diesen Funktionsausbildungen gehört die Ausbildung Selbstschutz Arbeit (SstA ->Verband öffentlicher Verkehr, n.D.). Mit der Ausbildung SstA darf eine Person den Gleisbereich allein (im Selbstschutz) betreten und Arbeiten ausführen. Der Grundkurs für die Ausbildung SstA findet an zwei Tagen statt. Hierbei werden sowohl die Theorie als auch die Praxis übermittelt. Der Kurs wird mit einer theoretischen Prüfung abgeschlossen, welche bei erfolgreichem Bestehen die Personen dazu befähigt, selbstständig im Gleisbereich arbeiten zu dürfen (Verband öffentlicher Verkehr, n.D.).

5.1.2 Checkliste Selbstschutz und Annäherungsdistanzen

Bevor der Gleisbereich für Arbeiten im Selbstschutz betreten werden darf, muss die Checkliste Selbstschutz (siehe Abbildung 5.1) ausgefüllt werden. Wie diese ausgefüllt werden muss, lernen die Teilnehmenden im SstA Kurs. Ein wichtiger Aspekt ist hierbei die Annäherungsdistanz, die vorab anhand der Geschwindigkeit des Zuges und der Räumungszeit berechnet wird. Das Thema der Annäherungsdistanz wird seitens der SBB als ein möglicher Grund genannt, warum die SstA-Prüfung nicht bestanden wird. Zum einen werden hier viele Punkte (4 Punkte) vergeben, jedoch die volle Punktzahl oft nicht erreicht. Dies scheint daran zu liegen, dass Trainierende das Teilgebiet nicht sonderlich gut im bisherigen Trainingssetting verstehen. Auf der nachfolgenden Abbildung 5.1 wird die notwendige Checkliste dargestellt und mit einer Räumungszeit von acht Sekunden gerechnet. Dies bedeutet, dass sich die Person innerhalb von acht Sekunden in den Fluchtraum zurückziehen können muss. Sofern die Züge auf den Gleisen, auf welchen gearbeitet wird, 80 km/h fahren, wird eine Annäherungsdistanz von 320 Metern berechnet. Dies bedeutet, dass die Person jederzeit über eine Sichtweite von mindestens 320

Metern verfügen muss. Ist diese Sichtweite nicht mehr gegeben, muss der Gleisbereich verlassen und sich in den Fluchraum zurückgezogen werden.

Checkliste Selbstschutz

Alarmstelle SBB: 0512250854
 Datum XX.XX.XXXX Tel.-Nr. Fdl 051 225 44 26
 Bahnhof/ Strecke/ Km Winterthur – Kempthal, GL 555, KM 24.0 – KM 23.9

Annäherungsdistanz

Räumungszeit: (Sicherheitsfrist:)	8 Sek. 12 Sek.	8 Sek. 14 Sek.	10 Sek. 18 Sek.)	
≤ 40 km/h	130 m	160 m	190 m	
≤ 60 km/h	200 m	240 m	270 m	<input type="checkbox"/> Gleisbegehung
≤ 80 km/h	270 m	320 m	360 m	<input type="checkbox"/> Gleisbegehung
≤ 100 km/h	340 m	390 m	450 m	<input type="checkbox"/> Arbeiten im Gleis
≤ 120 km/h	400 m	470 m	540 m	<input type="checkbox"/> Arbeiten im Gleis
≤ 160 km/h	540 m	630 m	720 m	<input type="checkbox"/> Arbeiten im Gleis

Gleisbegehung

- Kritische Zonen auf dieser Strecke? JA NEIN
- Ich überblicke jederzeit und uneingeschränkt die Annäherungsdistanzen
- Der Fluchraum ist jederzeit festgelegt und bekannt
- Ich selbst und mein Material sind jederzeit ausserhalb der Annäherungszone elektrischer Anlagen
- Bei Arbeiten, Abstand zum Gleis in Betrieb > 1.5 m (mit Absp.) > 3 m

Arbeiten im Gleis (nur SstA)

- Betriebliche Massnahmen Fahrrichtung festlegen Sperrern von Gleisen/Weichen
- Der Fluchraum ist festgelegt und bekannt JA
- Ich überblicke jederzeit und uneingeschränkt die Annäherungsdistanzen JA
- Ich arbeite ausserhalb der Annäherungszone elektrischer Anlagen JA NEIN
 Falls NEIN, sachverständige Person beigezogen? JA

Person 1 SstA SstB Unterschr. Visum 1
 Person 2 Erstinstruktion SstA SstB Unterschr. Visum 2

Abbildung 5.1. Checkliste Selbstschutz (SBB, n.D.)

Hierzu wurde die, in Studie III vorgestellte, digitale Trainingsumgebung gebaut. Mit Hilfe dieser kann die Sichtweite vom Trainer oder automatisiert eingestellt werden, indem eine Wetterveränderung durch Nebel oder Regen die Sicht auf bis zu 250m reduzieren kann.

5.1.3 Arbeitserfahrung

In der dritten Studie wurde das CAMIL Model und das TAM-Model vorgestellt. Beide Modelle sind hinsichtlich der Erfolgsfaktoren im immersiven virtuellen Lernraum wichtig um Stellschrauben für die Verbesserung einer iVR-Umgebung zu Beschreiben. In dieser letzten Studie des SUCAP Projekts wurde

aber ein anderer Schwerpunkt gelegt. Sowohl im Steeringboard als auch Advisoryboard wurde früh deutlich, dass sowohl die Gebrauchstauglichkeit, Nützlichkeit aber auch die arbeitsbezogene Einschätzung (der Mehrwert für die Arbeit an der Baustelle) im Vordergrund stehen und weniger die theoretischen Modelle des immersiven Lernens. Da jedoch für die Einschätzung des Sicherheitsempfindens bei der Arbeit am Gleis nach curricularem Kurs keine Kurzskalen existieren, wurde von der SBB (Paul Hügli) fünf zusätzliche Fragen entwickelt. Diese waren:

- Im Vergleich zu Schulungen ohne Lerntool fühle ich mich besser auf die Arbeit in der Praxis vorbereitet.
- Im Vergleich zu Schulungen ohne Lerntool fühle ich mich sicherer für die Arbeit in der Praxis vorbereitet.
- Die Gefahren im Gleisbereich konnten mir mit dem Lerntool sehr deutlich gemacht werden
- Das Lerntool hat mir eindrücklich die Gefahren im Gleisbereich vermittelt.
- Das Erlebnis mit dem Lerntool hat mein Situationsbewusstsein positiv beeinflusst.

Die Antwortdimensionen lauteten: 1 = «Stimme überhaupt nicht zu», 2= «Stimme nicht zu», 3= «Stimme weder zu noch lehne ich ab», 4= «Stimme zu» und 5 =«Stimme voll und ganz zu». Mit diesen Fragen sollte es den SstA-Teilnehmenden möglich sein, eine vorläufige Einschätzung für die zukünftige Arbeit zu geben. Mittels der theoretischen Vertiefung (siehe Kapitel III) und der praktischen Ausrichtung konnten nun die Forschungsfragen aufgestellt werden.

5.2 Forschungsfragen

Nach Identifizierung relevanter menschlicher Faktoren nach dem TAM Model (siehe Abschnitt 4.1.2) und praxisrelevanter Faktoren (Items von Paul Hügli) kann nun überprüft werden, ob die sechs verschiedenen Bedingungen (iVR- Singleplayer, iVR-Multiplayer, iVR-Multiplayer mit retrospektivem Perspektivwechsel, iVR-Classroom, Laptop PC und Kontrollgruppe) unterschiedlich stark die Wahrscheinlichkeit für das Erreichen der vollen Punktzahl in der Aufgabe zur Annäherungsdistanz/Checkliste sowie die Nutzererfahrung beeinflussen. Somit stellt sich für diese Teilstudie mit SstA Teilnehmenden folgenden Forschungsfragen:

F1: Unterscheidet sich die Wirkung von zusätzlichen digitalen Trainingseinheiten (iVR- Singleplayer, iVR-Multiplayer, iVR-Multiplayer mit retrospektivem Perspektivwechsel, iVR-Classroom, Laptop PC) auf die Prüfungsleistung (gemessen in der vollen Punktzahl in der Prüfungsfrage zu Annäherungsdistanz/Checkliste) im Vergleich zu einer Kontrollgruppe im SstA-Kurs?

F2: Unterscheiden sich die Ausprägungen der Fragen zur Arbeitseinschätzung in den unterschiedlichen, zusätzlichen digitalen Trainingseinheiten (iVR- Singleplayer, iVR-Multiplayer, iVR-Multiplayer mit retrospektivem Perspektivwechsel, iVR-Classroom, Laptop PC) im Vergleich zu einer Kontrollgruppe?

5.3 Methode

Die Umsetzung der Studie im SstA Setting kann einer Feldstudie mit Kontrollgruppe zugeordnet werden. Für die Durchführung dieser Feldstudie wurden die Teilnehmenden jeweils einer von sechs Gruppen zugeteilt. Die sechs Gruppen sind in der nachfolgenden Tabelle 5.1 ersichtlich.

Tabelle 5.1: Darstellung der sechs Gruppen

Regulärer Kurs	iVR-Singleplayer	Singleplayer Desktop	iVR-Multiplayer	iVR-Multiplayer mit retrospektivem Perspektivwechsel	iVR- Classroom
					
<i>Kontrollgruppe</i>	<i>Experimentalgruppe</i>	<i>Experimentalgruppe</i>	<i>Experimentalgruppe</i>	<i>Experimentalgruppe</i>	<i>Experimentalgruppe</i>
Der SstA Kurs fand wie gewohnt statt. Die Teilnehmenden sehen Filme und arbeiten an einer Holzeisenbahn. Am Ende des Tages erfolgte eine Befragung mittels Fragebogen	Zusätzlich zum regulären Kurs erlebten die Teilnehmenden allein mittels HMD vier Wetterbedingungen und Zugdurchfahrten. Mithilfe eines T-Schlüssels mussten sie blinkende Schrauben festdrehen. Unterstützt wurden sie durch eine Stimme respektive durch Text.	Zusätzlich zum regulären Kurs erlebten die Teilnehmenden allein mittels PC-Monitor vier Wetterbedingungen und Zugdurchfahrten. Mithilfe eines T-Schlüssels mussten sie blinkende Schrauben festdrehen. Unterstützt wurden sie durch eine Stimme respektive durch Text.	Zusätzlich zum regulären Kurs erleben die Teilnehmenden zusammen mit dem Trainer mittels HMD die virtuelle Welt. Der Trainer teilt die Personen ein lässt Züge durchfahren und verändert das Wetter analog zum Singleplayer (siehe Studie III)	Zusätzlich zum regulären Kurs erleben die Teilnehmenden zusammen mit dem Trainer mittels HMD die virtuelle Welt. Der Trainer teilt die Personen ein lässt Züge durchfahren und verändert das Wetter analog zum Singleplayer. Am Ende einer Wetterbedingung erleben die Teilnehmenden sich selbst am Gleis 30 Sekunden vorher durch den Blickwinkel aus dem Zugführerstand (siehe Studie III)	Zusätzlich zum regulären Kurs erleben die Teilnehmenden in Gruppen mittels HMD die virtuelle Welt. Der Trainer teilt, von der echten Welt aus die Personen ein, lässt Züge durchfahren und verändert das Wetter. Dies beobachtet der restliche Kurs von der echten Welt aus an einer Leinwand (siehe Studie III)

Die Gruppenzuteilung konnte nicht randomisiert erfolgen, da die Teilnahme freiwillig und ohne Incentivierung erfolgte. Alle Personen des jeweiligen SstA-Kurses, welche an der Erhebung teilnehmen wollten, wurden somit der gleichen Gruppe zugeordnet. Welche Gruppe dies sein wird, wurde bereits im Vorhinein festgelegt und ggf. angepasst. Dabei erfolgte dies anhand der Möglichkeit welcher Raum zur Verfügung stand und der laufenden Anzahl Personen, die bereits in einer bestimmten Bedingung teilgenommen hatten. Somit konnte gewährleistet werden, dass in allen Gruppen eine gleich grosse

Stichprobe vorhanden ist. Bei der vorliegenden Arbeit handelt es sich des Weiteren um ein Feldexperiment. Feldexperimente sind zugeschnitten auf konkrete Praxisprobleme und finden in der realen Umgebung statt. Die Untersuchungsbedingungen ähneln in einem Feldexperiment den Alltagsbedingungen (Döring & Bortz, 2016). Aus diesem Grund wird Feldexperimenten auch eine hohe externe Validität zugeschrieben (Brosius, Haas & Koschel, 2016). Ein Feldexperiment führt allerdings zu einer verminderten Kontrolle von Störvariablen wie z. B. unterschiedliche Räumlichkeiten oder die verschiedenen Trainer, was die kausale Interpretierbarkeit der Ergebnisse erschwert/teilweise unmöglich macht (Döring & Bortz, 2016). Sofern den unterschiedlichen Bedingungen verschiedene Personen zugeordnet werden, spricht man von einem Between-Subject-Design. In einem Experiment, welches als Between-Subject-Design angelegt ist, ist jede Person nur einer Bedingung ausgesetzt. Bei dieser Art von Experiment wird das Verhalten von Personen, welche sich in einer Versuchsbedingung befinden, mit dem Verhalten von Personen, welche sich in einer anderen Versuchsbedingung befinden, verglichen (Charness, Gneezy & Kuhn, 2011).

5.3.1 Stichprobe/Stichprobengewinnung

Mithilfe der Software G*Power (Faul, 2020) wurde a priori die benötigte Anzahl an Teilnehmenden berechnet. Für die gewählten statistischen Verfahren wurde mit sechs Gruppen, einer Teststärke von .95 und einer geschätzten mittleren Effektstärke von $f^2(V) = .15$ (Cohen, 1992) eine Stichprobengröße von 90 Personen berechnet.

Die Akquise der Teilnehmenden erfolgte über den Trainer, der via Mail eine Einladung an die Teilnehmer ca. eine Woche vor Kursbeginn verschickte. Diese Einladung musste aufgrund weniger Anmeldungen einmal angepasst werden. Die finale Version der Einladung ist in Abbildung 5.2 ersichtlich.

Guten Tag Herr/Frau ... (persönliche Anschreiben sind immer besser) wenn es nicht geht so lassen

Sie haben bald Ihren «Selbstschutz Arbeiten Sst A» Kurs. Wir wollen Ihnen hierzu in dieser Mail eine wichtige Zusatzinformation zukommen lassen:

Wie Sie wissen, dient der Kurs dazu, Ihnen ein sicheres Arbeiten am Gleis zu ermöglichen. Dabei ist das Verständnis zum Ausfüllen von einer Checkliste und das Verstehen der Annäherungsdistanz von Zügen von allerwichtigster Bedeutung, wenn es um Ihre Sicherheit geht. Um gerade dieses Wissen (dass auch über Leben und Tod entscheiden kann) vertieft zu üben, bieten wir am Ende von Tag 1 (Kurstagende 16:30 Uhr) eine weitere Trainingsstunde mit einer Virtual Reality Brille an (Ende 17:30h).

Durch die VR-Brille betreten Sie eine virtuelle Arbeitsstelle im Gleisbereich. Hier werden gelernte Inhalte in der virtuellen Welt dargeboten. Dabei ist das **Ziel**: Das Bewusstsein in Bezug auf **Sichtweite** und **Platzverhältnisse** zu schärfen.

Die Teilnahme an diesem VR-Zusatztraining ist freiwillig und die Anzahl Teilnehmer/innen ist auf max. 8 Personen beschränkt.

Bitte antworten Sie nicht auf diese E-Mail. Antwort bei Kursbeginn vor Ort an die Kursleitung ist völlig ausreichend. Über eine Teilnahme würden wir uns sehr freuen.

Freundliche Grüsse

--Name Kursleiter--
Kursleitung

Abbildung 5.2: Finaler Einladungstext:

Am ersten Trainingstag fragte der Trainer dann, wer gerne an der Untersuchung teilnehmen wollte. Diese Anzahl wurde dann am gleichen Tag an die FHNW rückgemeldet und je nach Anzahl Personen, die bereits ein Training absolviert hatten, wurde eine geplante Bedingung oder neue Bedingung gewählt. Die Erhebung startete am 19.12.2022 und endete am 27.04.2023 und bestand in diesem Zeitraum aus 34 Erhebungen.

5.3.2 Material

Für die immersiven Bedingungen (iVR-Singleplayer, iVR-Multiplayer, iVR-Multiplayer mit retrospektiven Perspektivwechsel und iVR-Classroom) wurde als HMD die im Workshop der Studie I ermittelte Meta Quest 2 im Standalone-Modus verwendet. Dies bedeutet, dass die Brille keine Verbindung zu einem Computer benötigte. Damit der Komfort erhöht werden konnte und auch Brillenträger ein angenehmeres Tragerlebnis haben wurden Elite-Straps und Brillen-Spacer benutzt. Über die eingebauten Raumklang-Lautsprecher konnten die Teilnehmenden das Audio wahrnehmen, was insbesondere beim Tutorial wichtig war. Die Teilnehmenden erhielten einen Controller pro Hand, mit welchem sie sich teleportieren sowie den T-Schlüssel greifen und benutzen konnten.

Für die Singleplayer Desktop Bedingung wurde ein Dell Precision 7770 (siehe Abbildung 5.5) verwendet. Zusätzlich zu jedem Computer wurden eine Maus und Kopfhörer zur Verfügung gestellt. Damit die Internetverbindung der VR-Brillen und der Computer während der ganzen Zeit gewährleistet werden konnte, wurde ein ASUS AX56 AX1800 Router eingesetzt.

5.3.3 Virtuelle Lernumgebungen

Die virtuelle Lernumgebung wurde basierend auf vorherigen Forschungsarbeiten (Studie I und Studie II) erarbeitet. Bilder zur Trainingsszene sind in Studie III ersichtlich. In der Studie II konnte gezeigt werden, wie ein Tutorial als Einführung in die iVR-Umgebung mit dem HMD am besten gestaltet werden sollte. Hierzu wurde eine virtuelle Agentin eingesetzt, welche die Teilnehmenden durch das Tutorial führt. Die virtuelle Agentin gibt über Textbausteine eingeblendete Erklärungen verbal von sich und zeigt die Funktionen vor. Zuerst wird mithilfe der virtuellen Agentin Emma gezeigt, wie sich die Teilnehmenden von einem Punkt zum anderen teleportieren können. Des Weiteren wird den Teilnehmenden durch die virtuelle Agentin Emma aufgezeigt, wie der T-Schlüssel gegriffen und benutzt werden kann. Das Anziehen der Schrauben mithilfe des T-Schlüssels wird im Tutorial dreimal wiederholt.

5.3.4 Datenerhebung

Seitens SBB wurden vorab Rahmenbedingungen für die Durchführung der Erhebung definiert. Eine wichtige Voraussetzung war, dass der SstA Kurs nicht verändert werden und erst nach Ende des Kurs-tages um ca. 16:30 Uhr die Untersuchung durchgeführt werden darf. Dies war wichtig, um gewährleisten zu können, dass alle Teilnehmenden des SstA Kurses den gesamten Lerninhalt vermittelt bekommen. Die Erhebung sollte zudem nicht länger als eine Stunde dauern. Miteingeschlossen in diese Stunde ist die Aufklärung/Einverständniserklärung mit Unterschrift und der Fragebogen (dessen Beantwortung optimalerweise nicht länger als 10 Minuten dauern sollte). Zusätzlich war es der SBB wichtig, dass in der Einführung zur Untersuchung darauf hingewiesen wird, dass die Teilnahme in keiner Verbindung mit dem Bestehen oder nicht Bestehen der Prüfung steht.

5.3.5 Durchführung der Datenerhebung

Wie bereits beschrieben, fand die Erhebung immer am ersten Tag nach dem regulären SstA Kurs statt. An der Erhebung waren immer mindestens zwei Versuchsleitende (FHNW) anwesend, um die Teilnehmenden sowie den Kursleiter unterstützten zu können. Dies einerseits bei Anliegen zum Experiment und andererseits bei technischen Fragen zum HMD oder dem PC-Monitor (siehe Abbildung 5.5). Die Teilnehmenden wurden ungefähr eine Woche vor dem Kurs durch den Kursleiter per E-Mail (siehe Abbildung 5.2) informiert, dass eine Erhebung stattfinden wird. Dies basierend auf dem Hintergrund, dass sich die Teilnehmenden durch die vorgängige Information bereits Gedanken machen konnten, ob sie

freiwillig am Experiment teilnehmen möchten. Durch die Vorinformation sollte daher gewährleistet werden können, dass die Teilnehmenden auch Zeit haben werden, um eine Stunde länger bleiben zu können. Aufgrund der Verfügbarkeiten der HMDs und der PC-Monitoren konnten bei der Bedingung iVR-Singleplayer jeweils acht Personen teilnehmen und in der Bedingung Singleplayer Desktop 6 Personen. Die maximale Anzahl Teilnehmenden in einem SstA Kurs beträgt 12 Personen. Bei der Kontrollgruppe konnten jeweils alle 12 Personen freiwillig an der Erhebung teilnehmen. Vor dem Start der Datenerhebung wurden sie umfassend über die Studie informiert und unterschrieben ihre Einverständniserklärung. Auf der nachfolgenden Abbildung 5.3 sind die Kursräume der Standorte Zürich und Olten ersichtlich, in welchen die Bedingung iVR-Singleplayer stattfand.

