



Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG

Entwicklung und Evaluation eines Leseassistenten in Virtual Reality

Bachelorarbeit im Studiengang Mensch-Computer-Interaktion

Autor:

Nathanael Strauß

Matrikel: 7158789

Human-Computer Interaction
Fachbereich Informatik

Erstgutachter: Prof. Dr. Frank Steinicke
Zweitgutachter: Jenny Gabel
Betreuer: Jenny Gabel

Hamburg, 15. September 2024

INHALTSVERZEICHNIS

1	Einleitung	7
2	Verwandte Arbeiten	8
3	Technische Implementation der Leseassistentenfunktionen	9
3.1	Technische Umsetzung generell	9
3.2	Scanpath (Einschbarer Blickverlauf)	9
3.2.1	Richtwerte	10
3.2.2	Implementationsdetails	11
3.3	Peripheral Blur (Zeilenfokus)	12
4	Benutzerstudie	14
4.1	Textparameter	14
4.2	Studiendesign	15
4.3	Ablauf der Studie	16
5	Ergebnisse	16
5.1	Demografie und Vorerfahrungen in VR	17
5.2	Leseverhalten	17
5.3	Quantitative Analyse	17
5.3.1	Simulator Sickness Questionnaire (SSQ)	18
5.3.2	Lesewiederholungen	18
5.3.3	Lesedauern	18
5.3.4	Antworten zu Textfragen	19
5.3.5	NASA RAW-TLX	19
5.3.6	Scanpath	20
5.3.7	Peripheral Blur	21
5.4	Qualitative Analyse	21
6	Diskussion	25
6.1	Simulator Sickness	25
6.2	NASA RAW-TLX	25
6.3	Inhaltlicher Abruf der Texte	25
6.4	Gesamtlesedauer	26
6.5	Lesewiederholungen	26
6.6	Korrelationen	26
6.6.1	Scanpath-Dauer	26
6.6.2	Scanpath-Häufigkeit	26
6.6.3	Peripheral Blur: Zeileneinstellungen	26
6.6.4	Peripheral Blur: Anzahl Zeileneinstellungen	26
6.7	Qualitative Ergebnisse	26
7	Limitationen und Future Work	27
8	Abschluss und Fazit	27

A	Anhang	30
A.1	Veranschaulichende Grafiken für Abläufe bzw. Aufbau der Anwendung	30
A.2	Boxplots: NASA RAW-TLX	35
A.3	Boxplots: Kondition Scanpath	36
A.4	Boxplots: Kondition Peripheral Blur	37
A.5	Residuen: SSQ: Übelkeits-Score (N)	38
A.6	Residuen: SSQ: Übelkeits-Score (O)	39
A.7	Residuen: SSQ: Übelkeits-Score (D)	40
A.8	Residuen: SSQ: Übelkeits-Score (TS)	41
A.9	Residuen: Lesewiederholungen	42
A.10	Residuen: Lesedauern	43
A.11	Residuen: Richtige Antworten gesamt	44
A.12	Residuen: NASA RAW-TLX: Subskala Geistige Anforderung	45
A.13	Residuen: NASA RAW-TLX: Subskala Körperliche Anforderung	46
A.14	Residuen: NASA RAW-TLX: Subskala Zeitliche Anforderung	47
A.15	Residuen: NASA RAW-TLX: Subskala Leistung	48
A.16	Residuen: NASA RAW-TLX: Subskala Anstrengung	49
A.17	Residuen: NASA RAW-TLX: Subskala Frustration	50
A.18	Dummy-Text aus Testrunden	51
A.19	Text 1	52
A.20	Text 2	53
A.21	Text 3	54

ABBILDUNGSVERZEICHNIS

1	Eine Scanpath-Variante. Kreise visualisieren Fixationen der Leser:innen an ihrer Position im Text. Farbe und Größe bilden die Dauer ab, Pfeile die Rücksprünge der Augen.	7
2	Ein Filter zum Scharfstellen der anvisierten Zeilen und Weichzeichnen des umliegenden Textes. Der fokussierte Bereich wird abhängig von der Kopfausrichtung im Text platziert und ist in seinen Zeilenhöhen verstellbar.	7
3	Exemplarischer Scanpath aus gleichnamiger Leseassistentz-Funktion. Kreise bilden in Größe und Farbe Fixationsdauern ab, Rückpfeile Regressionen. Über den linken Button kann die Seite direkt erneut und ohne Scanpath gelesen werden, während im Hintergrund neue Scanpath-Parameter erhoben werden. Erst mit Auswählen des rechten Buttons wird direkt zur nächsten Seite geblättert.	10
4	Leseassistentz-Funktion Scanpath und ihre Komponenten. (a): Das DataCapture-Package mit der modifizierten Schnittstelle InputCaptureSystem. (b): Pro Update()-Abruf wird über eine Liste aus Gazeays iteriert. (c) Taktung der Update()-Funktion nach der Abtaste von 90Hz. (d) Textplane mit Größe und z-Höhe des Texthintergrunds. (e) Raycast: Schnittpunkt aus Kollision des Gazeray mit der Textplane. (f) Winkel zwischen Gazerays (multipliziert mit 90 für °/s). (g) Schnittpunkte (rot) werden Fixationscluster (graue Ellipse) hinzugefügt. (h) Sakkaden-Kriterium erfüllt (schwarzer Schnittpunkt): Fixationscluster wird zu einer Position gemittelt (großes „X“), Fixationsdauer wird berechnet. (i) Visualisierung durch Auswählen des Buttons zum Wechsel auf nächste Seite. (j) Visualisierung der Regressionen.	13
5	Lesen mit Assistenz-Funktion Peripheral Blur. Die Person hat die Einstellung von drei fokussierten Zeilenhöhen ausgewählt und visiert den mittleren Teil des Textes an.	14
6	Leseassistentz-Funktion Peripheral Blur und ihre Komponenten. (a): Unity Shadergraph wird für den Blur-Effekt auf Material von zwei Shapes angewandt. Dieser macht die Shapes komplett transparent (grau für Illustration) und wirkt wie ein Weichzeichner auf dahinterliegende Objekte. (b): Beide Shapes werden übereinander angeordnet und einem gemeinsamen Parent Object zugeordnet. Dies ergibt den Blur-Filter. (c): Der Abstand zwischen den Shapes (fokussierter Bereich) ist über den Controller verstellbar. A-Button: Fügt eine Zeilenhöhe hinzu, B-Button: Nimmt eine Zeilenhöhe weg. (d): Der nose-direction-Ray schneidet abhängig von der Kopfdrotation (<i>pitch</i>) eine unendlich große Plane auf gleicher z-Höhe des Textpanelss (rotes Gitter). Der resultierende Schnittpunkt wird in seiner y- und z-Komponente auf den Transform des Blurfilters angewandt.	15
7	Einfache UI für den Seitenwechsel. Pfeilförmige Buttons wechseln zur vorherigen bzw. nächsten Seite. Die zentrierte Seitenanzeige ist im Bruchformat „X/4“.	15
8	Ablauf der Benutzerstudie als Flussdiagramm mit nachgestelltem VR-Setup und Screenshots aus der Anwendung.	16
9	Leseparameter: (a) Anzahl Lesewiederholungen und (b) Gesamtlesedauern in Sekunden	18
10	SSQ: Okulomotorik-Score (O)	18
11	Informationsabruf pro Frage: Anzahl richtiger und falscher Antworten pro Leseassistentz-Funktion bzw. Baseline-Kondition. Es wurde hierbei nicht zwischen Texten unterschieden, sondern nur nach Positionierung der Frage am Anfang (a), Mitte (b) oder Ende (c) des Textes gefiltert.	19
12	Subskalen des NASA RAW-TLX: Leistung (a), Anstrengung (b), Total Workload (c).	20
13	Kondition Peripheral Blur: Häufigkeiten für die letzte Zeileinstellung.	21
14	Leseassistentz-Funktion Scanpath und ihre Komponenten. (a): Das DataCapture-Package mit der modifizierten Schnittstelle InputCaptureSystem. (b): Pro Update()-Abruf wird über eine Liste aus Gazeays iteriert. (c) Taktung der Update()-Funktion nach der Abtaste von 90Hz. (d) Textplane mit Größe und z-Höhe des Texthintergrunds. (e) Raycast: Schnittpunkt aus Kollision des Gazeray mit der Textplane. (f) Winkel zwischen Gazerays (multipliziert mit 90 für °/s). (g) Schnittpunkte (rot) werden Fixationscluster (graue Ellipse) hinzugefügt. (h) Sakkaden-Kriterium erfüllt (schwarzer Schnittpunkt): Fixationscluster wird zu einer Position gemittelt (großes „X“), Fixationsdauer wird berechnet. (i) Visualisierung durch Auswählen des Buttons zum Wechsel auf nächste Seite. (j) Visualisierung der Regressionen.	30

15	Leseassistentz-Funktion Peripheral Blur und ihre Komponenten. (a): Unity Shadergraph wird für den Blur-Effekt auf Material von zwei Shapes angewandt. Dieser macht die Shapes komplett transparent (grau für Illustration) und wirkt wie ein Weichzeichner auf dahinterliegende Objekte. (b): Beide Shapes werden übereinander angeordnet und einem gemeinsamen Parent Object zugeordnet. Dies ergibt den Blur-Filter. (c): Der Abstand zwischen den Shapes (fokussierter Bereich) ist über den Controller verstellbar. A-Button: Fügt eine Zeilenhöhe hinzu, B-Button: Nimmt eine Zeilenhöhe weg. (d): Der nose-direction-Ray schneidet abhängig von der Kopffrotation (<i>pitch</i>) eine unendlich große Plane auf gleicher z-Höhe des Textpanelss (rotes Gitter). Der resultierende Schnittpunkt wird in seiner y- und z-Komponente auf den Transform des Blurfilters angewandt.	31
16	Ablauf der Benutzerstudie als Flussdiagramm mit nachgestelltem VR-Setup und Screenshots aus der Anwendung.	32
17	Proband in Seated VR (nachgestellt).	33
18	SSQ: Übelkeits-Score (N)	34
19	SSQ: Desorientierungs-Score (D)	34
20	SSQ: Gesamt-Score (TS)	34
21	NASA RAW-TLX: Subskala Geistige Anforderung	35
22	NASA RAW-TLX: Subskala Körperliche Anforderung	35
23	NASA RAW-TLX: Subskala Zeitliche Anforderung	35
24	NASA RAW-TLX: Subskala Frustration	35
25	Kondition Scanpath: Scanpath-Dauer	36
26	Kondition Scanpath: Scanpath-Häufigkeit	36
27	Kondition Peripheral Blur: Zeileneinstellung	37
28	Kondition Peripheral Blur: Anzahl Zeileneinstellungen	37
29	Residuen-Plot: SSQ: Übelkeits-Score (N)	38
30	Residuen-Histogramm: SSQ: Übelkeits-Score (N)	38
31	Normal-Q-Q-Plot: SSQ: Übelkeits-Score (N)	38
32	Residuen-Plot: SSQ: Okulomotorik-Score (O)	39
33	Residuen-Histogramm: SSQ: Okulomotorik-Score (O)	39
34	Normal-Q-Q-Plot: SSQ: Okulomotorik-Score (O)	39
35	Residuen-Plot: SSQ: Desorientierungs-Score (D)	40
36	Residuen-Histogramm: SSQ: Desorientierungs-Score (D)	40
37	Normal-Q-Q-Plot: SSQ: Desorientierungs-Score (D)	40
38	Residuen-Plot: SSQ: Gesamt-Score (TS)	41
39	Residuen-Histogramm: SSQ: Gesamt-Score (TS)	41
40	Normal-Q-Q-Plot: SSQ: Gesamt-Score (TS)	41
41	Residuen-Plot: Lesewiederholungen	42
42	Residuen-Histogramm: Lesewiederholungen	42
43	Normal-Q-Q-Plot: Lesewiederholungen	42
44	Residuen-Plot: Lesedauern	43
45	Residuen-Histogramm: Lesedauern	43
46	Normal-Q-Q-Plot: Lesedauern	43
47	Residuen-Plot: Richtige Antworten gesamt	44
48	Residuen-Histogramm: Richtige Antworten gesamt	44
49	Normal-Q-Q-Plot: Richtige Antworten gesamt	44
50	Residuen-Plot: NASA RAW-TLX: Subskala Geistige Anforderung	45
51	Residuen-Histogramm: NASA RAW-TLX: Subskala Geistige Anforderung	45
52	Normal-Q-Q-Plot: NASA RAW-TLX: Subskala Geistige Anforderung	45

53	Residuen-Plot: NASA RAW-TLX: Subskala Körperliche Anforderung	46
54	Residuen-Histogramm: NASA RAW-TLX: Subskala Körperliche Anforderung	46
55	Normal-Q-Q-Plot: NASA RAW-TLX: Subskala Körperliche Anforderung	46
56	Residuen-Plot: NASA RAW-TLX: Subskala Zeitliche Anforderung	47
57	Residuen-Histogramm: NASA RAW-TLX: Subskala Zeitliche Anforderung	47
58	Normal-Q-Q-Plot: NASA RAW-TLX: Subskala Zeitliche Anforderung	47
59	Residuen-Plot: NASA RAW-TLX: Subskala Leistung	48
60	Residuen-Histogramm: NASA RAW-TLX: Subskala Leistung	48
61	Normal-Q-Q-Plot: NASA RAW-TLX: Subskala Leistung	48
62	Residuen-Plot: NASA RAW-TLX: Subskala Anstrengung	49
63	Residuen-Histogramm: NASA RAW-TLX: Subskala Anstrengung	49
64	Normal-Q-Q-Plot: NASA RAW-TLX: Subskala Anstrengung	49
65	Residuen-Plot: NASA RAW-TLX: Subskala Frustration	50
66	Residuen-Histogramm: NASA RAW-TLX: Subskala Frustration	50
67	Normal-Q-Q-Plot: NASA RAW-TLX: Subskala Frustration	50
68	Konstruktion des Counterbalanced Latin Square nach der Methode von Lewis (1989, 1993).	55

TABELLENVERZEICHNIS

1	Allgemeine Kategorisierung des Feedbacks nach Nützlichkeit. Es wurde pro Funktion jede Person gezählt, bei der die entsprechende Feedback-Kategorie mindestens einmal vertreten war.	22
2	Für jede Erfahrungsstufe mit VR: Pro Leseassistentenfunktion wurden die Personen gezählt, welche die Funktion hilfreich, teilweise hilfreich und nicht hilfreich fanden.	22

Entwicklung und Evaluation eines Leseassistenten in Virtual Reality

Nathanael Strauß*
Universität Hamburg

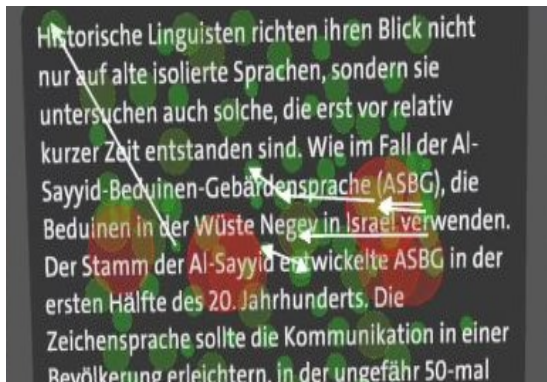


Abbildung 1: Eine Scanpath-Variante. Kreise visualisieren Fixationen der Leser:innen an ihrer Position im Text. Farbe und Größe bilden die Dauer ab, Pfeile die Rücksprünge der Augen.

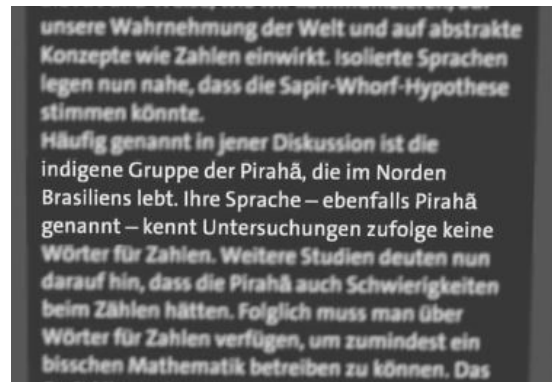


Abbildung 2: Ein Filter zum Scharfstellen der anvisierten Zeilen und Weichzeichnen des umliegenden Textes. Der fokussierte Bereich wird abhängig von der Kopfausrichtung im Text platziert und ist in seinen Zeilenhöhen verstellbar.

ZUSAMMENFASSUNG

Um implizite Hilfestellungen beim Lesen zu geben, die auf verschiedene Aspekte der typischen Leseproblematiken im akademischen Kontext abzielen, wurde eine Anwendung entwickelt, die eine Erfahrung der AR aufgrund ihrer technischen Beschränkungen in VR simuliert. Diese enthält zwei Leseassistenten-Funktionen: Scanpath und Peripheral Blur. Erstere lässt die Leser:innen ihren eigenen Lesefluss nach dem Lesen einer Seite nachvollziehen, um ggf. zu einer Anpassung im Leseverhalten zu animieren. Zweitere verstärkt den Effekt des verschwommenen Sehens in der Peripherie mit einem DR-Filter, um den Fokus auf aktuelle Textstellen zu erhöhen und Ablenkungen durch umliegenden Text zu reduzieren. Beide Funktionen wurden in einer Lesestudie in VR hinsichtlich inhaltlichen Abrufs (*Recall*), Leseparameter sowie Funktions-spezifischer Parameter und User Experience evaluiert. Der Großteil der Proband:innen nahm den Peripheral Blur als die hilfreichere Assistenzform wahr, während die meisten den Scanpath lediglich als interessant empfanden. Jedoch konnte keiner der statistischen Tests bestätigen, dass das assistierte Lesen in VR eine Verbesserung bezüglich genannter Merkmale im Vergleich zum Lesen in der VR ohne Leseassistenten bewirkte. Mit der Implementation und Evaluation zwei neuartiger Formen der Text-Augmentation leistet diese Arbeit einen Beitrag zu den assistiven Technologien zum Lesen durch feedbackbasierte bzw. unauffällige (*unobtrusive*) Visualisierungstechniken im Benutzungs-Kontext der Mixed Reality und liefert diesbezüglich Erkenntnisse über Usability-Aspekte und User Experience.

Index Terms: Human-centered computing—Human computer interaction (HCI)—HCI theory, concepts and models—; Human-centered computing—Human computer interaction (HCI)—Interactive systems and tools—; Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Virtual reality; Human-centered computing—Visualization—Visualization techniques—; Human-centered computing—Accessibility—Accessibility systems and tools—; Human-centered computing—Human computer interaction (HCI)—HCI design and evaluation methods—User studies;

1 EINLEITUNG

Da das Aneignen von Literatur eine der Schlüsselkompetenzen des Studiums darstellt, sind Student:innen häufig mit dem Lesen wissenschaftlicher Texte konfrontiert. Diese sind häufig auf akademischem Niveau und/oder englischsprachig, wobei das neue Fachvokabular und die komplexen Satzstrukturen den kognitiven und zeitlichen Aufwand erhöhen können, weshalb gerade diesem relevanten Prozess häufig mit Frustration begegnet wird.

Darüber hinaus – und dies entspricht der persönlichen Erfahrung – muss bei längeren und inhaltlich komplexeren Sätzen über einen größeren Zeitraum Konzentration aufgebracht werden, um den Textinhalt zusammenzuhalten. Andernfalls können vorherige Satzteile und somit die Zusammenhänge des Satzes entgleiten. Jedoch kann der Kontext oder Subtext von Sätzen ebenso verloren gehen, wenn Texte unaufmerksam überflogen werden, um nur die nötigste Information zu filtern, um z.B. eine konkrete Antwort auf eine gezielte Fragestellung (wie etwa für eine Hausarbeit) zu finden. Aber auch die Ermüdung über die Zeit kann zu einem unreflektierten Leseverhalten führen.

Genauso können Faktoren von außen einen störenden Einfluss auf das Lesen haben. Dazu gehören Ablenkungen auditiver Form, wie etwa Verkehrslärm oder visuelle Stimuli in der nahen Umgebung. Schon die umliegende Textmenge selbst kann hierbei die Augen ablenken, wodurch der Fokus auf den aktuellen Textabschnitt verloren geht. Auch die Leser:innen können von sich aus eine Anfälligkeit für diese Ablenkungen mitbringen.

Ein weiteres Problem, dem beim Lesen wissenschaftlicher Texte auch persönlich begegnet wurde, ist eine unproportionale Fixierung auf bestimmte Textabschnitte im Vergleich zum Rest des Textes, wenn Verständnisprobleme auftreten. So können komplexe oder mehrdeutige Formulierungen zu einem wiederholten Nachlesen führen, was eine weitere Frustrationsquelle beim Lesen darstellt. Oft wird die Relevanz einer Textstelle nicht hinterfragt, was zu diesem Problem erheblich beiträgt, da somit viel Zeit für weniger wichtige Teile verschwendet wird.

Im Vordergrund dieser Problematiken entstand die Idee eines Leseassistenten, welcher Leser:innen mit virtuellen Hilfestellungen

*e-mail: nathanael.strauss@studium.uni-hamburg.de

visueller Form durch den Text begleitet. Hierfür wurden zwei Leseassistenten-Funktionen in Virtual Reality (VR) konzipiert, welche verschiedene Aspekte der Leseproblematik adressieren sollen: *Scanpath* und *Peripheral Blur*.

Erstere gibt nach dem Lesen ein visuelles Feedback, welches einer Variante zu herkömmlichen Scanpaths entspricht und die Leser:innen zu einer Reflektion über das eigene Leseverhalten und im besten Fall zu einer Anpassung in diesem animieren soll. Somit würde die erste Leseassistenten-Funktion eine mehr explorative Methode darstellen, sich zusammen mit dem eigenen Leseverhalten tiefer mit dem Text auseinanderzusetzen im Tausch gegen eine längere Lesezeit und häufigeres Nachlesen. Ziel dessen ist es auch, ein Gefühl über die eigene Lesezeit zu bekommen und die Relevanz von Textstellen zu hinterfragen. So könnten sich Leser:innen im Fall einer Überfixierung eines Textabschnitts aufgrund von Verständnisproblemen gar nicht darüber bewusst sein, wie lange sie schon mit dieser Textstelle verbracht haben. Der Scanpath soll auch diese Verhaltensweisen vor Augen führen.

Die zweite Leseassistenten-Funktion ist inspiriert durch eine Eigenschaft des visuellen Systems, nämlich die Abnahme der räumlichen Auflösung in der Peripherie. Eine künstliche Verstärkung dieser Eigenschaft durch die Anwendung eines DR-Effekts auf umliegenden Text abhängig von der Kopfausrichtung soll den Fokus auf die anvisierte Textstelle erhöhen und auf diese Weise die angesprochene Problematik visueller Störreize innerhalb des Textes reduzieren. Somit soll schon während des Lesens eine Anpassung im Verhalten erfolgen, welche sich mehr auf die Aspekte der Effizienz beschränkt, d.h. kürzere Lesedauer und weniger Lesewiederholungen.

Da die Anwendung eine Unterstützung für das Lesen von Texten auf Papier oder Bildschirmen bieten sollte, war sie ursprünglich für die Augmented Reality (AR) vorgesehen. Jedoch wäre hierfür eine Texterkennung in Echtzeit notwendig gewesen. Eine weitere Schwierigkeit hätte der Echtzeitblur des Textes dargestellt, wobei ebenfalls nicht sicher war, ob die Deckkraft der Hologramme für ein Blur-Overlay (basierend auf Bildausschnitten) ausreichen würde. Aufgrund dieser technischen Beschränkungen wurde sich für die VR entschieden, da sich Text und Umgebung in einer komplett virtuellen Umgebung frei manipulieren ließen. Eine Umsetzung in AR wäre somit Gegenstand zukünftiger Arbeiten.

Auf der anderen Seite bringt Lesen in VR selbst Nachteile mit sich. So zeigten Studien, dass das Lesen auf Bildschirmen im Vergleich zu Lesen auf Papier zu Einbußen in Verständnis, Genauigkeit und Geschwindigkeit führt [10, 40]. Im Vergleich mit dem herkömmlichen LCD-Display für das Lesen am Desktop-PC konnten Rau et al. [48] zeigen, dass das Beantworten von Multiple Choice Fragen in VR 10% langsamere Antworten generierte, und in einer späteren Studie, dass zudem das Lesen in VR selbst länger dauerte als mit LCD-Displays [47]. Jedoch besteht mit den Leseassistenten-Funktionen wiederum das Potenzial, das Lesen in VR hinsichtlich Verständnis und Geschwindigkeit im Vergleich zum Lesen in VR ohne Assistenz zu verbessern.

Im Rahmen einer Lesestudie in VR mit Messwiederholung (*Within-Subjects-Design*) wurden beide Funktionen im Bezug auf die Leseparameter (Lesewiederholungen, Lesedauern), inhaltlichen Abruf (*Recall*) sowie auf Funktions-spezifische Merkmale evaluiert.

Es wurden hierfür folgende Forschungshypothesen aufgestellt und getestet:

H1(a): Lesen mit Assistenz-Funktion *Scanpath* führt zu einem signifikant besseren Abruf (*Recall*) der Texte als ohne Leseassistent.

H1(b): Lesen mit Assistenz-Funktion *Peripheral Blur* führt zu einem signifikant besseren Abruf (*Recall*) der Texte als ohne Leseassistent.

H2(a): Lesen mit Assistenz-Funktion *Scanpath* führt zu

einer **längeren** Lesedauer als das Lesen ohne Leseassistent.

H2(b): Lesen mit Assistenz-Funktion *Peripheral Blur* führt zu einer **kürzeren** Lesedauer als das Lesen ohne Leseassistent.

H3(a): Lesen mit Assistenz-Funktion *Scanpath* führt zu **mehr** Lesewiederholungen als das Lesen ohne Leseassistent.

H3(b): Lesen mit Assistenz-Funktion *Peripheral Blur* führt zu **weniger** Lesewiederholungen als das Lesen ohne Leseassistent.

2 VERWANDTE ARBEITEN

Im Folgenden werden Arbeiten genannt, welche sich auf die Leseassistenten in AR beziehen, da die Kernidee dieser Arbeit darin liegt, einen Leseassistenten zu entwickeln, der beim Lesen von Texten auf Papier oder Bildschirmen virtuelle Hilfestellungen anbietet. Die VR ist aufgrund technischer Beschränkungen seitens der AR lediglich das Mittel zur erstmaligen Umsetzung dieser Idee und ist daher für die verwandte Literatur nicht von Interesse.

Leseassistenten durch eyetracking-basierte Events. Hyrskykari [30] sowie Guo und Cheng [20] adressieren mit ihren Systemen die Probleme, die beim Lesen fremdsprachiger Texte auftauchen und bieten basierend auf einer Analyse der eyetracking-Daten Übersetzungen an, wenn sich daraus eine entsprechende Intention der Leser:innen ableiten lässt. *The Reading Assistant*, der von Sibert et al. [54] für Kinder mit Leseproblemen konzipiert wurde, erzielt eine Wiedererkennung von Wörtern durch synthetische Aussprache dieser, wenn das Scanmuster der Augen Verständnisschwierigkeiten impliziert.

Mit der rasanten Entwicklung im Bereich der KI innerhalb der letzten Jahre bietet es sich an, den Ansatz der soeben vorgestellten Systeme mit einem state-of-the-art Sprachmodell zu kombinieren, welches Wortbedeutungen und Übersetzungen nicht vom Rest des Textes isoliert anbietet, sondern in den Kontext des Textinhalts setzen kann. So erschien erst dieses Jahr eine Arbeit zu einem Prototypen eines headmounted AR-Leseassistenten mit gleichzeitiger Einbindung von KI: Der *Smart AI Reading Assistant* (SARA) [58]. Dieser gibt mithilfe von GPT-4 kontextuelle Wortdefinitionen und Übersetzungen anhand von Eyetracking-Metriken, die entweder auf ein Verständnisproblem auf Wort- oder Absatzebene hindeuten.

Leseassistenten für Sehingeschränkte. Im Bezug auf die perzeptive Verarbeitungsebene des Lesens gibt es AR-gestützte Hilfen zur Texterkennung für Menschen mit eingeschränktem Sehvermögen, sowohl für einfache Schildhinweise zur Navigation in unbekannten Umgebungen [28, 44], als auch für Fließtexte [57].

AR zu schulischen Illustrationszwecken. Im Bildungskontext wurden bisher mobile AR-Books und -Apps in den Schulalltag integriert, welche die Lernmaterialien mit 3D-Illustrationen, Animationen und anderen Medieninhalten anreichern und so den Schüler:innen einen stimulierenden Anreiz bieten, sich mit dem Lerninhalt eingängiger zu beschäftigen [29, 37, 63], oder ein Interesse am Lesen selbst zu entwickeln [51].

AR-basierte Interventionen für Leser:innen mit Neurodiversität. Auch für Personen mit Neurodiversität gibt es AR-Anwendungen, die den Leseprozess interaktiver gestalten sollen. So wurden bspw. Anwendungen bei Kindern bzw. Erwachsenen mit Dyslexie getestet, mit denen sie Textparameter (Schriftart/-größe, Zeilenabstände etc.) frei manipulieren und optional auf Text-to-Speech ausweichen können [21, 32, 46]. Aus dem Bestreben heraus ein Programm zur Lesekompetenz zu initiieren, das auf Kinder mit

ADHS zugeschnitten ist, wurde seitens der EU-Kommission das AHA-System (*ADHD Augmented*) [59] gefördert. Dieses integriert webbasiertes AR zum multisensorischen Lernen der Aussprache von Wörtern anhand von animierten 3D-Icons zusammen mit einer Online-Plattform zur Überwachung des Lernfortschritts seitens Pädagogen. Die dreidimensionalen AR-Objekte werden hierbei abhängig von ausgedruckten Markern platziert, die von einer Kamera erkannt werden.

AR zur Unterstützung von deep reading. Auch das vertiefte Lesen (*deep reading*) im Gegensatz zum oberflächlichen Lesen (*shallow reading*), welches bei Schüler:innen aufgrund des sprunghaften Konsumierens von Online-Content auftritt, kann mit AR gefördert werden, wie Li et al. [39] in ihrer Studie zeigen. Hierbei wurden in AR (und auch im VR-Modus) sukzessive Zeilen eines Overlays entfernt, welches den Text verdeckt. Damit wurde sichergestellt, dass die Zeilen von den Schüler:innen in der vorgegebenen Reihenfolge gelesen wurden.

In der Forschung besteht ein großer Mangel hinsichtlich AR-gestützter Leseassistenten, die den Text nicht nur um Medieninhalte ergänzt, sondern tatsächlich auch den Leseprozess durch direkte Textaugmentierungen unterstützt (wie z.B. in den Dyslexie-spezifischen Anwendungen). Das Problem bezieht sich auch auf solche Systeme, welche explizite statt implizite Hilfestellungen geben, wie z.B. Übersetzung und Zusammenfassung, was zwar für sich genommen beim Textverständnis hilft, aber gleichzeitig die Behandlung der Problemursache beim Lesen untergräbt. Ein weiteres Problem stellt die Unterrepräsentiertheit an erwachsenen Leser:innen als Zielgruppe dar, da sich die meisten AR-gestützten Interventionen meistens auf leseresistente Schüler:innen oder Kinder mit Neurodiversität beschränken und dabei viel mehr auf die Lesesakquisition als -Unterstützung abgezielt wird. Aufgrund dieses Mangels ist es das Anliegen dieser wissenschaftlichen Arbeit zur Forschung dieses so wenig behandelten Gebiets beizutragen.

3 TECHNISCHE IMPLEMENTATION DER LESEASSISTENZ-FUNKTIONEN

3.1 Technische Umsetzung generell

Als Head Mounted Display (HMD) wurde die Pico Neo 3 Pro Eye gewählt. Dieses verfügt über einen Eyetracker von Tobii¹, welcher eine Abtastrate von bis zu 90 Hz besitzt, auf die das Headset eingestellt wurde. Da die Leseaufgabe nur eine minimale Interaktion mit UI erforderlich machte, wurde für die Studie nur der rechte Controller verwendet (die Rechtshändigkeit stellte für die Proband:innen kein Problem dar).