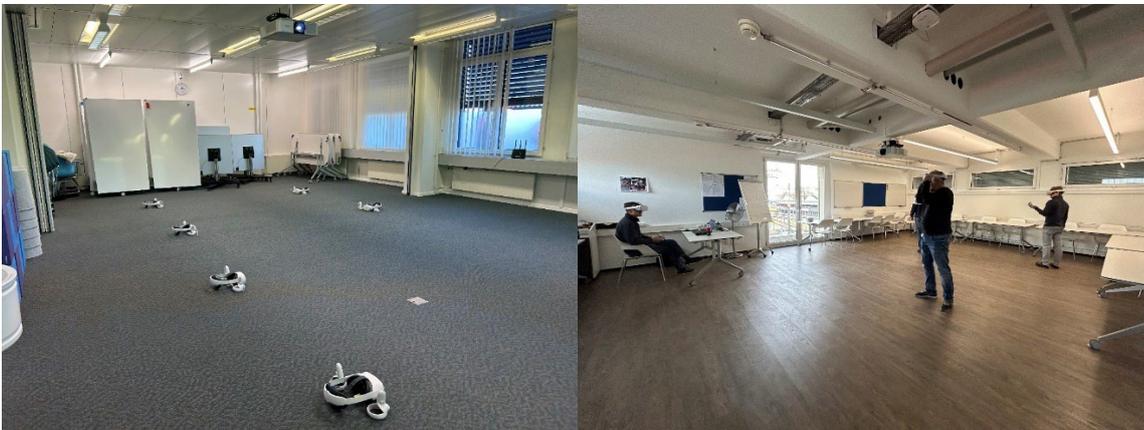


Abbildung 5.3. Erhebungen in Zürich Altstetten und Olten

5.3.6 Probleme bei der Datenerhebung

Bevor die Datenerhebung im Januar starten konnte, fiel auf, dass zwischen den beiden Teams Ost (Standort Zürich) und Mitte (Standorte Olten und Bern) grosse Unterschiede bestehen, was das Wissen über die Anwendung der HMDs und der Drehbücher anbelangt. Aus diesem Grund wurde im Dezember ein «Train-the-Trainer» Tag organisiert, an welchem alle Trainer zusammenkamen. Das Ziel hinter diesem Train-the-Trainer Tag war es, die Trainer im Umgang mit den HMDs vertraut zu machen und ihnen die Drehbücher zu erläutern. Gleichzeitig konnten sie die Anwendung selbst üben. Des Weiteren sollte dieser Tag die Möglichkeit geben, offene Fragen zu klären und allen Trainern den Ablauf des Experiments zu vermitteln. Im Februar wurde ein zusätzliches Training der Trainer durchgeführt, da zu Beginn der Erhebungen ersichtlich wurde, dass einige Trainer immer noch Schwierigkeiten bei der Durchführung und Bedienung der HMDs hatten. Dieser zweite Tag konnte den Trainern nochmals viel Sicherheit vermitteln, was sich positiv im weiteren Verlauf der Erhebungen ausgezeichnet hat.

5.3.7 Ablauf der Datenerhebung

Die Abläufe gestalteten sich je nach Bedingung leicht unterschiedlich. Die Abläufe werden nachfolgend aufbauend auf Tabelle 5.1 genauer beschrieben:

Bei allen iVR-Szenarien erhielten die Teilnehmenden zuerst eine Einführung auf der VR-Brille, um die Handhabung der Controller kennenzulernen und sich in der virtuellen Welt zurechtzufinden (siehe Abbildung 5.4). Die Teilnehmenden spielten hierzu das Tutorial selbstständig durch. Das Tutorial dauerte ungefähr 7 Minuten und wurde von der virtuellen Agentin Emma begleitet. In einem ersten Schritt zeigte Emma auf, wie die Fortbewegung mittels des Fortbewegungsstrahls funktioniert. Die Teilnehmenden konnten sich anschliessend in einer ersten Aufgabe an einen vorgegebenen Punkt fortbewegen. In der zweiten Aufgabe mussten die Teilnehmenden den T-Schlüssel, welcher sich auf einem Tisch befand, greifen und wieder loslassen. In der dritten und letzten Aufgabe lernten die Teilnehmenden, wie der T-Schlüssel zum Anziehen der losen Schrauben benutzt wird. Hierfür haben die Teilnehmenden drei Schrauben angezogen. Auf die Aufgabe mit dem Funkgerät wurde mangels Einsatzes im Training verzichtet.

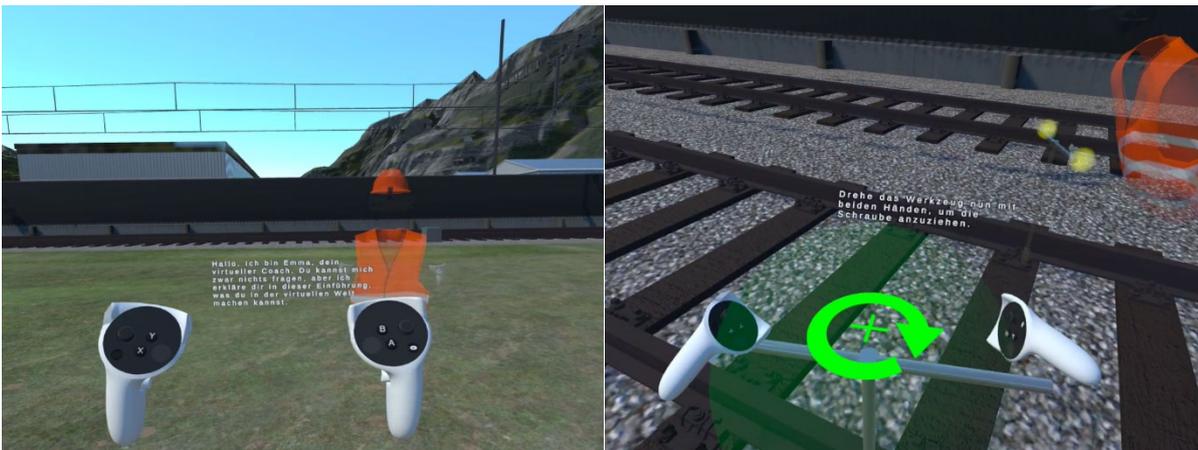


Abbildung 5.4. Aufgaben im Tutorial

In einem nächsten Schritt fand die Sensibilisierung der Teilnehmenden statt. Hierzu haben sich alle Teilnehmenden in die gleiche virtuelle Lernumgebung begeben und konnten sich auch gegenseitig sehen. Der Kursleiter forderte die Teilnehmenden dazu auf, sich in die Mitte der beiden Gleise zu begeben. Daraufhin erfolgte eine Zugsdurchfahrt bei schlechter Sicht mit 160 km/h. Das Ziel dieser Übung war die Sensibilisierung auf plötzlich ändernde Sichtverhältnisse. Nach dieser Sensibilisierung konnten die Teilnehmenden ihre VR-Brille ausziehen und unter Leitung des Kursleiters die Checkliste Selbstschutz gemäss einem projizierten Auftrag ausfüllen. Nach dem Ausfüllen und der Besprechung der Checkliste Selbstschutz betraten die Teilnehmenden die virtuelle Welt wieder mittels HMD. Sie spielten das iVR-Singleplayer Szenario selbstständig durch und wurden dabei durch eine Stimme respektive

durch Text unterstützt. In den anderen iVR-Bedingungen wurden sie vom Trainer angeleitet und bekamen von diesem Feedback. Die Teilnehmenden spielten die vier nachfolgenden Szenarien in unterschiedlicher Reihenfolge durch:

- Regenwetter setzt ein (Sicht 250m und Zug fährt durch)
- Regenwetter setzt ein (Sicht 320m und Zug fährt durch)
- Nebelwetter setzt ein (Sicht 250m und Zug fährt durch)
- Nebelwetter setzt ein (Sicht 320m und Zug fährt durch)

Die Aufgabe der Teilnehmenden war es, selbständig zu merken, wann sie aufgrund der Sichtverhältnisse oder eines sich nähernden Zuges die Gleise zu verlassen haben und sich in den Fluchraum zurückziehen müssen. Von Bedeutung war dabei die Einhaltung der vorher mittels Checkliste Selbstschutz berechneten Räumungszeit von 8 Sekunden sowie die daraus berechnete Mindestsichtweite von 320 Metern. Je nach Verhalten des Teilnehmenden, gab die virtuelle Agentin im IVR-Singleplayer oder der Desktopbedingung eine entsprechende Rückmeldung. Die gegebenen Rückmeldungen sind nachfolgend ersichtlich:

1. Negativ Sicht: Du hast dich auf dem Gleis befunden, obwohl die Sichtweite beim Regen nur 250 Meter betragen hat.
2. Negativ Sicht und Zug: Du bist zu langsam in den Fluchraum gelaufen und hast dich auf dem Gleis befunden, obwohl die Sichtweite nur 250 Meter betragen hat.
3. Negativ Fluchraum: Du hättest nicht in den Fluchraum gehen müssen, da ausreichend Sicht (mindestens 320m) vorhanden war.
4. Negativ Zug: Du bist zu langsam (mehr als 8 Sekunden) in den Fluchraum gegangen.
5. Positiv: Du hast alles richtig gemacht!

In den anderen iVR-Bedingungen erfolgte die Rückmeldungen durch den Trainer. Nachdem die Teilnehmenden alle vier Szenarien durchgespielt haben, erhielten sie eine Session-ID in der VR-Umgebung angezeigt. Die Session-ID wurde eingesetzt, um die Anonymität der Teilnehmenden zu gewährleisten. Zu Beginn des Fragebogens trugen die Teilnehmenden ihre Session-ID ein. Des Weiteren wurde die Kursliste der Teilnehmenden mit den Session-IDs ergänzt. So konnten die Trainer am zweiten Kurstag die Punktzahl für jede Person eintragen und anschliessend die Namen der Teilnehmenden entfernen. So konnte der Fragebogen mit der erzielten Punktzahl in Verbindung gebracht sowie die Anonymität der Teilnehmenden beibehalten werden.



Abbildung 5.5. Singleplayer Desktop Szenario

5.3.8 Fragebogen

Den Fragebogen, welchen die Teilnehmenden jeweils am Ende ausgefüllt haben, wurde mit der Software Enterprise Feedback Suite (Tivian XI GmbH, 2023) erstellt. Der Fragebogen enthielt zwei allgemeine Fragen, vier demographische Fragen sowie 49 Items aus den Modellen TAM, CAMIL und Fragen zur Arbeitseinschätzung der SBB, die für diese Studie relevant sind.

Für die Messung der Technologieakzeptanz als ein Mass der Arbeitseinschätzung im Lernkontext wurden Items zu den Dimensionen wahrgenommene Benutzerfreundlichkeit, wahrgenommener Nutzen sowie Nutzungsabsicht aus dem TAM verwendet, welche eine Reliabilität von Cronbachs $\alpha = .89$ beziehungsweise $.83$ aufweisen (Özgen, Afacan & Sürer, 2021). Weiterhin wurden Fragen zur Arbeitseinschätzung als Nutzen des jeweilig verwendeten Lernmediums in dem GK SstA in der Praxis von der SBB entwickelt (siehe oben) und erfasst. Die Fragen lassen sich dem Situationsbewusstsein und der Gefahrenwahrnehmung zuordnen. Fasst man die Items als Skala zusammen weist diese eine Reliabilität von Cronbachs $\alpha = .88$ auf. Auf eine Faktorenanalyse wurde angesichts der Stichprobengrösse von 127 Personen verzichtet. Dies aufgrund dessen, dass sowohl Comrey und Lee (2013) als auch Bühner (2011) erst eine Stichprobengrösse ab 200 Personen als angemessen für eine Faktorenanalyse beurteilen. In Tabelle 5.2 sind Beispielimite des Fragebogens aufgeführt. Der komplette Fragebogen ist in Anhang D einzusehen.

Beispielitems zum Fragebogen sind in Tabelle x dargestellt. Für die Beantwortung der Forschungsfragen relevante Skalen zur Auswertung sind rot umrandet.

Tabelle 5.2 Skalen und Item Beispiele mi für die Studie relevanten Items in rot umrandet

Skala	Beispielitem	Quelle
Vergnügen / Interesse	«Ich fand die Tätigkeiten mit dem Lerntool sehr interessant»	Wilde et al. (2009)
Wahr. Kompetenz	«Bei den Tätigkeiten mit dem Lerntool stellte ich mich geschickt an»	
Physische Präsenz	«Das Lerntool hat mich völlig in den Bann gezogen»	Makransky et al. (2017)
Handlungsfähigkeit	«Während der Arbeit mit dem Lerntool hatte ich das Gefühl, dass meine Erfahrungen und Handlungen von mir selbst gesteuert wurden»	Polito et al. (2013)
Selbstwirksamkeit	«Ich bin zuversichtlich, dass ich einem Freund/ einer Freundin die Grundlagen von Arbeitsstellensicherheit erklären kann»	Meyer et al. (2019)
ECL-I	«Die bei der Arbeit mit dem Lerntool verwendete Interaktionstechnik war sehr unklar»	Andersen & Makransky (2021)
ECL-E	«Das Lerntool war voll von unwichtigen Inhalten»	
Situatives Interesse	«War die Arbeit mit dem Lerntool für Sie spannend?»	Knogler et al. (2015)
Wahr. Benutzerfreundlichkeit	«Ich war bei der Arbeit mit dem Lerntool selten frustriert»	Özgen et al. (2021)
Wahr. Nützlichkeit	«Ich würde die Nutzung des Lerntools als Vorbereitung auf meine Arbeit nützlich finden»	
Nutzungsabsicht	«Sobald das Lerntool als Lernmethode in meiner Ausbildung verfügbar wäre, würde ich es nutzen»	
Arbeitseinschätzung	«Die Gefahren im Gleisbereich konnten mir mit dem Lerntool sehr deutlich gemacht werden»	Paul Hügli (SBB)

Englischsprachige Items wurden vorab auf Deutsch übersetzt. Einige Fragen wurden zugunsten der Zielgruppenangemessenheit umformuliert. Das Wort Simulation wurde auf Lerntool angepasst, um eine bessere Verständlichkeit herzustellen. Des Weiteren wurden Wörter wie ineffektiv und Interaktionstechnik jeweils ergänzend zu der Frage erklärt. Es wurde eine 5-Punkte-Likert-Skala benutzt, welche von 1 (stimme überhaupt nicht zu) bis 5 (stimme voll und ganz zu) reicht. Lediglich bei den Fragen zum Situationsbewusstsein und der Gefahrenwahrnehmung wurde eine 5-Punkte-Likert-Skala verwendet, welche von 1 (trifft überhaupt nicht zu) bis 5 (trifft vollumfänglich zu) reicht.

5.4 Ergebnisse

Insgesamt haben an 34 SstA-Kursen 127 Personen freiwillig an der Studie teilgenommen. Die nachfolgende Tabelle 5.3 gibt eine Übersicht über die Zusammensetzung der Stichprobe.

Tabelle 5.1 Übersicht Zusammensetzung der Stichprobe

Gruppe	N	M Alter/ Altersspanne	VR-Erfahrung	Geschlecht	Team (Personen)	Ort (Personen)
Regulärer Kurs	23	19,43 Jahre 15-41 Jahre	13 Personen	19 Männer 3 Divers 1 ohne Angabe	Ost (23)	Zürich (23)
iVR-Singleplayer	20	37,45 Jahre 16-53 Jahre	10 Personen	1 Frau 19 Männer	Mitte (13) Ost (7)	Olten (13) Zürich (7)
Singleplayer Desktop	20	39,40 Jahre 23-60 Jahre	7 Personen	1 Frau 18 Männer 1 ohne Angabe	Mitte (20)	Bern (2) Olten (18)
iVR-Multiplayer	23	35,04 Jahre 21-60 Jahre	5 Personen	4 Frau 18 Männer 1 ohne Angabe	Mitte (4) Ost (19)	Olten (4) Zürich (19)
iVR-Multiplayer mit retrospektivem Perspektivwechsel	21	38,33 Jahre 20-62 Jahre	8 Personen	1 Frau 20 Männer	Mitte (9) Ost (12)	Olten (9) Zürich (12)
iVR-Classroom	20	36,45 Jahre 23-59 Jahre	6 Personen	1 Frau 19 Männer	Mitte (15) Ost (5)	Olten (15) Zürich (5)
Alle Gruppen	127	34,05 Jahre 15-62 Jahre	49 Personen	8 Frauen 113 Männer 3 Divers 3 ohne Angabe	Mitte (61) Ost (66)	Bern (2) Olten (59) Zürich (19)

Zur Beantwortung der ersten Fragestellung wurde eine binominale logistische Regression mit kategori-alem Prädiktor gerechnet. Für die Beantwortung der zweiten Fragestellung erfolgte Kruskal-Wallis-Tests (KWTs) mit Einzelvergleichen.

5.4.1 Voraussetzungsprüfung der Daten

Die Voraussetzungen sind hierbei das die abhängige Variable nominalskaliert ist (was in unserem Fall volle Punktzahl oder Punktzahl kleiner 4 bedeutet), die unabhängige Variable nominalskaliert ist (Zuordnung zu den Trainingsbedingungen via Dummy-Codierung), Unabhängigkeit der Beobachtungen (ist gegeben, jede Person hat nur einmal teilgenommen) und weiter Voraussetzungen wie Multikollinearität, Linearität, Ausreisser etc. die im vorliegenden Fall keine Rolle spielen, da wir nur einen Prädiktor haben der Dummy-codiert wird. Berichtet werden die Güte des Modells und die Odds Ratio als Chance im Vergleich zur Kontrollgruppe volle Punktzahl zu bekommen. Deskriptiv wird die Anzahl voller Punktzahl vs. nicht volle Punktzahl als Prozentuales Verhältnis aufgezeigt.

Zur Beantwortung der zweiten Fragestellung wurde u.a. aufgrund fehlender Normalverteilung der Skalen non-parametrische Kruskal-Wallis Tests (KWT) angewendet und Einzelvergleiche inklusive Bonferroni-Korrektur ausgegeben. Berichtet werden die KWTs die gesamthaft überzufällig wurden. Deskriptiv werden die Einzelvergleiche mit einem Stern markiert deren Unterschiede innerhalb des KWTs signifikant wurde.

5.4.2 Punktzahl nach den Trainings

Die deskriptiven Daten weisen auf Unterschiede in den sechs verschiedenen Trainingsbedingungen hin. Es zeigt sich, dass die Anzahl der meisten Personen mit nicht voller Punktzahl (44%) im regulären Kurs zu finden sind. Gleichzeitig haben im regulären Kurs 56% der Teilnehmenden volle Punktzahl. Im Vergleich dazu zeigen alle zusätzlichen digitalen Interventionen eine Zunahme an Personen mit voller Punktzahl (blauer Kreisinhalt in Abbildung 5.6). Diese reichen von 87% volle Punktzahl in der iVR-Multiplayer Bedingung, 90,5%/ 90% Mehrspieler mit retrospektiven Perspektivwechsel/Desktop PC, sowie 95% volle Punktzahl in der iVR-Singleplayer sowie iVR-Classroom Bedingung (siehe Abbildung 5.6).

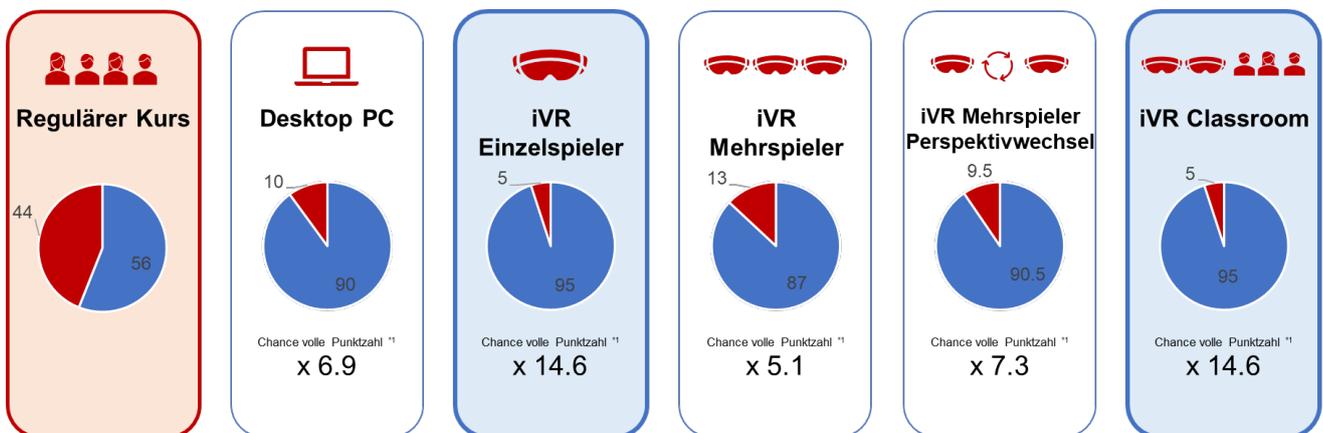


Abbildung 5.6: Prozentuale Rate an voller (blau) zu nicht voller (rot) Punktzahl betreffend die Frage zur Annäherungsdis-tanz/Checkliste in der SstA Prüfung in den verschiedenen Bedingungen. Odds Ratios als x- fache Chance (z.B. 14.6-fach höhere Chance im iVR-Classroom) volle Punktzahl zu erreichen im Vergleich zum regulären Kurs.

Die binominale logistische Regression mit kategorialem Prädiktor ergab ein signifikantes Modell (Chi-Quadrat=15.797, df= 5, p=.007) mit mittlerer Modellgüte (-2Log-Likelihood = 91.397, Cox & Snell R²=.117, Nagelkerkes R²=.205), dass eine erhöhte Chance volle Punktzahl in allen digitalen Interventionen im Vergleich zum regulären Kurs zu erreichen, aufzeigte. Die Daten in Tabelle 5.4 zeigen analog zu den Odds Ratios in Abbildung 5.6 die Odds Ratios. Hierbei haben iVR-Classroom und iVR-Singleplayer die höchste Chance (x14.6) volle Punktzahl zu bekommen. Die Werte steigen ab und zeigen

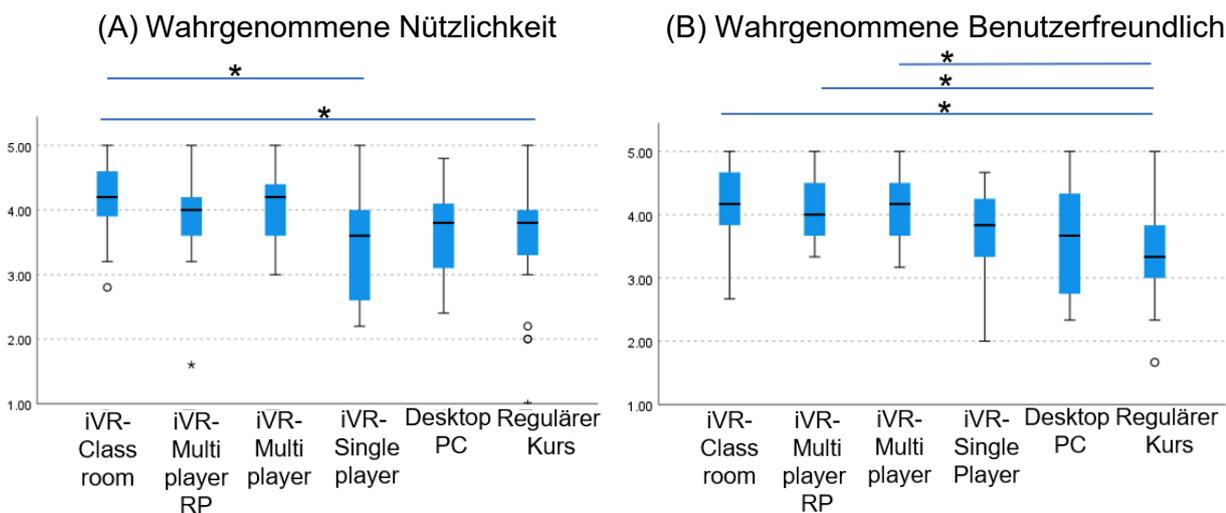
bei iVR- Multiplayer nur noch eine 5,1 fache erhöhte Chance im Vergleich zur Kontrollgruppe volle Punktzahl zu bekommen.

Tabelle 5.4: Koeffizienten der Bedingungen

Variablenausprägung	Regressionskoeffizient	Standardfehler	Wald	df	Sig.	Exp(B)/Odds ratio
iVR-Classroom	2.682	1.109	5.851	1	.016	14.615
iVR-Multiplayer mit RPW	1.989	.854	5.422	1	0.20	7.308
iVR-Multiplayer	1.635	.749	4.770	1	0.029	5.128
iVR-Singleplayer	2.682	1.109	5.851	1	0.016	14.615
Desktop-PC	1.935	.856	5.111	1	0.024	6.923

5.4.3 Arbeitseinschätzung TAM-Skalen

Die Einschätzung der TAM Skalen zeigten bei den KWTs für Wahrgenommene Nützlichkeit (Chi-Quadrat = 17.115, df=5, p=.004), Wahrgenommene Benutzerfreundlichkeit (Chi-Quadrat =23.644, df=5, p<.001) und bei Nutzungsabsicht (Chi-Quadrat =15.477, df=5, p<.009) überzufällige Unterschiede in den Trainingsbedingungen (siehe Abbildung 5.7A-C). Jedoch unterschieden sich nicht alle Bedingungen voneinander. Nach Anpassung der Einzelvergleiche via Bonferroni Korrektur scheint der Einzelvergleich von iVR-Classroom und dem regulären Kurs sich in allen Skalen als überzufällig abzubilden.



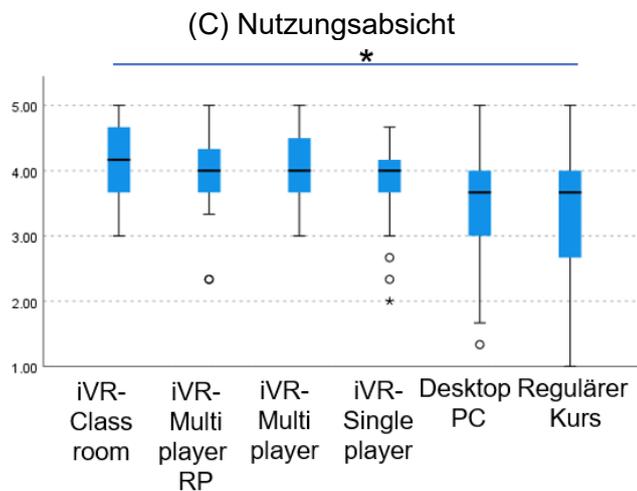


Abbildung 5.7A-C: Deskriptive Ergebnisse der signifikanten KWTs in den TAM Skalen Wahrgenommene Nützlichkeit, Wahrgenommene Nutzerfreundlichkeit und Nutzungsabsicht. Die Skalenmittelwerte (Einzelvergleiche) in den Trainingsbedingungen, die sich signifikant nach Bonferroni-Korrektur unterscheiden sind mit Sternchen markiert.

Gleichzeitig liegen deskriptiv die Skalenwerte der regulären Trainingsgruppe, Desktop PC und bei zwei Skalen auch der iVR-Singleplayer unter dem Wert «vier» und der Rest darüber.

5.4.4 Arbeitseinschätzung Situationsbewusstsein und Gefahrenwahrnehmung

In drei von fünf der von der SBB entwickelten Fragen (siehe Abbildung 5.8A-C) liessen sich ebenfalls überzufällige KWTs beobachten. Dies betraf das Item «Im Vergleich zu Schulungen ohne Lerntool fühle ich mich besser auf die Arbeit in der Praxis vorbereitet» ($\chi^2 = 20.099$, $df=5$, $p=.001$), das Item «Im Vergleich zu Schulungen ohne Lerntool fühle ich mich sicherer für die Arbeit in der Praxis vorbereitet.» ($\chi^2 = 15.972$, $df=5$, $p=.007$) und das Item «Die Gefahren im Gleisbereich konnten mir mit dem Lerntool sehr deutlich gemacht werden» ($\chi^2 = 16.887$, $df=5$, $p=.005$). In den Einzelvergleichen konnte nach Bonferroni-Korrektur wieder beobachtet werden, dass auch deskriptiv die Bewertungen vom iVR- Classroom am höchsten ausfielen. Gleichzeitig scheint sich iVR am häufigsten überzufällig von anderen Bedingungen zu unterscheiden. Diesmal vom Desktop, seltener vom regulären Kurs, der teilweise deskriptiv bessere Bewertungen in den Items erhielt als die digitalen Interventionen.

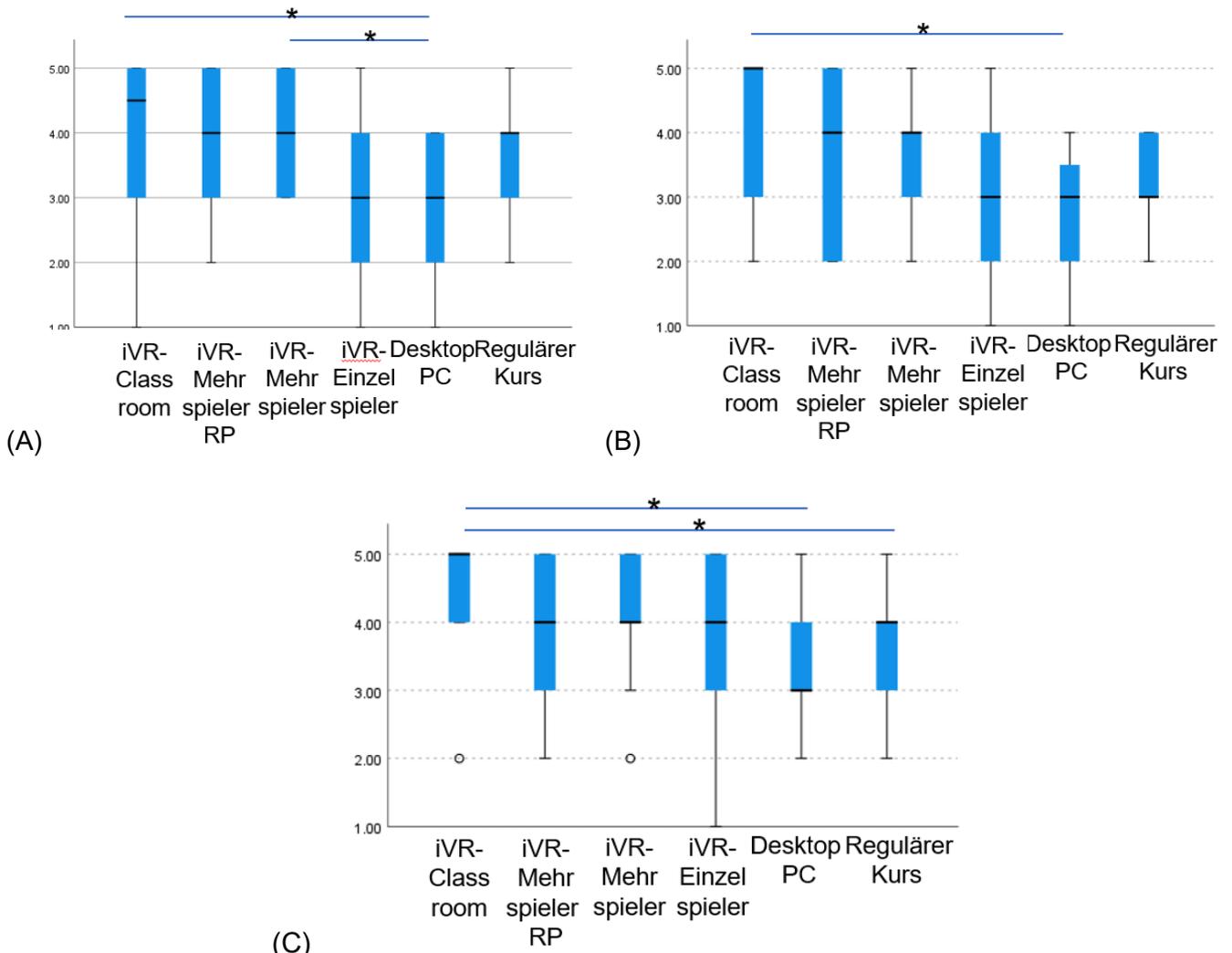


Abbildung 5.8A-C: Deskriptive Ergebnisse der signifikanten KWTs in den Items (A)«Im Vergleich zu Schulungen ohne Lern-tool fühle ich mich besser auf die Arbeit in der Praxis vorbereitet», (B)«Im Vergleich zu Schulungen ohne Lerntool fühle ich mich sicherer für die Arbeit in der Praxis vorbereitet.» und (C)«Die Gefahren im Gleisbereich konnten mir mit dem Lerntool sehr deutlich gemacht werden». Die Skalenmittelwerte (Einzelvergleiche) in den Trainingsbedingungen, die sich signifikant nach Bonferroni-Korrektur unterscheiden sind mit Sternchen markiert.

5.5 Diskussion

Mit dieser Hauptstudie wollten wir Antworten auf zwei Forschungsfragen finden: (1) Unterscheidet sich die Wirkung von zusätzlichen digitalen Trainingseinheiten (iVR- Singleplayer, iVR-Multiplayer, iVR-Multiplayer mit retrospektivem Perspektivwechsel, iVR-Classroom, Laptop PC) auf die Prüfungsleistung (gemessen in der vollen Punktzahl in der Prüfungsfrage zu Annäherungsdistanz/Checkliste) im Vergleich zu einer Kontrollgruppe im SstA-Kurs?

(2) Unterscheiden sich die Ausprägungen der Fragen zur Arbeitseinschätzung in den unterschiedlichen, zusätzlichen digitalen Trainingseinheiten (iVR- Singleplayer, iVR-Multiplayer, iVR-Multiplayer mit retrospektivem Perspektivwechsel, iVR-Classroom, Laptop PC) im Vergleich zu einer Kontrollgruppe (regulärer Kurs)?

Hierbei konnten wir zeigen, dass die digitalen Interventionen eine überzufällige Veränderung (Verbesserungen mit teils starken Effekten) bei den Kursteilnehmenden hinsichtlich der Prüfungsleistung bewirkten. Die stärksten Effekte im Vergleich zum regulären Training sind sowohl bei iVR-Singleplayer wie bei iVR-Classroom aufgetreten. Der kleinste Effekt ist bei iVR-Multiplayer aufgetreten. Dies zeigt zunächst, dass digitale Technik allein nicht besser unterrichtet, wie ein Trainer, der mit digitaler Technik unterrichtet. Beim Blick auf die TAM-Faktoren scheint sich nur die iVR-Classroom Bedingung im Vergleich zum regulären Training abzuheben, wenn auch mit kleinem Effekt. Hierbei könnte es eine Rolle spielen, dass der gemeinsame Moment im Unterricht auch ausserhalb des iVR-Raumes einen Einfluss auf die Bewertung haben kann. In unseren eigenen Studien (Christ et al, 2022) konnten wir zeigen, dass Teilnehmende, die ausserhalb der iVR-Welt etwas besprechen, gerne dies auch im iVR-Raum tun würden. Dies könnte zumindest die erhöhte Nutzungsabsicht erklären, da bei iVR-Classroom genau das getan wird. Pletz und Zinn (2018) stellten fest, dass Personen die Vorerfahrung mit iVR haben, die wahrgenommene Nützlichkeit und Nutzungsabsicht signifikant höher einschätzen. Hierzu können wir eher das Gegenteil aufzeigen, da bei uns mit 13 vorerfahrenen Personen der reguläre Kurs am meisten mit iVR vorerfahrene Personen aufwies. Auch die Ergebnisse von Özgen et al. (2021) können nur teilweise bestätigt werden. Sie fanden heraus, dass ein iVR-Training im Vergleich mit einem Paper-Based-Training ein signifikant höheres wahrgenommenes Vergnügen und Nutzungsabsicht bei iVR zur Folge hatte. Die erhöhte Nutzungsabsicht können auch wir bestätigen. Gleichzeitig zeigten die Autoren, dass bei der wahrgenommene Benutzerfreundlichkeit und wahrgenommene Nützlichkeit keine signifikanten Unterschiede auftraten (Özgen et al., 2021). Könnte es daran liegen, dass die Autoren keine Mehrspielmodi verwendeten und dieser jetzt bessere Werte erzeugt? In der ersten Multiplayer Studie von Pettersen et al. (2023) zeigt lernbezogene Vorteile von kollaborativen iVR-Erlebnissen gegenüber iVR. Einzelspieler. Jedoch wurden hier kein TAM-Daten erhoben. Es bleibt abzuwarten welche Faktoren diese Unterschiede von iVR Classroom auch in den Fragen zur Arbeitseinschätzung (Sicherheit etc.) bedingen. Hier konnte bei drei von fünf Fragen signifikante Unterschiede ausfindig gemacht werden, die vor allem durch den Unterschied von iVR Classroom zu Desktop-PC bedingt scheinen.

All dies zusammengefasst wirkt auf den ersten Blick erfreulich, muss aber auf den zweiten Blick mit Vorsicht interpretiert werden. Es ist zum jetzigen Zeitpunkt nicht möglich eine eindeutige Ursache für die Ergebnisse als Grund aufzuführen. Dies hat mehrere Gründe, die im Folgenden diskutiert werden.

5.5.1 Probanden

Die Probanden wurden nicht zufällig den Bedingungen zugeteilt. Der Grund hierfür lag in der Verfügbarkeit von Räumen in denen nur bestimmte Bedingungen getestet werden konnten (z.B. im Raum Neptun in Bern war kein iVR möglich oder Räume in Olten, bei denen die iVR Intervention Multiplayer immer noch einen zusätzlichen Raum gebraucht hat, dieser aber wegen anderen Kursen nicht zur Verfügung stand). Weiterhin hatte die Freiwilligkeit der Teilnahme die Konsequenz, dass das FHNW-Team nicht wusste, wieviel Probanden mitmachen werden. Dies hatte wiederum einen Einfluss auf welche Bedingungen noch Probanden benötigte. Somit wurden dann entsprechend dem «IST» (Anzahl an Probanden, die bereits erhoben wurden) mit weiteren Probanden aufgefüllt. Nur so konnten wir die notwendigen, etwa 20 Personen pro Bedingung, sicherstellen. Abschliessend bedingt die Freiwilligkeit auch ein wenig die «Besonderheit» der Teilnehmenden. Wir wissen nicht, ob diese Teilnehmenden eine besondere Affinität für digitale Medien haben, somit die Intervention besonders gut zum Gesamtsample gepasst haben. Hinsichtlich der Vorerfahrung haben 49 Personen von 127 Teilnehmenden Erfahrung mit einer VR-Brille gehabt. Dies sind 38%. In zukünftigen Studien sollten Fragen zur technischen Affinität beantwortet werden können, damit das Sample einordnungsfähig ist und man in Bezug auf mögliche, zukünftige Lernpopulation kein verzerrtes Bild erhält.

5.5.2 Interaktion Trainer und Technik

Es ist wichtig zu verstehen, dass wir bei jeder Erhebung, unabhängig von der Bedingung, den Einfluss der acht unterschiedlichen Trainer mitgemessen haben. Einige wirkten bis zum Schluss eher unsicher, andere sind während der Studiendauer über sich hinausgewachsen und wiederum andere waren bereits von Anfang an Ambassadeure. Auch war die Möglichkeit des Austausches in den Teams unterschiedlich, da manche Führungskräfte Fachführung praktizieren und somit inhaltlich beratschlagen und unterstützen konnten. Andere Führungskräfte arbeiten als Personalführung, ohne weiteren inhaltlichen Bezug und stehen somit einem inhaltlichen Austausch nicht zur Verfügung. Die unterschiedlichen Pensen einiger Personen (40% Trainer, 60% Arbeit am Gleis) machte ihren Einsatz als Trainer eher unwahrscheinlich andere hingegen durften mehrmals mitmachen. Hier wären am Ende Interviews schön gewesen, da wir nicht wissen, wie es den Trainern während der Erhebungsphase ging und was man in Zukunft verbessern sollte und was man beibehalten kann. Abschliessend ist festzuhalten, dass auch die Räume unterschiedlichen Einfluss auf die Teilnehmenden gehabt haben, da einige einfach, andere wiederum mit mehr Aufwand vorzubereiten waren. Dies kann beim Trainer unterschiedlich viel Stress ausgelöst haben, was sich wiederum in seine Performance im virtuellen Raum niederschlagen kann.