Die Anwendung wurde am PC in einer einzigen Unity-Szene (Version 2020.3.45f1) entwickelt. Als Skriptsprache wurde entsprechend C# gewählt. Für UI-Interaktionen wurde das XR Interaction Toolkit² installiert. Die UI sprach anfangs auf keine Controller-Eingaben an, weshalb Worldspace Canvasses aus der Szene des XRI Example Projekts übernommen und an das eigene Projekt angepasst wurden. Für die Einbindung des Eyetracking wurde die Tobii XR SDK in das Projekt geladen. Die Szene des zugehörigen Beispielprojekts *Getting Started*³ wurde zur Projektszene umgebaut, da auch das Eyetracking bei Interaktionen zunächst nicht registriert wurde. Weiterhin wurde ein Pico Link Kabel benutzt, zusammen mit der zugehörigen App für die Preview über Open-XR und Steam-VR. Die App SideQuest und ein Datenkabel ermöglichten die Installation der apk-Datei auf die

Brille. Zur Versionskontrolle wurde ein Repository in Git aufgesetzt.

Die virtuelle Umgebung sollte keine Ablenkungen bieten, weshalb die default Skybox durch ein neutrales, einfarbiges (#D2EAFB) Material ersetzt wurde. Bis auf einen Boden mit einfacher Gittertextur (aus *Getting Started* Projekt übernommen) und das Panel, welches den zu lesenden Text enthalten sollte (s. 4.1), war die Szene leer.

In jeder Kondition, was auch die Kondition *Baseline* ohne Leseassistenten-Funktion mit einschließt (s. 4.2), wurde die Lesezeit in Sekunden gemessen. Dies wurde über ein Zeitinkrement während der Raycast-Collision eines Gazeray (berechnet aus Eyetracking-Daten über Tobii XR SDK) mit einer Plane in der Größe und auf z-Höhe des Texthintergrunds realisiert. Mit jedem Seitenwechsel wurden die Lesedauer zusammen mit der Anzahl der Lesewiederholungen für die bisher gelesene Seite gespeichert. Als Lesewiederholung galt ein Zurück- oder Weiterblättern auf eine zuvor geöffnete Seite (Ausnahme bildet Scanpath, s. 3.2).

Jede Kondition mit Leseassistenten-Funktion bestand aus einem Testlauf und der tatsächlichen Runde (s. 4.3, Abbildung 8). Während der Testrunden wurden keine Leseparameter erhoben (Ausnahme bildet Peripheral Blur, s. Abschnitt 7), da sich dort mit der jeweiligen Leseassistenten-Funktion vertraut gemacht werden sollte. Wurde die richtige Runde beendet, wurden die Leseparameter für die zuletzt gelesene Seite gespeichert.

Da die Augen nicht innerhalb der Anwendung kalibriert werden konnten, musste dies in einer anderen App vor Start der Anwendung geschehen. Aus diesem Grund war es nicht vorgesehen, dass die Proband:innen das Headset während des VR-Teils absetzen. Da jedoch gleichzeitig Fragebögen zwischen den Konditionen erhoben werden sollten, mussten diese in die Anwendung integriert werden. Dies wurde über das 3D WebView Plugin⁴ realisiert, sodass in der Anwendung ein Browserfenster mit der Umfrage geladen werden konnte. Dies erforderte wiederum eine Internetverbindung des VR-Headsets.

Da Laien die Begriffe „Scanpath“ oder „Peripheral“ vermutlich nicht einzuordnen wissen, wurden Aliasse für beide Leseassistenten-Funktionen eingeführt. Innerhalb der Benutzerstudie wurde daher die Scanpath-Funktion „Einsehbarer Blickverlauf“ und Peripheral Blur „Zeilenfokus“ genannt, während sie hier weiterhin mit ihren eigentlichen Namen referenziert werden.

3.2 Scanpath (Einsehbarer Blickverlauf)

Die Idee für die erste Leseassistenten-Funktion beinhaltete eine feedbackbasierte Methode, die ein Bewusstsein für vielleicht unbewusste Verhaltensweisen beim Lesen schaffen und so zu einer Anpassung animieren sollte. Da sich Leseverhalten auf der obersten Ebene in Augenbewegungen äußert, ging es darum, für das Feedback eine Darstellungsform zu finden, welche die Historie der Augenbewegungen, d.h. den Pfad der Augen durch den Text abbildet. Dieser Idee kommt der Scanpath am nächsten, welcher in der Forschung verwendet wird, um Eyetracking-Daten einer Person visuell zu aggregieren. Auf diese Weise wird z.B. nachzuvollzogen, wie Personen mit Webseiten interagieren, um sie hinsichtlich Design und Nützlichkeit zu verbessern, wie es neben der Werbung [62] auch schon in der Mensch-Computer-Interaktion [13, 14] zum Einsatz kam. Aber auch in der Lesepsychologie [7, 12] hat diese Metrik Einzug gefunden, was ihre Relevanz auch in der Lese-Domäne unterstreicht.

Dieser ist aber für gewöhnlich nur aus Sicht der Forschenden von Interesse und auch nur dann, um die Daten einzelner Personen zu visualisieren, da sich aus Scanpaths mehrerer Proband:innen aufgrund ihrer Individualität und dadurch Variabilität schwer Erkenntnisse über eine gemeinsame Tendenz im Blickverhalten gewinnen lassen [14, 15]. Dieses Problem besteht jedoch nicht, wenn nur der eigene Scanpath von Interesse ist, was erneut für diese Form des

¹<https://www.tobii.com/products/integration/xr-headsets/device-integrations/pico-neo-3-pro-eye>

²<https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@3.0/manual/index.html>

³<https://developer.tobii.com/xr/develop/xr-sdk/getting-started/pico-neo-3-pro-eye/>

⁴<https://assetstore.unity.com/packages/tools/gui/3d-webview-for-android-and-ios-web-browser-135383>

Feedbacks spricht.

Nur wenige Studien [9, 61] beinhalten eine analytische Auseinandersetzung mit dem eigenen Scanpath. Doch auch dort ist der Hintergedanke zu testen, ob Proband:innen ihren eigenen Scanpath von denen anderer Personen unterscheiden können, womit der Selbstzweck der Analyse ausbleibt. Ebenso wurden hierbei keine Textinformationen, sondern Bilder betrachtet.

Ein Scanpath setzt sich in seiner Darstellung meistens aus Kreisen und Linien zusammen [5, 19]. Die Kreise bilden hierbei die Fixationen ab, zum einen in ihrer Position und zum anderen in ihrer Dauer über ihre Durchmesser; je größer der Durchmesser, desto größer die Dauer dieser Fixation. Oft sind die Kreise nummeriert, um die Reihenfolge der Fixationen anzugeben. Zwischen den Kreisen spannen sich Linien, welche die Sakkaden darstellen, d.h. die sprunghaften Bewegungen, die zwischen den Fixationen ausgeführt werden [49].

Für die gleichnamige Leseassistenten-Funktion sollte der Scanpath noch zugänglicher gestaltet werden, da er samt seiner Einfarbigkeit und Nummerierung noch zu abstrakt erscheinen könnte. Analog zur Größe kann so auch die Farbe die Fixationsdauer ähnlich zu Heatmaps abbilden. Diese greifen auf das geläufige Farbkontinuum zwischen Grün und Rot zurück, wobei grün mit einer geringeren und rot mit einer höheren Intensität assoziiert wird. Auch in der Eyetracking-Domäne kommen Heatmaps (bzw. Attention Maps) zum Einsatz, um Interessensregionen im Bezug auf die Verweildauer der Augen ähnlich zu einem Wärmebild hervorzuheben [5]. Daher wurde sich dazu entschieden, das Farbkontinuum der Heatmap auf die Fixationsdauern anzuwenden, wobei eine längere Fixation auf einen größeren Kreis abgebildet wird, dessen Farbe näher an Rot ist und eine kürzere Fixation auf einen kleineren Kreis, dessen Farbe näher an Grün ist. Einen ähnlichen Ansatz verfolgte Lankford [38], bei dem die Farbtransitionen über die Fixationen hinweg zeigten, wie schnell die Personen zwischen den Fixationen ihre Augen bewegten.

Da Laien die Nummerierung der Sakkaden unintuitiv erscheinen könnte, wurde statt dieser auf Pfeilspitzen zurückgegriffen, da diese anders als Zahlen auch eine Richtung und somit ebenfalls eine Reihenfolge implizieren. Werden alle Sakkaden als Pfeile dargestellt, führt dies zu vielen Überlagerungen und somit zu visuellem clutter, weshalb Sakkaden in vielen Studien von einer weiterführenden Analyse ausgeschlossen werden [35]. Da rückwärtsgerichtete Sakkaden (auch: Regressionen) innerhalb oder zwischen Zeilen ein Indiz für Verständnisschwierigkeiten an einer Textstelle sein können [49, 50], waren für eine Feedback-basierte Leseassistenten, welche Auffälligkeiten während des Lesens aufzeigt, ebendiese Rücksprünge von Interesse. Aus diesem Grund wurde der andere Sakkadentyp (vorwärtsgerichtete Sprünge der Augen) aus der Visualisierung ausgeschlossen. Die resultierende Darstellungsform der umgesetzten Leseassistenten-Funktion ist in Abbildung 3 zu sehen.

Durch die gefundene Darstellungsform könnte der Scanpath zu einem Benchmark des eigenen Leseverhaltens zweckentfremdet werden. So könnten Regionen aus rötlichen/großen Kreisen und evt. Rückpfeilen ein Verständnisproblem an der entsprechenden Textstelle signalisieren. Denn so erhöht sich mit der Anzahl an Regressionen auch die Dauer der Fixationen, wenn Verständnisschwierigkeiten beim Lesen auftreten [49]. Auf Basis dieser Visualisierung könnte die Entscheidung getroffen werden, die kritische Textstelle zu wiederholen oder die Relevanz der Textstelle zu hinterfragen, da für diese womöglich schon genug Lesezeit aufgewendet wurde. Längere bzw. häufigere Fixationen könnten jedoch auch für ein erhöhtes Interesse an dieser Textstelle sprechen, analog zu Attention Maps [5]. Daher kann der Scanpath auch als eine passive Markierung persönlich relevanter Textstellen gesehen werden, d.h. schon durch das Lesen selbst werden Markierungen z.B. durch rötliche Kreise und Rücksprünge vorgenommen. Ähnlich zu aktiven Textmarkierungen, die beim Lesen so häufig gemacht werden, könnten diese nach dem

Lesen schnell und einfach wiedergefunden und nachvollzogen werden. Auch das Ausbleiben solcher Regionen könnte bedeuten, dass eine Textstelle zu schnell überflogen und daher nicht bewusst aufgenommen wurde.

Da es sich, wie gesagt, um eine explorative Form der Leseassistenten handelt, musste sichergestellt werden, dass die Proband:innen auf diese mindestens einmal zugegriffen. War der Zugriff komplett freiwillig, bestand das Risiko, dass die Proband:innen den Scanpath kein einziges Mal nutzen würden, wodurch keine Vergleichbarkeit der Daten gewährleistet werden könnte. Als Kompromisslösung des unerzwungenen Feedbacks wurde der Scanpath immer dann sichtbar, wenn der Button für das Blättern auf die nächste Seite (s. Abbildung 7, rechts) ausgewählt wurde. Da der Button für das Weiterblättern somit ein Placebo-Button war und die Seitenwechsel-UI inaktiv gesetzt wurde, wurden zusammen mit dem Scanpath zwei Hilfs-Buttons eingeblendet (s. Abbildung 3). Der Button links des Textpanels realisierte das direkte Nachlesen der gleichen Seite (mit Ausblenden des Scanpath), während der rechte Button den tatsächlichen Wechsel zur nächsten Seite umsetzte. Mit jeder Lesewiederholung wurden im Hintergrund die Parameter für einen neuen Scanpath aufgezeichnet.



Abbildung 3: Exemplarischer Scanpath aus gleichnamiger Leseassistenten-Funktion. Kreise bilden in Größe und Farbe Fixationsdauern ab, Rückpfeile Regressionen. Über den linken Button kann die Seite direkt erneut und ohne Scanpath gelesen werden, während im Hintergrund neue Scanpath-Parameter erhoben werden. Erst mit Auswählen des rechten Buttons wird direkt zur nächsten Seite geblättert.

3.2.1 Richtwerte

Um einen Algorithmus zu entwickeln, welcher zwischen Fixationen und Sakkaden unterscheidet, mussten für beides geeignete Schwellenwerte gefunden werden. Laut Rayner [49] liegt die Dauer einer Fixation zwischen 200-300 Millisekunden, bei manchen Tätigkeiten (Musik hören, Tippen) reicht sie sogar bis zu 400 Millisekunden. Beim lautlosen Lesen liegt die Fixationsdauer um die 225 Millisekunden. Es gibt Befunde dafür, dass die Minimaldauer einer Fixation für Leseverständnis 200 Millisekunden beträgt [55], aber so kommen während des Lesens auch Fixationen zustande, die bei 50 Millisekunden liegen [49]. Um die gesamte Bandbreite der genannten Werte samt einem Toleranzbereich nach oben und unten abzudecken, wurde sich als Minimalwert für 50 und als Maximalwert einer Fixation für 400 Millisekunden entschieden. Demnach würden 50 Millisekunden im Scanpath als kleinstmöglicher Kreis, der in seinem Farbwert am nächsten an Grün liegt, abgebildet. Umgekehrt würde eine Fixation, welche 400 Millisekunden beträgt, auf den größtmöglichen Kreis, der in seinem Farbwert am nächsten an Rot liegt, abgebildet. Da die Augen während einer Sakkade innerhalb kürzester Zeit große Strecken zurücklegen, wurde die Geschwindigkeit als Metrik betrachtet, um Sakkaden von Fixationen trennen zu können. An dieser Stelle

sei zudem nochmal gesagt, dass die Augen auch während Fixationen nicht stillstehen, da auch hier ein minimaler Tremor, ein sogenannter Nystagmus besteht [49]. Als Maßeinheit für die Geschwindigkeit wurden die zurückgelegten Winkel pro Sekunde betrachtet, so wie es bei den geschwindigkeitsbasierten (*velocity based*) Verfahren der Klassifizierung von Augenbewegungen der Fall ist [52]. In Studien, in denen Eyetracking innerhalb der VR zum Einsatz kommt [26, 27], werden Sakkaden über den Schwellenwert von 75°/s bestimmt. In Anlehnung daran wurden Werte zwischen 60°/s und 120°/s getestet. Schließlich fiel die Wahl auf 100°/s, da dies nach eigenem Ermessen das beste Ergebnis generierte. Alle Werte unter diesem Schwellenwert würden als Fixation und alle darüber als Sakkade klassifiziert werden.

3.2.2 Implementationsdetails

Da ein Scanpath normalerweise im Millisekundenbereich operiert, musste direkt auf den Datenstream des Eyetrackers der Pico zugegriffen werden, um die tatsächliche Abtastrate von 90 Hz nutzen zu können. Hierfür wurde Tim Rolfs Data Capture Package⁵ in das Unity-Projekt eingebunden und in diesem speziell auf das Skript InputCaptureSystem als Schnittstelle zugegriffen (s. Abbildung 4(a)). In diesem werden die kombinierten Ursprungs- und Richtungsvektoren ausgelesen und an das Scanpath-Skript weitergereicht, aber nur dann, wenn beide Augen offen waren (kontrolliert auf Blinzel-Fall) und der Button zum Wechsel auf die nächste Seite noch nicht ausgewählt wurde, da währenddessen der Scanpath angezeigt wird.

Da im Folgenden näher auf die Implementation der Parameterberechnung der Scanpath-Funktion eingegangen wird, wird diese anhand von Pseudo-Code 1 erläutert. Aus Platzgründen verzichtet dieser auf Fallunterscheidungen, welche Fälle des ersten Schnittpunktes oder der ersten Klassifikation behandeln.

Mit jedem Frame-Update des Scanpath-Skripts wurden mehrere Ursprungs- und Richtungsvektoren über InputCaptureSystem ausgelesen und im Scanpath-Skript zu Gazerays kombiniert. Um die Update()-Funktion entsprechend der Abtastrate des Eyetrackers von 90Hz zu takten, wurden die konstruierten Gazerays einer Liste hinzugefügt, sodass sie innerhalb eines Update-Zyklus versetzt abgearbeitet werden konnten (s. Abbildung 4(b)). Für eine synchrone Zeitmessung wurde mit jeder Listeniteration ein Timer (was einem Ray, bzw. Sample entspricht) um 1/90s erhöht. Um in den Millisekundenbereich zu konvertieren, wurde der Zeitschritt mit dem Faktor 1000 multipliziert (Pseudocode: Z. 4). Demnach wurde alle 11ms ein Raycast mit dem aktuellen Gazeray durchgeführt (s. Abbildung 4(c)). Wurde hierbei eine Kollision mit einer Plane in der Größe und auf z-Höhe des Texthintergrunds (s. Abbildung 4(d),(e)) registriert, musste der resultierende Schnittpunkt als Teil einer Fixation oder Sakkade klassifiziert werden. Hierfür wurde das geschwindigkeitsbasierte Verfahren [52] in Anlehnung an den Tobii I-VT Classifier [43] umgesetzt: Ab dem zweiten Schnittpunkt mit der Plane wird paarweise der Winkel zwischen dem zugehörigen aktuellen und vorherigen Ray gebildet (Pseudocode: Z. 10, s. Abbildung 4 (f)).

An dieser Stelle muss darauf hingewiesen werden, dass in der Berechnung der Ray-Vektoren für die Winkelbildung ein Fehler im Code vorliegt, der im Pseudocode (Z. 6,7,10) bereits korrigiert wurde. Dieser äußert sich darin, dass der Vektor für den vorherigen Ray aus dem vorherigen Schnittpunkt mit der Textplane und dem Ursprung des aktuellen *statt* des vorherigen Ray aus der Liste gebildet wurde, wohingegen der Vektor für den aktuellen Ray richtigerweise aus dem aktuellen Schnittpunkt und dem Ursprung des aktuellen Rays berechnet wurde. Allerdings kann dieser Fehler nur minimal zu Verzerrungen im Scanpath beigetragen haben, da sich aufgrund der stationären VR in Sitzposition (auf einem Stuhl ohne Rollen) der Ursprung zwischen vorherigem und aktuellem Gazeray nur minimal verändern konnte (s. 4.2). Da das Textpanel groß genug war, um das gesamte Sichtfeld auszufüllen, waren die Proband:innen kaum

auf Kopfrotationen angewiesen. Bei entsprechend kleinen Kopfrotationen kann sich der Ursprungsvektor über die Rays hinweg kaum verändert haben. Nur Translationen hätten für eine erhebliche Lücke gesorgt, was sich zwar nicht komplett ausschließen ließ, aber durch das Studiendesign auf einem Minimum gehalten wurde.

Da somit lediglich der Winkel für 1/90s berechnet wurde, musste der Winkel für eine ganze Sekunde mit 90 multipliziert werden (Pseudocode: Z. 11). Liegt der resultierende Wert unter der festgelegten Schwelle (100°/s), wird der aktuelle Schnittpunkt als Teil einer Fixation klassifiziert und einem Fixations-Cluster zugeordnet (Pseudocode: Z. 12-14, s. Abbildung 4(g)). Andernfalls wird das Sample als Sakkade klassifiziert (Pseudocode: Z. 16).

Die aktuelle Klassifikation wird mit der vorherigen verglichen. Sind beide nicht gleich und wurde der vorherige Schnittpunkt als Fixation klassifiziert (Pseudocode: Z. 19 und 20), muss der aktuelle Schnittpunkt das Sakkaden-Kriterium (Pseudocode: Z. 15-16) erfüllt haben. Somit entspricht der bisherige Timer dem Zeitstempel für das erste Sakkadensample und der Zeitstempel für das letzte Fixationssample dem Timer davor (Pseudocode: Z. 21-22). Es wird die Dauer der (durch die Sakkade unterbrochene) Fixation berechnet und für die aktuelle Seite gespeichert (Pseudocode: Z. 23-26): Für den Startzeitpunkt der Fixation wird das Mittel aus dem Zeitstempel für das letzte Sample der vorherigen Sakkade und dem Zeitstempel für das erste Sample der bisherigen Fixation berechnet. Für den Endzeitpunkt wird das Mittel aus dem Zeitstempel für das letzte Sample der bisherigen Fixation und dem Zeitstempel für das erste Sample der neuen Sakkade gebildet. Die Dauer der beendeten Fixation ergibt sich letztendlich aus der Differenz aus End- und Startzeitpunkt. Anschließend wird über alle Schnittpunkte aus dem bisherigen Fixationscluster gemittelt und das Ergebnis für die aktuelle Seite gespeichert (Pseudocode: Z. 27-30, s. Abbildung 4(h)). Dies entspricht der aggregierten Abbildung einer Fixationsposition im Scanpath. Für ein neues Fixationscluster wird die zugehörige Liste geleert (Pseudocode: Z. 29).

Ab der zweiten aggregierten Fixationsposition des Scanpath wird auf das Regressionskriterium geprüft, da nur die Rücksprünge im Text visualisiert werden. Der Toleranzbereich für die Richtung einer Regression wurde auf alle Vektoren festgelegt, die nach links und/oder oben zeigen. Demnach wurde zunächst der Sakkadenvektor aus aktuellem und vorherigem Positionsaggregat der Fixationen berechnet (Pseudocode: Z. 31). Bildete dieser zusammen mit dem up-Vektor einen Winkel kleiner oder gleich 90 und entsprach die x-Komponente des Vektors einem negativen Wert (Pseudocode: Z. 33), galt die Sakkade als Regression. Solange sich die Augen weiterhin in einer rückwärtsgerichteten Bewegung befanden, wurden die Positionsaggregate einem entsprechenden Pfad in Form einer Liste hinzugefügt (Pseudocode: Z. 34-36). Wurde diese durch eine vorwärtsgerichtete Bewegung unterbrochen, wurde für die nächste Regression ein neuer Pfad für die aktuelle Seite angelegt (Pseudocode: Z. 39). Zusätzlich wurden für die Pfeilspitzen die Orientierungen gespeichert (Pseudocode: Z. 37).

Entspricht die Klassifikation des aktuellen Punktsamples nicht der vorherigen (Pseudocode: Z. 19), aber wurde das vorherige Punktsample nicht als Fixation klassifiziert (Pseudocode: Z. 40), so muss eine bisherige Sakkade durch eine Fixation unterbrochen worden sein. In diesem Fall werden nur der Startzeitpunkt der neuen Fixation und Endzeitpunkt der bisherigen Sakkade gespeichert (Pseudocode: Z. 41-42). Entspricht die aktuelle Klassifikation der vorherigen, wird nur der aktuelle Zeitstempel hinzugefügt (Pseudocode: Z. 44).

Da die eigentliche Visualisierung erst beim Auswählen des Buttons zum Blättern auf die nächste Seite stattfindet, wurde diese in das Skript zum Seitenwechsel ausgelagert: Kreise werden hierbei entsprechend der aggregierten Fixationspositionen gesetzt und gemäß der korrespondierenden Fixationsdauern in Größe und Farbe interpoliert mit 50ms als 0% und 400ms als 100% (s. Abbildung 4(i)). Dies galt allerdings nur für Fixationen, die mindestens 50ms dauerten.

⁵<https://git.informatik.uni-hamburg.de/hci/datacapture>

Ab dem Maximum von 400ms wurden die Werte abgeschnitten. Die Kreise wurden in Gimp mit einer Größe von 512x512 und einer Outline von 5 Pixeln erstellt und in Unity um den Faktor 0,03 skaliert (= Ausgangsgröße). Ebenfalls wurde ihre Deckkraft auf 20% heruntergesetzt, sodass der darunterliegende Text noch ausgemacht werden konnte.

Im Bezug auf die Farbe sollten zusätzlich Menschen mit Farbfeldsichtigkeit berücksichtigt werden. Somit musste ein alternatives Farbkontinuum zu grün bis rot gefunden werden. Dies wurde mit dem barrierefreien Farb-Tool von Adobe⁶ umgesetzt, welches verschiedene Arten von Farbfeldsichtigkeit simuliert. So konnten Farbgradienten gefunden werden, die auch im Falle einer Rot-/Grün- oder Blaublindheit als verschiedene Farbtemperaturen wahrgenommen werden können. Ein Vorliegen dieser Einschränkung wurde bereits in der Testrunde der Scanpath-Funktion abgefragt und der Gradient entsprechend an den Fall angepasst.

Zusammen mit den Kreisen werden mithilfe eines Linerenderers die Regressionspfade als weiße Linienzüge dargestellt, unterbrochen von entsprechend rotierten Pfeilspitzen ab dem zweiten Fixationspunkt des Pfades (s. Abbildung 4(j)).

Solange der Scanpath für eine bestimmte Seite sichtbar ist, wird die Zeit gemessen. Erst wenn über den Button, der zusammen mit dem Scanpath eingeblendet wird, zur nächsten Seite gewechselt wird, werden Scanpath-Dauer zusammen mit den Leseparametern und anderen Teilnehmerdaten für das Lesen der bisherigen Seite gespeichert. Dies gilt jedoch nur für den Fall des Vorwärtsblätterns, denn während des Zurückblätterns wird kein Scanpath generiert und die Teilnehmerdaten für das Lesen der bisherigen Seite werden direkt geloggt.

3.3 Peripheral Blur (Zeilenfokus)

Die Textmenge kann für sich genommen eine Frustrationsquelle darstellen, da sie auf die Leser:innen anfangs wie eine Reizüberflutung wirken kann. Informationen sind hierbei auf dichtem Raum gekapstelt, die anders als Bildinformation auf komplexeren, semantischen Verarbeitungsebenen entschlüsselt werden müssen; noch dazu ist dieser Prozess an die Abarbeitungsreihenfolge gebunden, die der Text mit seiner Struktur vorgibt. Zusätzlich ist die Schriftgröße vieler wissenschaftlicher Arbeiten oft sehr klein gewählt, was für noch mehr Information pro Seite führt. Dieser erste Gesamteindruck kann sich negativ auf die Erwartung, Motivation und Aufmerksamkeit ausüben, was den weiteren Lesefluss beeinträchtigt. Damit geht einher, dass umliegende Zeilen die eigenen Augen ablenken können (wie auch im Späteren aus Erfahrungsberichten hervorgeht); ein ähnliches Problem besteht im Verrutschen in der Zeile.

Somit sorgt viel Text für Orientierungslosigkeit und bietet ein hohes Ablenkungspotenzial, was der visuellen Aufmerksamkeit abträglich ist. Eine Lösung für dieses Problem bietet die *Diminished Reality* (DR). Abgeleitet aus Steve Mann's Mediated Reality Framework [41] werden mit dieser Störinformationen aus der realen Umgebung entfernt. Umgesetzt in der AR muss mithilfe von virtuellen Overlays eine Verdeckung unerwünschter Objekte durch ihren Hintergrund erzielt werden, um diese somit aus der Wahrnehmung der Nutzer:innen zu filtern. Da sich die technische Umsetzung hierbei als schwierig erweist (so wie bereits in Abschnitt 1 angesprochen), bietet es sich an DR in VR zu simulieren, so wie in der Studie von Cheng et al. [8]. Ausgehend von ihrem Verständnis des DR-Konzepts wählten sie für die VR-Szene Effekte, welche ablenkende Objekte nicht gänzlich entfernten, sondern in ihrer Salienz durch verschiedene Modifikationen reduzierten. Demnach kann DR auch in einer schwächeren Ausprägung angewandt werden. Unter den Effekten befand sich u.a. der Blur, welcher für diese Studie übernommen wurde, da er eine Analogie zum peripheren Sehen ist, bei dem mit zunehmendem Abstand zur Fovea die räumliche Auflösung abnimmt. Darüber hinaus wird mit dem Effekt keine

```

1 rayListToUpdate ← rayList mit neuen rays aus Daten von
  InputCaptureSystem;
2 rayList.Clear();

3 for jeder ray aus rayListToUpdate do
4   fixationTimer += (1 / 90) * 1000;
5   if Raycast(ray): Collision mit Text Plane, out: hit.point then
6     previousRay ← currentRay;
7     currentRay ← ray;
8     previousPosition ← currentPosition;
9     currentPosition ← hit.point;
10    angle ← Angle(previousPosition - previousRay.origin,
11                  currentRay - currentRay.origin);
12    degreesPerSecond ← angle * 90;
13    if degreesPerSecond < threshold then
14      isFixation = true;
15      positionList.Add(currentPosition);
16    else
17      isFixation ← false;
18    previousClass = currentClass;
19    currentClass = isFixation;
20    if previousClass != currentClass then
21      if previousClassification == true then
22        T_ersteSakkade ← fixationTimer;
23        T_letzteFixation ← vorheriger fixationTimer;
24        start ← mean(T_letzteSakkade, T_ersteFixation);
25        end ← mean(T_letzteFixation, T_ersteSakkade);
26        fixationDuration ← end - start;
27        füge fixationDuration aktueller Seite hinzu;
28        previousPosition_AVG ← currentPosition_AVG;
29        currentPosition_AVG ← avg(positionList);
30        positionList.Clear();
31        füge gemittelte Position aktueller Seite hinzu;
32        saccadeVector ← currentPosition_AVG -
33          previousPosition_AVG;
34        angle ← Angle(Vec3.up, saccadeVector);
35        if angle <= 90 && currentPosition_AVG.x <
36          previousPosition_AVG.x then
37          if arrowPath.Count == 0 then
38            arrowPath.Add(prevAVG);
39            arrowPath.Add(curAVG);
40            tipRotations.Add(angle);
41          else
42            erstelle neuen arrowPath und Liste für
43              tipRotations, füge beides aktueller Seite
44              hinzu;
45        else
46          T_ersteFixation ← fixationTimer;
47          T_letzteSakkade ← vorheriger fixationTimer;
48        else
49          füge fixationTimer hinzu;

```

Pseudo-Code 1: Update() des Scanpath-Skripts.

⁶<https://color.adobe.com/de/create/color-accessibility>

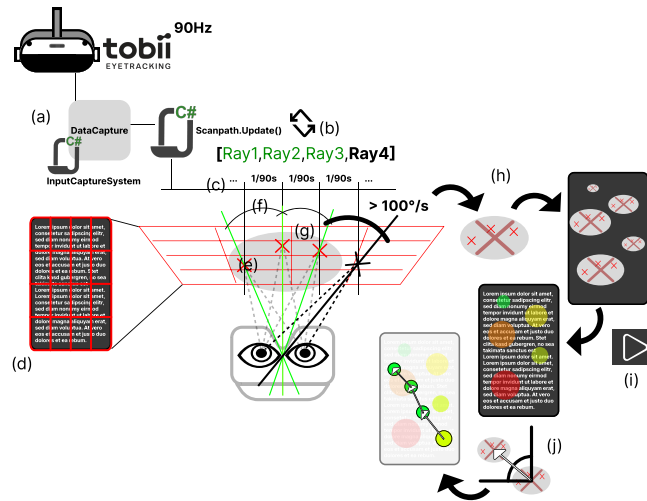


Abbildung 4: Leseassistenten-Funktion Scanpath und ihre Komponenten. (a): Das DataCapture-Package mit der modifizierten Schnittstelle InputCaptureSystem. (b): Pro Update()-Abruf wird über eine Liste aus Gazeays iteriert. (c) Taktung der Update()-Funktion nach der Abtastrate von 90Hz. (d) Textplane mit Größe und z-Höhe des Texthintergrunds. (e) Raycast: Schnittpunkt aus Kollision des Gazeray mit der Textplane. (f) Winkel zwischen Gazerays (multipliziert mit 90 für °/s). (g) Schnittpunkte (rot) werden Fixationscluster (graue Ellipse) hinzugefügt. (h) Sakkaden-Kriterium erfüllt (schwarzer Schnittpunkt): Fixationscluster wird zu einer Position gemittelt (großes „X“), Fixationsdauer wird berechnet. (i) Visualisierung durch Auswählen des Buttons zum Wechsel auf nächste Seite. (j) Visualisierung der Regressionen.

Änderung an Farbe, Transparenz oder Größe des Textes oder seines Hintergrunds vorgenommen, was an sich eine Ablenkung darstellen könnte.

Zugleich ist der Effekt durch seine Anpassbarkeit an das visuelle System sehr subtil. So wurde beispielsweise in der Eyetracking-Studie von Hata et al. [24] gezeigt, dass ein Blurfilter den Blick auf die einzigen scharfen Regionen in Bildern lenken kann, ohne dass die Personen merken, dass eine Manipulation ihrer visuellen Aufmerksamkeit stattfindet. Im Bezug auf Bilder mit Text argumentieren Azuma und Koike [2] gegen die Anwendung nach Hata, da die Anwendung eines Blur die Kanteninformation von Text so abschwächt, dass die Lesbarkeit darunter leidet. Allerdings wurde hierbei ein Filter gewählt, dessen fokussierter Highlight kreisförmig war und keine Rücksicht auf die Zeilenbreite oder -höhe nahm, weshalb Zeilen hinsichtlich beider Dimensionen abgeschnitten wurden. Dabei kann sich eine Hervorhebung relevanter Textstellen mit Berücksichtigung dieser Textdimensionen als äußerst effektiv herausstellen, wie Miksch et al. [42] in ihrer Untersuchung des *semantic depth of field* (SDOF) zeigen konnten. Dieses ist eine Erweiterung des Begriffs des *depth of field* (DOF) aus der Fotografie, wobei das Scharfstellen von Objekten im Vordergrund und gleichzeitiges Weichzeichnen von Objekten im Hintergrund auf die Relevanz der Elemente in Mengen übertragen wird.

Dementsprechend sollte für diese Studie ein Blur-Filter gebaut werden, der aus zwei rechteckigen Teilen besteht, sodass zwischen diesen genug Platz für den fokussierten Text liegt. Für den Blur-Effekt wurde ein Unity Shadergraph nachgebaut⁷ und auf das Material von zwei Shapes in der Breite des Textpanels angewandt (s. Abbildung 6(a)). Die Shapes wurden hierdurch komplett transparent und wirkten wie ein Weichzeichner auf dahinterliegende Objekte. Beide Shapes wurden übereinander angeordnet (mit dem Platz einer Zeilenhöhe dazwischen) und einem gemeinsamen Parent Object zugeordnet (s. Abbildung 6(b)). Dies ergab den Blur-Filter. Der Blur-Filter sollte in seiner Größe frei anpassbar sein, um so ein personalisiertes, ergonomisches Leseerlebnis zu garantieren. Um einen diskreten Wertebereich für die spätere Auswertung zu erzielen, sollte der Platz zwischen beiden Shapes einer gewünschten Anzahl

von Zeilenhöhen entsprechen. Für diese Einstellung wurden der A- und B-Button des Controllers auf eine gleichzeitige y-Translation beider Shapes gemappt, sodass mit jedem Auswählen der Taste beide Shapes ähnlich wie eine Kamerablende geweitet und verengt werden konnten. Mit jedem Auswählen des A-Buttons erhöhte sich der Platz zwischen den „Blenden“ um eine Zeilenhöhe, während der B-Button den fokussierten Bereich um eine Zeilenhöhe reduzierte (s. Abbildung 6(c)). Kleinstmögliche Zeilenhöhe sollte 1 sein, während es nach oben keine Begrenzung gab.

Über diese Anpassungsmöglichkeit soll ein flexibles Informations-Chunking stattfinden. Dabei würden Sinnabschnitte (z.B. Ab-/Sätze), die in ihrer Höhe variieren können, vom visuellen clutter des restlichen Textes isoliert. Nach dem divide-and-conquer-Prinzip sollen die Leser:innen auf diese Weise Kontrolle über die unüberschaubare Textmenge erlangen.

Als nächstes musste eine Eingabemodalität zur Steuerung des fokussierten Bereichs gefunden werden. Eine Möglichkeit lag auf der Hand: Der Blick selbst. So übernehmen die Augen die motorische Funktion beim Lesen und könnten demnach analog zu einem Scrolling-Rad einer Maus verwendet werden, um den fokussierten Bereich über den Text zu bewegen. Die visuelle Aufmerksamkeit würde demnach 1:1 durch die Position des fokussierten Bereichs abgebildet werden. Nur zeigt sich hierbei das innerhalb der Eyetracking-Community so häufig erwähnte Problem des *Midas Touch*, welches von Jacob [31] erstmals geprägt wurde. Da die Augen immer aktiv sind, würde jeder Blick eine Interaktion mit der virtuellen Umgebung darstellen; teils willentlich, teils unwillentlich. So verhält es sich auch bei der Positionierung des fokussierten Bereichs durch die Augen; eine andauernde Dringlichkeit, eine bewusste Entscheidung mit den Augen zu treffen, kann somit auf Dauer anstrengend sein. Ebenfalls könnten die schnellen, ruckartigen Bewegungen der Augen zu einem Flackern an den Rändern des Blur-Filters führen, welches wiederum die visuelle Aufmerksamkeit auf sich zieht. Wenn überhaupt könnte der fokussierte Bereich für die Dauer einer Zeile in seiner Höhe stabilisiert werden, bevor die Augen und so der gesamte fokussierte Bereich um die Höhe einer Zeile nach unten springen. Darüber hinaus sind die Augen von Natur aus Werkzeuge zur Aufnahme der Umwelt, statt zur Manipulation dieser.

⁷<https://www.youtube.com/watch?v=NLgrHFCF3oU>

Es kann jedoch schnell eine Alternative für die Eingabemodalität gefunden werden, da der Blick genauso durch die Kopfausrichtung eingestellt werden kann. Dies ist im Alltag sehr präsent, denn liegt ein auffälliger Reiz im Augenwinkel, wird einfach der Kopf entsprechend in dessen Richtung rotiert. Kopfbewegungen können zudem den Blick ähnlich wie ein Stativ einer Kamera komplett willentlich und kontrolliert einstellen, wohingegen viele Augenbewegungen komplett unbewusst geschehen. Dies bringt den weiteren Vorteil mit sich, dass der Filter abhängig von der Kopfausrichtung fest auf den zu lesenden Zeilen sitzt, während der Blick innerhalb des fokussierten Bereichs noch frei schweifen kann. Somit würde dem Gefühl nach immer noch der Blick das steuernde Element sein. Auch ein Zeilensprung kann erst dann stattfinden, wenn der gesamte eingestellte Bereich komplett gelesen wurde. Zusätzlich fühlt es sich beim Lesen natürlicher und angenehmer an, wenn der Kopf für die Augen innerhalb der obersten oder zumindest in den untersten Zeilen leicht nachjustiert.

Die Umsetzung der Positionierung des Filters auf dem Text geschah durch das Raycasting mit dem forward-Vektor, ausgehend von der virtuellen Kamera (nose direction). Da der Filter am oberen und unteren Ende des Textpanelss zusätzlichen Bewegungsspielraum haben sollte, wurde für das Raycasting eine unendlich große Plane auf gleicher z-Höhe wie die Textplane positioniert. Bei einer Kollision mit dem nose-direction-Ray wurde für die Position des Filters letztendlich nur die y- und z-Komponente des Schnittpunkts auf den Transform des Filters angewandt, sodass sein mittig sitzendes „Loch“ nur über die X-Rotationsachse des Kopfes (*pitch*) auf gleicher Höhe wie die anvisierten Zeilen positioniert werden kann (s. Abbildung 6(d)). Die Perspektive der Leser:innen bei Verwendung des Peripheral Blur kann in Abbildung 5 nachvollzogen werden. Somit wird mithilfe des Kopfes Zeilenblock für Zeilenblock gescannt und die Größe des fokussierten Bereichs ggf. mit dem Controller angepasst.

Mit dem Filter wurde somit eine Art verstellbarer field-of-view restrictor geschaffen. Dieser wird für gewöhnlich in VR-Anwendungen eingesetzt, um den 3D-Bewegungsfluss in der Peripherie zu reduzieren, da dieser sonst zu Motion Sickness führen kann [1]. Ebenfalls treten diese im Normalfall in Form von rund ausgeschnittenen Vignette-Overlays in Erscheinung, da sie keine Rücksicht auf Textdimensionen nehmen müssen. Übertragen auf den Bewegungsfluss in der zweidimensionalen Text-Umgebung (bzw. Lesefluss der Augen) könnte dies Implikationen für das Lesen in VR haben, welches bekanntlich länger dauert als Lesen auf Papier oder LCD-Bildschirmen (s. Abschnitt 1).

Der Peripheral Blur kann aufgrund seiner sequentiellen Hervorhebung von mehreren Wörtern zur Zeit ebenfalls mit einer *Rapid Serial Visual Presentation* (RSVP) [3,49] verglichen werden. Ein prominentes Beispiel hierfür lässt sich in Spritz Reading [56] finden. Hierbei wird ein passiver Lesefluss hergestellt, indem die einzelnen Wörter isoliert hintereinander eingeblendet werden. Dadurch kann eine feste Lesegeschwindigkeit in *words per minute* (wpm) eingehalten werden. Über einen sogenannten „Redicle“ (bestehend aus einem Buchstaben in rot und einem Rahmen) um jedes Wort wird das Auge auf den Buchstaben gelenkt, durch den das Wort schneller identifiziert werden kann. Da alle Wörter um diesen zentrierten Marker gesetzt werden, soll Augenbewegung und damit eine Ermüdung der Augen minimiert werden, während die Lesegeschwindigkeit erhöht werden soll. Allerdings bietet Spritz diesbezüglich keine wissenschaftliche Evidenz und zudem gehen Benedetto et al. [3] in ihrer kritischen Auseinandersetzung mit Spritz darauf ein, dass gerade ein Minimieren von Augenbewegungen zu einer Belastung der Augen durch z.B. reduzierte Blinzelrate führen kann. Der Peripheral Blur hingegen zielt nicht auf ein Minimieren der Augenbewegungen per se ab, sondern lediglich auf ein Minimieren möglicher Ablenkungen durch umliegenden Text. Ebenfalls können sich die Augen frei innerhalb des fokussierten Bereichs bewegen und sind nicht durchgängig

am gleichen Punkt verankert, wodurch die Lesegeschwindigkeit weiterhin normal variieren kann. Nichtsdestotrotz liegt die Absicht dieser Leseassistentz-Funktion darin, die Leseeffizienz, insbesondere die Lesegeschwindigkeit, zu erhöhen.

Da mit dem Seitenwechsel nur die letzte Zeileneinstellung gespeichert wurde, sollte zusätzlich eine Historie über alle Einstellungen pro Seite geführt werden, welche zudem protokollierte, wie oft Zeileneinstellungen pro Seite gemacht wurden. Über einen weiteren Raycast (Gazeray erzeugt aus den Eyetracking-Daten über Tobii XR SDK) wurde analog zur Leseassistentz-Funktion Scanpath die Lesezeit gemessen.

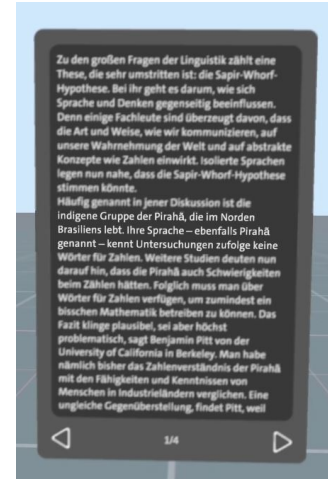


Abbildung 5: Lesen mit Assistenz-Funktion Peripheral Blur. Die Person hat die Einstellung von drei fokussierten Zeilenhöhen ausgewählt und visiert den mittleren Teil des Textes an.

4 BENUTZERSTUDIE

4.1 Textparameter

Um eine objektive Vergleichbarkeit der drei Texte zu garantieren, mussten sie neben dem Autoren, Thema und Länge auch im Lesbarkeitsindex Flesch Reading Ease (FRE) [17] Ähnlichkeiten aufweisen. Die Berechnung geschah online [53] nach der deutschen FRE-Formel [45], wie in Gleichung (1) zu sehen:

$$FRE_{Deutsch} = 180 - ASL - (58,5 * ASW) \quad (1)$$

wobei

ASL: Durchschnittliche Satzlänge, mittlere Wortanzahl pro Satz

ASW: Durchschnittliche Silbenanzahl pro Wort

Da die Proband:innen mindestens im ersten Semester studierten und dementsprechend die Mindestanforderung von Abitur-Absolvent:innen erfüllen sollten, wurde nach wissenschaftlichen Texten auf Deutsch mit einem FRE von 50 gesucht, da sein Anforderungsniveau genau auf der Schwelle zwischen beiden akademischen Graden liegt [45]. So fiel die Wahl auf einen längeren Artikel der wissenschaftlichen Zeitschrift Spektrum [60], welcher aufgrund seines konsistenten Schreibstils und Thematik in drei eigenständige Texte mit ähnlicher Länge (37-38 Sätze) und ähnlichem bis gleichen FRE (zwischen 49 und 50) zerlegt und passend gekürzt werden konnte (s. Anhang). Für den nachträglichen Abruf der Inhalte wurden pro Text drei Fragen formuliert, die jeweils nur eine konkrete Antwortoption zuließen (eine Variation im Wortlaut war möglich).

Um die Texte in die Anwendung miteinzubinden, wurde in Unity ein Textpanel erstellt. Dieses setzt sich aus zwei ineinander

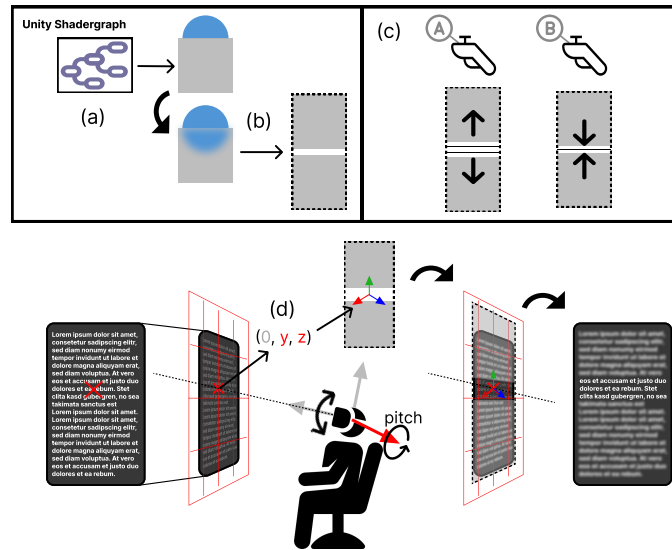


Abbildung 6: Leseassistenten-Funktion Peripheral Blur und ihre Komponenten. (a): Unity Shadergraph wird für den Blur-Effekt auf Material von zwei Shapes angewandt. Dieser macht die Shapes komplett transparent (grau für Illustration) und wirkt wie ein Weichzeichner auf dahinterliegende Objekte. (b): Beide Shapes werden übereinander angeordnet und einem gemeinsamen Parent Object zugeordnet. Dies ergibt den Blur-Filter. (c): Der Abstand zwischen den Shapes (fokussierter Bereich) ist über den Controller verstellbar. A-Button: Fügt eine Zeilenhöhe hinzu, B-Button: Nimmt eine Zeilenhöhe weg. (d): Der nose-direction-Ray schneidet abhängig von der Kopffrotation (*pitch*) eine unendlich große Plane auf gleicher z-Höhe des Textpanels (rotes Gitter). Der resultierende Schnittpunkt wird in seiner y- und z-Komponente auf den Transform des Blurfilters angewandt.

geschachtelten Canvasses (Worldspace, uniform skaliert) zusammen, wobei der innere Canvas dem Texthintergrund (ca. 1 x 1,6 Meter) und der äußere dem Rahmen des Textpanels (ca. 1,2 x 1,8 Meter) entspricht, welcher zudem die UI zum Seitenwechsel enthält. Da Studien in der Vergangenheit zeigten, dass heller Text auf dunklem Hintergrund bevorzugt wird, statt umgekehrt [11, 16], wurde als Textfarbe weiß (#FFFFFF) gewählt und für den Hintergrund ein dunkler Grauton (#373737). Dadurch bildeten Text und Hintergrund ein Kontrastverhältnis von 11,9 : 1, was den WCAG Standard nach Level AAA erfüllt [34]. Um Hintergrund und Textpanel-Rahmen visuell auseinanderzuhalten, wurde für letzteres ein hellerer Grauton (#565656) gewählt.

Das Textpanel wurde auf 1,75 Meter Entfernung zur Position der Leser:innen gesetzt [18]. Basierend auf den Ergebnissen von Kojić et al. [34] zu der präferierten Schriftgröße mit Einberechnung der Winkelausdehnung (*angular size*) bei längeren Texten in VR (Quest) wurde eine Schriftgröße gewählt, welche wie 17,87 mm aus 1 Meter Distanz wirkt, d.h. 17,87 Distance-Independent Millimeters (dmm) [25]. Um die Winkelausdehnung bei 1,75m Entfernung zu berücksichtigen, wurde der Wert folgendermaßen skaliert: $17,87\text{mm} \times 1,75 \approx 31,27\text{mm}$, was ca. der Schriftgröße 4,5 in der Hausschrift *TheSans UHH* mit normaler Schriftstärke (*Regular*) entspricht (ausgehend von der Höhe der Großbuchstaben).

Unter Berücksichtigung dieser Textparameter bot die Fläche des Texthintergrunds Platz für maximal 26 Zeilen. Da jeder Text den Hintergrund bis zu 4 mal auffüllen konnte, wurde jeder Text in 4 Parts bzw. Seiten heruntergebrochen.

Für den Seitenwechsel wurde eine einfache UI erstellt, welche aus zwei pfeilförmigen Buttons und einer Seitenanzeige im Bruchformat „X/4“ bestand (s. Abbildung 7). Beide Buttons realisierten den Seitenwechsel zur nächsten bzw. vorherigen Seite, indem sie das zuvor gelesene Viertel des Textes inaktiv und das nächste aktiv setzten (zusammen mit der Aktualisierung der Seitenanzeige).

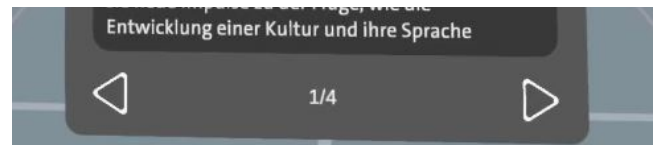


Abbildung 7: Einfache UI für den Seitenwechsel. Pfeilförmige Buttons wechseln zur vorherigen bzw. nächsten Seite. Die zentrierte Seitenanzeige ist im Bruchformat „X/4“.

4.2 Studiendesign

Für die Studie wurde ein Within-Subjects-Design gewählt. Es gab insgesamt drei Konditionen: Zwei Leseassistenten-Funktionen (wie in 3.2, 3.3 beschrieben) plus die Kontrollbedingung bzw. Baseline-Kondition. Auf jede der zwei Leseassistenten-Funktionen kam einer von drei Texten (immer ein anderer), während die Baseline-Kondition dem Lesen eines Textes ohne Leseassistenten entsprach. Um auf serielle Positionseffekte kontrollieren zu können, wurde pro Text eine Frage zum Anfang, Mittelteil und letzten Drittel des Textes gestellt. Mit den Texten entstand ein weiterer Randomisierungsfaktor, der vom Counterbalancing miteinbezogen werden musste. Aufgrund dieser erhöhten Komplexität erforderte die Konstruktion eines Latin Square zusätzlich zu der Methode von Bradley (nach dem Schema A, N, B, N - 1, C, N - 2, ...) [6] einen weiteren Schritt für eine gleichzeitige Randomisierung von Kondition und Textstimulus. Hierfür wurde das Verfahren von Lewis [64] gewählt, welches an die Schritte von Bradley anknüpft und die ungerade Anzahl an Konditionen berücksichtigt. Demnach wurden die resultierenden Strukturen nach Bradley auf Kondition und Text vertikal gespiegelt und über Kreuz gepaart. Daraus ergab sich ein Counterbalanced Latin Square bestehend aus sechs Zeilen und drei Spalten, mit je einem Konditions-Text-Paar in jeder Zelle (s. Anhang). Dies entspricht sechs verschiedenen Versuchsblöcken. Eine Power-Analyse vorab ergab eine Stichprobengröße von mindestens $N = 27$ zum

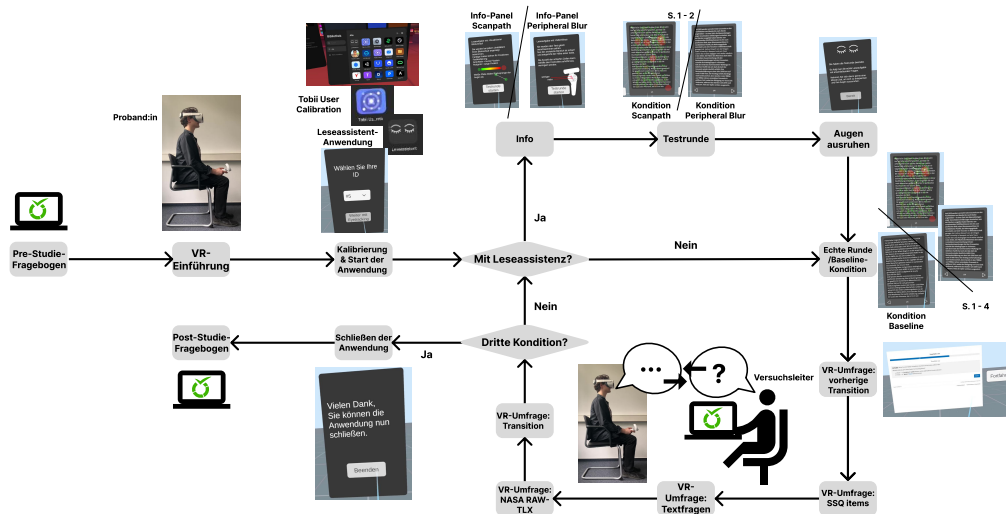


Abbildung 8: Ablauf der Benutzerstudie als Flussdiagramm mit nachgestelltem VR-Setup und Screenshots aus der Anwendung.

Nachweis mittlerer Effekte (Effektgröße 0,25, $\eta_p^2 = 0,06$) mit 80% Power für eine ANOVA mit einem alpha-Niveau von 0,05. Um die sechs möglichen Versuchsblöcke gleichmäßig aufzufüllen, wurde der Stichprobenumfang auf 30 erweitert und die IDs von oben nach unten aufsteigend angeordnet, sodass sich alle Versuchsblöcke nach jeder sechsten Proband:in für die nachfolgenden sechs wiederholen. Daraus ergaben sich fünf Säulen aus je sechs Proband:innen.

Als Umgebung für das Experiment wurde ein ruhiger, leerer Raum gewählt. Da das Lesen in VR bis auf Kopfbewegungen keine Bewegung erfordert, fand das gesamte VR-Experiment im Sitzen statt (*Seated VR*). Der hierfür verwendete Stuhl besaß keine Rollen, so dass während des Lesens keine zusätzlichen Bewegungen gemacht werden konnten und für alle Personen die gleichen Lesebedingungen galten (s. Abbildung 8).

4.3 Ablauf der Studie

Es wurden entsprechend des Studiendesigns 30 Personen über das universitätsinterne Online-Portal Sona rekrutiert oder aus dem nahen Umfeld eingeladen. Da das Experiment für 90 Minuten angesetzt war, wurden für die Teilnahme 1,5 Versuchspersonenstunden vergeben. Zugelassen wurden Personen, die eine normale oder korrigierte Sehstärke besitzen, davon ausgeschlossen waren Brillenträger. Für letzteres wurde eine Ausnahme gemacht, wenn die Dioptrie beider Augen ± 1 Dioptrie betrug und bei Vorführung einer Videodemo der Anwendung (in VR) klar kommuniziert wurde, dass der Text auf dem Panel ohne Probleme zu lesen ist.

Als Erstes wurden den Proband:innen die Einverständniserklärungen vorgelegt und nach Unterzeichnung die ersten Schritte in der VR (Kalibrierung der Augen) bis zum Öffnen der Anwendung, allgemeine Bedienung der UI und schließlich die Besonderheiten jeder Leseassistent-Funktion anhand von Screenshots erläutert. Es wurde ebenfalls auf die Fragebögen zwischen den Konditionen eingegangen und dass bzw. wie sie innerhalb der VR ausgefüllt werden müssen. Anschließend füllten die Proband:innen am Laptop den Fragebogen der Pre-Studie mit demografischen Daten und ihren Erfahrungen bezüglich Lesen und VR aus. Nachdem die stationäre Grenze in Relation zur aufrechten Sitzposition eingerichtet worden war, wurde den Personen das VR-Headset aufgesetzt und der Controller überreicht. Eine Nachstellung des VR-Setup und nachfolgenden Schritte der Benutzerstudie können anhand eines Flussdiagramms (Abbildung 8) visuell nachvollzogen werden. Beim Stabilisieren des Headsets am Kopf wurden die Proband:innen darauf hingewiesen, auf eine zentrierte Position der Augen vor den Gläsern zu achten.

Als nächstes erfolgte die Instruktion, den Kopf gerade auszurichten und die virtuelle Umgebung zu zentrieren. Anschließend wurde die Anwendung für die Augenkalibrierung geöffnet und nach Abschluss die Anwendung des Leseassistenten geöffnet. Mit Bestätigen der ID wurde das Eyetracking gestartet und die erste Kondition begonnen. Von nun an folgte die Anwendung immer dem gleichen Ablauf: Für alle Konditionen mit Leseassistent öffnete sich ein Panel mit Infos zur Bedienung bzw. Interpretation der Darstellungsform der Assistenzfunktion, bevor ein Testlauf mit gleicher Funktion und einem Dummy-Text (immer der gleiche, s. Anhang) durchgeführt wurde. Der Dummy-Text war auf zwei Seiten gekürzt und erforderte keinen inhaltlichen Abruf, da sich die Personen erst einmal mit der jeweiligen Funktionalität vertraut machen sollten. Mit dem Abschließen der Testrunde wurde den Proband:innen durch ein Fenster kommuniziert, dass sie ihre Augen entspannen konnten (ohne das Headset abzusetzen). Fühlten sich die Proband:innen bereit, starteten sie in den eigentlichen Durchlauf mit der gleichen Funktionalität und einem von drei vierseitigen Texten. Im Falle der Baseline-Kondition startete der Durchlauf direkt mit den vier Seiten, da diese Kondition ohne Leseassistent auf keine Testrunde angewiesen war. Hatten die Proband:innen den Text zuende gelesen, wurde das Beenden der Kondition über einen Button bestätigt, der ab der vierten Seite sichtbar blieb. Mit jedem Beenden einer Kondition öffnete sich das Fenster mit der Umfrage für den VR-Teil der Studie. Für jede abgeschlossene Kondition begann dieser mit den Elementen des Simulator Sickness Questionnaire (SSQ) [33]. Eine Transition innerhalb der Umfrage gab den Hinweis, dem Versuchsleiter bescheidzugeben und in der Umfrage zu verharren, da nun die drei inhaltlichen Fragen zum zuvor gelesenen Text folgen würden. Der Versuchsleiter las den Proband:innen die Fragen vor und protokollierte ihre Antworten in einer separaten Umfrage. Hierauf folgte ein NASA-TLX Fragebogen wieder in der VR-Umfrage [23]. Über eine weitere Transition konnte das Umfragefenster geschlossen und gleichzeitig in die nächste Kondition gestartet werden etc. Mit Beendigung der letzten Kondition und den zugehörigen Fragebögen konnte die Anwendung beendet, das VR-Headset abgesetzt und ein letzter Fragebogen mit Feedback zu beiden Leseassistent-Funktionen am Laptop ausgefüllt werden.

5 ERGEBNISSE

Es wurde eine quantitative und eine qualitative Analyse auf den Daten aus der VR und den restlichen Fragebögen durchgeführt. Vorab wird ein Überblick über die Demografie zusammen mit ihren VR-Erfahrungen und Lese-verhalten gegeben, anschließend werden

die statistischen Ergebnisse der quantitativen Analyse berichtet, bevor die qualitative Analyse mit einer inhaltlichen Zusammenfassung des Feedbacks zu den Leseassistenten-Funktionen folgt.

5.1 Demografie und Vorerfahrungen in VR

Von den 30 Proband:innen waren 7 männlich, 21 weiblich und 2 nicht-binär. Das Alter lag zwischen 20 und 44 Jahren, ca. 26 Jahre bildete den Durchschnitt ($M: 25,8$; $SD: 5,1$). Bis auf eine Person mit Russisch und Ukrainisch gaben alle Personen Deutsch als Muttersprache an. 19 Personen benötigten keine Sehkorrektur, 10 hingegen trugen Kontaktlinsen und 1 Person war Brillenträger, wurde aber nicht ausgeschlossen, da sie das erwähnte Ausnahmekriterium (s. 4.3) erfüllte. Keine:r der Proband:innen gab eine Farbfehlsichtigkeit an. Die Erfahrung mit VR-Systemen wurde über die Häufigkeit der Nutzung erhoben. So gaben 2 Personen an, mehr als 1x pro Woche VR-Geräte zu nutzen, 5 Personen 1x pro Monat, 1 Person 1x pro Vierteljahr, 4 Personen 1x pro Halbjahr, 2 Personen 1x pro Jahr und 16 Personen weniger als 1x pro Jahr. Unter den 30 Proband:innen fanden sich 26 Student:innen; hierbei vertreten waren die Studiengänge 2-Fach Lehramt (Deutsch/Biologie) mit einer Person, Psychologie mit 13 Personen und Mensch-Computer-Interaktion mit 12 Personen.

5.2 Leseverhalten

Die Einschätzung sowohl des eigenen Leseverhaltens als auch der gemachten Leseпаusen wurde auf einer Likert-Skala mit den Werten 1 (für „sehr langsam“ bzw. „keine Leseпаasen“) bis 7 (für „sehr schnell“ bzw. „viele Pausen“) angegeben. Für das Leseverhalten lag der Median bei 4 bzw. 3 für die Leseпаasen und der IQR in beiden Fällen bei 2. 19 Personen gaben an, pro Woche mehr als einmal zu lesen, 9 Personen einmal pro Woche und 2 Personen einmal pro Monat. 23 Personen von 30 gaben an, Probleme beim Lesen zu haben. Dabei wurden ihre expliziten Aussagen in Schrift ausgewertet und wie folgt kategorisiert und paraphrasiert (inhaltliche Überlappungen konnten nicht vermieden werden):

- **Müdigkeitsbedingt**
2 Personen gaben Müdigkeit als Beeinträchtigung beim Lesen an.
- **Bedingt durch Augenbelastung**
4 Personen sahen müde oder trockene Augen als Hindernis beim Lesen.
- **Aufmerksamkeitsbedingt**
In den Antworten war die Aufmerksamkeitsschwierigkeit mit 14 Personen am meisten vertreten. Dieses Problem äußert sich laut den Angaben darin, mit den eigenen Gedanken abzuschweifen, ablenkbar durch die Umgebung zu sein, durch unaufmerksames Überfliegen oder längeres Lesen über Zeilen zu stolpern und dementsprechend viele Rücksprünge im Text zu machen, das Gelesene zu Vergessen oder den Inhalt nicht bewusst aufzunehmen. In einem Fall wurde sogar ADHS erwähnt.
- **Textbedingt**
2 Personen führten die Problematik auf eine Irritation durch den zu lesenden Text zurück, die durch unbekannte Worte oder Grammatik- bzw. Rechtschreibfehler verursacht wird.
- **Verständnisbedingt**
2 Personen werden laut ihrer Aussage im Leseprozess gebremst, da sie einen Satz häufiger nachlesen müssen, um ihn zu verstehen.
- **Geschwindigkeitsbedingt**
3 Personen benötigen für die Informationsaufnahme mehr Zeit, lesen sehr sorgsam oder schätzen ihre Lesegeschwindigkeit als zu niedrig ein.

• Rücksprünge im Text

2 Personen gaben in ihren Antworten an, Wörter zu überlesen oder häufig Zeilensprünge zu machen (keine Begründung wie beim Punkt Aufmerksamkeitsbedingt).

Ebenfalls wurden die Proband:innen nach ihren eigenen Lesestrategien gefragt. 20 von 30 Personen machten hierzu Angaben, während der Rest keine Kenntnisse über eine eigene Lesestrategie hatte. Wie bei den Problemen beim Lesen zuvor wurden die expliziten Aussagen zusammengefasst und Kategorien zugeordnet (auch hier kam es in den Antworten zu inhaltlichen Überlappungen):

• Markieren

7 Personen verwenden laut eigener Aussage Markierungen innerhalb des Textes, um wichtige Passagen hervorzuheben oder die Aufmerksamkeit aufrechtzuerhalten.

• Nochmal Lesen

Die Methode, die mit 9 Personen am häufigsten aus den Antworten hervorging, stellte das wiederholte Nachlesen kritischer Textabschnitte dar.

• Sich Zeit lassen

3 Personen gaben in ihrer Antwort Zeit als Faktor an. Dies beinhaltet ein Verlangsamen des Lesens bei wichtigen Stellen aber auch die zeitliche Einplanung einer Lesephase selbst, um sich dabei entspannen zu können.

• Strukturiertes Lesen

3 Personen bevorzugten ein strukturiertes Lesen. Dabei wird der Text abschnittsweise gelesen oder bereits mit Stichpunkten abgearbeitet, ebenfalls wird nach jedem Abschnitt eine Pause gemacht. Zusätzlich ging aus einer Antwort hervor, dass beim ersten Überfliegen des Textes die grobe Textstruktur verinnerlicht wird, bevor genauer nachgelesen wird.