5.5.3 iVR Classroom

Neben den vielen Unsicherheiten, wie das eigentlich positive Ergebnis ursächlich verstanden werden kann, sticht eine iVR- Bedingung besonders hervor: iVR-Classroom. Bereits im Workshop der Studie I viel den Trainern diese Bedingung besonders positiv auf. In unseren Daten zeigen sich sowohl in den subjektiven Daten der Fragebögen zum TAM oder der Arbeitseinschätzung, die meisten überzufälligen Unterschiede zu Desktop Gruppe oder dem regulären Kurs. Gleichzeitig erzielte sie die höchste Anzahl an Personen mit voller Punktzahl und benötigt am wenigste Ressourcen. Hier empfiehlt es sich dieser didaktischen Variation ein verstärktes Augenmerk zukommen zu lassen und ggf. zusammen mit Trainern und unter Erwägung einer-Follow-up-Studie einen weiteren Einsatz zu planen. Da der wissenschaftliche Vergleich der iVR-Methoden komplettes Neuland auf wissenschaftlicher Ebene darstellt, können keine Vergleiche mit bestehender Literatur herangezogen werden. Jedoch sind die Ergebnisse es wert, weiter diese didaktische Form auch für andere iVR-Szenarien im Auge zu behalten.

5.6 Fazit

In vorliegender Studie hat sich der Lernerfolg in iVR-Lernsettings signifikant verbessert. Grundsätzlich wurde iVR im Vergleich zu konventionellen Methoden in dieser Studie eine gute Usability zugeschrieben und die Technologie wurde gut akzeptiert. Es ist jedoch fraglich, ob das iVR- Singleplayer-Setting für eine kurze Intervention, oder allgemein kürzere Schulungen, der optimale Ansatz ist. Bei den iVR-Mehrspielersettings werden möglicherweise aufgrund der unverzüglichen Unterstützung durch die Auszubildenden und der Peers der Einstieg und die Handhabung vergleichsweise erleichtert. In der Nutzungsabsicht haben sich die iVR-Lernsettings in dieser Studie nicht unterschieden, nur die iVR-Classroom Bedingung wies im Vergleich zum regulären Kurs eine signifikant höhere Nutzungsabsicht auf. Grundsätzlich konnte das Training mit iVR-Classroom im Vergleich zum regulären Kurs als ein herausstechenden didaktisches Verfahren beobachtet werden.

6 Ausblick

Das Projekt SUCAP hatte das Ziel mögliche Erfolgs- oder Misserfolgskriterien von immersiver virtueller Realität (iVR) im Bereich Gleissicherheitstrainings zu bestimmen. Hierbei sollen Spezifikationen wie das Lernsetting iVR- Einzelspieler, iVR-Mehrspieler, iVR-Mehrspieler mit retrospektivem Perspektivwechsel (Aufnahme eigener Handlung am Gleis und Wiedergabe 30 Sek. in der Vergangenheit mit Sicht durch das Zugführeresstand) iVR-Classroom und Hardware (VR-Brille, Laptop) in der Gleissicherheitsschulung (Aufgabe der Checkliste/Annäherungsdistanz) überprüft werden.

Wir konnten die Erfolgsfaktoren Präsenzzufriedenheit, Handlungsfähigkeit, Interesse und Wahrgenommene Kompetenz als wichtige Stellschrauben identifizieren, die es ermöglichen, menschliche Bewertungsdimensionen wie Nutzungsintention, Benutzerfreundlichkeit und wahrgenommene Nützlichkeit positiv vorhersagen können. Im Vergleich zu einem regulären Kurs zeigten alle zusätzlichen digitalen Interventionen in der Menge erreichter voller Punktzahl einen teilweisen grossen Zuwachs. Zu Beginn des Projekts wiesen uns die Experten darauf hin, dass die Teilprüfungsleistung Checkliste/Annäherungsdistanz meist nicht vollständig gelöst wird/ weniger oft volle Punktzahl erreicht wird. Sie schlossen weiterhin darauf, dass diese Teilaufgabe einen wesentlichen Beitrag zum nicht bestehen der gesamten Prüfung liefert. Betrachtet man hierzu nochmals die Daten von 2020 zur Prüfungsleistung zeigt sich folgendes Bild (siehe Abbildung 6.1)

Test	Number of participants	Number passed	Number failed
Sst B Gk	1096	1062	34
Sst A Gk	1726	1539	187
Sst A Wk	1173	842	331

Abbildung 6.1: Ausschnitt aus dem SUCAP-Projektantrag. 2020 fielen bei 2899 SstA-Kursen 518 Personen durch die Prüfung (17%). Bei etwa 200 CHF.- pro Wiederholungsprüfung machen das 103000.- CHF zusätzlichen Aufwand.

Würde man durch geeignete Massnahmen die Durchfallquote halbieren oder ganz verringern, wären Einsparpotentiale von 51500- 103000 CHF möglich.

Wir konnten für ein schwieriges Themenfeld der Gleissicherheit evidenzbasiert eine iVR- (Mehrspieler) Trainingsumgebung inklusive Tutorial und Desktopversion entwickeln. Dieses soll, nach Auskunft von Joe Scheidegger, das AppBackery Team der SBB als ersten Use Case auf eine neue Plattform der SBB migrieren. Mit dieser ersten Entwicklung konnten wir die Anzahl der Personen ohne volle Punktzahl auf bis zu 5% senken. Würde es uns gelingen die iVR- Umgebung in weiteren schwierigen Themen einzusetzen und damit die Gesamtprüfungsleistung ebenfalls auf 5% an Personen, die nicht bestehen, reduzieren, wären bei Betrachtung der Rechnung aus dem Jahr 2020 5% von 2899 145

Personen anstelle von 518 Personen durchgefallen. Dies würde bedeuten, dass bei $145 \cdot 200.-\text{CHF.} = 29\text{k.-CHF}$ Folgekosten den Betrag, der durch Wiederholungsprüfungen entsteht, um 71 % oder 74k.-CHF senken könnten. Eine Verbesserung dieser Situation würde also nicht nur dem Menschen beim Lernen unterstützen, sondern könnte auch Kosten von ca. 74k CHF.- einsparen. Wir wissen jedoch nicht, ob auch diese Ergebnisse in kontrollierten, randomisierten Studienbedingungen weiterbestehen würde. Darüber hinaus fehlen Interviewdaten von den Sicherheitstrainern, die zeigen, wie sie sich in den unterschiedlichen Bedingungen gefühlt haben und wo sie eventuell weitere Herausforderungen im Praxisalltag sehen. Gleichzeitig sind die Trainer eine sehr wertvolle Ressource, da (1) sie wussten, wo Handlungsbedarf besteht (siehe regulärer Kurs 44% keine volle Punktzahl), (2) sie wissen, was die iVR- Lernumgebung an Funktionen braucht (siehe Studie I) damit ein Erfolg möglich ist, der in diesem Projekt definitiv nicht zum Nachteil der SsttA-Teilnehmenden war. Es wäre also in Zukunft den Kursteilnehmenden geschuldet, dass sich die SBB verstärkt mit den Thema iVR in Schulungsszenarien auseinandersetzt.

7 Literatur

- Ai-Lim Lee, E., Wong, K. W. & Fung, C. C. (2010). How does desktop virtual reality enhance learning outcomes? A structural equation modeling approach. *Computers & Education*, 55(4), 1424–1442. <https://doi.org/10.1016/j.compedu.2010.06.006>
- Ajzen, I. (1985). From intentions to actions: A theory of planned behavior (pp. 11-39). Springer Berlin Heidelberg.
- Ajzen, I. & Fishbein, M. (1980). Understanding attitudes and predicting social behaviour. Englewood-Cliffs, NJ: Prentice Hall.
- Allcoat, D. & von Mühlengen, A. (2018). Learning in virtual reality: Effects on performance, emotion and engagement. *Research in Learning Technology*, 26. <https://doi.org/10.25304/rlt.v26.2140>
- Andersen, M. S. & Makransky, G. (2021). The validation and further development of a multidimensional cognitive load scale for virtual environments.pdf. *Journal of Computer Assisted Learning*, (37), 183–196. <https://doi.org/10.1111/jcal.12478>
- Ateş, C., Kaymaz, Ö., Kale, H. E. & Tekindal, M. A. (2019). Comparison of Test Statistics of Nonnormal and Unbalanced Samples for Multivariate Analysis of Variance in terms of Type-I Error Rates. *Computational and Mathematical Methods in Medicine*, 2019, 1–8. <https://doi.org/10.1155/2019/2173638>
- Austin, P. C., & Steyerberg, E. W. (2015). The number of subjects per variable required in linear regression analyses. *Journal of Clinical Epidemiology*, 68(6), 627–636. <https://doi.org/10.1016/j.jclinepi.2014.12.014>
- Bangor, A. (2009). Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *Journal of Usability Studies*, 4(3), 114–123.
- Bargas-Avila, J. A. & Hornbæk, K. (2011). Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (S. 2689–2698). Gehalten auf der CHI '11: CHI Conference on Human Factors in Computing Systems, Vancouver BC Canada: ACM. <https://doi.org/10.1145/1978942.1979336>
- Baylor, A. L. (2009). Promoting motivation with virtual agents and avatars: role of visual presence and appearance. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535), 3559–3565. <https://doi.org/10.1098/rstb.2009.0148>
- Baylor, A. L. & Ryu, J. (2003). THE PSYCHOMETRIC STRUCTURE OF PEDAGOGICAL AGENT PERSONA. *EdMedia+ innovate learning, Association for the Advancement of Computing in Education*, 448–451.

- Bohnsack, Ralf (2000). Gruppendiskussion. In Uwe Flick, Ernst von Kardoff & Ines Steinke (Hrsg.), *Qualitative Forschung. Ein Handbuch* (S.369-384). Reinbek: Rowohlt.
- Boyd, J., Barnett, S. W., Bodrova, E., Leong, D. J., & Gomby, D. (2005). Promoting children's social and emotional development through preschool education. In National Institute for Early Education Research *Preschool Policy Brief*. New Brunswick, NJ: Rutgers.
- Bowman, D. A., Gabbard, J. L. & Hix, D. (2002). A Survey of Usability Evaluation in Virtual Environments: Classification and Comparison of Methods. *Presence: Teleoperators and Virtual Environments*, 11(4), 404–424. <https://doi.org/10.1162/105474602760204309>
- Bowman, D. A., & McMahan, R. P. (2007). Virtual reality: How much immersion is enough? *Computer*, 40(7), 36–43. <https://doi.org/10.1109/MC.2007.257>
- Brehmer, J. & Becker, S. (2017). *E-Learning – ein neues Qualitätsmerkmal der Lehre*. Göttingen: Georg-August-Universität
- Brooke, J. (1986). *SUS - A quick and dirty usability scale*.
- Brosius, H., Haas, A. & Koschel, F. (2016). *Methoden der empirischen Kommunikationsforschung. Eine Einführung*. Wiesbaden: Springer Fachmedien.
- Brühlmann, F. & Schmid, G.-M. (2015). How to Measure the Game Experience?: Analysis of the Factor Structure of Two Questionnaires. *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems* (S. 1181–1186). Gehalten auf der CHI '15: CHI Conference on Human Factors in Computing Systems, Seoul Republic of Korea: ACM. <https://doi.org/10.1145/2702613.2732831>
- Brünken, R., Seufert, T. & Paas, F. (2010). Measuring Cognitive Load. In J.L. Plass, R. Moreno & R. Brünken (Hrsg.), *Cognitive Load Theory* (1. Auflage, S. 181–202). Cambridge University Press. <https://doi.org/10.1017/CBO9780511844744.011>
- Burghofer, K. & Lackner, C.K. (2014). Simulationstraining zwischen «human factors» und «technical skills». *Notfall Rettungsmed*, 17 (5), 386-392. DOI 10.1007/s10049-013-1801-z
- Buttussi, F. & Chittaro, L. (2021). Locomotion in Place in Virtual Reality: A Comparative Evaluation of Joystick, Teleport, and Leaning. *IEEE Transactions on Visualization and Computer Graphics*, 27(1), 125–136. <https://doi.org/10.1109/TVCG.2019.2928304>
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson.
- Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior & Organization*, 81(1), 1-8. <https://doi.org/10.1016/j.jebo.2011.08.009>
- Chen, T., Chen, J., Or, C. K., & Lo, F. P. (2022). Path analysis of the roles of age, self-efficacy, and TAM constructs in the acceptance of performing upper limb exercises through immersive virtual

- reality games. *International Journal of Industrial Ergonomics*, 91, 103360.
<https://doi.org/10.1016/j.ergon.2022.103360>
- Christ, O. & Hirschi, M. (2021). Virtual Reality in der Berufsbildung. *Transfer, Berufsbildung in Forschung und Praxis* (1/2021), SGAB, Schweizerische Gesellschaft für angewandte Berufsbildungsforschung.
- Coban, M., Bolat, Y. I., & Goksu, I. (2022). The potential of immersive virtual reality to enhance learning: A meta-analysis. *Educational Research Review*, 36, 100452.
<https://doi.org/10.1016/j.edurev.2022.100452>
- Comrey, A.L. and Lee, H.B. (2013) *A First Course in Factor Analysis*, 2nd Edition, Psychology Press, Hove.
- Cummings, J. J. & Bailenson, J. N. (2016). How Immersive Is Enough? A Meta-Analysis of the Effect of Immersive Technology on User Presence. *Media Psychology*, 19(2), 272–309.
<https://doi.org/10.1080/15213269.2015.1015740>
- Dan, A. & Reiner, M. (2017). EEG-based cognitive load of processing events in 3D virtual worlds is lower than processing events in 2D displays. *International Journal of Psychophysiology*, 122, 75–84. <https://doi.org/10.1016/j.ijpsycho.2016.08.013>
- Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3), 319. <https://doi.org/10.2307/249008>
- Deci, E. L., Eghrari, H., Patrick, B. C. & Leone, D. R. (1994). Facilitating Internalization: The Self-Determination Theory Perspective. *Journal of Personality*, 62(1), 119–142.
<https://doi.org/10.1111/j.1467-6494.1994.tb00797.x>
- Deci, E. L., Koestner, R. & Ryan, R. M. (2001). Extrinsic Rewards and Intrinsic Motivation in Education: Reconsidered Once Again. *Review of Educational Research*, 71(1), 1–27.
<https://doi.org/10.3102/00346543071001001>
- Dewe, H., Gottwald, J. M., Bird, L.-A., Brenton, H., Gillies, M., & Cowie, D. (2022). My virtual self: The role of movement in children's sense of embodiment. *IEEE Transactions on Visualization and Computer Graphics*, 28(12), 4061–4072. <https://doi.org/10.1109/tvcg.2021.3073906>
- Diekmann, A. (2007). *Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen* (4. Aufl.). Reinbek bei Hamburg: Rowohlt.
- Di Natale, A. F., Repetto, C., Riva, G., & Villani, D. (2020). Immersive virtual reality in K-12 and higher education: A 10-year systematic review of empirical research. *British Journal of Educational Technology*, 51(6), 2006–2033. <https://doi.org/10.1111/bjet.13030>
- Domin, M., Janneck, M. & Grimm, S. (2019). Altersbezogene Unterschiede bei der Interaktion mit einem Virtual-Reality-System. In K. T, S. E & K. N (Hrsg.), *Communities in New Media:*

- Researching the Digital Transformation in Science, Business, Education and Public Administration - Proceedings of 22nd Conference GeNeMe* (S. 24–34). Dresden: TUD-press. Verfügbar unter: <https://dl.gi.de/bitstream/handle/20.500.12116/34912/ge-neme2019-13.pdf?sequence=1>
- Döring, N. & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Berlin Heidelberg: Springer Verlag.
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53(1), 109–132.
- Facebook. (n. d.). *Meta Quest 2: Tech specs*. *Meta Quest 2: Tech specs*. Zugriff am 19.8.2022. Verfügbar unter: <https://store.facebook.com/ch/en/quest/products/quest-2/tech-specs#tech-specs>
- Faric, N., Potts, H. W. W., Hon, A., Smith, L., Newby, K., Steptoe, A. et al. (2019). What Players of Virtual Reality Exercise Games Want: Thematic Analysis of Web-Based Reviews. *Journal of Medical Internet Research*, 21(9), e13833. <https://doi.org/10.2196/13833>
- Farshid, M., Paschen, J., Eriksson, T. & Kietzmann, J. (2018). Go boldly! *Business Horizons*, 61(5), 657–663. <https://doi.org/10.1016/j.bushor.2018.05.009>
- Faul, F. (2020, März 17). G*Power. Kiel: Universität Kiel. Verfügbar unter: <https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower>
- Field, A. (2017). *Discovering statistics using IBM SPSS statistics* (5th edition.). Thousand Oaks, CA: SAGE Publications.
- Finch, H. (2005). Comparison of the Performance of Nonparametric and Parametric MANOVA Test Statistics when Assumptions Are Violated. *Methodology*, 1(1), 27–38. <https://doi.org/10.1027/1614-1881.1.1.27>
- Frese, M. & Zapf, D. (1994). Action as the core of work psychology: A German approach. In H.C. Triandis, M.D. Dunette & L.M. Hough (Hrsg.), *Handbook of Industrial and Organizational Psychology* (2. Auflage, S. 272–340). Palo Alto: Consulting Psychologists Press. Verfügbar unter: https://www.researchgate.net/publication/232492102_Action_as_the_core_of_work_psychology_A_German_approach
- Gaba, D.M. (2004). The future vision of simulation in health care. *BMJ Quality & Safety*, 13, 2-10. doi: 10.1136/qshc.2004.009878
- Gabbard, J. L. (1997). *A Taxonomy of Usability Characteristics in Virtual Environments*. Virginia: Virginia Polytechnic Institute and State University. Verfügbar unter: <https://vte-chworks.lib.vt.edu/bitstream/handle/10919/35731/etd.pdf>
- Goldin-Meadow, S., Nusbaum, H., Kelly, S. D. & Wagner, S. (2001). Explaining Math: Gesturing Lightens the Load. *Psychological Science*, 12(6), 516–522. <https://doi.org/10.1111/1467-9280.00395>

- Graesser, A. C., Moreno, K., Marineau, J., Adcock, A. & Olney, A. (2003). AutoTutor Improves Deep Learning of Computer Literacy: Is it the Dialog or the Talking Head?, 9.
- Green, S. B. (1991). How Many Subjects Does It Take To Do A Regression Analysis. *Multivariate Behavioral Research*, 26(3), 499–510. https://doi.org/10.1207/s15327906mbr2603_7
- Grolnick, W. S. & Ryan, R. M. (1987). Autonomy in Children's Learning: An Experimental and Individual Difference Investigation. *Journal of Personality and Social Psychology*, 52(5), 890–898.
- Hahs-Vaughn, D. L. (2016). *Applied Multivariate Statistical Concepts*. New York, NY : Routledge, 2017.: Routledge. <https://doi.org/10.4324/9781315816685>
- Hamilton, D., McKechnie, J., Edgerton, E., & Wilson, C. (2021). Immersive virtual reality as a pedagogical tool in education: A systematic literature review of quantitative learning outcomes and experimental design. *Journal of Computers in Education*, 8(1), 1–32. <https://doi.org/10.1007/s40692-020-00169-2>
- Harackiewicz, J. M., Smith, J. L., & Priniski, S. J. (2016). Interest matters: The importance of promoting interest in education. *Policy Insights From the Behavioral and Brain Sciences*, 3(2), 220–227. <https://doi.org/10.1177/2372732216655542>
- Harrell, F. E. (2001). *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. Springer. <https://doi.org/10.1007/978-1-4757-3462-1>
- Hart, S. G. & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology* (Band 52, S. 139–183). Elsevier. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- Hassenzahl, M., Burmester, M. & Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität (Berichte des German Chapter of the ACM). In G. Szwillus & J. Ziegler (Hrsg.), *Mensch & Computer 2003* (Band 57, S. 187–196). Wiesbaden: Vieweg+Teubner Verlag. https://doi.org/10.1007/978-3-322-80058-9_19
- Heimsch, F. M., Niederer, R. & Zöfel, P. (2018). *Statistik im Klartext: für Psychologen, Wirtschafts- und Sozialwissenschaftler (ps) (2., aktualisierte und erweiterte Auflage.)*. Hallbergmoos: Pearson.
- Hemmerich, W. (2016). StatistikGuru: Johnson Transformation berechnen. Verfügbar unter: <https://statistikguru.de/rechner/johnson-transformation-berechnen.html>
- Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, 41(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4
- Hill, T., & du Preez, H. (2021). A longitudinal study of students' perceptions of immersive virtual reality teaching interventions. In 2021 7th International Conference of the Immersive Learning Research Network (iLRN) (pp. 1-7). IEEE. <https://doi.org/10.23919/iLRN52045.2021.9459334>