• Wörter nachschlagen

2 Personen gaben an, Duden oder Übersetzer für unbekannte Wörter heranzuziehen.

• Eigene bzw. innere Stimme hören

3 Personen greifen nach ihrer Angabe auf die Technik zurück, den Text mit der eigenen Stimme zu wiederholen. Dies umfasst das Hören verschieden betonter Aussprachen des gleichen Satzes oder sogar die Wahrnehmung der „inneren Stimme“, d.h. das gedankliche Mitsprechen des Satzes.

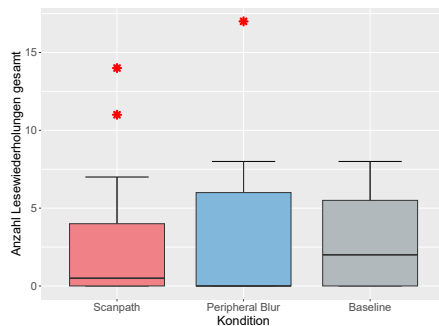
• Andere

3 Personen gaben spezifischere Strategien an. So hört 1 Person instrumentelle Musik, eine weitere gab gezieltes oder überfliegendes Lesen an und wieder eine andere sequentielles Lesen.

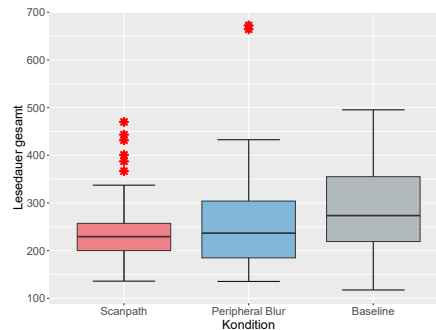
5.3 Quantitative Analyse

Für jede der Messungen werden zunächst über die Werte der deskriptiven Statistik berichtet und ggf. Auffälligkeiten grafisch anhand von Boxplots veranschaulicht, anschließend folgt die inferentielle Statistik mit ihren Ergebnissen aus den Signifikanztests.

Um zwischen den Konditionen auf signifikante Unterschiede zu testen, wurde aufgrund des Studiendesigns (> 2 Konditionen, Within-Subjects, 1 Faktor: Kondition) für alle intervallskalierten Merkmale die Normalitäts-Annahme einer einfaktoriellen Varianzanalyse mit Messwiederholung (oder: *one-way repeated measurements ANOVA*) geprüft. Hierfür wurde ein Shapiro-Wilk-Test auf den Residuen eines entsprechenden ANOVA-Modells durchgeführt und unterstützende Plots, Histogramme und Q-Q Plots erstellt (s. Anhang). Ausgehend



(a) Anzahl Lesewiederholungen



(b) Gesamtlese dauern in Sekunden

Abbildung 9: Leseparameter: (a) Anzahl Lesewiederholungen und (b) Gesamtlese dauern in Sekunden

von einer Erfüllung bzw. Verletzung der Normalitätsannahme wurde ein parametrischer bzw. nonparametrischer Test auf den Daten durchgeführt. Konnte keine Normalität angenommen werden, wurde der Median (*Mdn*) statt dem Mittelwert (*M*) betrachtet.

5.3.1 Simulator Sickness Questionnaire (SSQ)

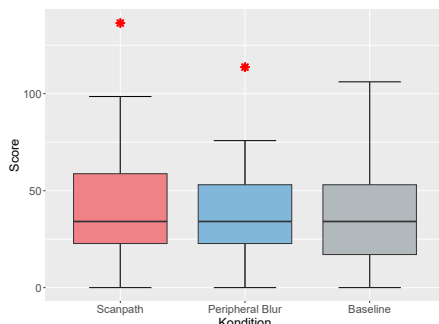
Der SSQ wurde nur nach den Konditionen erhoben. Eine Baseline-Messung vor dem VR-Teil der Studie blieb aus.

Für jede Proband:in wurden sowohl die Scores für die Subskalen als auch der Gesamtscore (*Total Score [TS]*), nach Anwendung der Formel so wie beschrieben von Bimberg et al. [4]) berechnet und ausgehend davon die Streuungs- und Lagemaße berechnet.

Hinsichtlich der Übelkeits-Subskala war der Interquartilsabstand bei allen drei Boxplots gleich (*IQR* = 19,08), ebenso die Mediane (Scanpath und Baseline: *Mdn* = 9,54) bis auf die Kondition Peripheral Blur (*Mdn* = 19,08).

Die Boxplots der Okulomotorik-Subskala zeigten neben einem leicht abweichenden Interquartilsabstand bei Peripheral Blur (*IQR* = 30,32; vgl. Scanpath und Baseline: *IQR* = 36,01) eine auffällige Streuung nach oben, was sich an den oberen Whiskern von Scanpath und Baseline (etwas höher mit *max* = 106,12) beobachten lässt, welche fast doppelt so lang sind wie bei Peripheral Blur (s. Abbildung 10). Ebenfalls besaßen Scanpath (*max* = 136,44) und Peripheral Blur (*max* = 113,7) je einen Ausreißer. Darüber hinaus war der Median bei allen Konditionen gleich (*Mdn* = 34,11).

Abbildung 10: SSQ: Okulomotorik-Score (O)



Die Desorientierungs-Subskala wies in ihren Boxplots einen höher gelegenen Ausreißer bei Kondition Scanpath auf (*max* = 222,72; vgl. Peripheral Blur und Baseline: *max* = 125,28). Kondition Scanpath und Peripheral Blur besaßen den gleichen Median und Interquartilsabstand (*Mdn* = 34,8; *IQR* = 38,28), nicht sehr viel anders verhielt

sich Kondition Baseline (*Mdn* = 27,84; *IQR* = 41,76).

Beim Gesamtscore ließ sich im Boxplot zu Kondition Scanpath ein weiter abseits gelegener Ausreißer verzeichnen (*max* = 1735,7) zusammen mit einem zweiten auf ähnlicher Höhe wie die Ausreißer von Peripheral Blur (*max* = 965,1) und Baseline (*max* = 972,5). Median und Interquartilsabstand von Scanpath (*Mdn* = 311,2; *IQR* = 297,8), Peripheral Blur (*Mdn* = 320,9; *IQR* = 285,89) und Baseline (*Mdn* = 293,4; *IQR* = 282,95) unterschieden sich augenscheinlich kaum voneinander.

Da hinsichtlich der Subskalen und Gesamt-Score keine Normalverteilung der Daten angenommen werden konnte ($p < 0,001$), wurde ein Friedman-Test auf allen Subskalen und Gesamt-Score durchgeführt. Es konnten keine signifikanten Unterschiede zwischen den Konditionen gezeigt werden, weder im Bezug auf die Subskalen (Übelkeit: $p = 0,51$; Okulomotorik: $p = 0,76$; Desorientierung: $p = 0,69$) noch auf den Gesamtscore ($p = 0,91$).

5.3.2 Lesewiederholungen

Pro Proband:in wurde je Kondition die Anzahl der Lesewiederholungen über alle Seiten aufaddiert, um so einen verwertbaren Score pro Person zu erhalten.

Die Boxplots von Peripheral Blur (*IQR* = 6) und Baseline (*IQR* = 5,5) zeigten einen ähnlichen Interquartilsabstand, Kondition Scanpath den kleinsten (*IQR* = 4). Beide Assistenzfunktionen besaßen Ausreißer mit 14 Lesewiederholungen bei Scanpath und 17 bei Peripheral Blur (s. Abbildung 9(a)). Der Median von Peripheral Blur lag bei 0 und unterschied sich bei Scanpath (*Mdn* = 0,5) und Baseline (*Mdn* = 2) nur geringfügig.

Da den Lesewiederholungen keine Normalverteilung nachgewiesen werden konnte ($p < 0,001$), wurde ein Friedman-Test durchgeführt. Hierbei konnten keine signifikanten Unterschiede zwischen den Konditionen bezüglich der Anzahl an Lesewiederholungen gefunden werden ($p = 0,94$).

5.3.3 Lesedauern

Pro Proband:in wurde pro Kondition die Gesamtlese dauern in Sekunden als Score berechnet.

Bei Betrachtung des Boxplots für die Baseline-Kondition (s. Abbildung 9(b)) wurde im Vergleich zu Scanpath und Peripheral Blur eine auffallend größere Spannweite beider Whisker deutlich (*min* = 117,5; *max* = 495,3 $\approx$ 8 Minuten), ebenfalls eine höhere Lage der Box (*Mdn* = 273,5 Sekunden $\approx$ 5 Minuten, *IQR* = 136,04), jedoch insgesamt nicht höher als die Mediane der anderen Gruppen (Scanpath: *Mdn* = 229,3 Sekunden $\approx$ 4 Minuten, Peripheral Blur: *Mdn* = 236,7 Sekunden $\approx$ 4 Minuten). Werden die Ausreißer der anderen Konditionen ignoriert, liegt bei Baseline die größte Streuung nach oben (und unten, vgl. Scanpath und Peripheral Blur: *min* = 136,1; *min* = 135,5) vor. Hierauf folgen mit absteigenden

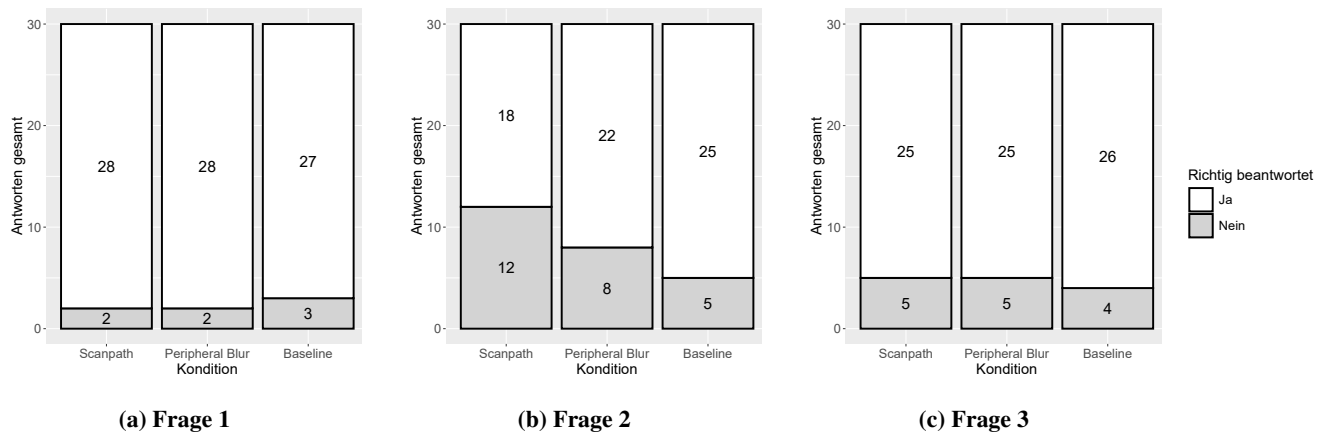


Abbildung 11: Informationsabruf pro Frage: Anzahl richtiger und falscher Antworten pro Lesedauerkondition bzw. Baseline-Kondition. Es wurde hierbei nicht zwischen den Positionen unterschieden, sondern nur nach Positionierung der Frage am Anfang (a), Mitte (b) oder Ende (c) des Textes gefiltert.

Werten Peripheral Blur und schließlich Scanpath (IQR = 57,18) mit der kleinsten Streuung. Peripheral Blur zeigt in seinem oberen Whisker eine nähere Höhe und Interquartilsabstand (IQR = 119,06) zur Baseline-Kondition, dafür eine tiefere Lage der Box. Wurden die Ausreißer betrachtet, ergab sich ein anderes Bild: Bei Scanpath lagen diese eng zwischen 350 Sekunden (ca. 6 Minuten) und 500 Sekunden (max = 470 Sekunden $\approx$ 8 Minuten). Bei Peripheral Blur gab es noch zwei größere, dicht beieinanderliegende Ausreißer (max = 672,5 Sekunden $\approx$ 11 Minuten).

Da eine Normalverteilung der Lesedauern nicht bestätigt werden konnte ($p < 0,001$), wurde ein Friedman-Test durchgeführt. Hinsichtlich der Gesamtlesedauern (in Sekunden) konnten mit diesem keine signifikanten Unterschiede zwischen den Konditionen nachgewiesen werden ($p = 0,53$).

5.3.4 Antworten zu Textfragen

Für den inhaltlichen Abruf der Textinhalte wurden die expliziten, protokollierten Antworten zu den Textfragen pro Proband:in gesichtet und auf einen binären Score 1 (Antwort richtig) oder 0 (Antwort falsch) abgebildet.

Um vorab auf Primäreffekte (*primacy effects*) und Rezenzeffekte (*recency effects*) zu kontrollieren, wurden die Antworten nach Positionierung der Frage am Anfang, in der Mitte und am Ende des Textes gefiltert, ohne dabei die Textzugehörigkeit zu berücksichtigen. Wie sich anhand von Abbildung 11(a) beobachten lässt, wurden auf die erste Frage die meisten richtigen Antworten gegeben, während Frage 2 die meisten Falsch-Antworten generierte, besonders Kondition Scanpath zeigte hierbei eine Tendenz nach oben. Für jede Frage wurde ein Cochran's Q-Test durchgeführt, um auf einen signifikanten Unterschied des Verhältnisses von richtigen zu falschen Antworten zwischen den Konditionen zu testen. Dieser war für keine der Fragen signifikant (Frage 1: $p = 0,87$; Frage 2: $p = 0,09$; Frage 3: $p = 0,93$).

Wurden die Gesamtsummen der richtigen Antworten für jede Kondition betrachtet, so wurden insgesamt bei der Baseline-Kondition die meisten richtigen Antworten gegeben ($N = 78$) und die meisten Falsch-Antworten bei Scanpath ($N = 19$, vgl. Baseline: $N = 12$).

Um weiterhin auf signifikante Unterschiede zwischen den Konditionen bezüglich der Anzahl richtiger Antworten zu testen, wurde die Normalitäts-Annahme einer einfaktoriellen Varianzanalyse mit Messwiederholung geprüft. Hierfür wurde pro Person für jede Kondition die Anzahl richtiger Antworten berechnet. Da für die Gesamtsummen der richtigen Antworten keine Normalverteilung gezeigt werden konnte ($p < 0,001$), erfolgte ein Friedman-Test, wel-

cher keine signifikanten Unterschiede zwischen den Konditionen hinsichtlich der Anzahl richtiger Antworten zeigen konnte ($p = 0,32$).

5.3.5 NASA RAW-TLX

Es wurde der RAW-TLX ohne Anwendung von Gewichten auf die Subskalen verwendet. Somit wurden die Scores zwischen 0 und 100 direkt übernommen und für den Total Workload ein normalisierter Wert durch das Mitteln aller sechs Subskalen berechnet [22].

Ausgehend von den Boxplots der geistigen Anforderungs-Subskala (*Mental Demand*) unterschieden sich die Mediane für Scanpath ($Mdn = 47,5$), Peripheral Blur ($Mdn = 42,5$) und Baseline ($Mdn = 47,5$) augenscheinlich nur gering voneinander. Die Interquartilsabstände lagen bei Scanpath (IQR = 43,75) und Peripheral Blur (IQR = 41,25) etwas höher als bei Baseline (IQR = 33,75).

Die körperliche Anforderung (*Physical Demand*) fiel bei ihren Boxplots durch mehrere Ausreißer auf (Scanpath, Peripheral Blur: max = 85; Baseline: max = 70). Scanpath und Baseline besaßen beide den gleichen Median ($Mdn = 5$) und Interquartilsabstand (IQR = 10), Peripheral Blur wich hiervon mit einem fast doppelt so großem Wert ab ($Mdn = 10$; IQR = 18,75). Letztere Kondition lag mit ihrem Median ($Mdn = 10$) auf gleicher Höhe wie das obere Quartil der beiden anderen Konditionen (3rd Qu. = 10).

Bei der Subskala der zeitlichen Anforderung (*Temporal Demand*) ließ sich ein höher gelegener Ausreißer (max = 90) in dem Boxplot von Scanpath beobachten. Ansonsten war der Interquartilsabstand bei allen Konditionen gleich (IQR = 20), die Mediane lagen bei den meisten Konditionen auf fast gleicher Höhe (Peripheral Blur: $Mdn = 12,5$ Baseline: $Mdn = 10$), bis auf Scanpath ($Mdn = 17,5$).

In der Leistungs-Subskala (*Performance*) zeigte Kondition Baseline in ihrem Boxplot eine auffällige Streuung nach oben, wie sich an dem oberen Whisker (max = 85, vgl. Scanpath: max = 70) in Abbildung 12(a) erkennen lässt. Eine weitere Exzentrik zeigte sich in den Ausreißern, die bei Peripheral Blur in einem Fall bis 100 (= max) streuen. Darüber hinaus waren die Interquartilsabstände bei Scanpath und Baseline gleich (IQR = 28,75). Bei Peripheral Blur fiel dieser etwas geringer aus (IQR = 22,5). Der Median von Peripheral Blur und Baseline waren gleich ($Mdn = 25$), dieser lag bei Scanpath etwas höher ($Mdn = 32,5$).

Bei Betrachtung der Anstrengungs-Subskala (*Effort*) zeigte sich in den Boxplots eine starke Abweichung in den oberen Whiskern für Scanpath und Peripheral Blur (max = 100, vgl. Baseline: max = 80), wie in Abbildung 12(b) zu sehen. Weiterhin waren die Interquartilsabstände bei Scanpath (IQR = 33,75) und Peripheral

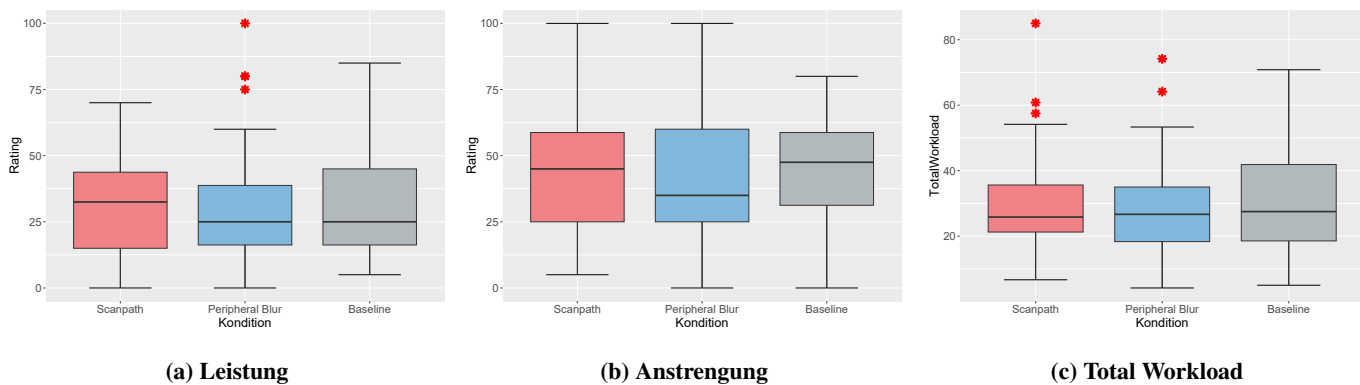


Abbildung 12: Subskalen des NASA RAW-TLX: Leistung (a), Anstrengung (b), Total Workload (c).

Blur (IQR = 35) sehr ähnlich, Baseline besaß den kleinsten Wert (IQR = 27,5). Zusätzlich lagen die Mittelwerte von Scanpath ($M = 44,33$; $SD = 23,55$) und Baseline ($M = 45$; $SD = 21,13$) sehr nah beieinander. Peripheral Blur belegte den kleinsten Wert ($M = 42$; $SD = 24,87$).

Die Subskala *Frustration* (gleichnamig) besaß in ihren Boxplots Ausreißer bis 100. Zudem zeigten sich größere Interquartilsabstände bei Scanpath (IQR = 30) und Baseline (IQR = 28,75) im Vergleich zu Peripheral Blur (IQR = 23,75). Die Mediane bei Scanpath und Baseline waren gleich ($Mdn = 17,5$) und leicht höher als bei Peripheral Blur ($Mdn = 12,5$).

Beim Total Workload hob sich der Boxplot der Baseline-Kondition von den anderen Konditionen mit einer höheren Streuung nach oben ab, wie an dem oberen Whisker ($\max = 70,83$) und einer breiteren Box (IQR = 23,33; besonders von $Mdn = 27,5$ bis zum dritten Quartil hin) in Abbildung 12(c) abzulesen ist. Scanpath (IQR = 14,38) und Peripheral Blur (IQR = 16,67) waren sich hinsichtlich der oberen Whisker und Interquartilsabstände sehr ähnlich. Die Mediane von Scanpath ($Mdn = 25,83$), Peripheral Blur ($Mdn = 26,67$) und Baseline ($Mdn = 27,5$) lagen sehr nah beieinander. Zudem fanden sich Ausreißer bei Scanpath ($\max = 85$) und Peripheral Blur ($\max = 74,17$).

Bis auf die Subskala für Anstrengung ($p = 0,41$) verletzten alle anderen Subskalen die Normalitätsannahme (Geistige Anforderung: $p = 0,006$; $p < 0,001$ sonst). Somit wurde hinsichtlich der Anstrengung die Normalitäts-Annahme für eine einfaktorielle Varianzanalyse mit Messwiederholung erfüllt, worauf folgend ein Mauchly's Test durchgeführt wurde. Dieser deutete auf keine Verletzung der Sphärizitäts-Annahme hin ($p = 0,86$). Somit konnte eine Varianzanalyse ohne Greenhouse-Geisser Korrektur durchgeführt werden, welche jedoch keine signifikanten Unterschiede in den Anstrengungs-Ratings zwischen den Konditionen bestätigte ($p = 0,73$). Alle Friedman-Tests für die übrigen Subskalen wiesen auf keine signifikanten Unterschiede zwischen den Konditionen hin.

Ebenfalls wurde auf Aspekte innerhalb der Leseassistenten-Funktionen ohne Vergleich mit den jeweils anderen Konditionen untersucht.

5.3.6 Scanpath

Da mit dem Scanpath eine Auseinandersetzung mit dem eigenen Leseverhalten beabsichtigt war, war die Gesamtdauer der Betrachtung der eigenen Scanpaths von Interesse. Daher wurde pro Person die Zeit (in Sekunden) über alle Seitenwechsel hinweg aufsummiert, d.h. die Zeit, die eine Person insgesamt für die Seitenwechsel brauchte, ab dem Zeitpunkt, an dem der Scanpath zu sehen war.

So gab es keine Person, welche den Scanpath auf jeder Seite direkt übersprang. Dies zeigte die Minimal-Gesamtdauer von 13,18 Sekun-

den (9,47 Sekunden für Seite 1). Ausgehend vom Boxplot lag der Median bei fast einer Minute ($Mdn = 57,2$) und der Interquartilsabstand bei 52,33. Ein Ausreißer lag bei 174,09 Sekunden, was fast 3 Minuten entspricht.

Weiterhin war von Interesse, ob ein Zusammenhang zwischen der Gesamtdauer der Scanpath-Betrachtung und der Anzahl richtiger Antworten bestand und wie stark dieser ausgeprägt war. Denn so könnte eine längere Auseinandersetzung mit dem eigenen Scanpath dafür sorgen, dass schon das wiederholte Betrachten der kritischen Regionen (gelbe bis rote Kreise und/oder Rücksprünge der Augen) im Text dafür sorgt, dass Inhalte besser erinnert und/oder sogar verstanden werden. Auch gerade das Fehlen oder eine geringe Anzahl solcher Hotspots (vereinzelte Pfade aus grünen Kreisen oder große Lücken in der Visualisierung) könnte für die Leser:innen ein Indiz dafür sein, dass eine Textstelle zu schnell bzw. unaufmerksam überflogen wurde. Dies kann wiederum zu einem Reflektieren über den Inhalt animieren, was letztendlich zu einer höheren Anzahl richtiger Antworten beitragen könnte.

Ein Shapiro-Wilk-Test auf den Gesamtdauern war signifikant ($p = 0,005$), was die Normalitäts-Annahme einer Pearson-Korrelation verletzte, weswegen eine Spearman-Rang-Korrelation auf der Gesamtdauer und der Anzahl richtiger Antworten berechnet wurde. Diese war nicht signifikant ($r_s = 0,31$; $p = 0,1$).

Aus den gleichen Gründen wie für die Gesamtdauer war auch die Häufigkeit, mit der die Scanpaths insgesamt aufgerufen wurden, hinsichtlich der tieferen Auseinandersetzung mit dem Textinhalt von Interesse. Denn so könnte auch die Beobachtung des eigenen Leseflusses über die Zeit zu einer Anpassung im Leseverhalten und einer Verbesserung in Verständnis und Erinnerung des Inhalts führen. Ebenso würde der Text wiederholt gelesen werden und sich somit besser einprägen.

Für die Gesamtanzahl an angezeigten Scanpaths wurden pro Proband:in die Anzahl der Zeilen gezählt, die sie mit jedem Seitenwechsel in der zugehörigen CSV-Datei generierten, abzüglich der Zeilen, welche eine Scanpath-Zeit von 0 Sekunden enthielten, da diese einen Seitenwechsel auf eine vorherige Seite bedeuteten, welcher ohne Verzögerung durch die Visualisierung stattfindet.

Eine minimale Häufigkeit von 3 impliziert, dass es mindestens eine Person gab, die pro Seite nur die Mindestanzahl an Scanpaths visualisierte (für die vierte Seite musste der Scanpath nicht zwangsläufig sichtbar gemacht werden, um das Experiment fortzusetzen, da mit der vierten Seite bereits der Button zum Beenden der Kondition erschien). Im Boxplot war ein Ausreißer ($\max = 12$) zu beobachten. Der Median betrug 4 und der Interquartilsabstand 1,75.

Abschließend wurde untersucht, ob und wie stark ein Zusammenhang zwischen der Häufigkeit des Sichtbarmachens der Scanpaths und der Anzahl richtig beantworteter Fragen vorliegt. Ein Shapiro-Wilk-Test auf der Gesamtzahl sichtbar gemachter Scanpaths war

signifikant ($p < 0,001$), was die Normalitäts-Annahme für eine Pearson-Korrelation verletzte. Eine Spearman-Rang-Korrelation erbrachte kein signifikantes Ergebnis ($r_s = 0,24$; $p = 0,2$).

5.3.7 Peripheral Blur

Da in der Kondition Peripheral Blur frei entschieden werden konnte, wieviele Zeilenhöhen der anvisierte Bereich enthalten sollte, war es erst einmal von Interesse, ob es Überschneidungen in der Präferenz für die Einstellungen gab. Bei der Betrachtung der ersten Datensätze stellte sich heraus, dass alle Einstellungen bereits in der Testrunde vorgenommen wurden, was auf den Umstand zurückgeführt werden kann, dass die letzte Zeileneinstellung aus der Testrunde für die eigentliche Runde gespeichert wird. Demnach implizieren die Datensätze, dass die Proband:innen mit dieser Einstellung zufrieden waren und sie in der nachfolgenden Runde nicht mehr verändern mussten. Daher sollte die Historie schon während der Testrunde erhoben werden. Für die Auswertung wurde pro Person der letzte Einstellungs-Wert extrahiert, da die meisten Einstellungen auf der ersten Seite gemacht wurden. Die kleinste Einstellung lag bei 1 Zeile, das Maximum bei 8 Zeilen. Abbildung 13 können die Häufigkeiten pro Zeileneinstellung entnommen werden. Hieran zeigt sich, dass 3 ($N = 11$) und 4 ($N = 10$) scharfgestellte Zeilen die häufigste Einstellung war, die zuletzt vorgenommen wurde. Die Einstellung von 2 scharfgestellten Zeilen belegte mit 5 Personen den dritten Platz. 1 scharfgestellte Zeile (Default-Einstellung) mit 2 Personen und 8 scharfgestellte Zeilen mit einer Person bildeten die Ausreißer. Darüber hinaus lag der Median bei 3 und der Interquartilsabstand bei 1.

Hinsichtlich des Zusammenhangs zwischen Anzahl scharfgestellter Zeilenhöhen und richtiger Antworten sind zwei verschiedene Richtungen möglich (wenn ein solcher überhaupt existieren sollte): Während auf der einen Seite weniger scharfgestellte Zeilen mehr visuellen Clutter in der Peripherie reduzieren und somit den Fokus auf die momentane Textstelle erhöhen könnten, könnte der Blur auf der anderen Seite selbst eine Ablenkung in der Peripherie darstellen, weshalb mehr scharfgestellte Zeilen zu einem erhöhten Fokus auf die aktuelle Textstelle beitragen könnten und somit zu mehr richtigen Antworten. Um auf die Signifikanz, Stärke und Richtung dieses Zusammenhangs zu untersuchen, wurde aufgrund der verletzten Normalitäts-Annahme (Shapiro-Wilk nicht signifikant: $p < 0,001$) eine Spearman-Rang-Korrelation berechnet. Diese war sehr schwach (ausgehend von dem guide von Kumar und Reddi [36]) und negativ, aber nicht signifikant ($r_s = -0,15$; $p = 0,42$).

Auch die Anzahl an gemachten Zeileneinstellungen wurde untersucht. Pro Person wurde für jede Seite (und evtl. Wiederholungen dieser Seite) der letzte Wert des Einstellungs-Zählers extrahiert und zu einem Wert aufaddiert. Da die letzte Zeileneinstellung aus der Testrunde für die richtige Runde gespeichert wurde, wurden auch die Werte aus der Testrunde in die Rechnung miteinbezogen. Die minimale Anzahl an Einstellungen lag bei 0, was sich mit der einen Person erklären lässt, welche die default-Einstellung von einer Zeile bis zum Ende nicht veränderte. Die Ausreißer (max = 67 Einstellungen) im zugehörigen Boxplot zeigen, dass die Grenzen nach oben ausgetestet, aber nicht als finale Einstellung beibehalten wurden, wie die deskriptive Statistik im Vorherigen zeigte. Weiterhin betrug der Median 5,5 und der Interquartilsabstand 8,25.

Auch hier wurde auf die Signifikanz und Stärke des Zusammenhangs von Anzahl der Zeileneinstellungen und Anzahl richtiger Antworten untersucht. So könnten z.B. mehr Einstellungen davon zeugen, dass der scharfgestellte Bereich in seiner Größe flexibel an verschieden große Sinnabschnitte (z.B. Absätze) angepasst wurde. Somit würden die Leser:innen beim strukturierten Lesen des Textes unterstützt werden und demnach den Text besser verinnerlichen. Auf der anderen Seite könnten viele Einstellungen auch bedeuten, dass die Grenzen der Funktion ausgetestet wurden. Ein Shapiro-Wilk-Test auf der Anzahl der Zeileneinstellungen war signifikant ($p < 0,001$). Aufgrund

der verletzten Normalitäts-Annahme wurde eine Spearman-Rang-Korrelation berechnet, welche nicht signifikant war ($r_s = 0,2$; $p = 0,29$).

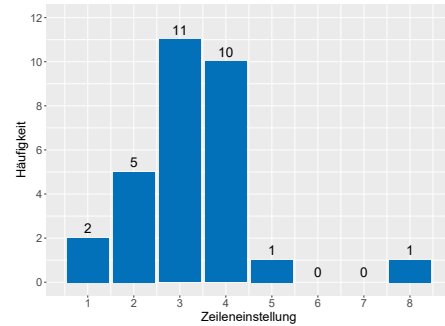


Abbildung 13: Kondition Peripheral Blur: Häufigkeiten für die letzte Zeileneinstellung.

5.4 Qualitative Analyse

Auf den Daten aus dem Fragebogen nach der Studie in VR wurde eine zusammenfassende, qualitative Inhaltsanalyse durchgeführt. Hierfür wurden die expliziten Aussagen der Proband:innen in Schrift gesichtet, kategorisiert und paraphrasiert. Die inhaltliche Kategorisierung der Antworten fand auf zwei Ebenen statt. Als erstes wurde das Feedback zu den beiden Leseassistenten-Funktionen auf saliente Muster hinsichtlich der Nützlichkeit untersucht, aus denen sich folgende Kategorien ergaben:

- **Irritierend / einschränkend**

In diese Kategorie fielen Aussagen, welche Aspekte der Leseassistenten-Funktionen enthielten, die das Leseerlebnis verschlechterten.

- **Nicht hilfreich / keine Wirkung**

Diese Kategorie umfasst Aussagen, welche die Leseassistenten-Funktionen als weder nützlich noch einschränkend beschreiben.

- **Nicht hilfreich aber Funktion verstanden**

Aussagen wurden dieser Kategorie zugeordnet, wenn sie das gleiche Kriterium der vorherigen Kategorie erfüllten, aber zusätzlich aus ihnen hervorging, dass die eigentliche Absicht der Leseassistenten-Funktion nachvollzogen wurde, auch wenn diese für die eigene Leseerfahrung keinen Mehrwert bot.

- **Hilfreich**

Hierzu zählten alle Aussagen, die den Leseassistentenfunktionen positive Aspekte zuschreiben, auch wenn diese nicht der primären Idee der Leseassistenten-Funktionen entsprachen.

Für jede der Kategorien wurde eine Farbe gewählt und bei jeder Person die Textabschnitte, die inhaltlich einer Kategorie zuzuordnen waren, in der zugehörigen Farbe markiert. Anschließend wurden für jede Kategorie pro Leseassistenten-Funktion alle Proband:innen zusammengezählt, bei denen die zugehörige Farbe mindestens einmal vorkam. Der folgenden Tabelle sind die Ergebnisse dieser Auszählung zu entnehmen:

Feedback-Kategorie	Leseassistentz-Funktion	Anzahl Personen
Irritierend / einschränkend	Scanpath	18
	Peripheral Blur	22
Nicht hilfreich / Keine Wirkung	Scanpath	13
	Peripheral Blur	0
Nicht hilfreich aber Funktion verstanden	Scanpath	9
	Peripheral Blur	0
Hilfreich	Scanpath	15
	Peripheral Blur	27

Tabelle 1: Allgemeine Kategorisierung des Feedbacks nach Nützlichkeit. Es wurde pro Funktion jede Person gezählt, bei der die entsprechende Feedback-Kategorie mindestens einmal vertreten war.

Somit entstand ein erstes grobes Bild, aus welchem sich ablesen ließ, dass fast alle Proband:innen Angaben machten, die für die Nützlichkeit des Peripheral Blur sprechen, während beim Scanpath diese Antwort-Kategorie nur bei der Hälfte aller Personen wiederzufinden war – fast genauso viele (aber weniger) wie diejenigen, aus deren Antworten hervorging, dass der Scanpath beim Lesen Irritation hervorruft. Auch die neutralen Kategorien (mit „nicht hilfreich...“) waren im Feedback zum Peripheral Blur nicht vertreten, dafür vereinzelt in den Antworten zum Scanpath. Auf der anderen Seite fanden sich bei der Mehrheit Aspekte zum Peripheral Blur, die dafür sprechen, dass die Leseaufgabe mit dieser Funktion am einschränkendsten empfunden wurde, mehr noch als beim Scanpath.

Da die Vorerfahrung mit VR-Systemen für das Lesen in VR von Relevanz sein kann, wurde untersucht, wie die Leseassistentz von den verschiedenen Gruppen nach Nutzungshäufigkeit aufgenommen wurde. So wurde die VR-Erfahrung bzw. die Nutzungshäufigkeit von VR in die Stufen „Viel“, „Wenig“ und „Keine“ zusammengefasst. Dabei galt eine Nutzung von mehr als einem Mal pro Woche als viel Vorerfahrung, die monatliche Nutzung als mittlere Vorerfahrung, eine Nutzung bis einschließlich zu einem Mal im Jahr als wenig Vorerfahrung und eine Nutzung unter einem Mal pro Jahr als keine Vorerfahrung. Pro Erfahrungsstufe wurde jeweils ausgezählt, wie viele Personen die entsprechende Leseassistentz-Funktion hilfreich, teilweise hilfreich und nicht hilfreich fanden. Daraus ergab sich folgende Übersicht:

Erfahrung mit VR	Leseassistentz-Funktion	Anzahl hilfreich	Anzahl teilweise hilfreich	Anzahl nicht hilfreich
Viel	Scanpath	2	0	0
	Peripheral Blur	2	0	0
Mittel	Scanpath	0	2	3
	Peripheral Blur	2	1	2
Wenig	Scanpath	2	2	3
	Peripheral Blur	4	2	1
Keine	Scanpath	4	1	11
	Peripheral Blur	7	5	4

Tabelle 2: Für jede Erfahrungsstufe mit VR: Pro Leseassistentzfunktion wurden die Personen gezählt, welche die Funktion hilfreich, teilweise hilfreich und nicht hilfreich fanden.

Es zeigt sich auch hier eine deutliche Dominanz des Peripheral Blur hinsichtlich der wahrgenommenen Nützlichkeit gegenüber dem Scanpath ab der mittleren Erfahrungsstufe. Auch mehr Personen mit wenig bis keiner VR-Erfahrung fanden den Peripheral Blur hilfreicher als den Scanpath. Zudem wurde er von mehr Personen aus dieser Nutzergruppe als hilfreich empfunden als von Personen aus der gleichen Gruppe, die ihn nicht hilfreich fanden, wohingegen mehr Personen mit wenig VR-Nutzung den Scanpath als nicht hilfreich einschätzten als diejenigen aus der Gruppe, die ihn hilfreich fanden. Interessant war auch zu sehen, dass beide Personen mit viel VR-Erfahrung beide Funktionalitäten als hilfreich empfanden.

Um ausgehend von der ersten Ebene der allgemeinen Kategorisierung den Schritt zur zweiten Ebene zu machen, wurde in den Antworten nach konkreteren Kategorien gefiltert und die Aussagen im Wortlaut gesammelt, ausgezählt und für eine verbesserte Lesbarkeit korrigiert.

So ergaben sich einmal für die Scanpath-Funktionalität folgende Themen-Cluster, welche die positiven wie auch negativen Aspekte der Leseassistentz-Funktionen wiedergaben:

Markierungen (Positiv). Wie ein Echo zu den Lesestrategien (s. 5.2) gaben 6 Personen an, den Scanpath als einen Textmarker zu bewerten und/oder als solchen zu verwenden:

„Den Blickverlauf empfand ich als hilfreich bei der Textwiederholung und habe diesen ähnlich wie Markierungen nutzen können.“ (VP 24)

Eine Personen ging sogar noch einen Schritt weiter und sprach von „mentalem Anmarkern“, d.h. ein momentanes, gedankliches Hervorheben relevanter Inhalte:

„[...] der Fokusreview [hat] meinen Fokusverlauf gezeigt und so praktisch mentalem [A]nmarkern geähnelt.“ (VP 4)

Näher konnten die gelben bis roten Hotspots die Leser:innen automatisch an schwierige oder unbekannte Wörter erinnern, da die Augen länger auf diesen verweilten:

„[...] Die gelben und roten Hotspots waren hilfreich für Fremdwörter und Eigennamen[,] um diese schnell wiederzufinden[...]“ (VP 21)

Demnach wurde der eigene Scanpath als 1:1-Abbildung der persönlichen Areas of Interest (AOI) im Text verstanden, wodurch Lesen und Markieren in einen Vorgang zusammenfallen. Und obwohl diese Funktion eine mehr Feedback-basierte Form der Leseassistentz darstellt, geht somit aus den Antworten der Proband:innen hervor, dass der Scanpath zu einer bereits etablierten Lesestrategie zweckentfremdet werden konnte.

Über Leseverhalten reflektieren / Optimierung (Positiv). 9 Personen gaben an, dass der Scanpath indirekt zu einem erneuten und mehr reflektierten Lesen animierte:

„[...] Es hat mir eher nochmal Zeit gegeben, um mir darüber Gedanken zu machen, ob ich wirklich alles verstanden habe, bevor ich 'gedankenlos' weiterlese.[...]“ (VP 15)

„[...] ich hatte das Gefühl, dass mir die Textabschnitte[,] die ich nicht direkt verstanden habe[,] mir rot angezeigt wurden und ich mir diese dann nochmal genauer anschauen konnte.“ (VP 13)

Auch hier lässt sich ein Rückbezug zu den in 5.2 genannten Lesestrategien herstellen, da der Scanpath von den Personen als Impulsgeber zum nochmaligen Lesen gesehen wurde.

Neben dem Anreiz, sich näher mit dem Text auseinanderzusetzen, wurde laut einer von besagten Personen ebenfalls der Anspruch an die eigene Lese-Effizienz erhöht. Dies schrieb sie dem Wunsch zu, einen Scanpath zu generieren, welcher kürzere Fixationszeiten und weniger Rücksprünge im Text abbildet. Aus diesem Grund verglich sie die Funktion mit einem „Lesecoach“:

„Es war wie ein Lesecoach, der einem auf die Augen geschaut hat, dass man richtig liest [...] weil man den [A]nspruch hatte, ohne [R]ücksprünge oder langes [V]erweilen auf einem [W]ort den [T]ext zu lesen“ (VP 14)

Demnach würde diese Funktion den Leser:innen eine Art Benchmark für den eigenen Lesefluss bieten, auf dessen Basis eine Optimierung stattfinden kann.

Seitenwechsel unterbricht Lesefluss (Negativ). 8 Personen gaben an, dass der Seitenwechsel für Verwirrung sorgte. Dadurch, dass die Scanpath-Funktionalität so konzipiert war, dass der Button zum Weiterblättern zunächst den Scanpath für die bisher gelesene Seite sichtbar machte, wurde die Erwartung verletzt, direkt weiterlesen zu können. Dies führte bei den Personen zu Irritiertheit und einem unterbrochenen Lesefluss. Verstärkt wurde dies dadurch, dass sich der letzte Satz einer Seite auf der nächsten fortsetzte, was durch Anzeigen des Scanpaths verzögert wurde:

„[...] ich fand den einsehbaren Blickverlauf nicht hilfreich. Eher umständlich, da der Satz am Ende der Seite noch nicht zu Ende war und man dann erst einmal unterbrochen wird, bevor man weiter lesen kann. Dadurch war ich jedes mal wieder aus dem Thema raus.“ (VP 16)

Fokus auf den Inhalt geht verloren, Verbesserungsdruck (Negativ). Bei 9 Personen animierte der Scanpath ebenfalls zum Nachlesen des Textes, allerdings wurde hierbei der Anreiz nicht als intrinsisch wahrgenommen, sondern vielmehr als extrinsisch, d.h. einen möglichst Hotspot-freien Scanpath ohne Rücksprünge zu generieren. Statt den konstruktiven Effekt eines „Lesecoach“ wahrzunehmen, empfanden die Personen einen Leistungsdruck, wodurch der Inhalt der Texte vernachlässigt wurde:

„[E]s war interessant zu sehen, wo ich Rücksprünge gemacht habe und auf welche Wörter ich länger geschaut habe, es hat aber auch ein bisschen den Druck aufgebaut, schnell und ohne Zurückblicken durch den Text zu kommen.“ (VP 12)

„Ja, er hat zum Nachlesen animiert. Jedoch nicht um des Inhalts wegen, sondern vielmehr, um ein 'besseres Ergebnis' (möglichst wenig rote oder stark grüne Felder) zu erzielen.“ (VP 28)

Irritiert durch Pfeile/Rücksprünge im Text. (Negativ.) 12 Personen gaben im Feedback zum Scanpath an, dass die Pfeile vom eigenen Leseempfinden abwichen:

„[...] Vor allem habe ich mich gefragt, warum ich oft von ziemlich weit unten nochmal ganz nach oben an den Anfang gesprungen bin. Das entspricht nicht meinem Leseempfinden“ (VP 9)

Auch für den Peripheral Blur zeigten sich thematische Überschneidungen im positiven und negativen Feedback der Proband:innen.

Präferenz in der Einstellung (Positiv). Was sich bereits in den quantitativen Daten bezüglich der Zeileneinstellungen zeigte,

spiegelt sich in den Antworten von 4 Personen wieder, welche von sich aus unaufgefordert die Angabe machten, 3 bzw. 4 Zeilen als Einstellung zu favorisieren. 2 Personen sahen 3 Zeilen als die beste Einstellung, da dies als Mindestanforderung für ein ergonomisches Leseempfinden wahrgenommen wurde, bei dem der Kontext nicht verloren geht, während 2 weitere Personen 4 scharfgestellte Zeilen bevorzugten, da diese ihnen dabei half sich auf den aktuellen Textabschnitt zu fokussieren, ohne dass die Augen vom umliegenden Text abgelenkt oder zu sehr beansprucht wurden:

„[...] ich [habe] immer mindestens drei Zeilen in Blick genommen. Eine Zeile war mir zu wenig, um das Gefühl zu haben, im Kontext zu lesen.“ (VP 8)

„[...] dadurch[,] das[s] der Fokus bei mir nur auf 4 Zeilen war, konnte ich mich besser konzentrieren und meine Augen ein wenig entspannen“ (VP 9)

Verhindert Ablenkung durch umliegenden Text, hilft beim Fokussieren (Positiv). Ein Aspekt, der im letzten Punkt vorweggegriffen wurde, ist die Wahrnehmung eines visual aid, um clutter auszublenden, was wiederum eine Ablenkung bzw. Belastung der Augen reduziert und damit den Fokus auf aktuelle Textstellen schärft. 19 Personen gaben an, dass der Peripheral Blur ihnen half, sich auf den aktuellen Textabschnitt zu konzentrieren oder sich nicht vom umliegenden Text ablenken zu lassen:

„Es hat mir beim Fokus[s]ieren auf die Textstelle sehr geholfen, weil es oft auch vorkommt, dass ich einfach ein paar Zeilen überspringe und da kurz ein paar Wörter lese, bevor ich bei meiner eigentlichen Textstelle weiterlese“ (VP 15)

„Super hilfreich. Ich tendiere dazu, eine Seite aus dem Augwinkel zu scannen und besonders lange und interessante Wörter können mich dabei ablenken. Da war es super, dass ich nur die aktuellen 'words of interest' erkennen konnte.“ (VP 20)

Laut diesen Aussagen würde der Peripheral Blur somit dem Problem der eingeschränkten Aufmerksamkeit entgegenwirken, das von den Proband:innen am häufigsten ($N = 14$) angegeben wurde.

Verhindert Verrutschen in der Zeile (Positiv). Eine weitere Leseproblematik, die in 5.2 angesprochen wurde, stellt das Verrutschen in der Zeile dar. Von 4 Personen wurde der Peripheral Blur als passendes Hilfsmittel gesehen:

„Der Zeilenfokus kann helfen[,] nicht in der Zeile zu verrutschen, das passiert beim normalen Lesen sonst häufiger“ (VP 14)

Um auch hier eine Analogie zu den Lesestrategien aus 5.2 herzustellen, kann das sequentielle Scannen der Zeilenblöcke mit dem strukturierten Lesen verglichen werden, bei dem Abschnitt für Abschnitt gelesen und zwischendrin Pausen eingelegt werden.

Überraschenderweise empfanden auch die Personen, welche in der deskriptiven Statistik Ausreißer hinsichtlich der Zeileneinstellungen darstellten ($N = 3$), den Zeilenfokus als hilfreich. 2 Personen, die bis zum Ende die Default-Einstellung von einer Zeile verwendeten, gaben an, sich besser fokussieren zu können, da sie sonst sehr ungenau lasen; es stabilisierte sozusagen ihren Blick.

Kopfbewegung verursacht Anstrengung (Negativ). Schon bei der allgemeinen Kategorisierung des Feedbacks zu den Funktionalitäten bezüglich ihrer Nützlichkeit zeichnete sich eine Dualität in der Rezeption des Peripheral Blur ab. So wurde dieser auf der einen Seite zwar mehrheitlich ($N = 27$) als hilfreich

empfundene, auf der anderen Seite ging aus dem negativen Feedback und seiner Häufigkeit eindeutig hervor, dass der Peripheral Blur auch als die Funktion wahrgenommen wurde, die beim Lesen die größte Anstrengung verursachte ($N = 22$). Zurückgeführt wurde dies auf die Scan-Bewegung, die mit dem Kopf samt befestigtem HMD durchgeführt wurde, um die scharfgestellten Zeilenblöcke in Sequenz zu lesen ($N = 12$):

„[...] Aufgrund des HMDs musste ich meinen Kopf unangenehm in den Nacken legen oder nach vorne kippen, um die jeweiligen Zeilen sehen zu können. Deshalb habe ich teilweise am Anfang oder Ende der Seite mehr Zeilen gebraucht, als in der Mitte, um den Winkel, mit dem ich meinen Kopf bewegen musste, etwas zu verkleinern.“ (VP 23)

Die gleiche Person bildete den größten Ausreißer im Bezug auf die Anzahl der Zeileneinstellungen, was sich jedoch nicht auf den angegebenen Grund der Person zurückführen ließ, wie die Videoaufzeichnung zeigte. Die meisten ihrer Einstellungen waren auf der gleichen Seite und in schneller Abfolge unabhängig der Kopfausrichtung.

„[...] Die Schwere des Headsets hat die Anstrengung nur verstärkt. Es fühlte sich insgesamt unnatürlich und unnötig an.“ (VP 21)

Vor allem in den Extremen hinsichtlich der Zeilen-Position wurde die Kopfbewegung als anstrengend empfunden:

„[...] Deshalb habe ich teilweise am Anfang oder Ende der Seite mehr Zeilen gebraucht, als in der Mitte, um den Winkel, mit dem ich meinen Kopf bewegen musste, etwas zu verkleinern.“ (VP 23)

Selbst Personen, die Angaben machten, die nicht explizit auf eine wahrgenommene Anstrengung hindeuten, konnten sich vorstellen, dass die gleiche Leseaufgabe bei längeren Texten zur Anstrengung führen könnte ($N = 3$):

„Nein, der Text war kurz genug, dass es nicht anstrengend war. Ich könnte mir allerdings vorstellen, dass es bei längeren Texten anstrengend werden könnte.“ (VP 8)

Kopf- statt Augenbewegung beim Lesen fühlt sich unnatürlich an (Negativ). Neben der Wahrnehmung einer unmittelbaren, körperlichen Belastung bestand ebenfalls das Empfinden, dass das aktive Nachjustieren des Kopfes zum Scharfstellen des nächsten Zeilenblocks den intuitiven Lesefluss der Augen beeinträchtigt ($N = 7$). Dies bremste eine Person beim Lesen aus und führte bei ihr sogar zu Ausreißern in der SSQ-Dimension der Okulomotorik, Desorientierung und im Gesamtscore:

„[...] Ich habe sowieso selten das Problem zwischen Zeilen aus Versehen zu springen und habe die verschwommenen Zeilen außerhalb meines Blickfeldes [sic] eher als störend wahrgenommen, weil ich nicht nur meinen Blick schweifen lassen konnte, sondern meinen ganzen Kopf bewegen musste.“ (VP 19)

Insbesondere bei den Extremen der Zeilenhöhe wurde bei einer von besagten Personen der Impuls beschrieben, einfach mit den Augen zum Ziel zu springen, statt eine längere Kopfbewegung durchzuführen:

„[...] ich hätte es als angenehmer empfunden[,] wen[n] durch Eyetracking die Zeilen fokussiert gewesen wären, auf denen meine Augen liegen und nicht den ganzen Kopf bewegen

zu müssen. Insbesondere oben und unten bei den Extremen.“ (VP 13)

Während der Peripheral Blur als Hilfe zur Orientierung im Text gedacht ist, kann die gleiche Assistenzfunktion bei Leser:innen, die sich schnell einen Überblick über den Text verschaffen, den gegenteiligen Effekt haben, wie der Bericht einer Person wiedergab:

„der blurr [sic] gefühl [sic] mir persönlich nicht[,] weil ich mir beim [L]esen einer [S]eite ein grobes [B]ild mache[,] wo welche [I]nfo zu finden ist, das hat der blurr [sic] sehr erschwert“ (VP 17)

Gleiche Person bildete den Ausreißer hinsichtlich der Lese-wiederholung bei Peripheral Blur, was nahelegt, dass diese wahrgenommene Einschränkung zu wiederholtem Nachlesen führte.

3 Personen wägen Vor- und Nachteile der Funktionalitäten gegeneinander ab. Dabei wurde wieder das allgemeine Muster wiedergegeben, dass der Peripheral Blur am hilfreichsten beim Lesen empfunden wurde, während der Scanpath als weniger hilfreich aber dafür interessanter wahrgenommen wurde:

„ich fand die Funktionalität mi[t] der Fixationszeit spannender, allerdings nicht hilfreicher beim Lesen, da hat mir eher die Funktionalität mit dem Weichzeichnen geholfen.“ (VP 11)

Darüber hinaus gab es auch Verbesserungsvorschläge zu den beiden Leseassistenten-Funktionen.

Zu dem Scanpath äußerten sich 7 Personen:

- **Erhol-Phase**

Es wurde sich von einer Person eine Art „cooldown“-phase zu Beginn einer Seite gewünscht, sodass das Korrigieren der Augen zur ersten Zeile unmittelbar nach dem Seitenwechsel nicht als ein Rücksprung gewertet wird. Damit würde der erste Pfad aus Rückpfeilen ausbleiben und nicht für Verwirrung sorgen.

- **Schusterjungen vermeiden**

Eine Person machte die Empfehlung, nur ganze Sätze am Seitenende zuzulassen, sodass der Scanpath erst nach Beendigung des letzten Satzes zu sehen ist.

- **Feedback nach Sinnabschnitten**

Zwei Personen führten den vorherigen Punkt weiter aus und machten den Vorschlag, den Scanpath nicht nur nach Beendigung des letzten Satzes einer Seite, sondern überhaupt nach jedem Sinnabschnitt anzuzeigen.

- **Seitenübergreifende Speicherung**

Eine Person gab Einblick in ihren Lesestil, welcher sich darin zeigt, den gesamten Text vorerst einmal durchzulesen um anschließend mehrere Seiten zurückzublättern, um sich vorherige Interessensregionen im Text nochmal anzuschauen. Für diesen Lesestil würde es sich laut der Person besser eignen, den Scanpath eines Abschnittes (aufbauend auf dem vorherigen Punkt) global zu speichern, um später nach dem Lesen weiterer Seiten zu diesem Abschnitt zurückzukehren und dort die früheren Markierungen einsehen zu können. Denn die bisherige Implementation sieht beim Seitenwechsel ein Verwerfen der letzten Scanpath-Parameter vor bzw. ein Überschreiben des Scanpath für die gleiche Seite beim wiederholten Lesen, was ein späteres Einsehen des letzten Scanpath einer Seite beim Zurückblättern verhindert.

- **Dauerhaftes Overlay**

Knüpft an die seitenübergreifende Speicherung an, nur dass der

letzte Scanpath einer Seite während des Nachlesens weiterhin sichtbar bleibt.

- **Anwendung bei fremdsprachigen Texten**

Aufgrund der Eigenschaft, dass Fremdwörter zu Hotspot-Regionen im Scanpath führen können (wie bereits im positiven Feedback erwähnt), sah eine Person den Scanpath für das Lesen von fremdsprachigen Texten als besser geeignet an.

- **Kalibrierung**

Eine Person ging von variierenden Fixationsdauern zwischen Personen aus und schlug demnach eine Kalibrierung der minimalen Fixationsdauer vor.

Für den Peripheral Blur zeigte sich in den Verbesserungsvorschlägen von 9 Personen ein homogeneres Muster, da die Bedienung dieser Funktion von den meisten Personen als sehr anstrengend und intuitiv wahrgenommen wurde:

- **Modalität / Sensibilität der Eingabe verändern**

5 Personen vertraten die Meinung, dass die Scanbewegung durch eine andere Eingabemodalität erfolgen sollte oder zumindest in der Empfindlichkeit erhöht werden (d.h. weniger Bewegung für Zeilensprung nötig), um das Leseerlebnis natürlicher und ergonomischer zu gestalten. Am häufigsten wurde der Bewegungsinput der Augen gefordert. Auch hier ging wieder die Anstrengung beim Lesen der ersten und letzten Zeilen hervor. In einem Fall wurde aber auch der klassische Cursor über den Raycast des Controllers vorgeschlagen.

- **Fokus an Textstruktur anpassen**

Analog zum Anzeigen des Scanpath nach Sinnabschnitten machten 2 Personen den Vorschlag, den Peripheral Blur in einer Form anzuwenden, die sich mehr an der Textstruktur orientiert. Dabei soll die Größe des fokussierten Bereichs statt einer fixen Anzahl aus Zeilenhöhen variabel an die Länge jedes Satzes angepasst werden, auf dem zur Zeit der Fokus liegt. Alternativ wurde ein Kompromiss genannt, welcher eine fixe Zeilenanzahl und einen übergangslosen Sprung zum nächsten Zeilenblock kombiniert.

- **Andere Ansicht**

Eine Person wünschte sich weiterhin eine Ansicht, bei der immer 2 aufeinanderfolgende Seiten nebeneinander angezeigt werden.

Darüber hinaus wurde im Allgemeinen eine Einstellung für die Textgröße gefordert und eine Berücksichtigung des Seitenwechsels im Hinblick auf die Beendigung des letzten Satzes für einen besseren Lesefluss betont.

6 DISKUSSION

6.1 Simulator Sickness

Keine Person brach die Studie ab, weder aus Gründen der Simulator Sickness Symptomatik noch aus anderen Gründen. Nur eine Person gab explizit im Post-Studien-Fragebogen an, dass sich bei ihr während der Peripheral Blur Kondition Anzeichen von Übelkeit zeigten. Vor allem bei dieser Kondition war eine solche Beobachtung im Vergleich zu den anderen Konditionen am ehesten zu erwarten, da nur der Peripheral Blur eine aktive Auf- und Abbewegung des Kopfes zum Scannen der Zeilenblöcke erforderlich machte und zudem einen Effekt verschwommenen Sehens simulierte. Jedoch bestand bei dieser Kondition wenn überhaupt für die okulomotorischen Items das Potenzial einer Verschlechterung gegenüber der Baseline-Kondition, da der VR-Teil komplett im Sitzen (auf einem Stuhl ohne Rollen) stattfand. Allerdings wurde im Vorfeld auch auf eine okulomotorische Belastung kontrolliert, indem Pausen zwischen Test- und eigentlicher Runde gemacht wurden, in denen sich

die Augen erholen konnten. Tatsächlich besaß die Kondition den kleinsten Interquartilsabstand in dieser Dimension und streute bis auf einen Ausreißer am wenigsten nach oben. Der Ausreißer könnte laut Angabe der zugehörigen Person daher rühren, dass das Fokussieren auf den Text durch die Kopffrotation statt durch die Augen als behindernd wahrgenommen wurde. Letztendlich zeigten sich zwischen den Konditionen hinsichtlich der SSQ-Symptomatik keine signifikanten Unterschiede, sodass das Studiendesign zumindest in Relation zur Baseline-Kondition, welche in der ursprünglichen Idee einer AR-Anwendung dem realen Lesen ohne Augmentationen entspricht, zu keiner Verschlechterung der Symptome führte. Insbesondere konnte somit auch keine Verschlechterung der okulomotorischen Symptomatik im Vergleich zur Baseline-Kondition angenommen werden.

6.2 NASA RAW-TLX

Der NASA RAW-TLX zeigte nur in den Subskalen Leistung (*Performance*), Anstrengung (*Effort*) und Total Workload visuelle Auffälligkeiten. Bezüglich der Anstrengungs-Subskala streuten beide Leseassistenten-Funktionen deutlich weiter nach oben als die Baseline-Kondition. Im Hinblick auf die Anstrengungs-Subskala wiesen beide Leseassistenten-Funktionen eine deutlich höhere Streuung nach oben auf als die Baseline-Kondition. Beim Scanpath ließe sich diese Beobachtung auf die angesprochene Unterbrechung des Leseflusses durch das unerwartete Anzeigen des Scanpath zurückführen, was vermutlich noch dadurch verstärkt wurde, dass die letzten Sätze der Seite abgeschnitten waren, wodurch das Beenden des Satzes auf der nächsten Seite durch Anzeigen des Scanpath verzögert wurde (Schusterjunge). Auch das Gegenteil der erhofften Wirkung einer intrinsisch motivierten Anpassung des Leseverhaltens trat ein, da laut den Angaben von 9 Personen das Einsehen des eigenen Scanpath den Druck aufbaute einen „erwünschten“ Scanpath zu generieren, der kaum rote Hotspots und/oder Rücksprünge enthält. Diese Wahrnehmung kann zu Stress und somit auch Anstrengung geführt haben. Auch die höheren Anstrengungs-Ratings der Kondition Peripheral Blur könnten sich mit den qualitativen Daten der Proband:innen erklären lassen. Denn obwohl der Peripheral Blur mehrheitlich als die hilfreichere Leseassistenten wahrgenommen wurde, ging aus dem Meinungsbild ebenso deutlich hervor, dass diese Funktion aufgrund ihrer Steuerung mit dem Kopf die größte Anstrengung verursachte. So fühlte sich die Kopfbewegung zur Höhenanpassung des Fokus laut den Aussagen mehrerer Proband:innen ungewohnt und unnatürlich an, da nicht die Augen das steuernde Element waren und der Kopf aktiv zum Lesen bewegt werden musste. Die wahrgenommene Anstrengung bei der Nutzung des Peripheral Blur kann sich auch auf die Wahrnehmung der eigenen Leistung ausgewirkt haben. Hierauf deuten die Ausreißer (bis 100) hin, wobei größere Werte eine negativere Selbsteinschätzung der Leistung bedeuten. Trotz dieser subjektiven Angaben und ihren Implikationen konnten für keine der Subskalen signifikante Unterschiede in den Ratings zwischen den Konditionen nachgewiesen werden.

6.3 Inhaltlicher Abruf der Texte

Bezüglich des inhaltlichen Abrufs der Texte (*Recall*) wurde durch die Verwendung des Leseassistenten eine höhere Anzahl richtiger Antworten angenommen (**H1(a)**, **H1(b)**). Es wurden Fragen zu Anfang, Mitte und Ende des Textes gestellt, um auf Primär und Rezenz-Effekte zu kontrollieren. Innerhalb der drei Fragen (ohne Differenzierung nach Textzugehörigkeit) konnten keine signifikanten Unterschiede zwischen den Konditionen hinsichtlich der Anzahl richtiger zu falschen Antworten gefunden werden. Auffällig hier war nur, dass bei der ersten von drei Textfragen insgesamt die meisten richtigen Antworten gegeben wurden, während die meisten falschen Antworten bei der zweiten Frage zu finden waren, insbesondere bei der Kondition Scanpath; dies galt auch für die Gesamtsumme der

richtigen zu falschen Antworten pro Kondition. Auch hinsichtlich der Gesamtanzahl richtiger Antworten pro Kondition konnte kein signifikanter Unterschied gefunden werden. Somit konnte weder Forschungshypothese **H1(a)** noch **H1(b)** angenommen werden. Darüber hinaus konnte somit nicht bestätigt werden, dass das in Abschnitt 1 angesprochene Potenzial des verbesserten Verständnisses durch Assistenz beim Lesen in VR erfüllt wurde.

6.4 Gesamtlesedauer

Da der Scanpath aufgrund seiner explorativen Eigenschaft zu einer eingängigeren Auseinandersetzung mit dem eigenen Leseverhalten und Text animieren sollte, wurde bei dieser Kondition von einer längeren Gesamtlesedauer als ohne Leseassistent ausgegangen (**H2(a)**), wohingegen der Peripheral Blur aufgrund seiner unmittelbaren Unterstützung des visuellen Fokus eine kürzere Lesezeit erwarten ließ als ohne Leseassistent (**H2(b)**). So zeigte die Baseline-Kondition tatsächlich die größte Streuung nach oben mit einem Maximum von 495,3 Sekunden (≈ 8 Minuten), womit augenscheinlich das in Abschnitt 1 erwähnte Potenzial einer verkürzten Lesezeit – zumindest bei Nutzung des Peripheral Blur – erfüllt werden würde. Mit Einbeziehen der Ausreißer wurde in der Kondition Peripheral Blur am längsten gelesen mit ca. 11 Minuten Länge. Die Ausreißer bei Scanpath (zwischen ca. 6 bis 8 Minuten) ließen sich analog zu den Ausreißern derselben Kondition innerhalb der Anstrengungs-Subskala erklären (s. 6.2). Denn so unterbrach laut den Proband:innen das Anzeigen des Scanpath zusammen mit den abgeschnittenen Sätzen den Lesefluss, wodurch vermutlich öfter zurückgeblättert werden musste, um den Anfang des Satzes nachzulesen. Dies könnte somit die Lesezeit verlängert haben. Kürzeste Gesamtlesedauer lag innerhalb der Baseline-Kondition bei ca. 2 Minuten. Trotz dieser Auffälligkeiten konnten keine signifikanten Unterschiede zwischen den Konditionen bezüglich der Gesamtlesedauer in Sekunden bestätigt werden. Daher konnten auch die Nullhypothesen zu **H2(a)** und **H2(b)** nicht abgelehnt werden. Zusätzlich konnte damit das augenscheinlich erfüllte Potenzial einer verkürzten Lesezeit durch Assistenz beim Lesen in VR nicht bestätigt werden.