- Horota, R. K., Rossa, P., Marques, A., Gonzaga, L., Senger, K., Cazarin, C. L., Spigolon, A., & Veronez, M. R. (2022). An Immersive Virtual Field Experience Structuring Method for Geoscience Education. *IEEE Transactions on Learning Technologies*.
<https://doi.org/10.1109/TLT.2022.3207089>
- Huang, W., Roscoe, R. D., Johnson-Glenberg, M. C. & Craig, S. D. (2021). Motivation, engagement, and performance across multiple virtual reality sessions and levels of immersion. *Journal of Computer Assisted Learning*, 37(3), 745–758. <https://doi.org/10.1111/jcal.12520>
- Huang, Y.-J., Liu, K.-Y., Lee, S.-S. & Yeh, I.-C. (2021). Evaluation of a Hybrid of Hand Gesture and Controller Inputs in Virtual Reality. *International Journal of Human–Computer Interaction*, 37(2), 169–180. <https://doi.org/10.1080/10447318.2020.1809248>
- Huang, X., Zou, D., Cheng, G., & Xie, H. (2021). A Systematic Review of AR and VR Enhanced Language Learning. *Sustainability*, 13(9), Article 9. <https://doi.org/10.3390/su13094639>
- IBM Corp. (2020). IBM SPSS Statistics for Windows. Armonk, NY: IBM Corp. Verfügbar unter: <https://www.ibm.com/analytics/spss-statistics-software>
- International Organization for Standardization. (2018). Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts. *ISO Standard No. 9241-11:2018*.
- International Organization for Standardization. (2019). Ergonomics of human-system interaction — Part 210: Human-centered design for interactive systems. *ISO Standard No. 9241-210:2019*.
- Irvine, J. (2018). A framework for comparing theories related to motivation in education. *Research in Higher Education Journal*, 35.
- Jacobson, M., & Ruddy, M. (2004). Open to outcome: A practical guide for facilitating & teaching experiential reflection. Wood'N'Barnes Publishing.
- Jackson, S. A. & Marsh, H. W. (1996). Development and Validation of a Scale to Measure Optimal Experience: The Flow State Scale. *Journal of Sport and Exercise Psychology*, 18(1), 17–35.
<https://doi.org/10.1123/jsep.18.1.17>
- jamovi. (2021). The jamovi project. jamovi. Verfügbar unter: <https://www.jamovi.org>
- Jeno, L. M., Vandvik, V., Eliassen, S. & Grytnes, J.-A. (2019). Testing the novelty effect of an m-learning tool on internalization and achievement: A Self-Determination Theory approach. *Computers & Education*, 128, 398–413. <https://doi.org/10.1016/j.compedu.2018.10.008>
- Johnson, N. L. (1949). Systems of Frequency Curves Generated by Methods of Translation. *Biometrika*, (1/2), 149–176.
- Johnson-Glenberg, M. (2018). Immersive VR and education: Embodied design principles that include gesture and hand controls. *Frontiers in Robotics and AI*, 5(81), 1–19.
<https://doi.org/10.3389/frobt.2018.00081>

- Jonassen, D. H. (1994). Thinking technology: toward a constructivist design model. *Educational Technology*, 34(4), 34–37.
- Kamińska, D., Zwoliński, G. & Laska-Leśniewicz, A. (2022). Usability Testing of Virtual Reality Applications—The Pilot Study. *Sensors*, 22(4), 1342. <https://doi.org/10.3390/s22041342>
- Kangas, J., Kumar, S. K., Mehtonen, H., Järnstedt, J. & Raisamo, R. (2022). Trade-Off between Task Accuracy, Task Completion Time and Naturalness for Direct Object Manipulation in Virtual Reality. *Multimodal Technologies and Interaction*, 6(1), 6. <https://doi.org/10.3390/mti6010006>
- Keebler, J. R., Shelstad, W. J., Smith, D. C., Chaparro, B. S. & Phan, M. H. (2020). Validation of the GUESS-18: A Short Version of the Game User Experience Satisfaction Scale (GUESS), 16(1), 15.
- Kilteni, K., Groten, R., & Slater, M. (2012). The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments*, 21(4), 373–387. https://doi.org/10.1162/PRES_a_00124
- Kim, Y., Baylor, A. L. & Shen, E. (2007). Pedagogical agents as learning companions: the impact of agent emotion and gender: Learning companions. *Journal of Computer Assisted Learning*, 23(3), 220–234. <https://doi.org/10.1111/j.1365-2729.2006.00210.x>
- Kirakowski, J. & Corbett, M. (1993). SUMI: the Software Usability Measurement Inventory. *British Journal of Educational Technology*, 24(3), 210–212. <https://doi.org/10.1111/j.1467-8535.1993.tb00076.x>
- Klepsch, M., Schmitz, F. & Seufert, T. (2017). Development and Validation of Two Instruments Measuring Intrinsic, Extraneous, and Germane Cognitive Load. *Frontiers in Psychology*, 8, 1997. <https://doi.org/10.3389/fpsyg.2017.01997>
- Klepsch, M. & Seufert, T. (2020). Understanding instructional design effects by differentiated measurement of intrinsic, extraneous, and germane cognitive load. *Instructional Science*, 48(1), 45–77. <https://doi.org/10.1007/s11251-020-09502-9>
- Knogler, M., Harackiewicz, J. M., Gegenfurtner, A., & Lewalter, D. (2015). How situational is situational interest? Investigating the longitudinal structure of situational interest. *Contemporary Educational Psychology*, 43, 39–50. <https://doi.org/10.1016/j.cedpsych.2015.08.004>
- Kolb, D. A. (2014). *Experiential learning: Experience as the source of learning and development*. FT press.
- Kuckartz, U. (2014). *Mixed Methods*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-531-93267-5>
- Kuckartz, U. (2014). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung (3. Aufl.)*. Weinheim: Beltz Juventa

- Laine, J., Lindqvist, T., Korhonen, T. & Hakkarainen, K. (2022). Systematic Review of Intelligent Tutoring Systems for Hard Skills Training in Virtual Reality Environments. *International Journal of Technology in Education and Science*, 6(2), 178–203. <https://doi.org/10.46328/ijtes.348>
- Lampen, E., Liersch, M. & Lehwald, J. (2020). Towards Motor Learning in Augmented Reality: Imitating an Avatar (Communications in Computer and Information Science). In C. Stephanidis, M. Antona & S. Ntoa (Hrsg.), *HCI International 2020 – Late Breaking Post-ers* (Band 1294, S. 181–188). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-60703-6_23
- van der Land, S., Schouten, A. P., Feldberg, F., van den Hooff, B. & Huysman, M. (2013). Lost in space? Cognitive fit and cognitive load in 3D virtual environments. *Computers in Human Behavior*, 29(3), 1054–1064. <https://doi.org/10.1016/j.chb.2012.09.006>
- Lee, K. M. (2004). Presence, explicated. *Communication theory*, 14(1), 27-50.
- Legris, P., Ingham, J., & Colletette, P. (2003). Why do people use information technology? A critical review of the technology acceptance model. *Information & management*, 40(3), 191-204. [https://doi.org/10.1016/S0378-7206\(01\)00143-4](https://doi.org/10.1016/S0378-7206(01)00143-4)
- Leppink, J., Paas, F., Van der Vleuten, C. P. M., Van Gog, T. & Van Merriënboer, J. J. G. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*, 45(4), 1058–1072. <https://doi.org/10.3758/s13428-013-0334-1>
- Lewis, J. R. (2018). The System Usability Scale: Past, Present, and Future. *International Journal of Human-Computer Interaction*, 34(7), 577–590. <https://doi.org/10.1080/10447318.2018.1455307>
- Leyman, E. L., Mirka, G. A., Kaber, D. B. & Sommerich, C. M. (2004). Cervicobrachial muscle response to cognitive load in a dual-task scenario. *Ergonomics*, 47(6), 625–645. <https://doi.org/10.1080/00140130310001629766>
- Liu, R., Wang, L., Lei, J., Wang, Q., & Ren, Y. (2020). Effects of an immersive virtual reality-based classroom on students' learning performance in science lessons. *British Journal of Educational Technology*, 51(6), 2034-2049.
- Longo, M. R., Schüür, F., Kammers, M. P. M., Tsakiris, M., & Haggard, P. (2008). What is embodiment? A psychometric approach. *Cognition*, 107(3), 978–998. <https://doi.org/10.1016/j.cognition.2007.12.004>
- Makransky, G. & Petersen, G. B. (2021). The Cognitive Affective Model of Immersive Learning (CAMIL): a Theoretical Research-Based Model of Learning in Immersive Virtual Reality. *Educational Psychology Review*, 33(3), 937–958. <https://doi.org/10.1007/s10648-020-09586-2>
- Makransky, G., Terkildsen, T. S. & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, 60, 225–236. <https://doi.org/10.1016/j.learninstruc.2017.12.007>

- Makransky, G., Lilleholt, L., & Aaby, A. (2017). Development and validation of the Multimodal Presence Scale for virtual reality environments: A confirmatory factor analysis and item response theory approach. *Computers in Human Behavior*, 72, 276-285. <https://doi.org/10.1016/j.chb.2017.02.066>
- Manero, B., Álvarez, I. M., Romero, A., & Cárdenas, M. M. (2022). Adopting an immersive virtual reality system to enhance Pre-service teachers' communicative competence. In 2022 International Conference on Advanced Learning Technologies (ICALT) (pp. 317-321). IEEE. <https://doi.org/10.1109/ICALT55010.2022.00101>
- Marangunić, N., & Granić, A. (2015). Technology acceptance model: a literature review from 1986 to 2013. *Universal access in the information society*, 14, 81-95. <https://doi.org/10.1007/s10209-014-0348-1>
- Martha, A. S. D. & Santoso, H. (2019). The Design and Impact of the Pedagogical Agent: A Systematic Literature Review. *The Journal of Educators Online*, 16(1). <https://doi.org/10.9743/jeo.2019.16.1.8>
- Mayer, R. E. (2014). *The Cambridge Handbook of Multimedia Learning* (2nd ed., Cambridge Handbooks in Psychology). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369>
- Mayer, R.E. & Moreno, R. (2003). Nine Ways to Reduce Cognitive Load in Multimedia Learning. *Educational Psychologist*, 38:1, 43-52, DOI: 10.1207/S15326985EP3801_6
- Meyer, O. A., Omdahl, M. K. & Makransky, G. (2019). Investigating the effect of pre-training when learning through immersive virtual reality and video: A media and methods experiment. *Computers & Education*, 140, 103603. <https://doi.org/10.1016/j.compedu.2019.103603>
- Microsoft. (2022). Microsoft Excel für Microsoft 365 MSO. Microsoft.
- Mikropoulos, T. A. & Natsis, A. (2011). Educational virtual environments: A ten-year review of empirical research (1999–2009). *Computers & Education*, 56(3), 769–780. <https://doi.org/10.1016/j.compedu.2010.10.020>
- Moon, J. A. (2013). *A handbook of reflective and experiential learning: Theory and practice*. Routledge.
- Mutschler, B. & Reichert, M. (2004). Usability-Metriken als Nachweis der Wirtschaftlichkeit von Verbesserungen der Mensch-Maschine-Schnittstelle, 12.
- Müser, S. & Fehling, C. D. (2022). AR/VR.nrw – Augmented und Virtual Reality in der Hochschullehre. *HMD Praxis der Wirtschaftsinformatik*, 59(1), 122–141. <https://doi.org/10.1365/s40702-021-00815-y>
- Mütterlein, J. (2018). The three pillars of virtual reality? Investigating the roles of immersion, presence, and interactivity. In Hawaii International Conference on System Sciences. <https://doi.org/10.24251/HICSS.2018.174>

- Niklas, S. (2014). Akzeptanz und Nutzung mobiler Applikationen. Wiesbaden: Springer.
- Nimon, K. F. (2012). Statistical Assumptions of Substantive Analyses Across the General Linear Model: A Mini-Review. *Frontiers in Psychology*, 3. <https://doi.org/10.3389/fpsyg.2012.00322>
- Noetel, M., Griffith, S., Delaney, O., Harris, N. R., Sanders, T. & Parker, P. (2022). Multimedia Design for Learning: An Overview of Reviews With Meta-Analysis. *Review of Educational Research*, 92(3), 413–454.
- Nyysönen, T., Helle, S., Lehtonen, T. & Smed, J. (2022). A Comparison of Gesture and Controller-based User Interfaces for 3D Design Reviews in Virtual Reality. Gehalten auf der Hawaii International Conference on System Sciences. <https://doi.org/10.24251/HICSS.2022.220>
- Oculus. (n. d.). Oculus Developer Blog. *Oculus Developer Blog*. Zugriff am 19.8.2022. Verfügbar unter: <https://developer.oculus.com/blog/>
- Oesterreich, R., Leitner, K. & Resch, M. (2000). *Analyse psychischer Anforderungen und Belastungen in der Produktionsarbeit*. Göttingen: Hogrefe.
- Özgen, D. S., Afacan, Y. & Sürer, E. (2021). Usability of virtual reality for basic design education: a comparative study with paper-based design. *International Journal of Technology and Design Education*, 31(2), 357–377. <https://doi.org/10.1007/s10798-019-09554-0>
- Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, 84(4), 429–434. <https://doi.org/10.1037/0022-0663.84.4.429>
- Paas, F., van Gog, T. & Sweller, J. (2010). Cognitive Load Theory: New Conceptualizations, Specifications, and Integrated Research Perspectives. *Educational Psychology Review*, 22(2), 115–121. <https://doi.org/10.1007/s10648-010-9133-8>
- Parong, J., & Mayer, R. E. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology*, 110(6), 785–797. <https://doi.org/10.1037/edu0000241>
- Parong, J., & Mayer, R. E. (2020). Cognitive and affective processes for learning science in Immersive Virtual reality. *Journal of Computer Assisted Learning*, 37(1), 226–241. <https://doi.org/10.1111/jcal.12482>
- Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18(4), 315–341. <https://doi.org/10.1007/s10648-006-9029-9>
- Pellas, N., Dengel, A., & Christopoulos, A. (2020). A Scoping Review of Immersive Virtual Reality in STEM Education. *IEEE Transactions on Learning Technologies*, 13(4), 748–761. <https://doi.org/10.1109/TLT.2020.3019405>

- Pelletier, L. G., Rocchi, M. A., Vallerand, R. J., Deci, E. L. & Ryan, R. M. (2013). Validation of the revised sport motivation scale (SMS-II). *Psychology of Sport and Exercise*, 14(3), 329–341. <https://doi.org/10.1016/j.psychsport.2012.12.002>
- Petersen, G. B., Stenberdt, V., Mayer, R. E., & Makransky, G. (2023). Collaborative generative learning activities in immersive virtual reality increase learning. *Computers & Education*, 207, 104931. <https://doi.org/10.1016/j.compedu.2023.104931>
- Phan, M. H., Keebler, J. R. & Chaparro, B. S. (2016). The Development and Validation of the Game User Experience Satisfaction Scale (GUESS). *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 58(8), 1217–1247. <https://doi.org/10.1177/0018720816669646>
- Pintrich, P. R., Smith, D. A. F., Garcia, T. & Mckeachie, W. J. (1993). Reliability and Predictive Validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement*, 53(3), 801–813. <https://doi.org/10.1177/0013164493053003024>
- Pituch, K. A. & Stevens, J. (2016). *Applied multivariate statistics for the social sciences: analyses with SAS and IBM's SPSS* (6th edition.). New York: Routledge/Taylor & Francis Group.
- Plechata, A., Sahula, V., Fayette, D. & Fajnerová, I. (2019). Age-Related Differences With Immersive and Non-immersive Virtual Reality in Memory Assessment. *Frontiers in Psychology*, 10, 1330. <https://doi.org/10.3389/fpsyg.2019.01330>
- Pletz, C., & Zinn, B. (2018). Technologieakzeptanz von virtuellen Lern- und Arbeitsumgebungen in technischen Domänen. *Journal of Technical Education (JOTED)*, 6(4). <https://doi.org/10.48513/joted.v6i4.143>
- Przyborski, A., & Riegler, J. (2010). Gruppendiskussion und Fokusgruppe. In G. Mey & K. Mruck (Hrsg.), *Handbuch qualitative Forschung in der Psychologie* (1. Aufl.). VS Verlag für Sozialwissenschaften.
- R Core Team. (2021). R: A Language and Environment for Statistical Computing. Wien: R Foundation for Statistical Computing. Verfügbar unter: <https://www.R-project.org/>
- Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, 147. <https://doi.org/10.1016/j.compedu.2019.103778>
- Raithel, J., Dollinger, B. & Hörmann, G. (2007). Einführung Pädagogik. Begriffe – Strömungen – Klassiker – Fachrichtungen (2., durchgesehene und erweiterte Aufl.). Wiesbaden: VS Verlag. <https://doi.org/10.1186/s41239-018-0123-0>
- Razali, N. M. & Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests.

- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78.
<https://doi.org/10.1037/0003-066X.55.1.68>
- Razali, N. M. & Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics*, 2(1), 21–33.
- Rojas-Sánchez, M. A., Palos-Sánchez, P. R., & Folgado-Fernández, J. A. (2023). Systematic literature review and bibliometric analysis on virtual reality and education. *Education and Information Technologies*, 28(1), 155–192. <https://doi.org/10.1007/s10639-022-11167-5>
- RStudio. (2022). RStudio. RStudio, PCB.
- Ryan, R. M. & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, 25(1), 54–67.
<https://doi.org/10.1006/ceps.1999.1020>
- Sarodnick, F. & Brau, H. (2016). *Methoden der Usability Evaluation: wissenschaftliche Grundlagen und praktische Anwendung* (3., unveränderte Auflage.). Bern: Hogrefe.
- Sayers, H. (2004). Desktop virtual environments: a study of navigation and age. *Interacting with Computers*, 16(5), 939–956. <https://doi.org/10.1016/j.intcom.2004.05.003>
- SBB. (n. d.). Personnel. *Facts and Figures - Personnel*. Zugriff am 19.8.2022. Verfügbar unter: <https://reporting.sbb.ch/en/personnel>
- SBB Lab. (n. d.). SUCAP - Success factors for training in virtual space: from simulation to practice. *Research Fund SBB Lab*. Zugriff am 19.8.2022. Verfügbar unter: <https://imp-sbb-lab.unisg.ch/de/forschungsfonds-sbb-eng>
- Schmidt, F. L. (1971). The Relative Efficiency of Regression and Simple Unit Predictor Weights in Applied Differential Psychology. *Educational and Psychological Measurement*, 31(3), 699–714.
<https://doi.org/10.1177/001316447103100310>
- Schöbel, S., Janson, A. & Mishra, A. (2019). A Configurational View on Avatar Design – The Role of Emotional Attachment, Satisfaction, and Cognitive Load in Digital Learning. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3524079>
- Schunk, D. H., & DiBenedetto, M. K. (2016). Self-efficacy theory in education. In *Handbook of motivation at school* (pp. 34-54). Routledge.
- Schumacher, V. & Martin, M. (2013). Lernen und Gedächtnis im Alter. In T. Bartsch & P. Falkai (Hrsg.), *Gedächtnisstörungen* (S. 31–39). Berlin, Heidelberg: Springer Berlin Heidelberg.
https://doi.org/10.1007/978-3-642-36993-3_3

- Sewell, J. L., Young, J. Q., Boscardin, C. K., ten Cate, O. & O'Sullivan, P. S. (2019). Trainee perception of cognitive load during observed faculty staff teaching of procedural skills. *Medical Education*, 53(9), 925–940. <https://doi.org/10.1111/medu.13914>
- Shiban, Y. (2015). The appearance effect: Influences of virtual agent features on performance and motivation. *Computers in Human Behavior*, 7
- Sitzmann, T., & Ely, K. (2011). A meta-analysis of self-regulated learning in work-related training and educational attainment: What we know and where we need to go. *Psychological Bulletin*, 137(3), 421–442. <https://doi.org/10.1037/a0022777>.
- Slobounov, S. M., Ray, W., Johnson, B., Slobounov, E. & Newell, K. M. (2015). Modulation of cortical activity in 2D versus 3D virtual reality environments: An EEG study. *International Journal of Psychophysiology*, 95(3), 254–260. <https://doi.org/10.1016/j.ijpsy-cho.2014.11.003>
- Stolz, S. A. (2015). Embodied learning. *Educational Philosophy and Theory*, 47(5), 474–487. <https://doi.org/10.1080/00131857.2013.879694>
- Sutcliffe, A. G. & Kaur, K. D. (2000). Evaluating the usability of virtual reality user interfaces. *Behaviour & Information Technology*, 19(6), 415–426. <https://doi.org/10.1080/014492900750052679>
- Sutcliffe, A. & Gault, B. (2004). Heuristic evaluation of virtual reality applications. *Interacting with Computers*, 16(4), 831–849. <https://doi.org/10.1016/j.intcom.2004.05.001>
- Sweller, J., van Merriënboer, J. J. G. & Paas, F. (2019). Cognitive Architecture and Instructional Design: 20 Years Later. *Educational Psychology Review*, 31(2), 261–292. <https://doi.org/10.1007/s10648-019-09465-5>
- Sweller, J. (2005). Implications of cognitive load theory of multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 19–30). New York: Cambridge University Press.
- Sweller, J. (2010). Cognitive load theory: Recent theoretical advance. In J. L. Plass, R. Moreno & R. Brünken (Eds.), *Cognitive load theory* (pp. 29–47). New York: Cambridge University Press.
- Sweller, J., van Merriënboer, J. J. G. & Paas, F. G. W. C. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, 10(3), 251–296.
- Sweller, J., Ayres, P. & Kalyuga, S. (2011). Measuring Cognitive Load (Explorations in the Learning Sciences, Instructional Systems and Performance Technologies). In J. Sweller, P. Ayres & S. Kalyuga (Hrsg.), *Cognitive Load Theory* (S. 71–85). New York, NY: Springer. https://doi.org/10.1007/978-1-4419-8126-4_6
- Tabachnick, B. G. & Fidell, L. S. (2013). *Using multivariate statistics* (Always learning) (6. ed., international ed.). Boston Munich: Pearson.

- Tcha-Tokey, K., Loup-Escande, E., Christmann, O. & Richir, S. (2016). A questionnaire to measure the user experience in immersive virtual environments. *Proceedings of the 2016 Virtual Reality International Conference* (S. 1–5). Gehalten auf der VRIC '16: Virtual Reality International Conference - Laval Virtual 2016, Laval France: ACM. <https://doi.org/10.1145/2927929.2927955>
- Tenemaza, M., Ramírez, J., & de Antonio, A. (2016). Acceptability of an A2R application: Analysis of correlations between factors in a TAM. In 2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct) (pp. 178-183). IEEE. <https://doi.org/10.1109/ISMAR-Adjunct.2016.0071>
- Tivian XI GmbH. (2022). Enterprise Feedback Suite. Tivian XI GmbH. Verfügbar unter: <https://qbdocs.atlassian.net/wiki/spaces/DOC/pages/920748282/Enterprise+Feed-back+Suite>
- VERBI GmbH. (2022). MAXQDA Analytics Pro 2022. Berlin: VERBI GmbH. Verfügbar unter: <https://www.maxqda.de/produkte/maxqda-analytics-pro>
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425-478. <https://doi.org/10.2307/30036540>
- Verband öffentlicher Verkehr (n. d.). R RTE 20100 Sicherheit bei Arbeiten im Gleisbereich. Verfügbar unter: <https://www.voev.ch/de/Technik/Themen-Technik-Bahn-und-RTE/Themenbereiche-Technik/R-RTE-20100-Sicherheit-bei-Arbeiten-im-Gleisbereich/Ausbildung-R-RTE-20100>
- Verma, J. P. (2015). *Repeated Measures Design for Empirical Researchers*. Wiley.
- Verneau, M., van der Kamp, J., Savelsbergh, G. J. P. & de Looze, M. P. (2014). Age and Time Effects on Implicit and Explicit Learning. *Experimental Aging Research*, 40(4), 477–511. <https://doi.org/10.1080/0361073X.2014.926778>
- Villena-Taranilla, R., Tirado-Olivares, S., Cózar-Gutiérrez, R. & González-Calero, J. A. (2022). Effects of virtual reality on learning outcomes in K-6 education: A meta-analysis. *Educational Research Review*, 35, 100434. <https://doi.org/10.1016/j.edurev.2022.100434>
- Webster, R. (2016). Declarative knowledge acquisition in immersive virtual learning environments. *Interactive Learning Environments*, 24, 1319–1333.
- Wenk, N., Penalver-Andres, J., Buetler, K. A., Nef, T., Müri, R. M. & Marchal-Crespo, L. (2021). Effect of immersive visualization technologies on cognitive load, motivation, usability, and embodiment. *Virtual Reality*. <https://doi.org/10.1007/s10055-021-00565-8>
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R. et al. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, 4(43), 1686. <https://doi.org/10.21105/joss.01686>
- Wickham, H., Miller, E. & Smith, D. (2022). haven: Import and Export „SPSS“, „Stata“ and „SAS“ Files. Verfügbar unter: <https://CRAN.R-project.org/package=haven>

- Wienrich, C., Döllinger, N., Kock, S., Schindler, K. & Traupe, O. (2018). Assessing User Experience in Virtual Reality – A Comparison of Different Measurements (Lecture Notes in Computer Science). In A. Marcus & W. Wang (Hrsg.), *Design, User Experience, and Usability: Theory and Practice* (Band 10918, S. 573–589). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-91797-9_41
- Wilde, M., Bätz, K., Kovaleva, A. & Urhahne, D. (2009). Überprüfung einer Kurzsкала intrinsischer Motivation (KIM). *Zeitschrift für Didaktik der Naturwissenschaften*, 15, 31–45.
- Witte, K. (2018). *Grundlagen der Sportmotorik im Bachelorstudium (Band 1)*. Berlin, Heidelberg: Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-57868-1>
- Wu, B., Yu, X. & Gu, X. (2020). Effectiveness of immersive virtual reality using head-mounted displays on learning performance: A meta-analysis. *British Journal of Educational Technology*, 51(6), 1991–2005. <https://doi.org/10.1111/bjet.13023>
- Zhou, X., Teng, F., Du, X., Li, J., Jin, M. & Xue, C. (2022). H-GOMS: a model for evaluating a virtual-hand interaction system in virtual environments. *Virtual Reality*. <https://doi.org/10.1007/s10055-022-00674-y>
- Ziegler, C., Papageorgiou, A., Hirschi, M., Genovese, R. & Christ, O. (2020). Training in Immersive Virtual Reality: A Short Review of Presumptions and the Contextual Interference Effect (Advances in Intelligent Systems and Computing). In T. Ahram, R. Taiar, V. Gremeaux-Bader & K. Aminian (Hrsg.), *Human Interaction, Emerging Technologies and Future Applications II* (Band 1152, S. 328–333). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-44267-5_49
- Zimmerman, B. J. (2013). Theories of self-regulated learning and academic achievement: An overview and analysis. *Self-regulated learning and academic achievement*, 1-36.

8 Anhang

Anhang A

Studie 1 Bestimmung der Lernszene: Definitionen der Haupt- und Subkategorien aus der qualitativen Auswertung

Fluchtraum	
Der Fluchtraum bezeichnet den Raum, welcher von den Mitarbeitenden aufgesucht ist, wenn sich während der Gleisarbeiten sich ein Zug nähert. Bevor die Arbeitstätigkeiten im Gleis aufgenommen werden, wird der Fluchtraum vorgängig besprochen und definiert. Je nach Situation kann der Ort des Fluchtraums variieren, da für dessen Festlegung räumliche Kriterien besonders relevant sind.	
Sich verändernder Fluchtraum bei bewegten Arbeitsauftrag	Mitarbeitende müssen bei ihrer Arbeit im Gleis beachten, dass bereits nach 1 Meter gehen, sich der Fluchtraum verkleinern kann. Diese Veränderung muss berücksichtigt werden beim Aufsuchen eines geeigneten Fluchtraums.
Fluchtraum zwischen 2 Gleisen	Zwischen 2 Gleisen muss genügend Platz vorhanden sein. Um sich ausreichend von einem vorbeifahrenden Zug zu schützen, wird mindestens einen Raum von 4.20 Meter zwischen den Gleisen benötigt, damit diese Stelle als Fluchtraum gilt.
Bahnhof als Fluchtraum	Der Bahnhofplatz kann für Arbeiten vor Ort als Fluchtraum definiert und genutzt werden, wenn sich ein Zug der Arbeitsstelle nähert.
Fluchtraum neben Lärmschutzwand	Bei Gleisarbeiten, bei denen Lärmschutzwänden vorhanden sind, muss ein geeigneter Fluchtraum vordefiniert werden. Wenn zwischen dem Gleis und der Lärmschutzwand nicht ausreichend Platz ist, kann als Fluchtraum ein Platz neben der Lärmschutzwand definiert werden.
Zu enge Fluchträume bzw. Platzverhältnisse	Fluchträume, welche durch Hindernisse (z.B. Lärmschutzwand) zu eng sind bzw. nicht ausreichend Platz bieten, dürfen die Mitarbeitenden nicht verwenden.
Fluchtraum beim Überqueren einer Brücke	Beim Überqueren einer Brücke bestehen für die Mitarbeitenden nur eine beschränkte Anzahl von Fluchträumen. Diese Räume müssen vor einer Brückenüberquerung definiert werden.
Räumungszeit	Die Räumungszeit definiert, wie viel Zeit benötigt wird, um von einem bestimmten Standort in den Fluchtraum zu flüchten. Je nach vorhandenem Zeitfenster müssen die Mitarbeitenden ein passender Fluchtraum festlegen.
Tätigkeiten zur Aufrechterhaltung der Sicherheit	
Damit die SBB-Mitarbeitenden während den Arbeiten im Gleis sicher sind, müssen verschiedene Sicherheitsmassnahmen ergriffen werden. Diese Massnahmen werden entweder vor oder während der Arbeitstätigkeit umgesetzt, um so die Sicherheit am Gleis aufrechtzuerhalten.	
Beschaffung notwendiger Hintergrundinformationen	Bevor die Mitarbeitenden der SBB mit ihrer Arbeit beginnen, sammeln diese notwendige Hintergrundinformationen. Die Informationen helfen den Mitarbeitenden dabei, Risiken

	besser einzuschätzen oder weiter Sicherheitsmassnahmen zu treffen.
Vorgängige Sperrung der Gleise	Grundsätzlich können die Mitarbeitenden ein Gleis vor einer Gleisbegehung sperren, umso für mehr Sicherheit an der Arbeitsstelle zu sorgen.
Ausfüllen Checkliste Selbstschutz	Vor einer Gleisbegehung müssen die SBB-Mitarbeitenden eine Checkliste ausfüllen, worin wichtige Sicherheitsaspekte gemeinsam besprochen und definiert werden. Je nachdem, wie sich eine Situation verändert, muss eine neue Checkliste ausgefüllt werden.
Selbstständiges Einschätzen von Distanzen	Sicherheitsrelevante Distanzen können sich beim Verlassen des Arbeitsortes, durch Hindernisse, den weiteren Gleisverlauf oder durch einen Wetterwechsel verändern. Um die Sicherheit während der Gleisarbeiten aufrechtzuerhalten, müssen Distanzen regelmässig neu beurteilt und eingeschätzt werden.
Kontrolle der Zugannäherung	Wenn die Mitarbeiter im Gleis eine Arbeit ausführen, ist es notwendig, dass die Gleise durchgehend beobachtet werden. Deshalb müssen die Personen in 3-Sekunden-Intervallen nach links oder rechts blicken, um herannahende Züge rechtzeitig zu identifizieren.
Positionierung auf Gleis zur Kontrolle der Zugannäherung	Beim Ausführen einer Arbeitstätigkeit im Gleis ist es wichtig, dass die SBB-Mitarbeitenden sich richtig Positionieren, um Züge rechtzeitig wahrzunehmen. Relevant ist hierbei die Personenkonstellation, da die Person entweder nur für eine oder für zwei Fahrrichtungen sich absichern muss.
Rückkehr auf Gleis	Die SBB-Mitarbeitenden können dann wieder an den Arbeitsplatz zurückkehren, wenn der Zug die Strecke passiert hat und genug weit weg ist von Arbeitsort.
Berechnung von Überquerungszeiten	Falls Personen mehrere Gleise überqueren müssen, um an den eigentlichen Arbeitsort zu gelangen, ist die Berechnung der Überquerungsdauer notwendig. Je nach Überquerungstrecke bzw. -dauer steigt das Risiko, dass sich ein Zug nähert und die Mitarbeitenden sich schnellstmöglich in den Fluchtraum begeben müssen.
Fortlaufendes Einschätzen von Annäherungsdistanzen	Damit sich die SBB-Mitarbeitenden bspw. an mehreren Orten Arbeiten im Gleis erledigen müssen, ist das fortlaufende Einschätzen von Annäherungsdistanzen notwendig. Durch das Verlassen des alten Arbeitsplatzes können die Bedingungen vor Ort die vorherigen Einschätzungen verfälschen.
Fluchtraum aufsuchen	Die Mitarbeitenden verlassen beim Annähern eines Zuges sofort den Arbeitsort und begeben sich in den vordefinierten Fluchtraum. Sobald der Zug vorbeigefahren ist, kann der Fluchtraum wieder verlassen werden.
Bewegungsraum klären für Standortwechsel	Für einen sicheren Standortwechsel muss vorgängig der Raum von den Mitarbeitenden definiert werden, welcher sie nutzen können, um an den nächsten Standort zu gelangen. Bei zu engen oder unübersichtlichen Gleisabschnitten muss dieser Raum sorgsam definiert werden.

Gleis verlassen	Sobald sämtliche Arbeiten im Gleis von den SBB-Mitarbeitenden erledigt worden sind, sollen diese die Gleise verlassen, umso die Wahrscheinlichkeit eines Unfalles zu reduzieren.
Fluchtraum identifizieren	Egal ob allein oder zu zweit, die Mitarbeitenden müssen bei jeder Gleisbegehung vorgängig einen geeigneten Fluchtraum definieren, welcher von allen genutzt wird. Falls die Mitarbeitenden sich von einer Arbeitsstelle zu einer anderen bewegen, muss der nächste Fluchtraum festgelegt werden.
Nachträglich Gleissperrung veranlassen	Um die Sicherheit der Mitarbeitenden während den Gleisarbeiten zu gewährleisten, kann das betroffene Gleis nachträglich gesperrt werden, damit dieser Streckenabschnitt temporär nicht von Zügen befahren wird.
Arbeitskoordination	Falls mehrere Personen an einem Arbeitsauftrag teilnehmen, können die verschiedenen Aufgaben koordiniert werden. Das heisst, während einige Personen sich um den eigentlichen Auftrag kümmern, sorgt die andere Person für den Schutz und Sicherheit der anderen.
Kommunikation	Kommunikation untereinander
Sichtweite	
Zur Einschätzung von Distanzen spielt die Sichtweite eine wichtige Rolle. Durch eine eingeschränkte Sicht müssen die sicherheitsrelevanten Distanzen neu beurteilt werden. Für die Beurteilung dieser Sicherheitsdistanzen sowie als Orientierungshilfe der Sichtweite dienen Hinweistafeln, welche entlang der Gleise aufgestellt sind.	
Kilometrierungs-Tafeln	Die Kilometrierungs-Tafeln geben den Mitarbeitenden einerseits eine Orientierung, auf welchem Abschnitt sie sich gegenwärtig befinden. Andererseits können aufgrund dieser Orientierung die weiteren Kilometrierungs-Tafeln für die Berechnung von sicherheitsrelevanten Distanzen verwendet werden.
Freie Sicht	Am Arbeitsort finden die Mitarbeitenden weder schlechte Umweltbedingungen oder Hindernisse vor, welche die Sichtweite beeinträchtigen könnten.
Annäherungsdistanz	Die Annäherungsdistanz bezeichnet die Distanz, welche ersichtlich sein muss, wenn ein Zug mit einem bestimmten Fortbewegungstempo sich dem Arbeitsplatz nähert. Ebenso muss dabei die Räumungszeit und Sicherheitsfrist berücksichtigt werden, um die Annäherungsdistanz abschliessend zu bestimmen.
Ort der Arbeitsstelle	
Der Ort für Gleisarbeiten sich je nach Auftrag verschieden. Je nach Arbeitsort gilt es dir vor Ort anzutreffenden Bedingungen zu berücksichtigen, welche einen Einfluss auf die Sicherheit haben. Das Vorhandensein von Lärmschutzwänden oder mehreren Gleisen haben einen direkten Einfluss auf das Sicherheitsverhalten der SBB-Mitarbeitenden.	
Arbeitsstelle hinter Brücke	Eine Arbeitsstelle, welche sich hinter einer Eisenbahnbrücke befindet, muss unter Berücksichtigung hoher Sicherheitsmassnahmen erreicht werden. Je nach Situation hat es auf der Brücke keine Fluchräume. Deshalb muss vorgängig definiert werden, wie am besten solche Arbeitsstellen erreicht werden.

Zwischen Gleis und Hindernissen	Der geringe Abstand zwischen einem Gleis und einem Hindernis wird als kritische Zone beurteilt, da sich dort für die Mitarbeitenden kein geeigneter Fluchraum befindet. Wenn der Arbeitsort eine solche kritische Zone ist, muss vorher ein geeigneter Fluchraum bestimmt werden.
Zwischen den Gleisen	An vielen Streckenabschnitten der SBB befinden sich mehrere Gleise nebeneinander. Arbeitsorte können sich dabei zwischen diesen Gleisen befinden und haben so direkten Einfluss auf die Sicherheitsmassnahmen und -verhaltensweisen der Mitarbeitenden.
Gleis mit Lärmschutzwand	Arbeitstätigkeiten fallen auch auf Streckenabschnitte, an denen vor Ort Lärmschutzwände vorhanden sind. An solchen Orten müssen die räumlichen Bedingungen zwischen Gleis und Schutzwand berücksichtigt werden, um die Sicherheit während der Arbeit aufrechtzuerhalten.
Beschaffenheit des Bodens	Die Beschaffenheit des Bodens kann dafür sorgen, dass die Mitarbeitenden keinen sicheren Stand haben. Zum Beispiel können die Mitarbeitenden bei Schotter stolpern und sich nicht in Sicherheit begeben, wenn sich ein Zug nähert.
In einem Gleis	Die SBB-Mitarbeitenden arbeiten innerhalb eines Gleises. Obwohl sich bei einem herannahenden Zug nur die Schienen zusätzlich überschritten werden müssen, gilt es, diese Extra-Distanz zu berücksichtigen.
Arbeiten am Gleis	
Für die SBB-Mitarbeitenden gibt es verschiedene Tätigkeiten, welche sie am Gleis durchführen. Darunter zählen Kontroll- und Vermessungsaufgaben. Je nach Streckenabschnitt müssen nebeneinanderliegende Gleise überschritten werden, wobei Sicherheitsmassnahmen berücksichtigt werden müssen.	
Kontrollen an einzelnen / mehreren Stellen am Gleis (Standortwechsel)	Damit der Gleisverkehr reibungslos funktioniert, müssen die SBB-Mitarbeitenden unterschiedliche Kontrollen vornehmen. Dafür begeben sich die Mitarbeitenden direkt ans Gleis, um die entsprechenden Gleisabschnitte und dessen Infrastruktur zu kontrollieren.
Weg von Ausgangsposition zu Arbeitsstelle	Bevor die Arbeit aufgenommen wird an der entsprechenden Arbeitsstelle, müssen die SBB-Mitarbeitenden einen sicheren Weg dorthin finden, um mögliche Risiken und Gefahren zu minimieren.
Über längere Distanz den Gleisen entlang gehen	Je nach Situation müssen die Mitarbeitenden Kontrollen an mehreren Orten durchführen und müssen deshalb über eine längere Distanz am Gleis entlang gehen. Neben dem eigentlichen Arbeitsauftrag müssen die Personen auch während der Fortbewegungszeit für ihre Sicherheit sorgen.
Vermessung durchführen am Gleis	Zur Gewährleistung eines funktionierenden Gleisverkehrs müssen auch diverse Vermessungen vorgenommen werden. Neben den Kontrollaufgaben helfen Vermessungen dabei, die Sicherheit und Qualität von Gleisabschnitten sicherzustellen.
Stromstellungen kontrollieren mit Befestigungsschlüssel	Eine weitere Arbeit von SBB-Mitarbeitenden ist es, Stromstellungen an den Gleisen zu kontrollieren. Anhand von