6.5 Lesewiederholungen

Die Annahmen zu der Anzahl an Lesewiederholungen einer Seite verhielten sich analog zu denen der Lesedauer. So wurde erwartet, dass der Scanpath aufgrund der Reflektion über das eigene Leseverhalten (in Form des eigenen Scanpath) – und einer daraus folgenden vertieften Auseinandersetzung mit dem Text – zu einem häufigeren Nachlesen einer Seite führte als ohne Leseassistent **H3(a)**. Der Peripheral Blur hingegen sollte wegen seiner unmittelbaren Assistenz durch Ausblenden störenden Textes in der Peripherie während des Lesens zu einer entsprechend niedrigeren Anzahl an Lesewiederholungen führen als ohne Leseassistent **H3(b)**. Die deskriptiven Daten gaben dies kaum wieder, da Peripheral Blur und die Baseline-Kondition ähnlich nach oben streuten, die Daten von Scanpath hingegen am wenigsten. Zusätzlich besaß Peripheral Blur einen Ausreißer mit 17 Lesewiederholungen. Dieser geht laut zugehöriger Person darauf zurück, dass der Überblick über den Textinhalt durch den Blurfilter eingeschränkt wurde. Auch hier konnten letztendlich keine signifikanten Unterschiede zwischen den Konditionen gefunden werden, weshalb **H3(a)** und **H3(b)** nicht angenommen werden können.

6.6 Korrelationen

Darüber hinaus wurden innerhalb der Konditionen mit Leseassistenten-Funktion Korrelationen berechnet.

6.6.1 Scanpath-Dauer

Die Scanpath-Zeit spiegelte wieder, wie lange die Proband:innen ihren Scanpath sahen, bevor sie die Seite erneut lasen oder auf die nächste Seite wechselten. Mit einer längeren Auseinandersetzung des Scanpath, der ihnen zeigte, wo und wie lange sie mit ihren

Augen verweilten, könnte die Aufmerksamkeit auf Regionen im Text gelenkt werden, die mit längeren Fixationszeiten und/oder Rücksprüngen (abgebildet als rötliche Kreise und Rückpfeile) von Verständnisproblem zeugen könnten. Eine erneute Betrachtung dieser hervorgehobenen Stellen könnte dazu führen, dass ihr Inhalt besser im Gedächtnis bleibt. Genauso könnten Regionen mit wenig grünen bis gar keine Kreise (denn unter 50 Millisekunden wurde eine Fixation nicht im Scanpath abgebildet) darauf hindeuten, dass eine Textstelle nur flüchtig mit den Augen überflogen und daher nicht bewusst aufgenommen wurde. Dies könnte die Leser:innen dazu animieren an dieser Stelle bewusster nachzulesen, was aufgrund der transparenten Fixationskreise auch während der Anzeige des Scanpath möglich war. Aufgrund dieser Implikationen wurde ein Zusammenhang zwischen der Betrachtungszeit des Scanpath und der Anzahl richtiger Antworten getestet. Eine signifikante Korrelation konnte jedoch nicht nachgewiesen werden.

6.6.2 Scanpath-Häufigkeit

Ebenso von Interesse wie die Scanpathzeit war die Häufigkeit, mit der Scanpaths angezeigt wurden, da eine wiederholte Betrachtung des eigenen Leseflusses zu einer Anpassung im Leseverhalten geführt haben kann, ebenso würde so der Textinhalt vermehrt wiederholt werden, wodurch er sich besser einprägen kann. So wurde auch auf eine Korrelation zwischen der Anzahl der Scanpaths und Anzahl richtiger Antworten getestet, welche ebenfalls nicht signifikant ausfiel.

6.6.3 Peripheral Blur: Zeileneinstellungen

Beim Peripheral Blur wurde auf einen Zusammenhang zwischen den Zeileneinstellungen und der Anzahl richtiger Antworten untersucht. Sollte ein solcher bestehen, waren zwei Richtungen möglich: So konnten weniger fokussierte Zeilen mehr ablenkenden Text innerhalb des visuellen Umfelds abdämpfen, was den Fokus auf die anvisierte Textstelle erhöhen und so zu einer höheren Anzahl richtiger Antworten führen könnte. Auf der anderen Seite hätte der Blur-Effekt an sich eine Ablenkungs- oder Störquelle für die Proband:innen beim Lesen darstellen können, weshalb mehr scharfgestellte Zeilen zu mehr richtigen Antworten führen würden. Die Korrelation war negativ, allerdings sehr schwach und zudem nicht signifikant.

6.6.4 Peripheral Blur: Anzahl Zeileneinstellungen

Auch der Zusammenhang zwischen der Anzahl an gemachten Einstellungen und der Anzahl richtiger Antworten wurde betrachtet. Denn so können viele Zeileneinstellungen eine Anpassung an verschieden große Sinnabschnitte implizieren, auf der anderen Seite aber auch ein Austesten der Grenzen der Funktionalität. Ebenso kann eine hohe Zahl an scharfgestellten Zeilen gewählt worden sein, da wie bereits erwähnt der Blur selbst beim Lesen stören könnte. Allerdings fiel die hierfür berechnete Korrelation nicht signifikant aus.

6.7 Qualitative Ergebnisse

Aus den qualitativen Ergebnissen geht hervor, dass der Peripheral Blur insgesamt die beliebteste und hilfreichste Leseassistenten darstellt, auch unter den Proband:innen mit wenig bis keiner VR-Erfahrung. Er half den Proband:innen, nicht in der Zeile zu verrutschen und den Fokus auf der aktuell gelesenen Textstelle zu behalten, während ablenkender Text in der Peripherie in seiner Salienz reduziert wurde. Gleichzeitig wurde seine Bedienung von den meisten Proband:innen als anstrengend und unintuitiv empfunden, da zum einen der mangelnde Trage-Komfort und zum anderen die Kopfbewegung zum Fokussieren des Textes den freien Lesefluss einschränkten. Gerade die Personen, die nicht in der Zeile verrutschen und sich schnell in Texten zurechtfinden, empfanden die Assistenzform als Hindernis beim Lesen. Die beliebteste Einstellung von drei

bis vier fokussierten Zeilen sorgte laut den Proband:innen für ein angenehmes Leseerlebnis, während laut manchen Proband:innen mehr Zeilen nötig waren, um eine unangenehme Kopfausrichtung in den obersten und untersten Zeilen zu kompensieren.

Der Scanpath wurde zwar überwiegend als interessant, aber weniger hilfreich wahrgenommen. Jedoch konnte er bei manchen Proband:innen laut Feedback zur erhofften Wirkung des reflektierten Leseverhaltens beitragen und wurde als eine Art aufmerksamkeitsbasierte Markierung von Interessensregionen verstanden. Mit der Hervorhebung von Fremdwörtern besteht somit auch ein Potenzial in der Anwendung auf fremdsprachige Texte. Auf der anderen Seite führte das Anzeigen des Scanpath zwischen den Seiten bei manchen Proband:innen für eine Unterbrechung im Lesefluss und beförderte zum Teil ein Gefühl von Leistungs- bzw. Zeitdruck, was am Ziel einer intrinsischen Anpassung im Leseverhalten vorbeiführt.

7 LIMITATIONEN UND FUTURE WORK

Die ursprüngliche Idee für eine virtuelle Leseassistentz enthielt zwar eine Anwendung in AR, allerdings wurde aufgrund von technischen Limitationen wie der Occlusion und Texterkennung auf die VR zurückgegriffen, um eine Augmentation eines Textes auf Papier oder Bildschirmen zu simulieren. Mit Verbesserungen hinsichtlich der Deckkraft von Hologrammen bzw. Overlays könnte die Anwendung im Rahmen weiterer Forschung in der ursprünglich beabsichtigten Form umgesetzt und getestet werden.

Darüber hinaus wurden von Proband:innen Hardware- und Softwarespezifische Einschränkungen bemängelt. Hinsichtlich der Leseassistentz-Funktion Scanpath sorgten manche Regressionspfeile für Verwirrung, da sich an einen entsprechenden Rücksprung der Augen nicht erinnert werden konnte. Sehr wahrscheinlich ist dieses Problem darauf zurückzuführen, dass beim Wechsel auf die nächste oder vorherige Seite die Augen auf den unteren Teil des Textpanels gerichtet waren und somit beim Sprung auf die nächste Seite eine kurze Weile auf den unteren Text der neuen Seite gerichtet waren, bevor sie zur ersten Zeile des Textes zurückkorrigierten. Diese Korrektur der Augen wurde dann bereits als Regressionspfad gewertet und die Parameter für einen Pfad aus Rückpfeilen gespeichert. Zur Lösung dieses Problems machte eine Person den Vorschlag, eine Art Erholphase einzuführen, sodass für eine kurze Dauer, in der die Augen an den Anfang des Textes der nächsten oder vorherigen Seite springen können, keine Parameter für einen neuen Scanpath aufgezeichnet werden. Auch ein Anzeigen des Scanpath nach einem beendeten Sinnabschnitt statt einer ganzen Seite wurde sich gewünscht. Ein Anzeigen des Scanpath zu einem beliebigen Zeitpunkt wäre umsetzbar, jedoch in einem Studiendesign schwer kontrollierbar, da wie bereits erwähnt nicht garantiert werden kann, dass die Proband:innen den Scanpath dann überhaupt nutzen würden. Allerdings könnte eine Kombination aus dem Anzeigen zu einem beliebigen Zeitpunkt und beim Seitenwechsel realisiert werden. Ebenfalls könnten absatzweise Collider eingebaut werden, die ein Eintreten und Verlassen des Blickes registrieren, wodurch nach dem Lesen eines Absatzes ein Scanpath mit den erhobenen Parametern dieser Textstelle gezeichnet wird. Kritisiert wurden zudem Schusterjungen, welche den Lesefluss zusammen mit dem Anzeigen des Scanpath unterbrechen. Somit könnten die Textparameter für eine weitere Studie so angepasst werden, dass der letzte Satz auf jeder Seite nicht abgeschnitten wird. Darüber hinaus fiel bei der Betrachtung der Werte der Proband:innen auf, dass sie häufig auf der vierten Textseite keinen Scanpath visualisierten, was vermutlich daran lag, dass ein Wechsel auf die nächste Seite am Ende des Textes unintuitiv und unnötig erscheint. Denn so konnte nur durch einen Wechsel zur nächsten Seite der eigene Scanpath eingesehen werden. Ebenfalls fiel nach Beenden der Studie ein Fehler bei der Berechnung des Vektors zur Winkelbildung für die Maßeinheit Winkel pro Sekunde auf, so wie in 3.2.2 beschrieben. Allerdings kann dieser Fehler nur minimal zu Verzerrungen im Scanpath geführt haben, da das Studiendesign und Wahl der

Textparameter kaum Kopfrotationen geschweige denn Translationen zuließen. Im Bezug auf den Peripheral Blur wurde vor allem die Bedienung als unangenehm und unnatürlich empfunden, was vermutlich zum Großteil auf die schlechte Polsterung und Mangel an Belüftung der Pico zurückzuführen ist. Aufgrund des Eyetracking durfte das Headset während des gesamten VR-Teils nicht abgesetzt werden und drückte somit lange Zeit auf Stirn und Nase, was neben Druckstellen auch für viel Schwitzen sorgte. Zusammen mit der Auf- und Abbewegung entstand somit physischer Stress, was an der beabsichtigten ergonomischen Erfahrung vorbeiging. Für zukünftige Studien könnte daher auf eine VR- bzw. AR-Brille zurückgegriffen werden, die den Nutzer:innen einen höheren Komfort bietet, um das volle Potenzial dieser Leseassistentz ausschöpfen zu können. Ebenfalls wurde die Kopfbewegung als Eingabemodalität als unintuitiv wahrgenommen, da der Blick außerhalb des Blur-Filters nicht frei schweifen konnte, weshalb sich entsprechend der Blick als steuerndes Element gewünscht wurde. Es müsste somit zwischen dem Problem des Midas Touch und der Intuitivität abgewogen werden. Allerdings hätte auch wieder das unergonomische Headset zu dieser Wahrnehmung beitragen können. In weiterführenden Arbeiten könnten beide Methoden gegeneinander getestet werden. Zusätzlich führte die Speicherung der Zeileneinstellung zwischen Test- und eigentlicher Runde zu Komplikationen bei der Auswertung. Denn so waren Proband:innen mit der letzten Einstellung aus der Testrunde anscheinend bereits zufrieden, sodass sie in der eigentlichen Runde keine weiteren Einstellungen vornahmen. Aus diesem Grund musste nach den ersten zwei Proband:innen der Code so angepasst werden, dass auch schon während der Testrunde Zeileneinstellungen in die Historie geloggt wurden. Bei der weiterführenden Auswertung musste dann zwischen Test und eigentlicher Runde unterschieden werden. Von daher würde es sich für eine weitere Studie anbieten, die Einstellung vor der eigentlichen Runde wieder auf den default einer fokussierten Zeile zurückzusetzen.

8 ABSCHLUSS UND FAZIT

Das Lesen wissenschaftlicher Texte wird häufig mit viel Anstrengung und einem hohen Zeitaufwand verbunden. Dabei wandert der Blick unablässig über schwierige Textstellen und bleibt dort hängen, Satzteile entgleiten aufgrund von abnehmender Aufmerksamkeit oder werden unachtsam überflogen. Ebenso kann allein die Textmenge irritieren und die Augen ablenken oder es wird in der Zeile verrutscht. Aus der Motivation heraus eine Hilfestellung während und nach dem Lesen zu bieten, wurde eine Anwendung entwickelt, die beide Assistenzformen enthält: Scanpath und Peripheral Blur. Erstere gibt nach dem Lesen in der Form eines angepassten Scanpath Einblick in den eigenen Lesefluss und soll zu einer Auseinandersetzung mit dem eigenen Leseverhalten und evtl. zu einer Änderung führen, die ein bewussteres Lesen enthält. Die zweite Leseassistentz-Funktion greift während des Lesens auf einen DR-Filter angelehnt an das visuelle System zurück, welcher eine frei anpassbare Anzahl an Zeilen scharfstellt, während der umliegende Text weichgezeichnet wird, was eine Reduktion in seiner Salienz und seinem Ablenkungspotenzial beabsichtigt. Aufgrund der technischen Beschränkungen der AR hinsichtlich der Occlusion wurde die Anwendung in VR simuliert. Beide Funktionen wurden innerhalb einer Lese-Studie in stationärer Seated VR mit Within-Subjects-Design und drei Konditionen (2 Leseassistentzen plus Kontrolle) hinsichtlich inhaltlichen Abrufs, Leseparameter sowie funktionsinterner Merkmale evaluiert. Es wurde eine quantitative Analyse auf den Daten aus der Benutzerstudie durchgeführt. Weder im Bezug auf die Leseparameter noch auf den inhaltlichen Abruf der Textinhalte konnten signifikante Unterschiede zwischen den Konditionen nachgewiesen werden. Auch die Korrelationen innerhalb der Konditionen deuteten auf keinen signifikanten Zusammenhang zwischen den funktionspezifischen Merkmalen und Anzahl richtiger Antworten hin. Schließlich wurde eine qualitative Inhaltsanalyse auf den

freien Angaben der Proband:innen durchgeführt. Die Proband:innen empfanden die Assistenz mit Peripheral Blur am hilfreichsten, aber aufgrund ihrer Bedienung auch am anstrengendsten, während sie den Scanpath für interessant, aber weniger hilfreich hielten. Abschließend lässt sich sagen, dass keine der Leseassistenten-Funktionen zu einer Verbesserung in Leseverhalten oder -Effizienz im Vergleich zum Lesen in VR ohne Assistenz beitrugen, ebensowenig konnte diesbezüglich eine Verschlechterung nachgewiesen werden. Dennoch erlaubte die qualitative Inhaltsanalyse zahlreiche Einblicke in das Leseverhalten der Proband:innen und ihre Wahrnehmung der Leseassistenten-Funktionen. So ließen sich in den Beschreibungen der Proband:innen teilweise Parallelen zwischen den Leseassistenten-Funktionen und den angegebenen Lesestrategien erkennen; sogar die beabsichtigte Wirkung der Assistenzfunktionen spiegelte sich in zahlreichen Antworten wieder. Die zwei neuartigen Leseassistenten-Funktionen könnten Grundlage zukünftiger Arbeiten im Bereich der Assistive Technologies innerhalb der Lesedomäne bilden, wo sie mithilfe der vorgebrachten Verbesserungsvorschläge zu einer AR-Anwendung weiterentwickelt werden und bezüglich ihres ganzen Potenzials untersucht werden können.

LITERATUR

- [1] I. B. Adhanom, N. N. Griffin, P. MacNeilage, and E. Folmer. The effect of a foveated field-of-view restrictor on vr sickness. In *2020 IEEE conference on virtual reality and 3D user interfaces (VR)*, pp. 645–652. IEEE, 2020.
- [2] K. Azuma and H. Koike. A study on gaze guidance using artificial color shifts. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces*, pp. 1–5, 2018.
- [3] S. Benedetto, A. Carbone, M. Pedrotti, K. Le Fevre, L. A. Y. Bey, and T. Baccino. Rapid serial visual presentation in reading: The case of spritz. *Computers in Human Behavior*, 45:352–358, 2015.
- [4] P. Bimberg, T. Weissker, and A. Kulik. On the usage of the simulator sickness questionnaire for virtual reality research. In *2020 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW)*, pp. 464–467. IEEE, 2020.
- [5] T. Blascheck, K. Kurzhals, M. Raschke, M. Burch, D. Weiskopf, and T. Ertl. State-of-the-art of visualization for eye tracking data. In *Eurovis (stars)*, p. 29, 2014.
- [6] J. V. Bradley. Complete counterbalancing of immediate sequential effects in a latin square design. *Journal of the American Statistical Association*, 53(282):525–528, 1958.
- [7] M. P. Bucci, N. Nassibi, C.-L. Gerard, E. Bui-Quoc, and M. Seassau. Immaturity of the oculomotor saccade and vergence interaction in dyslexic children: evidence from a reading and visual search study. *PloS one*, 7(3):e33458, 2012.
- [8] Y. F. Cheng, H. Yin, Y. Yan, J. Gugenheimer, and D. Lindlbauer. Towards understanding diminished reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pp. 1–16, 2022.
- [9] A. D. Clarke, A. Mahon, A. Irvine, and A. R. Hunt. People are unable to recognize or report on their own eye movements. *Quarterly Journal of Experimental Psychology*, 70(11):2251–2270, 2017.
- [10] A. Dillon. Reading from paper versus screens: A critical review of the empirical literature. *Ergonomics*, 35(10):1297–1326, 1992.
- [11] T. Dingler, K. Kunze, and B. Outram. Vr reading uis: Assessing text parameters for reading in vr. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–6, 2018.
- [12] T. Dingler, A. S. Shirazi, K. Kunze, and A. Schmidt. Assessment of stimuli for supporting speed reading on electronic devices. In *Proceedings of the 6th Augmented Human International Conference*, pp. 117–124, 2015.
- [13] S. Eraslan, Y. Yesilada, and S. Harper. Eye tracking scanpath analysis on web pages: how many users? In *Proceedings of the ninth biennial ACM symposium on eye tracking research & applications*, pp. 103–110, 2016.
- [14] S. Eraslan, Y. Yesilada, and S. Harper. Scanpath trend analysis on web pages: Clustering eye tracking scanpaths. *ACM Transactions on the Web (TWEB)*, 10(4):1–35, 2016.
- [15] S. Eraslan, Y. Yesilada, V. Yaneva, and S. Harper. Eye-tracking scanpath trend analysis for autism detection. *ACM SIGACCESS Accessibility and Computing*, (128):1–8, 2020.
- [16] A. Erickson, K. Kim, G. Bruder, and G. F. Welch. Effects of dark mode graphics on visual acuity and fatigue with virtual reality head-mounted displays. In *2020 IEEE Conference on virtual reality and 3D user interfaces (VR)*, pp. 434–442. IEEE, 2020.
- [17] R. Fleisch. A new readability yardstick. *Journal of applied psychology*, 32(3):221, 1948.
- [18] J. Gabel, M. Ludwig, and F. Steinicke. Immersive reading: Comparison of performance and user experience for reading long texts in virtual reality. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–8, 2023.
- [19] J. H. Goldberg and J. I. Helfman. Visual scanpath representation. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, pp. 203–210, 2010.
- [20] W. Guo and S. Cheng. An approach to reading assistance with eye tracking data and text features. In *Adjunct of the 2019 International Conference on Multimodal Interaction*, pp. 1–7, 2019.
- [21] T. Gupta, L. Aflatoony, and L. Leonard. Augmentally: A reading assistant application for children with dyslexia. In *Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 1–3, 2021.
- [22] S. G. Hart. Nasa-task load index (nasa-tlx); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, vol. 50, pp. 904–908. Sage publications Sage CA: Los Angeles, CA, 2006.
- [23] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, vol. 52, pp. 139–183. Elsevier, 1988.
- [24] H. Hata, H. Koike, and Y. Sato. Visual guidance with unnoticed blur effect. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, pp. 28–35, 2016.
- [25] R. Hinojosa. Sizing UIs in VR with Distance-Independent Millimeters. Website. <https://www.ryanhinojosa.com/2018/01/08/device-independent/>, 2018. Letzter Zugriff am 14.08.2024.
- [26] Z. Hu, A. Bulling, S. Li, and G. Wang. Fixationnet: Forecasting eye fixations in task-oriented virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, 27(5):2681–2690, 2021.
- [27] Z. Hu, S. Li, C. Zhang, K. Yi, G. Wang, and D. Manocha. Dgaze: Cnn-based gaze prediction in dynamic scenes. *IEEE transactions on visualization and computer graphics*, 26(5):1902–1911, 2020.
- [28] J. Huang, M. Kinader, M. J. Dunn, W. Jarosz, X.-D. Yang, and E. A. Cooper. An augmented reality sign-reading assistant for users with reduced vision. *PloS one*, 14(1):e0210630, 2019.
- [29] L. A. Huisinga. *Augmented reality reading support in higher education: Exploring effects on perceived motivation and confidence in comprehension for struggling readers in higher education*. PhD thesis, Iowa State University, 2017.
- [30] A. Hyrskykari. *Eyes in attentive interfaces: Experiences from creating iDict, a gaze-aware reading aid*. Tampere University Press, 2006.
- [31] R. J. Jacob. Eye tracking in advanced interface design. *Virtual environments and advanced interface design*, 258(288):2, 1995.
- [32] L. Kasa. Augmented reality as an educational support tool for dyslexic students in higher education: Exploring potentials of augmented reality to support reading processes, 2022.
- [33] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, 3(3):203–220, 1993.
- [34] T. Kojić, D. Ali, R. Greinacher, S. Möller, and J.-N. Voigt-Antons. User experience of reading in virtual reality—finding values for text distance, size and contrast. In *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*, pp. 1–6. IEEE, 2020.
- [35] T. Kübler, W. Fuhl, R. Rosenberg, W. Rosenstiel, and E. Kasneci. Novel methods for analysis and visualization of saccade trajectories. In *Computer Vision—ECCV 2016 Workshops: Amsterdam, The Netherlands, October 8–10 and 15–16, 2016, Proceedings, Part I 14*, pp. 783–797. Springer, 2016.
- [36] K. P. Kumar and V. Reddi. Significance of spearman’s rank correlation

- coefficient. *International Journal For Multidisciplinary Research*, 5(4), 2023.
- [37] A.-F. Lai, C.-H. Chen, and G.-Y. Lee. An augmented reality-based learning approach to enhancing students' science reading performances from the perspective of the cognitive load theory. *British Journal of Educational Technology*, 50(1):232–247, 2019.
- [38] C. Lankford. Gazetracker: software designed to facilitate eye movement analysis. In *Proceedings of the 2000 symposium on Eye tracking research & applications*, pp. 51–55, 2000.
- [39] X. Li, Y. Han, Y. Wu, K. Yue, and Y. Liu. User study of an ar reading aid system to promote deep reading. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 611–612. IEEE, 2021.
- [40] A. Mangen, B. R. Walgermo, and K. Brønnick. Reading linear texts on paper versus computer screen: Effects on reading comprehension. *International journal of educational research*, 58:61–68, 2013.
- [41] S. Mann. Mediated reality. *Linux Journal*, 1999(59es):5–es, 1999.
- [42] S. Miksch and H. Hauser. Semantic depth of field. In *Proc. IEEE Symp. Information Visualization.*, pp. 97–104. Citeseer, 2001.
- [43] A. Olsen. The tobii i-vt fixation filter. *Tobii Technology*, 21:4–19, 2012.
- [44] I. Ouali, M. B. Halima, and W. Ali. Augmented reality for scene text recognition, visualization and reading to assist visually impaired people. *Procedia Computer Science*, 207:158–167, 2022.
- [45] pagerangers. Der Flesch Reading Ease – Was er ist und wie man ihn berechnet. Website. <https://pagerangers.com/glossar/flesch-reading-ease/>, 2020. Letzter Zugriff am 13.08.2024.
- [46] A. Pettersen et al. Effective use of augmented reality to support reading for people with dyslexia. Master's thesis, University of Agder, 2024.
- [47] P.-L. P. Rau, J. Zheng, and Z. Guo. Immersive reading in virtual and augmented reality environment. *Information and Learning Sciences*, 122(7/8):464–479, 2021.
- [48] P.-L. P. Rau, J. Zheng, Z. Guo, and J. Li. Speed reading on virtual reality and augmented reality. *Computers & Education*, 125:240–245, 2018.
- [49] K. Rayner. Eye movements in reading and information processing: 20 years of research. *Psychological bulletin*, 124(3):372, 1998.
- [50] E. D. Reichle, K. Rayner, and A. Pollatsek. The ez reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and brain sciences*, 26(4):445–476, 2003.
- [51] K. Rush. *Augmented reality (AR): a school library app to engage high school reluctant readers to read for pleasure*. Northern Illinois University, 2017.
- [52] D. D. Salvucci and J. H. Goldberg. Identifying fixations and saccades in eye-tracking protocols. In *Proceedings of the 2000 symposium on Eye tracking research & applications*, pp. 71–78, 2000.
- [53] P. Schöll. Flesch-Index berechnen. Website. <https://www.fleschindex.de/>, 2024. Letzter Zugriff am 14.08.2024.
- [54] J. L. Sibert, M. Gokturk, and R. A. Lavine. The reading assistant: eye gaze triggered auditory prompting for reading remediation. In *Proceedings of the 13th annual ACM symposium on User interface software and technology*, pp. 101–107, 2000.
- [55] G. D. Spache. Is this a breakthrough in reading? *The Reading Teacher*, 15(4):258–263, 1962.
- [56] Spritz. Spritz. Website. <https://spritz.com/>, 2019. Letzter Zugriff am 23.08.2024.
- [57] L. Stearns, L. Findlater, and J. E. Froehlich. Design of an augmented reality magnification aid for low vision users. In *Proceedings of the 20th international ACM SIGACCESS conference on computers and accessibility*, pp. 28–39, 2018.
- [58] E. Thaqi, M. O. Mantawy, and E. Kasneci. Sara: Smart ai reading assistant for reading comprehension. In *Proceedings of the 2024 Symposium on Eye Tracking Research and Applications*, pp. 1–3, 2024.
- [59] C. Tosto, T. Hasegawa, E. Mangina, A. Chifari, R. Treacy, G. Merlo, and G. Chiazese. Exploring the effect of an augmented reality literacy programme for reading and spelling difficulties for children diagnosed with adhd. *Virtual Reality*, 25:879–894, 2021.
- [60] A. Valentino. Isolierte Sprachen: Die Geheimnisse der Überlebenden. Website. <https://www.spektrum.de/news/isolierte-sprachen-die-geheimnisse-der-ueberlebenden/2184438>, 2023. Letzter Zugriff am 14.08.2024.
- [61] M. van Wermeskerken, D. Litchfield, and T. van Gog. What am i looking at? interpreting dynamic and static gaze displays. *Cognitive science*, 42(1):220–252, 2018.
- [62] M. Wedel and R. Pieters. A review of eye-tracking research in marketing. *Review of marketing research*, pp. 123–147, 2017.
- [63] R. M. Yilmaz and H. Bursaliis. Effect of augmented reality applications on secondary school students' reading comprehension and learning permanency. 2019.
- [64] R. Zeelenberg and D. Pecher. A method for simultaneously counterbalancing condition order and assignment of stimulus materials to conditions. *Behavior research methods*, 47:127–133, 2015.

A ANHANG

A.1 Veranschaulichende Grafiken für Abläufe bzw. Aufbau der Anwendung

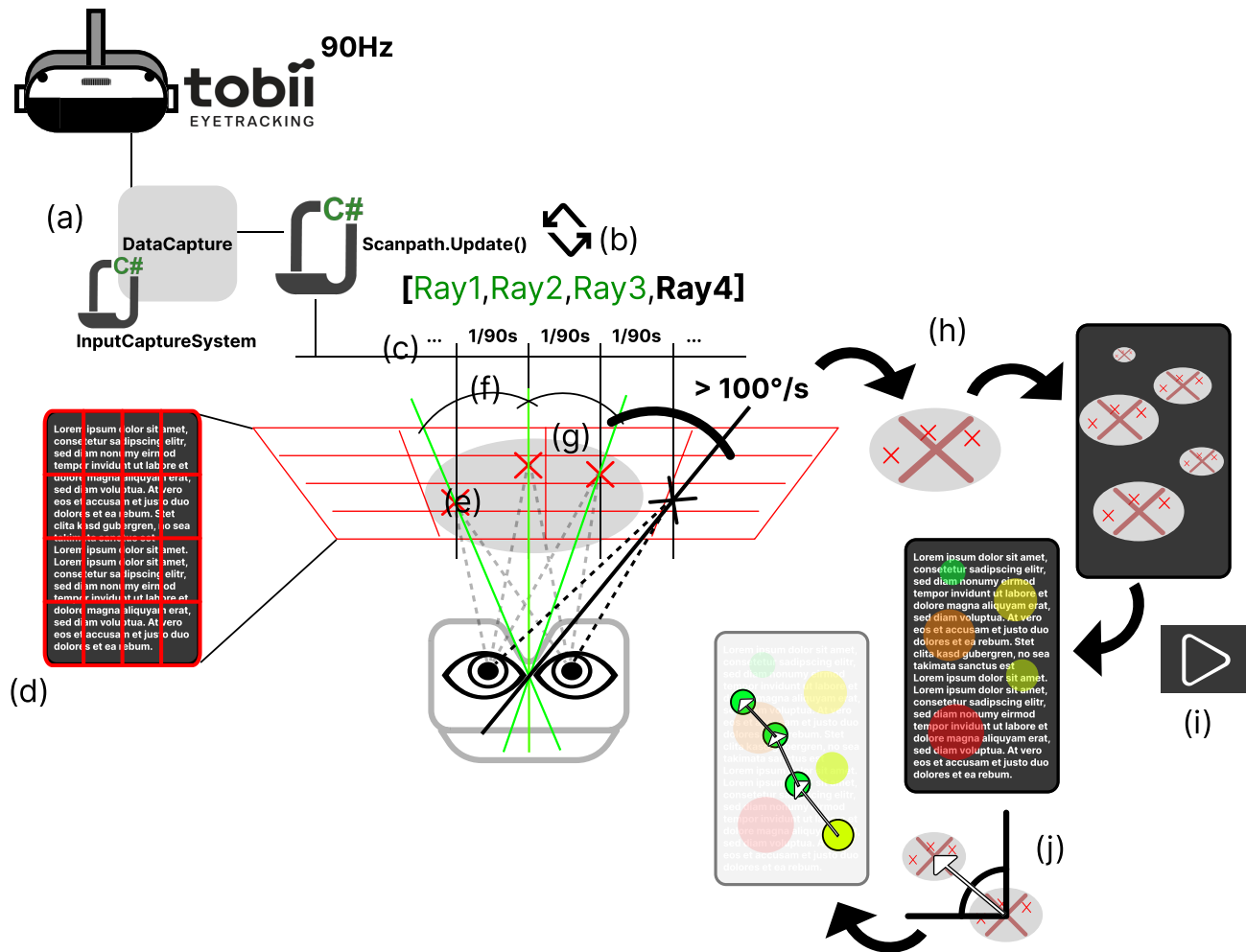


Abbildung 14: Leseassistenten-Funktion Scanpath und ihre Komponenten. (a): Das DataCapture-Package mit der modifizierten Schnittstelle InputCaptureSystem. (b): Pro Update()-Abruf wird über eine Liste aus Gazeays iteriert. (c) Taktung der Update()-Funktion nach der Abtastrate von 90Hz. (d) Textplane mit Größe und z-Höhe des Texthintergrunds. (e) Raycast: Schnittpunkt aus Kollision des Gazeray mit der Textplane. (f) Winkel zwischen Gazerays (multipliziert mit 90 für °/s). (g) Schnittpunkte (rot) werden Fixationscluster (graue Ellipse) hinzugefügt. (h) Sakkaden-Kriterium erfüllt (schwarzer Schnittpunkt): Fixationscluster wird zu einer Position gemittelt (großes „X“), Fixationsdauer wird berechnet. (i) Visualisierung durch Auswählen des Buttons zum Wechsel auf nächste Seite. (j) Visualisierung der Regressionen.

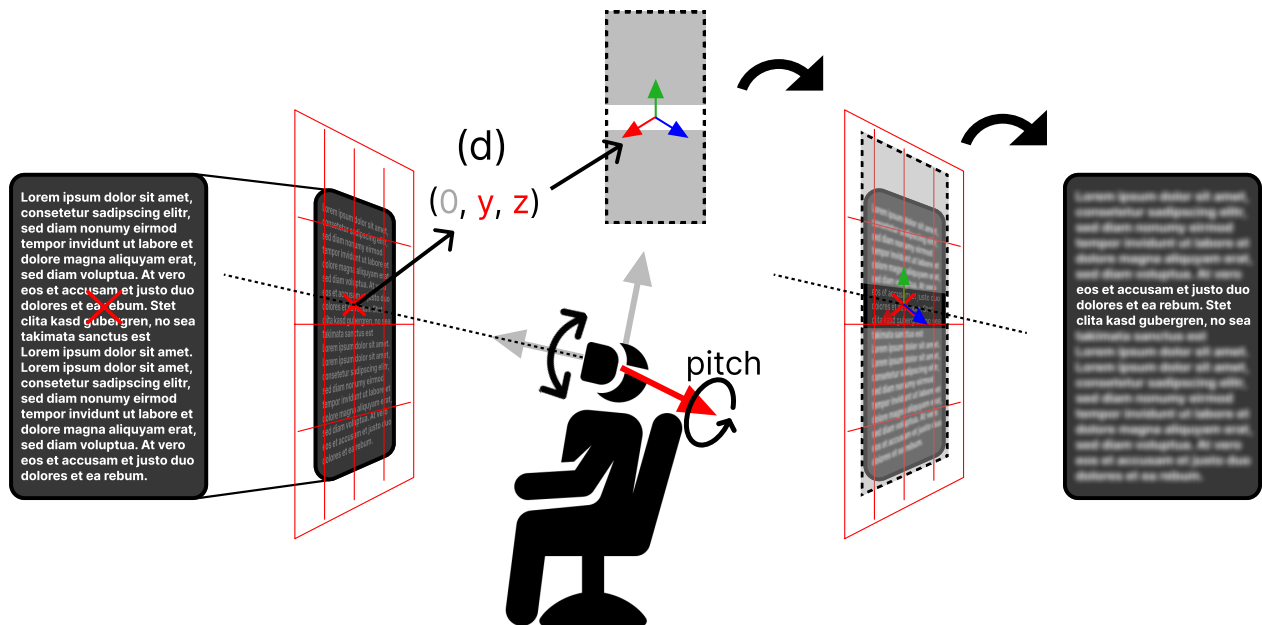
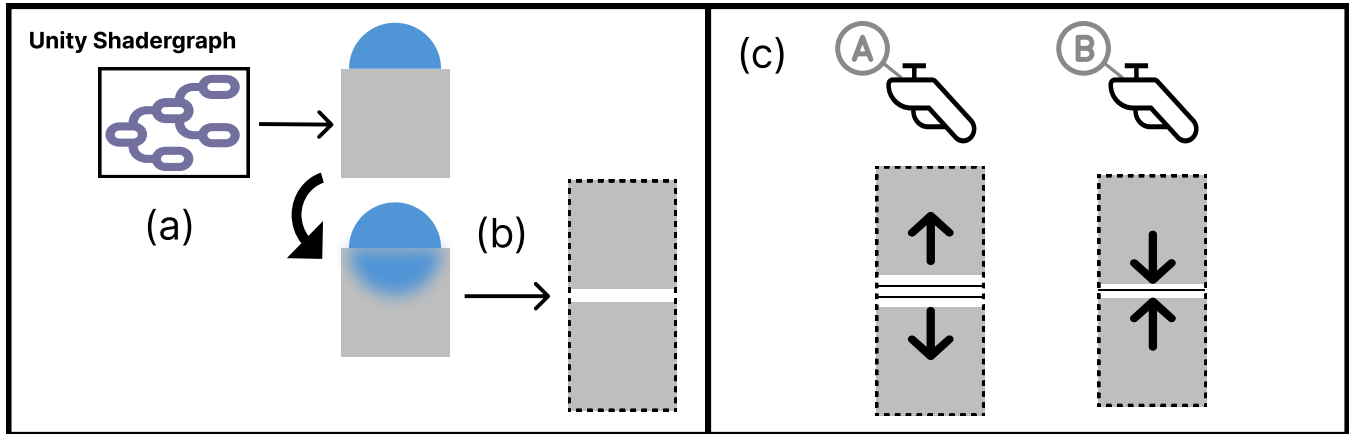


Abbildung 15: Leseassistenz-Funktion Peripheral Blur und ihre Komponenten. (a): Unity Shadergraph wird für den Blur-Effekt auf Material von zwei Shapes angewandt. Dieser macht die Shapes komplett transparent (grau für Illustration) und wirkt wie ein Weichzeichner auf dahinterliegende Objekte. (b): Beide Shapes werden übereinander angeordnet und einem gemeinsamen Parent Object zugeordnet. Dies ergibt den Blur-Filter. (c): Der Abstand zwischen den Shapes (fokussierter Bereich) ist über den Controller verstellbar. A-Button: Fügt eine Zeilenhöhe hinzu, B-Button: Nimmt eine Zeilenhöhe weg. (d): Der nose-direction-Ray schneidet abhängig von der Kopffrotation (*pitch*) eine unendlich große Plane auf gleicher z-Höhe des Textpanelss (rotes Gitter). Der resultierende Schnittpunkt wird in seiner y- und z-Komponente auf den Transform des Blurfilters angewandt.

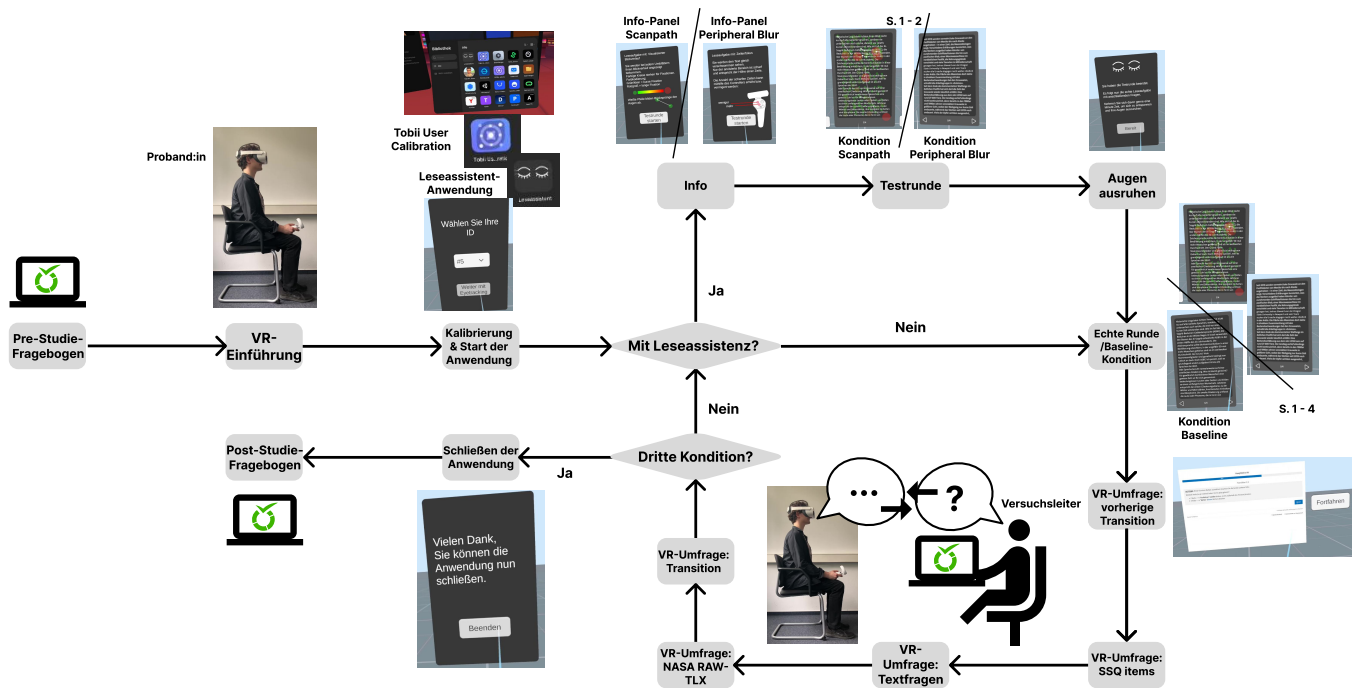


Abbildung 16: Ablauf der Benutzerstudie als Flussdiagramm mit nachgestelltem VR-Setup und Screenshots aus der Anwendung.



Abbildung 17: Proband in Seated VR (nachgestellt).

Boxplots: SSQ

Abbildung 18: SSQ: Übelkeits-Score (N)

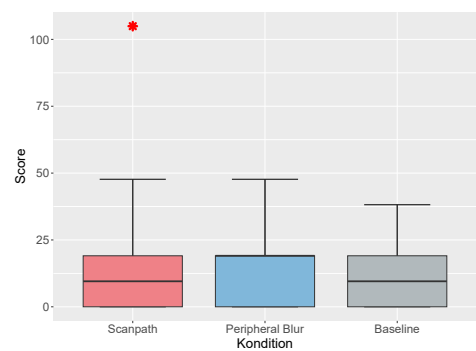


Abbildung 19: SSQ: Desorientierungs-Score (D)

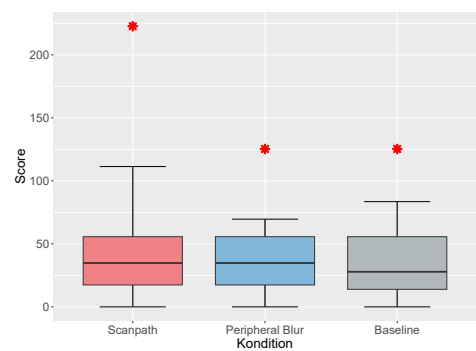
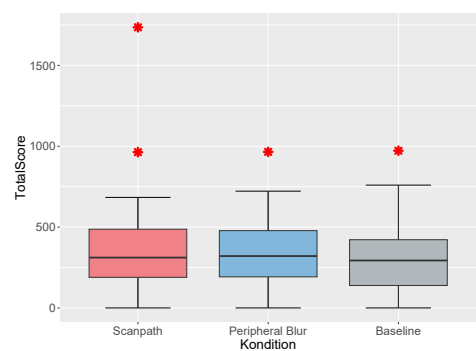


Abbildung 20: SSQ: Gesamt-Score (TS)



A.2 Boxplots: NASA RAW-TLX

Abbildung 21: NASA RAW-TLX: Subskala Geistige Anforderung

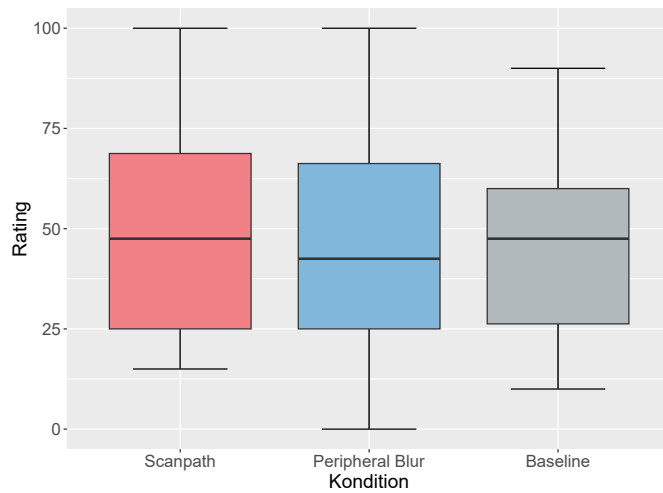


Abbildung 22: NASA RAW-TLX: Subskala Körperliche Anforderung

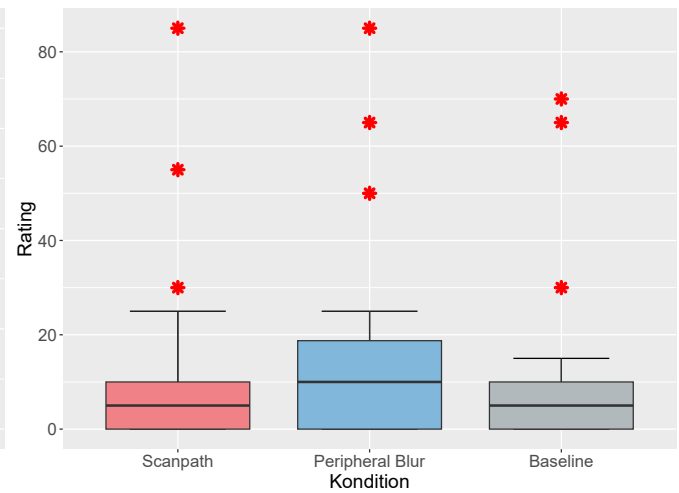


Abbildung 23: NASA RAW-TLX: Subskala Zeitliche Anforderung

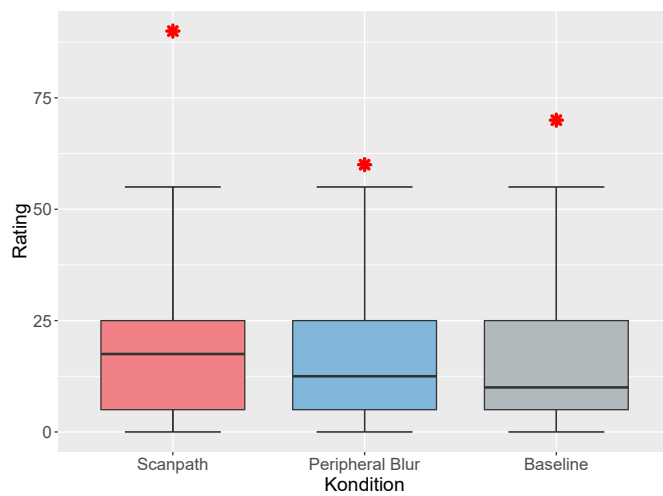
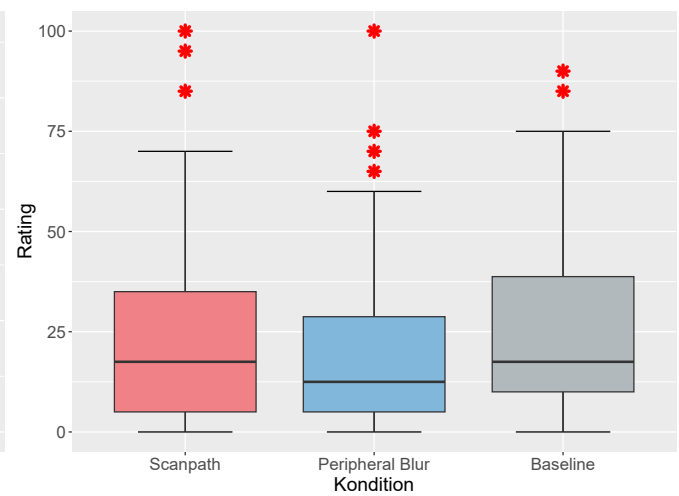


Abbildung 24: NASA RAW-TLX: Subskala Frustration



A.3 Boxplots: Kondition Scanpath

Abbildung 25: Kondition Scanpath: Scanpath-Dauer

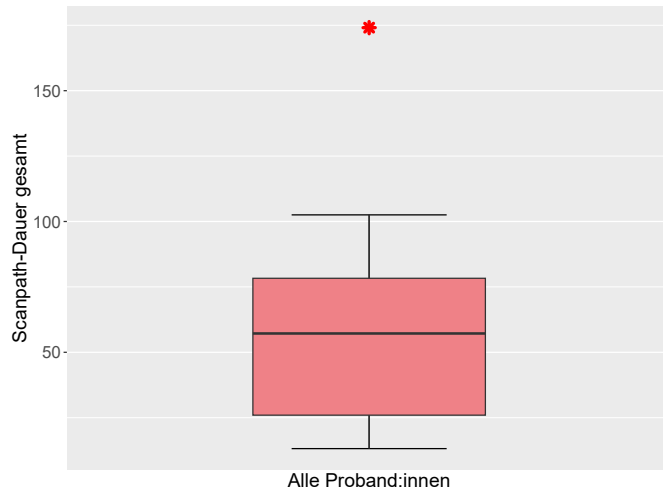
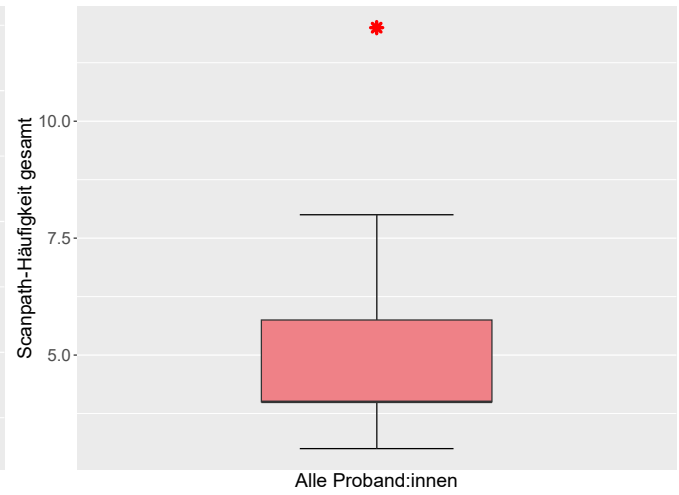


Abbildung 26: Kondition Scanpath: Scanpath-Häufigkeit



A.4 Boxplots: Kondition Peripheral Blur

Abbildung 27: Kondition Peripheral Blur: Zeileneinstellung

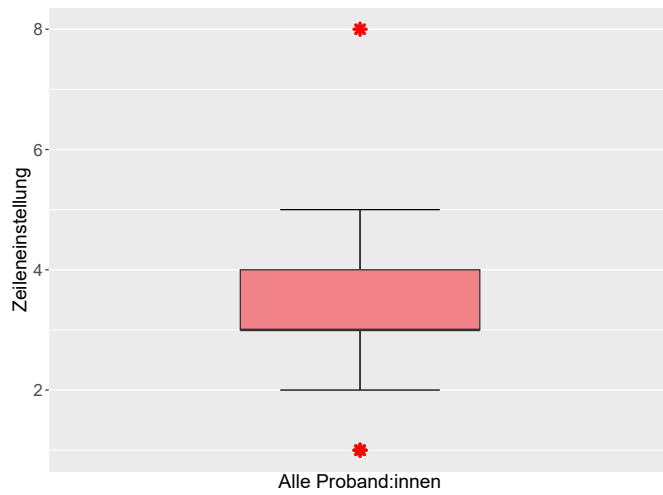
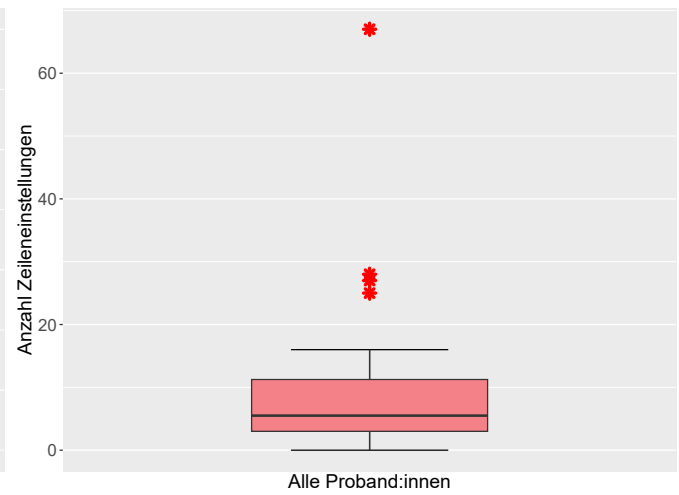


Abbildung 28: Kondition Peripheral Blur: Anzahl Zeileneinstellungen



A.5 Residuen: SSQ: Übelkeits-Score (N)

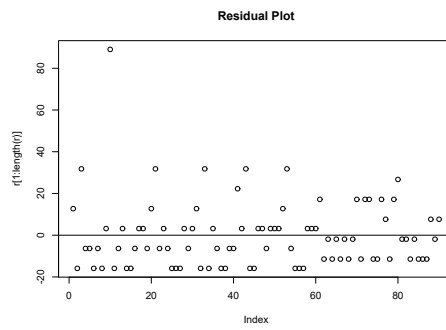


Abbildung 29: Residuen-Plot: SSQ: Übelkeits-Score (N)

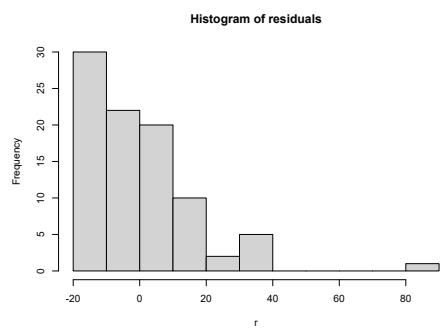


Abbildung 30: Residuen-Histogramm: SSQ: Übelkeits-Score (N)

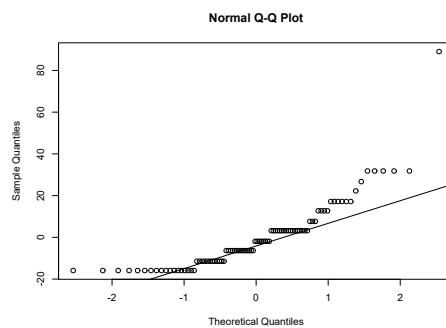


Abbildung 31: Normal-Q-Q-Plot: SSQ: Übelkeits-Score (N)

A.6 Residuen: SSQ: Übelkeits-Score (O)

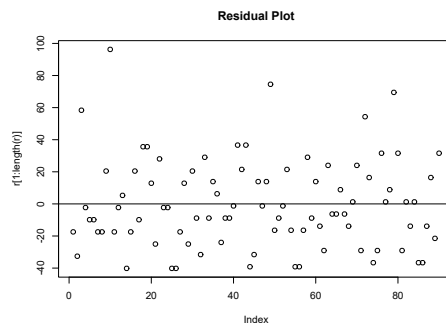


Abbildung 32: Residuen-Plot: SSQ: Okulomotorik-Score (O)

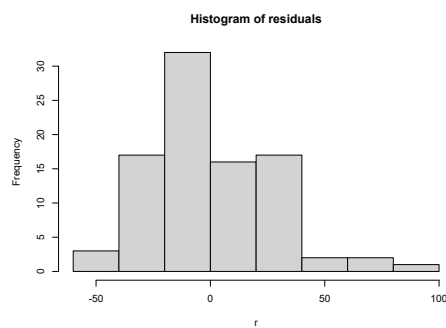


Abbildung 33: Residuen-Histogramm: SSQ: Okulomotorik-Score (O)

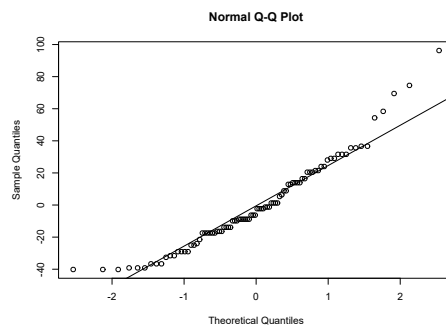


Abbildung 34: Normal-Q-Q-Plot: SSQ: Okulomotorik-Score (O)

A.7 Residuen: SSQ: Übelkeits-Score (D)

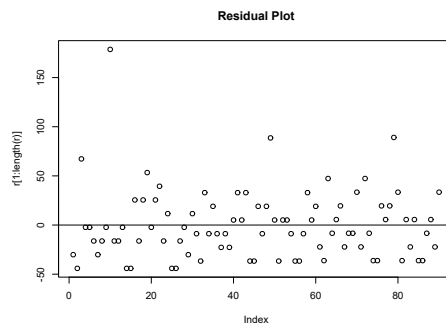


Abbildung 35: Residuen-Plot: SSQ: Desorientierungs-Score (D)

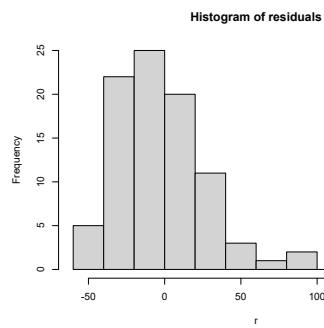


Abbildung 36: Residuen-Histogramm: SSQ: Desorientierungs-Score (D)

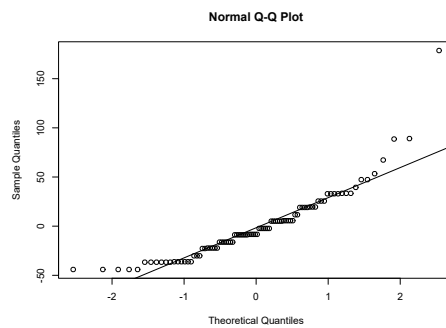


Abbildung 37: Normal-Q-Q-Plot: SSQ: Desorientierungs-Score (D)

A.8 Residuen: SSQ: Übelkeits-Score (TS)

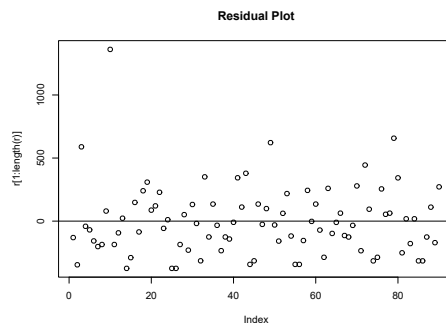


Abbildung 38: Residuen-Plot: SSQ: Gesamt-Score (TS)

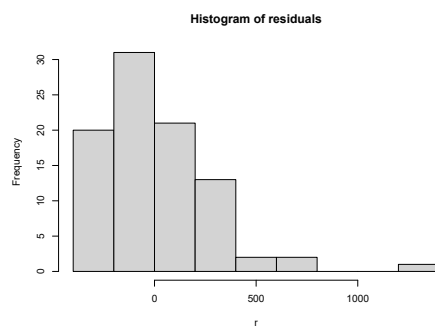


Abbildung 39: Residuen-Histogramm: SSQ: Gesamt-Score (TS)

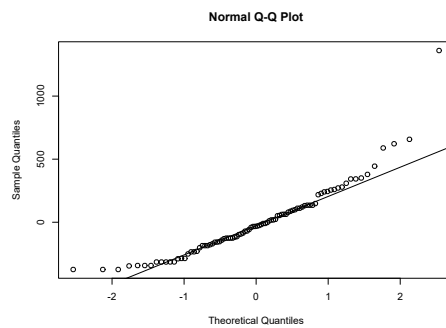


Abbildung 40: Normal-Q-Q-Plot: SSQ: Gesamt-Score (TS)

A.9 Residuen: Lesewiederholungen

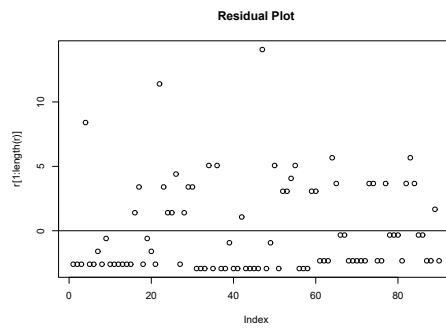


Abbildung 41: Residuen-Plot: Lesewiederholungen

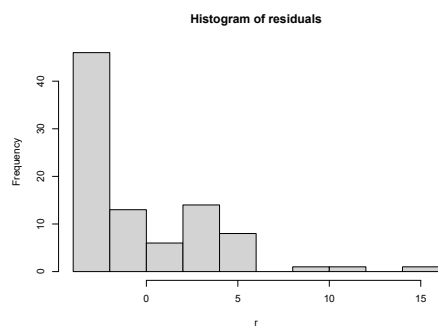


Abbildung 42: Residuen-Histogramm: Lesewiederholungen

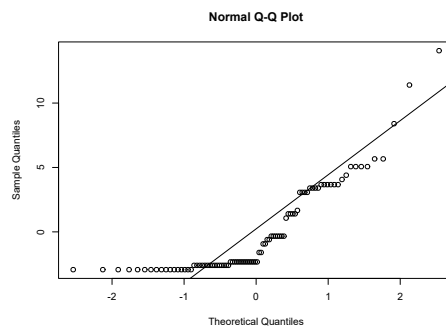


Abbildung 43: Normal-Q-Q-Plot: Lesewiederholungen

A.10 Residuen: Lesedauern

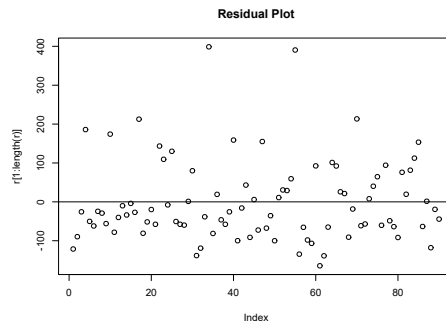


Abbildung 44: Residuen-Plot: Lesedauern

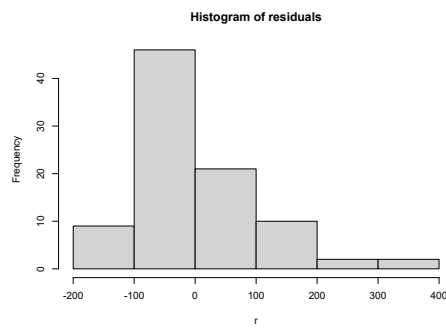


Abbildung 45: Residuen-Histogramm: Lesedauern

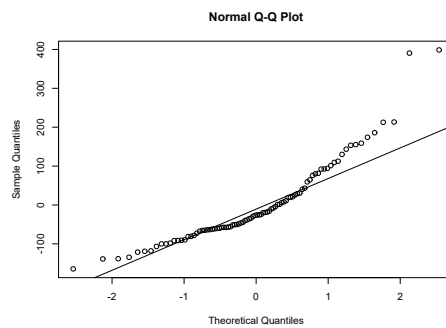


Abbildung 46: Normal-Q-Q-Plot: Lesedauern

A.11 Residuen: Richtige Antworten gesamt

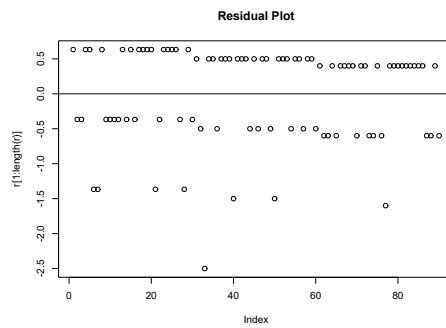


Abbildung 47: Residuen-Plot: Richtige Antworten gesamt

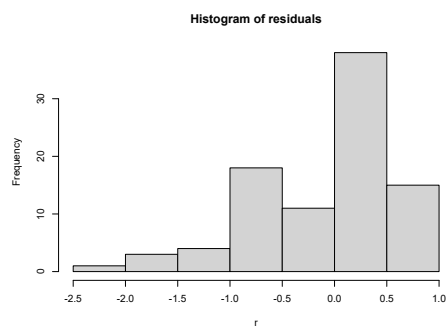


Abbildung 48: Residuen-Histogramm: Richtige Antworten gesamt

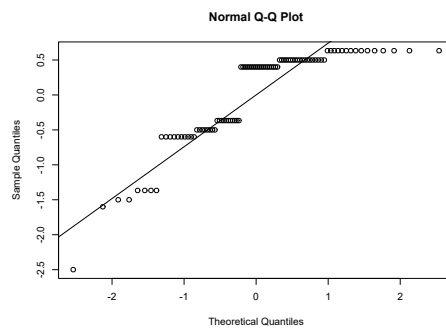


Abbildung 49: Normal-Q-Q-Plot: Richtige Antworten gesamt

A.12 Residuen: NASA RAW-TLX: Subskala Geistige Anforderung

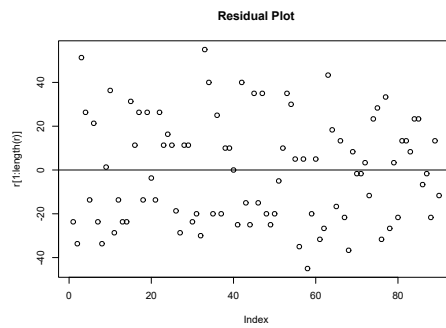


Abbildung 50: Residuen-Plot: NASA RAW-TLX: Subskala Geistige Anforderung

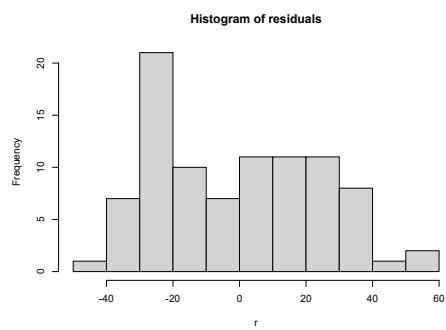


Abbildung 51: Residuen-Histogramm: NASA RAW-TLX: Subskala Geistige Anforderung

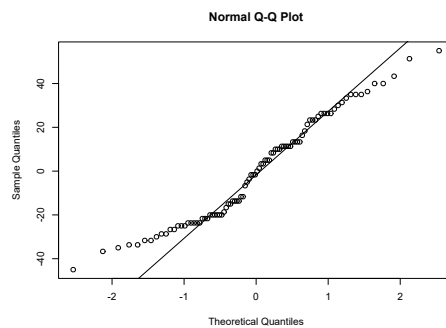


Abbildung 52: Normal-Q-Q-Plot: NASA RAW-TLX: Subskala Geistige Anforderung

A.13 Residuen: NASA RAW-TLX: Subskala Körperliche Anforderung

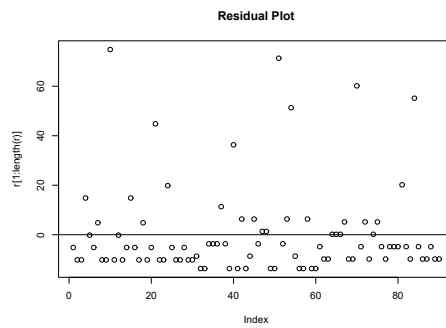


Abbildung 53: Residuen-Plot: NASA RAW-TLX: Subskala Körperliche Anforderung

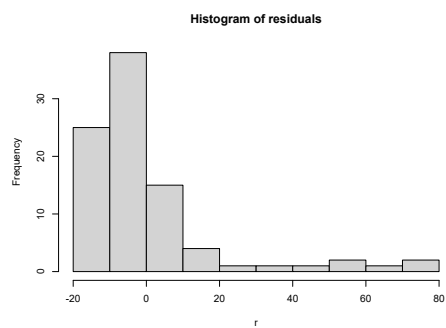


Abbildung 54: Residuen-Histogramm: NASA RAW-TLX: Subskala Körperliche Anforderung