	Befestigungsschlüssel werden diese geprüft, ob alles in Ordnung ist.
Überqueren von mehreren Gleisen	Bei gewissen Aufgaben müssen sich die SBB-Mitarbeitenden an mehreren Orten Kontrollen oder Vermessungen vornehmen. Um an diesen Ort zu gelangen, müssen die Mitarbeitenden je nachdem mehrere Gleise überqueren. Beim Überqueren gilt es, Sicherheitsmassnahmen zu berücksichtigen, um sicher an den nächsten Arbeitsort zu gelangen.
Anziehen einer Befestigung	Damit Befestigungen am Gleis sich nicht lösen, müssen die SBB-Mitarbeitenden diese regelmässig anziehen. Bei dieser Tätigkeit müssen die Gleise immer im Auge behalten werden, um sich rechtzeitig in Sicherheit zu begeben.
Anziehen einer Schraube	Zur richtigen Befestigung der Schienen müssen die Mitarbeitenden die Schrauben prüfen, ob sie korrekt angezogen sind. Beim Anziehen ist es notwendig, die Sicherheitsmassnahmen einzuhalten und herannahende Züge rechtzeitig zu identifizieren.
Wetter	
Bei Arbeiten am Gleis haben die vor Ort vorhandenen Umweltbedingungen einen Einfluss auf die Aufrechterhaltung der Sicherheit. Das Wetter oder ein Wetterwechsel kann dazu führen, dass die Sichtweite von SBB-Mitarbeitenden plötzlich eingeschränkt oder sogar die auditive Wahrnehmung verändert wird.	
Nebel	Das Auftreten von Nebel an einem bestimmten Arbeitsort führt zu einer eingeschränkten Sichtweite. Diese Einschränkung ist für die SBB-Mitarbeitenden insofern relevant, dass sicherheitsrelevante Distanzen neu beurteilt werden müssen.
Tageslicht	Je nach Jahreszeit besitzen die Tage mehr oder weniger lang Tageslicht. Zudem können zum Beispiel Schnee oder Nebel das Tageslicht dämpfen. Beides beeinflusst die visuelle Wahrnehmung der Mitarbeitenden, was dessen Sicherheit bei Arbeiten im Gleis beeinträchtigt.
Starker Regen	Durch starken Regen kann es ebenso zu einer eingeschränkten Sichtweite führen, was sich negativ auf die Sicherheit der Mitarbeitenden auswirkt. Ebenso kann starker Regen neue Lärmquellen schaffen und so die auditive Wahrnehmung eines herannahenden Zuges beeinflussen.
Starker Schneefall	Ebenso kann starker Schneefall neben der visuellen auch die auditive Wahrnehmung beeinträchtigen, da Züge auf schneebedeckten Gleisen eine veränderte Lärmkulisse erzeugen und wenig gut sichtbar sind.
Zug	
Bei Arbeiten am Gleis stellen Züge für die SBB-Mitarbeitenden eine grosse Gefahr dar. Je nach Gleisabschnitt besitzen die Züge eine andere Reisegeschwindigkeit. Ausserdem muss auch die Fahrtrichtung von potenziell herannahenden Zügen beurteilt werden. Beide Faktoren haben einen Einfluss auf die Sicherheitsmassnahmen, welche während den Arbeitstätigkeiten am Gleis zu berücksichtigen sind.	
Zuggeschwindigkeit	Die Züge unterscheiden sich in ihrer Geschwindigkeit je nachdem, wo sie sich gerade befinden. Deshalb ist für die SBB-Mitarbeitenden bzw. für dessen Sicherheit notwendig,

	in Erfahrung zu bringen, welche Geschwindigkeiten die Züge an einem bestimmten Einsatzort aufweisen.
Fahrtrichtung	Je nach Gleisabschnitt können die Züge aus verschiedenen Richtungen auftauchen. Um die Sicherheit der SBB-Mitarbeitenden zu gewährleisten, muss deshalb die korrekte Fahrtrichtung von herannahenden Zügen bestimmte werden.
Gleichzeitiges Vorbeifahren zweier Züge	An vielen Streckenabschnitten befinden sich zwei Gleise nebeneinander. Umso grösser ist deshalb das Sicherheitsrisiko für SBB-Mitarbeitende, wenn sich an ihrem Arbeitsort zwei Züge kreuzen.
Kein Alarmsignal im Selbstschutz	Je nach Situation fahren Züge an Arbeitsorten vorbei, wo keine Alarmsignale vorhanden sind, und für die dortigen Mitarbeitenden ausgelöst werden können. Deshalb müssen den Zügen darauf achten, wie die Mitarbeitenden beim Annähern bzw. Vorbeifahren reagieren und im Notfall selbst Alarmsignal geben.
Zug verlässt Annäherungsdistanz	Ein Zug, welche die Annäherungsdistanz bereits verlassen hat, stellt für die Mitarbeitenden kein weiteres Risiko dar. Dennoch müssen die Personen in ihren Fluchräumen genug lange warten, da diese Wartezeit je nach Tempo variieren kann.
Hindernisse	
Hindernisse haben einen Einfluss auf die Sichtweite und somit auf die Sicherheit der SBB-Mitarbeitenden während ihrer Tätigkeiten im Gleis. Je nach Arbeitsort können diese Hindernisse variieren und müssen für die Aufrechterhaltung der Sicherheit entsprechend berücksichtigt werden.	
Bagger	Grosse Gerätschaften wie Bagger können bei Gleisarbeiten der SBB oder bei anderen Baustellen zum Einsatz kommen. Durch die Grösse kann je nach Arbeitsort die Sichtweite beeinträchtigt werden, was die Sicherheit von SBB-Mitarbeitenden gefährdet.
Wald	Da die Zugstrecken der SBB durch die Landschaft der Schweiz verlaufen, sind an gewissen Streckenabschnitten neben Wäldern vorzufinden. Die unzähligen Bäume können dabei die Sichtweite beeinträchtigen und stellen dadurch ein Risiko dar für die Gewährleistung der Sicherheit.
Bäume mit / ohne Blätter	Neben einem ganzen Wald stellen auch einzelne Bäume (mit / ohne Blätter) Hindernisse dar, welche die Sichtweite von den SBB-Mitarbeitenden beeinträchtigen und somit die Sicherheit.
Versperrung innerhalb des Fluchtraums	Hindernisse wirken sich nicht nur auf die Sichtweite aus, sondern können auch benötigte Fluchräume versperren. Dadurch kann es sein, dass die SBB-Mitarbeitenden sich zu einem Fluchraum begeben müssen, welcher weiter weg ist von der eigentlichen Arbeitsstelle. Das wiederum kann sich negativ auf die Sicherheit der Mitarbeitenden auswirken.
Lärmschutzwand	Bei Gleisbegehungen an einem bestimmten Gleisabschnitt stellen die vorzufindenden Lärmschutzwände ein Hindernis dar. Da sich Lärmschutzwände teilweise über längere

	Distanzen erstrecken, kann das die Sichtweite und Sicherheit beeinträchtigen.
Brücke	Brücke können je nach Grösse und Aussehen herannahende Züge verdecken. Je nach Streckenabschnitt verlieren die Mitarbeitenden durch die Brücke wertvolle Meter an Sichtweite. Ebenso kann eine Brücke ein Hindernis dafür sein, um schnell und sicher an einen anderen Ort zu gelangen. Für eine Brückenüberquerung müssen auch geeignete Fluchräume definiert werden.
Öffentliche Wege	Öffentliche Wege sind eine Möglichkeit, wie potenzielle Hindernisse wie beispielsweise Brücken umgangen werden können. Anstatt sich auf einer Brücke von Punkt A zu Punkt B zu bewegen, können die SBB-Mitarbeitende solche Wege nutzen, um sich sicher fortzubewegen.
Baustelle	Ein weiteres Hindernis stellen Baustellen an sich dar. Je nach Grösse der Baustelle und die Fahrzeuge, welche sich vor Ort befinden, kann die Sichtweite von den SBB-Mitarbeitenden beeinträchtigt werden und somit auch das Sicherheitsrisiko.
Ablenkungen	
Während der Gleisarbeiten kann es zu Ablenkungen der SBB-Mitarbeitenden kommen. Die Ablenkungen führen dazu, dass die Aufmerksamkeit der Mitarbeitenden für eine gewisse Zeitspanne auf andere Faktoren gerichtet wird und so Sicherheitsrisiken für die Mitarbeitenden entstehen.	
Telefon klingelt	Die SBB-Mitarbeitenden müssen je nach Arbeitstätigkeit miteinander über weite Distanzen kommunizieren. Trotz dieser Notwendigkeit kann ein klingelndes Telefon die Mitarbeitenden beeinträchtigen bei der Aufrechterhaltung von Sicherheitsmassnahmen, indem der Fokus auf das Telefon gerichtet wird anstatt auf potenziell herannahende Züge.
Starker Regenfall	Durch den starken Regenfall kann die auditive Wahrnehmung gestört bzw. abgelenkt werden, da viel Regentropfen, die auf den Helm prallen, Lärm erzeugen. Anstatt dass sich die Mitarbeitenden auf relevante Geräusche fokussieren können, werden sie vom Regen abgelenkt.
Jemand / Etwas fällt zu Boden	Mitarbeitende oder Gegenstände, welche zu Boden fallen, können die Aufmerksamkeit von SBB-Mitarbeitenden auf sich lenken. Für diesen Zeitraum werden die Personen vom eigentlichen Geschehen abgelenkt und können herannahende Züge nicht rechtzeitig wahrnehmen.
Baustellenlärm	Baustellen können verschiedene Lärmquellen erzeugen und bei der Lautstärke variieren. Je nach Art des Lärms kann es zur Ablenkung von SBB-Mitarbeitenden kommen, wobei die Aufmerksamkeit für einen gewissen Zeitraum nicht auf potenziell herannahende Züge gerichtet wird.
Fotografieren	Wenn eine Teilaufgabe darin besteht, dass bestimmte Teile eines Streckenabschnittes vor, während oder nach anderen Arbeitstätigkeiten fotografiert werden müssen, kann diese Ablenkung die Sicherheit von Personen gefährden. Anstatt

	auf herannahende Züge, konzentrieren sich die SBB-Mitarbeitenden darauf, ein angemessenes Foto zu machen.
Zustand des Gleisabschnitts	
Bei Gleisarbeiten bestimmen die Mitarbeitenden den Zustand des Gleises. Neben der Bestimmung, auf welchem Gleis tatsächlich der Arbeitsauftrag ausgeführt wird, erhalten die SBB-Mitarbeitenden die Information hinsichtlich der Gleissperrung.	
Rangiergleis	Rangiergleise stellen jene Gleise dar, welche genutzt werden, um Wagons oder Züge temporär zu platzieren sowie um Zugkompositionen zu bilden. Bei einem Streckenabschnitt mit Rangiergleisen, befinden sich mehrere Gleise nacheinander, welche ineinandergreifen. Beim Überqueren solcher Gleise gilt es die Berücksichtigung der Sicherheitsmassnahmen.
Schotterkronen	Schotterkronen bilden Streckenabschnitte, wobei die Gleise auf künstlich angelegten Absätzen gelegt werden, die aus Schotter bestehen. Die Höhe solcher Schotterkronen kann variieren. Die Beschaffenheit und Höhe von solchen Höhen können für die Mitarbeitenden ein Risiko sein, wenn diese rasch die Gleise verlassen müssen.
gesperrt/ungesperrt	Bei Gleisarbeiten erhalten die SBB-Mitarbeitenden Informationen dazu, ob das Gleis, auf dem gearbeitet wird, gesperrt oder ungesperrt ist. Je nach Zustand hat das einen Einfluss auf die Sicherheitsmassnahmen und -verhaltensweisen der Mitarbeitenden.
Doppelspur	Bei einer Doppelspur befinden sich nebeneinander zwei Gleise.
Betriebsgleis	Ein Betriebsgleis ist ein Gleis, welches der Betriebserfüllung bzw. dem Transport von Menschen und Waren dient. Somit kann ein Betriebsgleis gleichzeitig auch ein Arbeitsgleis sein.
Kurve	Gleise, welche eine Biegung nach links oder rechts in ihrem Verlauf aufweisen, werden Kurven genannt. Je nach Grad dieser Biegung hat es einen Einfluss auf die Sichtweite der SBB-Mitarbeitenden und somit auch auf die Sicherheit.
Arbeitsgleis	Das Arbeitsgleis definiert das Gleis, auf dem die Mitarbeitenden ihre Arbeitsaufgaben erfüllen müssen. Durch diese Festlegung können Sicherheitsmassnahmen an die vorzufindenden Bedingungen adaptiert werden, um so die Sicherheit aufrechtzuerhalten.
Nachbargleis	Das Nachbargleis ist das Gleis, welches sich neben dem Arbeitsgleis befindet. Wenn ein Nachbargleis bei einer Gleisbegehung gesperrt ist, kann es beispielsweise als Fluchraum dienen, wenn genügend Distanz dazwischen ist.
Unterführung	Bei einer Unterführung absolvieren die Züge einen Teil der Zugfahrtstrecke unterirdisch. Züge, welche plötzlich aus einer Unterführung auftauchen, können SBB-Mitarbeitende überraschen und so dessen Sicherheit beeinträchtigen.
Abstand zwischen 2 Gleisen	Je nach Abstand zwischen zwei Gleisen kann, dass andere Gleis als Fluchraum genutzt werden. Ist jedoch nur Platz

	vorhanden, müssen die SBB-Mitarbeitenden einen anderen Fluchraum definieren.
Personenkonstellation	
Die Arbeitstätigkeiten der SBB-Mitarbeitenden unterscheiden sich auch dadurch, wie viele Personen teilnehmen. Die Personenkonstellation hat auch einen Einfluss auf die Umsetzung von Sicherheitsmassnahmen.	
Alleine	Personen, welche eine Gleisarbeit allein durchführen, müssen selbstständig und fortlaufend für ihre Sicherheit sorgen, indem sie die entsprechenden Sicherheitsmassnahmen befolgen.
Zu zweit	SBB-Mitarbeitende, welche zu zweit eine Gleisarbeit durchführen, müssen vorgängig absprechen, wie sie vorgehen wollen, um die Aufrechterhaltung der Sicherheit zu gewährleisten.
Aufteilung der Aufgaben	Die Arbeitsaufteilung muss klar geregelt sein, wenn mehrere Personen einer Arbeitstätigkeit nachgehen. Vorgängige Besprechungen müssen deshalb stattfinden, damit jeder SBB-Mitarbeitende weiss, welche Rolle jeder einnimmt. Die Rollen können sich dabei darin unterscheiden, dass jemand für die Auftragserledigung zuständig ist, während die andere Person für die Sicherheit sorgt.

Anhang B
Fragebogenitems der Studie II – Bedienungsparadigma

Kategorie	Item Original	Item DE & angepasst	Quelle
Usability	I think that I would like to use this system frequently.	Ich denke, dass ich die Simulation häufig nutzen möchte.	SUS A quick and dirty usability scale
	I found the system unnecessarily complex.	Ich fand die Simulation unnötig kompliziert.	
	I thought the system was easy to use.	Ich fand die Simulation einfach zu bedienen.	
	I think that I would need the support of a technical person to be able to use this system.	Ich glaube, ich bräuchte die Unterstützung einer technischen Person, um die Simulation nutzen zu können.	
	I found the various functions in this system were well integrated.	Ich fand, dass die verschiedenen Funktionen in der Simulation gut integriert waren.	
	I thought there was too much inconsistency in this system.	Ich fand, dass die Simulation zu inkonsequent war.	
	I would imagine that most people would learn to use this system very quickly.	Ich könnte mir vorstellen, dass die meisten Menschen sehr schnell lernen würden, mit der Simulation umzugehen.	
	I found the system very cumbersome to use.	Ich fand die Simulation sehr umständlich zu bedienen.	
	I felt very confident using the system.	Ich fühlte mich sehr sicher im Umgang mit der Simulation.	
	I needed to learn a lot of things before I could get going with this system.	Ich musste eine Menge Dinge lernen, bevor ich mit der Simulation loslegen konnte.	

Kategorie	Item Original	Item DE & angepasst	Quelle
Motivation / Interesse Vergnügen	Die Tätigkeit in der Ausstellung hat mir Spaß gemacht	Die Tätigkeit in der Simulation hat mir Spass gemacht	KIM Kurzsкала intrinsischer Motivation
	Ich fand die Tätigkeit in der Ausstellung sehr interessant.	Ich fand die Tätigkeit in der Simulation sehr interessant.	
	Die Tätigkeit in der Ausstellung war unterhaltsam.	Die Tätigkeit in der Simulation war unterhaltsam.	
Motivation / Wahrgenommene Kompetenz	Mit meiner Leistung in der Ausstellung bin ich zufrieden.	Mit meiner Leistung in der Simulation bin ich zufrieden.	
	Bei der Tätigkeit in der Ausstellung stellte ich mich geschickt an.	Bei der Tätigkeit in der Simulation stellte ich mich geschickt an.	
	Ich glaube, ich war bei der Tätigkeit in der Ausstellung ziemlich gut.	Ich glaube, ich war bei der Tätigkeit in der Simulation ziemlich gut.	
Motivation / Wahrgenommene Wahlfreiheit	Ich konnte die Tätigkeit in der Ausstellung selbst steuern	Ich konnte die Tätigkeit in der Simulation selbst steuern	
	Bei der Tätigkeit in der Ausstellung konnte ich wählen, wie ich es mache	Bei der Tätigkeit in der Simulation konnte ich wählen, wie ich es mache	
	Bei der Tätigkeit in der Ausstellung konnte ich so vorgehen, wie ich es wollte	Bei der Tätigkeit in der Simulation konnte ich so vorgehen, wie ich es wollte	
Motivation / Druck Anspannung	Bei der Tätigkeit in der Ausstellung fühlte ich mich unter Druck.	Bei der Tätigkeit in der Simulation fühlte ich mich unter Druck.	
	Bei der Tätigkeit in der Ausstellung fühlte ich mich angespannt.	Bei der Tätigkeit in der Simulation fühlte ich mich angespannt.	
	Ich hatte Bedenken, ob ich die Tätigkeit in der Ausstellung gut hinbekomme.	Ich hatte Bedenken, ob ich die Tätigkeit in der Simulation gut hinbekomme.	

Kategorie	Item Original	Item DE & angepasst	Quelle
Cognitive Load / EL env	The elements in the virtual environment made the learning very unclear.	Die Elemente in der Simulation machten das Lernen sehr unklar	MCLSVE The validation and further development of a multidimensional cognitive load scale for virtual environments
	The virtual environment was, in terms of learning, very ineffective.	Die Simulation war, was das Lernen angeht, sehr ineffektiv.	
	The virtual environment was full of irrelevant content.	Die Simulation war voll von unwichtigen Inhalten.	
	It was difficult to find the relevant learning information in the virtual environment.	Es war schwierig, die relevanten Lerninformationen in der Simulation zu finden.	
Cognitive Load / GL	The simulation really enhanced my understanding of the topics covered		
	The simulation really enhanced my knowledge and understanding of lab safety		
	The simulation really enhanced my understanding of the procedures covered		
	The simulation really enhanced my understanding of concepts and definitions		

Kategorie	Item Original	Item DE & angepasst	Quelle
Cognitive Load / IL	The topic covered in the simulation was very complex.		MCLSVE The validation and further development of a multidimensional cognitive load scale for virtual environments
	The simulation covered procedures that I perceived as very complex.		
	The simulation covered concepts and definitions that I perceived as very complex.		
Cognitive Load / EL ins	The instructions and/or explanations used in the simulation were very unclear.	Die in der Simulation verwendeten Anweisungen und/oder Erklärungen waren sehr unklar.	
	The instructions and/or explanations used in the simulation were, in terms of learning, very ineffective.	Die in der Simulation verwendeten Anweisungen und/oder Erklärungen waren im Hinblick auf den Lernerfolg sehr ineffektiv.	
	The instructions and/or explanations used in the simulation were full of unclear content.	Die in der Simulation verwendeten Anweisungen und/oder Erläuterungen waren inhaltlich unklar.	
Cognitive Load / EL Int	The interaction technique used in the simulation was very unclear.	Die in der Simulation verwendete Interaktionstechnik war sehr unklar.	
	The interaction technique used in the simulation was, in terms of learning, very ineffective.	Die in der Simulation verwendete Interaktionstechnik war im Hinblick auf das Lernen sehr ineffektiv.	
	The interaction technique used in the simulation made it harder to learn.	die in der Simulation verwendete Interaktionstechnik erschwerte das Lernen.	
	The interaction technique used in the simulation was difficult to master.	Die in der Simulation verwendete Interaktionstechnik war schwer zu meistern.	