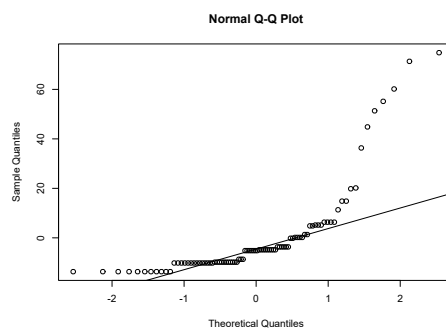


Abbildung 55: Normal-Q-Q-Plot: NASA RAW-TLX: Subskala Körperliche Anforderung

A.14 Residuen: NASA RAW-TLX: Subskala Zeitliche Anforderung

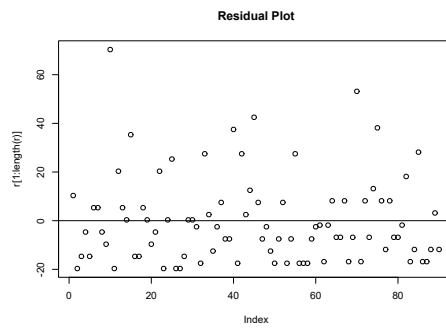


Abbildung 56: Residuen-Plot: NASA RAW-TLX: Subskala Zeitliche Anforderung

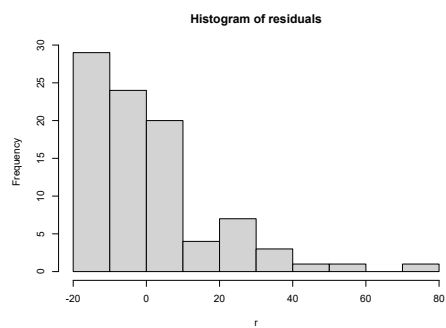


Abbildung 57: Residuen-Histogramm: NASA RAW-TLX: Subskala Zeitliche Anforderung

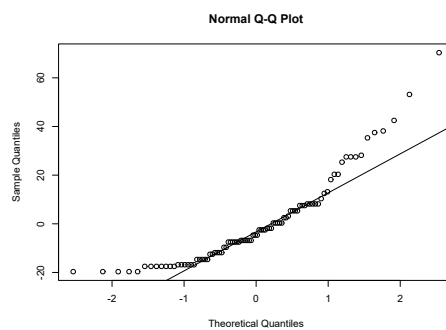


Abbildung 58: Normal-Q-Q-Plot: NASA RAW-TLX: Subskala Zeitliche Anforderung

A.15 Residuen: NASA RAW-TLX: Subskala Leistung

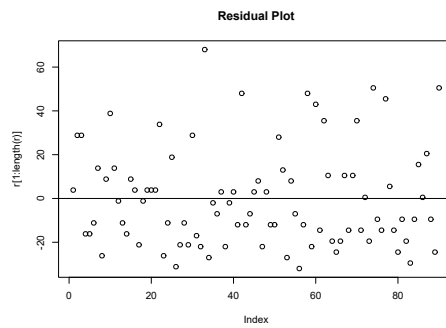


Abbildung 59: Residuen-Plot: NASA RAW-TLX: Subskala Leistung

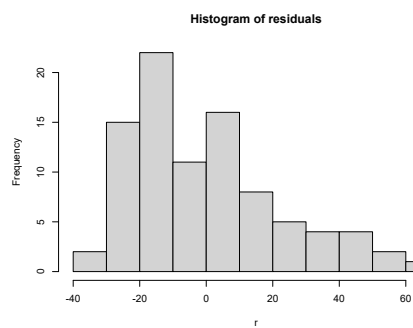


Abbildung 60: Residuen-Histogramm: NASA RAW-TLX: Subskala Leistung

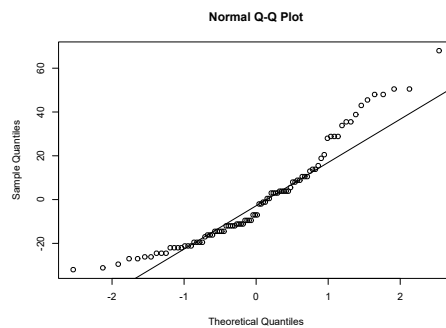


Abbildung 61: Normal-Q-Q-Plot: NASA RAW-TLX: Subskala Leistung

A.16 Residuen: NASA RAW-TLX: Subskala Anstrengung

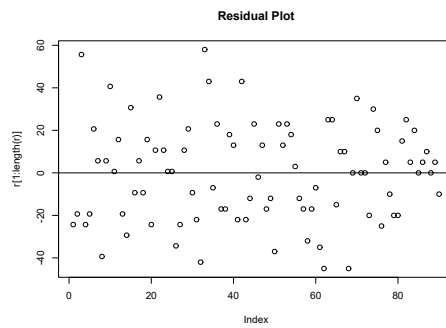


Abbildung 62: Residuen-Plot: NASA RAW-TLX: Subskala Anstrengung

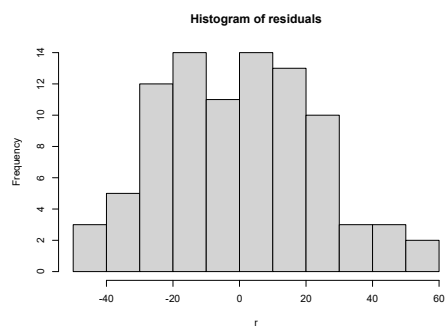


Abbildung 63: Residuen-Histogramm: NASA RAW-TLX: Subskala Anstrengung

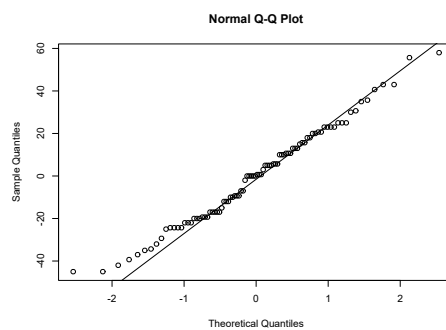


Abbildung 64: Normal-Q-Q-Plot: NASA RAW-TLX: Subskala Anstrengung

A.17 Residuen: NASA RAW-TLX: Subskala Frustration

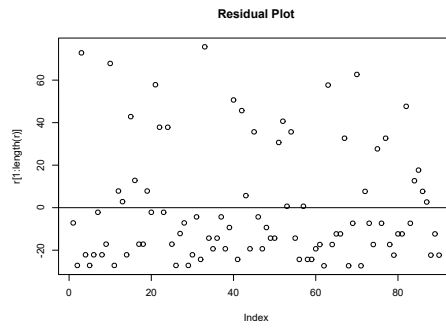


Abbildung 65: Residuen-Plot: NASA RAW-TLX: Subskala Frustration

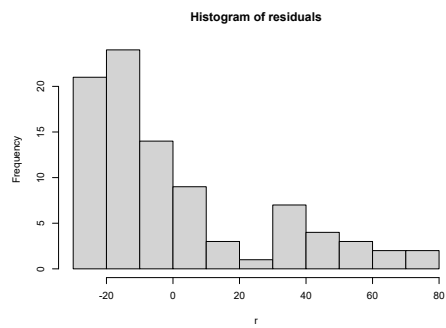


Abbildung 66: Residuen-Histogramm: NASA RAW-TLX: Subskala Frustration

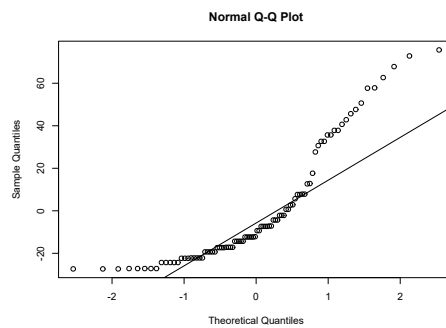


Abbildung 67: Normal-Q-Q-Plot: NASA RAW-TLX: Subskala Frustration

A.18 Dummy-Text aus Testrunden

Seit 2019 werden vermehrt tote Grauwale an den Pazifikküsten von Mexiko bis nach Alaska angetrieben – in einer Zahl, die Meeresbiologen sorgt. Verschiedene Erklärungen kursierten, was das Sterben ausgelöst haben könnte: von zunehmenden Schiffskollisionen bis hin zum pazifischen Blob, einer Warmwasserblase im nordöstlichen Pazifik, die Nahrungsgründe verschiebt und viele Tierarten in Mitleidenschaft gezogen hat. Joshua Stewart von der Oregon State University in Newport und sein Team suchen die Ursache dagegen noch weiter nördlich in der Arktis. Die Fläche des Meereises dort stehe in direktem Zusammenhang mit den Bestandsschwankungen bei den Grauwalen, schreibt die Arbeitsgruppe in »Science«.

Seit dem Ende des kommerziellen Walfangs im östlichen Pazifik hat sich dort die Zahl der Grauwale wieder deutlich erhöht. Eine Bestandsschätzung aus dem Jahr 2016 kam auf rund 27 000 Tiere. Der Anstieg verlief allerdings nicht kontinuierlich, denn bereits in den 1980er und 1990er Jahren verendeten Grauwale in größerer Zahl, wobei der Rückgang nur kurze Zeit andauerte, während das Sterben seit 2019 noch andauert. Viele der Opfer wirkten ausgezehrt, allerdings galt das nicht für alle gestrandeten Tiere.

Stewart und Co blickten daher genauer auf die Entwicklung des Meereises in der Arktis, wohin die Wale im Sommer ziehen, um zu fressen. Den Winter verbringen sie dann dagegen entlang der Küste von Baja California, wo sie ihre Jungen zur Welt bringen – eine Wanderung über 12 000 Kilometer. Der Vergleich der Meereisdaten mit den Populationsschwankungen der Wale zeigte ein klares Bild: Wenn die eisbedeckte Fläche stark zurückgeht, setzt ein Walsterben ein. Das galt für die 1980er und 1990er Jahre ebenso wie für die aktuelle Entwicklung.

Der Zusammenhang ist allerdings nicht ganz so einfach, wie es zuerst erscheint. Die Grauwale fressen bevorzugt bestimmte Krustentiere, die sich wiederum von Algen ernähren, die direkt unter dem Eis wachsen. Je mehr Eis vorhanden ist, desto mehr dieser Algen schwimmen im Wasser und desto größer ist die Biomasse der davon lebenden Flohkrebse. Schmilzt das Eis, können die Grauwale auf einer größeren Fläche fressen und ihnen steht mehr Nahrung zur Verfügung.

Doch die Hausse ist nur kurzfristig. Erholt sich das Eis nicht, schrumpft die Biomasse der Algen und damit ebenso die der Krustentiere. Stattdessen machen sich andere Arten breit, die den Walen weniger Energie liefern. Zudem verteilen Meeresströmungen in den eisfreien Regionen die Algen und Flohkrebse, so dass diese weniger konzentriert vorkommen. »Weniger Eis bedeutet weniger Algen, was sich negativ auf die Beute der Grauwale auswirkt«, sagt Stewart. »All diese Faktoren kommen zusammen und verringern die Qualität und Verfügbarkeit der Nahrung, auf die sie angewiesen sind.«

Für die Grauwale führt die geringere Verfügbarkeit von Beutetieren letztendlich zum Tod. Das Ereignis seit 2019 ist noch nicht abgeschlossen und dauert schon wesentlich länger an als die beiden früheren Wellen, was mit dem negativen Eistrend in der Arktis zusammenhängt. Die Region heizt sich durch den Klimawandel schneller auf als viele andere Regionen der Erde, und bald könnte sie im Sommer gänzlich meereisfrei bleiben.

»Wir befinden uns jetzt auf unbekanntem Gelände. Die beiden vorangegangenen Ereignisse dauerten trotz ihrer Bedeutung und Dramatik nur ein paar Jahre«, so Stewart. »Das jüngste Sterbeereignis hat sich verlangsamt und es gibt Anzeichen für eine Trendwende, aber die Population ist bislang weiter zurückgegangen. Ein Grund für die verzögerte Erholung könnte der Klimawandel sein, der zu einem langfristigen Trend hin zu weniger nahrhaften Beutetieren beiträgt.«

Neueste Schätzungen gehen von 14 500 Grauwalen im nordöstlichen Pazifik aus. Ein Aussterben der Population in der Region befürchten die Wissenschaftler dennoch nicht; sie gehen eher von einer Stabilisierung auf niedrigerem Niveau aus.

A.19 Text 1

Im Februar 2022 starb in Chile eine 93-jährige Frau. Mit ihr verstummte eine ganze Sprache. Denn Cristina Calderón war die letzte Sprecherin von Yaghan, das einst in ganz Feuerland, an der Spitze Südamerikas, verbreitet war. Mit Yaghan ging nicht nur eine Fülle an Wissen verloren, sondern auch eine isolierte Sprache – eine, zu der es auf der gesamten Welt nichts Vergleichbares gab und die keinen Bezug zu einer der bekannten Sprachfamilien erkennen ließ.

Von ungefähr 7000 existierenden Sprachen stellen etwa 200 isolierte Sprachen dar – und viele von ihnen sind vom Aussterben bedroht. Schätzungen zufolge werden 30 bis 50 Prozent aller Sprachen bis zum Ende des 21. Jahrhunderts verschwunden sein, darunter vor allem solche, derer nur wenige hundert Menschen mächtig sind.

Doch je gefährdeter die Existenz einer Sprache ist, desto intensiver forschen Fachleute über sie und erkennen: Isolierte Sprachen bergen Indizien, mit denen die Bandbreite menschlicher Kommunikation und Kognition genauer zu erfassen ist. In den vergangenen Jahren lieferten sie neue Impulse zu der Frage, wie die Entwicklung einer Kultur und ihre Sprache miteinander zusammenhängen. Und sie lieferten neuen Stoff für die umstrittene These, dass die Sprache beeinflusst, wie Menschen ihre Umgebung wahrnehmen und sie in Worte fassen. »Jede isolierte Sprache öffnet ein neues Fenster ins menschliche Denken«, sagt Lyle Campbell, Linguist an der University of Hawai'i at Mānoa. Forscherinnen und Forscher wie Campbell erkunden daher auch Strategien, um seltene Sprachen vor dem Aussterben zu bewahren.

Die Letzten einer Linie

Die Sprachgeschichte ist im Grunde ein Familienstammbaum. Die Ähnlichkeiten zwischen Dänisch und Niederländisch, zwischen Swahili und Zulu oder Türkisch und Usbekisch bezeugen, dass sich Sprachen eigentlich nicht voneinander getrennt entwickelt haben. Vielmehr verteilen sie sich auf etwa 140 Abstammungslinien – ausgenommen die so genannten Isolate. Von jenen ist wohl Baskisch am bekanntesten, das Menschen an der Grenze zwischen Frankreich und Spanien sprechen. Die Unterschiede zu den benachbarten Sprachen sind frappant: Während »vin« ein beliebtes französisches Getränk ist und auch in Spanien viele Menschen gerne ein Glas »vino« genießen, ordert man in einem baskischen Lokal »ardoa«. Und spricht man in Spanien und Frankreich von »agricultura« beziehungsweise »agriculture« – der Landwirtschaft –, heißt es im Baskenland »nekazaritza«.

Sicher ist: Isolate waren irgendwann einmal Teil einer größeren Sprachfamilie. Doch mit der Zeit sind die anderen Mitglieder verschwunden, übrig blieb ein einziges. Das sibirische Ket zum Beispiel gehörte einst zur kleinen Familie jenesischer Sprachen. Heute ist es das letzte Überbleibsel jener altsibirischen Idiome. Womöglich weil das Russische alle anderen verdrängte, als die Russen zwischen dem 16. und 18. Jahrhundert Stück für Stück Sibirien eroberten.

Das Überbleibsel einer Weltanschauung

Für Linguisten sind isolierte Sprachen besonders aufschlussreich. Denn jede von ihnen zeigt, welche Varianten der Kommunikation die Menschheit entwickelt hat. »Die Art und Weise, wie Menschen mit ihren Sprachen eine kognitive Welt konstruieren, ist extrem vielfältig«, sagt Anna Belew von der University of Hawai'i at Mānoa, die wie Campbell im Endangered Languages Project leitend tätig ist. »Und isolierte Sprachen – allein auf Grund der Tatsache, dass sie keine Schwestersprachen haben – fassen eine Weltanschauung meist auf einzigartige Weise in Worte.«

Indem Forscher und Forscherinnen die Strukturen jener Idiome entschlüsseln, decken sie auch bislang unbekannte Sprachregeln auf. Kutenai zum Beispiel ist ein Isolat, das eine indigene Gemeinschaft mit gleichem Namen spricht. Ihr Stammesgebiet liegt an der Grenze zwischen den USA und Kanada. Und Kutenai ist nahezu einzigartig: Es kennt eine grammatikalische vierte Person. Das heißt, die Kutenai können in einem Satz wie »Er hielt seinen Hut« genau spezifizieren, wem der Hut gehört – dem Subjekt »er« oder einer anderen männlichen Person. Kusunda, das in Nepal gesprochen wird, kommt hingegen ohne Elemente aus, die in den meisten Sprachen absolut wesentlich sind: Wörter wie »nein«.

• Frage 1: Welche Konsequenz brachte der Tod der 93-jährigen Frau aus Chile mit sich?

Mögliche Antwort-Optionen: Der Verlust einer ganzen Sprache / Mit ihr starb eine Sprache / Sie war die letzte Sprecherin ihrer Sprache (bzw. Yaghan, falls sich wer erinnert)

• Frage 2: Wie nennt man die Sprachen, die sich unabhängig von den 140 Abstammungslinien entwickelt haben?

Mögliche Antwort-Optionen: Isolate / Isolierte Sprachen

• Durch welche Besonderheit kann die Kutenai-Sprache z.B. den Besitz eines Hutes genauer spezifizieren?

Mögliche Antwort-Optionen: Die vierte Person / eine zusätzliche Person in der Grammatik

A.20 Text 2

Historische Linguisten richten ihren Blick nicht nur auf alte isolierte Sprachen, sondern sie untersuchen auch solche, die erst vor relativ kurzer Zeit entstanden sind. Wie im Fall der Al-Sayyid-Beduinen-Gebärdensprache (ASBG), die Beduinen in der Wüste Negev in Israel verwenden. Der Stamm der Al-Sayyid entwickelte ASBG in der ersten Hälfte des 20. Jahrhunderts. Die Zeichensprache sollte die Kommunikation in einer Bevölkerung erleichtern, in der ungefähr 50-mal mehr Menschen gehörlos sind als im weltweiten Durchschnitt. Der Grund: Viele Stammesmitglieder sind genetisch bedingt von Geburt an taub. Doch ASBG ist speziell, weil es grundlegend anders aufgebaut ist als alle Sprachen der Welt.

Jede Sprache beruht normalerweise auf einer zweifachen Gliederung. Was ist damit gemeint? Für gewöhnlich kombinieren Menschen eine gewisse Zahl an für sich genommen bedeutungslosen Lauten oder Gesten und bilden so einen umfangreichen Wortschatz. Letzterer entspricht der ersten Gliederungsebene, zu der Wörter und Sätze zählen, ihre kleinsten Einheiten sind Morpheme. Die zweite Gliederung umfasst die Laute oder Phoneme, die in Form von Buchstaben niedergeschrieben werden können. Sie allein tragen keine Bedeutung, aber aus ihnen lassen sich bedeutungsvolle Wörter konstruieren. Ein Beispiel: Ich habe Bauchweh. Diese Kombination vermittelt eine eindeutige Information. Und zwar genau diese Information. Aus den Bestandteilen des Satzes oder den einzelnen Wörtern ließen sich aber auch andere bedeutungstragende Aussagen zusammenstellen. Wohingegen die einzelnen Komponenten der zweiten Gliederung, bei Bauchweh etwa »b«, »au« oder »w«, für sich genommen nicht signifikant sind.

Viel zu komplex!

Die zweifache Gliederung bringt Vorteile mit sich, weil sie das Gehirn nicht überstrapaziert. Man stelle sich vor, es gäbe ein Kommunikationssystem, in dem der spezifische Schmerz für Bauchweh einem bestimmten Schrei entspräche. Dann bräuhete man ein System, das »eine so beträchtliche Anzahl verschiedener Zeichen aufweisen müsste, dass das menschliche Gedächtnis sie nicht fassen könnte«, schrieb der Linguist André Martinet (1908–1999), der den Begriff der »double articulation«, der zweifachen Gliederung, einführte.

Eine Sprache, in der jedes Wort einzigartig ist und nicht in kleinere, festgelegte Einheiten zerlegt werden kann, wäre demnach viel zu komplex. Menschen müssten sich merken, dass ein Zischen in einer bestimmten Klangfarbe »Frühstück« hieße oder ein leichtes Krächzen »Toilette«. Doch genau so funktioniert ASBG. Anders als alle bekannten gesprochenen und Gebärdensprachen müssen sich die Al-Sayyid für jedes Wort eine neue und eindeutige Geste einprägen. ASBG bezeugt damit, dass die zweifache Gliederung kein universelles Sprachmerkmal ist. Und die Gebärden der Al-Sayyid liefern Hinweise darauf, warum dem so ist.

Die beiden Linguisten Simon Kirby von der University of Edinburgh und Monica Tamariz von der dortigen Heriot-Watt University verglichen ASBG mit einer nicht verwandten, aber ähnlich jungen Gebärdensprache – dem israelischen Shassi. Diese ist zweifach gegliedert. Kirby und Tamariz fiel auf, dass Schüler sie in Gruppen lernen, wie es auch für viele andere Gebärdensprachen der Fall ist. Und immer dann, wenn sich gleichaltrige Schüler und Schülerinnen miteinander unterhielten, nutzten sie die zweifache Gliederung – weil es den Austausch erleichterte. Eine Rolle spielt dabei auch, dass Lernende einer neuen Sprache stets dazu tendieren, es sich möglichst einfach zu machen. Wenn dann andere Schüler die simpleren Konstruktionen übernehmen, gehen diese allmählich in den Sprachschatz der gesamten Gemeinschaft über.

Im Fall von ASBG verläuft der Lernprozess dagegen nicht über Gleichaltrige, sondern die Älteren der Al-Sayyid geben die Gebärden vor. Selbst wenn die Sprachschüler vereinfachte Zeichen entwickelten, dann würden sie die Älteren nicht adaptieren – weil es für sie Kindersprech wäre. Die Folge: Neue, simple Gebärden finden keinen Eingang in die Sprache, damit bleibt sie komplex.

- **Frage 1: Bezüglich welchen Defizits hebt sich der israelische Al-Sayyid-Stamm vom weltweiten Durchschnitt ab?**

Mögliche Antwort-Optionen: Genetisch bedingte Taubheit / Taubheit / Taubheit von Geburt an / (Anteil) gehörlose(r) Menschen

- **Frage 2: Worauf basieren Sprachen, in denen sinnhaltige Wörter aus bedeutungslosen Lauten oder Gesten konstruiert werden?**

Mögliche Antwort-Optionen: Zweifache Gliederung / double articulation / zwei Gliederungsebenen

- **Frage 3: Wieso adaptieren die Älteren bei ASBG die simpleren Gebärden der jüngeren Altersgruppe nicht?**

Mögliche Antwort-Optionen: Antwort (Variationen im Wortlaut möglich): Sie sehen sie als „Kinder-Sprech“

A.21 Text 3

Zu den großen Fragen der Linguistik zählt eine These, die sehr umstritten ist: die Sapir-Whorf-Hypothese. Bei ihr geht es darum, wie sich Sprache und Denken gegenseitig beeinflussen. Denn einige Fachleute sind überzeugt davon, dass die Art und Weise, wie wir kommunizieren, auf unsere Wahrnehmung der Welt und auf abstrakte Konzepte wie Zahlen einwirkt. Isolierte Sprachen legen nun nahe, dass die Sapir-Whorf-Hypothese stimmen könnte.

Häufig genannt in jener Diskussion ist die indigene Gruppe der Pirahã, die im Norden Brasiliens lebt. Ihre Sprache – ebenfalls Pirahã genannt – kennt Untersuchungen zufolge keine Wörter für Zahlen. Weitere Studien deuten nun darauf hin, dass die Pirahã auch Schwierigkeiten beim Zählen hätten. Folglich muss man über Wörter für Zahlen verfügen, um zumindest ein bisschen Mathematik betreiben zu können. Das Fazit klinge plausibel, sei aber höchst problematisch, sagt Benjamin Pitt von der University of California in Berkeley. Man habe nämlich bisher das Zahlenverständnis der Pirahã mit den Fähigkeiten und Kenntnissen von Menschen in Industrieländern verglichen. Eine ungleiche Gegenüberstellung, findet Pitt, weil sich die Lebensweise in einem Land wie den USA stark von der im Amazonasgebiet unterscheide.

Deshalb suchte Pitt nach einer Gemeinschaft, die einer ähnlichen Kultur wie die Pirahã angehört und Zahlwörter nutzt oder auch nicht. Fündig wurden Pitt und sein Team ebenfalls im Amazonasgebiet: bei den Tsimané im bolivianischen Tiefland. Die Sprache jener Menschen ist isoliert – und sie verwenden Zahlwörter. Allerdings kennen nicht alle Tsimané gleich viele davon: Einige können bis ins Unendliche zählen, andere hingegen geraten bei zehn oder zwölf ins Stocken.

Pitt und seine Kollegen baten 30 Tsimané, von denen 15 problemlos zählen konnten, mit ihnen zusammenzuarbeiten. Sie sollten Zählaufgaben lösen, nämlich neben eine Reihe Spielsteine exakt dieselbe Menge von Steinchen auslegen. Die meisten Probanden bekamen das problemlos hin – wenn sie die Spielsteine parallel zur Reihe auf den Tisch setzen konnten. Die Aufgabe war leicht, vermutet Pitt, weil man dafür eigentlich nicht zählen können muss.

Zählen bis vier ist universell

Als Nächstes bat Pitts Team die Probanden und Probandinnen, die Steinchen lotrecht zur Reihe anzulegen. Nun waren Zählkenntnisse erforderlich. Das zeigte auch der Versuch. Bis zu vier Steinchen konnten alle Teilnehmenden anlegen. Doch umfasste die vorgegebene Reihe mehr als vier Objekte, gelang es ihnen nur, die Senkrechte so weit zu führen, wie ihre Kenntnis der Zahlwörter reichte. Wer beispielsweise die Zahlen bis 15 beherrschte, konnte maximal eine ebenso lange Reihe bilden. Ein Zahlenverständnis bis vier haben Wissenschaftlerinnen und Wissenschaftler ebenso bei verschiedenen Tierarten beobachtet wie Küken, Bienen oder Pferden. Es scheint eine universelle Fähigkeit zu sein. Einige Fachleute gehen hingegen davon aus, dass Menschen auch ohne spezifische Wörter, nur kraft ihrer Kognition weiter zählen können. Pitts Forschungen legen allerdings das Gegenteil nahe. Mit Hilfe der Sprache, vor allem für den Bereich jenseits der Zahl Vier, eröffnen sich Menschen »neue konzeptionelle Fähigkeiten«, sagt Pitt, die ihnen anderenfalls nicht zur Verfügung stünden.

Isolierte Sprachen wie die der Tsimané bergen Erkenntnisse über das menschliche Denken. Weil ihre Strukturen in keiner zugehörigen Sprachfamilie mehr erhalten sind, ist Forschern und Forscherinnen sehr daran gelegen, sie zu untersuchen. »Wir müssen unbedingt Informationen über isolierte Sprachen gewinnen«, sagt Campbell. »Denn wenn wir sie verlieren, sind diese Informationen für immer verloren.«

- **Frage 1: Wieso wird in der Diskussion um die Sapir-Whorf-Hypothese häufig die indigene Gruppe der Pirahã betrachtet?**

Mögliche Antwort-Optionen: Man schaut, ob das Fehlen von Zahlwörtern das Denken (Rechnen) beeinflusst (und umgekehrt) / Es fehlen ihnen Wörter für Zahlen, was heißen könnte, dass auch nicht richtig gezählt bzw. gerechnet werden kann / Studien zeigen, dass die Pirahã Probleme beim Zählen haben, da ihnen Wörter für Zahlen fehlen

- **Frage 2: Wieso wählte Benjamin Pitt für seine Studie eine Gemeinschaft aus, die eine ähnliche Kultur wie die Pirahã hat?**

Mögliche Antwort-Optionen: Er sah den Vergleich mit Industrieländern als unverhältnismäßig / man hat die Pirahã davor nicht mit ähnlichen Kulturen verglichen / eine ähnliche Kultur ist für eine Gegenüberstellung mit der Pirahã sinnvoller

- **Frage 3: Welche universelle Fähigkeit zeichnete sich im Experiment mit den Tsimane ab?**

Mögliche Antwort-Optionen: Zählen bis vier / Zahlenverständnis bis vier

Abbildung 68: Konstruktion des Counterbalanced Latin Square nach der Methode von Lewis (1989, 1993).

A	C	B		1	3	2
B	A	C		2	1	3
C	B	A		3	2	1

B	C	A		2	3	1
C	A	B		3	1	2
A	B	C		1	2	3

Teilnehmer					Kondition x Stimulus -Reihenfolge		
#1	#7	#13	#19	#25	A2	C3	B1
#2	#8	#14	#20	#26	B3	A1	C2
#3	#9	#15	#21	#27	C1	B2	A3
#4	#10	#16	#22	#28	B1	C3	A2
#5	#11	#17	#23	#29	C2	A1	B3
#6	#12	#18	#24	#30	A3	B2	C1

Ich bin damit einverstanden, dass meine Arbeit in den Bestand der Bibliothek eingestellt wird.

Hamburg, 13.09.2024

Ort, Datum



Unterschrift

Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit im Studiengang Mensch-Computer-Interaktion selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel — insbesondere keine im Quellenverzeichnis nicht benannten Internet-Quellen — benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen entnommen wurden, sind als solche kenntlich gemacht. Ich versichere weiterhin, dass ich die Arbeit vorher nicht in einem anderen Prüfungsverfahren eingereicht habe und die eingereichte schriftliche Fassung der auf dem elektronischen Speichermedium entspricht.

Hamburg, 13.09.2024

Ort, Datum



Unterschrift