Kategorie	Item Original	Item DE	Quelle
Usability /Playability	I find the controls of the game to be straightforward	Ich fand die Steuerung der Simulation sehr einfach	GUESS-18 Validation of the GUESS-18: A Short Version of the Game User Experience Satisfaction Scale (GUESS) Experience Satisfaction Scale
	I find the game's interface to be easy to navigate.	Ich finde, dass die Benutzeroberfläche der Simulation einfach zu navigieren war.	
Narratives	I am captivated by the game's story from the beginning.		
	I enjoy the fantasy or story provided by the game.		
Creative Freedom	I feel the game allows me to be imaginative		
	I feel creative while playing the game.		
Audio Aesthetics	I enjoy the sound effects in the game.		
	I feel the game's audio (e.g., sound effects, music) enhances my gaming experience.		
Social Connectivity	I find the game supports social interaction (e.g., chat) between players		
	I like to play this game with other players.		
Enjoyment	I think the game is fun.	Ich denke, die Simulation machte Spass.	
	I feel bored while playing the game. (REVERSE CODE)	Ich fühlte mich in der Simulation gelangweilt.	
Personal Gratification	I am very focused on my own performance while playing the game	Ich war sehr auf meine eigene Leistung während der Simulation konzentriert	
	I want to do as well as possible during the game.	Ich wollte während der Simulation so gut wie möglich abschneiden.	

Kategorie	Item Original	Item DE	Quelle
Visual Aesthetics	I enjoy the game's graphics		GUESS-18 Validation of the GUESS-18: A Short Version of the Game User Experience Satisfaction Scale (GUESS) Experience Satisfaction Scale
	I think the game is visually appealing.		
Play Engrossment	I feel detached from the outside world while playing the game.	Ich fühlte mich von der Aussenwelt abgeschnitten, während ich in der Simulation war.	
	I do not care to check events that are happening in the real world during the game.	Es war mir egal, welche Ereignisse in der realen Welt während der Simulation stattfanden.	

Kategorie	Item Original	Item DE	Quelle
Facilitating Learning	The agent led me to think more deeply about the presentation.	Der virtuelle Coach hat mich dazu gebracht, tiefer über das Tutorial nachzudenken.	API The API (Agent Persona Instrument) for Assessing Pedagogical Agent Persona
	The agent made the instruction interesting.	Der virtuelle Coach machte das Tutorial interessant.	
	The agent encouraged me to reflect what I was learning.	Der virtuelle Coach ermutigte mich, das im Tutorial Gelernte zu reflektieren.	
	The agent kept my attention.	Der virtuelle Coach hat meine Aufmerksamkeit aufrechterhalten.	
	The agent presented the material effectively.	Der virtuelle Coach hat das zu Lernende effektiv präsentiert.	
	The agent helped me to concentrate on the presentation.	Der virtuelle Coach half mir, mich auf das Tutorial zu konzentrieren.	
	The agent focused me on the relevant information.	Der virtuelle Coach hat mich auf die relevanten Informationen konzentriert.	
	The agent improved my knowledge of the content.	Der virtuelle Coach hat mein Wissen über das Tutorial verbessert.	
	The agent was interesting.	Der virtuelle Coach war interessant.	
	The agent was enjoyable.	Der virtuelle Coach war unterhaltsam.	
Credible	The agent was knowledgeable.	Der virtuelle Coach war kompetent.	
	The agent was intelligent.	Der virtuelle Coach war intelligent.	
	The agent was useful.	Der virtuelle Coach war nützlich.	
	The agent was helpful.	Der virtuelle Coach war hilfreich.	
	The agent was instructor-like.	Der virtuelle Coach war wie eine Lehrerin.	
Human-like	The agent has a personality.		
	The agent's emotion was natural.		
	The agent was human-like.		
	The agent's movement was natural.		
	The agent showed emotion.		
Engaging	The agent was expressive.		
Engaging			
	The agent was enthusiastic.		API The API (Agent Persona Instrument) for Assessing Pedagogical Agent Persona
	The agent was entertaining.		
	The agent was motivating.		
	The agent was friendly.		

Anhang C

Online Fragebogen Studie III – iVR-Multiplayer-Training mit Auszubildenden

Fragebogen

1 Demografische Daten

Alter

Bitte gebe dein Alter im Eingabefeld in ganzen Zahlen an (z.B. 21).

Geschlecht

Bitte gebe an mit welchem Geschlecht du dich identifizieren.

- Männlich
- Weiblich
- Divers

2 VR Modus

Was für eine Virtual Reality Schulung hast du soeben besucht?

Falls du dir unsicher bist, frage bei einer zur Verfügung stehenden Person nach.

- Multiplayer
- Multiplayer mit Bodyswap
- Classroom

3 Wahrgenommene Nutzerfreundlichkeit

Wahrgenommene Nützlichkeit

Bitte gebe an inwieweit du den folgenden Aussagen zu stimmst.

	1 Stimme überhaupt nicht zu	2 Stimme nicht zu	3 Stimme weder zu noch lehne ich ab	4 Stimme zu	5 Stimme voll und ganz zu
Meine Interaktion mit der Simulation war klar und verständlich.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Anwendung der Simulation ist einfach zu bedienen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Bedienung der Simulation zu erlernen war für mich einfach.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich war in der Simulation selten verwirrt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich habe in der Simulation selten Fehler gemacht.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich war in der Simulation selten frustriert.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

4 Wahrgenommener Nutzen

Wahrgenommener Nutzen

Bitte gebe an inwieweit du den folgenden Aussagen zu stimmst.

	1 Stimme überhaupt nicht zu	2 Stimme nicht zu	3 Stimme weder zu noch lehne ich ab	4 Stimme zu	5 Stimme voll und ganz zu
Ich würde ein Virtual Reality Training als Vorbereitung auf meine Arbeit nützlich finden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ich ein Virtual Reality Training als Vorbereitung auf meine Arbeit einsetzen würde, könnte ich meine Aufgaben schneller erledigen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Verwendung eines Virtual Reality Trainings würde meine Arbeitsleistung verbessern.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich fand den Einsatz des Virtual Reality Tools hilfreich.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Instruktionen zur Bedienung des Virtual Reality Tools waren hilfreich.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

5 Nutzungsabsicht

Nutzungsabsicht

Bitte gebe an inwieweit du den folgenden Aussagen zu stimmst.

	1 Stimme überhaupt nicht zu	2 Stimme nicht zu	3 Stimme weder zu noch lehne ich ab	4 Stimme zu	5 Stimme voll und ganz zu
Sobald ein Virtual Reality System als Lernmethode in meiner Ausbildung verfügbar wäre, würde ich es nutzen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Angenommen, ich hätte Zugang zu einem Virtual Reality System, würde ich es nutzen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich denke, dass ich Virtual Reality Systeme häufig nutzen werde.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

6 Vergnügen

Vergnügen

Bitte gebe an inwieweit du den folgenden Aussagen zu stimmst.

	1 Stimme überhaupt nicht zu	2 Stimme nicht zu	3 Stimme weder zu noch lehne ich ab	4 Stimme zu	5 Stimme voll und ganz zu
Die Tätigkeiten in der Simulation haben mir Spass gemacht.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich fand die Tätigkeiten in der Simulation sehr interessant.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Tätigkeiten in der Simulation waren unterhaltsam.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

7 Wahrgenommene Kompetenz

Wahrgenommene Kompetenz

Bitte gebe an inwieweit du den folgenden Aussagen zu stimmst.

	1 Stimme überhaupt nicht zu	2 Stimme nicht zu	3 Stimme weder zu noch lehne ich ab	4 Stimme zu	5 Stimme voll und ganz zu
Mit meiner Leistung in der Simulation bin ich zufrieden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Bei den Tätigkeiten in der Simulation stellte ich mich geschickt an.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich glaube, ich war bei den Tätigkeiten in der Simulation ziemlich gut.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

8 Autonomie

Autonomie

Bitte gebe an inwieweit du den folgenden Aussagen zu stimmst.

	1 Stimme überhaupt nicht zu	2 Stimme nicht zu	3 Stimme weder zu noch lehne ich ab	4 Stimme zu	5 Stimme voll und ganz zu
Während der Simulation hatte ich das Gefühl, dass ich meine Erfahrungen und Handlungen unter Kontrolle hatte.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Während der Simulation hatte ich das Gefühl, dass meine Erfahrungen und Handlungen nicht von mir gesteuert wurden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Während der Simulation hatte ich das Gefühl, dass meine Erfahrungen und Handlungen von mir selbst gesteuert wurden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

9 Physische Präsenz

Physische Präsenz

Bitte gebe an inwieweit du den folgenden Aussagen zu stimmst.

	1 Stimme überhaupt nicht zu	2 Stimme nicht zu	3 Stimme weder zu noch lehne ich ab	4 Stimme zu	5 Stimme voll und ganz zu
Die Simulation kam mir real vor.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich hatte den Eindruck, in der virtuellen Umgebung handeln zu können, anstatt etwas von aussen zu bedienen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Meine Erlebnisse in der Simulation schienen mit denen in der realen Welt übereinzustimmen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich hatte das Gefühl, in der Simulation "präsent" zu sein.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Simulation hat mich völlig in ihren Bann gezogen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

10 Selbstwirksamkeit

Selbstwirksamkeit

Bitte gebe an inwieweit du den folgenden Aussagen zu stimmst.

	1 Stimme überhaupt nicht zu	2 Stimme nicht zu	3 Stimme weder zu noch lehne ich ab	4 Stimme zu	5 Stimme voll und ganz zu
Ich bin zuversichtlich, dass ich die grundlegenden Konzepte von Gleissicherheit verstehen kann.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bin zuversichtlich, dass ich auch die komplexen Sachverhalte zu Gleissicherheit verstehen kann.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bin zuversichtlich, dass ich bei Aufgaben und Tests über Gleissicherheit eine gute Leistung erbringen kann.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bin zuversichtlich, dass ich einem Freund/ einer Freundin die Grundlagen von Gleissicherheit erklären kann.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

11 Kognitive Belastungen

Kognitive Belastungen (Interaktion)

Bitte gebe an inwieweit du den folgenden Aussagen zu stimmst.

	1 Stimme überhaupt nicht zu	2 Stimme nicht zu	3 Stimme weder zu noch lehne ich ab	4 Stimme zu	5 Stimme voll und ganz zu
Die in der Simulation verwendete Interaktionstechnik war sehr unklar.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die in der Simulation verwendete Interaktionstechnik war im Hinblick auf das Lernen sehr ineffektiv.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die in der Simulation verwendete Interaktionstechnik erschwerte das Lernen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die in der Simulation verwendete Interaktionstechnik war schwer zu meistern.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

12 Kognitive Belastungen 2

Kognitive Belastungen (Umfeld)

Bitte gebe an inwieweit du den folgenden Aussagen zu stimmst.

	1 Stimme überhaupt nicht zu	2 Stimme nicht zu	3 Stimme weder zu noch lehne ich ab	4 Stimme zu	5 Stimme voll und ganz zu
Die Elemente in der Simulation machten das Lernen sehr unklar.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
In Bezug auf das Lernen war die Simulation ineffektiv.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Simulation war voll von unwichtigen Inhalten.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Es war schwierig, die relevanten Informationen in der Simulation zu finden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

13 Interesse

Interesse

Bitte gebe an inwieweit du den folgenden Aussagen zu stimmst (1= keineswegs, 4= sehr).

	1	2	3	4
Hat die Lektion deine Neugierde geweckt?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Hat die Lektion deine Aufmerksamkeit geweckt?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Warst du auf die Lektion konzentriert?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
War die Lektion für dich unterhaltsam?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Hattest du während der Lektion Spass?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
War die Lektion für dich spannend?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

14 Abschluss

Das möchte ich euch noch mitteilen:

15 Endseite

Vielen Dank für deine Teilnahme!

Du kannst die Umfrage nun schliessen.

16 Endseite

Anhang D
Items zur Studie IV – Ergänzung SstA mit iVR

Item Original	Item Deutsch & angepasst	Quelle
Kognitiv-affektives Modell des immersiven Lernens (CAMIL)		Petersen et al. (2022)
Physical Presence (P)	Physische Präsenz	Makransky et al. (2017)
The virtual environment seemed real to me.	Das Lerntool kam mir real vor	
I had a sense of acting in the virtual environment, rather than operating something from outside.	Ich hatte den Eindruck, mit dem Lerntool handeln zu können, anstatt etwas von aussen zu bedienen	
My experience in the virtual environment seemed consistent with my experiences in the real world.	Meine Erlebnisse mit dem Lerntool schienen mit denen in der realen Welt übereinzustimmen	
I had a sense of “being there” in the virtual environment.	Ich hatte das Gefühl, im Lerntool präsent zu sein	
I was completely captivated by the virtual world.	Das Lerntool hat mich völlig in den Bann gezogen	
Agency (A)	Handlungsfähigkeit	Polito et al. (2013)
During the lesson, my experiences and actions were under my control.	Während der Arbeit mit dem Lerntool hatte ich das Gefühl, dass ich meine Erfahrungen und Handlungen unter Kontrolle hatte	
During the lesson, I felt that my experiences and actions were not caused by me. (R)	Während der Arbeit mit dem Lerntool hatte ich das Gefühl, dass meine Erfahrungen und Handlungen nicht von mir gesteuert wurden (R)	
During the lesson, my experiences and actions felt self-generated.	Während der Arbeit mit dem Lerntool hatte ich das Gefühl, dass meine Erfahrungen und Handlungen von mir selbst gesteuert wurden	
Self-Efficacy (S)	Selbstwirksamkeit	Meyer et al. (2019)
I'm confident I can understand the basic concepts of viral diseases.	Ich bin zuversichtlich, dass ich die grundlegenden Konzepte von Arbeitsstellensicherheit verstehen kann	
I'm confident I can understand the most complex material regarding viral diseases.	Ich bin zuversichtlich, dass ich auch die komplexen Sachverhalte zur Arbeitsstellensicherheit verstehen kann	

I'm confident I could do an excellent job on assignments and tests about viral diseases.	Ich bin zuversichtlich, dass ich bei Aufgaben und Tests über die Arbeitsstellensicherheit eine gute Leistung erbringen kann	
I'm confident I could explain the basic causes of epidemics and pandemics to a friend.	Ich bin zuversichtlich, dass ich einem Freund/ einer Freundin die Grundlagen von Arbeitsstellensicherheit erklären kann	
Extraneous Cognitive Load Interaction (ECLI)	Extrinsische kognitive Belastung Interaktion	Andersen und Makransky (2021)
The interaction technique used in the simulation was very unclear.	Die bei der Arbeit mit dem Lerntool verwendete Interaktionstechnik war sehr unklar	
The interaction technique used in the simulation was, in terms of learning, very ineffective.	Die beim Lerntool verwendete Interaktionstechnik war im Hinblick auf das Lernen sehr ineffektiv	
The interaction technique used in the simulation made it harder to learn.	Die bei der Arbeit mit dem Lerntool verwendete Interaktionstechnik erschwerte das Lernen	
The interaction technique used in the simulation was difficult to master.	Die bei der Arbeit mit dem Lerntool verwendete Interaktionstechnik war schwer zu meistern	
Extraneous Cognitive Load Environment (ECLE)	Extrinsische kognitive Belastung Umwelt	Andersen und Makransky (2021)
The elements in the virtual environment made the learning very unclear.	Die Elemente des Lerntools machten das Lernen sehr unklar	
The virtual environment was, in terms of learning, very ineffective.	In Bezug auf das Lernen war das Lerntool ineffektiv	
The virtual environment was full of irrelevant content.	Das Lerntool war voll von unwichtigen Inhalten	
It was difficult to find the relevant learning information in the virtual environment.	Es war schwierig, die relevanten Informationen in der Simulation zu finden	
Situational Interest (SI)	Situatives Interesse	Knogler et al. (2015)
Did the lesson spark your curiosity?	Hat die Arbeit mit dem Lerntool Ihre Neugierde geweckt?	

Did the lesson capture your attention?	Hat die Arbeit mit dem Lerntool Ihre Aufmerksamkeit geweckt?	
Where you concentrated on the lesson?	Waren Sie auf die Arbeit mit dem Lerntool konzentriert?	
Was the lesson entertaining for you?	War die Arbeit mit dem Lerntool für Sie unterhaltsam?	
Did you have fun during the lesson?	Hatten Sie während der Arbeit mit dem Lerntool Spass?	
Was the lesson exciting for you?	War die Arbeit mit dem Lerntool für Sie spannend?	
Kurzskala intrinsischer Motivation (KIM)		
Enjoyment / Interest (E)	Vergnügen / Interesse	Wilde et al. (2009)
Die Tätigkeit in der Ausstellung hat mir Spass gemacht.	Die Tätigkeiten mit dem Lerntool haben mir Spass gemacht	
Ich fand die Tätigkeit in der Ausstellung sehr interessant.	Ich fand die Tätigkeiten mit dem Lerntool sehr interessant	
Die Tätigkeit in der Ausstellung war unterhaltsam.	Die Tätigkeiten mit dem Lerntool waren unterhaltsam	
Perceived Competence (PC)	Wahrgenommene Kompetenz	Wilde et al. (2009)
Mit meiner Leistung in der Ausstellung bin ich zufrieden.	Mit meiner Leistung bei der Arbeit mit dem Lerntool bin ich zufrieden	
Bei der Tätigkeit in der Ausstellung stellte ich mich geschickt an.	Bei den Tätigkeiten mit dem Lerntool stellte ich mich geschickt an	
Ich glaube, ich war bei der Tätigkeit in der Ausstellung ziemlich gut.	Ich glaube, ich war bei den Tätigkeiten mit dem Lerntool ziemlich gut	
Modell der Technikakzeptanz (TAM)		
Perceived Ease of Use (PEOU)	Wahrgenommene Benutzerfreundlichkeit	Özgen et al. (2021)
My interaction with the virtual reality has been clear and understandable	Meine Interaktion mit dem Lerntool war klar und verständlich	

Overall, the virtual reality is easy to use	Die Anwendung des Lerntools ist einfach zu bedienen	
Learning to operate the virtual reality was easy for me	Die Bedienung des Lerntools zu erlernen war für mich einfach	
I rarely become confused when I use the virtual reality	Ich war bei der Arbeit mit dem Lerntool selten verwirrt	
I rarely make errors when using the virtual reality	Ich habe bei der Arbeit mit dem Lerntool selten Fehler gemacht	
I am rarely frustrated when using the virtual reality	Ich war bei der Arbeit mit dem Lerntool selten frustriert	
Perceived Usefulness (PU)	Wahrgenommene Nützlichkeit	Özgen et al. (2021)
Virtual reality would be useful as part of my work.	Ich würde die Nutzung des Lerntools als Vorbereitung auf meine Arbeit nützlich finden	
Virtual reality enables me to accomplish task more quickly	Wenn ich das Lerntool als Vorbereitung auf meine Arbeit einsetzen würde, könnte ich meine Aufgaben schneller erledigen	
Virtual reality has improved my quality of work	Die Verwendung des Lerntools würde meine Arbeitsleistung verbessern	
Virtual reality makes it easier to do my job	Ich fand den Einsatz des Lerntools hilfreich	
Virtual reality gives me greater control over my job	Die Instruktionen zur Bedienung des Lerntools waren hilfreich	
Intention to Use (I)	Nutzungsabsicht	
I intend to use any system using virtual reality when it becomes available in basic design courses	Sobald das Lerntool als Lernmethode in meiner Ausbildung verfügbar wäre, würde ich es nutzen	Özgen et al. (2021)
Assuming I had access to the virtual reality system, I intend to use it	Angenommen, ich hätte Zugang zu dem Lerntool, würde ich es nutzen	
I predict that I would use virtual reality frequently	Ich denke, dass ich das Lerntool häufig nutzen werde	
Fragen der SBB zu Risikowahrnehmung und Vorbereitung auf die Arbeit		SBB

-	Im Vergleich zu Schulungen ohne dem Lerntool fühle ich mich besser auf die Arbeit in der Praxis vorbereitet	SBB (Paul Hügli)
-	Im Vergleich zu Schulungen ohne dem Lerntool fühle ich mich sicherer für die Arbeit in der Praxis vorbereitet	
-	Das Lerntool hat mir eindrücklich die Gefahren im Gleisbereich vermittelt	
-	Das Erlebnis mit dem Lerntool hat mein Situationsbewusstsein positiv beeinflusst	
-	Die Gefahren im Gleisbereich konnten mir mit dem Lerntool sehr deutlich gemacht werden	
Abschliessendes Feedback oder Anmerkungen (offene Frage)		
-	Haben Sie Anmerkungen zum heutigen Training mit dem Lerntool?	-